Title
The Political Geography of Territorial Control in Africa

Author
Seese, Michael F.

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

The Political Geography of Territorial Control in Africa

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Political Science and International Affairs

by

Michael F. Seese

Committee in charge:

Professor Karen Ferree, Chair
Professor Claire Adida
Professor Prashant Bharadwaj
Professor Jesse Driscoll
Professor David Lake

2023

The Dissertation of Michael F. Seese is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Many friends and colleagues provided help and support throughout this process. Scott Simpson has been my constant sounding board and advisor on both technological and personal issues since our undergraduate days, and has proven a diverting travel companion through more countries than I can count in the decades since. Mayra Gonzalez and Michelle Cash remain my dual anchors in San Diego, and my primary link to the world outside the Academy. Their friendship means the world to me. Matthew Wagner has been a source of indispensable professional, substantive, and research advice for many years. My time traveling with Matt through Latin America and Southeast Asia helped to expand my academic worldview beyond the African continent, and gave me a profound appreciation for the old school, "deep" fieldwork valued by area specialists and others of his tribe. Ana García Hernández has been a wonderful friend, colleague, and collaborator, as well as travel buddy, Spanish language tutor, brunch companion, and flatmate during my time in Europe.

This dissertation owes much to the intellectual contributions of my "UCSD Brain Trust:" Julia Clark and John Porten. Over the past decade, I spent many delightful hours over tea and shisha with Julia, on no less than three separate continents. Her clarity and analytical insight were invaluable, both to the dissertation and to every other project I have pursued (or considered and ultimately rejected) throughout graduate school. Over the years, John has assumed the role of collaborator and "work husband," balancing my empiricism and introversion with his talents for theory building and the more social aspects of research and field work that I find fear-inducing or tedious. I look forward to our continued friendship and collaboration. As comparativists, it was often difficult for the three of us to be in the same room—not least the same country—at the same time, particularly over the past couple of years as our fieldwork progressed and our lives changed, but I count our times together among the great highlights of graduate school.

I am lucky to have spent time at UCSD in the company of many other brilliant and talented graduate students, including Taylor Carlson, Christina Cottiero, Brian Engelsma, Charlie McClean, Zoe Neveremer, Stan Oklobdzija, Gregoire Phillips, Inbok Rhee, Luke Sanford, Rachel Schoner, Liesel Spangler, Alex Verink, Walker Vincoli, and Devin Wright. Many parts of this dissertation were improved by their feedback, both in workshop and informally over countless cups of coffee.

In the Fall of 2017, I was lucky to find an intellectual home at the WZB Berlin Social Science

Center. Macartan Humphreys brought me to Berlin as a predoctoral research fellow and founding member of the Institutions and Political Inequality Research Group, and provided me with a set of personal, professional, and research opportunities that were foundational in shaping my research agenda and academic trajectory. I am immensely lucky to count Macartan as a mentor and friend. My colleagues at WZB served, for several years, as a daily source of encouragement and diversion. Manuel Bosancianu made certain that I ate a square meal each day in the WZB canteen, and that I occasionally left my work cave for drink or two. Clara Bicalho Correia, Ana García Hernández, and Lily Medina were generous in their time and friendship. Bernd Beber and Alex Scacco helped to socialize me into the discipline, and ensured that I took active steps to network myself in the European and U.S. academic communities. Marion Obermaier and Reinhild Wagner did everything they could to insulate me from the more mundane aspects of the WZB administration, and helped me adapt to life in Germany and at the WZB. Jan-Hinrich Wagner guided me through the formidable German state bureaucracy; without him, I would certainly have ended up either homeless or deported.

I owe a great deal to my many mentors at UCSD. Karen Ferree has been an incredible advisor and chair. Her encouragement and understanding got me through the most difficult times of the PhD program; her insight and advice greatly improved all aspects of my dissertation. Karen is the academic and professional role model that I strive to emulate. Jesse Driscoll was my first mentor at UCSD—he brought me on as a research assistant on my first real research project while I was in the master's program at IRPS. Our conversations over the years helped to define the way I think about everything from politics to pedagogy. Claire Adida was a constant source of encouragement during my time at UCSD, particularly during the height of the pandemic. Claire is responsible for "showing me the ropes" of how a research project evolves from pre-analysis plan to publication. David Lake provided exceptionally clear feedback throughout the dissertation process, and taught me a great deal about the discipline and the profession. Prashant Bharadwaj generously gave his time and perspective, and I am immensely thankful for his support. Clark Gibson served as my sounding board throughout graduate school; he supplied me with both coffee and clarity when I needed them most. Barbara Walter and Emilie Hafner-Burton encouraged me to pursue a PhD after completing my master's degree at UCSD. Their support during the application process literally changed my life.

Ariane Parkes, Julie Choi, and Kathleen Smith helped me to navigate the complexities of

Graduate Affairs and University administration. They also ensured that I had stipend check in my account each month. It is no exaggeration to say that this dissertation would not have been written without their support. Several institutions helped me throughout this process as well. I am especially grateful to the Friends of the UCSD International Center and the Development Impact Lab for providing the funding for my first real fieldwork experience in Kenya during the Summer of 2014.

My greatest thanks, however, are reserved for my wonderful family. My aunt, Virginia Rich, provided a comfortable bed and a warm meal each time I circulated through Southern California while I was living abroad. Watching her write her own dissertation as a young child sparked my interest in the Academy as a vocation. My sister, Lisa, has looked after me since childhood, and has shouldered more than her fair share of familial burdens as I was off trekking through various corners of the world. The most recent Seese, my nephew Kaleo, was born as I started my first master's program. It has been my great privilege to watch him grow up and prepare to embark on his own college journey, and I look forward to his promising future.

The majority of the credit for who I've become and what I've achieved go to my parents, Craig and Estrella Seese. My mother, from a very early age, instilled in me the value of hard work and self-discipline. Her belief that education was the greatest gift a parent could leave a child was central, in many ways, to both my sister's and my own career choices. Her encouragement and sacrifice opened up the world for me, and I love her dearly. My father has been a parent, friend, and role model since the day I was born. He taught me kindness, and imparted on me a sense of adventure and curiosity that define who I am today. But he also taught me strength and survival; as a child, in the woods, and as a man, in the face of a terminal diagnosis of ALS. It is to my parents that I dedicate this dissertation.

VITA

2012	Master of Pacific International Affairs University of California San Diego
2017-2019	Predoctoral Research Fellow Wissenschaftszentrum Berlin für Sozialforschung (WZB Berlin Social Science Center)
2023	Doctor of Philosophy in Political Science and International Affairs University of California San Diego

ABSTRACT OF THE DISSERTATION

The Political Geography of Territorial Control in Africa

by

Michael F. Seese

Doctor of Philosophy in Political Science and International Affairs

University of California San Diego, 2023

Professor Karen Ferree, Chair

Many African states fail to exercise meaningful control over the entirety of the territory defined by the state's de jure borders. This dissertation seeks to map the current geographic reach of the African state, and assess the impacts of living in either state-consolidated or unconsolidated territory on the lives of African citizens.

The first part of the dissertation presents a novel mapping of governed and ungoverned space in Africa, and asks why state leaders choose to exert control over certain regions of their countries and not others. Using a series of supervised machine learning algorithms, I find that the spatial distribution of state authority is strongly correlated with market access and economic productivity, as well as areas with high concentrations of critical infrastructure, suggesting that rent extraction and strategic concerns are two core motivations in the decision of where to locate government assets.

The second part of the dissertation looks explicitly at the effects of living inside and outside of state-controlled territory. I examine outcomes such as the public's attitudes towards traditional authorities, and the incidence of transmissible diseases like Malaria and HIV, and find marked differences between areas in which the state is present and areas in which the state is effectively absent.

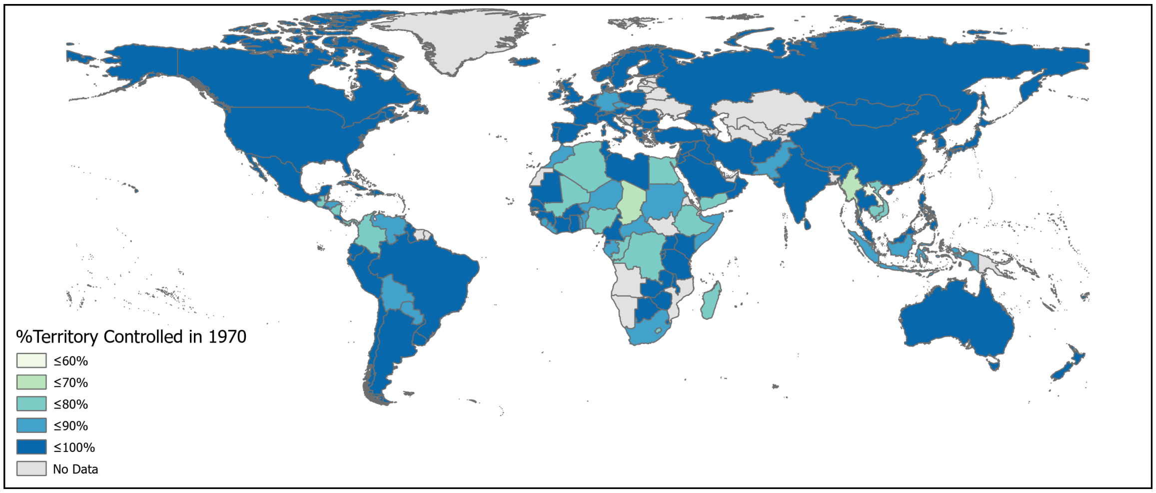
Chapter 1

Introduction

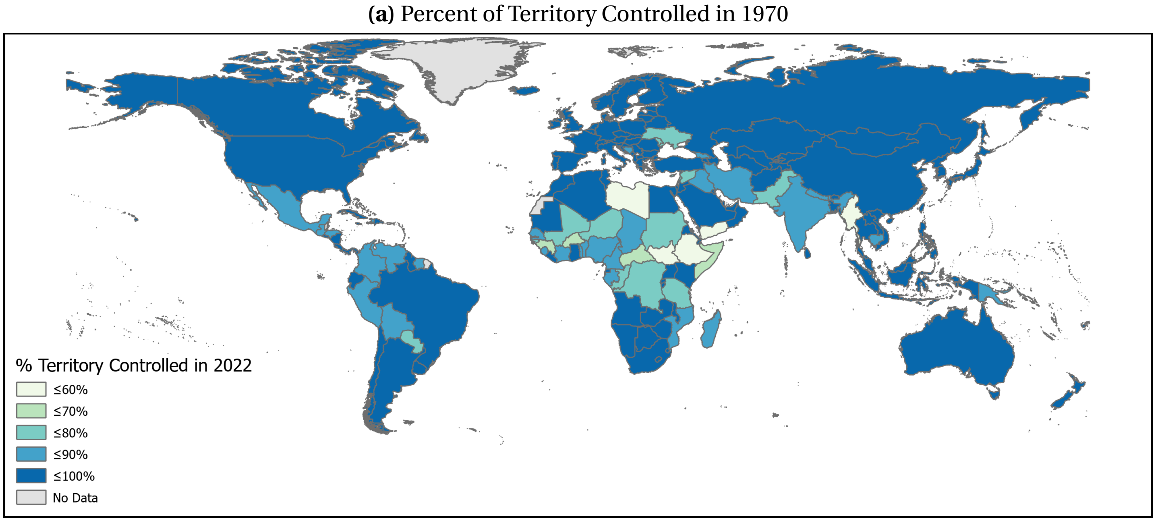
1.1 The "Weak" African State

Nearly 40 years ago, Jackson and Rosberg (1982, 1) observed that Africa's states were among the weakest in the world—many of these newly independent countries were plagued by ineffective institutions, political instability, and governments that exercised "only tenuous control over the people, organizations, and activities within their territorial jurisdictions" or that have "periodically ceased to control substantial segments of their country's territory and population." This problem of incomplete control that Jackson and Rosberg highlight persists to the present day. Figure 1.1, for example, maps data from the Varieties of Democracy (V-Dem) Project that estimates the percentage of a country's territory that is under the control of the central government in the years 1970 (the year by which a majority of African states achieved independence) and 2022 (the final year for which V-Dem has data). Although several African states have apparently succeeded in their state consolidation efforts (e.g., Algeria, Egypt, South Africa), a comparison of the two maps in Figure 1.1 suggests very little change in terms of the extensive margin of territorial control that we observe in most African states over the past half century. As it was in the immediate post-independence period, Africa today remains a hotspot of weak and failed states.

This dissertation explores the phenomenon of state weakness and incomplete state consolidation in the African context. Like most Africanist scholars, I understand the relative "strength" or "weakness" of a state to be largely a function of *control*: a state is strong to the extent that it exercises effective control over territory—and the people, agents, and organizations that exist on that territory—and weak to the extent that it lacks this control (Jackson and Rosberg 1982; Herbst 1996;



(a) Percent of Territory Controlled in 1970



(b) Percent of Territory Controlled in 2022

Figure 1.1. Varieties of Democracy Project estimates of the percentage of a country's territory controlled by the central government in 1970 and 2022 (V-Dem variable `v2svstterr`).

Reno 2001; Wang 2003; Rotberg 2004; Herbst 2014). In what follows, I focus on two broad questions concerning the control that African states are able assert and maintain. First, which areas within the de jure borders of a state is the central government able to control? V-Dem estimates that Senegal was able to exercise control over 86.83% of its territory in 2022, while Kenya was able to exercise control over 90.80% of its territory. It is unclear, however, which specific territories these assessments refer to, or where these territories are situated. By developing a subnational measure of territorial control, I am able to improve upon these blunt, methodologically opaque estimates and explore within-state and between-state variation in state strength. I am also able to test certain hypotheses about why states tend to concentrate control in some areas of their jurisdictions but not others.

The second question I address is whether or not control actually matters to ordinary citizens: do we see differential outcomes for Africans living in state-incorporated versus unincorporated territory? Received wisdom in the social sciences (e.g., Skocpol 1985) suggests that the state plays an important, and often beneficial role in the lives of its citizens; we should therefore expect residents of state-incorporated territory to enjoy higher levels of welfare and public goods provision than residents of territory outside the control of the central government.[1] In Chapters 5 and 7, I interrogate this assumption by looking at outcomes such as public health and governance.

1.2 Territorial Control

The two questions that form the basis of this dissertation each revolve around the concept of territorial control, which I define as *the degree to which a state is able to influence, direct, or restrict the behavior of those people, agents, and organizations that exist within some arbitrarily defined territorial space within the polity*. Several facets of this definition are worth highlighting. First, this definition centers on the powers and capabilities exercised by the formal state, or the de jure central government of a given polity. Typically, the "state" refers to the decision-making and administrative apparatus associated with the political organization that maintains at least a nominal hold over the capital or primate city. It is important to note, however, that the state's level of control within the capital may itself vary, as we see in the case of the Transitional Federal Government (TFG) in Mogadishu between 2004 and 2012, or in the current government of Sudan in Khartoum. What is

1. Many scholars (e.g., Reno 1997; Meagher 2012; Lust 2022) also underscore the role of non-state and informal institutions in promoting security, development, and other pro-social outcomes.

important for definitional purposes is not the extent to which this organization has consolidated power over the capital (or the periphery), but rather that its nominal hold on power is *privileged*, insomuch as it affords this organization international legitimacy and the assortment of benefits that flow from state capture (e.g. taxation, foreign aid, policy making, etc.).[2]

Second, I define territorial control in terms of the state's potential, whether latent or manifest, to influence, direct, or restrict behavior. Control is often understood in terms of its observable impact on human behavior. Agnew (1999, 502), for example, argues that control is only expressed when it results in behavioral changes; an authority's capacity to influence does not, in itself, constitute control. States, however, do not always actively exercise control, even if they have the ability to do so. There are many reasons to believe that the state's latent capacity to observe or coerce are sufficient to alter behavior in meaningful ways. The sociologist Erving Goffman makes this point in his 1966 exposition on human behavior in public places. Goffman suggests that merely being observed by authority, combined with one's awareness of being observed, tends to discourage deviant behavior and reenforce patterns of conduct consistent with social and legal norms. Thus, the mere existence of a state in a certain region may be sufficient to influence the behavior of the region's inhabitants.

Finally, this definition is agnostic to the spatial extent in which the central government's control operates, so long as that extent is contained within the de jure borders of the polity. Control may vary across undefined geographic space—within cities and villages, and occasionally from block to block and from street to street within the same neighborhood—rather than by city, village, or formal administrative district. This is an integral component of the definition, as there is no reason to believe that control is either confined to the physical extent of existing political borders, or uniform within those borders. I do, however, impose the restriction that control can only be exercised within the confines of the state's internationally recognized borders, as such a restriction is implied by traditional notions of Westphalian sovereignty.[3]

2. Control, of course, can be exercised by any number of *non-state* actors, including informal political organizations (e.g., clans, tribal elders, etc.), rebel groups, and pretender states. The simultaneous exercise of control by these non-state actors is a condition that Tilly (1975, 1996) refers to as "contested" or "multiple sovereignty." Occasionally, different factions within the internationally recognized government will compete for primacy, as was the case in South Sudan between 2013 and 2014, during which time the President, Salva Kiir Mayardit, and Vice President, Riek Machar, both held legitimate executive office, though each maintained a loyal faction of the Sudan People's Liberation Army (SPLA). In such cases, it is difficult to accurately identify which actor represents the "state"—the question becomes which faction constitutes the de jure, internationally-recognized national authority.

3. In practice, states often do extend their control in an extraterritorial fashion, as exemplified during the colonial and

1.2.1 Territoriality and the Primacy of Space

Why must the concept of control be so closely linked with that of territory? After all, many of the most common indicators of political control, such as tax compliance, corruption indices, or crime statistics, are not overtly spatial measures. The modern state, however, is an unequivocally territorial institution. Nearly all definitions of the state include some reference to territory (e.g., Weber 2021), and possession of territory is widely understood to be a necessary precondition for state sovereignty in the contemporary international system (e.g., Krasner 1988, 2001).

The modern, territorially-bound state is not the only method of structuring authority. Political authorities can be classified by the degree to which their control is exercised in a territorial or non-territorial fashion. As Branch (2013, 20) explains, "The primary distinction is between authorities defined in territorial or spatial terms and authorities defined without reference to space or place." Non-territorial authorities claim control over individuals or collections of individuals without reference to their spatial location or distribution. Examples of such authority structures include historical Christendom—a socioreligious polity in which geographically disconnected, multi-ethnic populations fell under the ecumenical authority of the Holy See—and the Hansa, a loose confederation of merchant guilds that existed in the Late Middle Ages, whose authority was confined to matters of trade, protection, and logistics. In these examples, the extent of political control is not characterized by territory, but rather by identity or issue domain.

A second model of political authority is what we might call "radiant" or "semi-spatial" authority. Classic examples include the Roman and Chinese Empires, in which control radiates outward from a strong center of power and decays toward a periphery (21). Herbst (2014, 45) notes that this was the dominant model of authority in pre-colonial Africa, explaining that "power was (quite realistically) conceived of as a set of concentric circles radiating out from the core." These semi-spatial authority structures have two characteristics that distinguish them from contemporary territorial states. First, there tends to exist overlapping systems of power and control. In the early 19^{th} Century, for example, the European kingdoms of France and Spain coexisted with the Holy Roman Empire, and in the African context, Ashanti authority (particularly in the southern provinces) was divided between the political ruler—the Asantahene, who held jurisdiction over land and property—and the

mandate periods, during war time, or during instances of external intervention and subversion (see Lee 2018).

Fanti, who held jurisdiction over people Herbst (2014, 40). These overlapping systems of authority made it difficult to distinguish between external and internal affairs, and allowed subjects to "forum shop," either through petition or exit. Second, rulers did not necessarily conceptualize their rule in terms of exhaustive claims based on discrete territorial divisions. These rulers lacked the legible, geometric mappings of terrain and people that Scott (1999) argues is central to modern statecraft.

The territorial basis for control is a relatively modern concept; Branch (2013, 32–33) traces the origins of the territorial state to 1815 and the conclusion of the Congress of Vienna. And as Herbst (2014, 36) points out, the linkage between control and territory was uncommon in pre-colonial Africa. Contemporary authority structures are defined by clear boundaries that delineate ostensibly homogeneous space, characterized by nominally undifferentiated control by a single political organization. Territorially-bound political authorities are currently the predominant method of organizing political markets, and they now serve as the central building blocks of the current international system. Because these organizations ground their rule in territorial claims, it is appropriate to assess their relative strength or weakness on territorial grounds.

1.2.2 Presence

In order to effectively exercise territorial control, states must maintain presence throughout their geographies. Presence refers to the physical existence of agents of the state in a specific locale. These agents include any individual or organization employed by or loyal to the central government, who in turn formally some basic administrative function, including (though not limited to) security, tax collection, or service provision.

Presence is necessary for states to exercise territorial control for a couple of reasons. The first deals with monitoring and surveillance, or what Scott (1999) terms "legibility." One Scott's core arguments is that in order to control territory, the territory must first be made legible to the center. Basic data ("mētis") must be collected, so that the center has an understanding of both what (or whom) is to be controlled, and the means and costs necessary to control it (them). To collect this information, the state relies on agents situated across a state's jurisdiction. They record vital statistics, such as population censuses and cadasters; they determine who is eligible for benefits and other entitlements, such as passports, voting rights, and and access to social services; and perhaps most importantly, they are instrumental in re-forming physical and social geographies to improve monitoring and extraction activities by the center.[4] Although many of these agents, particularly in high-income states, also provide social services and other public goods, their first order objective in the context of territorial control is surveillance of land and populations.

Presence is also a necessary precursor to two core functions of the state: exclusion (controlling the flow of people, goods, capital, and information throughout a territory) and extraction (the state's ability to generate revenue through either taxation or the direct sale of resources). Exclusion, for example, typically requires the presence of security forces, which are able to remove threats to the state (e.g., rebel groups, criminal enterprise, external states) through physical force, or to deter these threats from forming. The state's ability to extract resources from a given area, and to monitor the inhabitants of a particular space, are also made easier when agents of the state are at hand. Extraction requires either a bureaucratic infrastructure to monitor capital flows and levy taxes, or a logistics infrastructure of transportation networks and ports to move commodities to market.

The importance of physical presence in exerting control over territory is well-known to state-builders and military strategists alike. In fact, presence is typically the first step that claimants to power make when attempting to expand their authority. There are a number of historical examples from the European colonial period: In 1494, the Portuguese Empire used its foothold in the Cape Verde Islands to successfully demarcate its territorial control along a meridian 370 leagues from Ribeira Grande (now Cidade Velha) in the Treaty of Tordesillas. Similarly, in the late 19$^{\text{th}}$ Century, the British Raj demarcated control over various dominions (as contrasted with Suzerainties, which were princely states ruled by a de facto indigenous vassal ruler loyal to the British Crown) based on the physical presence security forces and administrative officers in cities such as Guwahati, Madras, and Calcutta. In the modern era, governments have spent a great deal of fiscal and military resources to extend physical presence into contested territory; examples of such efforts include the 2016-2017 battle to retake the Iraqi city of Mosul from the Islamic State, and the ongoing military operations in northeastern Nigeria to retake territory from Boko Haram.

4. See Scott (1999, 12–22) for an example of German foresters in early modern European states

1.3 State Consolidation and the Allocation of Control

Where do states choose to maintain their presence? Scholars generally address this question in the context of "state consolidation," a complex process by which political organizations come to monopolize the means of violence and a variety of associated functions—including taxation, social ordering, and the development and institutionalization of bureaucracies—and expand this monopoly across a given population or territory.[5] As the monopoly on violence becomes increasingly entrenched over time, these organizations take steps to insulate themselves from internal and external rivals, and to ensure the future survival of the organization. Such steps include the legitimization of state domination (to minimize internal threats), the hardening of geographic boundaries (to minimize external threats and interference by foreign actors), and the rationalization and depersonalization of rule (which increases the probability of the state's survival from one period to the next).

Social scientists have been interested in the process of state consolidation for several decades. In the 1980s, scholarship in political science, economics, and sociology began to converge on a set of interrelated explanations of the origins and development of the modern territorial state. These explanations revolve primarily around the necessity of raising capital to finance state activities, including war (Bean 1973; Cohen et al. 1981; Tilly et al. 2017) and the provision of property rights and other public goods (Levi 1981; North 1982; Bates et al. 2002). There is some controversy, however, on the utility of these theories in explaining the development of non-Western states and states that formed, or gained independence, after the Second World War (Herbst 2014). In this dissertation, I black box some these broader theories, and look at the state consolidation process at a smaller spatial scale, asking why states choose to extend their control to one region over another, rather than why they engage in this process at all.

Specifically, I argue that state leaders *actively* and *purposively* decide where (and when) to allocate resources in an effort to develop and extend basic mechanisms of control. There are several reasons to believe that the distribution of control across a polity is deliberate, the first of which is efficiency. Power is expensive to broadcast over distance, and capital—both political and fiscal—is scarce. Rational rulers should therefore choose to allocate resources to control territory where they

5. I use the term "state consolidation" interchangeably with "state incorporation," "state building," and "state formation."

expect net positive, or at least net neutral, returns on their investment. Second, control seems to correlate with factors that are not stochastic. Herbst (2014) argues that control is positively associated with population density. Fearon and Laitin (2003) argue that control is negatively correlated with rough terrain, as inhospitable geographies pose logistical and supply-chain problems for the state's security apparatus. There is also some consensus that patterns of control in Africa and elsewhere tend to follow an urban-rural divide, and that control is more concentrated around economically important road networks and transportation hubs (Herbst 2014; Müller-Crepon et al. 2021). Finally, control is necessary to state survival. The collapse of a central government's authority, and its loss of monopoly over the means of coercion and taxation is one of the key indicators of state failure (North 1982; Bates 2008).

I argue that this allocation decision represents a straightforward spatial selection problem, in which leaders select a subset of spatial entities, which maximize (a) the expected rents generated, (b) the strategic advantage derived from controlled territory, and (c) the accessibility of this territory, subject to some cost constraint. In subsequent periods (typically the state's fiscal year), leaders revisit this decision, and re-optimize based on their updated beliefs about the benefits of retaining or extending control over specific territory, and the costs associated with controlling that territory. In Chapter 4, I flesh out this theory in a bit more detail, and examine the social and political environments that are most conducive to state control.

1.4 Outline of Dissertation

The dissertation proceeds in two parts. The first part explores variation in territorial control across the African continent. I leverage publicly available geospatial data on the locations of government facilities in Africa to estimate the extent of de facto state control in 52 African countries—to essentially map areas where the state exists, and areas where the state is effectively absent. In Chapter 3, I use this new measure to empirically test some of the hypotheses generated by Jeffrey Herbst in his seminal 2000 study, *States and Power in Africa*. In Chapter 4, I attempt to explain this variation by asking which geographical, demographic, and economic characteristics influence a state's decision to control a particular piece of territory. The chapter then employs a set of supervised machine learning algorithms and other geospatial techniques to predict where states will elect to locate their assets and materiel. I find that African governments are most likely to control areas that are either strategically important to the survival of the incumbent regime, or areas with a high density of economic productivity and market activity.

The second half of the dissertation looks explicitly at the effects of living in state-controlled territory. In Chapter 5, I evaluate African's perceptions of traditional authorities (e.g., tribal chiefs and associated institutions) and the influence that these authorities wield in local communities. Using geolocated data from AfroBarometer Round 8, I find that perceptions of traditional authorities tend to be more positive, and the influence of these authorities stronger, in unincorporated regions than they are elsewhere.

In Chapter 6, I examine rates of common endemic diseases in Africa, and find that individuals living "inside" the state exhibit lower rates of malaria infection than those living "outside" of the the state. This is true even in some of the comparatively weak states of sub-Saharan Africa, suggesting that there are tangible welfare benefits to living in state-consolidated territory. Surprisingly, I find that these benefits do not exist in the case of HIV-AIDS; rates of HIV tend to be systematically lower in areas outside of state control. I attribute these differential disease dynamics to anti-HIV stigma. Because HIV-vulnerable individuals fear government sanctioning for their often-illicit high-risk behaviors, transmission rates will tend to be higher in state-consolidated areas.

1.5 Works Cited

Agnew, John. 1999. "Mapping Political Power Beyond State Boundaries: Territory, Identity, and Movement in World Politics." *Millennium* 28 (3): 499–521.

Bates, Robert, Avner Greif, and Smita Singh. 2002. "Organizing Violence." *Journal of Conflict Resolution* 46 (5): 599–628.

Bates, Robert H. 2008. "The Logic of State Failure: Learning from Late-Century Africa." *Conflict Management and Peace Science* 25 (4): 297–314.

Bean, Richard. 1973. "War and the Birth of the Nation State." *The Journal of Economic History* 33 (1): 203–221.

Branch, Jordan. 2013. *The Cartographic State: Maps, Territory, and the Origins of Sovereignty.* Cambridge University Press.

Cohen, Youssef, Brian R. Brown, and A. F. K. Organski. 1981. "The Paradoxical Nature of State Making: The Violent Creation of Order." *American Political Science Review* 75 (4): 901–910.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75–90.

Goffman, Erving. 1966. *Behavior in Public Places: Notes on the Social Organization of Gatherings.* Reissue edition. New York: Free Press.

Herbst, Jeffrey. 1996. "Responding to State Failure in Africa." *International Security* 21 (3): 120–144.

———. 2014. *States and Power in Africa: Comparative Lessons in Authority and Control - Second Edition.* Princeton University Press.

Jackson, Robert H., and Carl G. Rosberg. 1982. "Why Africa's Weak States Persist: The Empirical and the Juridical in Statehood." *World Politics* 35 (1): 1–24.

Krasner, Stephen D. 1988. "Sovereignty: An Institutional Perspective." *Comparative Political Studies* 21 (1): 66–94.

———. 2001. "Sovereignty." *Foreign Policy,* no. 122, 20–29. JSTOR: 3183223.

Lee, Melissa M. 2018. "The International Politics of Incomplete Sovereignty: How Hostile Neighbors Weaken the State." *International Organization* 72 (2): 283–315.

Levi, Margaret. 1981. "The Predatory Theory of Rule." *Politics & Society* 10 (4): 431–465.

Lust, Ellen M. 2022. *Everyday Choices: The Role of Competing Authorities and Social Institutions in Politics and Development.* Cambridge University Press.

Meagher, Kate. 2012. "The Strength of Weak States? Non-State Security Forces and Hybrid Governance in Africa." *Development and Change* 43 (5): 1073–1101.

Müller-Crepon, Carl, Philipp Hunziker, and Lars-Erik Cederman. 2021. "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa." *Journal of Conflict Resolution* 65 (2-3): 563–590.

North, Douglass C. 1982. *Structure and Change in Economic History.* Unknown edition. New York: W. W. Norton & Company.

Reno, William. 1997. "War, Markets, and the Reconfiguration of West Africa's Weak States." *Comparative Politics* 29 (4): 493–510. JSTOR: 422016.

———. 2001. "How Sovereignty Matters: International Markets and the Political Economy of Local Politics in Weak States." In *Intervention and Transnationalism in Africa: Global-Local Networks of Power,* edited by Thomas M. Callaghy, Ronald Kassimir, and Robert Latham, 135:197–215. Cambridge University Press Cambridge.

Rotberg, Robert I. 2004. *State Failure and State Weakness in a Time of Terror.* Rowman & Littlefield.

Scott, James C. 1999. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed.* 0 edition. New Haven, CT London: Yale University Press.

Skocpol, Theda. 1985. "Bringing the State Back In: Strategies of Analysis in Current Research." In *Bringing the State Back In,* edited by Dietrich Rueschemeyer, Peter B. Evans, and Theda Skocpol, 3–38. Cambridge: Cambridge University Press.

Tilly, Charles. 1975. "Reflections on the History of European State-Making." In *The Formation of National States in Western Europe,* edited by Charles Tilly, 3–83. Princeton, NJ: Princeton Univ. Press.

———. 1996. *European Revolutions, 1492 - 1992.* Oxford, UK Cambridge, Mass., USA: Wiley-Blackwell.

Tilly, Charles, Ernesto Castañeda, and Cathy Schneider. 2017. "War Making and State Making as Organized Crime." In *Collective Violence, Contentious Politics, and Social Change.* Routledge.

Wang, Shaoguang. 2003. "The Problem of State Weakness." *Journal of Democracy* 14:36–42.

Weber, Max. 2021. *Politics As a Vocation.* Creative Media Partners, LLC.

Chapter 2

Mapping Territorial Control in Africa

2.1 Introduction

In the early morning hours of Tuesday, 28 July 2015, I set off from the Eastleigh neighborhood of Nairobi on a six hour drive to Garissa, a small town in northeastern Kenya, roughly 200 kilometers from the Somali border. Garissa County is one of the most sparsely populated regions in East Africa. It is also one of the most insecure. Three months prior to my visit, four gunmen stormed Garissa University College, killing 179 people and sparking a low intensity conflict between the al Qaeda-linked militant group al Shabaab and the Kenyan government. Despite the threat of ongoing violence, travel through this region did not invoke the sense of isolation and vulnerability that one typically experiences in a conflict zone. Trappings of the state were conspicuous—the government-funded A3 motorway was in good repair, I stopped for lunch at a restaurant adjacent to the police barracks at Mwingi, and encountered no less than six security checkpoints along the route.

A decade earlier, I embarked on a similar journey from Mopti, in central Mali, to the city of Timbuktu. Like Garissa County, the Tombouctou Région is among the least densely populated regions of West Africa, and one of the most unstable. Tensions between Taureg separatists and Mali's central government had simmered for decades, leading to sporadic outbreaks of violence in Gao, Kidal, and Timbuktu. Yet, despite the obvious parallels to Garissa, the passage to Timbuktu was a markedly different experience. Tombouctou seemed, at least at the time, completely decoupled from Bamako, the political and economic nucleus of the country. Overland access to the city requires traversing poorly-maintained desert trails in a 4×4. While a single serviceable motorway (the RN33) does exist, transportation infrastructure is so poor in the region that most travelers opt for the five-day

Niger River voyage between Koulikoro and Koriome rather than traveling overland. Not once did we encounter a single government building or agent of the state—whether security officer, elected official, or public servant—either en route or within the city limits.

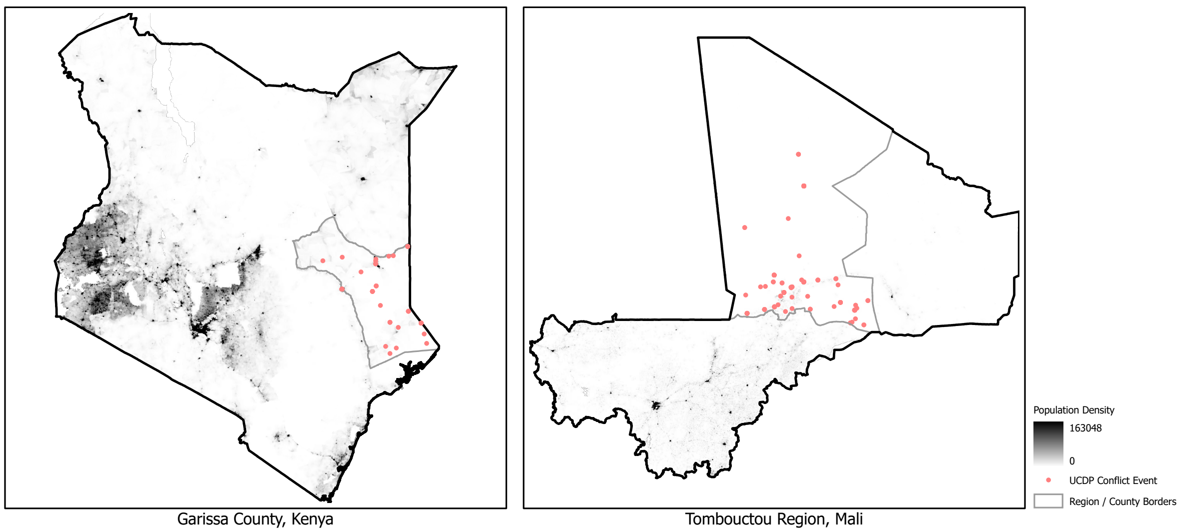


Figure 2.1. Comparison of Garissa County, Kenya and Tombouctou Région, Mali showing country-level population density (2020 estimates) and regional conflict events (2010–2020).

The differences between rural Mali and rural Kenya—two areas that are at least superficially similar by nearly all observable metrics—are quite palpable. Even to the casual observer, the grip of the state *feels* more untethered in Tombouctou than in Garissa County. These impressions are consistent with both the policy-oriented and academic literatures on state building in Africa. A 2007 RAND report prepared for the U.S. Air Force, for example, includes northern Mali and the broader Sahel as an example of "ungoverned space"—a region in which state infrastructure is absent and where the state is either unwilling or unable to execute basic functions (Rabasa et al. 2007). Garissa, on the other hand, is cited in a 2010 report by the Feinstein International Center at Tufts University as a "successful" example of state consolidation in Kenya, in which a once peripheral region was effectively "pacified" and brought under the direct control of the central government through a process of "villagization" during the 1960s and 1970s (Bradbury and Kleinman 2010, 20).[1]

Jeffrey Herbst (2014), in what is perhaps the most widely-cited study of state building in Africa, echoes these assessments, attributing differences in territorial control largely to aggregate

1. "Villagization" involves "the forced movement of pastoralists into 'protected' villages and the confiscation of their livestock" as part of the Manyatta Strategy implemented under President Jomo Kenyatta (Bradbury and Kleinman 2010, 20; see also De Waal 1997, 40).

patterns of population density in each country. Herbst classifies Mali as a "hinterland" country—a state whose geography is characterized by small areas of high population density, and large areas in which few people live (2014, 152). In Herbst's typology of African national design, these hinterland geographies pose significant challenges to the central government in its attempts to extend state authority over vast expanses of largely empty territory that are geographically removed from the capital. These challenges, he argues, explain the lack of state penetration in Tombouctou. Kenya, by contrast, is characterized by what Herbst refers to as a "neutral" geography, in which population density is dispersed, but not discontinuous (2014, 152). Hinterlands do exist in neutral states, as in the case of Garissa County, but these regions and their relatively sparse populations are not so far removed from the capital and other centers of state power that governance is untenable.

Like Herbst, most Africanists agree that there exists a great deal of variation in the patterns of territorial control that we observe across the continent. This is, in fact, the starting point of many of the classic works on state building in the African context. Yet, despite the emergence of an extensive literature that seeks to explain this variation (e.g., Jackson and Rosberg 1982; Boone 1998, 2003; Englebert 2009; Thies 2007, 2009; Herbst 2014), there have been few attempts to date to rigorously quantify it in a transparent, replicable, and valid manner—arguably the necessary first step in articulating any sort of theory to explain inter- and intrastate differences in territorial control, and to empirically test any hypotheses that emerge from that theory.[2] The focus of this chapter, then, is to measure and describe within-country and between-country spatial variation in territorial control across African states, in order to assess whether the perceived differences between Tombouctou and Garissa County—or between the Tambacounda Région of Senegal and Nigeria's Borno State—exist solely in the mind of the intrepid traveler, or whether there is some *systematic* evidence to support these perceptions. In what follows, I review the ways in which researchers have sought to measure territorial control and related constructs in the past. I then outline a new measurement strategy, rooted in the theoretical work of Mann (2012a, 2012b), Migdal (1988), and Scott (1999), which employs geospatial data on the physical presence of government facilities and other infrastructure throughout a state's territorial jurisdiction.[3] Section 2.3 describes the data

2. Two notable exceptions to this measurement deficit are Tao et al. (2016) and Müller-Crepon (2021), discussed in Section 2.2.1 below.

3. See Soifer (2008) and Soifer and vom Hau (2008) for a discussion of Mann's notion of "infrastructural power."

and methods I use to construct this measure. Section 2.4 presents the results of this exercise: a novel mapping of territorial control onto the African continent using cross-sectional geospatial data collected between 2018 and 2021.

2.2 Measuring Territorial Control

2.2.1 Existing Measures

Over the past several decades, social scientists have assembled a number of indirect measures of territorial control, primarily for use in empirical studies of civil war (Rueda 2017; Rubin 2020; Anders 2020), economic development, and public health (Koehnlein and Koren 2022). These measures typically fall into two broad categories. The first is a family of *state capacity* variables, such as tax revenue, government spending, and net exports, each generally measured at the country level.[4] Tax variables, including tax revenue per capita (Fearon and Laitin 2003) and tax revenue as a percentage of GDP (Besley and Persson 2008, 2009; Thies 2010), measure fiscal capacity, or the state's ability to extract rents and raise revenue from its population (Levi 1988). Tax variables are also positively correlated with a state's overall budget, which is in turn an indicator of the strength of the state's security apparatus and its administrative or bureaucratic capacity (Besley and Persson 2008; Cárdenas 2010; Dincecco and Prado 2012; Dincecco and Katz 2016). Variables such as total military personnel and military spending are used by Walter (2006, 2019), Gennaioli and Voth (2015), and Hanson and Sigman (2019) as a more direct proxy for the efficacy of state's security apparatus and the strength of its monopoly on force. Finally, resource wealth, which tends to augment the state's budget, is also a widely used measure of state capacity; Collier and Hoeffler (2004) employ the ratio of primary commodity exports to GDP, while M. Humphreys (2005) uses annual oil production and proven oil reserves as an indicator of "state strength" in their respective studies of civil conflict.

More recent work seeks to measure state capacity at the subnational level, using spatially variant indicators such as rough or mountainous terrain (Fearon and Laitin 2003; Hendrix 2011), distance to the national or administrative capital (Centeno 2002), and road or rail density (Acemoglu

4. McAdam et al. (2001, 78) define state capacity as the "degree of control that state agents exercise over persons, activities, and resources within their government's territorial jurisdiction," which is conceptually similar to the definition of territorial control I offer in Chapter 1. For more a more detailed discussion of the definitions of state capacity and associated measures, see Hendrix (2010), Cingolani (2013), Savoia and Sen (2015), and Hanson and Sigman (2021).

et al. 2015; Müller-Crepon 2021). Rugged topographies, such as jungles and mountainous areas, and remote regions with limited access to transportation networks (and thus other state-related facilities) are thought to be negatively correlated with state capacity, as governments have more difficulty penetrating these areas due to the high costs and complex logistics of access (Fearon and Laitin 2003, 80; Soifer 2008; Nunn and Puga 2012).

A second category involves a diverse set of *governance* indicators, or what Cingolani (2013) classifies as measures of "bureaucratic" or "administrative" state capacity. Governance broadly entails the ability of the state to excute policies and enforce laws (Fukuyama 2014, 9), which Skocpol (1985, 16) argues is only possible after the establishment of sovereignty and stable control over territory. Measures of this type include tax compliance (Benson and Kugler 1998; Besley and Persson 2008; Buhaug 2010; Dincecco and Prado 2012; Ottervik 2013; Wang and Hu 2015), levels corruption (Bäck and Hadenius 2008; Fortin 2010; Bersch et al. 2017), bureaucratic and regulatory quality (Williams 2021), and contract enforcement (Besley and Persson 2009). Governance is also measured using a number of expert-coded indices of government quality and performance; common indices include the Polity 5 and V-Dem datasets, and (in the African context) the Ibrahim Index of African Governance. Many contemporary analyses also incorporate a variety of macroeconomic indicators, such as GDP per capita, annual GDP growth, trade statistics, and public goods provision as broad measures of governance quality, as state capacity and territorial control are generally understood to promote economic development (Skocpol and Finegold 1982; Geddes 1994; Dincecco 2017) and the the provision of public goods (Tilly 1992; Acemoglu et al. 2001; Acemoglu et al. 2011; Besley and Persson 2009). It is worth noting that there are very few measures of this type coded at the subnational level in widespread use. Two exceptions are the G-Econ dataset (Nordhaus 2006), which measures gross economic output (gross cell product) at 1° resolution ($\approx 100 \times 100$km), and the geographic extent of malarial risk, which McArthur and Sachs (2001) and Sachs et al. (2004) show is correlated with development outcomes, and may thus be indicative of within-state variation in governance.[5]

Although these various indicators of state capacity and governance are widely used as proxies of territorial control, they suffer from a number of issues. First, both state capacity and governance are conceptually distinct from territorial control, though scholars and policy makers often conflate

5. See also Chang and Wei (2019) on the relationship between state capacity as proxied by natural resource wealth and Malaria risk.

these constructs. Territorial control, as I discuss in Chapter 1, is effectively a behavioral phenomenon, which many of these existing measures fail to adequately capture. State capacity variables like tax revenue or government expenditures, for example, contain little information on the state's ability to influence, direct, or restrict behavior in a given location. Governance indicators, such as tax compliance and contract enforcement, arguably do provide a more direct gauge of behavior, though they tend to capture the intensive margin of control rather than the extensive margin (particularly in geographically aggregated measures). Second, the abundance of variables used by scholars to operationalize territorial control and other closely related attributes of "stateness" makes it difficult to compare results (and mechanisms) across different studies (Hanson and Sigman 2021, 1496). The lack of consensus around a standard measurement strategy allows for the proliferation of ad hoc operationalizations with varying degrees of construct validity (Ottervik 2013; Thomas 2010), and many of these operationalizations are overly tailored to explain a particular outcome (Lindvall and Teorell 2016). Finally, the majority of these variables are coded at the national level, and only occasionally at the district level. This means that the most widely-used indicators of territorial control do not vary within the confines of a state's borders, obfuscating important subnational variation and limiting the types of empirical analyses that can be conducted.[6]

Cognizant of these issues, researchers have started to develop a new set of innovative measurement strategies to capture the concept of territorial control. Several of these are worth discussing, as they represent significant departures from the more traditional approaches outlined above. The first is a 2017 paper by Lee and Zhang, in which the authors operationalize Scott's (1999) concept of legibility—the "breadth and depth of a state's knowledge of its citizens and their activities"—as the accuracy of population age distributions reported in official national censuses (Lee and Zhang 2017, 119). While not an explicit measure of territorial control, Lee and Zhang argue that this is a valid measure of the central government's access to administratively useful information, which is necessary for the state to monitor its citizens and to enforce compliance with rules and regulations—two functions that are directly related to control. Tao et al. (2016) propose a more direct measure of territorial control in conflict zones in sub-Saharan Africa. These authors combine battle outcomes

6. Goodwin (2001) and Kalyvas (2006) underscore the theoretical importance of subnational variation in territorial control in explaining outcomes such as political violence; both authors argue that uneven patterns of territorial control influence the likelihood of revolution or the types of violence we observe.

with spatial data on terrain, population density, and transportation networks, and estimate the plausible service area controlled by the victor of a given battle based on the effective ease with which agents of the prevailing side can access surrounding territory as a function of time, cost, and distance. Although this measurement strategy has a number of appealing qualities, such as construct validity and subnational variation, its implementation is limited to areas of active conflict. This reduces the measure's utility in more general social science applications. In a similar vein, Müller-Crepon (2021) and Müller-Crepon et al. (2021) estimate the accessibility and connectedness of certain African regions from road network data culled from historical Michelin Atlases, and generate service areas similar to those created by Tao et al. (2016). This strategy has the advantage of functioning outside of active conflict zones; it also allows the authors the assess temporal as well as spatial variation in territorial control. Finally, Luna and Soifer (2017) use 2014 AmericasBarometer Survey data to measure state capacity at the local level. The authors construct a localized index based on a series of survey items that probe state reach, tax compliance, and property rights protections; the index can then be mapped to specific locations within each country.[7]

2.2.2 Territorial Control as Presence

In Chapter 1, I define territorial control as *the degree to which a state is able to influence, direct, or restrict the behavior of those people, agents, and organizations that exist within some arbitrarily defined territorial space within the polity.* Operationalizing this definition requires us to identify features of the state that allow the center to shape the behavior of its subjects. In recent decades, many wealthy countries have been able to achieve this objective through *indirect* means, by utilizing information and communications technologies (ICT) to broadcast power over distance. Social media, mobile telephony, and remote surveillance methods are used by countries such as China, Russia, and the United States to regulate behavior and enforce compliance among the population (Livingston and Walter-Drop 2014; Warren 2014). These emergent technologies greatly reduce geographic and social distances, minimize the information and transaction costs associated with the control of outlying territory (and, in some cases, extraterritorial jurisdictions), and more generally allow the state to maintain a "virtual" foothold throughout the polity—a phenomenon that Rosenau and

7. Driscoll and Seese (2023) use a similar approach to map the extent of government control in 2012 in Mogadishu, Somalia.

Czempiel (1992) refer to as "governance without government."

At its core, though, territorial control is most efficiently achieved through *direct* contact and intervention—in particular, the strategic placement of state agents, materiel, and other physical infrastructure in a given locale. Political and military leaders throughout history have long recognized the importance of locating state assets in regions they wished to control, and therefore invested considerable resources in developing, supplying, and staffing "outposts of the empire." The Romans, for example, established settlements and military fortifications as far away as Newcastle (then known as Pons Aelius) in the northeast of England, and at Carthage in present-day Tunisia. Kublai Khan moved the capital of the Mongol Empire from Karakorum (Övörkhangai Province, Mongolia) to the more centrally-located Khanbaliq (present-day Beijing) in order to maintain control over territories captured during the defeat of the Song Dynasty in 1279 (Man 2012, ch. 6). In the late 19$^{\text{th}}$ Century, cities such as Nairobi, Luanda, and Cape Town were settled by Europeans to ease the administration of the African colonies.[8] By establishing an enduring presence in a given region, the state (or the metropole) is in a position to directly monitor behavior, enforce edicts, and extract rents.

Why is presence so central to the exercise of control? A well-developed literature in sociology suggests that (co-)presence tends to induce behavioral change, whether or not agents of the state take active steps to promote or induce these changes (Mead 1934; Cooley 1956; Goffman 1966; Zhao 2003).[9] Goffman (1966, 243), for example, finds that the mere occurrence of an observer is enough to alter behavior in ways inconsistent with the observee's motives or intentions, whether or not the observer and the observee actually interact in any meaningful way. This behavioral change is not necessarily provoked by coercion or fear of punishment, but rather by the observee's attempt to manage situational norms or "proprieties." In order to avoid the appearance of deviance, Goffman (1966, 4) explains, individuals will engage in "approved" acts and refrain from "acts that are felt to be improper."[10] One of the primary effects of presence, then, is to make the dyadic other (i.e., the

8. See Herbst (2014, 73–80) and Crowder (2023) for a discussion of European administrative presence in Africa in the wake of the 1884–1885 Berlin Conference.

9. Subramaniam et al. (2013, 480), drawing on work by Goffman (1964, 1966), define co-presence as "co-location in space–time that allows for instantaneous and reciprocal human interaction." An interaction, in turn, is "an environment of mutual monitoring possibilities, anywhere within which an individual will find himself accessible to the naked senses of all others who are 'present,' and similarly find them accessible to him" (Goffman 1964, 135).

10. According to Goffman (1966, 5), the propriety or impropriety of an act is determined by the judgement of a particular social group. While there may be some dissensus on what constitutes an improper act, basic rules of conduct are generally "few and clear."

state) more salient, while the lack of presence leads to diminished awareness of the state-as-observer, uninhibited behavior, and reduced responsiveness.

While presence itself may be sufficient to induce state-sanctioned behaviors among the populace, it also serves a positive function—it allows the state to actively monitor its citizens, to employ coercive force to compel compliance with prescribed behaviors, and to mete punishment for any infractions.[11] To implement these tactics, nearly all modern states seek to extend the presence of the central government throughout their de jure territories. This is a state building strategy that Coleman (1977, 3), refers to as "political penetration," which he defines as the steps taken by a state to "establish an effective and authoritative central presence through its geographical and sectoral peripheries, and acquire a capacity for the extraction and mobilization of resources to implement its policies and pursue its goals, however these may be determined." Because physical manifestations of the state are so integral to both the state's ability to influence behavior and the state's broader goals of maximizing internal and external sovereignty (Coleman 1977), I operationalize territorial control as the *physical presence of state agents and infrastructure in a given region.* This type of operationalization seems to be gaining some traction with social science researchers. Schönholzer and François, for example, in a 2023 working paper, measure early state formation using geospatial data on ancient government-affiliated buildings from the Atlas of World Archaeology and the Seshat Global History Databank. Similarly, Jensen and Ramey (2020) and Rogowski et al. (2022) measure historical state capacity in the United States using longitudinal GIS data on the locations of U.S. Post Offices.

2.2.3 Conceptual Issues

A presence-based operationalization of territorial control has a number of desirable qualities. First, it provides a reasonable proxy for the behavioral aspects of control—if the state's ability to influence, direct, or restrict behavior is largely determined by the presence of, or proximity to, agents of the state, we should expect a strong positive correlation between between the location of state assets and the degree of control exercised by the central government. Second, because the locations of state assets can be represented by discrete points in space, it is possible to aggregate these points

11. See, for example, Goffman's (1966, 22–23) discussion of "public order." He notes that co-presence is generally sufficient to regulate behavior, though adherence to basic rules is reinforced by police authority.

and generate a single measure that varies across a country's territory. While a presence-based operationalization captures the behavioral dimensions of territorial control, it does tend to obscure important aspects of "stateness," such as state capacity and Weberian legitimacy (Weber 2021), which are implicit in other measures. Additionally, while presence can be measured at any arbitrary territorial unit—allowing researchers to assess subnational *spatial* variation—it does fail to account for *temporal* variation in territorial control. I discuss each of these limitations below.

State Capacity

While most definitions of state capacity do emphasize the behavioral and territorial aspects of control (see, for example, the definition offered by McAdam et al. (2001) in Footnote 4), state capacity is a complex, multidimensional concept involving attributes ranging from coercive and bureaucratic capabilities to fiscal and productive resources (Hendrix 2010; Cingolani 2013). To draw a clear distinction between control and capacity, I define state capacity as *the financial and political resources available to state leaders to carry out the administrative and service functions of a state*. Under this definition, state capacity is understood to be fungible across different domains. Leaders may choose to allocate resources to control, but also to public goods provision, entitlements, and other development projects (e.g., "guns vs. butter"). State capacity effectively represents the state's budget constraint in allocating control. As state capacity increases, leaders are able to invest more resources into the control of territory; some simultaneity therefore exists between state capacity and territorial control. While increased state capacity should allow leaders to increase the amount of territory controlled, greater territorial coverage should also increase aggregate levels of state capacity, as extending coverage should theoretically result in a larger and more compliant tax base.

Compliance, Participation, and Legitimacy

As I discuss in Chapter 1, the primary observable effect of territorial control is behavioral change among those subject to the state's jurisdiction. One factor that determines whether a state is able to effect behavioral change is an individual's propensity towards compliance. Indeed, Migdal (1988, 32) argues that the extent of a state's control is reflected in three closely-related indicators: compliance (conformance to state demands), participation (voluntary use of state-sanctioned institutions), and legitimation (acceptance and approval of state rule as "appropriate" or "right"). To the extent that an individual is disinclined to comply with government edicts, there is very little that the state can do to induce or prevent certain actions, short of exercising force. Because civilian compliance may be low even in state incorporated territory, behavioral measures of control are largely untenable. This is the primary reason that my conceptual definition of control focuses on the state's *potential* to alter behavior, rather than a more direct behavioral outcome.

Legitimacy is one factor that influences how compliant or participatory individuals are likely to be; populations that view government at legitimate are more likely to acquiesce to laws and mandates put forth by the center. Additionally, legitimacy allows states to effectively steer the activities of individuals and organizations without the necessity of constant coercion (Wang 1995, 89). Unfortunately, presence cannot directly capture individual compliance or the degree of legitimacy governments hold among their citizens. Rather than incorporating compliance and legitimacy directly into my operational definition of control, I examine them as outcomes that are potentially explained by whether or not an individual lives in state-controlled territory. Chapter 5, for example, looks explicitly at the legitimacy of both traditional authorities and the central government in state incorporated and unincorporated areas.

Temporal Variation

The data described in Section 2.3 below are cross-sectional, and thus do not allow me to explore temporal variation in territorial control. This obviously limits the types of analyses that can be conducted using this particular measure. However, increasing interest in open source and community mapping over the past five years have made it significantly easier to obtain geolocation data on the presence of state infrastructure in many countries. Going forward, it should be possible to generate a yearly measure territorial control using a variety publicly available data products. The cross-sectional data and analyses in this dissertation can therefore be considered a "proof-of-concept" for subsequent research.

What is more pertinent in the context of this dissertation is *short-term* temporal variation in territorial control. Goffman (1966) and L. Humphreys (1975), for example, argue that a single physical space can take on different characteristics at different times of day. The social interactions that are able to take place are dependent on which actors are present. If the degree to which a which a state can influence, direct, or restrict behavior varies over time within a 24 hour period, this potentially poses some challenges my measurement strategy. These challenges are mitigated, however, by two considerations. First, state infrastructure tends to be a durable reminder of state-presence, whether or not agents of the state are physically present at a given time. Second, none of the outcome variables I explore in this dissertation are the result of very short term (<24 hour) processes.

Endogeneity

The most pressing conceptual issue with a presence-based measure of territorial control is endogeneity—are state agents and infrastructure present in a given locale because the state is opting to situate its assets in territory it already controls, or does control stem from the placement of assets in a given locale? It is unfortunately not possible to resolve this issue using the data I have at my disposal, so I remain cautious about the claims I make throughout the dissertation using this particular measure. I would also note that this endogeneity concern is not unique to my measurement strategy; nearly all of the existing measures described in Section 2.2.1 share this issue.

2.3 A Presence-Based Measure of Territorial Control

2.3.1 State Infrastructure

Presence involves the physical manifestations of the state in a given place. Formally, I define presence as the occurrence, or density, of various agents of the state in an arbitrarily selected locale. The state is understood to be "present" to the extent that agents exist in that location, or to the extent that they effectively exercise their respective functions in that location and those proximate to it. Agents of the state are a diverse set. As countries develop, they begin to take on a wider range of responsibilities and services. Concomitant with these expanded functions, individuals and entities employed by the center have come to include an array of bureaucrats and administrators from various professions, from public health practitioners and other service providers, to security personnel and tax assessors. I focus on the presence of those agents that are directly tied to the core functions of a state outlined by Scott (1999): security, taxation, and conscription. These agents, and their associated facilities, include:

Single or Multi-Use Government Facilities. These include all brick-and-mortar buildings that are owned or operated by the central government for any purpose. Single use buildings may include police headquarters, military bases, or national courts. These single use facilities may also include the offices and residences of territorial governors or other representatives of the national executive. The analogous edifice to a multi-use property within the United States is the Federal Building, which occurs in most state capitals and other primate cities. These buildings are in many ways the central government's "embassy" to the given locality, representing the center's interests, providing "consular" services to the civilian population, and serving as a focal point for central government operations in key locales.

I focus on general purpose government buildings for a couple of reasons, the first which is that they represent a significant outlay of financial capital for the central government, which gives some indication of the state's intent to control a specific area. The second reason is that they are a key component of presence. That agents of the state exist in a given location is a strong indicator of territorial control. Note that I deliberately exclude locations that are used primarily as the district-level offices of elected officials or political parties, as these offices tend to focus on constituent services, rather than the broader functions of the central government.

Courts and Magistrates. Judicial facilities, like the single and multi-use buildings described above, represent significant investments on the part of the central government. They also serve a key function in enforcing compliance with laws and regulations, and adjudicating disputes among citizens.

Post Offices. Post offices are perhaps the single location type included in this measure that that does not relate directly to the core functions of the state. I include these offices because they represent an enormous fiscal outlay for the central government. In many countries, postal services are either state-sponsored or de facto monopolies, due in large part to the logistical complexity inherent in collecting and distributing post. Physical locations for collection and sorting of mail need to be established, and these locations need to be serviced by government personnel. Transportation and logistics routes must be established, which requires detailed cartography and complex routing procedures to ensure service. Postal delivery also relies on a complex network of international treaties and alliances to function. It is worth noting that interagency cooperation between national post offices often functions even when other relations fail. The sheer capacity needed to maintain and keep track of postal offices scattered across a polity is a monumental task, which is one of the reasons that the post is so ineffective in most of the developing world. That these offices exist in even remote areas of a country's hinterland, however, is indicative of some basic level of government presence.

Police Stations. One of the key functions of a state is to provide order and security, and to ensure compliance with government sanctioned laws and regulations. States employ a variety of agents to accomplish these tasks, including national and municipal police forces. In areas that police forces maintain a physical presence, they are able to influence the behavior of those individuals in their immediate vicinity through either coercion or the threat of force and punishment. Police are also required to remove unwanted persons, capture unwanted goods and capital, and prevent access to illicit information. Moreover, their very existence often serves to deter these flows, as proximity increases the probability of interception or capture.

Military Bases. Like the police and other domestic security forces, militaries and their bases extend the reach of the governmental power through their role as a coercive or violent entity. Militaries tend to fall under the strict purview of the central government, though it is possible that paramilitaries run by municipalities, or similar organizations loosely affiliated with the central government (e.g., the Wagner Group) will exist. In many jurisdictions, the role of the military is outward looking—that is, their mandate is to project force outside of or along the polity's borders. In other countries, however, militaries also have the authority to enforce domestic laws, and are often used to control civilian populations in territory that is otherwise difficult for the state to access.

Border Crossings, Internal Checkpoints, and Roadblocks. These facilities are designed explicitly to control the movement of migrants, foreign persons, and illicit goods into and throughout state territory.

Land Use and Agricultural Administration Offices. These bureaucracies manage physical territory, allocate state and private property to specific purposes, and ensure the property rights of land users by maintaining cadasters and adjudicating land disputes. Such offices also play an important role in state extraction. In the United States, for example, the U.S. Bureau of Land Management assigns mining rights on public lands, as with alluvial gold in national parks or natural gas in federal land holdings. Land use agencies also ensure that extraction is done in a sustainable manner, and that natural commodities are used efficiently, in line with Olson's theory of the stationary bandit—the goal of most governments is not total extraction, but instead to maximize the total extraction in each round (Olson 1993). In this sense, these types of offices manage the extractive potential of the state, and preserve its core economic interests.

Identity Card, Passport, Licensing, and Birth Registration Offices. Identification bureaus are necessary in the surveillance and monitoring of civilian populations. Administrative offices such as birth registries and passport bureaus not only tabulate the civilian population in the interval between national censuses, but they also have the power to decide which persons are eligible for state services and entitlements. These offices effectively determine who is "inside the state" and who is excluded. In many countries, national identity cards or internal passports are required for movement within the country, employment, and welfare benefits.

Tax assessors. Tax assessors serve to both extract rents on behalf of the state, and to catalogue and monitor citizens and corporate entities within the state's jurisdiction. While their primary function is to collect taxes and other rents for the central government, these bureaucracies are also tasked with understanding which persons and enterprises are present in a particular area in order to effectively levy taxes, as well as regulating economic activity to ensure accountability and minimize tax evasion.

2.3.2 Data

The data presented in this chapter are assembled from both commercial and public sources, collected over a period of four years. Primary sources include the Google Maps platform and Open Street Map; auxiliary sources include Foursquare, Bing Maps, and Wikimapia. Data scraped from these sources were reconciled to remove duplicate observations, and to merge similar observations. Steps were taken to validate the data using publicly available databases, including lists and websites published by the news media, national and local governments, and private citizens. Given the inconsistent and unreliable web presence of many African state agencies, however, the overwhelming majority of locations could not be verified by a secondary source.

The raw GIS dataset includes 19,419 individual "control points," which are mapped in Appendix Figure 2.12. Roughly 36% of control points are classified as security infrastructure, which includes police stations, military bases, and other emergency services. 32% involve non-security-related government offices of varying types, 22% are post offices, and an additional 4% of locations consist of judicial facilities such as courts, magistrates, and prisons. Appendix Figure 2.10 provides a more detailed breakdown of the distribution.[12]

The data are presented graphically as a modified kernel density surface, plotted in Figure 2.3. To generate this surface, a smoothed line is fit over each control point, i (i.e., each location in Figure 2.12), and the surrounding density is estimated with:

$$\text{Density} = \frac{1}{\text{Bandwidth}^2} \sum_{i=1}^{n} \left(\frac{3}{\pi} \left(1 - \left(\frac{\text{Distance}_i}{\text{Bandwidth}} \right)^2 \right)^2 \right) \tag{2.1}$$

Where "Bandwidth" is an optimal search radius defined by Silverman (1986), and "Distance" is the distance between the surface cell to be estimated and control point i. The local density value is highest at a given control point, and gradually decays over space in a radial manner, which accounts for the leopard–print pattern that characterizes the map in Figure 2.3. For ease of interpretation, estimated densities are then multiplied by the cell size (1km^2) to provide an estimated count of the number of control points in a given cell. The resulting surfaces are then stitched together into a single high-resolution raster to provide a continuous estimate for the entire continent. Figure 2.3 shows the results of this estimation procedure; the figure is discussed at length in Section 2.4.

Density estimation is carried out on a country-by-country basis in order to prevent control point locations in one country from influencing density estimates in a neighboring country. [13] The

12. This distribution is similar to that of other developing contexts; analogous data from Mexico is comprised of 38% security infrastructure, 44% government services, 16% post offices, and 2.5% judicial facilities. The inverse ratio of security facilities to government services in Mexico likely reflects the more stable security situation in that country, as compared to many African states, as well as a more robust government bureaucracy that employs a greater number of staff and requires more numerous facilities.

13. This modeling decision reflects the norm of Westphalian sovereignty, broadly conceived as non-interference into the domestic affairs of other states. African states generally respect this norm; instances of overt extraterritorial incursion are rare. Herbst (2014, 25) argues that the territorial boundaries between states provide a useful buffer mechanism that insulates polities from international pressures, though examples of foreign meddling and foreign intervention do exist.

estimation procedure therefore assumes that the control exercised by a given state ends at that state's official borders. By confining the estimation procedure to the formal boundaries of a given state, we are able to detect a set of political discontinuities that arise when transitioning from one country to another. Figure 2.2 depicts several of these discontinuities along the South Sudanese border. In quadrant B2, for example, we see an area of comparatively high control in the South Sudanese state of Northern Bahr el Ghazal, centered in the town of Aweil and the outlying communities of Nyamliell and Winejok. We see very little state presence in the neighboring Sudanese states (Wilāyat) of South Darfur and South Kordofan. Although Aweil is a relatively minor settlement situated in an area of low population density, this territory borders the disputed Abyei Area, which is effectively a formal condominium jointly controlled by South Sudan and Sudan on a temporary basis, pursuant to the Abyei Protocol of the 2005 Naivasha Agreement.[14] A speculative account of this discontinuity might point to the Sudanese government's inability to wrest control of large portions of South Kordofan from the SPLM-N. Active resistance from the SPLM-N and a recalcitrant population may make it difficult to situate state assets in the area. By contrast, the South Sudanese government has incentive to maintain a strong presence in the region to minimize conflict contagion, and to position itself to administer the disputed territory in the event of Sudanese withdrawal.

A similar discontinuity exists along South Sudan's border with Uganda. Quadrants C4 and D4 of Figure 2.2 show a high concentration of Ugandan control in northern Gululand, which radiates from the urban centers of Gulu Town and Arua. The densities we observe in northwestern Uganda may reflect the residual presence of state security forces deployed during Operation Iron Fist, at the height of the LRA Conflict in 2002–2005.[15] The extent of Ugandan control is delimited by the country's northern border; the data show relatively sparse South Sudanese presence in the frontier districts of Yei and Kajo Kaii in Central Equatoria. This international boundary reflects very real differences in the patterns of control exercised by each state, despite the similar physical geographies that exist on either side of the line, and the presence of trans-border communities such as Moyo and

14. The map in Figure 2.2 depicts the disputed Abyei Area and the Kafia Kingi Area as Sudanese territory.

15. The zone of high-density Ugandan control in quadrants C4 and D4 includes the Mount Kei White Rhino Sanctuary in the northwestern extreme of the country. The presence of endangered species, including the Black Rhinoceros, may also explain the high concentration of Ugandan security forces and other state infrastructure (e.g., the Ugandan Wildlife Authority) in the region. Renewed efforts to protect the Rhino and other species from poaching, coupled with the increased security associated with Western-focused ecotourism, have led to an increase in funding for security and wildlife services in Uganda. Unfortunately, it is not possible to confirm this theory with the existing data.

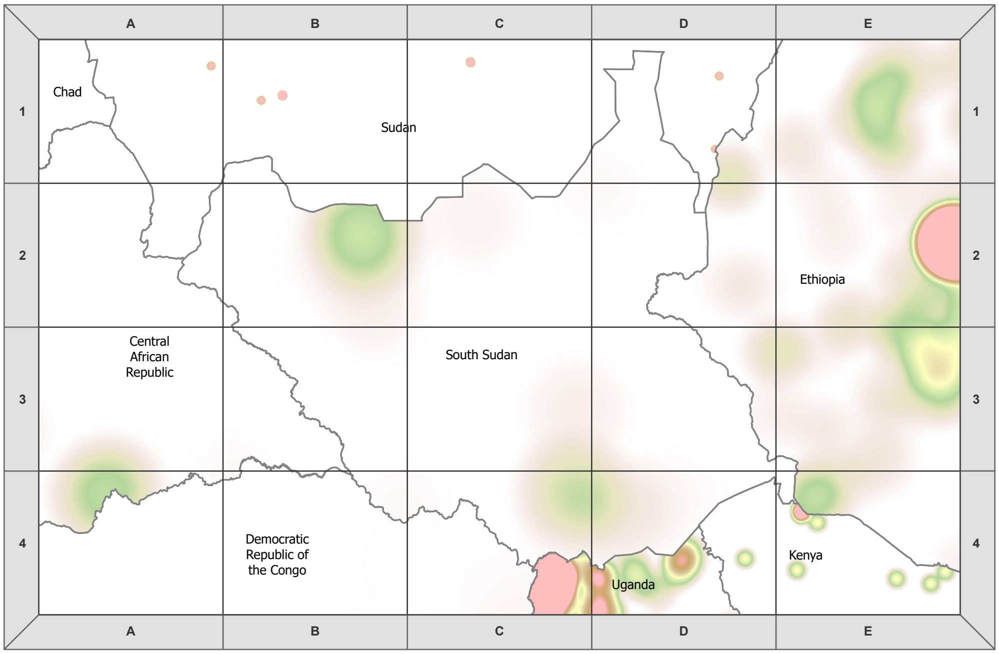


Figure 2.2. Detail of South Sudan's borders with neighboring states.

Nimule, whose populations intermingle across the extremely porous border in this region.

2.3.3 Data Issues

The central concern with this type of publicly sourced GIS data is missingness. Relevant points of interest may be excluded either intentionally (e.g., classified military bases) or unintentionally from source records. Because there is no census of state infrastructure on the continent, it is not possible to determine rates and patterns of missing data. It is plausible, however, to assume that most data are missing at random, as missingness is likely correlated with location. Publicly sourced geographic data, which is used by all of the sources scraped for this analysis, tend to suffer from what geographers call the "Starbucks bias." Points of interest, such as places of worship, commercial centers, and government buildings, are more likely to be captured by citizen contributors and recorded in databases such as Google Maps and OSM when they lie in close proximity to a Starbucks or to a similar outlet, as these establishments are concentrated in highly trafficked, affluent areas, where patrons are more likely to possess the technology necessary to catalog GIS coordinates and supplementary descriptives. Given the socio-demographic geography of many African states, the Starbucks bias effectively translates to an urban bias in the African context. We expect data from urban areas to be more complete and more accurate than data from rural areas.

Although missing data does pose certain challenges for analysis, there is reason to believe that any bias that results from missingness is minimal. The locations of government infrastructure are less likely to be excluded from the source data than other types of of location data. The importance of government services in the lives of ordinary citizens suggests that there is some incentive to make the locations of these types of facilities public. The majority of residents of a given county, for example, may require the services of a county identification card or tax office, while comparatively fewer residents are interested in the the café next door. Citizen contributors are therefore more likely to visit and to record the details of these locations as a form public good. Frequent directional searches and human traffic to these locations inform algorithmically driven data collection. Repositories of this information, including Google and OSM, seem to have an interest in government location data as well. Although Google and other corporate repositories keep their sourcing methods proprietary, and open source repositories are not usually specific about their data acquisition priorities, all repositories have remarkably specific tags for government locations (e.g., "Police Station" or "Government Identification Office"), suggesting that these locations are important. Additionally, patterns of missingness are likely themselves indicative of government control. Locations of state infrastructure are more likely to be reported if contributors are secure in their personal safety (e.g., they can geotag locations using their mobile devices without fear of theft of personal injury). Absence of data, then, may itself be indicative of absence of territorial control.

2.4 A Political Topography of the African State

In her 2003 book, Catherine Boone coins the term "political topographies" to describe variation in the centralization of state power across rural Africa. For Boone, this variation is the result of institutional configurations and the types of power-sharing relationships that exist between central and local governments. Boone is able to classify various stretches of West African territory according to the degree of state incorporation at two distinct time periods—the late colonial period through independence, and the 1960s through the 1980s—and map these political topographies as they existed during these periods.

This section is in many ways an empirical extension of Boone's work. Using the data described above, I am able to chart subnational variation in the centralization of state power at a highly granular level. The coverage includes the entire continent, and depicts nuanced variation that other estimates are unable to detect. Unlike Boone, however, what follows is a single, relatively recent snapshot of patterns of territorial control across Africa.

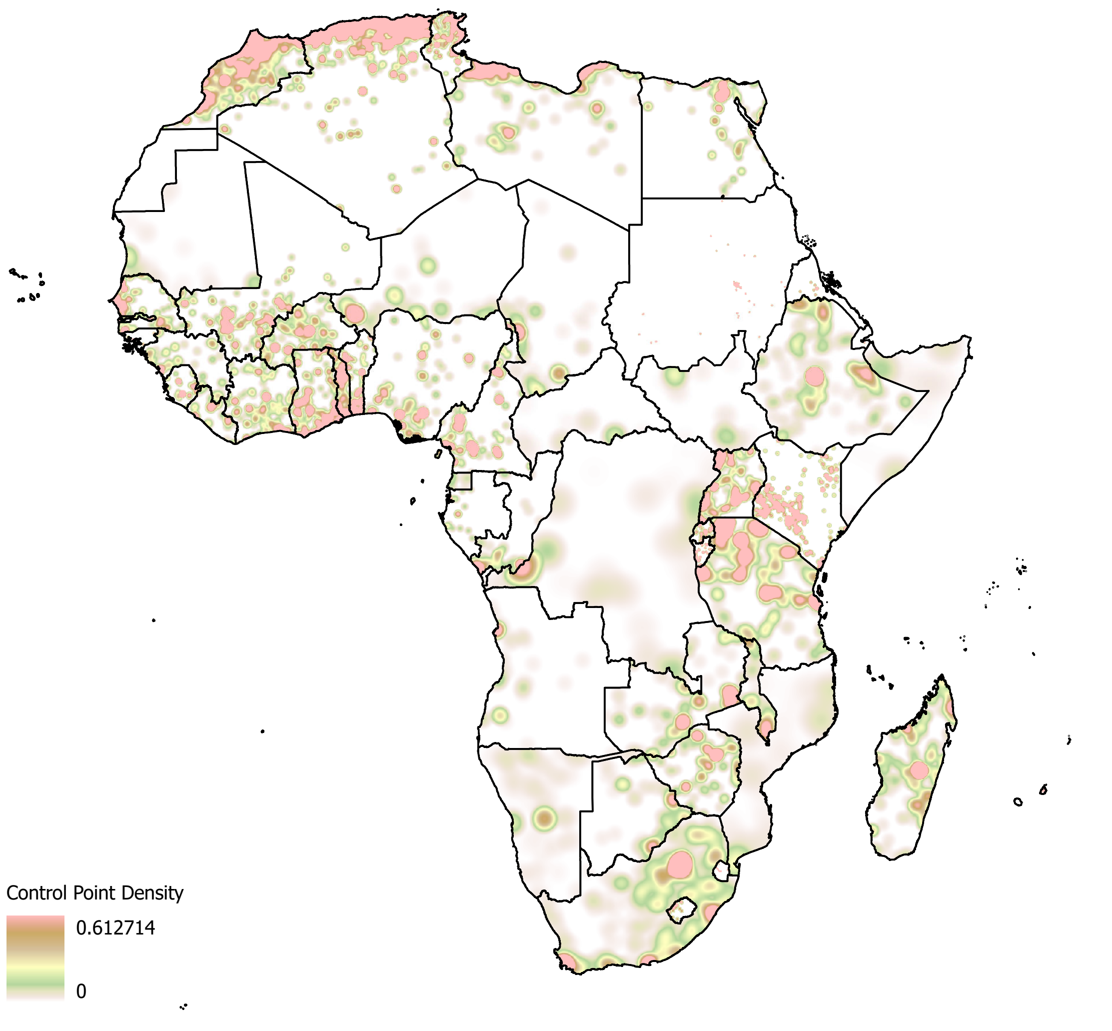


Figure 2.3. Control point density across the African continent. Each color interval represents roughly $\frac{1}{2}$ standard deviation.

2.4.1 The View from Space

The map in Figure 2.3 presents a graphical representation of the density surface described in Section 2.3.2. The map depicts the estimated count of control point locations in each 1km^2 cell for the entirety of continental Africa, with the exception of the disputed territory of Western Sahara.

Each color interval on the map represents a roughly $\frac{1}{2}$ standard deviation change in the overall distribution. Table 2.1 provides summary statistics for this raster data set. Note that even in high density settings, the expected count is less than one control point per square kilometer (≈ 0.61). This is itself an unexpected finding; even in the heart of Nairobi's Central Business District or Abuja's Three Arms Zone—the nuclei of national politics in Kenya and Nigeria—government infrastructure is still quite sparse on average. It is difficult to draw an appropriate comparison, but the average density of religious institutions and places of worship in Nairobi is 22.52 per 1km^2, and the average density of coffee shops in the city is 0.11 per 1km^2 (compare with the 19.86 Starbucks locations per 1km^2 in San Francisco). Places of worship are probably the closest analog. Unlike coffee shops and other retail outlets, places of worship are not necessarily in direct competition with each other for clientele, so spatial optimization is about providing adequate pastoral coverage for a given population rather than solving Hotelling's Game (Hotelling 1929).

Table 2.1. Territorial Control Raster Summary Statistics

Pixel Size	n, Cells	μ	σ	Min	Median	Max
1km^2	29031588	0.000544804	0.003927179	0.000000000	0.000004890	0.612714112

Another feature of these data worth underscoring is the intense right skew of the distribution (skewness = 62.5, kurtosis = 6465.6); the vast majority of cells have an estimated control point density of zero, and mean control point density across the continent is roughly 5 per 1000 square kilometers. Functionally, this means that there are vast swathes of the continent with little to no government presence. This is not surprising; scholars and policy makers are acutely aware of the governance deficit on the continent. What is notable, however, is the spatial distribution of these ungoverned spaces—they tend to fall within the *interior* regions of a state, rather than the border peripheries. Indeed, only the five Sahelian countries—Mauritania, Mali, Niger, Chad, and Sudan—and Congo-Kinshasa have large stretches of uncontrolled border territory, all along the northern borders, which fall either in the middle of the Sahara, or the middle of the jungles of the Congo basin. These are all regions were we might expect government presence to be low, given their remote and difficult locations, though these borders are not completely devoid of government presence. Outposts such as Bordj Badji Mokhtar and In Guezzam in Algeria, Abu Simbel in Egypt, and Bangassou in Central

African Republic seem to be positioned at strategic locations along these highly permeable borders—each of these areas of high control point density are positioned along a major transnational highways (RN19, N1, and N4 respectively). The other exception to this pattern is Angola, whose land borders are completely devoid of government presence.

The distribution of uncontrolled territory is somewhat inconsistent with Herbst's hypotheses. Herbst (2014, 49, 134) argues that, during both the colonial and early post-colonial eras, African governments focused on consolidating control over "core" areas of the state, by which he means primate cities in coastal and capital regions. Because the "hardness" of African borders preserves the integrity of the state, African governments could afford differentiated control in border regions. This, according to Herbst, is why African states did not invest in improving tax collection in these outlying areas (which, by extension, implies investing in physical infrastructure we would detect), or tying remote populations to the center with symbolic politics (134). The northern borders of the Sahelian states are probably the most direct example of Herbst's theory. The Sahara represents a very real buffer between Sahelian and Mediterranean countries. But Herbst's theory runs into a couple of shortcomings: First, it is unclear whether the border itself or the harsh conditions of the Sahara are responsible for the lack of government presence in this region. Second, the theory would predict *symmetric* absence of government control on either side of the border. A border, as a buffer, should insulate both states equally. Yet we see that higher capacity states, such as Algeria and Egypt, have attempted to control border choke points in the Sahara region. The same is true for a relatively low-capacity state, Central African Republic, whose territorial control seems to cover the extent of the country's southern border with Congo-Kinshasa and western borders with Cameroon and Congo-Brazzaville. Angola is also problematic—the buffer mechanism that Herbst speculates, and the apparent absence of state control along the country's borders, did not prevent external intervention during the country's decades long civil war. This was one of the most internationalized conflicts in contemporary African history.

Contrary to certain observations made by Herbst, we find that in a majority of African states, control seems to be concentrated along the perimeter of the jurisdiction. This is not necessarily a function of the placement of capital cities like Bangui, Kinshasa, Brazzaville, Lomé, N'Djamena, and Gaborone, all of which are adjacent to an international border. Nor is it necessarily a function of population density. Aside from the Great Lakes region, KwaZulu-Natal, the Kinshasa–Brazzaville metropolitan area, and the Accra–Lagos corridor, population densities of African countries tend not to be concentrated along borders. Yet in many instances, control point density is either higher or more pervasive in border regions than in a country's interior, or in capitals or other major cities. Central African Republic is case in point. Although we can clearly locate the outline of Bangui in Figure 2.3, control point density is highest in Bangassou—a relatively minor prefectural capital that is strategically situated on the north bank of the Mbomou River. Senegal, Liberia, and Uganda show similar patterns of high-density control over border regions, with relatively low levels of state infrastructure in the interior.

High control point density regions do tend to fall in highly populated areas—particularly those along the Mediterranean Coast, the Gulf of Guinea, and the Great Lakes. Major metropolitan areas are evident in Figure 2.3, though the data do not strictly reflect the high service density we expect in urban locales. In fact, only 18.47% of control point locations fall within 10 kilometers of a city with a population greater than 500,000, which effectively means that more than 80% of the observable state infrastructure in Africa exists *outside* of the exurban range of Africa's 100 largest cities.[16] In fact, average control point density by country tends to correlate more with common measures of state capacity than population size (see Appendix Figure 2.13).

Perhaps the most pervasive pattern of control point density across countries we see in Figure 2.3 is what we might describe as a port–and–outpost configuration, in which control is most heavily concentrated in port cities—many of which are located along coastlines or international borders—and in the occasional "outpost" of the country's interior. Namibia is the clearest example of this phenomenon. Control point density is highest at the coast, in the Atlantic port of Walvis Bay, and at the country's capital of Windhoek in the geographic center of the country. Note that the second largest city, Rundu, which is an important transit hub and the effective terminus of the Caprivi Strip, barely registers on the map. This pattern is not unique; we see it replicated in throughout the continent in countries like Mauritania, Angola, Congo-Brazzaville, and even to some extent in Nigeria and Kenya, where outpost cities are more prevalent than in some of the other examples.

16. Zanzibar City, Tanzania is the continent's 100th largest city, with an estimated population of 501,459. Figure 2.11 in the Appendix replicates Figure 2.4 for cities of different population sizes.

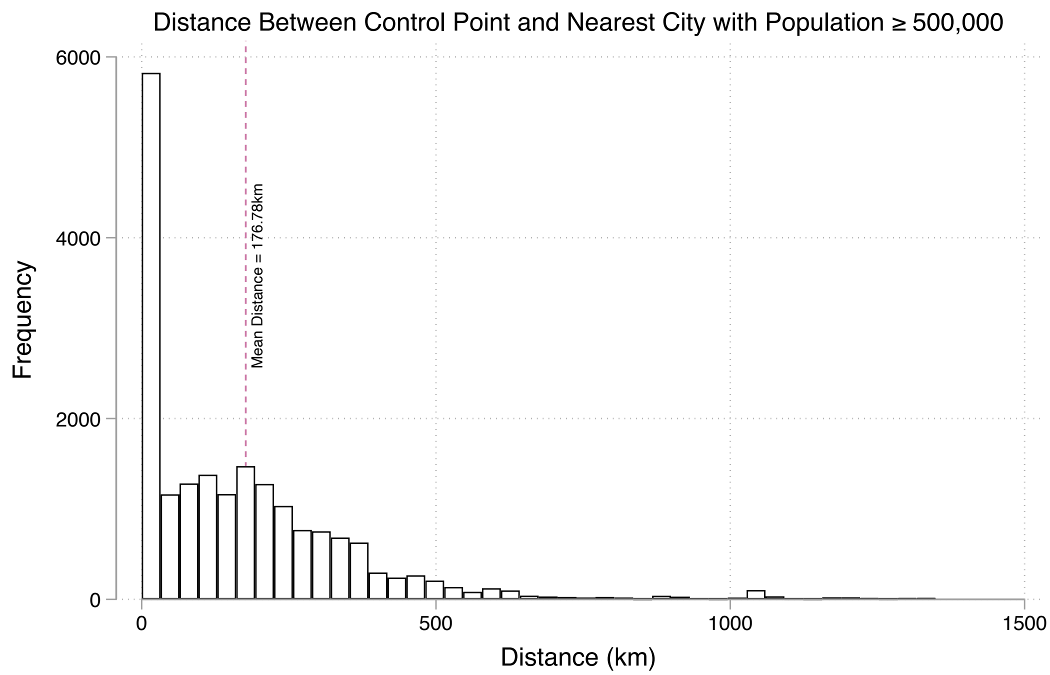


Figure 2.4. Frequency of control points by distance to nearest major city.

2.4.2 Zooming In

In this section, I drill down into four case study countries: Mali, Senegal, Nigeria and Kenya, and we pay special attention to Garissa County and the Tombouctou Région of Mali. These cases allow us to compare cross-country variation in the distribution of territorial control. Figure 2.5 maps control point density in each of the four countries on a single, harmonized scale. The most striking feature of this plot is that the degree of control (i.e., the proportion of covered territory) increases as the size of the country decreases, which is consistent with our theory—the costs of consolidating control should increase in the size of the country to the advantage of smaller states. Although land area is a reasonable explanation for the coverage of control, it does not explain the placement of control, why density exists in some places but not in others. The subsequent chapter looks at some of the correlates a bit more systematically, but in broadly descriptive terms, there are no geographic features to explain the differences in patterns of control that these countries exhibit.

In terms of physical geography, three of these countries—Mali, Nigeria, and Senegal—are predominately flat and covered in tropical savanna (Köppen Aw/As) or semi-arid steppe (Köppen

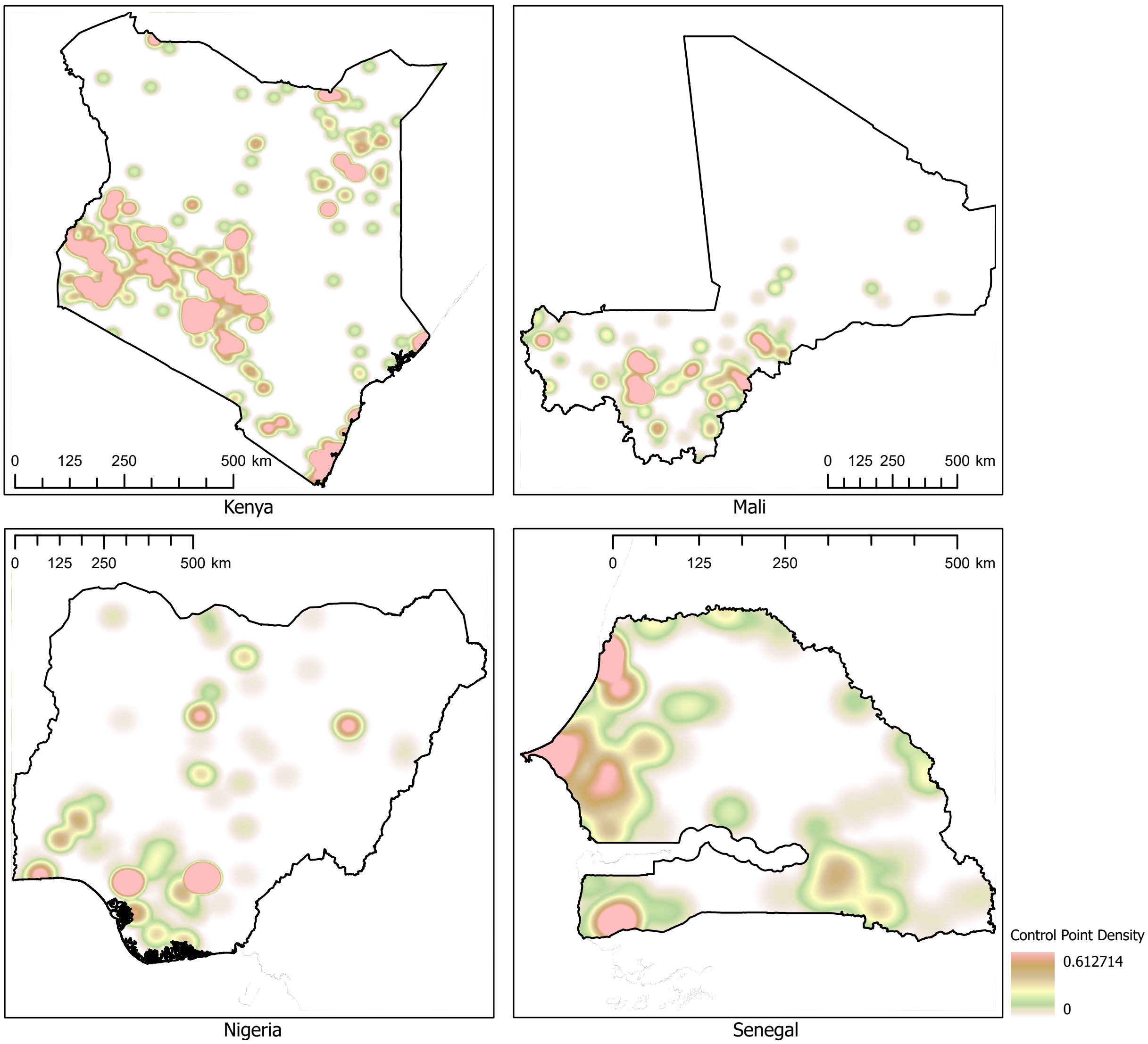


Figure 2.5. Comparison of territorial control across four countries with a harmonized scale.

BS), with small portions of hot desert climate (Köppen BWh) in the far northern reaches of Mali and Nigeria. Kenya has the most diverse physical geography of the four cases, ranging from a more temperate oceanic climate (Köppen Cfb) in the central highlands to hot desert (Köppen BWh) across the Awara and Ngangerabeli Plains in the northeastern Mandera, Wajir, and Garissa counties. On naïve visual inspection, there is no obvious correlation between control and either topography or climate. Areas of high control point density are just as likely to occur at high elevations (e.g., Nairobi, Kisumu) as low elevations (e.g., Lagos, Bamako). Extreme climates do not seem to affect the spatial distribution of control point density either—we see significant areas of control in the warmer desert climates of Kenya (e.g., the northern portions of Wajir County), Mali (Saharan villages throughout Kidal and Gao régions), and Nigeria (Kano, Kaduna, and Northeastern States). In Senegal, control point density is highly concentrated in the warm, arid regions north of the Gambia, rather than the comparatively verdant Casamance.

The lack of apparent correlation between control and environment seems to suggest that control is not strictly a function of physical geography. It is not, however, strictly a function of human geography either. While control does occur more frequently in densely populated areas, it is not coterminous with population centers. Nigeria and Mali provide the clearest examples of this among our four case studies. 22% of Nigeria's population lives within 100 kilometers of the Atlantic Coast; roughly one in eight Nigerians live within the Lagos conurbation alone. Yet 80% of the territory with greater than 1 standard deviation of control falls outside of the high urban density coastal region. 45% of high control ($\geq 1\sigma$) locations occur in locations with population density less than or equal to 100 people per square kilometer. Lagos is a particularly interesting example, despite being one of the largest cities in Africa, the the coverage and density of control is on par with cities like Kano and Port Harcourt, which are less than a tenth the size. In Nigeria, control is highly dispersed throughout the spatial population distribution: there is broad coverage on the extensive margin, but not deep coverage at the intensive margin. Mali exhibits the inverse pattern. If we take into account the significantly lower population densities as compared to Nigeria, we find that control point density substantially under-covers population hotspots, particularly in régions like Sikasso, Kayes, and the south and west border areas of Tombouctou (low coverage at the extensive margin). Yet, despite the comparatively low populations in even urban areas such as Bamako and Kayes, we see extremely high state penetration, on par with Nigerian cities with populations that are orders of magnitude greater.

Despite low spatial correlation with physical or human geography, Figure 2.5 does suggest some correlation between control point density and network geography. Control seems to be concentrated across central "node" cities across all four cases. These nodes tend to have two defining features. First, they are of great historical importance. These are either some of the earliest established cities in their respective regions, such as as Timbuktu, which became a permanent settlement in the 12th century, and Kano, which was an important waypoint on the trans-Saharan trade routes since the early 11th century, and Mombasa along the Swahili coast. Or they are important European settlements from the colonial era, such as Nairobi, which originated as a rail station in the 1890s, and Dakar, an important entrepôt of African-European trade dating back to the mid-1400s. The second feature is that these nodes exhibit a high degree of network centrality. They tend to lie at the intersections of major transit routes, or in areas of high density of arterial roads, and thus represent major choke points in the ground-based transportation network of each country.

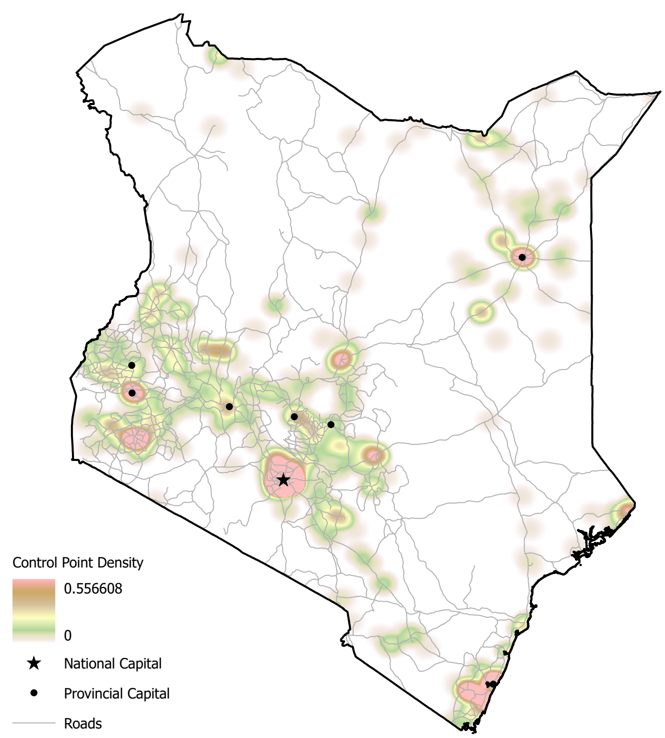


Figure 2.6. Transportation network, urban geography, and the distribution of territorial control within Kenya.

Figure 2.6 details the transportation network and urban geography of Kenya, and illustrates the importance of these nodes in the overall distribution of control density in the country. The most densely controlled regions are centered in historically important cities such as Mombasa, Kisumu, and Wajir[17] which are two of the oldest extant human settlements in East Africa, but also important colonial cities, such as Nairobi, Kisii, settled by British soldiers during the First World War (Kisii was originally called Bosongo (or Abasongo) which is the Abagusii word for "the place where white people settled"), and Isiolo, which originated as a military bivouac during the First World War. Mwingi is the one possible exception to this pattern. Mwingi is a minor village along the A3 motorway that stretches from Nairobi to Garissa. The village seems to have originated as factory town in the late 19th Century that produced chalk, whitening powders, and other mineral pigments. The area was surveyed by British colonial authorities as early as 1902 as a site of possible coal and garnet deposits (Walker 1903; Crowther 1957), and as a possible link in a planned railroad to British outposts in the the Northern Province along the border with Italian Somaliland. In fact, surveying and grading work for this railway was carried out as far as Mwingi before the project was ultimately abandoned. The village currently serves as a regional cattle market and the primary link between the former Northeast Province and the transportation hub at Thika. It is worth noting that, with the exception of Nairobi and Kisumu, none of these towns are major population centers, particularly in modern times, as rural Kenyans have migrated in large numbers to urban agglomerations such as Nairobi.

Every one of these nodes, however, is located at the intersection of major land-transport routes. In fact, the colonial rail hierarchies along the Kenyan trunk of the Uganda Railway are evident in patterns of control density throughout Kenya. Figure 2.7 overlays the control point density from Figure 2.6 onto a mid-century rail traffic density chart from the East African Railways and Harbours Corporation (EAR&H). Rail terminus points at Mombasa, Nairobi, and Kisumu display high levels of control point density, while intermediate stops at Mtito Andei (Tsavo), Nanyuki, Nakuru, and Kitale display lesser but significant degrees of control point density. As the primacy of rail diminished beginning in the 1960s, we see similar patterns begin to emerge in the arterial road network in Figure 2.6, particularly in the regions surrounding Lake Victoria and in the extreme northeast of the country.

17. Wajir is from the Borana word for "coming together," and refers to the fact that diverse semi-nomadic pastoral Somali clans would come to this location to water their herds. It is the capital of the former Somali Ajuran Sultanate, a empire that ruled over much of the Horn of Africa in the late middle ages, and dominated trade in the North Indian Ocean. In 1912, it became the British colonial headquarters in the region.

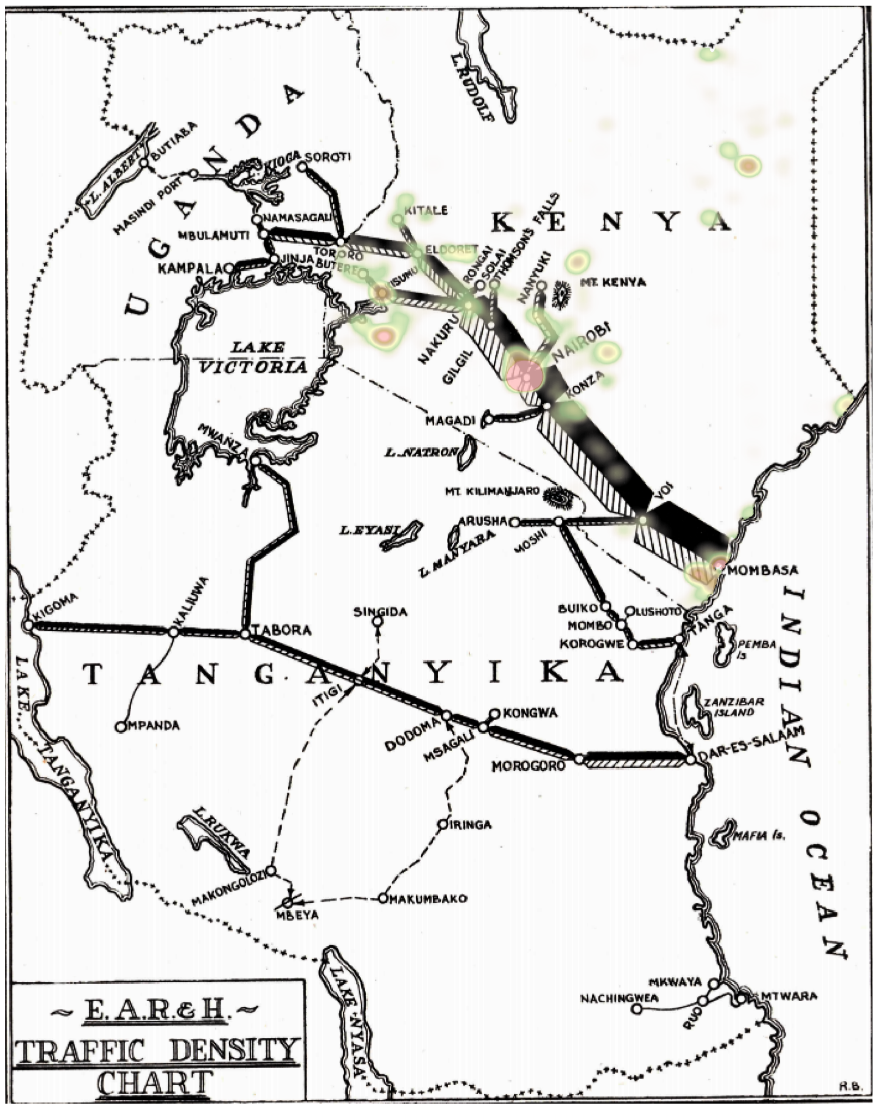


Figure 2.7. EAR&H map of Ugandan Railway (c. 1948) overlaid with Kenyan control point density from Figure 2.6.

Areas of high control occur at almost all major intersections of arterial roads in the country, regardless of whether or not this intersection is colocated with any significant population center. Wajir is case in point, despite its small population size, inhospitable climate, and distance from the capital, it lies at the intersection of three trunk roads (B9, C80, D570) and a half dozen other major roads that connect Wajir to Dif, Dasheq, and a handful of other settlements in northern Kenya. Wajir is also the location of the only non-military airstrip in the northeast corner of the country. The network density of arterial roads increases moving southwest, as does control point density, but interestingly, density is highest surrounding road network vertices with high valency—these are nodes with high number of road connections. But another pattern is noticeable as well—control occasionally occurs on roads in the absence of any network node, or at nodes with low valency. Examples occur along the A2 and B9 motorways at villages and permanent settlements that barely register on Google Maps and other archives (e.g., the source may catalog a police station, but not the name of the village). These high control point density areas, however, represent network choke points—or the only paved overland passage—to the hinterlands of Marsabit, Samburu, Isiolo, and Garissa counties.

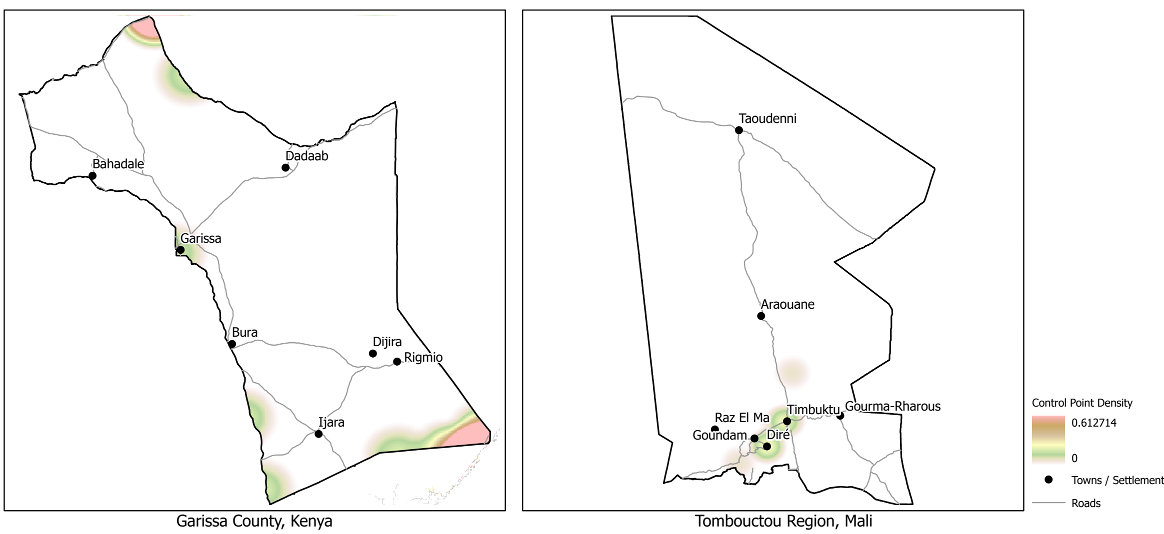


Figure 2.8. Territorial control in Garissa County, Kenya and Tombouctou Région, Mali.

As we zoom down even further to the county level, we find that these country level patterns do not scale down. Figure 2.8 shows control point density in Garissa Country and the Tombouctou Région on a harmonized scale. Other than high concentrations in the administrative capitals, there is no discernible pattern to the densities we see in these subregions. Indeed, it appears that most of the control in Garissa county actually radiates from villages in Tana River County to the west, and from the town of Habaswein in Wajir County. There is no overt correlation with nodes in the road network, or county-specific patterns of population density. The same is true of Tombouctou, though both the coverage and intensity of control point density is higher in Garissa than Tombouctou, which would seem to validate some of the claims made at the top of the chapter.

2.4.3 Generalized Patterns of Territorial Control

The discussion above focuses on the location of high density areas of control within a given polity. This section focuses on their spatial distribution. We can categorize patterns of territorial control according to two different dimensions: coverage and intensity. Coverage refers to the overall proportion of territory with some non-negligible government presence, while intensity refers to the concentration of control point density within these areas.

Table 2.2. State-Level Distributional Patterns

Intensity		Coverage: Low	Coverage: High
	High	Archipelago	Consolidated
	Low	Unconsolidated / City State	Lo-Fi

Table 2.2 classifies different distributional patterns according to these two characteristics. Low intensity – low coverage are what we might call "unincorporated" or "city state" countries. Examples include Somalia (particularly if you exclude Somaliland), Angola, and Mauritania. These countries exhibit very low levels of control within their borders, though one or two regions within the country—usually the capital or primate city—may exhibit higher densities than the surrounding territory. Such countries have weak governments centered in their capitals, and have little to no control outside of the city limits. Low intensity – high coverage countries have a greater extent of control within their borders, yet government is not strong enough to consolidate control over the entirety of the territory. This pattern is referred to as "Low-Fi" in Table 2.2 due to the low quality of control exercised by the government. Examples include Namibia, Mozambique, and Zambia.

High intensity – low coverage states are characterized by tight control in very limited locales. These areas of control resemble an Archipelago: islands of control in a sea of ungoverned territory. Sudan, Djibouti, and Eritrea all exhibit this type of pattern. The final category—consolidated states—are those countries that have both high intensity and high coverage. Examples include Tanzania, Uganda, and Togo. Figure 2.9 shows the empirical distribution of these patterns.

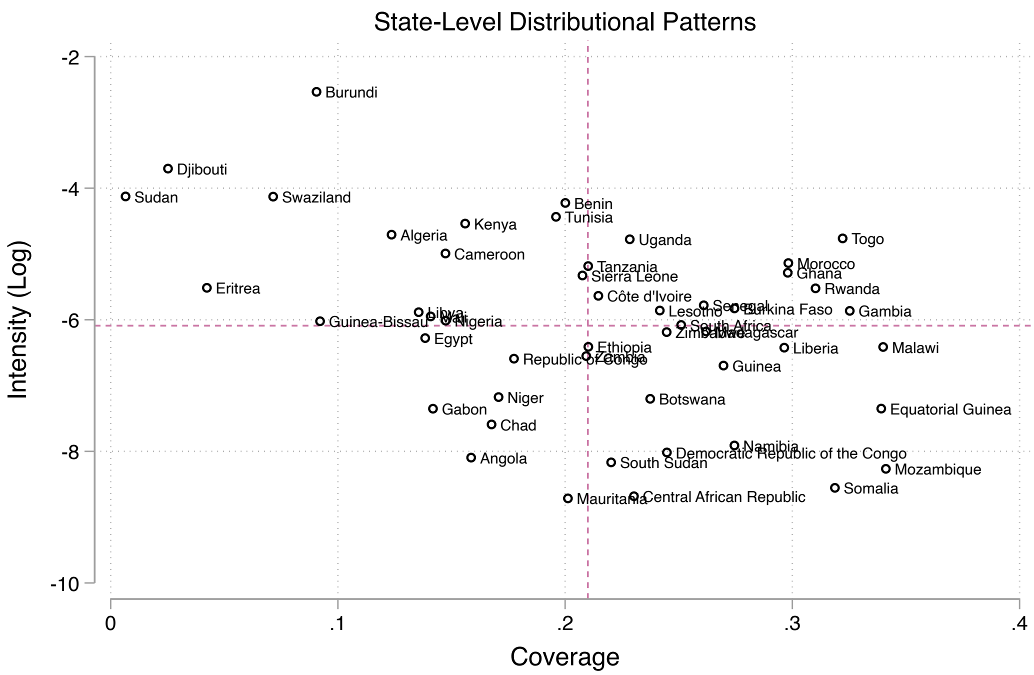


Figure 2.9. African states by coverage and intensity. Coverage is defined as the proportion of a country's total area with ≥ (country) average control point density. Intensity is the average control point density in "covered" regions.

2.5 Discussion

This chapter outlines a new measure of territorial control that is based on the presence of state-related facilities in specific localities. Using publicly-sourced GIS data from a variety of sources, I generate a cross-sectional measure of control point density that depicts the extent and intensity of territorial control across the continent between 2018 and 2021. Plotting this measure reveals a number of interesting spatial patterns in territorial control—the presence of marked discontinuities in control across international borders, and concentrations of territorial control along transportation networks and international borders. In Chapter 3, I explore these patterns in greater detail, by assessing how territorial control correlates with various geographic features, including population distribution, regional accessibility, and a country's extractive capacity.

2.6 Appendix

2.6.1 Control Points

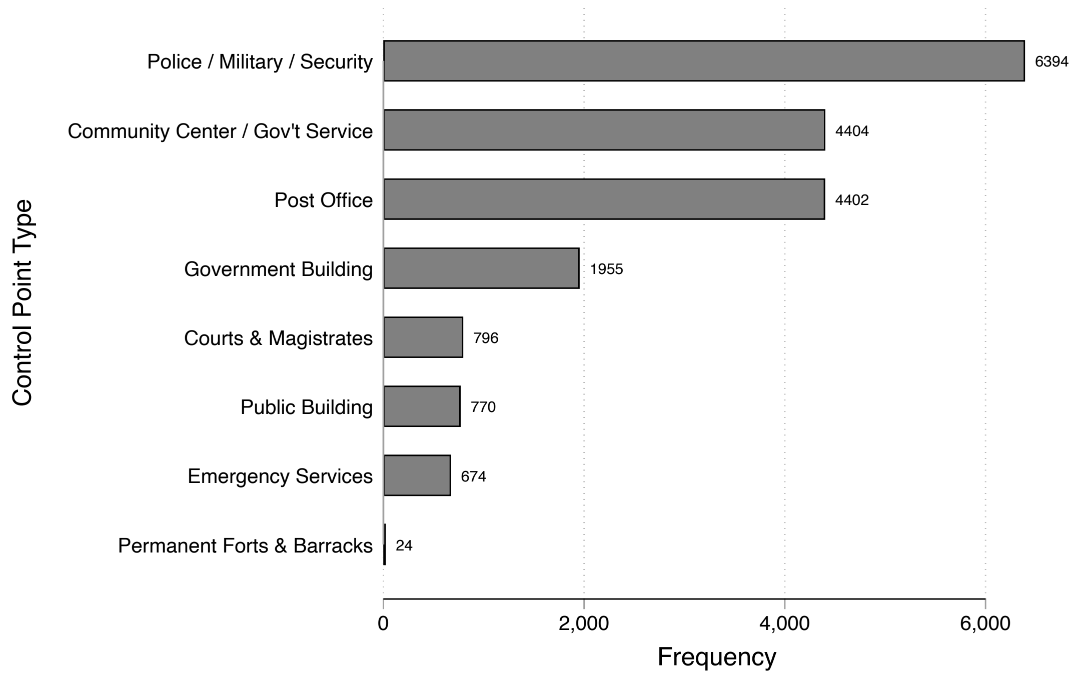


Figure 2.10. Frequency of control points by type.

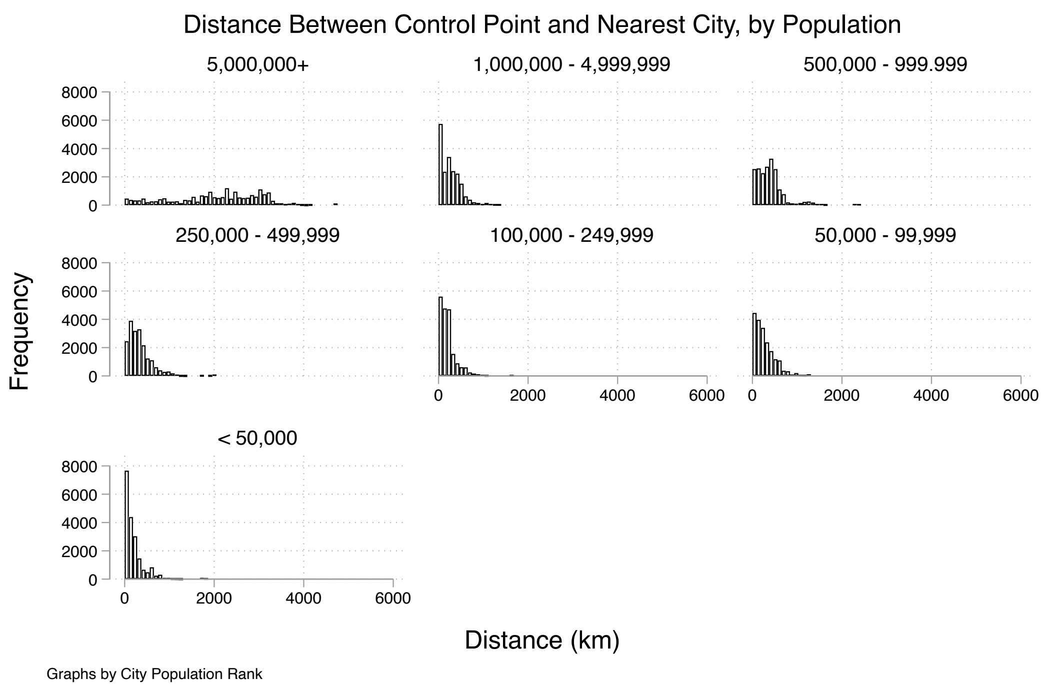


Figure 2.11. Frequency of control points by distance to nearest city of a population given size.

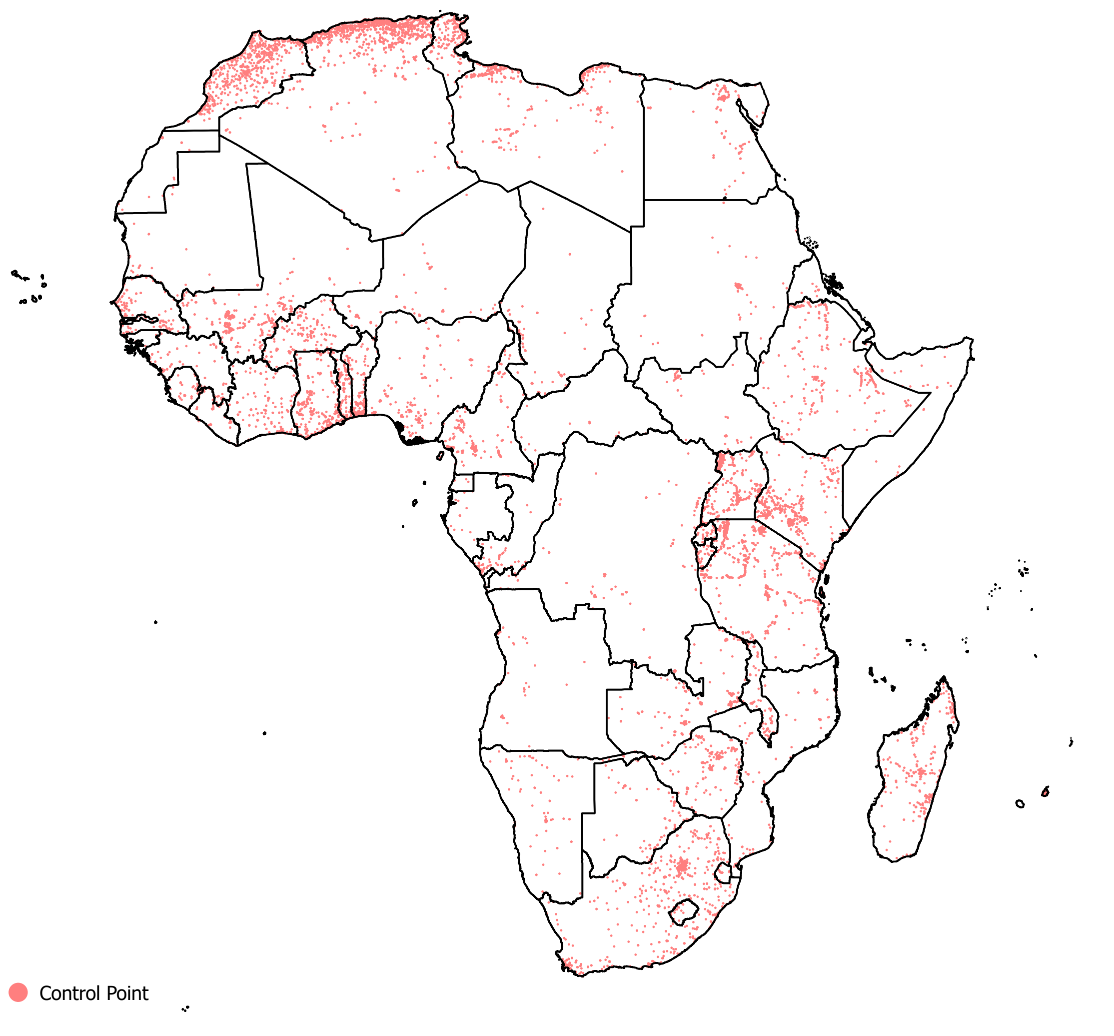


Figure 2.12. Control point locations used to generate smoothed surface.

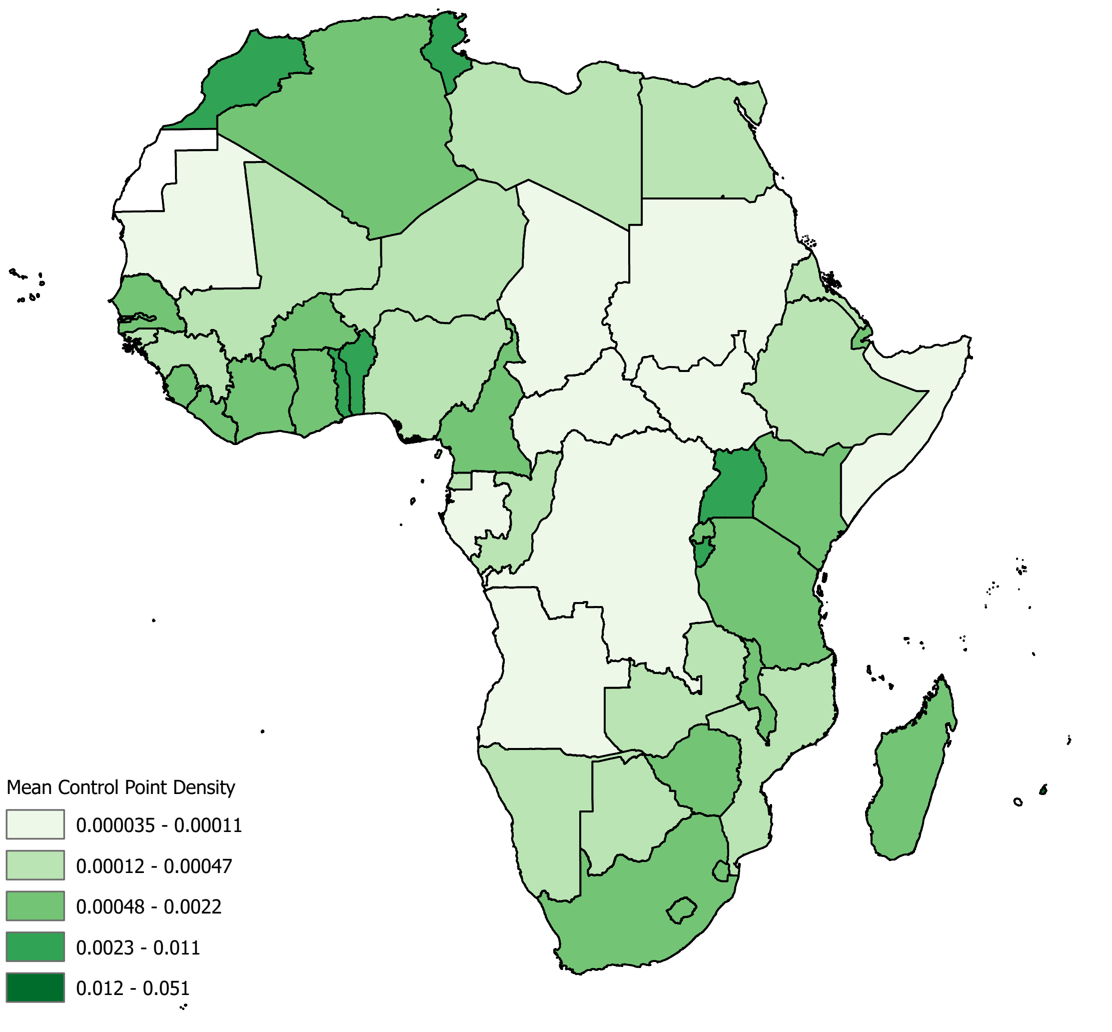


Figure 2.13. Average control point density by country.

2.6.2 Units of Analysis

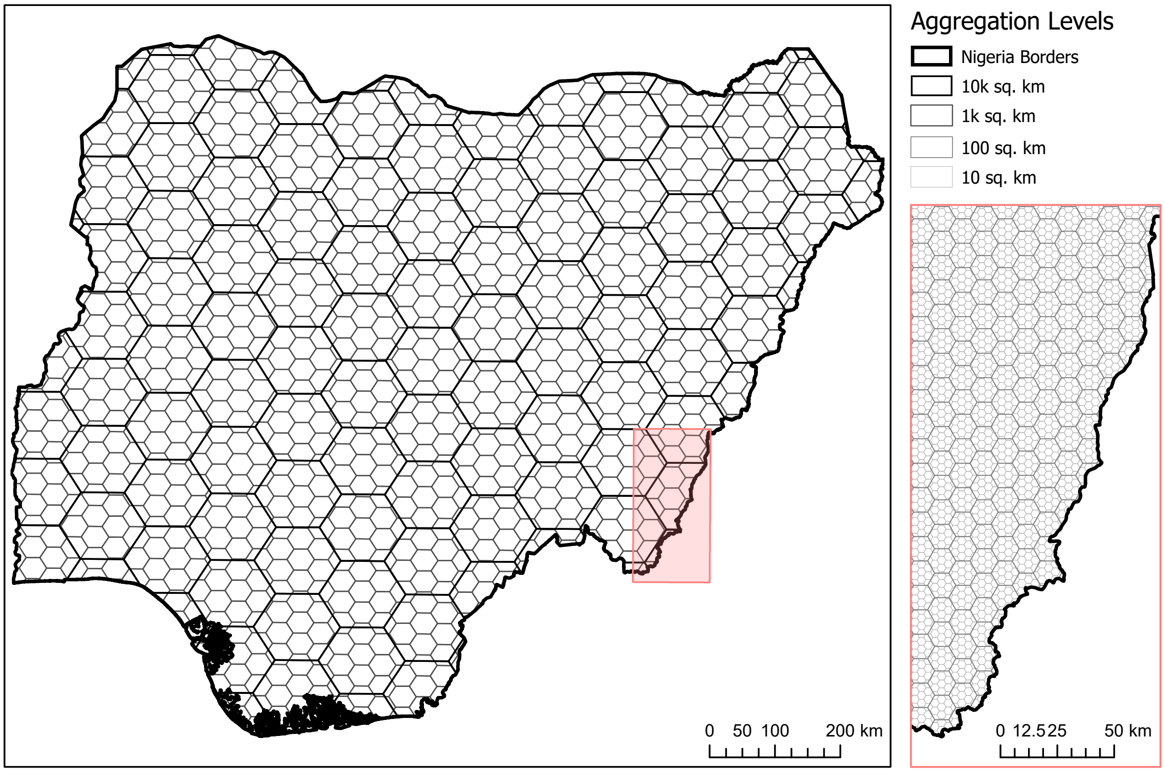


Figure 2.14. Hexagonal lattices overlaid on Nigeria; inset detail shows 100km^2 and 10km^2 cells.

2.7 Works Cited

Acemoglu, Daron, Davide Cantoni, Simon Johnson, and James A. Robinson. 2011. “The Consequences of Radical Reform: The French Revolution.” *American Economic Review* 101 (7): 3286–3307.

Acemoglu, Daron, Camilo García-Jimeno, and James A. Robinson. 2015. “State Capacity and Economic Development: A Network Approach.” *American Economic Review* 105 (8): 2364–2409.

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review* 91 (5): 1369–1401.

Anders, Therese. 2020. “Territorial Control in Civil Wars: Theory and Measurement Using Machine Learning.” *Journal of Peace Research* 57 (6): 701–714.

Bäck, Hanna, and Axel Hadenius. 2008. “Democracy and State Capacity: Exploring a J-Shaped Relationship.” *Governance* 21 (1): 1–24.

Benson, Michelle, and Jacek Kugler. 1998. “Power Parity, Democracy, and the Severity of Internal Violence.” *Journal of Conflict Resolution* 42 (2): 196–209.

Bersch, Katherine, Sérgio Praça, and Matthew M. Taylor. 2017. “State Capacity, Bureaucratic Politicization, and Corruption in the Brazilian State.” *Governance* 30 (1): 105–124.

Besley, Timothy, and Torsten Persson. 2008. “Wars and State Capacity.” *Journal of the European Economic Association* 6 (2-3): 522–530.

———. 2009. “The Origins of State Capacity: Property Rights, Taxation, and Politics.” *American Economic Review* 99 (4): 1218–1244.

Boone, Catherine. 1998. “State Building in the African Countryside: Structure and Politics at the Grassroots.” *The Journal of Development Studies* 34 (4): 1–31.

———. 2003. *Political Topographies of the African State: Territorial Authority and Institutional Choice.* Cambridge University Press.

Bradbury, Mark, and Michael Kleinman. 2010. *Winning Hearts and Minds? Examining the Relationship between Aid and Security in Kenya.* Medford, MA: Feinstein International Center.

Buhaug, Halvard. 2010. “Dude, Where’s My Conflict?: LSG, Relative Strength, and the Location of Civil War.” *Conflict Management and Peace Science* 27 (2): 107–128.

Cárdenas, Mauricio. 2010. “State Capacity in Latin America.” *Economía* 10 (2): 1–45. JSTOR: 25800045.

Centeno, Miguel Angel. 2002. *Blood and Debt: War and the Nation-State in Latin America.* Penn State Press.

Chang, Wen-Yang, and Dan Wei. 2019. “Natural Resources and Infectious Diseases: The Case of Malaria, 2000–2014.” *The Social Science Journal* 56 (3): 324–336.

Cingolani, Luciana. 2013. *The State of State Capacity: A Review of Concepts, Evidence and Measures.* Technical report. UNU-MERIT.

Coleman, James S. 1977. “The Concept of Political Penetration.” In *Government and Rural Development in East Africa: Essays on Political Penetration,* edited by L. Cliffe, J. S. Coleman, and M. R. Doornbos, 3–18. Institute of Social Studies. Dordrecht: Springer Netherlands.

Collier, Paul, and Anke Hoeffler. 2004. “Greed and Grievance in Civil War.” *Oxford Economic Papers* 56 (4): 563–595.

Cooley, Charles Horton. 1956. *The Two Major Works of Charles H. Cooley: Social Organization and Human Nature and the Social Order.* Glencoe, Ill: Free Press.

Crowder, Michael. 2023. *West Africa Under Colonial Rule.* Taylor & Francis.

Crowther, A. F. 1957. *Geology of the Mwingi Area, North Kitui: Degree Sheet 45, South-west Quarter (with Coloured Geological Map).* Geological Survey of Kenya.

De Waal, Alexander. 1997. *Famine Crimes: Politics & the Disaster Relief Industry in Africa.* Indiana University Press.

Dincecco, Mark. 2017. *State Capacity and Economic Development: Present and Past.* Cambridge University Press.

Dincecco, Mark, and Gabriel Katz. 2016. “State Capacity and Long-Run Economic Performance.” *The Economic Journal* 126 (590): 189–218.

Dincecco, Mark, and Mauricio Prado. 2012. “Warfare, Fiscal Capacity, and Performance.” *Journal of Economic Growth* 17 (3): 171–203.

Driscoll, Jesse, and Michael Seese. 2023. “Exiting Anarchy: Institutional Correlates of Civilian Welfare in 2012 Mogadishu.”

Englebert, Pierre. 2009. “Africa: Unity, Sovereignty, Sorrow.” In *Africa.* Lynne Rienner Publishers.

Fearon, James D., and David D. Laitin. 2003. “Ethnicity, Insurgency, and Civil War.” *American Political Science Review* 97 (1): 75–90.

Fortin, Jessica. 2010. “A Tool to Evaluate State Capacity in Post-Communist Countries, 1989–2006.” *European Journal of Political Research* 49 (5): 654–686.

Fukuyama, Francis. 2014. *State-Building: Governance and World Order in the 21st Century.* Cornell University Press.

Geddes, Barbara. 1994. “Challenging the Conventional Wisdom Economic Reform and Democracy.” *Journal of Democracy* 5 (4): 104–118.

Gennaioli, Nicola, and Hans-Joachim Voth. 2015. “State Capacity and Military Conflict.” *The Review of Economic Studies* 82 (4): 1409–1448.

Goffman, Erving. 1964. “The Neglected Situation.” *American Anthropologist* 66 (6): 133–136. JSTOR: 668167.

———. 1966. *Behavior in Public Places: Notes on the Social Organization of Gatherings.* Reissue edition. New York: Free Press.

Goodwin, Jeff. 2001. *No Other Way Out: States and Revolutionary Movements, 1945-1991.* Cambridge University Press.

Hanson, Jonathan K., and Rachel Sigman. 2019. "State Capacity and World Bank Project Success."

———. 2021. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." *The Journal of Politics* 83 (4): 1495–1510.

Hendrix, Cullen S. 2010. "Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict." *Journal of Peace Research* 47 (3): 273–285.

———. 2011. "Head for the Hills? Rough Terrain, State Capacity, and Civil War Onset." *Civil Wars* 13 (4): 345–370.

Herbst, Jeffrey. 2014. *States and Power in Africa: Comparative Lessons in Authority and Control - Second Edition.* Princeton University Press.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal* 39 (153): 41–57. JSTOR: 2224214.

Humphreys, Laud. 1975. *Tearoom Trade, Enlarged Edition: Impersonal Sex in Public Places.* Transaction Publishers.

Humphreys, Macartan. 2005. "Natural Resources, Conflict, and Conflict Resolution: Uncovering the Mechanisms." *Journal of Conflict Resolution* 49 (4): 508–537.

Jackson, Robert H., and Carl G. Rosberg. 1982. "Why Africa's Weak States Persist: The Empirical and the Juridical in Statehood." *World Politics* 35 (1): 1–24.

Jensen, Jeffrey L., and Adam J. Ramey. 2020. "Early Investments in State Capacity Promote Persistently Higher Levels of Social Capital." *Proceedings of the National Academy of Sciences* 117 (20): 10755–10761.

Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War.* Cambridge University Press.

Koehnlein, Britt, and Ore Koren. 2022. "COVID-19, State Capacity, and Political Violence by Non-State Actors." *Journal of Peace Research* 59 (1): 90–104.

Lee, Melissa M., and Nan Zhang. 2017. "Legibility and the Informational Foundations of State Capacity." *The Journal of Politics* 79 (1): 118–132.

Levi, Margaret. 1988. *Of Rule and Revenue.* University of California Press.

Lindvall, Johannes, and Jan Teorell. 2016. *State Capacity as Power: A Conceptual Framework,* 1v2016, Lund.

Livingston, Steven, and Gregor Walter-Drop, eds. 2014. *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood.* Illustrated edition. Oxford ; New York: Oxford University Press.

Luna, Juan Pablo, and Hillel David Soifer. 2017. "Capturing Sub-National Variation in State Capacity: A Survey-Based Approach." *American Behavioral Scientist* 61 (8): 887–907.

Man, John. 2012. *Kublai Khan.* Random House.

Mann, Michael. 2012a. *The Sources of Social Power: Volume 1, A History of Power from the Beginning to AD 1760.* Cambridge University Press.

Mann, Michael. 2012b. *The Sources of Social Power: Volume 2, The Rise of Classes and Nation-States, 1760-1914.* Cambridge University Press.

McAdam, Doug, Sidney Tarrow, and Charles Tilly. 2001. *Dynamics of Contention.* Cambridge University Press.

McArthur, John W., and Jeffrey D. Sachs. 2001. *Institutions and Geography: Comment on Acemoglu, Johnson and Robinson (2000).* Working Paper, 8114. National Bureau of Economic Research: 8114.

Mead, G.H. 1934. *Mind, Self, and Society from the Standpoint of a Social Behaviorist.* Mind, Self, and Society from the Standpoint of a Social Behaviorist. University of Chicago Press: Chicago.

Migdal, Joel S. 1988. *Strong Societies and Weak States.* Princeton, N.J: Princeton University Press.

Müller-Crepon, Carl. 2021. "State Reach and Development in Africa since the 1960s: New Data and Analysis." *Political Science Research and Methods,* 1–10.

Müller-Crepon, Carl, Philipp Hunziker, and Lars-Erik Cederman. 2021. "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa." *Journal of Conflict Resolution* 65 (2-3): 563–590.

Nordhaus, William D. 2006. "Geography and Macroeconomics: New Data and New Findings." *Proceedings of the National Academy of Sciences* 103 (10): 3510–3517.

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics* 94 (1): 20–36.

Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87 (3): 567–576.

Ottervik, Mattias. 2013. *Conceptualizing and Measuring State Capacity.* Technical report 2013:20. Gothenburg: The Quality of Goverment Institute.

Rabasa, Angel, Steven Boraz, Peter Chalk, Kim Cragin, Theodore W. Karasik, Jennifer D. P. Moroney, Kevin A. O'Brien, and John E. Peters. 2007. *Ungoverned Territories: Understanding and Reducing Terrorism Risks.* Technical report. RAND Corporation.

Rogowski, Jon C., John Gerring, Matthew Maguire, and Lee Cojocaru. 2022. "Public Infrastructure and Economic Development: Evidence from Postal Systems." *American Journal of Political Science* 66 (4): 885–901.

Rosenau, James N., and Ernst-Otto Czempiel. 1992. *Governance Without Government: Order and Change in World Politics.* Cambridge University Press.

Rubin, Michael A. 2020. "Rebel Territorial Control and Civilian Collective Action in Civil War: Evidence from the Communist Insurgency in the Philippines." *Journal of Conflict Resolution* 64 (2-3): 459–489.

Rueda, Miguel R. 2017. "Popular Support, Violence, and Territorial Control in Civil War." *Journal of Conflict Resolution* 61 (8): 1626–1652.

Sachs, Jeffrey, John W McArthur, Guido Schmidt-Traub, Margaret Kruk, Chandrika Bahadur, Michael Faye, and Gordon McCord. 2004. "Ending Africa's Poverty Trap." *Brookings Papers on Economic Activity* 2004 (1): 117–240.

Savoia, Antonio, and Kunal Sen. 2015. "Measurement, Evolution, Determinants, and Consequences of State Capacity: A Review of Recent Research." *Journal of Economic Surveys* 29 (3): 441–458.

Schönholzer, David, and Pieter François. 2023. *Environmental Circumscription and the Origins of the State.*

Scott, James C. 1999. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed.* New Haven, CT London: Yale University Press.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis.* CRC Press.

Skocpol, Theda. 1985. "Bringing the State Back In: Strategies of Analysis in Current Research." In *Bringing the State Back In,* edited by Dietrich Rueschemeyer, Peter B. Evans, and Theda Skocpol, 3–38. Cambridge: Cambridge University Press.

Skocpol, Theda, and Kenneth Finegold. 1982. "State Capacity and Economic Intervention in the Early New Deal." *Political Science Quarterly* 97 (2): 255–278. JSTOR: 2149478.

Soifer, Hillel. 2008. "State Infrastructural Power: Approaches to Conceptualization and Measurement." *Studies in Comparative International Development* 43 (3): 231–251.

Soifer, Hillel, and Matthias vom Hau. 2008. "Unpacking the Strength of the State: The Utility of State Infrastructural Power." *Studies in Comparative International Development* 43 (3): 219–230.

Subramaniam, Niran, Joe Nandhakumar, and João Baptista (John). 2013. "Exploring Social Network Interactions in Enterprise Systems: The Role of Virtual Co-Presence." *Information Systems Journal* 23 (6): 475–499.

Tao, Ran, Daniel Strandow, Michael Findley, Jean-Claude Thill, and James Walsh. 2016. "A Hybrid Approach to Modeling Territorial Control in Violent Armed Conflicts." *Transactions in GIS* 20 (3): 413–425.

Thies, Cameron G. 2007. "The Political Economy of State Building in Sub-Saharan Africa." *The Journal of Politics* 69 (3): 716–731.

———. 2009. "National Design and State Building in Sub-Saharan Africa." *World Politics* 61 (4): 623–669.

———. 2010. "Of Rulers, Rebels, and Revenue: State Capacity, Civil War Onset, and Primary Commodities." *Journal of Peace Research* 47 (3): 321–332.

Thomas, M A. 2010. "What Do the Worldwide Governance Indicators Measure?" *The European Journal of Development Research* 22 (1): 31–54.

Tilly. 1992. *Coercion, Capital, and European States, A.D. 990-1990.* Revised edition. Cambridge, MA: Wiley-Blackwell.

Walker, E. Eaton. 1903. *Reports on the Geology of the East Africa Protectorate.* Africa (Great Britain. Colonial Office), no. 11. London: H.M. Stationery Office.

Walter, Barbara F. 2006. "Building Reputation: Why Governments Fight Some Separatists but Not Others." *American Journal of Political Science* 50 (2): 313–330.

———. 2019. "Explaining the Number of Rebel Groups in Civil Wars." *International Interactions* 45 (1): 1–27.

Wang, Shaoguang. 1995. "The Rise of the Regions: Fiscal Reform and the Decline of Central State Capacity in China." In *The Waning of the Communist State: Economic Origins of Political Decline in China and Hungary,* edited by Andrew G. Walder, 87–113. Berkeley, Los Angeles, Oxford: University of California Press.

Wang, Xiaohu (Shawn), and An'gang Hu. 2015. *The Chinese Economy in Crisis: State Capacity and Tax Reform.* Routledge.

Warren, T. Camber. 2014. "Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty." *International Organization* 68 (1): 111–141.

Weber, Max. 2021. *Politics As a Vocation.* Creative Media Partners, LLC.

Williams, Martin J. 2021. "Beyond State Capacity: Bureaucratic Performance, Policy Implementation and Reform." *Journal of Institutional Economics* 17 (2): 339–357.

Zhao, Shanyang. 2003. "Toward a Taxonomy of Copresence." *Presence* 12 (5): 445–455.

Chapter 3

Geographic Correlates of Territorial Control

3.1 Introduction

In his 2014 book, *States and Power in Africa*, Jeffrey Herbst makes a number of claims about where, within a given African state, we should expect territorial control to concentrate. These claims are primarily informed by the spatial distribution of a country's population, and the relative (in)accessibility of various regions within a country. In this chapter, I empirically test some of Herbst's core hypotheses using the territorial control measure described in Chapter 2. Importantly, however, this chapter does not engage directly with Herbst's broader theories about state consolidation in the African context—in particular, the argument that consolidation efforts are primarily influenced by the costs of extending control, the configuration of international borders across the continent, and the nature of the state system that African leaders constructed in the post-colonial era. Instead, this chapter focuses on empirical claims that Herbst makes about geography in Chapter 5, "National Design and the Broadcasting of Power."

3.2 Population Distribution

Herbst argues that one of the distinguishing features of state building in Africa is the distribution of the population across the continent. Unlike Europe, whose population is relatively concentrated in across compact stretches of territory, African populations are sparsely distributed across vast territories. How leaders respond to particular patterns of population density explain patterns of authority and control. While he underscores two other salient geographical features—ecological variation and transportation networks (2014, 12)—he dismisses both the size and shape of countries as unimportant (c.f. Alesina et al. 2005), instead arguing that "what is critical is the particular population distribution that they present to national leaders. As Gottmann (1975, 145) noted in what became a classic study, 'it is the *organization of a territory by its population* that counts more than any other feature of it.'" In this section, I provide both direct and indirect tests of Herbst's population hypotheses.

3.2.1 Population Density

Here, I test this proposition empirically using high resolution population data from the Facebook Connectivity Lab and the Center for International Earth Science Information Network (CIESIN) at Columbia University. Facebook uses machine learning algorithms to identify structures, such as homes and apartment complexes, from satellite imagery provided by Maxar Technologies. CIESIN then combines these data with national census data to produce high resolution (30 meter) population density estimates.[1] As a robustness check, I replicate the analysis using the coarser (30 arc-second) WorldPop population density data.

I estimate the naïve correlation using OLS:

$$\text{Territorial Control}_i = \alpha + \beta\left(\text{Population Density}_i\right) + \epsilon_i \tag{3.1}$$

Where α is the expected territorial control value in uninhabited or very low population density areas of the continent, β is our coefficient of interest, and ϵ is a robust error term clustered at the country level. All variables are standardized ($\mu = 0$, $\sigma = 1$) for ease of interpretation. The unit of analysis, i, is an arbitrarily defined spatial grid cell (see Appendix Figure 2.14). Territorial control and population density values are averaged within each grid cell. Spatial phenomena are susceptible to what geographers call the modifiable areal unit problem (MAUP), in which the measurement of a phenomenon (such as territorial control or population density) is influenced by the shape and size of the aggregation unit. To address this potential bias, I construct a series of hexagonal lattices with cell sizes ranging from 10 square kilometers to 10,000 square kilometers and replicate the analysis at each level of aggregation. Table 3.1 presents the results.

1. Because FaceBook–CIESIN only reports data for inhabited territory, I impute population density = 0 for all missing raster values. This has the effect of decreasing average density per grid cell. Appendix Tables 3.8 through 3.12 report summary statistics for both the raw and corrected data.

Table 3.1. Correlation With Population Density

	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)
Mean Facebook Pop. Dens.	0.109**	0.0932*	0.0791*	0.0778	0.0939					
	(0.0510)	(0.0489)	(0.0400)	(0.0481)	(0.136)					
Mean WorldPop Pop. Dens.						0.211***	0.262***	0.349***	0.385***	0.781*
						(0.0765)	(0.0903)	(0.100)	(0.131)	(0.450)
Constant	-0.000104	-0.000202	-0.000324	-0.000770	-0.00332	-0.00145	-0.000820	-0.000370	0.000940	0.108
	(0.0250)	(0.0262)	(0.0324)	(0.0444)	(0.135)	(0.0246)	(0.0257)	(0.0310)	(0.0396)	(0.166)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Observations	2,989,475	307,105	33,091	4,067	52	2,972,830	305,826	33,002	4,057	52
R^2	0.012	0.009	0.006	0.006	0.008	0.045	0.069	0.123	0.149	0.431
Adj. R^2	0.0120	0.00878	0.00626	0.00585	-0.0115	0.0449	0.0693	0.123	0.149	0.419
Clusters	52	52	52	52	–	52	52	52	52	–
F Statistic	4.550	3.631	3.903	2.616	0.474	7.593	8.438	12.14	8.595	3.007
Prob > F	0.0377	0.0623	0.0536	0.112	0.494	0.00811	0.00542	0.00102	0.00504	0.0890

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Results suggest that there is, in fact, a statistically significant positive correlation between between population density and territorial control. At the smallest level of aggregation ($10km^2$), a one standard deviation increase in population density is associated with a 0.1 to 0.2 standard deviation increase in territorial control, which seems to confirm some of the descriptive observations above that control is more entrenched in urban and other high population density areas.[2] Given the relatively small magnitude of these effects, and the small R-Square values, we can assume that control is not collinear with population density. It is not the case that population density is the sole (or even primary) predictor of the intensity of territorial control across the continent. One caveat to interpreting these results, though, is that the control point density data does not contain any information on the service load, or efficiency, of individual control points. A police station or tax assessor's office in high density urban areas, for example, may be equipped (i.e., through personnel or technology) to handle many more cases than an analogous office in in a rural area—effectively exerting a higher degree of control from a single location. Chapter 6 evaluates this possibility using historical budgetary and personnel data from Kenya and Senegal.

3.2.2 The Urban–Rural Divide

A more generalized measure of population density is whether territory can be classified as either rural or urban. Here, we look at the correlation between urban areas and territorial control, using data from the Global Rural-Urban Mapping Project (GRUMP v1.02). The GRUMP data combines population data with nighttime light emission data to identify areas of the Earth's surface that appear to be urbanized (i.e., settlements that contain a population ≥ 5000). We estimate the correlation between urban areas and territorial control with:

$$\text{Territorial Control}_i = \alpha + \beta\,(\text{Urban}_i) + \epsilon_i \tag{3.2}$$

Where "Urban" is a binary indicator of an urban area, and ϵ is a robust error term clustered at the country level. Functionally, the "Urban" variable represents the proportion of urban territory within a given grid cell. All variables are standardized for ease of interpretation. Results are in Table 3.2.

2. Results do not change dramatically when excluding island Africa (Seychelles, Mauritius). Cape Verde, Comoros, Mayotte, Reunion, Saint Helena, and São Tomé and Príncipe are excluded from the analysis.

Table 3.2. Correlation with Urban Territory

	(1)	(2)	(3)	(4)	(5)
Mean Urban–Rural Extent	0.229***	0.303***	0.453***	0.600***	1.082***
	(0.0601)	(0.0824)	(0.124)	(0.150)	(0.123)
Constant	4.08e-06	0.000242	0.00110	0.00808	0.156***
	(0.0230)	(0.0232)	(0.0276)	(0.0352)	(0.0540)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Observations	2,989,475	307,105	33,091	4,067	52
R^2	0.052	0.091	0.201	0.333	0.913
Adj. R^2	0.0524	0.0911	0.201	0.332	0.911
Clusters	52	52	52	52	–
F Statistic	14.54	13.56	13.45	15.97	76.81
Prob > F	0.000372	0.000558	0.000585	0.000207	0

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.2 shows point estimates of much higher magnitude than those in Table 3.1. A one standard deviation increase in the the proportion of urban territory is associated with between $\frac{1}{4}$ and one standard deviation increase in territorial control, depending on the level of aggregation. The relationship between urban territory and territorial control is particularly robust at the country level; highly urban countries, such as South Africa (5.02% urban), Tunisia (8.18% urban), and Gambia (8.97% urban), have a predicted territorial control value that is $\frac{1}{3}$ standard deviations above mean, which is roughly 3 times the effect of population density by itself (as measured by the highly granular FaceBook data). The disconnect between population density and urban areas is likely a function of the relatively low populations of African cities in relation to area. While high density African cities, such as Kinshasa, Lagos, and Nairobi, do exist, the median density of African cities is less than one half the median density of European cities. In effect, then, urban territory is a less noisy measure of population because it truncates some of the extreme variation induced by the granularity of the FaceBook data. This is why these estimates are more consistent with the WorldPop models.

In general, results are consistent with the observations of Livingston and Walter-Drop (2014, 5): "When one ventures to the edges of Kinshasa or Kabul, one has reached the outer limits of the state's governance reach. Beyond the city borders lies a vast stretch of territory where the state is weak or altogether absent."

3.2.3 African National Design

Herbst (2014, Ch. 5) hypothesizes that the pattern of population density within a state's borders determines the pattern of territorial control within that state. Using these new territorial control data, we can directly test this hypothesis. Herbst categorizes African states into four distinct types of political geography based on population densities and land area (see Appendix Table 3.13). Difficult geographies, found in countries like Angola and the DRC, are defined by non-contiguous regions of high population density scattered throughout a large state. Hinterland countries, such as Chad, Mali, and other Sahelian states, contain one or two urban agglomerations with relatively high population density and large hinterland regions with low population density. Neutral geographies, such as Cameroon and Côte d'Ivoire, contain dispersed populations and relatively small hinterland regions. Countries with favorable geographies, such as Benin and Botswana, are relatively small states whose population distributions follow a pattern of radial decay from the capital.

Based on Herbst's analysis, we should expect countries with more favorable geographies to exhibit higher aggregate levels of territorial control than countries with more difficult geographies. Hinterland countries and countries with neutral geographies should fall somewhere between the two extremes. To test this hypothesis, I estimate the following model using OLS:

$$\text{Territorial Control}_i = \alpha + \beta\left(\text{Country Type}_i\right) + \epsilon_i \tag{3.3}$$

Where "Territorial Control" is the mean feature density in country i, "Country Type" is Herbst's ordinal typology, and ϵ is a robust error term. Results are given in Table 3.3. I estimate the model twice, using both the raw and standardized outcome data. Because the outcome variable is standardized in columns (3) and (4), β is interpreted as the standard deviation change in territorial control when moving from one country type to another. Columns (1) and (3) show the results using an ordinal typology, while columns (2) and (4) estimate Equation 3.3 with country type as a factor variable.

Results of models (1) and (3) indicate that there is a significant difference between country types; control tends to increase when moving from difficult geographies to more favorable geographies. Models (2) and (4) show a slightly more nuanced picture. Estimated control point density by country type is given in Figure 3.1 for models (3) and (4). The factor model (model (4)) indicates that

Table 3.3. Relationship Between Herbst's Classification and Territorial Control

	Mean		z(Mean)	
	(1)	(2)	(3)	(4)
Typology (Ordinal)	0.000355**		0.0493**	
	(0.000163)		(0.0226)	
Hinterland Countries		-0.000217		-0.0301
		(0.000163)		(0.0226)
Neutral Geography		0.000954***		0.132***
		(0.000306)		(0.0425)
Favorable Geography		0.000962*		0.133*
		(0.000491)		(0.0681)
Constant	-3.88e-05	0.000380**	-0.304***	-0.245***
	(0.000294)	(0.000147)	(0.0408)	(0.0204)
Observations	40	40	40	40
R^2	0.091	0.112	0.091	0.112
Adj. R^2	0.0673	0.0377	0.0673	0.0377
F Statistic	4.745	7.899	4.745	7.899
Prob > F	0.0357	0.000356	0.0357	0.000356

Robust standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$,*$p < 0.1$

hinterland countries have the lowest estimated average control point density, significantly lower, in fact, than countries with difficult geographies. This may due, in part, to the size of these hinterland countries. Chad, Malia, Mauritnia, and Niger, have some of the largest land area out of any country on the continent. Assuming a constant cost per unit of territory controlled, and no gains from efficiency, the size alone of these countries suggests that territorial control should be more difficult to achieve. The results of model (4) in Figure 3.1 also indicate that that there is no real difference between countries with neutral geographies and countries with favorable geographies. Estimated control point density is largely identical (though neutral countries have marginally higher estimated control point density than countries with favorable geographies), though these estimates have fairly large standard errors, which is driven by the high levels of variation within these groups, both in terms of measured territorial control, but also just descriptively: Botswana, Gambia, and Eswatini are not similar countries in many respects, yet they are all classified by Herbst as possessing favorable geographies (see Figure 3.3). Estimating the models with a logged outcome variable as in Figure 3.4 reduces the standard errors, though point estimates are a similar. Neutral countries have a slightly, though significantly higher estimated control point density than favorable countries.

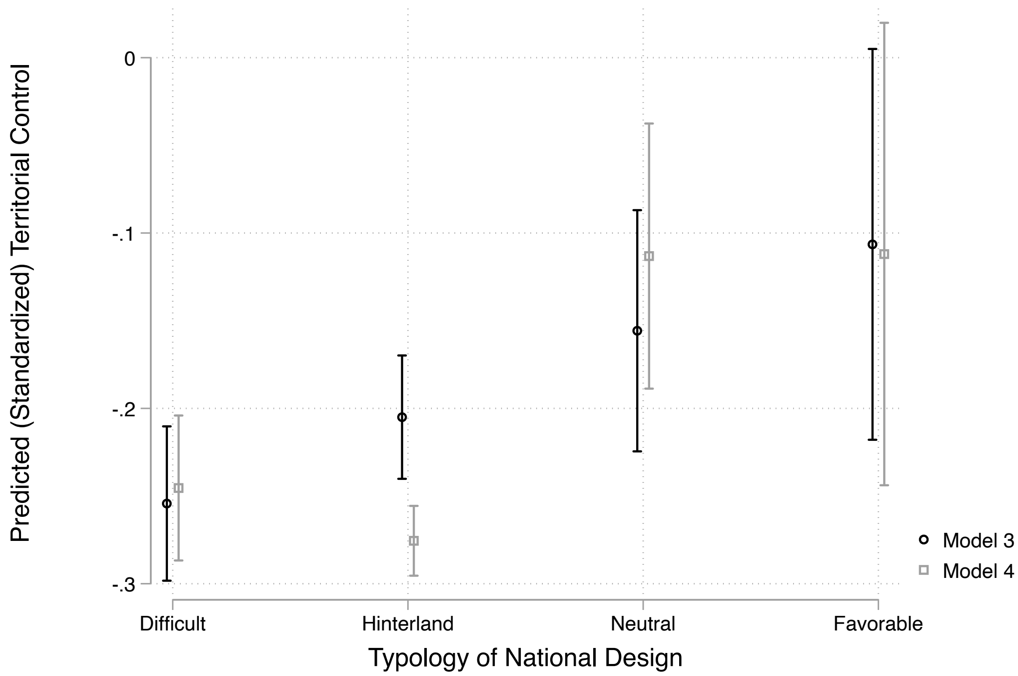


Figure 3.1. Predicted (standardized) territorial control value by country type.

3.3 Accessibility

In this section, I look at the accessibility of territory as a predictor of control. The civil war literature suggests that the central government's ability to penetrate territory is indicative of territorial control (Fearon & Laitin 2003, p. 80; Kalyvas; other citations). As noted above, one of the most common proxy measures for control is the ruggedness of terrain within a given territory. Fearon and Laitin (2003), Collier and Hoeffler (2004), Collier et al. (2009), and Miguel et al. (2004) use the percent of mountainous terrain in a country defined by Gerrard (2000). Nunn and Puga (2012) and Kalyvas and Kocher (2009) use a more disaggregated approach pioneered by Riley et al. (1999), and refined by Shaver et al. (2019). These new territorial control data allow us to asses the relationship between rugged terrain territorial control at a more granular, subnational level. Using elevation data from NASA-JPL's Shuttle Radar Topography Mission (SRTM), I estimate:

$$\text{Territorial Control}_i = \alpha + \beta\,(\text{Elevation}_i) + \epsilon_i \tag{3.4}$$

Where "Elevation" is either the cell mean elevation or the cell standard deviation of elevation. All variables are standardized. Standard deviation of elevation is the preferred metric, as it represents the variability in elevation within a given grid cell, and is thus a direct measure of the ruggedness of the terrain in a given locale. It is also the closest measure of ruggedness to Riley et al. (1999), which is essentially the deviation in elevation between a given square grid cell and its eight contiguous neighbors. Table 3.4 provides the results.

The results here are counterintuitive, and somewhat surprising. We find that the standard deviation of elevation is positively and significantly correlated with territorial control. A one standard deviation increase in in the standard deviation of elevation is associated with a 0.05 to 0.1 standard deviation *increase* in territorial control. This relationship is significant across all levels of aggregation except for the country level, which does not attain statistical significance. This suggests a marginal increase in levels of state penetration in even the most rugged of terrain. This is inconsistent with the theory promulgated in the civil war literature that mountainous areas are difficult for governments to penetrate (Fearon and Laitin 2003; Collier and Hoeffler 2004), though this theory is disputed in the African context by Buhaug and Rød (2006). Results further show an absence of any systematic relationship between mean elevation and territorial control. Theory suggests a negative and significant relationship. But it is possible that the relationship may not be linear. To account for this possibility, I fit the following fractional polynomial model:

$$\log(\text{Territorial Control})_i = \alpha + \beta_1 \left(\text{Elevation}_i^{(p1)}\right) + \beta_2 \left(\text{Elevation}_i^{(p2)}\right) + \epsilon_i \tag{3.5}$$

Where $p1$ and $p2$ are powers chosen from the set S = {-2, -1, -0.5, 0, 0.5, 1, 2, 3}. The model is estimated for 36 power combinations (28 unique combinations and 8 repeated power models, in which $p1 = p2$); the most parsimonious model is then selected by default to minimize both deviance and model complexity (Royston and Sauerbrei 2008). Fractional polynomials allow for more flexibility and a greater range of functional forms than a standard quadratic model, and are widely used in the epidemiological literature to model nonlinear relationships (Royston and Altman 1994; Royston and Sauerbrei 2008; Baneshi et al. 2013).

Results of the fractional polynomial model detailed in Equation 3.5 are given in Appendix

Table 3.4. Correlation with Elevation

	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)
Standard Deviation of Elevation	0.0559*	0.0673*	0.0867*	0.0974*	-0.0816					
	(0.0293)	(0.0350)	(0.0437)	(0.0487)	(0.0829)					
Mean Elevation						0.0207	0.0165	0.00459	-0.0280	-0.114
						(0.0200)	(0.0207)	(0.0239)	(0.0335)	(0.128)
Constant	-3.26e-05	-3.08e-05	4.81e-05	0.000674	-0.00605	-4.25e-05	-1.40e-05	9.80e-05	0.000546	0.00473
	(0.0245)	(0.0252)	(0.0302)	(0.0401)	(0.134)	(0.0252)	(0.0262)	(0.0316)	(0.0419)	(0.144)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Observations	2,983,010	306,431	33,017	4,059	52	2,983,010	306,431	33,017	4,059	52
R^2	0.003	0.004	0.007	0.009	0.006	0.000	0.000	0.000	0.001	0.014
Adj. R^2	0.00312	0.00449	0.00742	0.00899	-0.0136	0.000426	0.000269	-9.24e-06	0.000538	-0.00588
Clusters	52	52	52	52	–	52	52	52	52	–
F Statistic	3.636	3.693	3.946	3.991	0.968	1.061	0.634	0.0367	0.697	0.788
Prob > F	0.0622	0.0603	0.0524	0.0511	0.330	0.308	0.430	0.849	0.408	0.379

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

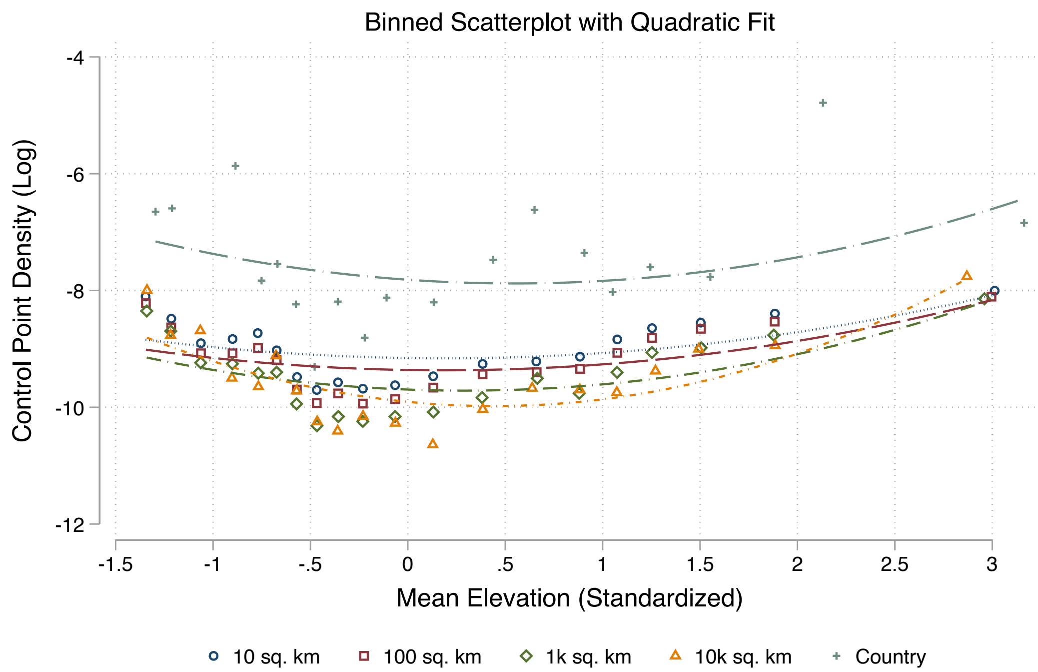


Figure 3.2. Binned scatterplots depicting relationship between (log) control point density and elevation.

Section 3.6.2. Because of the difficulty inherent in interpreting polynomial coefficients, I plot the quadratic fit with a binned scatterplot of the data in Figure 3.2. The figure shows an upward facing parabolic curve at all levels of aggregation, suggesting that density is highest at elevation extremes. These curves are mirrored by the fractional polynomial regressions, shown in Appendix Figure 3.5. This is a somewhat puzzling finding; existing theory would suggest that control is more difficult to exert in mountainous regions (Fearon and Laitin 2003) and "sparsely wooded lowlands" (Buhaug and Rød 2006, 327). However, these expectations do not necessarily accord with the topography of the African continent. Consider the comparison between Colombia in South America to Kenya. Because of the climate extremes with jungle in low lying and coastal regions and the high altitude of the Andes, human settlements in Colombia are centered in the mid-elevation regions—in cities such as Bogotá and Medellín, which are not only population centers, but the core economic and political engines of the country, suggesting a downward facing parabola in which control is centered at middle elevations. Kenya (and other African states), by contrast, does not have such topographical extremes. Thus, human settlements tend to exist along the coasts due to the historical importance of trade (e.g. Mombasa) and at higher elevations, where colonial powers established administrative outposts in more temperate locations insulated from malaria-carrying insects (Acemoglu et al. 2001). Mid-level elevations on the continent (i.e., between -0.5 and 0.5 standard deviations from the continental mean elevation) are generally arid with low population densities. Examples include the Chalbi Desert and Awara Plain in Kenya, and stretches of the Malian Sahara extending south from the Adrar des Ifoghas. Both regions are at mean elevation in their respective countries ($\approx 400-500$ meters) and have extremely low measures of territorial control. Nairobi, by contrast, is situated at 1795 meters ($\approx 2\sigma$ above mean) and Bamako is situated at 350 meters ($\approx 1\sigma$ below mean).

Although important, elevation is not the only geographical feature that influences accessibility. Transportation infrastructure—and roads in particular—allow governments to extend control throughout the state. Herbst quotes Heggie (1995, 1): "Efficient transportation consolidates political areas, whether the Roman Empire or the United States of America. The lack of ready means of circulation is a source of political weakness whatever the density of the population" (Herbst 2014, 162). In the previous chapter, I underscored that territorial control is primarily concentrated in areas of high road network density. I test this empirically by estimating:

$$\text{Territorial Control}_i = \alpha + \beta \left(\text{Road Density}_i\right) + \epsilon_i \tag{3.6}$$

Results are given in Table 3.5. The patterns described in the previous chapter seemed to be confirmed. A standard deviation increase in mean road density is associated with a 0.06 standard deviation increase in territorial control across all levels of aggregation except the country level.

3.4 Resources and Extractive Capacity

The most common measures state capacity and territorial control in the literature tend to be economic variables. These are believed to be important because extractive capacity is both an impetus for control, as states try to maximize rents, and also a prerequisite of control, providing the resources necessary for the state employ and deploy agents across territory. Unfortunately, subnational tax and expenditure data is unavailable for most African states, so we turn to two other proxy measures: the G-Econ dataset, which is a cell-based GPD estimate developed by Nordhaus

Table 3.5. Correlation with Road Density

	(1)	(2)	(3)	(4)	(5)
Mean Road Density	0.0653***	0.0649***	0.0663**	0.0566*	-0.134
	(0.0236)	(0.0236)	(0.0249)	(0.0285)	(0.155)
Constant	-0.000385	-0.000394	-0.000477	-0.000743	0.0153
	(0.0262)	(0.0272)	(0.0325)	(0.0419)	(0.154)
Unit of Analysis (Cell Size)	10km^2	100km^2	$1{,}000\text{km}^2$	$10{,}000\text{km}^2$	Country
Observations	2,989,002	307,053	33,088	4,066	52
R^2	0.004	0.004	0.004	0.003	0.018
Adj. R^2	0.00429	0.00423	0.00439	0.00298	-0.00203
Clusters	52	52	52	52	–
F Statistic	7.666	7.554	7.104	3.949	0.742
Prob > F	0.00783	0.00826	0.0103	0.0523	0.393

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

and Chen (2011), and the Visible and Infrared Imaging Suite (VIIRS) nocturnal light emissions data from the Earth Observation Group at the Colorado School of Mines. Economists have established a robust relationship between nighttime light emissions and economic development (Bluhm and Krause 2022). Harbers (2015) uses nighttime lights data as a measure of tax extraction and state capacity in Ecuador, while Koren and Sarbahi (2018) argues that "nighttime lights are reflective of the state's penetration and presence." We therefore expect a strong positive correlation between territorial control and these two indicators. To test this hypothesis, I estimate:

$$\text{Territorial Control}_i = \alpha + \beta \left(\text{Economic Activity}_i\right) + \epsilon_i \tag{3.7}$$

Where "Economic Activity" is measured either as mean cell-level GDP (PPP in 2005) or mean luminosity. Results are given in Tables 3.6 and 3.7.

Results show a small, though significant, positive correlation between territorial control and disaggregated GDP, a one standard deviation increase in GDP is associated with 0.2 standard deviation increase on average in territorial control. The results of the VIIRS analysis is less equivocal. A one standard deviation increase in nighttime lights emissions is associated with 0.1 to 1.2 standard deviation increase in territorial control, depending on the level of aggregation.

Table 3.6. Correlation with Economic Activity

	(1)	(2)	(3)	(4)	(5)
Mean Gridded GDP (PPP, 2005)	0.196*	0.208*	0.257*	0.286	0.0582***
	(0.106)	(0.113)	(0.149)	(0.175)	(0.0192)
Constant	-0.00421	-0.00485	-0.00582	-0.0155	-0.165***
	(0.0234)	(0.0243)	(0.0290)	(0.0362)	(0.0261)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Observations	2,807,199	288,370	31,057	3,808	50
R^2	0.039	0.044	0.070	0.123	0.093
Adj. R^2	0.0391	0.0440	0.0704	0.123	0.0741
Clusters	50	50	50	50	–
F Statistic	3.419	3.375	2.988	2.682	9.216
Prob > F	0.0705	0.0722	0.0902	0.108	0.00387

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.7. Correlation with Luminosity

	(1)	(2)	(3)	(4)	(5)
Mean VIIRS Nighttime Lights	0.122**	0.234***	0.444***	0.424***	1.198***
	(0.0489)	(0.0561)	(0.109)	(0.102)	(0.279)
Constant	-5.10e-05	-8.07e-07	-3.55e-05	0.00193	0.175*
	(0.0246)	(0.0243)	(0.0266)	(0.0350)	(0.100)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Observations	2,989,279	307,079	33,089	4,067	52
R^2	0.015	0.055	0.199	0.176	0.761
Adj. R^2	0.0150	0.0550	0.199	0.176	0.756
Clusters	52	52	52	52	–
F Statistic	6.272	17.39	16.44	17.30	18.49
Prob > F	0.0155	0.000118	0.000172	0.000122	7.90e-05

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

3.5 Discussion

Overall, the correlations provided in this chapter tend to support many of the empirical claims that Herbst (2014) makes in Part Three of *States and Power in Africa*—I find that territorial control is highly correlated with population density, urban territory, and transportation accessibility. I also show that Herbst's typology of African national design is broadly predictive of territorial control, providing empirical support for the classification scheme that Herbst outlines in Chapter 5. Perhaps the one surprising result from these analyses is the weak correlation between rugged terrain and territorial control. While the much of the literature in political science would lead us to believe that inaccessibility, as proxied by elevation, should decrease the levels of control that a state exercises in remote areas, it is possible that certain geographic idiosyncrasies of the African continent encourage states to situate assets at middling elevations with more hospitable climates—a proposition I examine in more detail in Chapter 4. It is also possible that there is some interaction between elevation and transportation networks, which these naïve correlations do not pick up, that attenuates accessibility issues in rugged terrain.

3.6 Appendix

Table 3.8. Summary Statistics By Aggregation Level (10km^2 Hexagonal Grid)

10km^2 Hexagonal Grid								
Variable	n, Cells	μ	σ	Min	25%	Median	75%	Max
Mean Control Point Density	2989475	0.00	0.00	0.00	0.00	0.00	0.00	0.59
Mean Facebook Pop. Dens.	64338	10.26	14.28	0.00	5.17	8.01	12.39	2028.74
Mean Facebook Pop. Dens.	3031098	0.22	2.55	0.00	0.00	0.00	0.00	2028.74
Mean WorldPop Pop. Dens.	3003098	37.72	342.89	0.00	0.35	4.27	18.20	80016.59
Mean Urban–Rural Extent	3019814	0.01	0.09	0.00	0.00	0.00	0.00	1.00
Mean Road Density	3019079	0.08	0.08	0.00	0.03	0.06	0.09	0.50
Mean Elevation	3013282	628.08	448.56	-154.00	315.36	494.20	906.67	5423.55
Standard Deviation of Elevation	3013282	17.76	28.47	0.00	3.84	8.34	17.92	513.58
Mean Gridded GDP (PPP, 2005)	2835022	0.70	3.08	0.00	0.03	0.10	0.38	85.04
Mean VIIRS Nighttime Lights	3019447	0.10	2.90	0.00	0.00	0.00	0.00	2141.01

Table 3.9. Summary Statistics By Aggregation Level (100km^2 Hexagonal Grid)

100km^2 Hexagonal Grid								
Variable	n, Cells	μ	σ	Min	25%	Median	75%	Max
Mean Control Point Density	307105	0.00	0.00	0.00	0.00	0.00	0.00	0.52
Mean Facebook Pop. Dens.	34873	10.58	16.15	0.00	5.67	8.55	12.73	2028.74
Mean Facebook Pop. Dens.	311507	1.18	6.35	0.00	0.00	0.00	0.00	2028.74
Mean WorldPop Pop. Dens.	309023	39.75	300.14	0.00	0.39	5.06	21.38	31617.95
Mean Urban–Rural Extent	310315	0.01	0.08	0.00	0.00	0.00	0.00	1.00
Mean Road Density	310239	0.08	0.08	0.00	0.03	0.06	0.09	0.50
Mean Elevation	309632	624.36	446.84	-120.81	312.73	492.25	901.64	4197.83
Standard Deviation of Elevation	309632	32.61	48.33	0.00	7.75	16.16	34.58	851.43
Mean Gridded GDP (PPP, 2005)	291221	0.70	3.05	0.00	0.03	0.10	0.39	85.04
Mean VIIRS Nighttime Lights	310272	0.11	1.69	0.00	0.00	0.00	0.00	474.62

Table 3.10. Summary Statistics By Aggregation Level (1,000km^2 Hexagonal Grid)

1,000km^2 Hexagonal Grid								
Variable	n, Cells	μ	σ	Min	25%	Median	75%	Max
Mean Control Point Density	33091	0.00	0.00	0.00	0.00	0.00	0.00	0.20
Mean Facebook Pop. Dens.	10984	11.55	15.87	0.00	6.29	9.39	13.77	1144.89
Mean Facebook Pop. Dens.	33606	3.77	10.57	0.00	0.00	0.00	6.08	1144.89
Mean WorldPop Pop. Dens.	33384	44.50	276.42	0.00	0.54	6.82	26.70	20393.79
Mean Urban–Rural Extent	33477	0.01	0.06	0.00	0.00	0.00	0.00	1.00
Mean Road Density	33473	0.08	0.08	0.00	0.03	0.06	0.09	0.50
Mean Elevation	33403	614.82	441.79	-93.77	305.85	486.95	890.79	3730.08
Standard Deviation of Elevation	33403	57.07	77.09	0.00	14.60	29.69	64.60	1027.42
Mean Gridded GDP (PPP, 2005)	31364	0.72	3.01	0.00	0.03	0.11	0.41	85.04
Mean VIIRS Nighttime Lights	33471	0.12	1.07	0.00	0.00	0.00	0.00	56.30

Table 3.11. Summary Statistics By Aggregation Level (10,000km^2 Hexagonal Grid)

10,000km^2 Hexagonal Grid								
Variable	n, Cells	μ	σ	Min	25%	Median	75%	Max
Mean Control Point Density	4067	0.00	0.00	0.00	0.00	0.00	0.00	0.08
Mean Facebook Pop. Dens.	2316	11.46	8.78	0.00	6.20	9.59	14.11	104.45
Mean Facebook Pop. Dens.	4147	6.40	8.69	0.00	0.00	3.29	10.41	104.45
Mean WorldPop Pop. Dens.	4117	49.31	185.24	0.00	0.96	10.04	36.13	7302.99
Mean Urban–Rural Extent	4129	0.02	0.06	0.00	0.00	0.00	0.01	1.00
Mean Road Density	4127	0.08	0.08	0.00	0.03	0.06	0.09	0.48
Mean Elevation	4121	592.74	424.15	-55.18	293.98	471.38	844.09	2776.71
Standard Deviation of Elevation	4121	96.31	113.60	0.00	25.70	54.02	119.66	870.21
Mean Gridded GDP (PPP, 2005)	3844	0.75	2.96	0.00	0.04	0.13	0.47	83.78
Mean VIIRS Nighttime Lights	4129	0.14	0.79	0.00	0.00	0.00	0.02	19.66

Table 3.12. Summary Statistics By Aggregation Level (Country Level)

Country Level								
Variable	n, Cells	μ	σ	Min	25%	Median	75%	Max
Mean Control Point Density	52	0.00	0.01	0.00	0.00	0.00	0.00	0.05
Mean Facebook Pop. Dens.	54	11.02	5.92	0.00	7.13	10.33	13.60	34.14
Mean Facebook Pop. Dens.	58	10.26	6.37	0.00	5.62	9.82	13.58	34.14
Mean WorldPop Pop. Dens.	58	96.91	130.46	1.82	19.40	49.47	107.16	532.00
Mean Urban–Rural Extent	58	0.05	0.10	0.00	0.00	0.01	0.03	0.59
Mean Road Density	58	0.11	0.10	0.00	0.04	0.07	0.13	0.38
Mean Elevation	58	583.85	432.90	23.25	267.36	452.70	858.31	2214.79
Standard Deviation of Elevation	58	276.54	165.93	15.61	140.15	254.04	371.16	711.75
Mean Gridded GDP (PPP, 2005)	51	0.83	0.92	0.04	0.18	0.46	1.07	3.93
Mean VIIRS Nighttime Lights	58	0.20	0.40	0.00	0.02	0.05	0.16	1.89

3.6.1 African National Design

Table 3.13. Herbst's Typology of African National Design

Difficult Geography	Hinterland Countries	Neutral Geography	Favorable Geography
Angola	Chad	Cameroon	Benin
Congo-Kinshasa	Mali	Ivory Coast	Botswana
Ethiopia	Mauritania	Ghana	Burkina Faso
Mozambique	Niger	Kenya	Burundi
Namibia		Malawi	Central African Republic
Nigeria		Uganda	Congo-Brazzaville
Senegal		Zambia	Equatorial Guinea
Somalia			Eritrea
Sudan			Gabon
Tanzania			Gambia
			Guinea
			Guinea-Bissau
			Lesotho
			Liberia
			Rwanda
			Sierra Leone
			Swaziland (Eswatini)
			Togo
			Zimbabwe
Herbst (2014, Table 5.1)			

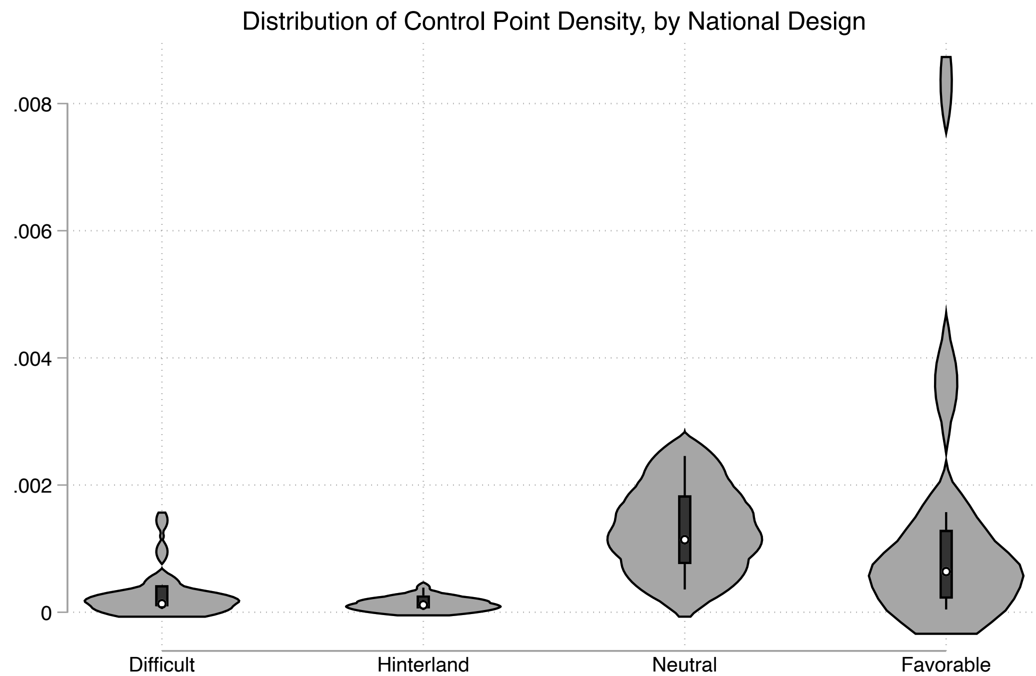


Figure 3.3. Control point density by African national design.

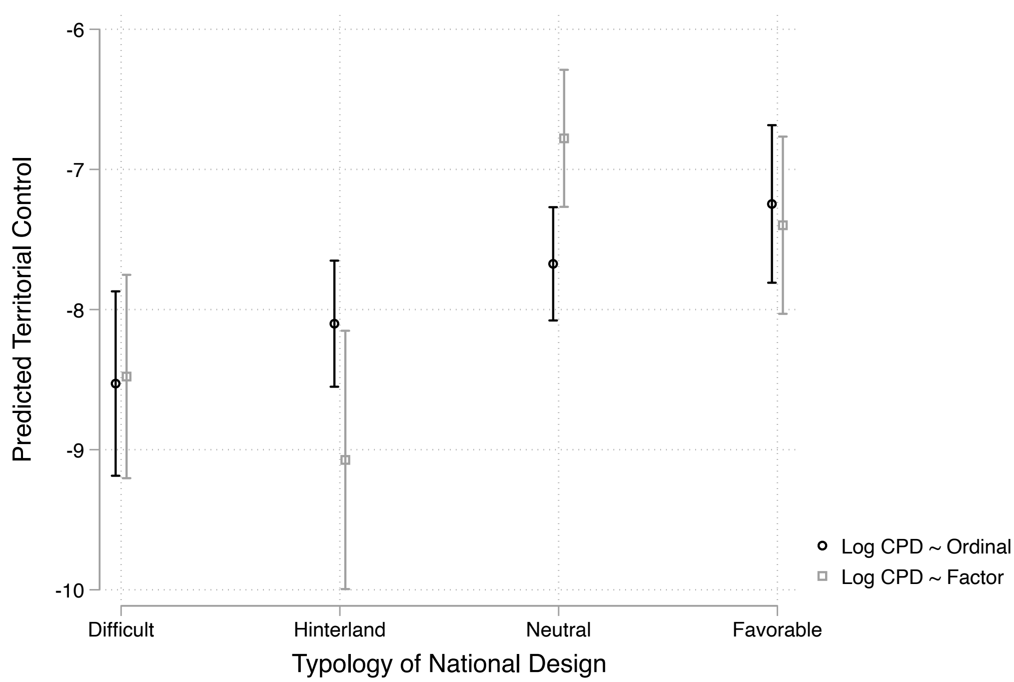


Figure 3.4. Control point density (log) by African national design.

3.6.2 Fractional Polynomial Models

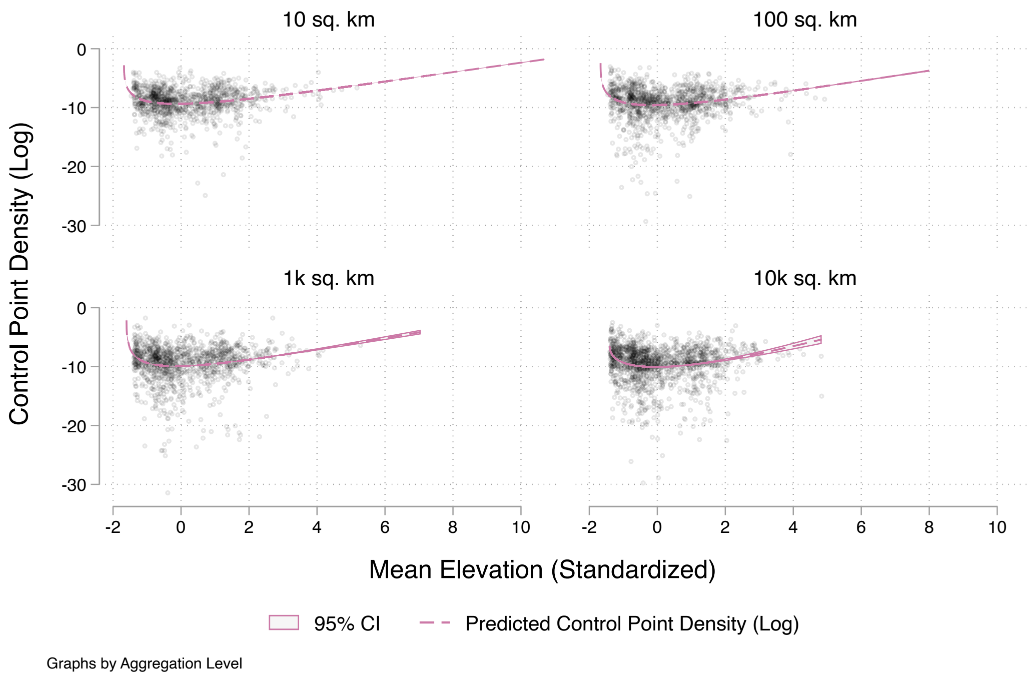


Figure 3.5. Fitted fractional polynomial models at different units of analysis. 2,000 randomly sampled observations are plotted at each level of aggregation to provide a sense of the bivariate distribution.

Table 3.14. Fractional Polynomial Models

	(1)	(2)	(3)	(4)	(5)
Mean Elevation (*p*1)	-1.055***	-7.021***	-7.679***	-6.108***	-1.322*
	(0.0163)	(0.101)	(0.297)	(0.497)	(0.764)
Mean Elevation (*p*2)	8.561***	2.910***	3.243***	2.625***	0.924*
	(0.0374)	(0.0402)	(0.121)	(0.213)	(0.484)
Constant	-2.664***	-2.428***	-2.134***	-6.513***	-6.534***
	(0.0295)	(0.0995)	(0.287)	(0.269)	(0.716)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1{,}000km^2$	$1{,}0000km^2$	Country
Observations	1,626,746	173,104	20,449	2,948	52
R^2	0.033	0.031	0.035	0.050	0.071
Adj. R^2	0.0326	0.0307	0.0345	0.0490	0.0335
F Statistic	27379	2743	366	76.93	1.884
Prob > F	0.000	0.000	0.000	0.000	0.163
Powers	{0.5, 0.5}	{0.5, 0.5}	{0.5, 0.5}	{0.5, 1}	{1, 1}

Standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

3.6.3 Fixed Effects Models

Table 3.15. Correlation With Population Density (Facebook)

	(1)	(2)	(3)	(4)	(5)
Mean Facebook Pop. Dens.	0.101**	0.0837*	0.0770*	0.124**	0.0939
	(0.0482)	(0.0446)	(0.0431)	(0.0511)	(0.136)
Constant	-9.68e-05**	-0.000181*	-0.000316*	-0.00123**	-0.00332
	(4.61e-05)	(9.65e-05)	(0.000177)	(0.000506)	(0.135)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Country FE	✓	✓	✓	✓	
Observations	2,989,475	307,105	33,091	4,067	52
R^2	0.011	0.007	0.006	0.012	0.008
Adj. R^2	0.0107	0.00699	0.00554	0.0119	-0.0115
F Statistic	4.414	3.522	3.189	5.926	0.474
Prob > F	0.0406	0.0663	0.0801	0.0185	0.494

Robust standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Table 3.16. Correlation With Population Density (WorldPop)

	(1)	(2)	(3)	(4)	(5)
Mean WorldPop Pop. Dens.	0.204***	0.255***	0.340***	0.352***	0.781*
	(0.0752)	(0.0893)	(0.100)	(0.121)	(0.450)
Constant	-0.00145***	-0.000816***	-0.000366***	0.000860***	0.108
	(4.63e-05)	(5.18e-05)	(4.89e-05)	(0.000292)	(0.166)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Country FE	✓	✓	✓	✓	
Observations	2,972,830	305,826	33,002	4,057	52
R^2	0.043	0.067	0.120	0.143	0.431
Adj. R^2	0.0428	0.0668	0.120	0.143	0.419
F Statistic	7.327	8.155	11.48	8.391	3.007
Prob > F	0.00922	0.00620	0.00136	0.00554	0.0890

Robust standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

3.6.4 Controlling for Island Countries

The following results include a binary control for island countries.

Table 3.17. Correlation with Road Density

	(1)	(2)	(3)	(4)	(5)
Mean Road Density	0.124**	0.122**	0.124**	0.122**	-0.134
	(0.0489)	(0.0487)	(0.0516)	(0.0501)	(0.155)
Constant	-0.000716**	-0.000738**	-0.000906**	-0.00149**	0.0153
	(0.000277)	(0.000293)	(0.000382)	(0.000575)	(0.154)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Country FE	✓	✓	✓	✓	
Observations	2,989,002	307,053	33,088	4,066	52
R^2	0.006	0.006	0.006	0.007	0.018
Adj. R^2	0.00607	0.00598	0.00645	0.00707	-0.00203
F Statistic	6.419	6.270	5.806	5.902	0.742
Prob > F	0.0144	0.0155	0.0196	0.0187	0.393

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.18. Correlation with Elevation (Standard Deviation)

	(1)	(2)	(3)	(4)	(5)
Standard Deviation of Elevation	0.0407	0.0513	0.0724	0.0881	-0.0816
	(0.0347)	(0.0429)	(0.0559)	(0.0652)	(0.0829)
Constant	1.09e-05	6.83e-06	6.37e-05	0.000631**	-0.00605
	(9.90e-05)	(0.000101)	(6.08e-05)	(0.000308)	(0.134)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Country FE	✓	✓	✓	✓	
Observations	2,983,010	306,431	33,017	4,059	52
R^2	0.001	0.002	0.004	0.007	0.006
Adj. R^2	0.00144	0.00217	0.00416	0.00658	-0.0136
F Statistic	1.376	1.428	1.674	1.825	0.968
Prob > F	0.246	0.238	0.202	0.183	0.330

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.19. Correlation with Elevation (Mean)

	(1)	(2)	(3)	(4)	(5)
Mean Elevation	0.0168	0.0126	-0.00138	-0.0179	-0.114
	(0.0264)	(0.0272)	(0.0307)	(0.0309)	(0.128)
Constant	-1.13e-05	1.92e-05	0.000156	0.000426	0.00473
	(0.000217)	(0.000234)	(0.000297)	(0.000367)	(0.144)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Country FE	✓	✓	✓	✓	
Observations	2,983,010	306,431	33,017	4,059	52
R^2	0.000	0.000	0.000	0.000	0.014
Adj. R^2	0.000157	8.53e-05	-2.92e-05	-3.72e-05	-0.00588
F Statistic	0.406	0.215	0.00202	0.335	0.788
Prob > F	0.527	0.645	0.964	0.566	0.379

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.20. Correlation With Population Density (Facebook)

	(1)	(2)	(3)	(4)	(5)
Mean Facebook Pop. Dens.	0.108**	0.0927*	0.0794*	0.0841*	0.0849
	(0.0507)	(0.0486)	(0.0400)	(0.0485)	(0.0738)
Island Nation	8.213***	4.895*	2.859	3.750	4.265**
	(2.980)	(2.779)	(2.141)	(2.600)	(1.886)
Constant	-0.00113	-0.00162	-0.00344	-0.0184	-0.167***
	(0.0249)	(0.0261)	(0.0322)	(0.0418)	(0.0301)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Observations	2,989,475	307,105	33,091	4,067	52
R^2	0.020	0.016	0.015	0.071	0.694
Adj. R^2	0.0204	0.0157	0.0151	0.0710	0.682
Clusters	52	52	52	52	–
F Statistic	6.070	3.339	2.767	2.238	3.234
Prob > F	0.00432	0.0434	0.0723	0.117	0.0480

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.21. Correlation With Population Density (WorldPop)

	(1)	(2)	(3)	(4)	(5)
Mean WorldPop Pop. Dens.	0.210***	0.262***	0.348***	0.376***	0.417**
	(0.0763)	(0.0901)	(0.0999)	(0.125)	(0.195)
Island Nation	8.100***	4.779*	2.760	3.518	3.445**
	(2.938)	(2.717)	(2.045)	(2.420)	(1.569)
Constant	-0.00247	-0.00221	-0.00338	-0.0156	-0.0748
	(0.0245)	(0.0257)	(0.0309)	(0.0371)	(0.0646)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Observations	2,972,830	305,826	33,002	4,057	52
R^2	0.053	0.076	0.131	0.206	0.785
Adj. R^2	0.0532	0.0759	0.131	0.206	0.776
Clusters	52	52	52	52	–
F Statistic	7.678	5.808	6.956	5.223	6.308
Prob > F	0.00122	0.00534	0.00213	0.00864	0.00365

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.22. Correlation with Urban Territory

	(1)	(2)	(3)	(4)	(5)
Mean Urban–Rural Extent	0.225***	0.299***	0.449***	0.569***	1.283***
	(0.0595)	(0.0818)	(0.123)	(0.132)	(0.228)
Island Nation	7.248**	3.837	1.835	2.080	-1.008
	(2.724)	(2.292)	(1.340)	(1.603)	(1.042)
Constant	-0.000905	-0.000873	-0.000905	-0.00204	0.224**
	(0.0230)	(0.0232)	(0.0278)	(0.0346)	(0.0916)
Unit of Analysis (Cell Size)	10km^2	100km^2	1,000km^2	10,000km^2	Country
Observations	2,989,475	307,105	33,091	4,067	52
R^2	0.059	0.095	0.204	0.352	0.920
Adj. R^2	0.0590	0.0953	0.204	0.352	0.916
Clusters	52	52	52	52	–
F Statistic	11.56	8.834	8.517	9.916	69.03
Prob > F	7.22e-05	0.000508	0.000643	0.000230	0

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.23. Correlation with Road Density

	(1)	(2)	(3)	(4)	(5)
Mean Road Density	0.0663***	0.0663***	0.0695***	0.0749***	0.0558
	(0.0236)	(0.0237)	(0.0250)	(0.0259)	(0.0443)
Island Nation	8.394***	4.996*	2.922	3.798	4.334**
	(2.978)	(2.808)	(2.156)	(2.620)	(1.934)
Constant	-0.00144	-0.00185	-0.00368	-0.0187	-0.173***
	(0.0262)	(0.0272)	(0.0324)	(0.0392)	(0.0229)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Observations	2,989,002	307,053	33,088	4,066	52
R^2	0.013	0.011	0.014	0.070	0.690
Adj. R^2	0.0131	0.0115	0.0136	0.0695	0.678
Clusters	52	52	52	52	–
F Statistic	7.870	5.474	4.745	5.083	3.255
Prob > F	0.00105	0.00702	0.0129	0.00971	0.0471

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.24. Correlation with Elevation (Standard Deviation)

	(1)	(2)	(3)	(4)	(5)
Standard Deviation of Elevation	0.0553*	0.0669*	0.0877**	0.104**	0.0310
	(0.0293)	(0.0349)	(0.0436)	(0.0490)	(0.0246)
Island Nation	8.337***	4.913*	2.876	3.759	4.290**
	(2.957)	(2.799)	(2.141)	(2.608)	(1.933)
Constant	-0.00107	-0.00146	-0.00309	-0.0169	-0.163***
	(0.0245)	(0.0252)	(0.0300)	(0.0368)	(0.0281)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Observations	2,983,010	306,431	33,017	4,059	52
R^2	0.012	0.011	0.016	0.075	0.688
Adj. R^2	0.0117	0.0115	0.0164	0.0745	0.676
Clusters	52	52	52	52	–
F Statistic	5.849	3.407	2.868	3.268	3.247
Prob > F	0.00517	0.0408	0.0660	0.0462	0.0474

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.25. Correlation with Elevation (Mean)

	(1)	(2)	(3)	(4)	(5)
Mean Elevation	0.0218	0.0182	0.00860	-0.00630	0.0528
	(0.0200)	(0.0207)	(0.0237)	(0.0293)	(0.0447)
Island Nation	8.392***	4.952*	2.861	3.711	4.327**
	(2.958)	(2.810)	(2.157)	(2.622)	(1.931)
Constant	-0.00110	-0.00147	-0.00306	-0.0171	-0.169***
	(0.0251)	(0.0262)	(0.0314)	(0.0386)	(0.0253)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Observations	2,983,010	306,431	33,017	4,059	52
R^2	0.009	0.007	0.009	0.064	0.690
Adj. R^2	0.00916	0.00737	0.00884	0.0639	0.678
Clusters	52	52	52	52	–
F Statistic	4.607	1.929	0.935	1.033	3.161
Prob > F	0.0145	0.156	0.399	0.363	0.0511

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3.26. Correlation with Luminosity

	(1)	(2)	(3)	(4)	(5)
Mean VIIRS Nighttime Lights	0.122**	0.233***	0.443***	0.412***	0.787**
	(0.0485)	(0.0557)	(0.109)	(0.0923)	(0.325)
Island Nation	8.411***	4.875*	2.708	3.416	2.057
	(2.941)	(2.810)	(1.999)	(2.340)	(1.368)
Constant	-0.00108	-0.00140	-0.00298	-0.0141	0.0359
	(0.0246)	(0.0242)	(0.0264)	(0.0319)	(0.107)
Unit of Analysis (Cell Size)	$10km^2$	$100km^2$	$1,000km^2$	$10,000km^2$	Country
Observations	2,989,279	307,079	33,089	4,067	52
R^2	0.024	0.062	0.207	0.230	0.831
Adj. R^2	0.0237	0.0618	0.207	0.229	0.824
Clusters	52	52	52	52	–
F Statistic	7.192	10.18	9.106	10.23	12.16
Prob > F	0.00177	0.000191	0.000415	0.000183	5.15e-05

Robust standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

3.7 Works Cited

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Alesina, Alberto, Enrico Spolaore, and Romain Wacziarg. 2005. "Trade, Growth and the Size of Countries." In *Handbook of Economic Growth,* edited by Philippe Aghion and Steven N. Durlauf, 1:1499–1542. Elsevier.

Baneshi, Mohammad Reza, Fatemeh Nakhaee, and Matthew Law. 2013. "On the Use of Fractional Polynomial Models to Assess Preventive Aspect of Variables: An Example in Prevention of Mortality Following HIV Infection." *International Journal of Preventive Medicine* 4 (4): 414–419.

Bluhm, Richard, and Melanie Krause. 2022. "Top Lights: Bright Cities and Their Contribution to Economic Development." *Journal of Development Economics* 157:102880.

Buhaug, Halvard, and Jan Ketil Rød. 2006. "Local Determinants of African Civil Wars, 1970–2001." *Political Geography* 25 (3): 315–335.

Collier, Paul, and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56 (4): 563–595.

Collier, Paul, Anke Hoeffler, and Dominic Rohner. 2009. "Beyond Greed and Grievance: Feasibility and Civil War." *Oxford Economic Papers* 61 (1): 1–27.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75–90.

Gerrard, A. J. W. 2000. *What Is a Mountain?*

Gottmann, Jean. 1975. "The Evolution of the Concept of Territory." *Social Science Information* 14 (3): 29–47.

Harbers, Imke. 2015. "Taxation and the Unequal Reach of the State: Mapping State Capacity in Ecuador." *Governance* 28 (3): 373–391.

Heggie, Ian G. 1995. *Manangement and Financing of Roads. An Agenda for Reform.* Technical report Paper 275. Washington, DC: World Bank.

Herbst, Jeffrey. 2014. *States and Power in Africa: Comparative Lessons in Authority and Control - Second Edition.* Princeton University Press.

Kalyvas, Stathis N., and Matthew Adam Kocher. 2009. "The Dynamics of Violence in Vietnam: An Analysis of the Hamlet Evaluation System (HES)." *Journal of Peace Research* 46 (3): 335–355.

Koren, Ore, and Anoop K Sarbahi. 2018. "State Capacity, Insurgency, and Civil War: A Disaggregated Analysis." *International Studies Quarterly* 62 (2): 274–288.

Livingston, Steven, and Gregor Walter-Drop, eds. 2014. *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood.* Illustrated edition. Oxford ; New York: Oxford University Press.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112 (4): 725–753.

Nordhaus, William, and Xi Chen. 2011. *Geographically Based Economic Data (G-Econ).*

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *The Review of Economics and Statistics* 94 (1): 20–36.

Riley, S. J., Stephen D. DeGloria, and Robert Elliot. 1999. "A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity." *Intermountain Journal of Sciences* 5 (1-4): 23–27.

Royston, Patrick, and Douglas G. Altman. 1994. "Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling." *Journal of the Royal Statistical Society Series C: Applied Statistics* 43 (3): 429–453.

Royston, Patrick, and Willi Sauerbrei. 2008. *Multivariable Model - Building: A Pragmatic Approach to Regression Anaylsis Based on Fractional Polynomials for Modelling Continuous Variables.* John Wiley & Sons.

Shaver, Andrew, David B. Carter, and Tsering Wangyal Shawa. 2019. "Terrain Ruggedness and Land Cover: Improved Data for Most Research Designs." *Conflict Management and Peace Science* 36 (2): 191–218.

Chapter 4

Political Ecology of the African State

Abstract

States seldom control the entirety of the territory defined by their de jure borders. African states in particular are known for their incomplete consolidation; in many African countries, there exist large swathes of territory in which the state is either entirely absent, or exercises minimal control. I explore this phenomenon of incomplete consolidation by addressing two fundamental questions in comparative politics: First, where do states choose to locate their assets and materiel, and thus exercise direct control over territory? And second, why do states elect to control certain territories within their jurisdictions, but not others? To answer these questions, I combine the dataset described in Chapter 2 with a variety of geospatial data from publicly available sources, and use these data to characterize the "ecological niche" of the African state. This niche is defined by a set of spatially-variant strategic, demographic, and economic features that make a given geography conducive to state colonization. I find that a core set of features—economic activity, population density, and proximity to other state resources—are strong predictors of the overall spatial distribution of state infrastructure across the continent. Geographically localized analyses, however, reveal a more nuanced picture, one which underscores the importance of market access, agricultural productivity, and durable infrastructure in determining where government assets are situated in a given region.

4.1 Introduction

Authorities in Africa have long struggled to exercise control over the large, sparsely populated regions that comprise the majority of the African continent (Herbst 2014). Even today, many African states remain poorly consolidated. Vast stretches of the Sahel, the Congo Basin, and the Horn of Africa are effectively ungoverned space—areas in which the central government is unable to maintain a monopoly on violence or to enforce basic political decisions (Livingston and Walter-Drop 2014). Despite the widespread recognition by scholars, policymakers, and aid practitioners that such "areas of limited statehood" exist, and that they pose significant challenges to regional security, economic development, and public goods provision, there are very few systematic attempts to identify and map these areas, or to explain how they are distributed across a state's geography.[1]

This chapter presents one potential mapping of governed and ungoverned space in Africa. Using publicly-sourced data on the location of state infrastructure across the continent, I develop a density-based measure of territorial control, which allows researchers to visualize the contours of state power at a very high level of resolution. This approach is not entirely new. A string of recent papers in comparative politics use geospatial data on the location of government facilities to measure state capacity and control. Cappelen and Hariri (2022), for example, estimate the reach of the early modern European state by looking at the proportion of medieval castles in a given area that are controlled by the Crown. Proximity to such installations, they argue, is a reasonable proxy of the state's local monopoly on violence. Müller-Crepon (2021), Henn (2022), and Fergusson et al. (2022), in studies of public goods provision, operationalize state capacity as distance to a national or provincial capital. Much of this work, however, takes the location of government facilities as exogenous.[2] Where castles and capitals are located is treated as an accident of history, or an arbitrary decision on the part of state leaders. This chapter contributes to this emerging literature by explicitly considering *site choice* as an outcome of interest: Why do states elect to situate a military base, a capital, or some other piece of infrastructure in one place rather than another? Or more generally, what factors influence the distribution of state assets across territory?

1. A number of problems are either explicitly or implicitly attributed to the presence of ungoverned space. Fearon and Laitin (2003) argue that insurgent violence tends to occur in rugged terrain, beyond the reach of the central government. Besley and Persson (2010), Hendrix (2010), and Kalyvas (2015) find that ungoverned space inhibits a state's ability to deter violent challenges to governmental authority. Müller-Crepon (2021) shows that economic development and the provision of public goods and social services are correlated with state presence.

2. There is an older, more qualitative literature on the location of national capitals; see Spofford (1881). There is also a more recent literature on national capital relocation (Potts 1985, which focuses on Lilongwe; Schatz 2004; Rossman 2018; Ishenda and Guoqing 2019; Rachmawati et al. 2021), though these studies tend to focus on the *decision* to relocate rather than the actual site selection. The most advanced research in this vein comes from the field of strategic studies. Military tacticians have long been interested in the placement of military assets in theatre (Bell and Griffis 2015), and the extent of territory controlled by such assets (e.g., Tao et al. 2016).

To address this question, I begin with the simple premise that the geographic patterns of power and authority within a given country are not arbitrary. As Herbst (2014) makes clear, the exercise of control comes at a significant cost to the state. The construction of new administrative offices and security installations, the maintenance of existing facilities, and the continuous deployment of soldiers and civilian personnel represent a substantial—and generally persistent—capital investment on the part of the government. State leaders must therefore make canny decisions about how best to allocate control across territory, given their respective political and financial constraints. I argue that, in making these types of decisions, leaders will tend to prioritize regions that are either strategically important to the survival of the state and the incumbent regime, or regions rich in taxable resources, such as labor, market transactions, and agricultural production and extractive reserves. Because this theory posits that various *geographical* characteristics (i.e., qualities that vary across physical space) influence the *spatial distribution* of state control, I take an overtly geospatial approach to identify the environs amenable to state presence, and to test various hypotheses about the correlates of subnational state consolidation in Africa.

This chapter has two main goals. The first is strictly descriptive—to map the "political topographies" of the African state (Boone 2003), and to characterize the conditions favorable to state colonization. In particular, I employ a technique known as ecological niche modeling (ENM) to identify a set of geographic features useful in predicting the occurrence and frequency of state assets (or what I refer to as "control points") in a given location. This approach is widely used in the fields of ecology, bio-geography, and urban planning, though rarely applied in political science or economics. The second goal is empirical. After identifying a set of salient strategic, economic, and demographic characteristics, I estimate their influence on the location-specific occurrence (i.e., the extensive margin of state control) and frequency (i.e., the intensive margin) the of state assets, and evaluate how their influence changes across space, both between and within formally defined administrative units. I find that the most robust predictors of territorial control tend to be access to markets and other economic assets, particularly distance to airports and major cities, agricultural production, and localized economic activity. However, the relative influence of these factors differs by location, suggesting that states have diverse priorities in determining patterns of control within their respective borders, and that national leaders may even assign varying degrees of importance to individual factors at the subnational level in response to local conditions. Finally, I find that those features that predict the occurrence of state assets are not necessarily the same predictors of the frequency of state assets. In some areas, a single outpost of the state is sufficient to fulfill a government's objectives, while other areas seem to require a greater concentration of resources to achieve an optimal level of governance.

The remainder of the chapter is organized as follows: Section 4.2 outlines a theory of state consolidation and proposes a set of hypotheses derived from this framework. Section 4.3 summarizes the data used to test these hypotheses. Sections 4.4 through 4.6 engage in a number of empirical exercises, beginning with a series of conventional models that serve as a baseline for the geospatial analyses presented in Sections 4.5 and 4.6. In Section 4.5, I detail two continent-scale ecological niche models that predict, with a high degree of accuracy, the location-specific occurrence and frequency of state infrastructure across the continent. Section 4.6 uses geographically weighted regression to provide a direct test of the hypotheses laid out in Section 4.2. Section 4.7 concludes with a discussion of my results and potential avenues for further research.

4.2 Theoretical Framework

I begin with the assumption that leaders *actively* and *purposively* decide where to position state assets in their efforts to extend control over the territory contained within a country's *de jure* borders. Power is expensive to broadcast over distance, and capital—both political and financial—is scarce, even in high capacity states like the U.S. and Germany (Hendrix 2010; Herbst 2014; Warren 2014). Rational leaders should therefore choose to allocate resources to areas in which they expect a net positive, or at least a net neutral, return on their investment. This return is generally understood to be pecuniary, taking the form of taxation (Friedman 1977; North 1982; Lake 1992; Olson 1993). But leaders are also concerned with the survival of the state and the incumbent regime. This suggests that there is some value derived from mitigating the internal and external threats posed by rebel groups, foreign actors, and other political organizations that compete with the central government for influence and resources. I argue that these economic and strategic considerations shape not only the patterns of power and authority that we observe in modern states, but also the spatial distribution of physical infrastructure—the facilities and agents that states employ to surveil populations, extract rents, and induce compliance (Scott 1999).

Framed in this way, the process of state consolidation is essentially a textbook example of a multi-objective spatial optimization problem, similar to those faced by retail firms as they decide where to locate a brick-and-mortar outlet, such as a coffee shop or a department store. Just as retailers make decisions about which markets to capture, and where to locate their physical and human infrastructure to maximize profits, state builders will attempt to partition the polity into sets of governed and ungoverned spaces, and to situate the state's physical and human infrastructure in a manner that both maximizes rents and minimizes security threats and management costs.[3] Indeed, these maximands—expected rents, the strategic advantages derived from controlling a particular territory, and the contiguity and coverage of that controlled territory—represent the three main criteria that leaders consider in their site selection decisions. Rent extraction is one of the core features of contemporary states; prevailing microeconomic theories of the state conceive of the state as a profit maximizing firm that trades protection services for revenue (Lake 1992). The objective of capturing strategically-important regions follows from the state's desire to achieve an internal monopoly on violence, and the geographic properties of contiguity and coverage relate to the costs associated with maintaining control. Contiguous territory reduces transportation and communication outlays, while greater coverage allows states to benefit from economies of scale.

In practice, state builders generally forego the sophisticated GIS and multi-objective optimization tools used by contemporary firms to guide their decision-making processes. Leaders instead muddle through, using an "adjust and evaluate" approach to arrive at a reasonable solution to the spatial optimization problem. The obvious starting point is to grow the state from a set of preexisting "seed locations." These are areas in which some degree of territorial control is inherited from the previous period. When Kenya achieved independence in 1963, for example, the ruling KANU-KADU coalition did not begin building a state from scratch; Kenya was not *terra nullius*. Rather, the nascent government inherited a number of control points—facilities such as police stations, administrative offices, and other premises—from the British colonial authorities, and built out this infrastructure incrementally over time. Our baseline expectation, then, is that state

3. In other work, I incorporate time into the theory. Territorial control is a normal good, and the persistence of ungoverned space is generally seen as undesirable. I model the process of state consolidation as a dynamic market entry game in which leaders attempt to maximize expected inter-temporal profits. Based on a set of parameters, leaders will elect to either enter a space, or write it off until a later period.

infrastructure will tend to cluster in space. More formally, we should see a high degree of spatial autocorrelation in the distribution of control points:

Hypothesis 1: *The occurrence and frequency of control points in a given neighborhood is positively and significantly correlated with the occurrence and frequency of control points in adjacent neighborhoods.*

Importantly, this hypothesis is not simply a restatement of Tobler's first law of geography.[4] In most countries, there is no overt policy that requires government facilities to cluster in this manner. By situating new assets proximate to existing ones, states are able to gradually increase their reach through a process of accretion, in a manner similar to the "oil-spot" counterinsurgency strategies employed by U.S. forces in Iraq and Afghanistan. The colocation of control points also allows for states to reduce administrative overhead, simplify supply chains, and ease access to complementary government resources, such as technocratic expertise and security services.

Because a state's budget constraints preclude the possibility of expanding indiscriminately, we expect leaders to concentrate resources in areas that are economically and strategically important to the survival of the state. As profit-maximizing entities, a state's location-allocation decisions are driven in large part by the flow of tax revenue to the central government. Regions with a high concentration of labor, firms, and natural resources are especially attractive, given the broad tax base:

Hypothesis 2: *Control points are likely to occur more often and with increased frequency in areas with high-value economic assets, and less often and with decreased frequency in areas with few resources or little economic activity.*

Situating government facilities in these economically salient areas has two major advantages: First, it facilitates tax collection. When agents of the state are physically present in a given location, they are better able to monitor firms and other market participants to ensure accurate measures of taxable activity (Scott 1999; Gordon and Li 2009; Balán et al. 2022); they are also able to induce tax compliance through messaging, audits, and sanctioning (Tilly 1992; Pomeranz and Vila-Belda 2019). Second, it helps to secure property rights, which are integral to economic development and the smooth functioning of markets. This, in turn, fosters investment and grows the regional tax base, ensuring maximal rents over time (North 1982; Olson 1993; Besley 1995).

4. "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970).

Profit maximization is futile if the state's time horizons are short. Leaders will not find it worthwhile to invest in territorial control if they expect a competitor to capture or replace the state.[5] To maintain their rule, leaders will attempt to eliminate internal and external threats by positioning state assets in strategically important regions, especially those home to recalcitrant populations capable of mounting violent collective action against the incumbent regime, or those prone to incursion by other states and non-state actors.

Hypothesis 3: *Control points are likely to occur more often and with increased frequency in areas of high strategic importance to the state, and less often and with decreased frequency in areas of low strategic importance.*

State presence in these areas is important insomuch as it increases the costs of insurgency, thereby deterring rebellion and foreign meddling. By positioning instruments of surveillance and repression in areas susceptible to colonization by competing institutions, states are better able to engage in exclusionary practices—essentially managing the flow of people, goods, capital, and information throughout its territory.

4.3 Data

Although recent work emphasizes the role of soft power in consolidating control over territory, I focus on the more traditional operationalization of state strength as the material resources used to coerce the compliance of a population (Warren 2014). In most modern states, power and authority are broadcast across space from a constellation of "control points," or government-owned facilities that serve to surveil local populations, enforce edicts, and regulate behavior. These control points are, in many ways, analogous to the networks of cellular towers operated by mobile carrier services that receive and transmit signals within a surrounding catchment area, and relay information to a central exchange, which in turn allocates resources and moderates informational traffic throughout the system. Similar to a mobile carrier service, the state is concerned with both *coverage* and *signal strength*. Coverage is determined by the presence of a cellular tower in a given neighborhood, while

5. The African independence period provides a compelling example of this dynamic. During the late 1950s and early 1960s, the French *corps des administrateurs coloniaux* (Colonial Administration Service), which administered colonies in tropical Africa, began to draw down overseas staffing levels in anticipation of the transfer of power to newly independent African states. By 1960, thousands of French colonial administrators were forcibly retired, or reassigned to other departments in the French civil service (Dimier 2004).

signal strength is generally determined by the frequency of towers in a given neighborhood. This mobile carrier service analogy suggests that a density-based measure, which incorporates both the presence and frequency of control points across a state's jurisdiction, may be an appropriate measure of territorial control.

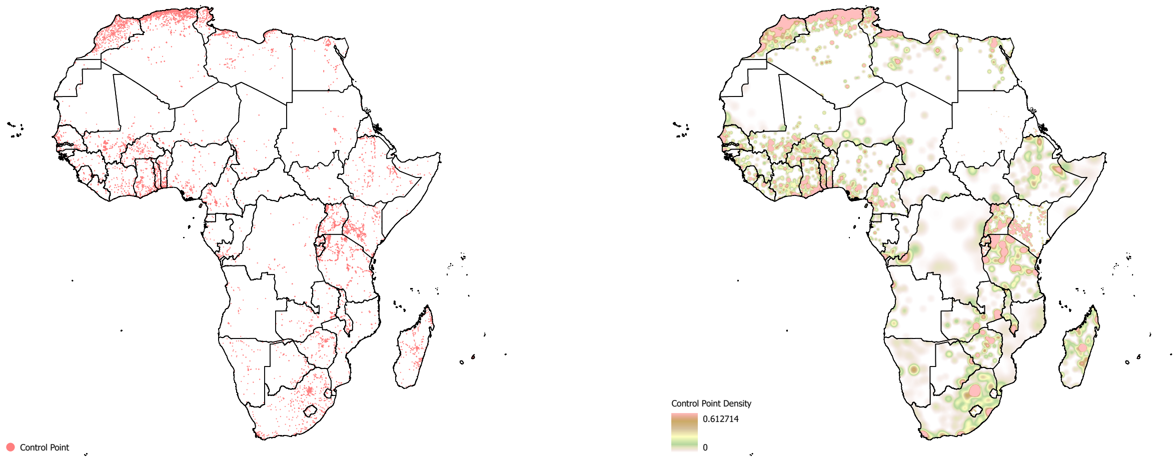
To construct such a measure, I scraped several commercial and public repositories of GIS data, including Google Maps, Bing Maps, Open Street Map, and Wikimapia, for the locations of government-controlled facilities throughout the African continent. Data scraped from these sources were then reconciled to remove duplicate observations, and to merge similar observations. Steps were taken to validate the data using publicly available databases, including lists and websites published by the news media, national and local governments, and private citizens. Given the patchy web presence of many African state agencies, however, the overwhelming majority of locations could not be verified by a secondary source. Included facilities fall into seven broad categories:

1. Police, military, and security services,
2. General government services (e.g., identity card offices and tax assessors),
3. Post offices,
4. General purpose government / public buildings (e.g., city halls, federal buildings),
5. Courts and magistrates,
6. Emergency services,
7. Forts, barracks, bivouacs, and training facilities.

It is worth noting that the selection criteria deliberately excludes facilities that are primarily geared towards public services and entitlements, such as social welfare offices and schools, instead prioritizing locations that discharge the first order concerns of the state—monitoring, enforcement, and extraction.

The resulting dataset includes 19,419 individual control points, which are mapped in Figure 4.1a. Roughly 36% of control points are classified as security infrastructure, which includes police stations, military bases, and other emergency services. An additional 32% involve non-security-related government offices of varying types, 22% are post offices, and 4% of locations consist of judicial facilities such as courts, magistrates, and prisons.

The raw data are used to generate a modified kernel density surface, shown in Figure 4.1b, which represents governed and ungoverned space across the continent. To create this surface, a smoothed line is fit over each control point, i (i.e., each location in Figure 4.1a), and the surrounding



(a) Control Point Locations

(b) Control Point Density (Raw)

Figure 4.1. Density based measure of governed and ungoverned space in Africa.

density is estimated with:

$$Density = \frac{1}{Bandwidth^2} \sum_{i=1}^{n} \left(\frac{3}{\pi} \left(1 - \left(\frac{Distance_i}{Bandwidth} \right)^2 \right)^2 \right), \tag{4.1}$$

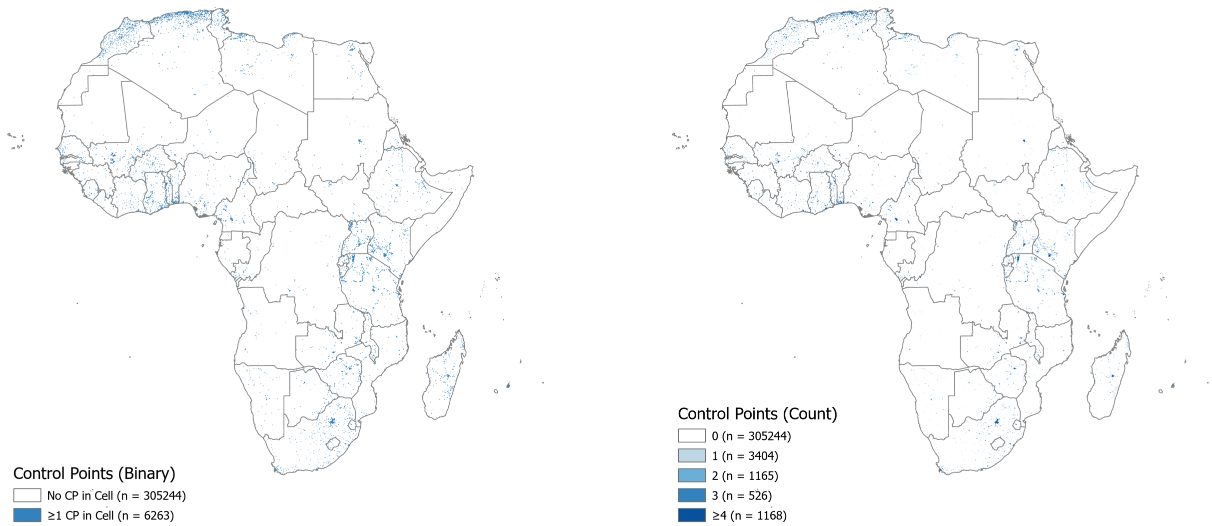
where *Bandwidth* is an optimal search radius defined by Silverman (1986), and *Distance* is the geodesic distance between the surface pixel to be estimated and control point i. The local density value is highest at a given control point, and gradually decays over space in a radial manner, which accounts for the leopard–print pattern that characterizes the map in Figure 4.1b.[6]

4.3.1 Outcome Variables & Features

Because I am interested in modeling both the *presence* and *frequency* of control points across the continent, I generate two primary outcome variables for this analysis—a binary presence indicator and a frequency variable. I construct these variables by overlaying a 100 square kilometer hexagonal grid across the study area, and summing all control points within each cell. Cells containing at least one control point are coded as $1 =$ state presence, and 0 otherwise (state absence). The frequency variable is the raw count of control points within each cell. In the Appendix, I also show two alternative measures of the dependent variable; the first is the log count of control points, which I use as an

6. Density estimation is carried out on a country-by-country basis in order to prevent control point locations in one country from influencing density estimates in a neighboring country. The resulting surfaces are then stitched together into a single high-resolution (1km) raster to provide a continuous estimate for the entire continent.

alternative specification for the random forest models, and the second is the control point density measure detailed above. For reasons explained later in the paper, I use this density variable to estimate the country-level geographically weighted regression models in Section 4.6.[7]



(a) Control Points (Binary)

(b) Control Points (Count)

Figure 4.2. Spatial distribution of the two primary outcome variables. Note that the unit of analysis is the 100km^2 hexagonal grid cell. Cells containing at least one control point are quite rare; they make up only 2% of the total.

To test the three hypotheses detailed in Section 4.2, I assemble a set of predictor variables (or "features") from a variety of sources. Summary statistics and data sources are given in Table 4.3 in the Appendix. All features are aggregated to the same 100 square kilometer hexagonal cell used to construct the outcome variables for the main analyses, though I use a less coarse aggregation level for the continent-scale GWR models in Section 4.6, for reasons explained below.[8]

I operationalize economically salient areas as urban agglomerations with high levels commercial activity, as well as logistics hubs with air and maritime ports capable of transporting goods to domestic and international markets. I also include several standard measures of economic productivity: nighttime light emissions, which proxies for economic activity and wealth (Mellander

7. Because the data include observations for which Control Points (Count) = 0, I calculate the log frequency as ln(Control Points (Count) + 1). This transformation has two advantageous properties: it retains 0 count cells (as $\ln(1) = 0$), and this type of transformation has been shown to improve prediction accuracies for Random Forest models of count data (Stevens et al. 2015).

8. To deal with missing data issues and maximize the number of observations in each model, I use mean imputation to estimate missing values for the World Bank data.

et al. 2015; Weidmann and Schutte 2017); total agricultural production values by area; and distance to proven mineral and energy reserves and refinement or processing facilities. As a coarse measure of limited economic activity, I include local poverty headcount ratio estimates.

Strategic areas are more difficult to operationalize. I focus on three core attributes: political geography, human geography, and proximity to durable infrastructure. A country's political geography is defined by the borders that delineate the polity, the location of the capital city, and the incidence of conflict and political violence. I therefore include the distances to each of these attributes, as well as the location-specific count of various types of conflict events. I capture various facets of a country's human geography with measures of population density, the number of politically relevant ethnic groups in a given area, the number of politically excluded groups in a given area, and the presence of trans-border ethnic kin. Because human settlement patterns in Africa have traditionally favored low-lying coastal regions and more temperate regions not subject to periods of extreme heat and drought, I also include an indicator for Köppen–Geiger climate classification. Finally, I consider three measures of durable infrastructure: road density, and distance to power plants and high voltage power transmission lines.

4.3.2 Spatial Autocorrelation

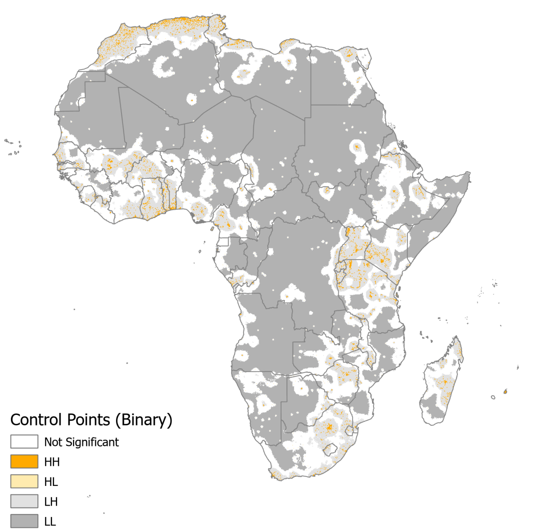
Because we expect some degree of spatial autocorrelation in the data (i.e., geographically adjacent observations have similar values), I calculate Global Moran's I and Local Indicator of Spatial Autocorrelation (LISA) statistics (Anselin 1995) for the outcome variables. Tests for Moran's I, shown in Table 4.1, reveal positive and significant levels of spatial autocorrelation, indicating a high degree of spatial clustering among control points.

Table 4.1. Spatial Autocorrelation of Outcome Variables

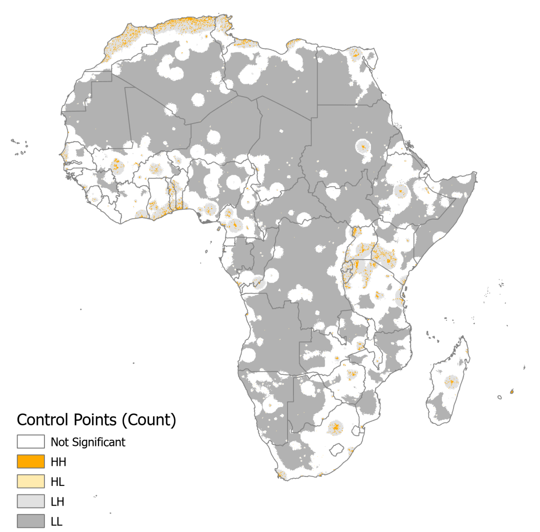
Outcome Variable	Moran's I	z-Score
Binary	0.110413***	739.52
Frequency	0.056389***	383.09
Log Frequency	0.121026***	810.73
Mean Control Point Density	0.277596***	1874.01

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

The LISA statistics (Anselin's Local Moran's I), which are mapped in Figure 4.3, are slightly more informative; these statistics provide a measure of local spatial autocorrelation and non-stationarity, and allow us to visualize clusters and outliers in the outcome data. The maps in Figure 4.3 show a series of High-High clusters (high control point density cells bordering other high control point density cells) in the Great Lakes, the Congo Basin, the Mediterranean coast, and along the Gulf of Guinea, as well as the Addis Ababa, Pretoria-Johannesburg, Antananarivo, and Khartoum-Omdurman capital regions. These clusters are surrounded by Low-High and High-Low (low [high] control point density cells bordering high [low] control point density cells) outlier regions. These are essentially transition zones, in which areas of concentrated state presence give way to hinterlands with low densities of state assets. Areas with no significant spatial autocorrelation (those shaded white) are areas in which control points are effectively randomly distributed—there is no statistically discernible pattern to either the occurrence or frequency of control points in these locales.



(a) Control Points (Binary)



(b) Control Points (Count)

Figure 4.3. Local cluster maps for the two primary outcome variables. Dark orange indicates statistically significant clusters of high values (e.g., presence and frequency of control points), while dark gray indicates statistically significant clusters of low values (e.g., absence of control points). Light orange and light gray indicate statistically significant outlier regions (e.g., transition zones).

The high degree of spatial autocorrelation is notable for three reasons. First, significant spatial autocorrelation in the dependent variable implies that state infrastructure is *not* randomly distributed throughout space. This supports the core assumption of this chapter that state assets are not arbitrarily scattered across a country's geography. Second, it provides some confirmation of

Hypothesis 1—that policy makers will tend to "grow" the state incrementally from a set of seed points, situating state assets in close proximity to existing assets. Finally, spatial autocorrelation complicates the spatial methods used throughout this paper. Random forest models, like those presented in Section 4.5, are sensitive to spatial autocorrelation (Dormann et al. 2007; Lichstein et al. 2002; Sinha et al. 2019). Spatial autocorrelation is also problematic for the geographically weighted regression (GWR) models presented in Section 4.6, as GWR builds a local regression equation for each location in the dataset, which gives rise to issues of local multicollinearity. When there is insufficient local variation in either the dependent or the independent variables, it becomes difficult to estimate local coefficients. I take a number of steps to mitigate this issue, which I describe below.

4.3.3 Mahalanobis Distance

Before I formally evaluate the hypotheses laid out in Section 4.2, I first want to get a sense of whether or not there is any discernible similarity in the areas in which we tend to observe control points. In other words, do state leaders favor a particular *type* of geography in which to locate state assets? To answer this question, I calculate Mahalanobis D^2, which is the standardized multivariate distance between the values of the features (the predictor variables) measured at a given place, and the mean values for those same features across all observed control points. The more "similar" the conditions in a given cell are to the mean conditions of all cells that contain a control point, the smaller the D^2 distance, and thus the more "suitable" that point is for colonization. The underlying intuition is that the the closer a point lies to another in n-dimensional space, the more likely these two points are to belong to the same set. This is similar to other multidimensional scaling methods used in political science, such as such as the NOMINATE method for assessing legislative roll call votes (Poole and Rosenthal 1985), or propensity score matching when assignment to treatment is not randomized.[9] Just as we expect co-partisan legislators to cluster together in n-dimensional policy space, we might expect control points to cluster together in areas with similar geographic characteristics.

9. The use of Mahalanobis metric matching was one of the earlier methods developed for matched-sample research (see Rubin 1980).

Mahalanobis distance is given by:

$$D^2(\boldsymbol{y}) = (\boldsymbol{y} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{y} - \boldsymbol{\mu}), \tag{4.2}$$

which is the multivariate distance of point y's features from the mean of all treated points' features ("treated" in this context indicates that the cell contains at least one control point). Anther way to conceptualize this distance is as the deviation of point y from the hypothesized "optimum" environment for the state to exist (i.e., its "ecological niche"). I estimate this distance using the full set of variables in Table 4.3. Because $D^2(\boldsymbol{y})$ approximates a $\chi^2_{(\boldsymbol{y})}$ distribution under multivariate normal assumptions, we can rescale $D^2(\boldsymbol{y})$ to the unit interval. These rescaled distances may be interpreted as a posterior probability resulting from a logistic regression or a Bayes discriminant function (Dunn and Duncan 2000; Rotenberry et al. 2006). Thus, a $\chi^2_{(\boldsymbol{y})}$ value of 0.75 can be interpreted as $\frac{3}{4}$ chance that a given cell belongs to the "treated" group, or that the given cell is 75% similar to the average "treated" cell.

The map in Figure 4.4 shows the rescaled Mahalanobis distance for each 100 square kilometer cell in the dataset. Note that these distances are estimated at the continent-scale, rather than on a country-by-country basis. Comparing Figure 4.4 to Figure 4.2 above, we see that there are quite a few areas of the continent that are ostensibly conducive to colonization by the state—at least based on observed their observed geographical characteristics—but that contain few, if any, control points. We see relatively high similarity scores throughout much of South Africa, along with portions of Namibia, Angola, and DR Congo. Surprisingly, there are also significant portions of the Sahel (northern Mali in particular) with high $\chi^2_{(\boldsymbol{y})}$ scores, despite the fact that there is almost no measurable state presence in these regions. It is not possible to decompose the Mahalanobis distances to get a sense of which features are driving these unexpected scores, but it seems likely that high-similarity regions in the Sahel are due in part to the presence of excluded ethnic groups and trans-border ethnic kin, while some of the other high similarity regions in the south of the continent may be driven by population density and economic activity.

The inverse pattern is also apparent in Figure 4.4—there are a number of regions with low similarity scores that do, in fact, host state assets. Figure 4.5 shows this a bit more clearly; this figure highlights Kenya and Senegal, and overlays grid cells containing observed control points in red. While these pattern raise interesting questions about why certain high-similarity regions contain so few control points, and why certain low-similarity regions contain so many, the core take away is that an ecological niche does seem to exist, and that this niche is captured by the features identified by the theory laid out in Section 4.2.

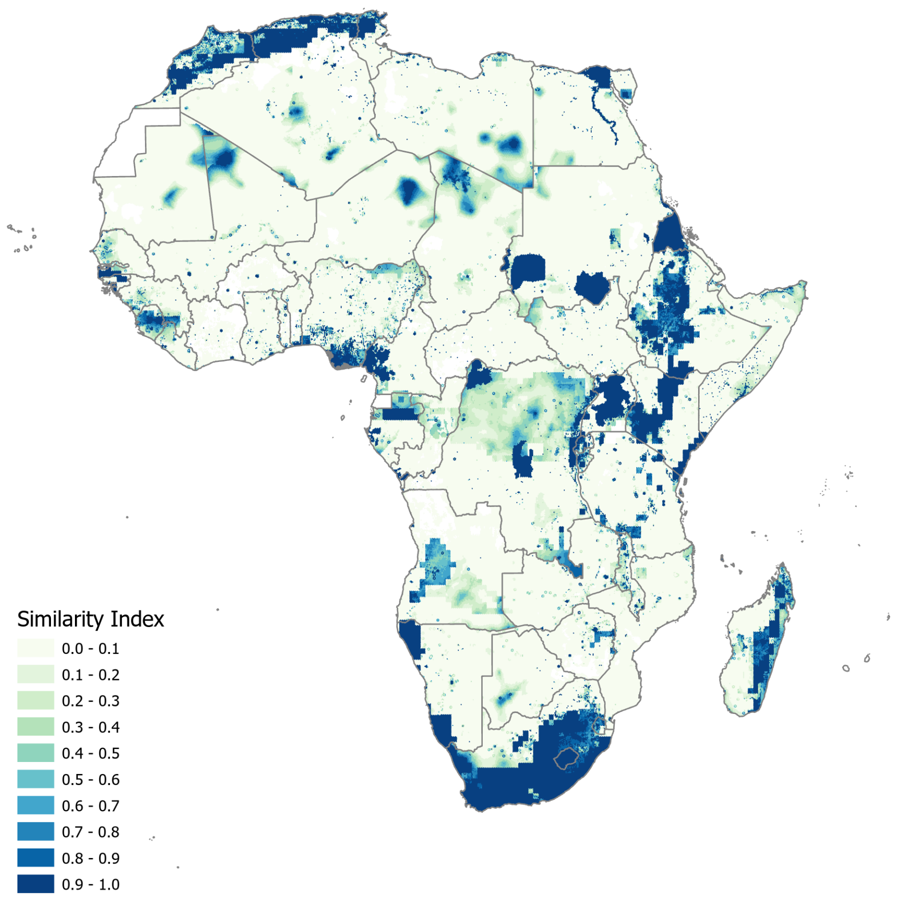
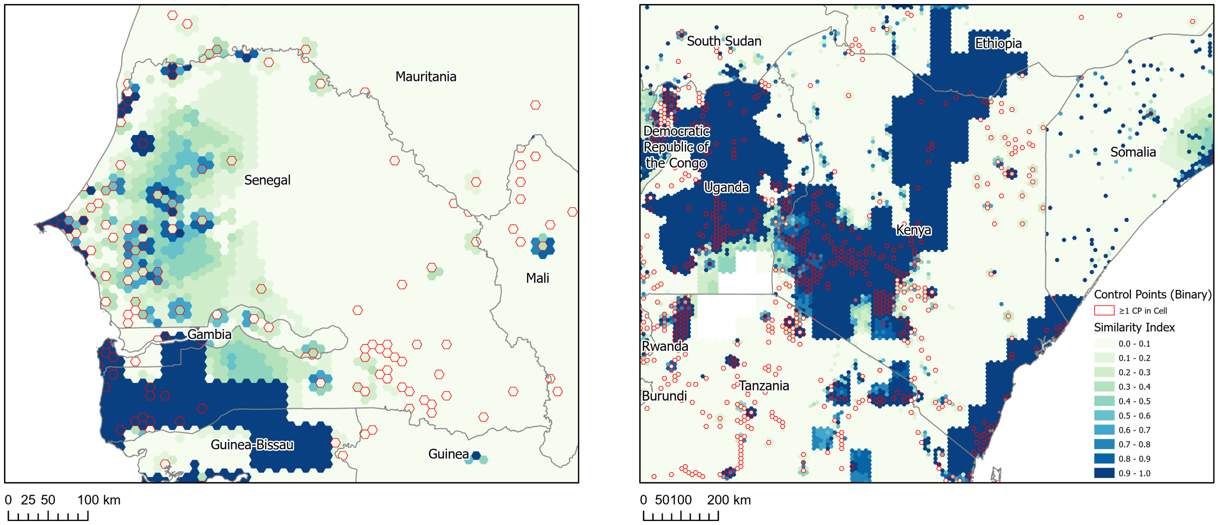


Figure 4.4. Similarity index based on a χ^2 transformation of Mahalanobis D^2.

4.4 Conventional Models

As an preliminary exercise, I begin with a series of conventional regression models to assess the predictive power of the feature variables described in Section 4.3 above. These models focus on two primary outcome variables—a binary occurrence variable and a frequency variable. I include the full set of features, including those measured at the country (rather than subnational) level. Results



(a) Detail of Senegal

(b) Detail of Kenya

Figure 4.5. Senegal and Kenya details of similarity index from Figure 4.4 with cells containing control points (e.g., "treated" cells) shown in red. Note that each map is drawn to scale; the hexagonal cells in both panels are of equal size (100km^2).

are given in Appendix Table 4.4; columns (1) and (2) provide the results of a linear probability model (OLS) and a logistic model estimated for the binary outcome variable, respectively. The logit performs fairly well. Using 0.5 as a cutoff, the logistic model correctly classifies 98.08% of observations (see Table 4.2). Because the groups are unbalanced (control points (CP) are quite rare across the continent, thus 97.74% of observations are CP = 0), and logistic models tend to do a better job at predicting members of the larger group, 99.77% of CP = 0 observations are correctly classified compared to only 24.61% of CP = 1. The false negative rate for true CP = 1 observations is 75.39%.

Table 4.2. Logit Classification Matrix

	CP=1	CP=0	Total
Pr(CP=1)≥0.5	1521	604	2125
Pr(CP=1)<0.5	4660	267545	272205
Total	6181	268149	274330
Correctly Classified	24.61% (Pr(CP=1)≥0.5 | CP=1)	99.77% (Pr(CP=1)<0.5 | CP=0)	98.08% (Overall)

Columns (3) through (5) of Table 4.4 model the frequency of control points, beginning with an OLS estimation, followed by a negative binomial and a Poisson. Due to the high concentration of zero count observations, and the the high variance in relation to the mean of the outcome variable (frequency of control points), the negative binomial is the more appropriate specification. I include the Poisson as both a robustness check and to facilitate a direct comparison with the geographically weighted Poisson I estimate in Section 4.6.

Both sets of models return similar results. We find that there is some evidence to support the hypotheses laid out in Section 4.2, particularly Hypothesis 1. Consistent with the findings in Section 4.3.2, control points in adjoining cells is a strong predictor of both outcome variables. For the binary outcome variable, we find that each additional control point in an adjoining cell is associated with a roughly 5% increase in the odds of finding a control point in the observed cell. Similarly, with the Poisson model, we see that each additional control point in an adjoining cell is associated with a 0.007 increase in the log count (a 0.7% increase in the raw count) of control points in the observed cell. Both of these parameters are statistically significant at the 99% confidence level. These results not only provide support for Hypothesis 1, they also suggest that the spatial methods described in Sections 4.5 and 4.6, which are designed to account for the spatial relationships between variables, may be more appropriate estimators for the geographic distribution of state-affiliated locations.

Evidence for Hypothesis 2 is quite compelling. The occurrence and frequency of control points are both strongly correlated with total agricultural production values and nighttime light emissions, suggesting that state assets are generally situated in more prosperous regions with high levels of economic activity. We also see a very strong relationship with the market access variables—distance to cities, airports and maritime ports. However, the distance measures to mineral and energy facilities returned mixed results. With the exception of the distance to major mineral deposits, all of these variables are either not significant or signed inconsistently with the hypothesis. The presence and frequency of control points increases the farther you get from mineral refineries, ore processing plants, and oil and gas pipelines. This may be due, in part, to zoning ordinances, which prevent these types of facilities from being constructed near population centers. Finally, I do uncover a statistically significant negative relationship between poverty and state presence.

The results are less conclusive for Hypothesis 3. Distances to the national capital, major cities, and international borders are negatively correlated with both outcome variables, suggesting that control points tend to *increase* the farther you get from these strategic locations. Results do show, however, that control points are more likely to occur in areas that that experience political violence and demonstrations. The human geography variables are generally significant and signed consistently with the Hypothesis 3. The odds of observing a control point are four times higher in urban areas, and 1% increase in population density is associated with a 0.6 increase in the log count of control points in a given cell. The number of ethnic groups in a given cell is positively correlated with the occurrence and frequency of control points, but the findings on trans-border ethnic kin is mixed. Although the point estimate for the binary variable implies a positive correlation, it is not significant at conventional levels, suggesting that the presence of trans-border ethnic kin does not influence whether or not the state is present in a given cell. However, we do see a relatively large, statistically significant effect of trans-border ethnic kin on the frequency of control points' the Poisson estimates a 0.5 increase in log count for cells containing trans-border ethnic kin. I also find that control points are more frequent in temperate climates and climates generally found at higher elevations—areas of the continent that are conducive to human settlement. The odds of observing a control point increase by 800% in Mediterranean climates (coastal Morocco, Algeria, Tunisia) and 1000% in cold desert climates (South Africa, the northern Maghreb). Results for the infrastructure variables also provide mixed support for Hypothesis 3. Control points are more frequent in areas proximate to electric power facilities, including power generation facilities and high voltage power lines. The one surprise is the negative correlation with road density. This runs counter to all expectations, and seems to be a robust finding, as it is replicated in the spatial models below.

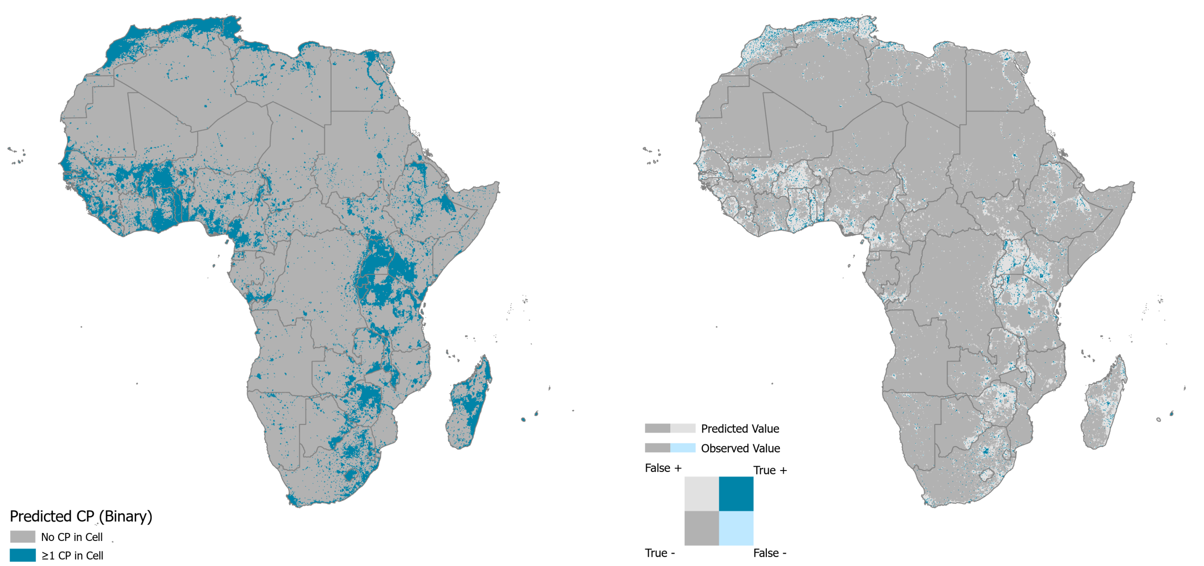
4.5 Geographical Random Forest Models

In this section, I move beyond the naïve correlates of state presence and attempt to estimate how control points are spatially distributed across the continent. I employ a machine learning technique called Ecological Niche Modeling (ENM), which is used to estimate the geographic range of a given phenomenon—typically the distribution of a particular species in ecological applications (i.e., "species distribution" modeling), or human population estimates in demographic applications—based on a set of predictor variables (e.g., climate, elevation, ground cover, and distance-based measures). ENM allows researchers to ascertain the observable conditions conducive to a particular phenomenon (e.g., which environmental conditions, or "ecological niche," does a species or phenomenon favor), and to extrapolate a geographic range for that species or phenomenon (e.g., where else does this niche exist?). In this particular context, I use ENM to model the occurrence and frequency of control points across territory within African states. The primary goal of this section is to identify a set of features that are the strongest predictors of state control in a given location. Once I have identified a set of influential predictors, I estimate a series of geographically weighted regression (GWR) models to asses where these variables are successful in predicting state control, and where they are not. This is a novel approach, in that it combines a two geospatial methods that are seldom used together, and that are both exceedingly rare in political science research.

I use a random forest (RF) technique to model the spatial distribution of control points. RF is not the only method of modeling spatial distributions, though it is preferred in many cases, as it reduces the possibility of overfitting the model to the training data, provides the best discrimination between presence and absence, and marginally outperforms other modeling techniques in terms of predictive power (Oppel et al. 2012). The RF procedure is fairly straightforward. This is a supervised machine learning technique that generates a series of decision trees to form a prediction about an outcome. Data are partitioned into a training set and a testing set. The training procedure takes random samples (with replacement) from the training set, and determines which features (i.e., predictor variables) will partition the observations in such a way that maximizes between group variation and minimizes within group variation. For classification tasks (e.g., the binary state presence variable), a majority vote is taken to determine the predicted outcome, while for regression-like tasks (e.g., the control point frequency variable), the predictions of individual trees are averaged to form a final prediction.

I estimate two separate RF models, one for each outcome variable shown in Figure 4.2. The first uses the binary control point outcome for classification. The model creates 500 decision trees, 25% of the data is excluded for validation purposes, and I initiate 25 validation runs. The mean accuracy rate is 0.859 ($\sigma = 0.002$). In terms of sensitivity, 86% of observed zeros (state absence) in the data are correctly predicted, while 89% of observed ones (state presence) are correctly predicted. The model's prediction surface is shown in Figure 4.6a, and a confusion surface showing the locations of incorrect predictions is shown in Figure 4.6b.

Based on the maps in Figure 4.6, we see that that the algorithm tends to over-predict state presence (CP = 1). This is consistent with the $\chi^2_{(\boldsymbol{y})}$ similarity index shown in Figure 4.4; estimates of cell-specific Mahalanobis distances suggest that the state is not present in locations that are suitable to colonization, and the RF algorithm appears to be replicating this pattern. Although we do not currently observe control points in these locations, these false positives represent our best guess as to where the state is *likely* to expand in the future. In practical terms, these false positives may indicate that the model is missing some sort of constraining feature—an observable characteristic that limits state expansion into areas where we we might otherwise expect to see control points. The theory of spatial optimization outlined in Section 4.2 suggests that a cost constraint may be the limiting factor, though this pattern may also indicate a saturation effect—control points are absent in these locations because the state does not need additional facilities to exercise control in these neighborhoods.



(a) Random Forest Prediction

(b) Classification Matrix

Figure 4.6. Results of a random forest classification model to predict the occurrence of control points across the continent. Panel 4.6a shows cells predicted to contain at least one control point in blue. Panel 4.6b classifies each cell by the accuracy of the prediction; dark gray and dark blue indicate correct predictions, while lighter shades indicate incorrect predictions.

Figure 4.7 provides the importance scores of the various features according to a Gini-based metric (a full summary is available in Table 4.5 in the Appendix). Essentially, this is a measure of how often a particular feature is responsible for a split in a decision tree (and the impact of that split) divided by the total number of trees. The range of values comes from the 25 validation runs.

The scale is irrelevant; only the relative values are informative. Note that these are not traditional coefficients—they cannot be used to determine if a feature decreases or increases the likelihood of a location being classified in a certain way. This is the primary reason that I compliment this analysis with the geographically weighted regression models in Section 4.6 below.

Figure 4.7 indicates that the most important predictor of occurrence is population density, which is consistent with Hypothesis 3. Herbst (2014) argues that the state is interested in controlling people rather than territory, so it make sense that control points tend to occur in densely populated areas. Recall, though, that the control points deliberately exclude government locations that are explicitly positioned based on population density, such as schools and social services, which tend to be assigned to a population-based catchment area. Economic variables round out rest of the top performers. Nighttime light emissions, which are a measure of overall wealth and economic activity, and total agricultural production value score quite highly, as do variables that measure access to markets, such as distance to major cities and airports. Of interest, though, is the high degree of importance assigned to demonstration events. Although Hypothesis 3 posits that the state will be more scarce in areas prone to political violence, it easy to see why demonstrations would occur in areas "close" to the state. In many instances, the target audience of a demonstration is the central government, so these events are likely to occur in locations visible to the center. The least important variables are some of the political and demographic variables: a country's former metropole and Herbst's classification of African national design are among the poorest performing features, as are variables that code ethnic demography, such as the number and composition of ethnic groups and the presence of trans-border ethnic kin.

Figure 4.6 shows the results and standardized residuals of a random forest regression model predicting the frequency of control points at specific locations across the continent. Because details are difficult to make out in Figure 4.8, Figure 4.9 zooms in on Nigeria. This model uses the same parameters as the classification model. In general, the RF algorithm does not perform as well in predicting frequencies as it does in predicting the occurrence outcome. The average R^2 of the 25 validation runs is 0.52 ($\sigma = 0.069$). This is quite a bit lower than the Poisson model in Section 4.4 (pseudo $R^2 = 0.60$), though it is worth noting that the RF model is not only trying to predict arbitrary frequencies, but also *where* these frequencies are situated in space. Again, though, the RF model

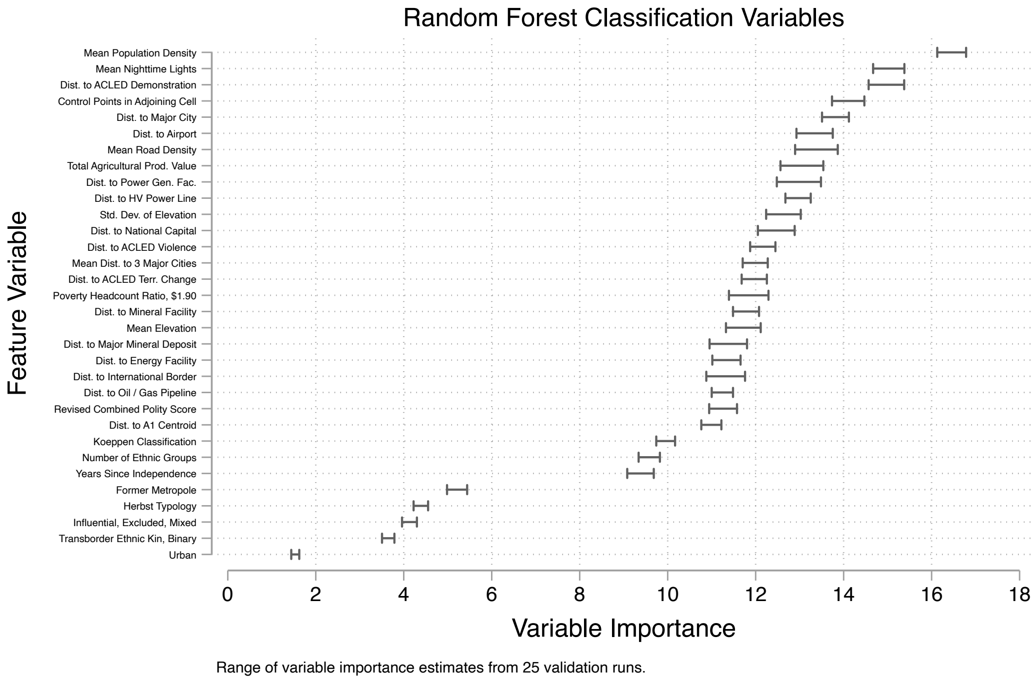
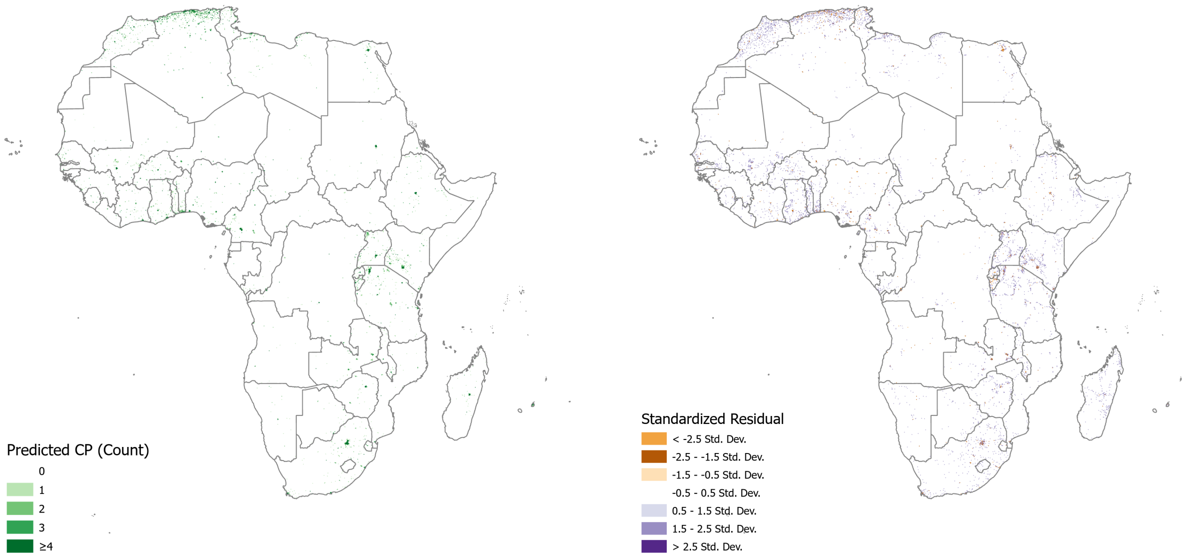


Figure 4.7. Gini-based variable importance measures for random a forest classification model. See Table 4.5 in the Appendix for detailed statistics.

tends to over-predict frequencies in specific locations (see Figure 4.8b), suggesting that the state is not only underperforming at the extensive margin, but at the intensive margin as well. Figure 4.10 provides the variable importance scores for the RF count model. The most important features are roughly the same as the occurrence model, though some of the violence measures have increased importance, particularly the distance to violent events and demonstrations. The economic variables also tend not to perform as well in this model, though the distance variables to major cities do rank highly.

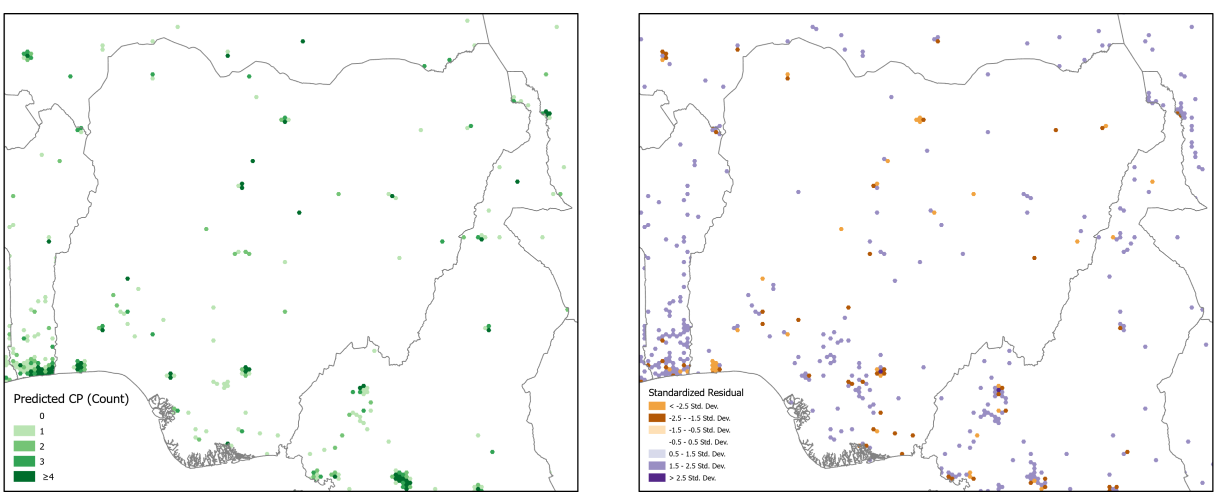
Overall, a relatively small set of features allow us to predict state presence with fairly high levels of accuracy. These features constitute the ecological niche of the African state—a set of environmental conditions favorable to state presence. Based on the variable importance scores in Figures 4.7 and 4.10, the niche is defined by four broad attributes: adjacency to other state assets, densely populated human settlements, access to markets, and high levels of economic and agricultural production. These models suggest, however, that there are different attributes that drive the occurrence and frequency of control points. Coverage tends to coincide with the presence



(a) Random Forest Prediction

(b) Standardized Residuals

Figure 4.8. Results of a random forest regression model to predict the frequency of control points across the continent. Panel 4.8a provides the predicted count of control points per cell. Panel 4.8b shows standardized residuals.



(a) Random Forest Prediction

(b) Standardized Residuals

Figure 4.9. Nigeria detail of results of a (continent-scale) random forest regression model to predict the frequency of control points across the continent. Panel 4.9b shows standardized residuals.

of human and economic capital in a region, while the state will concentrate its resources in areas where coverage already exists, and which have ready access to large markets and formal political institutions, such as the national capital and other major cities.

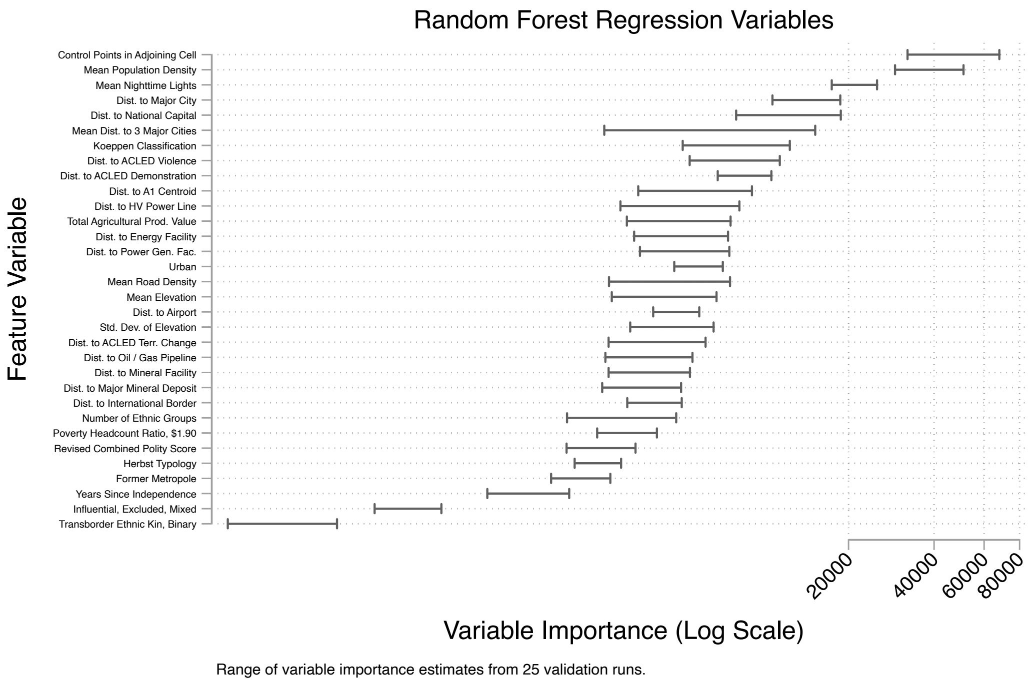
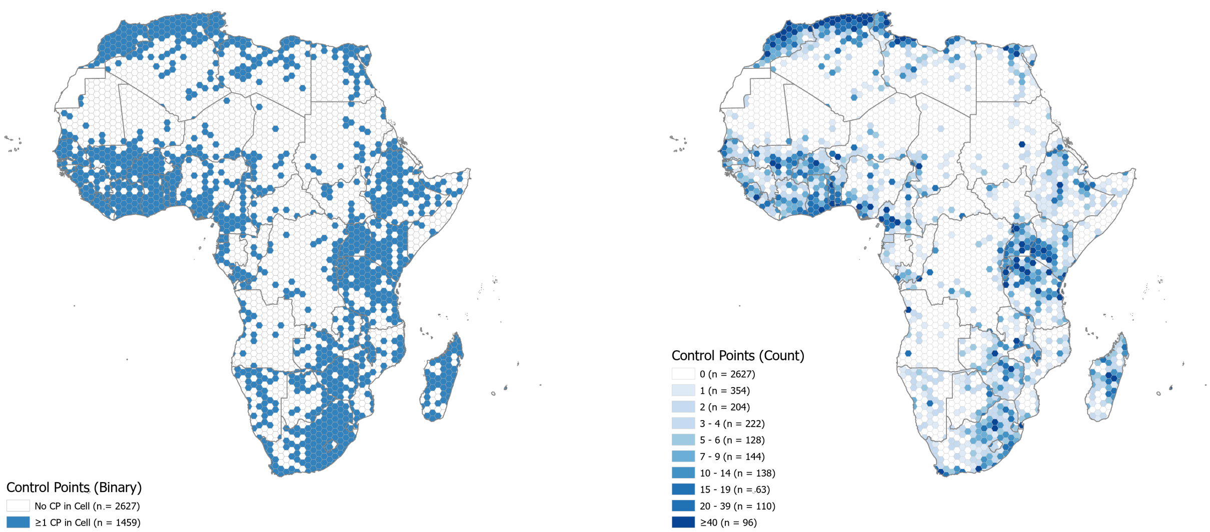


Figure 4.10. Gini-based variable importance measures for a random forest regression model. See Table 4.6 in the Appendix for detailed statistics.

4.6 Geographically Weighted Regression Models

The random forest models presented in Section 4.5 help us to isolate a set of features that characterize the ecological niche of the African state, and to make highly accurate predictions about the spatial distribution of state assets in Africa. However, the RF models do not allow us to test explicit hypotheses about the relationship between our explanatory features and the location-specific occurrence and frequency of control points. In this section, I turn to geographically weighted regression to test the hypotheses laid out in Section 4.2 by estimating the relative influence of various features in a given locale.

The models in this section differ from those in Sections 4.4 and 4.5, in that they allow for spatial non-stationarity—instances in which the relationship between the explanatory and dependent variables changes by geographic location. Global models, such as the conventional logit and the random forest classifier, assume that the data-generating process is stationary over space, such that a single coefficient can capture the relationship between each explanatory variable and the dependent variable. This is seldom the case with spatial phenomena. Take energy resources in the DRC, for example. Current oil and gas production is located almost exclusively in the Congo River Delta and in offshore fields along the Atlantic coast. We might plausibly hypothesize, then, that in coastal areas of the country, proximity to proven energy reserves is a viable predictor of state presence, as the Nigerian government has incentive to protect this valuable infrastructure, and ensure that energy firms are complying with environmental and tax regulations. In interior regions of the country, where energy production is non-existent, this relationship is unlikely to obtain. State presence is more likely to be determined by other factors, such as population density or durable infrastructure. Geographically weighted models relax the stationarity assumption, and allow relationships between the explanatory and dependent variables to vary by locality.



(a) Control Points (Binary) **(b)** Control Points (Count)

Figure 4.11. Spatial distribution of the two primary outcome variables aggregated to the 10,000km^2 hexagonal grid cell.

Geographically weighted regression models estimate a set of location-specific of coefficients for each observation in the dataset using the features from that observation's immediate neighbors. This enables us to map local coefficients to evaluate the spatial variability of coefficient values and identify potential clustering. These models suffer from one important limitation, however: spatial autocorrelation among both explanatory and dependent variables reduces neighborhood variation, thus posing collinearity issues. This essentially precludes the use of categorical variables and variables measured at the country or district level, such as Polity scores and climate classification. I take two approaches to mitigate this issue. The first is to aggregate the outcome variables and explanatory features to a coarser (i.e., larger) unit of analysis—the 10,000 square kilometer hexagonal grid cell, shown in Figure 4.11. Aggregation increases the amount of spatial variation, which helps to minimize collinearity issues. Second, I estimate country-specific regressions for two case studies—Kenya and Senegal—using the the density measure from Figure 4.1b (which has a high degree of variability) as an outcome.

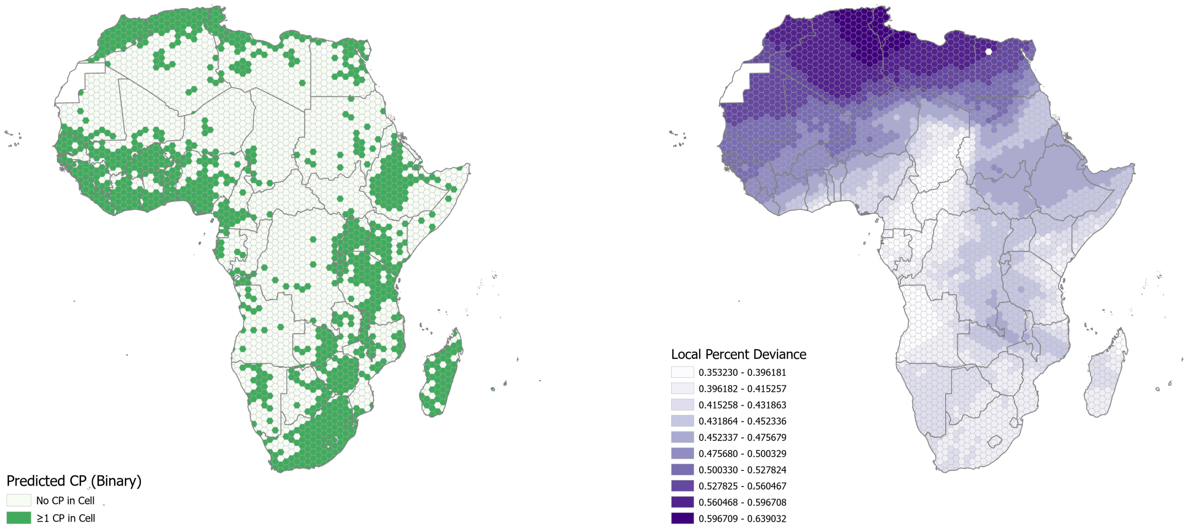
4.6.1 Geographically Weighted Logistic Regression – Continent-Scale

To model the binary occurrence outcome, I estimate:

$$\ln\left(\frac{p_i}{1-p_i}\right) = \sum_{k=1}^{M} \beta_k(u_i, v_i) x_{ik} + \epsilon_i, \tag{4.3}$$

where $\ln\left(\frac{p_i}{1-p_i}\right)$ is the predicted odds for observation i at coordinates (u_i, v_i), and $\beta_k(u_i, v_i)$ is the local coefficient estimate for the kth explanatory, x, at location i. Figure 4.12 shows the resulting binary prediction surface and local goodness of fit statistics. In general, the model performs quite well—the included features accurately predict 81.65% of cells in which the state is present (CP = 1), and 86.07% of cells in which the state is absent (CP = 0). This is a substantial improvement over the conventional logit (see Table 4.2), which struggled to identify cells containing control points. Figure 4.12b indicates that the model has the most predictive power in the Maghreb, in coastal West Africa, and through the Great Lakes.

Figure 4.13 provides select coefficient estimates; the full set of estimates are available in the Appendix. Several core findings are worth highlighting: Unlike the preceding analyses, this model does not provide support for Hypothesis 1. In Figure 4.13a, we see that the occurrence of control points is positively and significantly correlated with control points in adjoining cells in only the Atlantic coastal regions of the continent, stretching from Guinea to Côte d'Ivoire. These results



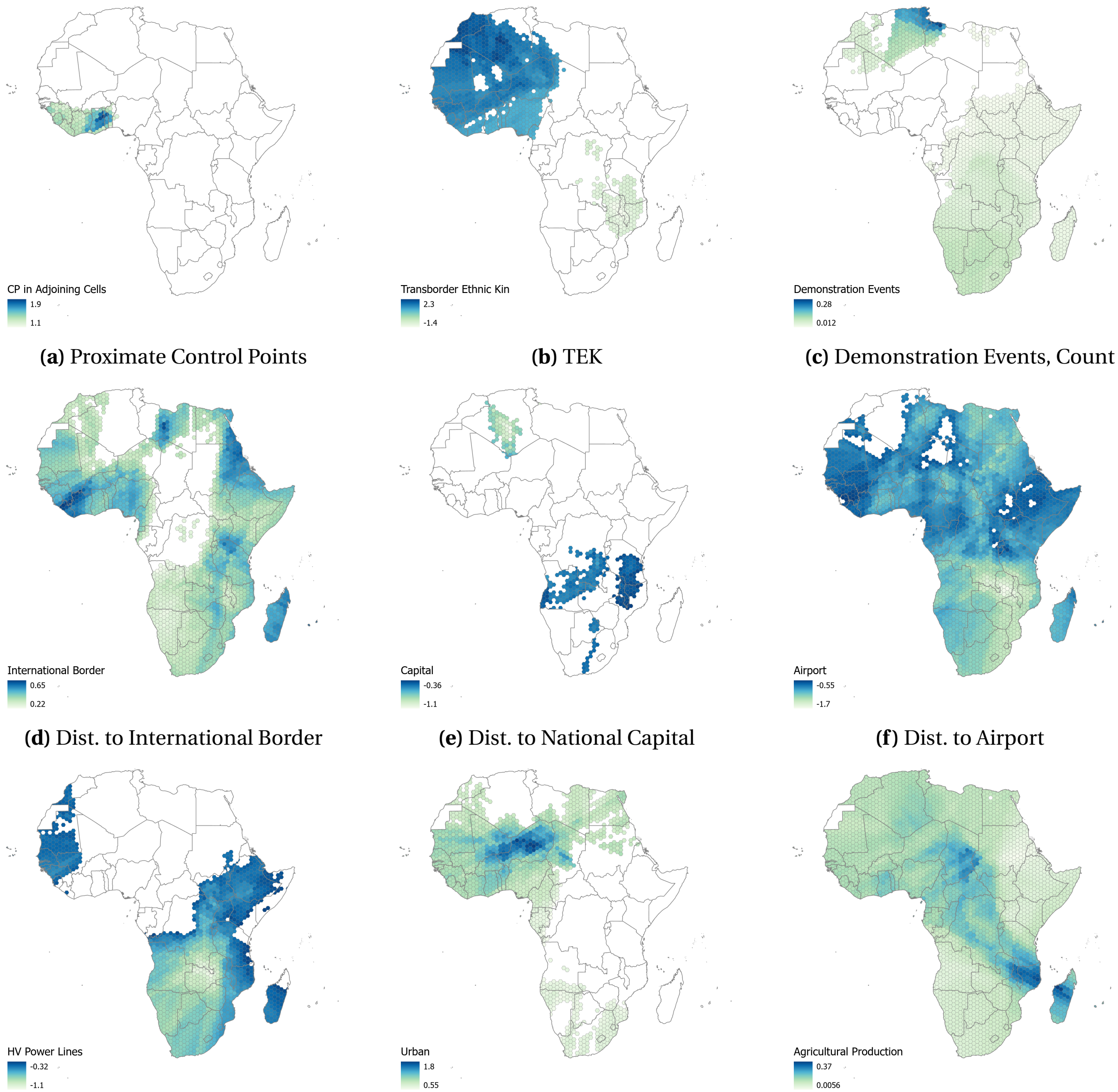
(a) GWR Binary Prediction

(b) Goodness of Fit

Figure 4.12. Results of a geographically weighted logistic regression, estimated at the continent-scale using the 10,000km^2 hexagonal grid cell as the unit of analysis. Panel 4.12a shows predicted values, while Panel 4.12b plots the local percent deviance, a goodness of fit statistic similar to R^2.

suggest that other factors, aside from proximity to existing state assets, is driving location decisions at the local-level.

Results do provide support for Hypothesis 2. Throughout the continent, coverage is highly correlated with features such as agricultural production and distance to airports, which are proxies for economic activity and access to markets. Figure 4.2110 in the Appendix underscores the importance of market access; we see that in coastal West Africa, and throughout the Congo River Basin, state presence is strongly correlated with proximity to maritime and riverine ports. Evidence for Hypothesis 3 is mixed. On one hand, Figures 4.13g and 4.13h underscore the importance of proximity to strategic locations such as high voltage power lines (c.f., Ukraine) and urban agglomerations; coefficients are significant and signed consistently with Hypothesis 3 in large areas of the continent. However, coefficients for the variables measuring proximity to the national capital and international borders run counter to expectations. Capital distance appears to be a poor predictor of presence throughout much of the continent. We also see that, throughout the continent, the coefficients on border distance are positive and significant, indicating that the occurrence of control points increases with distance to the border. While this finding is inconsistent with my hypotheses, it does lend support to an argument made by Herbst (2014) that, because borders serve as effective buffer zones Africa,



(a) Proximate Control Points **(b)** TEK **(c)** Demonstration Events, Count

(d) Dist. to International Border **(e)** Dist. to National Capital **(f)** Dist. to Airport

(g) Dist. to HV Power Trans. Line **(h)** Percent Urban Territory **(i)** Total Agricultural Production

Figure 4.13. Select parameter estimates for the geographically weighted logistic regression model. Only coefficients with a t value significant at the 95% confidence level are shown in each map. The full set of coefficient estimates are available in the Appendix.

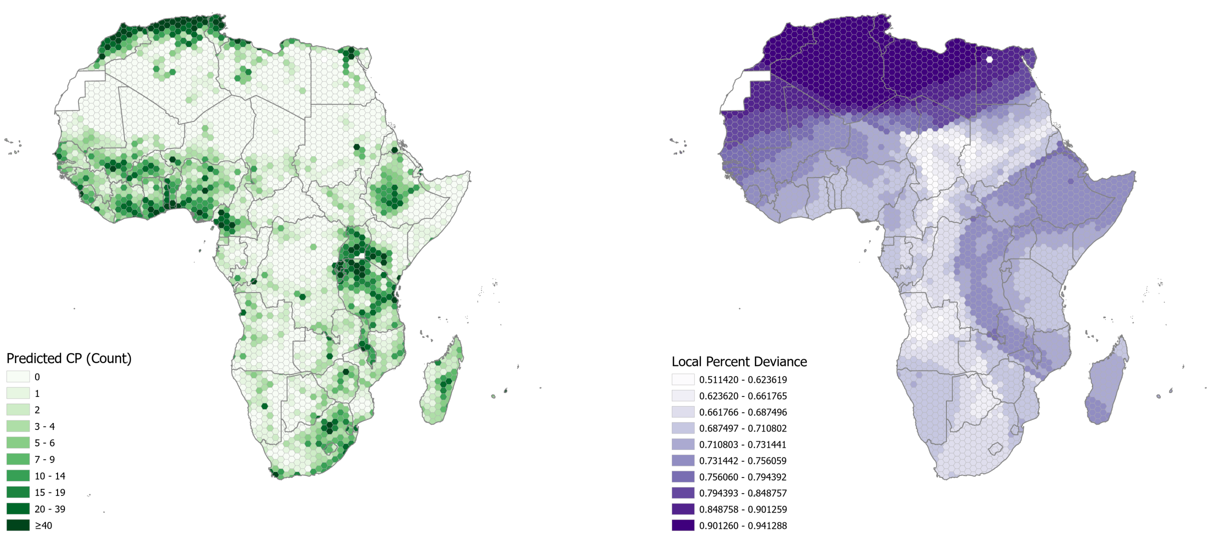
there is little need to police these areas.[10]

4.6.2 Geographically Weighted Poisson Regression – Continent-Scale

I model the frequency outcome using the following Poisson specification:

$$y_i \sim \text{Poisson}\left[\exp\left(\alpha(u_i, v_i) + \sum_{k=1}^{M} \beta_k(u_i, v_i)x_{ik}\right)\right]. \tag{4.4}$$

Figure 4.14 shows the predicted counts generated by the model, along with a local goodness of fit statistic. This spatial model improves substantially on the conventional Poisson in Table 4.4; the local model explains roughly 80% of the variance in the frequency outcome.



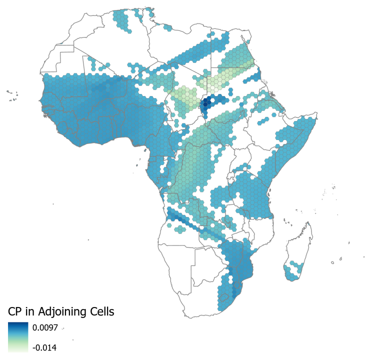
(a) GWR Poisson Prediction

(b) Goodness of Fit

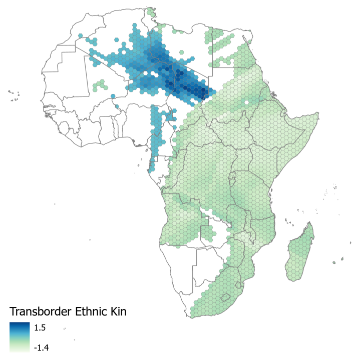
Figure 4.14. Results of a geographically weighted Poisson regression, estimated at the continent-scale using the 10,000km^2 hexagonal grid cell as the unit of analysis. Panel 4.12a shows predicted values, while Panel 4.12b plots the local percent deviance.

Selected coefficient estimates are shown in Figure 4.15; the Appendix contains the full set of estimates. Results provide support for Hypothesis 1 throughout much of the continent. Coefficient estimates suggest a roughly 1 to 1 mapping of the number of control points in one cell to the number of control points in adjoining cells throughout West Africa, Central Africa, and along the Swahili Coast. Results are broadly consistent with Hypotheses 2 and 3 as well. Agricultural production values

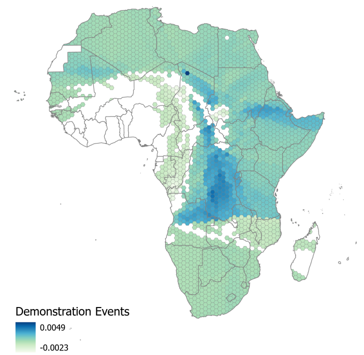
10. There may also be a mechanical effect at play—there is greater area in the interior of a country to locate a state facility than along its border



(a) Proximate Control Points

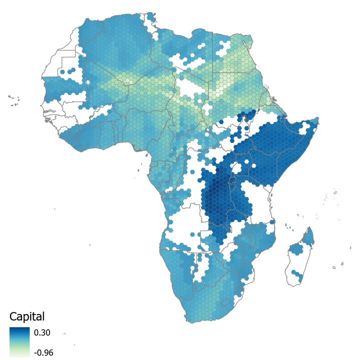


(b) TEK

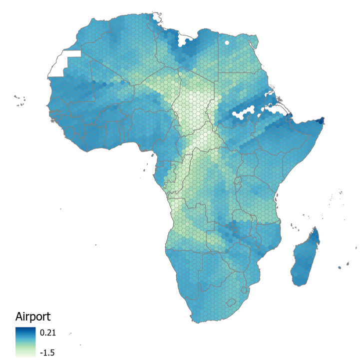


(c) Demonstration Events, Count

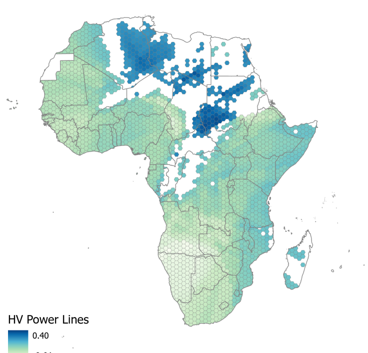
(d) Dist. to International Border



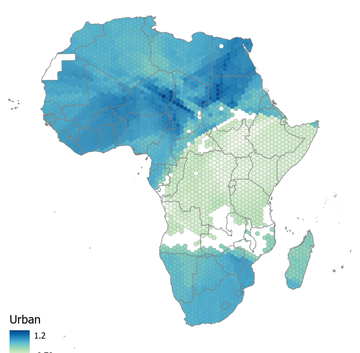
(e) Dist. to National Capital



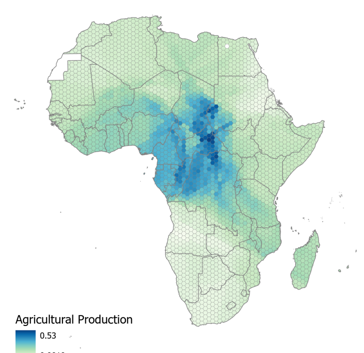
(f) Dist. to Airport



(g) Dist. to HV Power Line



(h) Percent Urban Territory



(i) Total Agricultural Production

Figure 4.15. Select parameter estimates for geographically weighted Poisson regression model. Only coefficients with a t value significant at the 95% confidence level are shown in each map. The full set of coefficient estimates are available in the Appendix.

and economic activity, as measured by nighttime light emissions, are associated with a substantial increase (between 18% and 89%) in the frequency of control points in a particular cell; this is true throughout almost the entirety of the continent. Proximity to strategic locations such as the national capital, airports, and high voltage power lines tends to increase the frequency of control points in a particular region. However, the Poisson replicates the negative correlation between border distance and the frequency of control points.

4.6.3 Geographically Weighted Regression – Kenya & Senegal

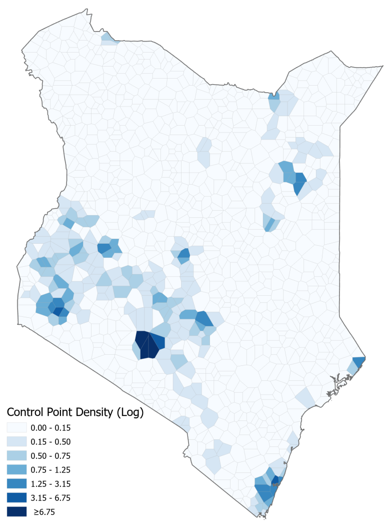
Because spatial dynamics tend to differ at different geographic scales, I drill down into two country case-studies—Kenya and Senegal—to evaluate how well various features perform at the hyper-local level. Estimating geographically weighted models using these smaller units of analysis exacerbates the problems of spatial autocorrelation and local multicollinearity. To minimize these issues, I employ two strategies. The first is to use control point density as an outcome variable (see Figure 4.1b). Because density is continuous, this measure has greater spatial variation than either the binary or count variables used for the continent-scale analyses. Second, I randomly select 20% of the 100 square kilometer hexagonal grid cells, and estimate the model on a sample of observations.

My primary specification is as follows:

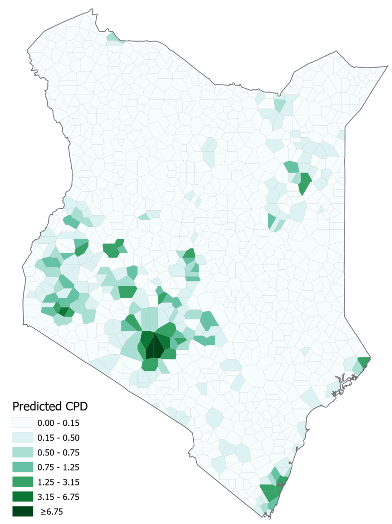
$$y_i = \alpha(u_i, v_i) + \sum_{k=1}^{M} \beta_k(u_i, v_i) x_{ik} + \epsilon_i, \tag{4.5}$$

where y is the outcome of interest and $\alpha(u_i, v_i)$ and $\beta_k(u_i, v_i)$ are the set of local intercepts and coefficients at location i, defined by latitude and and longitude coordinates (u_i, v_i). Predicted densities and model diagnostics are mapped in Figures 4.16 and 4.17. The full set of coefficient estimates are available in the Appendix. For ease of interpretation and general aesthetic purposes, I generate a Voronoi tessellation of the sampled observations, and present the results within these polygons.

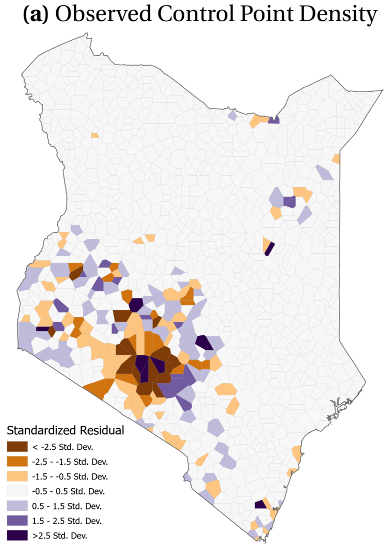
Results of these models are quite compelling; they highlight how the relationship between various features and the density of control points varies across space, within a given country. Figures 4.3324 and 4.3326, for example, indicate that a country's human geography—including urban territory and population density—is only a significant predictor of control point density in the southeast corner of Kenya, which is the region in which the capital of Nairobi is situated. The coefficients in Figure 4.318, however, indicate that the influence of the capital wanes the farther you travel from Nairobi. In other regions of the country, other features, such as proximate control points and economic activity, are better indicators of state presence. Results from Senegal show the decaying influence of various features quite clearly. For example, control point density is highly correlated with proximity to maritime ports along the coast, but this correlation decreases the further you travel inland, until it disappears completely about 450 kilometers from the ocean.



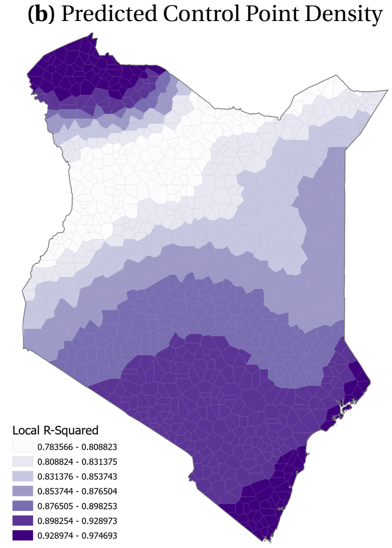
(a) Observed Control Point Density



(b) Predicted Control Point Density

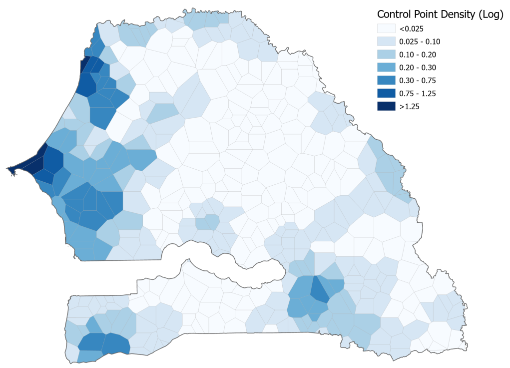


(c) Standardized Residuals

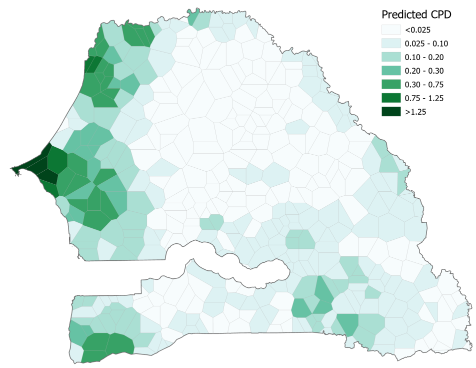


(d) Local R^2

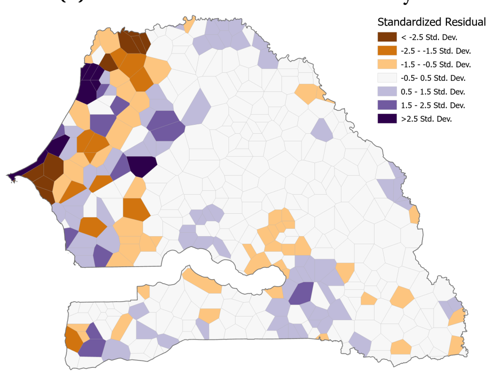
Figure 4.16. Results of a GWR model estimated for Kenya.



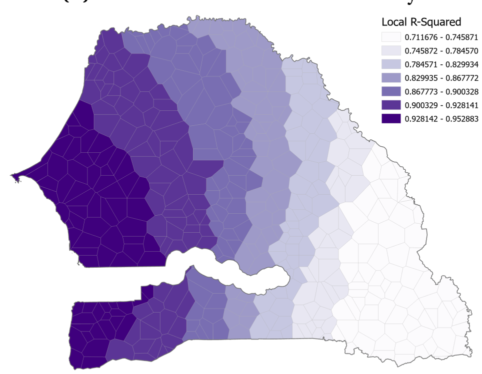
(a) Observed Control Point Density



(b) Predicted Control Point Density



(c) Standardized Residuals



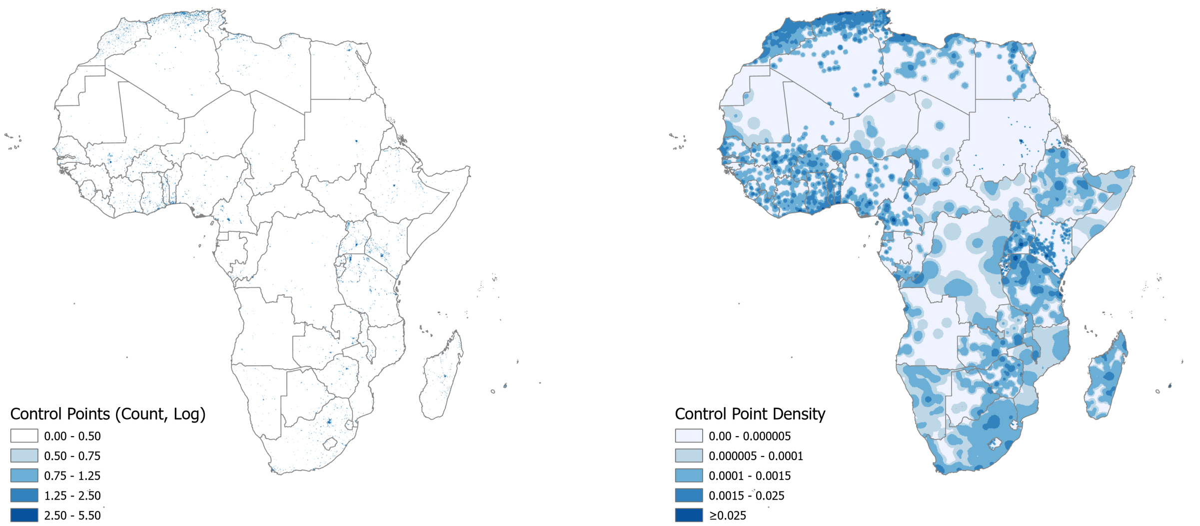
(d) Local R^2

Figure 4.17. Results of a GWR model estimated for Senegal.

4.7 Discussion

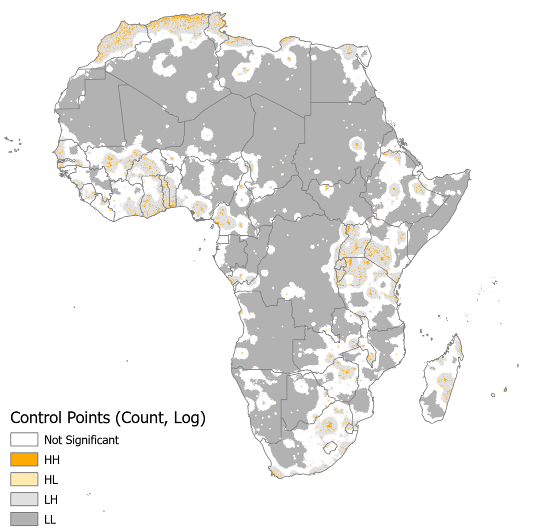
This chapter provides a novel mapping of territorial control in Africa, and develops a theory to explain the spatial distribution of governed and ungoverned space across the continent. Using a variety of geospatial methods, I find that state assets tend to concentrate in certain regions—particularly those proximate to other state infrastructure, markets, and population centers, as well as areas that host high levels of economic activity. These features allow us to predict the spatial situation of state assets with a high degree of accuracy at the macro level. My analyses show, however, that the factors that influence state presence vary across space, and at different geographic scales. At the national and local levels, the idiosyncratic factors that drive a state's decision on to where to allocate resources is place-specific and variable by region.

4.8 Appendix

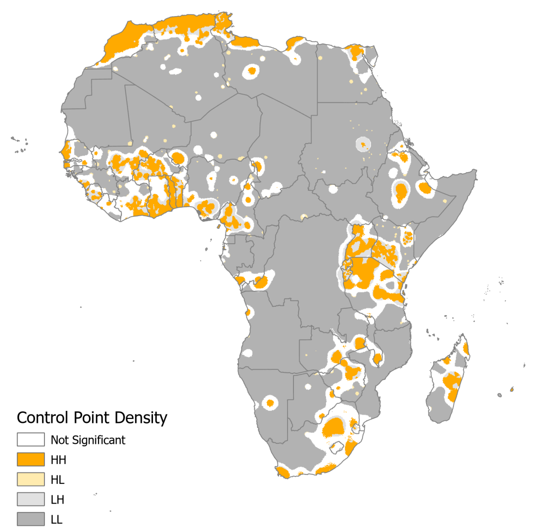


(a) Control Points (Count, Log) **(b)** Control Point Density

Figure 4.18. Spatial distribution of the supplementary outcome variables. Note that the unit of analysis is the 100km^2 hexagonal grid cell.



(a) Control Points (Count, Log)



(b) Control Point Density

Figure 4.19. Local cluster maps of the supplementary outcome variables. Dark orange indicates statistically significant clusters of high values (e.g., presence and frequency of control points), while dark gray indicates statistically significant clusters of low values (e.g., absence of control points). Light orange and light gray indicate statistically significant outlier regions (e.g., transition zones).

Table 4.3. Descriptive Statistics, 100km^2 Hexagonal Cells

Variable	Data Source	n	μ	σ	Min	25%	Median	75%	Max
Control Points (Binary)		311507	0.02	0.14	0.0	0.0	0.0	0.0	1.00
Control Points (Count)		311507	0.06	1.11	0.0	0.0	0.0	0.0	198.00
Control Points (Count, Log)		311507	0.02	0.18	0.0	0.0	0.0	0.0	5.29
Control Point Density		307105	0.00	0.00	0.00	0.00	0.00	0.00	0.52
Control Points in Adjoining Cells, Count	Hand-Coded	311507	0.35	3.45	0.00	0.00	0.00	0.00	325.00
Control Points in Adjoining Cells, Binary	Hand-Coded	311507	0.08	0.27	0.00	0.00	0.00	0.00	1.00
Control Points in Adjoining Cells, Count (Log)	Hand-Coded	311507	0.10	0.40	0.00	0.00	0.00	0.00	5.79
Years Since Independence	Hand-Coded	299838	64.11	13.58	42.00	60.00	62.00	64.00	112.00
Country Population (Millions)	Various	311437	38.17	39.28	0.09	12.08	26.57	47.62	190.60
Island Country, Binary	Hand-Coded	311437	0.00	0.03	0.00	0.00	0.00	0.00	1.00
Revised Combined Polity Score	CSP	308475	1.83	4.27	-9.00	-2.00	2.00	6.00	10.00
Dist. to International Border	GADM	311507	130.95	117.23	0.00	38.08	98.72	193.86	656.01
Dist. to A1 Centroid	GADM	311507	147.41	106.06	0.03	70.77	121.76	194.74	1113.33
Dist. to National Capital	Hand-Coded	311437	620.01	409.15	0.67	306.42	510.87	872.76	1969.41
Dist. to Airport	WFP	311507	98.37	85.56	0.23	43.76	75.15	122.94	2821.80
Dist. to H2O Port	WFP	311507	387.08	285.64	0.18	154.30	324.80	561.99	2737.23
Dist. to Major City	Various	311507	171.81	128.92	0.10	78.73	138.02	230.19	2610.98
Mean Dist. to 3 Major Cities	Various	311507	248.54	142.87	8.61	142.24	216.95	323.74	2723.17
Dist. to Major Mineral Deposit	USGS	311507	240.80	180.09	0.21	101.89	200.07	333.79	2836.62
Dist. to Power Gen. Fac.	USGS	311507	114.02	101.05	0.09	45.52	84.01	147.59	2825.26
Dist. to Non-Energy Mineral Facility	USGS	311507	247.60	211.27	0.01	86.10	186.33	352.89	1279.95
Dist. to Energy Extraction / Refinement Facility	USGS	311507	416.56	399.21	0.00	129.57	302.54	567.89	2307.72
Dist. to ACLED Violence	ACLED	311507	62.87	64.83	0.01	18.05	40.81	86.83	2817.91
Dist. to ACLED Demonstration	ACLED	311507	77.20	83.23	0.01	24.38	49.65	96.69	2591.96
Dist. to ACLED Territorial Change	ACLED	311507	261.14	278.41	0.10	75.71	170.47	342.56	3902.78
ACLED Violence Events, Count	ACLED	311507	0.48	14.67	0.00	0.00	0.00	0.00	6526.00
ACLED Demonstration Events, Count	ACLED	311507	0.32	8.77	0.00	0.00	0.00	0.00	1519.00
ACLED Territorial Change Events, Count	ACLED	311507	0.02	0.47	0.00	0.00	0.00	0.00	64.00
Number of Ethnic Groups	EPR	311507	4.37	3.96	0.00	0.00	4.00	7.00	15.00
Transborder Ethnic Kin, Binary	EPR	311507	0.65	0.48	0.00	0.00	1.00	1.00	1.00
Poverty Headcount Ratio, \$1.90	World Bank	311115	34.39	29.18	0.00	6.39	26.55	62.13	97.55
Dist. to Oil / Gas Pipeline	USGS	311507	445.48	353.47	0.00	145.57	379.97	673.70	2815.54
Dist. to HV Power Transmission Line	USGS	311507	178.19	199.51	0.00	29.83	93.14	266.32	2826.17
Total Agricultural Production Value (in Millions of USD)	SPAM	311507	0.47	2.17	0.00	0.00	0.00	0.16	154.10
Mean Elevation (Meters)	NASA-JPL SRTM	311507	623.03	446.84	-120.81	311.45	491.17	899.42	4197.83
Std. Dev. of Elevation	NASA-JPL SRTM	311507	32.61	48.24	0.00	7.78	16.24	34.60	851.43
gEcon Mean	gEcon	311175	0.69	3.00	0.00	0.03	0.11	0.37	85.04
Percent Urban Territory	GRUMP	311507	0.01	0.08	0.00	0.00	0.00	0.00	1.00
Mean Nighttime Lights	VIIRS	311507	0.11	1.69	0.00	0.00	0.00	0.00	474.62
Mean Population Density	WorldPop	311507	40.26	300.21	0.00	0.39	5.18	21.80	31617.95
Mean Road Density	OSM	311507	0.08	0.08	0.00	0.03	0.06	0.09	0.50

Note: All distance measures expressed in kilometers.

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
Control Points in Adjoining Cells, Count	0.00542***	1.046***	0.119***	0.0619***	0.00665***
	(0.00111)	(0.00552)	(0.0107)	(0.00661)	(0.000944)
Country Population	-0.00202	0.779***	-0.00249	-0.262***	-0.224***
	(0.00293)	(0.0239)	(0.00469)	(0.0350)	(0.0412)
Revised Combined Polity Score	0.000144	1.030***	0.00145	0.0144**	0.00827
	(0.000421)	(0.00534)	(0.000879)	(0.00626)	(0.00661)
Dist. to International Border	-0.000418	0.989	-0.00281	-0.0788***	0.0196
	(0.000466)	(0.0131)	(0.00167)	(0.0155)	(0.0177)
Dist. to A1 Centroid	-0.00300**	0.965*	-0.0123**	-0.0407*	-0.0588**
	(0.00133)	(0.0200)	(0.00579)	(0.0233)	(0.0284)
Dist. to National Capital	0.000839	1.063**	-0.0169*	0.112***	0.0560**
	(0.00254)	(0.0274)	(0.00984)	(0.0308)	(0.0232)
Dist. to Airport	-0.00134	0.820***	0.00562	-0.216***	-0.214***
	(0.00146)	(0.0184)	(0.00555)	(0.0266)	(0.0240)
Dist. to H20 Port	-0.00538***	0.851***	-0.00807	-0.114***	-0.128***
	(0.00192)	(0.0175)	(0.00844)	(0.0232)	(0.0284)
Dist. to Major City	-0.00419*	1.015	-0.0153	-0.00156	-0.110***
	(0.00229)	(0.0280)	(0.0166)	(0.0318)	(0.0258)
Mean Dist. to 3 Major Cities	0.00865**	1.181***	0.0284	0.0488	0.0324
	(0.00423)	(0.0636)	(0.0196)	(0.0871)	(0.0782)

Continued on next page

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
Dist. to Major Mineral Deposit	-0.000160	0.933***	-0.000136	-0.0583**	-0.0683***
	(0.00163)	(0.0188)	(0.00417)	(0.0295)	(0.0249)
Dist. to Power Gen. Fac.	0.00161	0.924***	0.00886**	-0.133***	-0.0476**
	(0.00139)	(0.0199)	(0.00351)	(0.0236)	(0.0237)
Dist. to Non-Energy Mineral Facility	0.00208*	1.086***	0.0163***	0.111***	0.0980***
	(0.00104)	(0.0214)	(0.00470)	(0.0244)	(0.0273)
Dist. to Energy Extraction / Refinement Facility	0.00250	1.086***	0.00777**	0.0769***	0.0307
	(0.00156)	(0.0179)	(0.00323)	(0.0183)	(0.0215)
Dist. to ACLED Violence	-0.00177	0.949**	-0.00527	-0.0918***	-0.0625**
	(0.00123)	(0.0217)	(0.00469)	(0.0279)	(0.0273)
Dist. to ACLED Demonstration	-0.00931***	0.633***	0.00143	-0.444***	-0.335***
	(0.00226)	(0.0153)	(0.00404)	(0.0287)	(0.0259)
Dist. to ACLED Territorial Change	0.00309**	1.163***	0.00481	0.165***	0.140***
	(0.00150)	(0.0226)	(0.00617)	(0.0266)	(0.0304)
ACLED Violence Events, Count	7.01e-05	1.000	0.00112	0.00108	0.000406***
	(4.76e-05)	(0.000918)	(0.00150)	(0.00155)	(0.000145)
ACLED Demonstration Events, Count	0.000667***	1.016***	0.0361***	0.00128	0.000914***
	(9.47e-05)	(0.00317)	(0.0100)	(0.000870)	(0.000227)
ACLED Territorial Change Events, Count	0.00394***	1.043*	-0.0339	0.0554**	-0.0226*
	(0.00135)	(0.0241)	(0.0314)	(0.0252)	(0.0137)
Number of Ethnic Groups	0.000545	1.059***	0.00345**	0.0697***	0.0676***

Continued on next page

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
	(0.000513)	(0.00854)	(0.00143)	(0.0104)	(0.0139)
Transnational Ethnic Kin, Binary	-0.00957	1.197	0.00736	0.297**	0.542***
	(0.00893)	(0.134)	(0.0131)	(0.126)	(0.132)
Poverty Headcount Ratio, $1.90	-0.000213***	0.995***	-0.000218	-0.00308**	-0.00185
	(7.80e-05)	(0.000926)	(0.000169)	(0.00132)	(0.00184)
Dist. to Oil / Gas Pipeline	0.00222*	1.019	0.00735**	0.00859	-0.00364
	(0.00120)	(0.0151)	(0.00346)	(0.0165)	(0.0183)
Dist. to HV Power Transmission Line	-0.00378***	0.864***	0.00345	-0.148***	-0.0996***
	(0.00106)	(0.0103)	(0.00314)	(0.0165)	(0.0187)
Total Agricultural Production Value	0.000240	1.028***	-0.000873	0.0319***	0.0259***
	(0.000198)	(0.00376)	(0.000580)	(0.00433)	(0.00490)
Mean Elevation	0.00146	1.053**	0.0121***	0.0800***	0.137***
	(0.00179)	(0.0221)	(0.00400)	(0.0250)	(0.0306)
Std. Dev. of Elevation	1.75e-05	1.001**	-2.65e-05	8.91e-05	0.000850***
	(2.21e-05)	(0.000265)	(3.77e-05)	(0.000308)	(0.000303)
gEcon Mean	-0.000708**	0.984***	-0.0120***	-0.0101**	-0.00771**
	(0.000328)	(0.00348)	(0.00335)	(0.00409)	(0.00315)
Percent Urban Territory	0.523***	4.286***	1.592***	1.517***	0.985***
	(0.0858)	(0.500)	(0.479)	(0.122)	(0.134)
Mean Nighttime Lights	0.00158	1.014*	0.0411**	-0.0113***	-0.0103***
	(0.00114)	(0.00826)	(0.0177)	(0.00365)	(0.00294)

Continued on next page

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
Mean Population Density	0.00615***	1.679***	0.0103***	0.507***	0.599***
	(0.00145)	(0.0286)	(0.00258)	(0.0181)	(0.0215)
Mean Road Density	-0.0486**	0.146***	-0.0643	-1.537***	-1.933***
	(0.0221)	(0.0344)	(0.0533)	(0.374)	(0.419)
Former Metropole					
Belgium	-0.0101*	0.322***	-0.0223	-0.974***	-0.568**
	(0.00557)	(0.0506)	(0.0238)	(0.248)	(0.234)
France	0.00691	1.017	0.0339***	-0.435**	0.363*
	(0.00789)	(0.129)	(0.0125)	(0.209)	(0.187)
Germany	0.00471	0.898	-0.00225	-0.591**	0.0178
	(0.0110)	(0.192)	(0.0237)	(0.239)	(0.278)
Italy	0.0105	3.322***	0.0615***	0.934***	1.263***
	(0.0134)	(0.547)	(0.0158)	(0.190)	(0.205)
Portugal	0.00255	0.583***	0.0146	-1.126***	-0.627***
	(0.00601)	(0.0875)	(0.0189)	(0.169)	(0.201)
UK	0.00365	0.725***	0.00965	-0.546***	0.0136
	(0.00776)	(0.0861)	(0.0125)	(0.174)	(0.170)
Herbst Typology					
Difficult	-0.00704	0.567***	0.0151	-0.694***	-0.308***
	(0.00714)	(0.0408)	(0.0113)	(0.117)	(0.109)
Favorable	0.00385	1.173**	-0.00133	0.0535	0.0233

Continued on next page

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
	(0.00660)	(0.0773)	(0.0112)	(0.0838)	(0.0907)
Hinterland	0.0124*	1.286***	0.0125	0.124	0.299***
	(0.00731)	(0.0919)	(0.0173)	(0.0838)	(0.0944)
Influential, Excluded, or Mixed Groups					
Influential	-0.00414	0.564***	-0.0245*	-0.823***	-0.791***
	(0.00917)	(0.0717)	(0.0145)	(0.152)	(0.151)
Excluded	0.000989	0.941	-0.0276	-0.398	-1.091***
	(0.0128)	(0.267)	(0.0165)	(0.291)	(0.235)
Mixed	0.00389	0.686***	-0.0255*	-0.607***	-0.774***
	(0.00926)	(0.0917)	(0.0142)	(0.153)	(0.166)
Köppen Climate Classification Code					
Af - Tropical Rainforest	0.0362***	5.593**	0.102**	1.489***	1.822***
	(0.0126)	(4.120)	(0.0468)	(0.434)	(0.665)
Am - Tropical Monsoon	0.0227*	4.081*	0.117**	1.307***	1.988***
	(0.0114)	(2.999)	(0.0500)	(0.421)	(0.681)
As - Savanna, Dry Summer	0.0369***	6.750***	0.128***	1.899***	1.914***
	(0.0120)	(5.002)	(0.0461)	(0.443)	(0.664)
Aw - Savanna, Dry Winter	0.0231**	4.869**	0.103**	1.546***	1.846***
	(0.0105)	(3.560)	(0.0475)	(0.412)	(0.651)
BWk - Cold Desert	0.0303***	10.55***	0.0990**	2.588***	2.312***
	(0.0104)	(7.861)	(0.0461)	(0.450)	(0.670)

Continued on next page

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
BWh - Hot Desert	0.0202*	3.647*	0.0821*	1.357***	1.645**
	(0.0103)	(2.676)	(0.0465)	(0.418)	(0.650)
BSk - Cold Semi-Arid	0.0349**	6.645**	0.0718	1.803***	2.007***
	(0.0134)	(4.892)	(0.0499)	(0.421)	(0.657)
BSh - Hot Semi-Arid	0.0213**	4.517**	0.0857*	1.367***	1.612**
	(0.00967)	(3.305)	(0.0470)	(0.414)	(0.650)
Cfa - Humid Subtropical	0.00986	3.341	-0.0242	0.909**	1.283*
	(0.0125)	(2.505)	(0.0750)	(0.438)	(0.670)
Cfb - Oceanic	0.0192	3.666*	0.0635	0.954**	1.322**
	(0.0122)	(2.699)	(0.0594)	(0.431)	(0.670)
Csa - Hot Summer Mediterranean	0.119***	8.112***	0.115	1.901***	1.948***
	(0.0269)	(5.978)	(0.0955)	(0.421)	(0.655)
Csb - Warm Summer Mediterranean	0.00542	3.651*	-0.271***	1.298***	1.675**
	(0.0172)	(2.870)	(0.0843)	(0.479)	(0.682)
Cwa - Monsoon	0.0201**	3.743*	0.0743	1.091**	1.275*
	(0.00978)	(2.751)	(0.0496)	(0.424)	(0.656)
Cwb - Subtropical Highland	0.00916	3.281	0.0771	0.912**	1.602**
	(0.0105)	(2.418)	(0.0468)	(0.429)	(0.663)
Constant	0.0911	19.65***	-0.426*	5.544***	3.353***
	(0.0762)	(20.85)	(0.212)	(1.142)	(1.109)
Observations	274,330	274,330	274,330	274,330	274,330

Continued on next page

Table 4.4. Results of Conventional (Non-Spatial) Models

	Control Points (Binary)		Control Points (Count)		
	(1)	(2)	(3)	(4)	(5)
Variable	OLS	Logit (Odds Ratio)	OLS	Negative Binomial	Poisson
R^2	0.221		0.349		
Pseudo R^2		0.391		0.279	0.601
χ^2		13260		15600	24646
Dispersion Parameter (α)				5.290	
Log Likelihood		-17997		-29059	-38051

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. All distance measures are in meters and log transformed.

Table 4.5. Variable Importance — Random Forest Classification

Variable	Rank	n	μ	σ	Min	0.25	Median	0.75	Max
Mean Population Density	1	25	16.41	0.20	16.13	16.25	16.41	16.55	16.78
Mean Nighttime Lights	2	25	15.11	0.19	14.67	14.97	15.11	15.27	15.38
Dist. to ACLED Demonstration	3	25	14.92	0.19	14.57	14.77	14.91	15.03	15.38
Control Points in Adjoining Cells, Count	4	25	14.01	0.20	13.74	13.86	13.99	14.12	14.47
Dist. to Major City	5	25	13.76	0.15	13.51	13.63	13.74	13.87	14.12
Dist. to Airport	6	25	13.32	0.19	12.93	13.23	13.33	13.43	13.75
Mean Road Density	7	25	13.29	0.26	12.90	13.17	13.25	13.39	13.87
Total Agricultural Production Value (in Millions of USD)	8	25	13.23	0.24	12.56	13.09	13.28	13.43	13.54
Dist. to Power Gen. Fac.	9	25	13.03	0.19	12.48	12.96	13.03	13.13	13.48
Dist. to HV Power Transmission Line	10	25	12.97	0.15	12.68	12.89	12.96	13.05	13.25
Std. Dev. of Elevation	11	25	12.63	0.22	12.24	12.45	12.63	12.81	13.03
Dist. to National Capital	12	25	12.53	0.18	12.05	12.42	12.49	12.65	12.89
Dist. to ACLED Violence	13	25	12.18	0.16	11.87	12.11	12.21	12.26	12.45
Mean Dist. to 3 Major Cities	14	25	12.04	0.14	11.71	11.96	12.03	12.12	12.28
Dist. to ACLED Territorial Change	15	25	11.98	0.15	11.68	11.87	12.02	12.07	12.25
Poverty Headcount Ratio, $1.90	16	25	11.81	0.24	11.39	11.68	11.77	12.00	12.29
Dist. to Non-Energy Mineral Facility	17	25	11.76	0.16	11.49	11.67	11.75	11.85	12.08
Mean Elevation	18	25	11.69	0.17	11.33	11.55	11.70	11.82	12.11
Dist. to Major Mineral Deposit	19	25	11.44	0.18	10.95	11.39	11.42	11.54	11.80
Dist. to Energy Extraction / Refinement Facility	20	25	11.42	0.15	11.01	11.34	11.44	11.52	11.66
Dist. to International Border	21	25	11.31	0.21	10.88	11.15	11.31	11.49	11.76
Dist. to Oil / Gas Pipeline	22	25	11.30	0.14	11.00	11.20	11.31	11.41	11.49
Revised Combined Polity Score	23	25	11.25	0.19	10.94	11.10	11.27	11.37	11.58
Dist. to A1 Centroid	24	25	10.99	0.13	10.76	10.87	11.03	11.07	11.22
Koeppen Classification	25	25	9.96	0.12	9.74	9.87	9.95	10.04	10.17
Number of Ethnic Groups	26	25	9.61	0.13	9.34	9.50	9.65	9.69	9.82
Years Since Independence	27	25	9.35	0.18	9.08	9.21	9.36	9.46	9.68
Former Metropole	28	25	5.22	0.08	4.98	5.18	5.22	5.26	5.44
Herbst Typology	29	25	4.39	0.09	4.22	4.30	4.39	4.45	4.55
Influential, Excluded, Mixed	30	25	4.10	0.08	3.96	4.04	4.11	4.14	4.30
Transborder Ethnic Kin, Binary	31	25	3.63	0.09	3.51	3.54	3.64	3.68	3.79
Urban	32	25	1.54	0.04	1.44	1.51	1.53	1.57	1.62

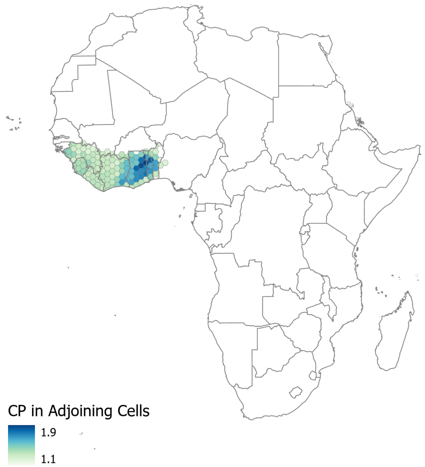
Table 4.6. Variable Importance — Random Forest Regression

Variable	Rank	n	μ	σ	Min	0.25	Median	0.75	Max
Control Points in Adjoining Cells, Count	1	25	48369.47	7509.57	32262.07	44109.44	48889.91	52559.02	67897.42
Mean Population Density	2	25	41907.70	6154.78	29151.79	38191.46	42600.50	47136.82	50786.67
Mean Nighttime Lights	3	25	21643.13	1943.83	17452.73	20298.74	21607.01	22705.61	25200.57
Dist. to Major City	4	25	14776.08	1843.99	10791.05	13906.18	14683.02	15637.07	18705.70
Dist. to National Capital	5	25	14400.47	2971.10	8043.02	13191.81	15215.56	16188.97	18782.01
Mean Dist. to 3 Major Cities	6	25	11201.72	3684.54	2761.61	9145.15	11831.96	13845.61	15264.03
Koeppen Classification	7	25	9914.33	2004.83	5214.37	9511.96	10477.42	11245.42	12423.55
Dist. to ACLED Violence	8	25	8896.85	1626.69	5516.00	7628.86	9346.57	10135.27	11453.94
Dist. to ACLED Demonstration	9	25	8781.45	1110.65	6926.35	7710.86	9046.33	9571.02	10700.14
Dist. to A1 Centroid	10	25	7034.07	1671.75	3637.23	6418.35	7464.54	8319.85	9140.72
Dist. to HV Power Transmission Line	11	25	6657.84	1302.19	3147.35	5935.82	6692.33	7616.78	8254.52
Total Agricultural Production Value	12	25	6155.35	1277.68	3316.47	6158.11	6486.71	6796.35	7684.29
Dist. to Energy Extraction / Refinement Facility	13	25	6097.32	1113.53	3518.23	5603.70	6131.42	6964.03	7531.53
Dist. to Power Gen. Fac.	14	25	5785.80	928.79	3687.19	5338.23	6049.09	6277.55	7609.01
Urban	15	25	5777.28	522.04	4869.65	5422.55	5813.52	6097.48	7218.07
Mean Road Density	16	25	5447.57	1036.98	2869.43	5111.25	5661.05	6155.83	7652.38
Mean Elevation	17	25	5277.97	874.57	2934.41	4782.43	5458.93	5789.93	6860.08
Dist. to Airport	18	25	4953.52	429.45	4105.39	4793.89	4898.27	5085.62	5962.75
Std. Dev. of Elevation	19	25	4683.45	753.96	3407.78	4096.24	4721.99	5240.17	6698.51
Dist. to ACLED Territorial Change	20	25	4624.10	851.79	2860.01	4359.16	4880.85	5233.41	6274.69
Dist. to Oil / Gas Pipeline	21	25	4363.88	576.06	2786.16	3988.86	4411.71	4665.40	5643.99
Dist. to Non-Energy Mineral Facility	22	25	4288.17	679.88	2860.15	3832.32	4228.21	4785.28	5529.67
Dist. to Major Mineral Deposit	23	25	4086.74	570.12	2715.62	3968.66	4150.01	4324.69	5147.75
Dist. to International Border	24	25	3995.68	398.71	3326.76	3752.05	3896.37	4210.53	5173.80
Number of Ethnic Groups	25	25	3746.69	827.18	2042.18	3182.28	3853.02	4326.47	4946.52
Poverty Headcount Ratio, $1.90	26	25	3339.02	397.85	2606.77	3084.56	3344.23	3485.49	4230.64
Revised Combined Polity Score	27	25	2762.39	373.42	2034.36	2472.85	2721.40	3029.45	3557.04
Herbst Typology	28	25	2674.98	312.26	2172.02	2486.88	2627.72	2943.60	3165.24
Former Metropole	29	25	2301.50	314.78	1794.56	1996.44	2233.85	2543.07	2900.55
Years Since Independence	30	25	1460.17	257.45	1070.16	1279.11	1438.53	1631.60	2078.14
Influential, Excluded, Mixed	31	25	592.20	82.64	429.20	552.46	601.09	652.12	737.54
Transborder Ethnic Kin, Binary	32	25	191.59	52.82	130.59	146.98	177.30	220.44	316.52

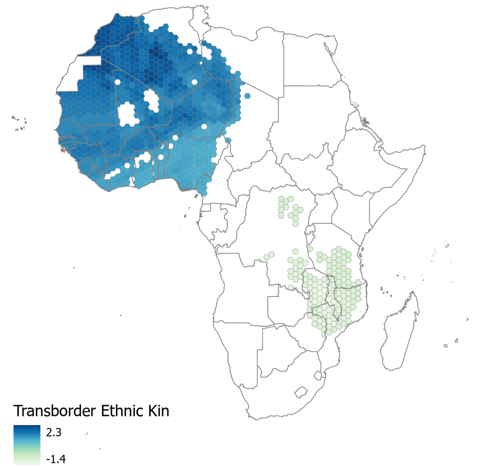
Table 4.7. Summary of Local Parameter Estimates, Continent-Scale Geographically Weighted Logistic Regression

Variable	n, Sig. Local Coeff.	μ	σ	Min	Max
Control Points in Adjoining Cells	217	1.38758	0.16039	1.08336	1.90100
Transborder Ethnic Kin, Binary	1373	1.11674	0.95662	-1.44269	2.33433
Number of Ethnic Groups	866	-0.11901	0.06049	-0.20336	0.22500
ACLED Violence Events, Count	2	-0.00623	0.00463	-0.00951	-0.00295
ACLED Demonstration Events, Count	2410	0.05059	0.03915	0.01199	0.28111
ACLED Territorial Change, Count	61	0.09548	0.19736	-0.61174	0.24405
Dist. to International Border	2988	0.40613	0.08068	0.22331	0.64916
Dist. to National Capital	376	-0.53413	0.16304	-1.09803	-0.35673
Dist. to Airport	3714	-0.95667	0.21675	-1.68212	-0.54857
Dist. to H20 Port	1122	0.50201	0.28941	-1.01250	0.90104
Dist. to Major City	706	-1.19013	0.27259	-1.94629	-0.77358
Dist. to Major Mineral Deposit	329	-0.51865	0.06745	-0.75576	-0.41742
Dis. to Power Generation Facility	1910	-0.69037	0.11542	-1.11429	-0.44371
Dist. to Non-Energy Mineral Facility	625	0.52731	0.06266	0.31397	0.73009
Dist. to Energy Extraction / Refinement Facility	30	-0.35698	0.06965	-0.43990	-0.22293
Dist. to ACLED Violence	29	-0.46379	0.58061	-0.91453	0.70242
Dist. to ACLED Demonstration	139	-0.82102	0.11533	-1.13138	-0.64614
Dist. to ACLED Territorial Change	96	0.55079	0.04925	0.46301	0.69494
Poverty Headcount Ratio	1271	-0.64158	0.19749	-1.09201	0.75600
Dist. to Oil / Gas Pipeline	716	0.33742	0.32568	-0.55909	0.67573
Dist. to HV Power Transmission Line	1825	-0.60631	0.17619	-1.12372	-0.32480
Mean Elevation	573	-0.27588	0.58482	-1.28054	0.99571
Std. Dev. Elevation	169	0.41193	0.05583	0.32605	0.54999
Percent Urban Territory	1849	0.93196	0.22418	0.55217	1.76312
Nighttime Lights	563	-4.64608	2.07693	-13.65054	-1.71596
Mean Population Density	780	0.36697	0.21935	-0.56623	0.70696
Mean Road Density	447	-8.29758	5.11786	-16.63575	22.69799
Total Agricultural Production	4086	0.12529	0.05750	0.00561	0.36544

Note: Only statistically significant local parameter estimates included in table.



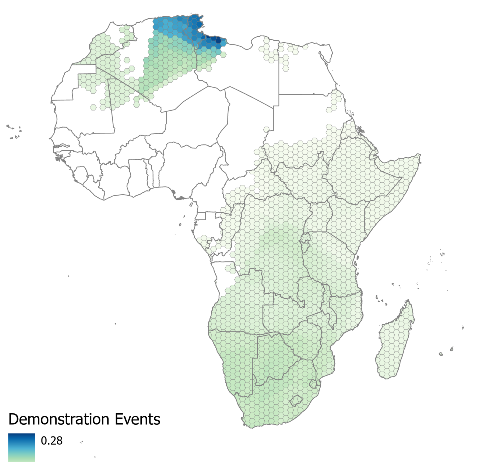
(1) Control Points in Adjoining Cells



(2) Transborder Ethnic Kin, Binary

(3) Number of Ethnic Groups

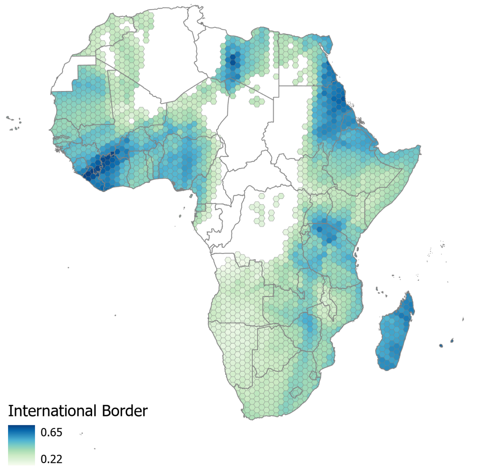
(4) ACLED Violence Events, Count



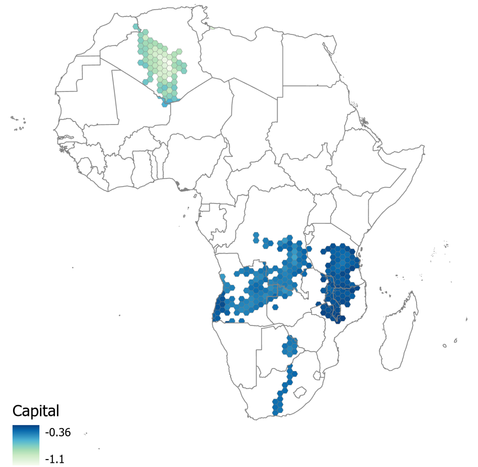
(5) ACLED Demonstration Events, Count

(6) ACLED Territorial Change Events, Count

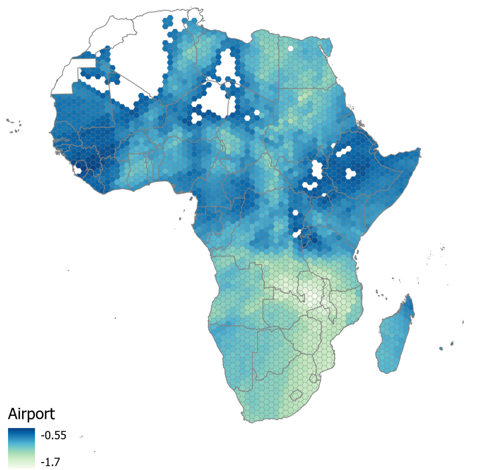
Figure 4.20. Local parameter estimates for a continent-scale geographically weighted logistic regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.21.



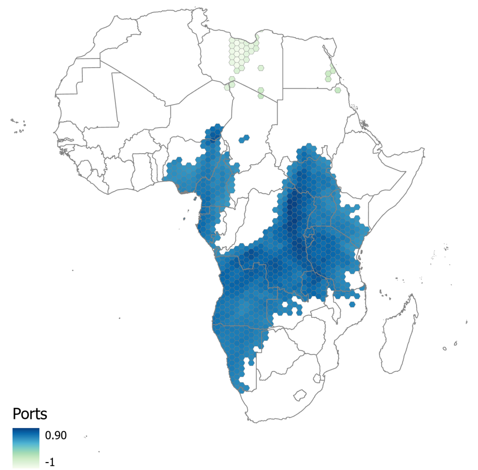
(7) Dist. to International Border



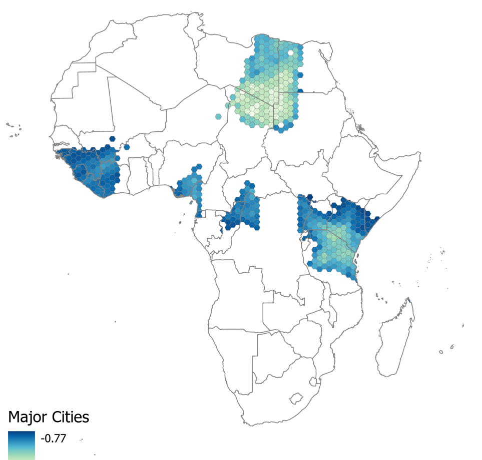
(8) Dist. to Capital



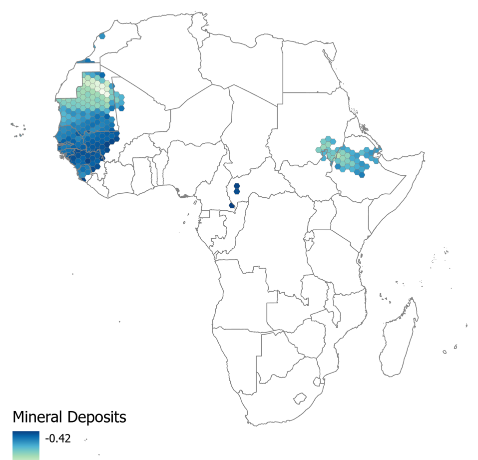
(9) Dist. to Airport



(10) Dist. to H20 Port

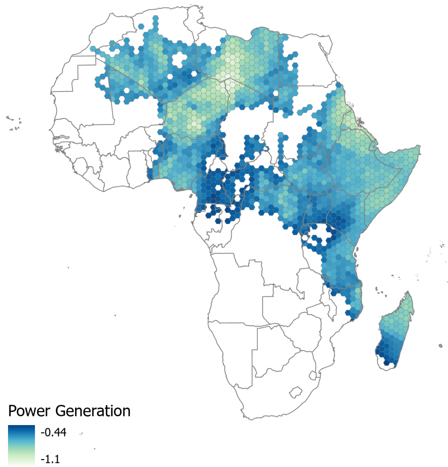


(11) Dist. to Major City

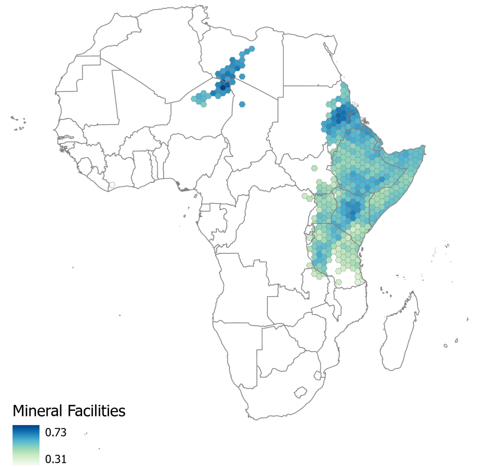


(12) Dist. to Major Mineral Deposit

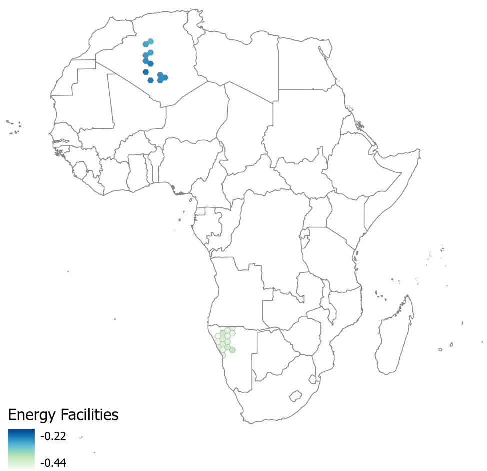
Figure 4.21. Local parameter estimates for a continent-scale geographically weighted logistic regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.22.



(13) Dist. to Power Generation Facility

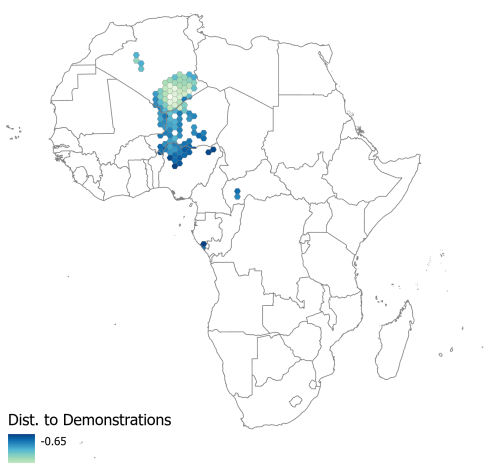


(14) Dist. to Non-Energy Mineral Facility

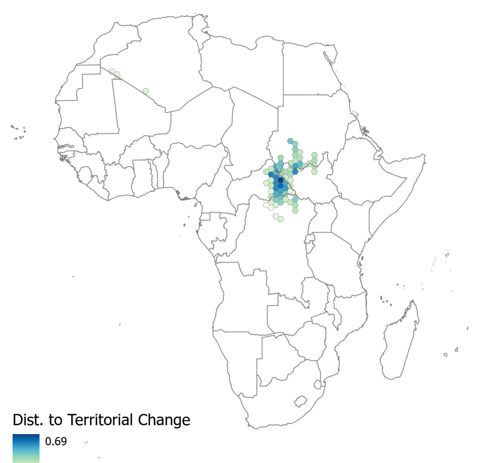


(15) Dist. to Energy Extraction / Refinement Facility

(16) Dist. to ACLED Violence

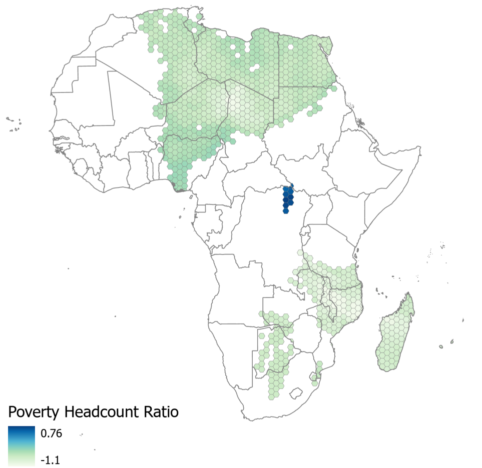


(17) Dist. to ACLED Demonstration

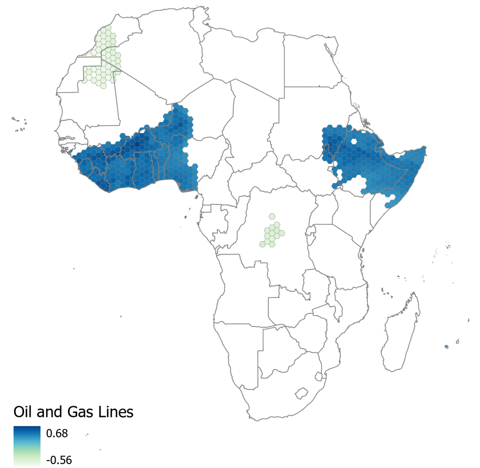


(18) Dist. to ACLED Territorial Change

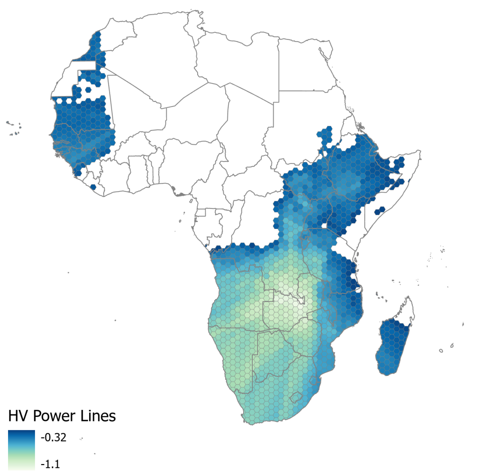
Figure 4.22. Local parameter estimates for a continent-scale geographically weighted logistic regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.23.



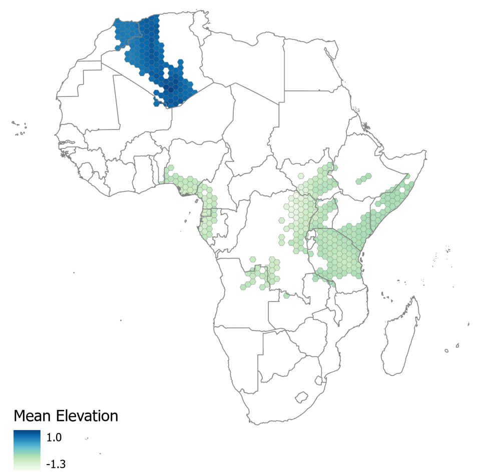
(19) Dist. to Power Generation Facility



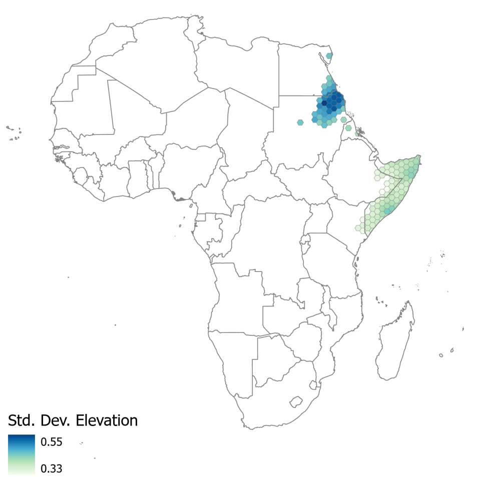
(20) Dist. to Oil / Gas Pipeline



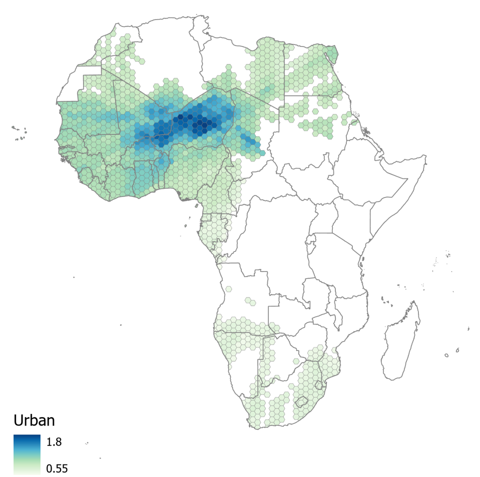
(21) Dist. to HV Power Transmission Line



(22) Mean Elevation

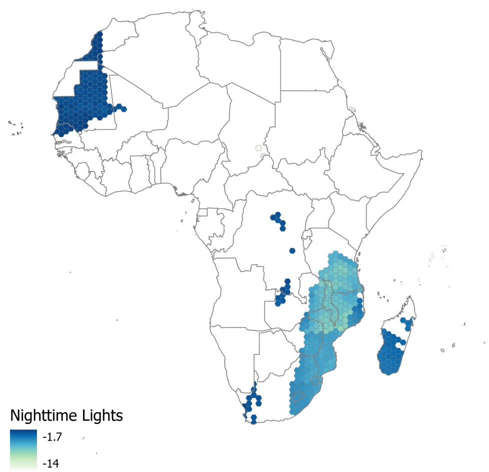


(23) Std. Dev. Elevation

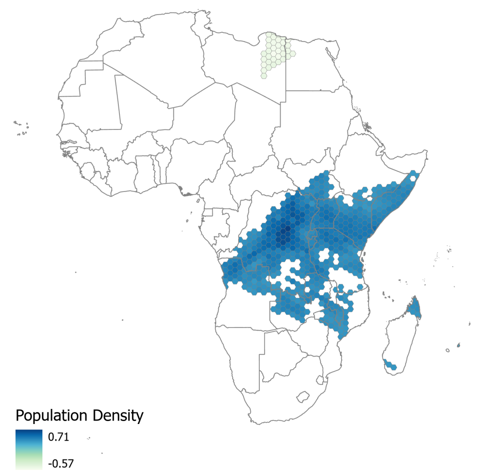


(24) Percent Urban Territory

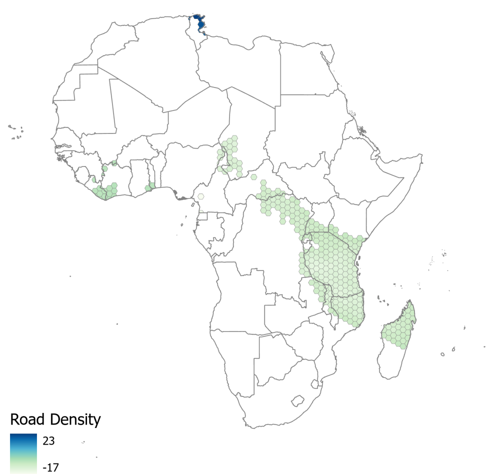
Figure 4.23. Local parameter estimates for a continent-scale geographically weighted logistic regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.24.



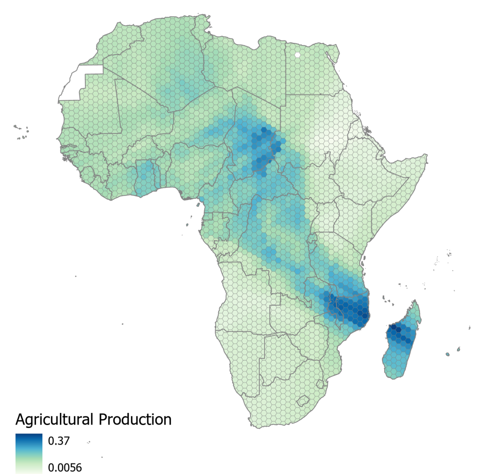
(25) Nighttime Lights



(26) Mean Population Density



(27) Mean Road Density



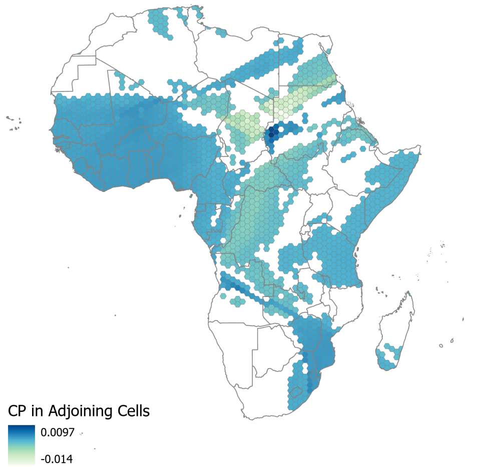
(28) Total Agricultural Production

Figure 4.24. Local parameter estimates for a continent-scale geographically weighted logistic regression model. Only coefficients with a t value significant at the 95% confidence level are shown.

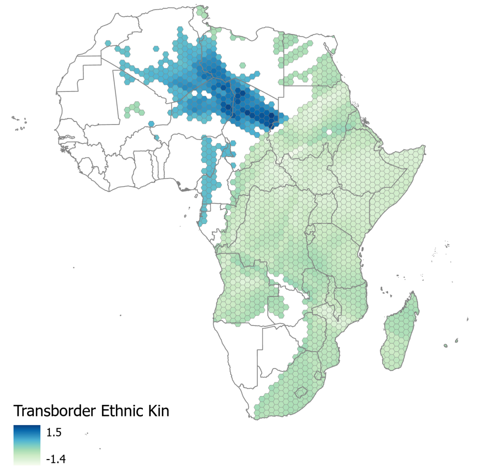
Table 4.8. Summary of Local Parameter Estimates, Continent-Scale Geographically Weighted Poisson Regression

Variable	n, Sig. Local Coeff.	μ	σ	Min	Max
Control Points in Adjoining Cells	2187	0.00035	0.00231	-0.01434	0.00974
Transborder Ethnic Kin, Binary	2385	-0.43692	0.51220	-1.38349	1.47944
Number of Ethnic Groups	3126	0.06353	0.07905	-0.16026	0.29196
ACLED Violence Events, Count	2988	0.00045	0.00048	-0.00208	0.00194
ACLED Demonstration Events, Count	3180	0.00075	0.00085	-0.00233	0.00495
ACLED Territorial Change, Count	2510	-0.01555	0.03453	-0.31036	0.09507
Dist. to International Border	3756	0.22658	0.15026	-0.45788	0.44104
Dist. to National Capital	3184	-0.16114	0.19556	-0.95795	0.29927
Dist. to Airport	4028	-0.52372	0.24170	-1.49187	0.20604
Dist. to H20 Port	2683	-0.05049	0.22749	-0.78273	0.50390
Dist. to Major City	2951	-0.43357	0.33453	-1.43002	0.82031
Dist. to Major Mineral Deposit	2528	0.10650	0.18868	-0.60276	0.68656
Dis. to Power Generation Facility	2749	-0.20896	0.21992	-0.71512	0.53305
Dist. to Non-Energy Mineral Facility	3001	0.20319	0.17426	-0.49113	0.54700
Dist. to Energy Extration / Refinement Facility	3294	0.06354	0.16607	-0.30644	0.46934
Dist. to ACLED Violence	3149	0.27548	0.40232	-0.62863	1.16716
Dist. to ACLED Demonstration	2944	-0.30879	0.36913	-1.42699	0.47112
Dist. to ACLED Territorial Change	2858	0.11792	0.27210	-0.52243	0.90982
Poverty Headcount Ratio	3266	-0.23547	0.32561	-0.96654	0.94266
Dist. to Oil / Gas Pipeline	2913	0.14643	0.18940	-0.36763	0.68431
Dist. to HV Power Transmission Line	3190	-0.20951	0.18144	-0.64368	0.39880
Mean Elevation	2962	0.00194	0.25973	-1.35754	0.64561
Std. Dev. Elevation	2486	0.00048	0.21955	-0.60478	0.42592
Percent Urban Territory	3762	0.30679	0.43185	-0.69730	1.24137
Nighttime Lights	2837	0.60648	1.48524	-4.88601	4.33611
Mean Population Density	3754	0.43635	0.16636	-0.35627	1.00515
Mean Road Density	3366	-2.55057	4.69586	-13.83195	14.35736
Total Agricultural Production	4086	0.17123	0.09525	0.00193	0.53363

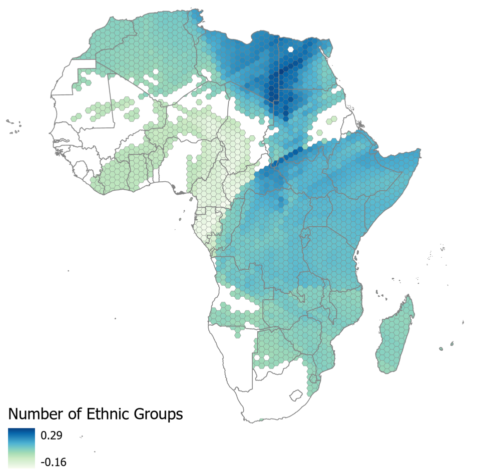
Note: Only statistically significant local parameter estimates included in table.



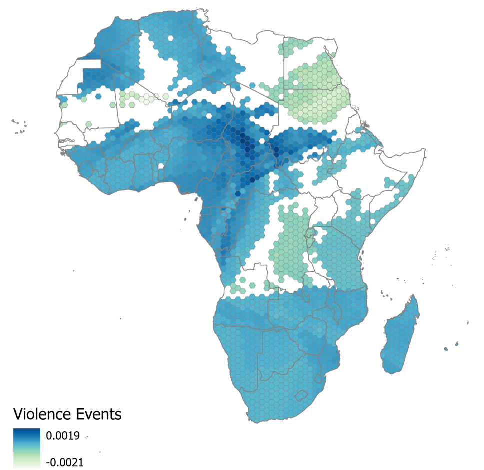
(1) Control Points in Adjoining Cells



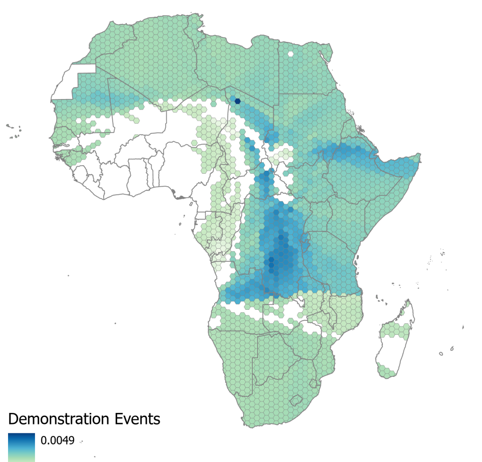
(2) Transborder Ethnic Kin, Binary



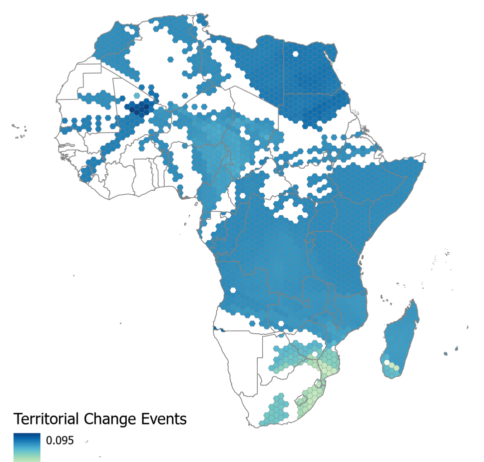
(3) Number of Ethnic Groups



(4) ACLED Violence Events, Count



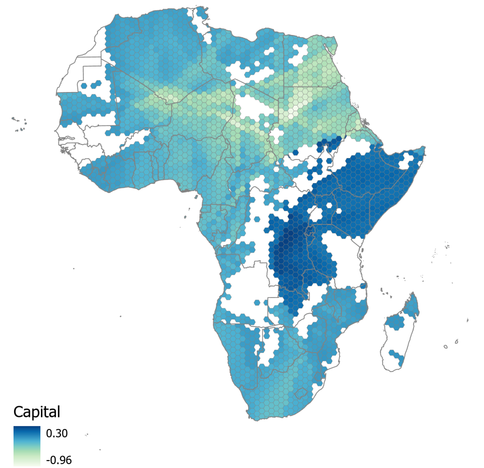
(5) ACLED Demonstration Events, Count



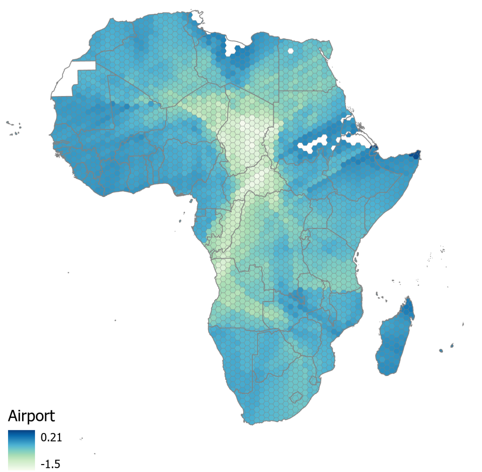
(6) ACLED Territorial Change Events, Count

Figure 4.25. Local parameter estimates for a continent-scale geographically weighted Poisson regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.26.

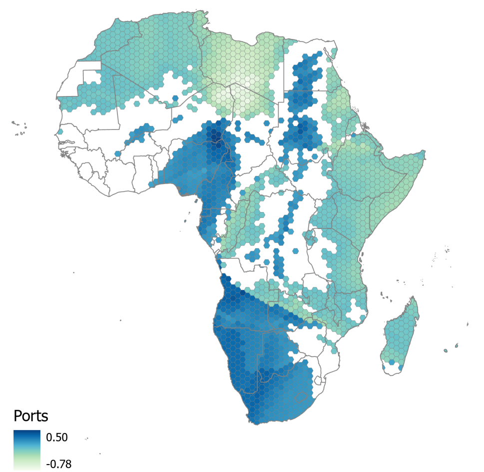
(7) Dist. to International Border



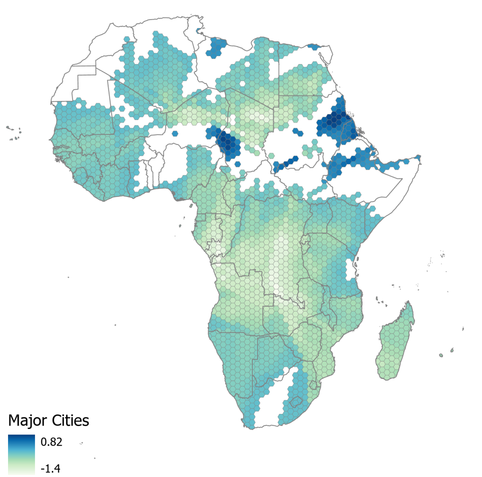
(8) Dist. to Capital



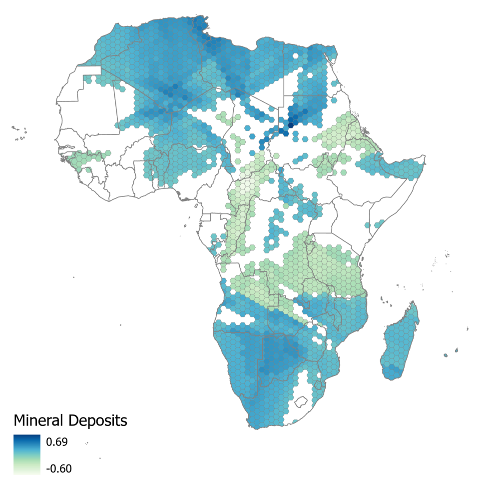
(9) Dist. to Airport



(10) Dist. to H20 Port

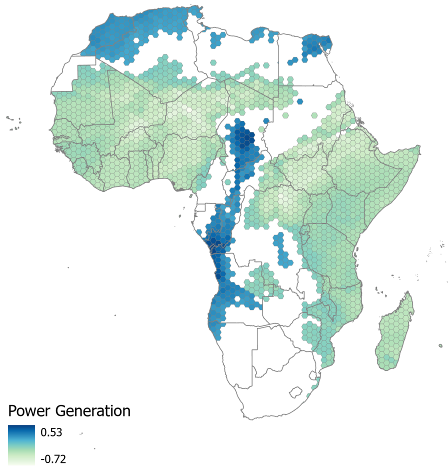


(11) Dist. to Major City

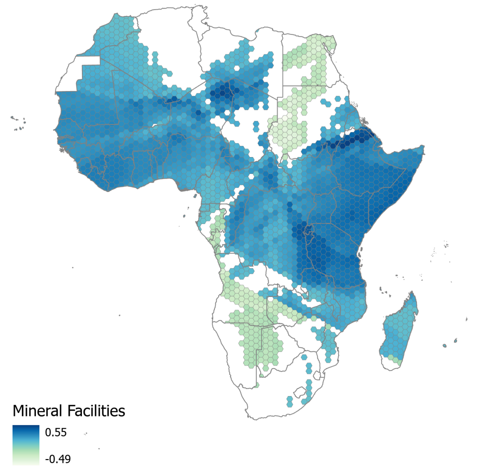


(12) Dist. to Major Mineral Deposit

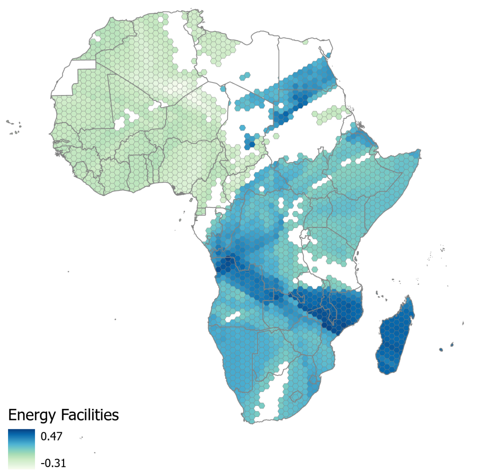
Figure 4.26. Local parameter estimates for a continent-scale geographically weighted Poisson regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.27.



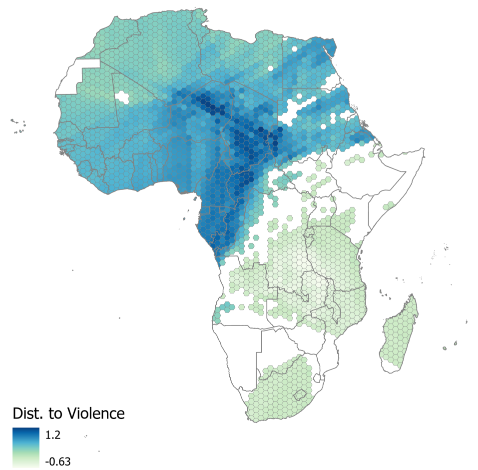
(13) Dist. to Power Generation Facility



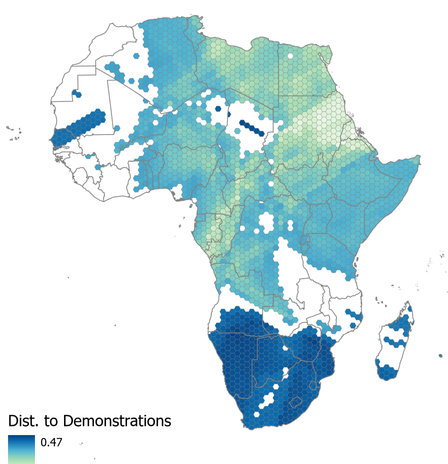
(14) Dist. to Non-Energy Mineral Facility



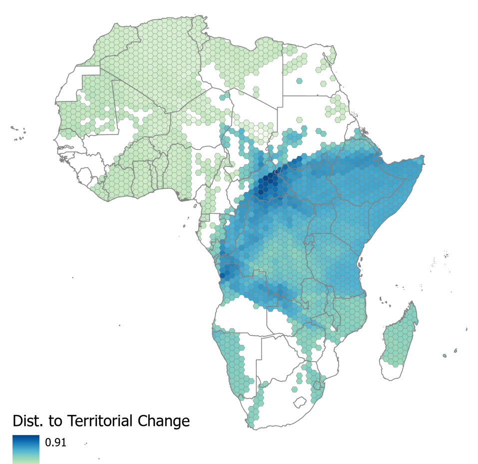
(15) Dist. to Energy Extraction / Refinement Facility



(16) Dist. to ACLED Violence

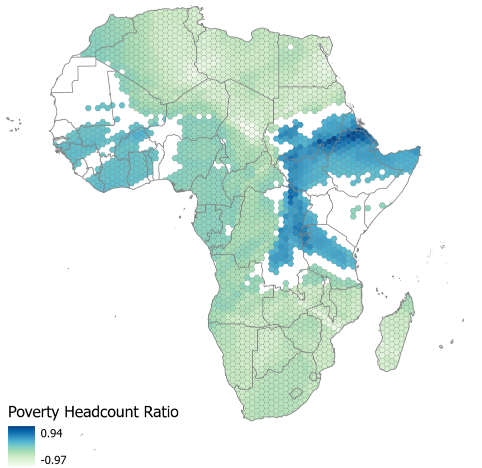


(17) Dist. to ACLED Demonstration

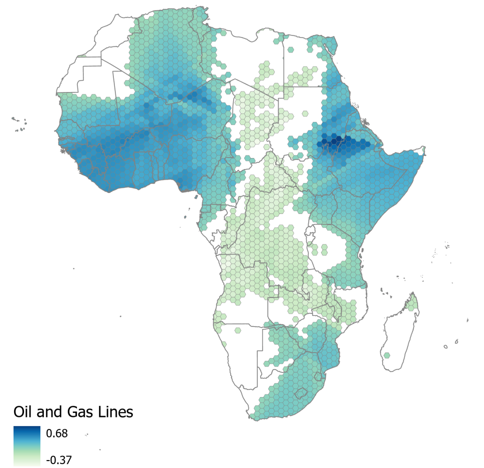


(18) Dist. to ACLED Territorial Change

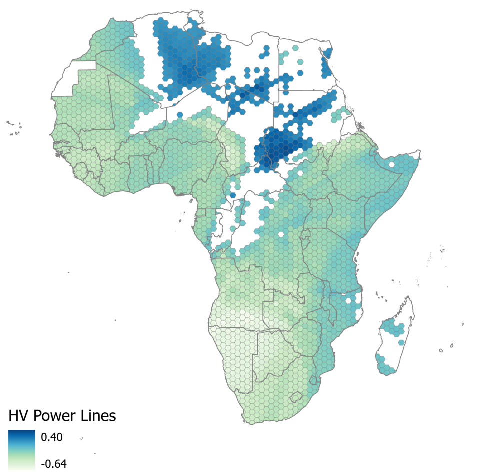
Figure 4.27. Local parameter estimates for a continent-scale geographically weighted Poisson regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.28.



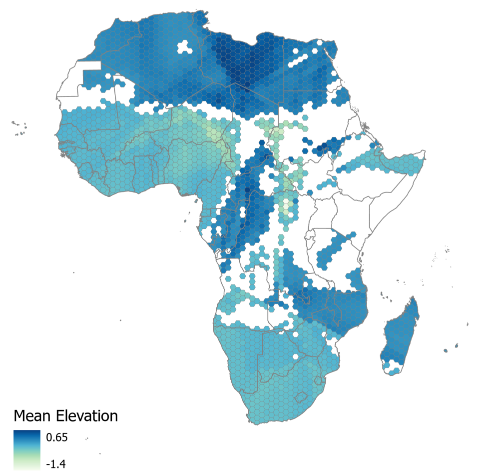
(19) Dist. to Power Generation Facility



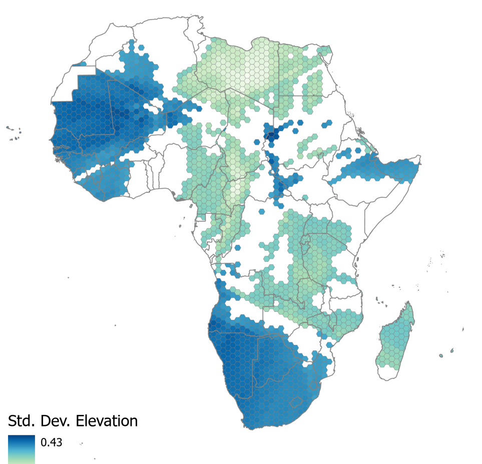
(20) Dist. to Oil / Gas Pipeline



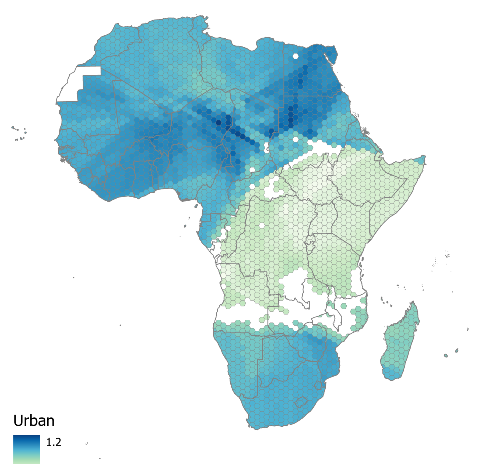
(21) Dist. to HV Power Transmission Line



(22) Mean Elevation

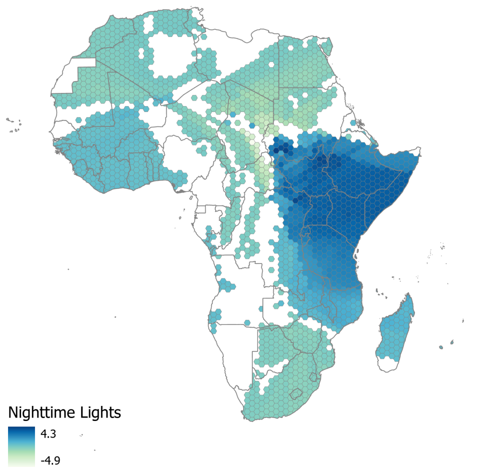


(23) Std. Dev. Elevation

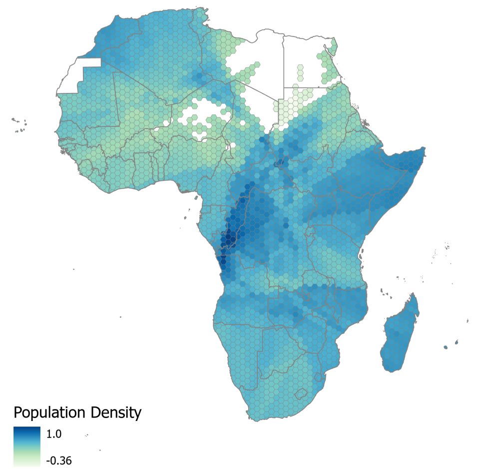


(24) Percent Urban Territory

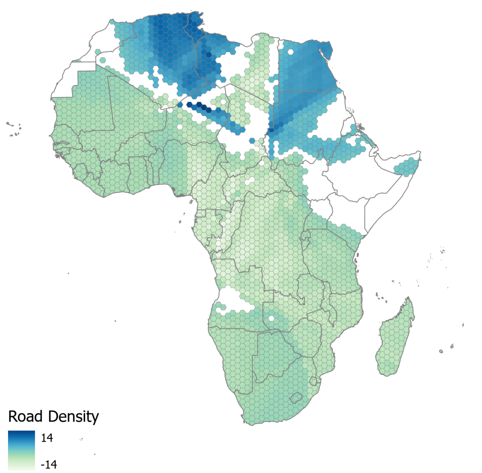
Figure 4.28. Local parameter estimates for a continent-scale geographically weighted Poisson regression model. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.29.



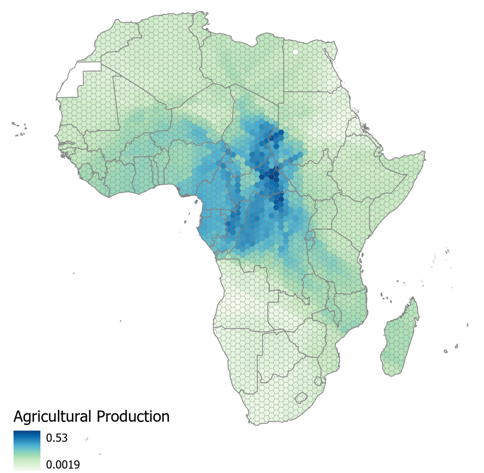
(25) Nighttime Lights



(26) Mean Population Density



(27) Mean Road Density



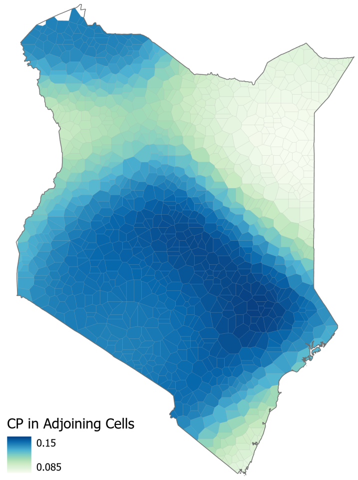
(28) Total Agricultural Production

Figure 4.29. Local parameter estimates for a continent-scale geographically weighted Poisson regression model. Only coefficients with a t value significant at the 95% confidence level are shown.

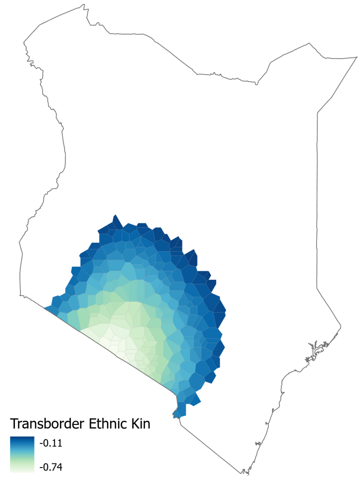
Table 4.9. Summary of Local Parameter Estimates, Kenya Sample Geographically Weighted Regression

Variable	n, Sig. Local Coeff.	μ	σ	Min	Max
Control Points in Adjoining Cells	1216	0.12388	0.01963	0.08534	0.15014
ACLED Violence Events, Count	171	-0.03445	0.06100	-0.11337	0.10097
ACLED Demonstration Events, Count	284	-0.02594	0.01302	-0.06318	-0.00877
ACLED Territorial Change, Count	132	-0.48182	1.68250	-3.45342	2.51513
Number of Ethnic Groups	145	0.02547	0.00755	0.01341	0.04152
Transborder Ethnic Kin, Binary	227	-0.34887	0.16837	-0.73624	-0.11227
Poverty Headcount Ratio	255	0.01047	0.00545	0.00299	0.02167
Dist. to International Border	148	-0.03800	0.05176	-0.07822	0.08058
Dist. to National Capital	655	-0.67551	0.28106	-1.24675	-0.21454
Dist. to Airport	70	-0.12414	0.03096	-0.17424	-0.06498
Dist. to H20 Port	389	-0.43817	0.20334	-0.79325	-0.12204
Dist. to Major City	33	-0.41740	0.07308	-0.52997	-0.30494
Dist. to Major Mineral Deposit	0	—	—	—	—
Dis. to Power Generation Facility	120	-0.08944	0.01873	-0.11834	-0.05565
Dist. to Non-Energy Mineral Facility	88	0.12490	0.01785	0.09224	0.15846
Dist. to Energy Extration / Refinement Facility	320	0.38602	0.66648	-0.39101	2.27306
Dist. to ACLED Violence	234	-0.10438	0.01893	-0.13352	-0.06212
Dist. to ACLED Demonstration	88	0.09563	0.01951	0.06406	0.13298
Dist. to ACLED Territorial Change	271	-0.23779	0.09712	-0.43318	-0.07467
Dist. to Oil / Gas Pipeline	163	0.08025	0.00975	0.06226	0.10075
Dist. to HV Power Transmission Line	0	—	—	—	—
Total Agricultural Production	162	0.01406	0.00354	0.00735	0.01996
Mean Elevation	139	0.22484	0.06371	0.10734	0.33422
Std. Dev. Elevation	44	-0.06202	0.00541	-0.06921	-0.05149
Percent Urban Territory	196	0.15964	0.06099	0.06789	0.30067
Nighttime Lights	525	1.42197	0.72186	0.34972	3.32865
Mean Population Density	85	-0.04681	0.00602	-0.05885	-0.03770
Mean Road Density	290	-4.57210	1.95220	-9.55899	-1.66928

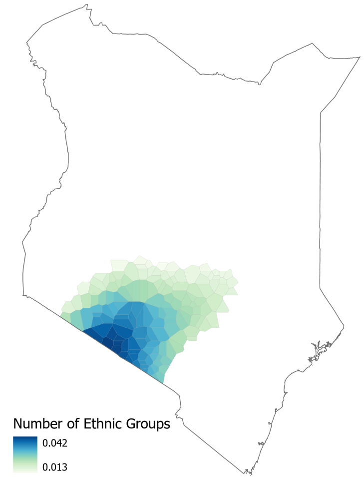
Note: Only statistically significant local parameter estimates included in table.



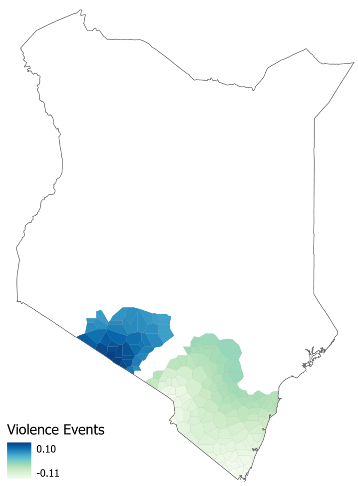
(1) Control Points in Adjoining Cells



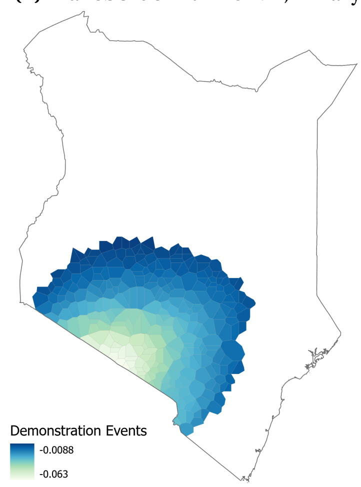
(2) Transborder Ethnic Kin, Binary



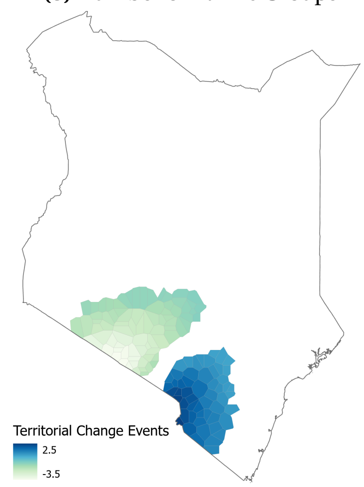
(3) Number of Ethnic Groups



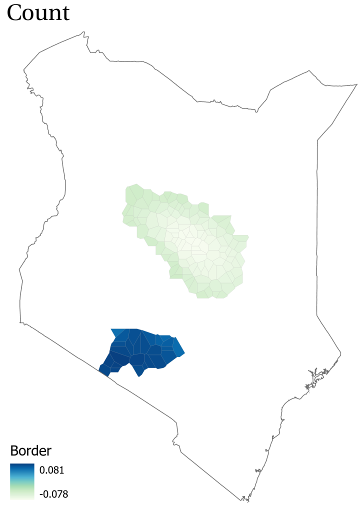
(4) ACLED Violence Events, Count



(5) ACLED Demonstration Events, Count

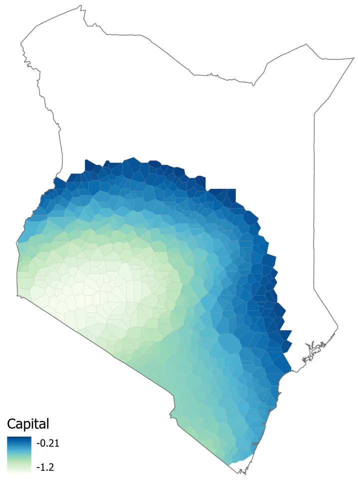


(6) ACLED Territorial Change Events, Count

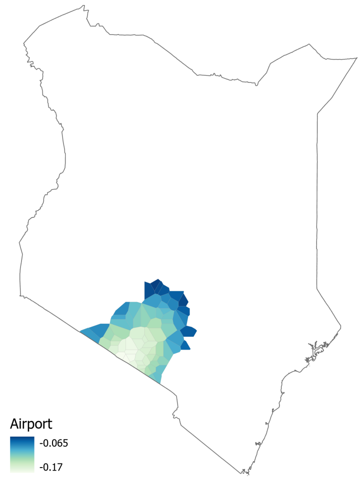


(7) Dist. to International Border

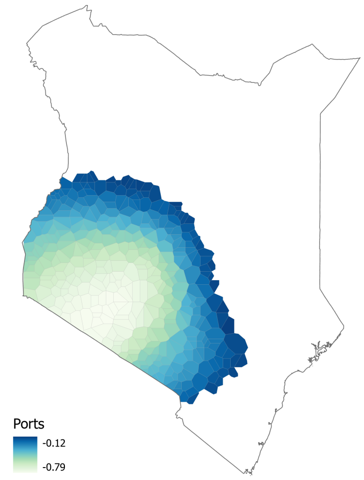
Figure 4.30. Local parameter estimates for the Kenya sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.31.



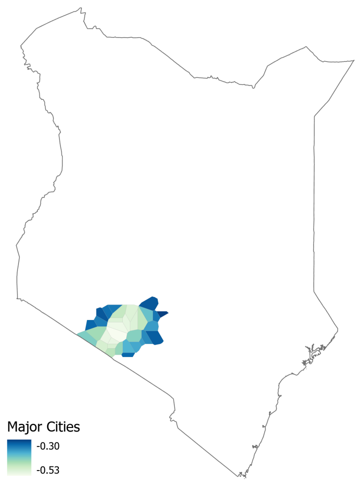
(8) Dist. to Capital



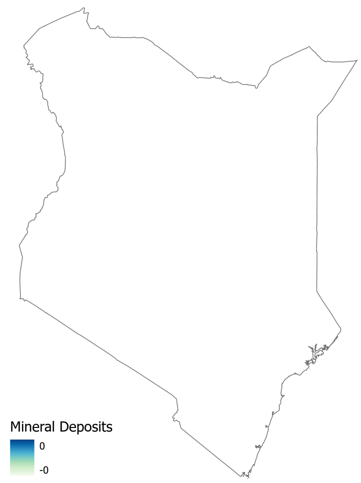
(9) Dist. to Airport



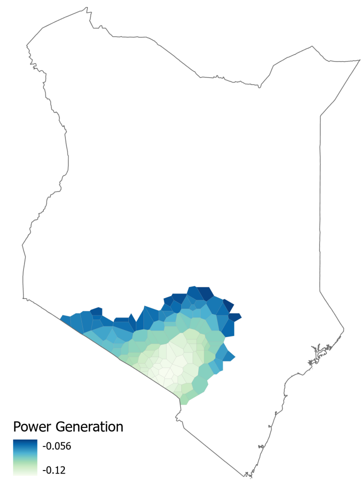
(10) Dist. to H20 Port



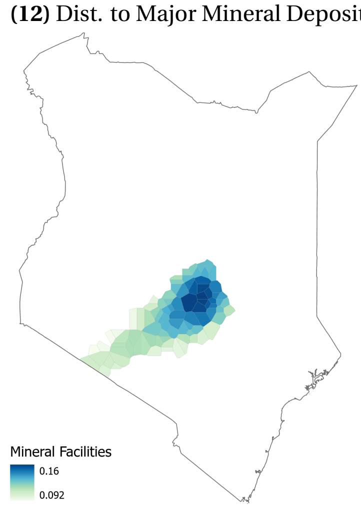
(11) Dist. to Major City



(12) Dist. to Major Mineral Deposit

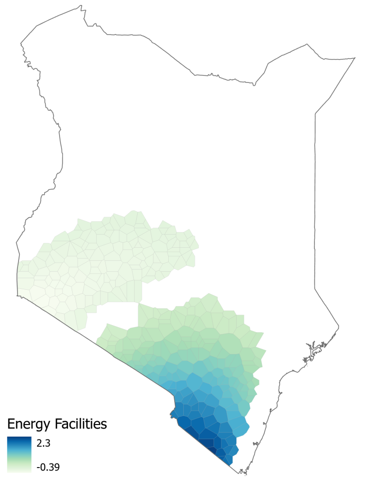


(13) Dist. to Power Generation Facility

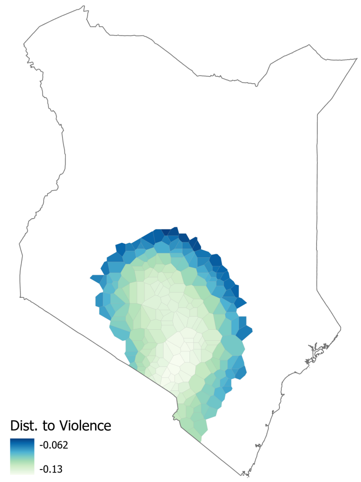


(14) Dist. to Non-Energy Mineral Facility

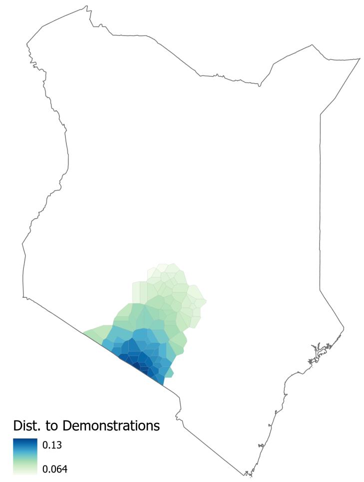
Figure 4.31. Local parameter estimates for the Kenya sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.32.



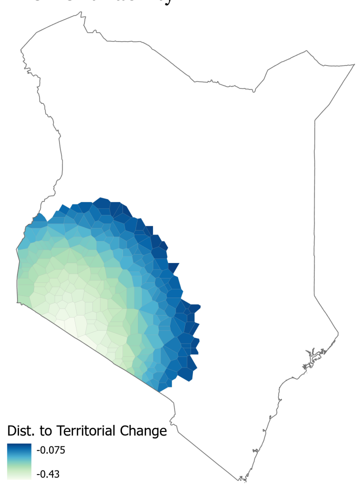
(15) Dist. to Energy Extraction / Refinement Facility



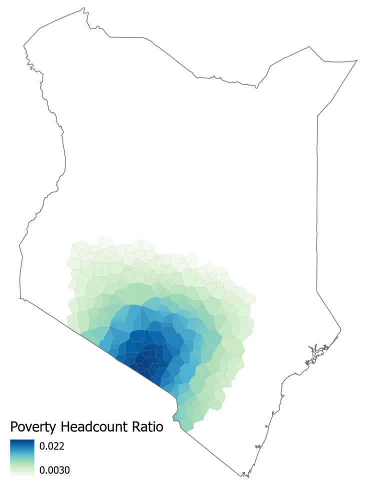
(16) Dist. to ACLED Violence



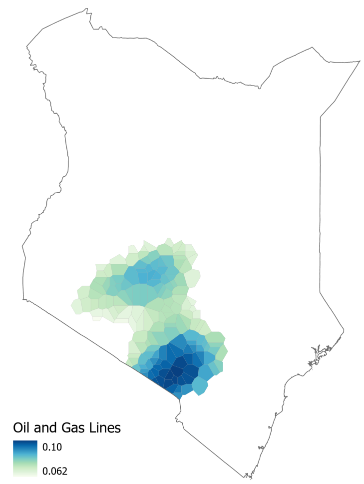
(17) Dist. to ACLED Demonstration



(18) Dist. to ACLED Territorial Change



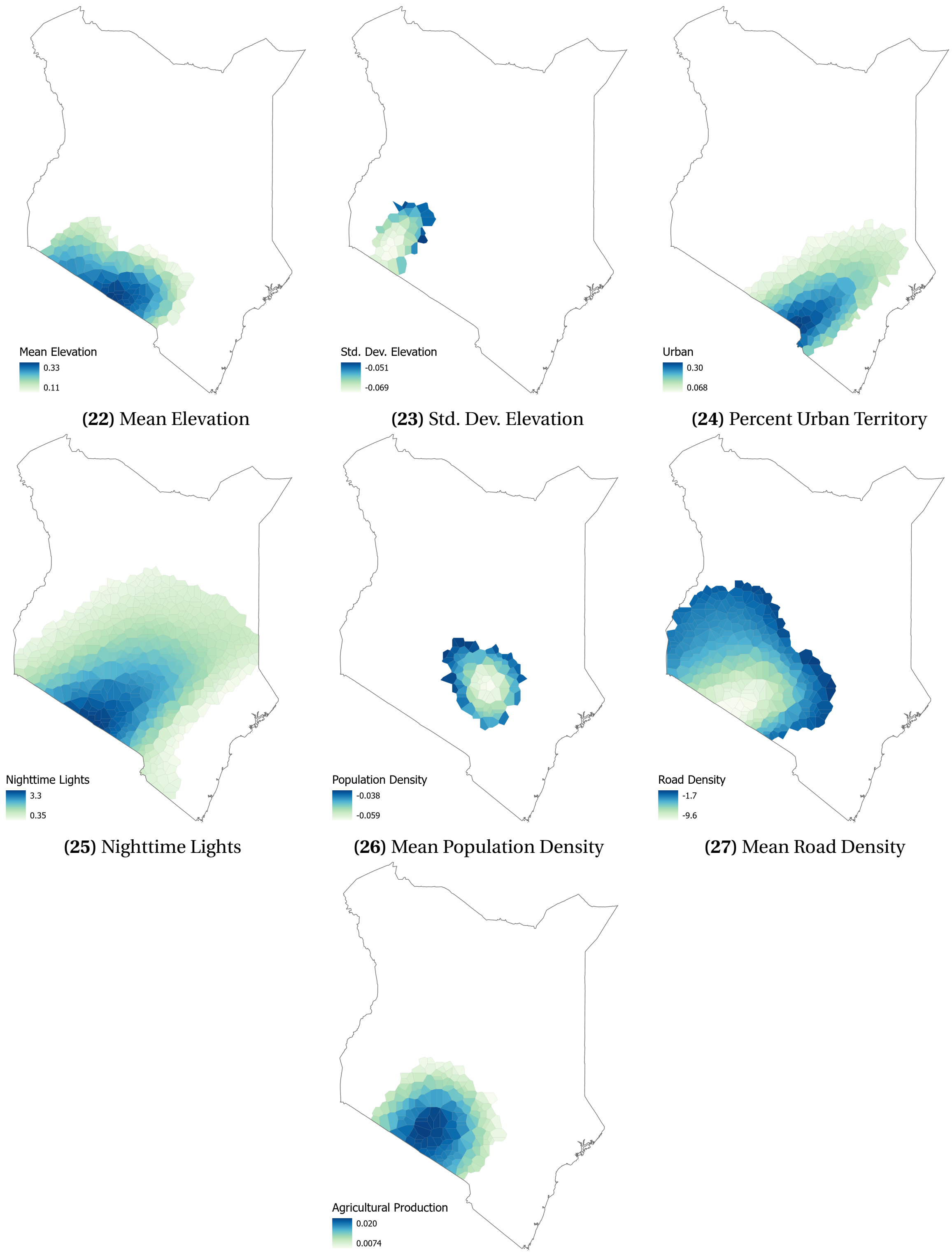
(19) Dist. to Power Generation Facility



(20) Dist. to Oil / Gas Pipeline

(21) Dist. to HV Power Transmission Line

Figure 4.32. Local parameter estimates for the Kenya sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.33.



(22) Mean Elevation **(23)** Std. Dev. Elevation **(24)** Percent Urban Territory

(25) Nighttime Lights **(26)** Mean Population Density **(27)** Mean Road Density

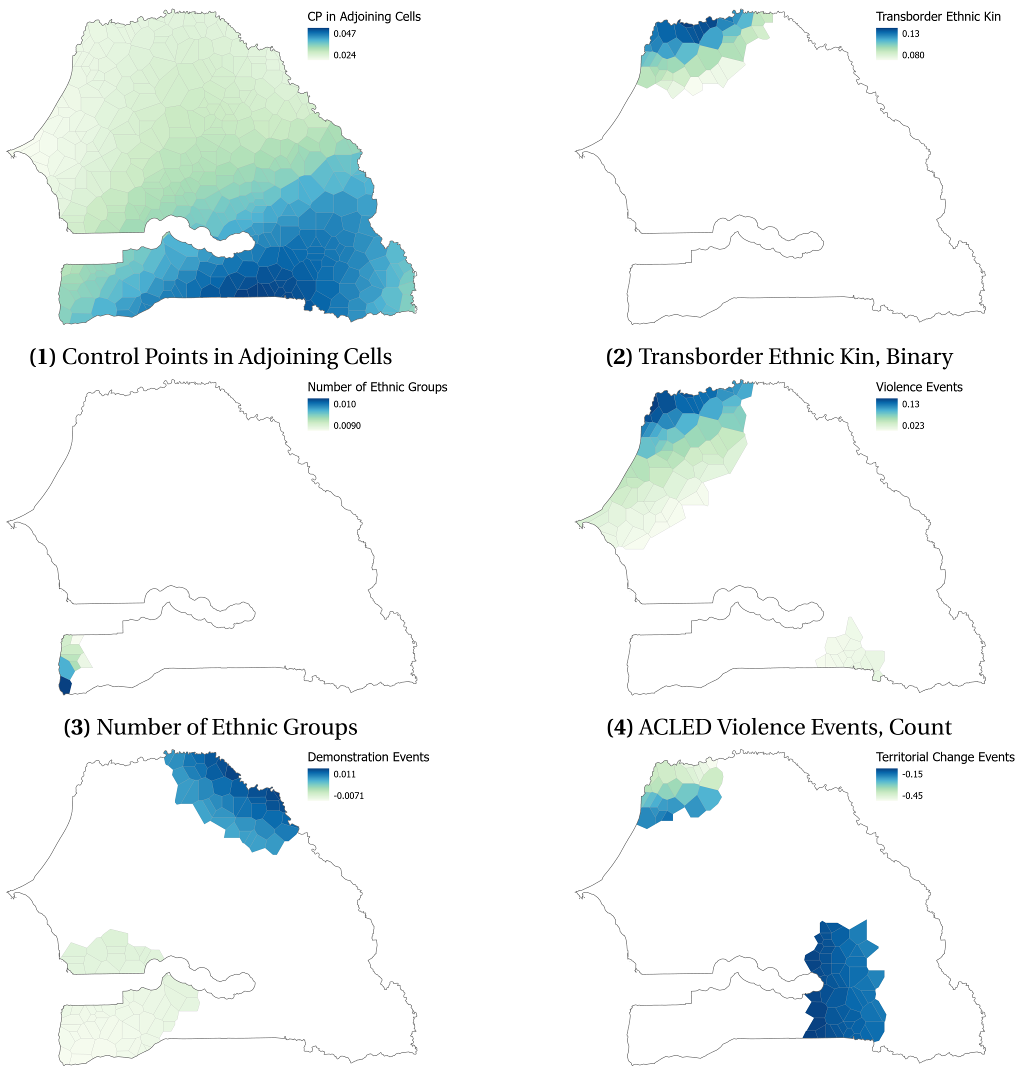
(28) Total Agricultural Production

Figure 4.33. Local parameter estimates for the Kenya sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown.

Table 4.10. Summary of Local Parameter Estimates, Senegal Sample Geographically Weighted Regression

Variable	n, Sig. Local Coeff.	μ	σ	Min	Max
Control Points in Adjoining Cells	428	0.03345	0.00641	0.02372	0.04710
ACLED Violence Events, Count	91	0.05751	0.03233	0.02267	0.12798
ACLED Demonstration Events, Count	98	-0.00026	0.00689	-0.00712	0.01111
ACLED Territorial Change, Count	76	-0.22271	0.08057	-0.44569	-0.14955
Number of Ethnic Groups	8	0.00940	0.00041	0.00896	0.01022
Transborder Ethnic Kin, Binary	44	0.10669	0.01574	0.07985	0.13355
Poverty Headcount Ratio	132	-0.00203	0.00045	-0.00287	-0.00115
Dist. to International Border	48	-0.02376	0.00567	-0.03535	-0.01419
Dist. to National Capital	348	6.61070	1.29237	3.60041	9.60033
Dist. to Airport	134	-0.04265	0.01495	-0.07778	-0.02118
Dist. to H20 Port	266	-0.10954	0.03488	-0.20353	-0.03473
Dist. to Major City	372	-0.14296	0.05043	-0.24793	-0.07336
Dist. to Major Mineral Deposit	228	0.06176	0.03030	0.02665	0.14901
Dis. to Power Generation Facility	35	0.03446	0.00637	0.02222	0.04611
Dist. to Non-Energy Mineral Facility	127	0.11685	0.04566	0.03651	0.19432
Dist. to Energy Extration / Refinement Facility	332	-6.78390	1.41444	-10.68876	-4.11070
Dist. to ACLED Violence	47	-0.03017	0.00469	-0.03716	-0.02051
Dist. to ACLED Demonstration	137	0.02852	0.00527	0.01921	0.03861
Dist. to ACLED Territorial Change	109	0.05576	0.08242	-0.08750	0.21403
Dist. to Oil / Gas Pipeline	29	0.84389	0.15347	0.57988	1.11326
Dist. to HV Power Transmission Line	85	0.01869	0.00446	0.01129	0.02726
Total Agricultural Production	0	—	—	—	—
Mean Elevation	285	-0.05369	0.02955	-0.08943	0.05252
Std. Dev. Elevation	41	-0.03706	0.00596	-0.05245	-0.02846
Percent Urban Territory	237	-0.04112	0.00625	-0.05766	-0.02830
Nighttime Lights	348	0.38238	0.07512	0.19134	0.51372
Mean Population Density	37	0.02927	0.00556	0.02104	0.04204
Mean Road Density	334	-0.93235	0.35131	-1.69618	-0.39826

Note: Only statistically significant local parameter estimates included in table.



(1) Control Points in Adjoining Cells

(2) Transborder Ethnic Kin, Binary

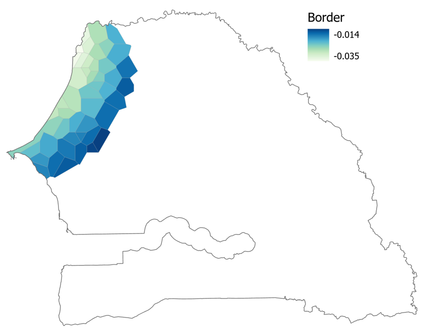
(3) Number of Ethnic Groups

(4) ACLED Violence Events, Count

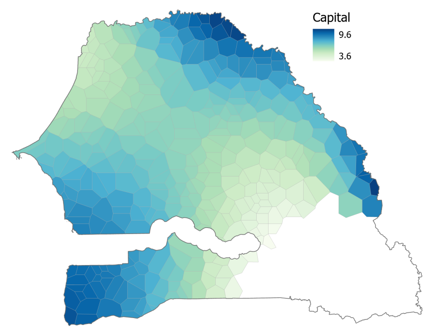
(5) ACLED Demonstration Events, Count

(6) ACLED Territorial Change Events, Count

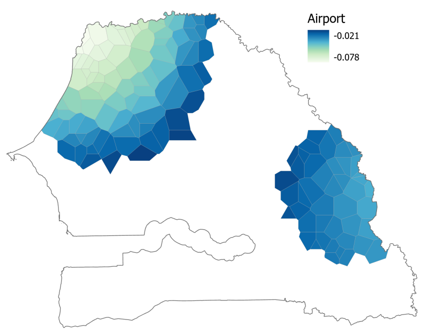
Figure 4.34. Local parameter estimates for the Senegal sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.35.



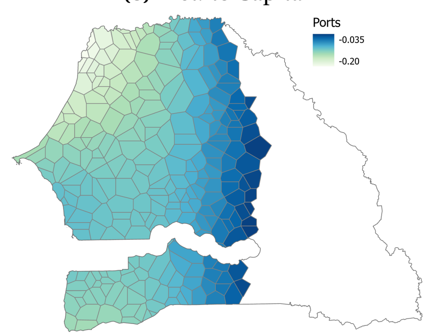
(7) Dist. to International Border



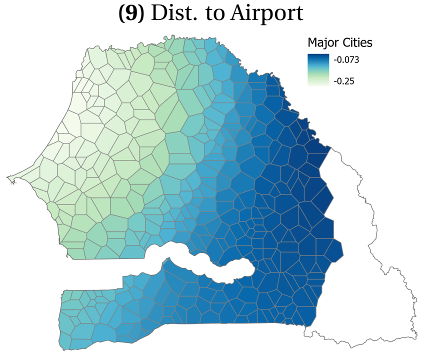
(8) Dist. to Capital



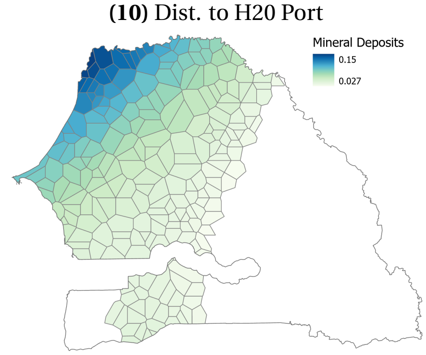
(9) Dist. to Airport



(10) Dist. to H20 Port

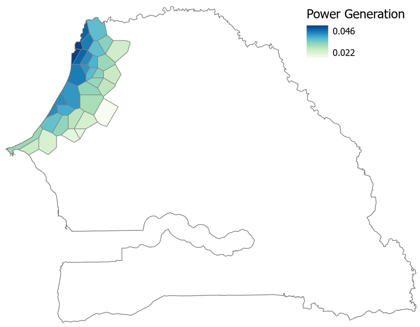


(11) Dist. to Major City

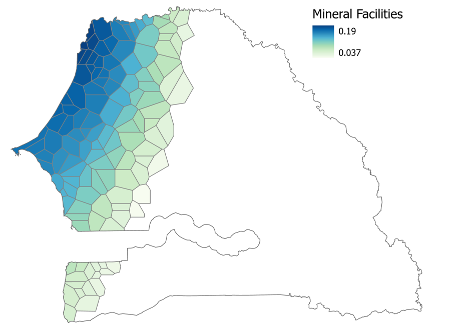


(12) Dist. to Major Mineral Deposit

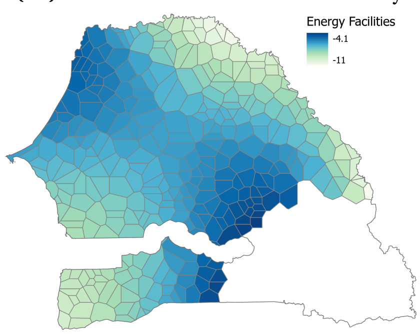
Figure 4.35. Local parameter estimates for the Senegal sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.36.



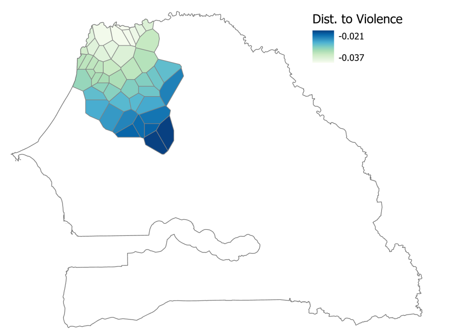
(13) Dist. to Power Generation Facility



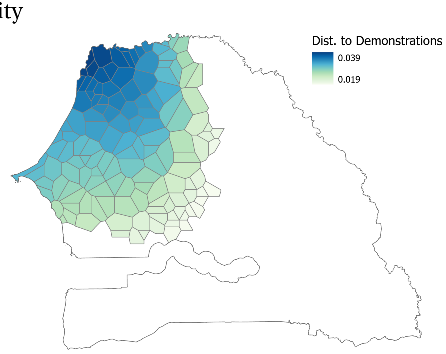
(14) Dist. to Non-Energy Mineral Facility



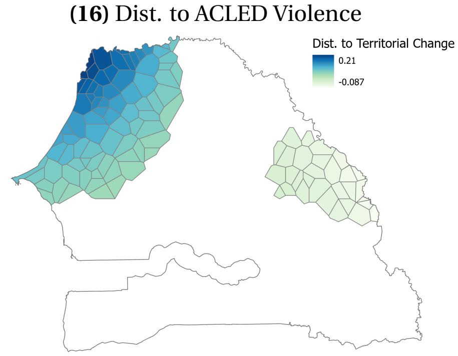
(15) Dist. to Energy Extraction / Refinement Facility



(16) Dist. to ACLED Violence

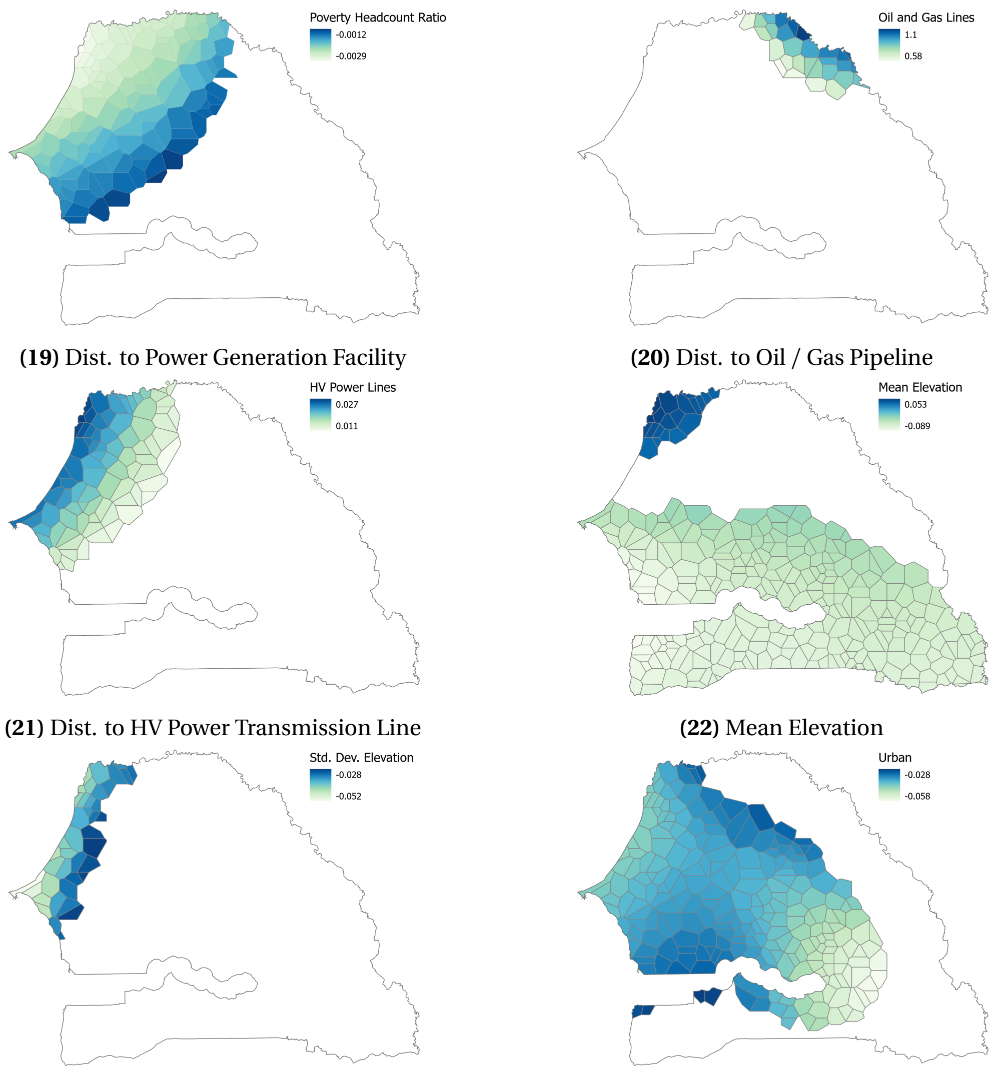


(17) Dist. to ACLED Demonstration



(18) Dist. to ACLED Territorial Change

Figure 4.36. Local parameter estimates for the Senegal sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.37.



(19) Dist. to Power Generation Facility

(20) Dist. to Oil / Gas Pipeline

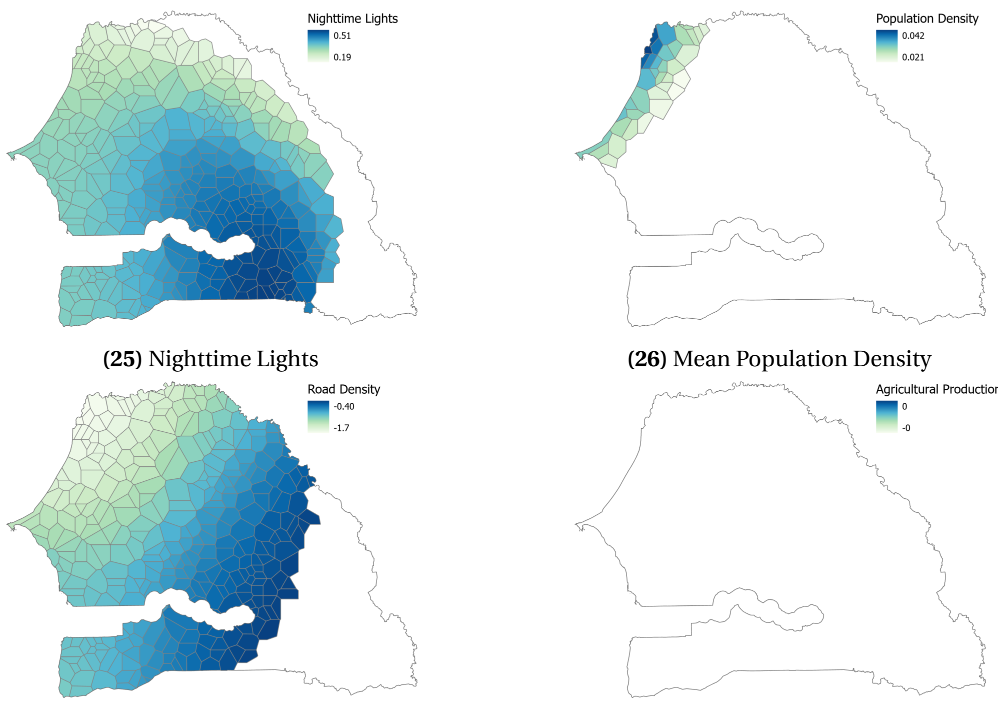
(21) Dist. to HV Power Transmission Line

(22) Mean Elevation

(23) Std. Dev. Elevation

(24) Percent Urban Territory

Figure 4.37. Local parameter estimates for the Senegal sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown. Continued in Figure 4.38.



(25) Nighttime Lights

(26) Mean Population Density

(27) Mean Road Density

(28) Total Agricultural Production

Figure 4.38. Local parameter estimates for the Senegal sample GWR. Only coefficients with a t value significant at the 95% confidence level are shown.

4.9 Works Cited

Anselin, Luc. 1995. “Local Indicators of Spatial Association—LISA.” *Geographical Analysis* 27 (2): 93–115.

Balán, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel. 2022. “Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo.” *American Economic Review* 112 (3): 762–797.

Bell, John E., and Stanley E. Griffis. 2015. “Military Applications of Location Analysis.” In *Applications of Location Analysis,* edited by H. A. Eiselt and Vladimir Marianov, 403–433. International Series in Operations Research & Management Science. Cham: Springer International Publishing.

Besley, Timothy. 1995. “Property Rights and Investment Incentives: Theory and Evidence from Ghana.” *Journal of Political Economy* 103 (5): 903–937.

Besley, Timothy, and Torsten Persson. 2010. “State Capacity, Conflict, and Development.” *Econometrica* 78 (1): 1–34.

Boone, Catherine. 2003. *Political Topographies of the African State: Territorial Authority and Institutional Choice.* Cambridge University Press.

Cappelen, Christoffer, and Jacob Gerner Hariri. 2022. *Tracing the Origins of the Early Modern State: Introducing the Castles Data.* SSRN Scholarly Paper, 4212429, Rochester, NY.

Dimier, Veronique. 2004. “For a New Start: Resettling French Colonial Administrators in the Prefectoral Corps.” *Itinerario* 28 (1): 49–66.

Dormann, Carsten, Jana M. McPherson, Miguel B. Araújo, Roger Bivand, Janine Bolliger, Gudrun Carl, Richard G. Davies, Alexandre Hirzel, Walter Jetz, W. Daniel Kissling, Ingolf Kühn, Ralf Ohlemüller, Pedro R. Peres-Neto, Björn Reineking, Boris Schröder, Frank M. Schurr, and Robert Wilson. 2007. “Methods to Account for Spatial Autocorrelation in the Analysis of Species Distributional Data: A Review.” *Ecography* 30 (5): 609–628.

Dunn, J. E., and L. Duncan. 2000. “Partitioning Mahalanobis D2 to Sharpen GIS Classification.” In *Management Information Systems,* 195–204.

Fearon, James D., and David D. Laitin. 2003. “Ethnicity, Insurgency, and Civil War.” *American Political Science Review* 97 (1): 75–90.

Fergusson, Leopoldo, Horacio Larreguy, and Juan Felipe Riaño. 2022. “Political Competition and State Capacity: Evidence from a Land Allocation Program in Mexico.” *The Economic Journal* 132 (648): 2815–2834.

Friedman, David. 1977. “A Theory of the Size and Shape of Nations.” *Journal of Political Economy* 85 (1): 59–77.

Gordon, Roger, and Wei Li. 2009. “Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation.” *Journal of Public Economics* 93 (7): 855–866.

Hendrix, Cullen S. 2010. “Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict.” *Journal of Peace Research* 47 (3): 273–285.

Henn, Soeren J. 2022. "Complements or Substitutes? How Institutional Arrangements Bind Traditional Authorities and the State in Africa." *American Political Science Review,* 1–20.

Herbst, Jeffrey. 2014. *States and Power in Africa: Comparative Lessons in Authority and Control - Second Edition.* Princeton University Press.

Ishenda, Doris Kokutungisa, and Shi Guoqing. 2019. "Determinants in Relocation of Capital Cities." *Journal of Public Administration and Governance* 9 (4): 200–220.

Kalyvas, Stathis N. 2015. "How Civil Wars Help Explain Organized Crime—and How They Do Not." *Journal of Conflict Resolution* 59 (8): 1517–1540.

Lake, David A. 1992. "Powerful Pacifists: Democratic States and War." *American Political Science Review* 86 (1): 24–37.

Lichstein, Jeremy W., Theodore R. Simons, Susan A. Shriner, and Kathleen E. Franzreb. 2002. "Spatial Autocorrelation and Autoregressive Models in Ecology." *Ecological Monographs* 72 (3): 445–463.

Livingston, Steven, and Gregor Walter-Drop. 2014. *Bits and Atoms: Information and Communication Technology in Areas of Limited Statehood.* Oxford University Press.

Mellander, Charlotta, José Lobo, Kevin Stolarick, and Zara Matheson. 2015. "Night-Time Light Data: A Good Proxy Measure for Economic Activity?" *PLOS ONE* 10 (10): e0139779.

Müller-Crepon, Carl. 2021. "State Reach and Development in Africa since the 1960s: New Data and Analysis." *Political Science Research and Methods,* 1–10.

North, Douglass C. 1982. *Structure and Change in Economic History.* Unknown edition. New York: W. W. Norton & Company.

Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87 (3): 567–576.

Oppel, Steffen, Ana Meirinho, Iván Ramírez, Beth Gardner, Allan F. O'Connell, Peter I. Miller, and Maite Louzao. 2012. "Comparison of Five Modelling Techniques to Predict the Spatial Distribution and Abundance of Seabirds." *Biological Conservation,* Seabirds and Marine Protected Areas Planning, 156:94–104.

Pomeranz, Dina, and José Vila-Belda. 2019. "Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities." *Annual Review of Economics* 11 (1): 755–781.

Poole, Keith T., and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29 (2): 357–384. JSTOR: 2111172.

Potts, Deborah. 1985. "Capital Relocation in Africa: The Case of Lilongwe in Malawi." *The Geographical Journal* 151 (2): 182–196. JSTOR: 633532.

Rachmawati, Rini, Eko Haryono, and Amandita Ainur Rohmah. 2021. "Developing Smart City in the New Capital of Indonesia." In *2021 IEEE International Smart Cities Conference (ISC2),* 1–7.

Rossman, Vadim. 2018. *Capital Cities: Varieties and Patterns of Development and Relocation.* Routledge.

Rotenberry, John T., Kristine L. Preston, and Steven T. Knick. 2006. "Gis-Based Niche Modeling for Mapping Species' Habitat." *Ecology* 87 (6): 1458–1464.

Rubin, Donald B. 1980. “Bias Reduction Using Mahalanobis-Metric Matching.” *Biometrics* 36 (2): 293–298. JSTOR: 2529981.

Schatz, Edward. 2004. “What Capital Cities Say about State and Nation Building.” *Nationalism and Ethnic Politics* 9 (4): 111–140.

Scott, James C. 1999. *Seeing like a State: How Certain Schemes to Improve the Human Condition Have Failed.* 0 edition. New Haven, CT London: Yale University Press.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis.* CRC Press.

Sinha, Parmanand, Andrea E. Gaughan, Forrest R. Stevens, Jeremiah J. Nieves, Alessandro Sorichetta, and Andrew J. Tatem. 2019. “Assessing the Spatial Sensitivity of a Random Forest Model: Application in Gridded Population Modeling.” *Computers, Environment and Urban Systems* 75:132–145.

Spofford, Ainsworth Rand. 1881. *The Founding of Washington City.* J. Murphy & Company.

Stevens, Forrest R., Andrea E. Gaughan, Catherine Linard, and Andrew J. Tatem. 2015. “Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data.” *PLOS ONE* 10 (2): e0107042.

Tao, Ran, Daniel Strandow, Michael Findley, Jean-Claude Thill, and James Walsh. 2016. “A Hybrid Approach to Modeling Territorial Control in Violent Armed Conflicts.” *Transactions in GIS* 20 (3): 413–425.

Tilly. 1992. *Coercion, Capital, and European States, A.D. 990-1990.* Revised edition. Cambridge, MA: Wiley-Blackwell.

Tobler, W. R. 1970. “A Computer Movie Simulating Urban Growth in the Detroit Region.” *Economic Geography* 46 (sup1): 234–240.

Warren, T. Camber. 2014. “Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty.” *International Organization* 68 (1): 111–141.

Weidmann, Nils B, and Sebastian Schutte. 2017. “Using Night Light Emissions for the Prediction of Local Wealth.” *Journal of Peace Research* 54 (2): 125–140.

Chapter 5

Traditional Authorities and the State

Abstract

Despite recent trends in government centralization across the African continent, traditional authorities remain an important institution in African politics and society. In this chapter, I ask *where* support for traditional authorities is most concentrated on the continent. Using geolocated data from Round 8 of the AfroBarometer survey and a novel measure of state incorporation, I find that support for traditional authorities, as well as the influence that traditional authorities wield over the lives of African citizens, is much greater in unincorporated regions than it is in state-incorporated regions. Surprisingly, I find that the same is true of formal authorities as well. These results hold in both a correlational and a quasi-experimental framework.

5.1 Introduction

After several decades of post-independence instability, the late 1980s and early 1990s marked the beginnings of a gradual process of state consolidation in many African countries. During this period, state leaders—often in concert with international financial institutions such as the World Bank and the International Monetary Fund—sought to improve stable and effective governance, and made substantial investments in development programs, infrastructure, and bureaucratic capacity (Bratton and van de Walle 1997; Boone 1997; Kpundeh and Levy 2004, 7–11).[1] These modernization

1. There seems to be some consensus that these programs had a positive effect on Africa's development outcomes (Mosley and Weeks 1993; Bates and Block 2013). By 1995, the average per capita GDP growth rate in sub-Saharan Africa began to steadily increase, poverty rates began to decline sharply, and almost all measures of good governance (e.g., economic freedom, electoral constraints, executive competitiveness) improved significantly (see Fosu (2018) for a summary). Sub-Sharan Africa's average Polity score improved from < -5 in 1990 (autocracy) to > 0 in 1995 (open anocracy).

efforts typically involved a shift in power away from traditional authorities towards more formal, Westphalian governance structures (van der Windt et al. 2019). The incorporation of traditional societies into the modern state, however, often led to the marginalization of traditional authorities and the erosion of their power and influence (Kyed and Buur 2007; Von Trotha 1996). This was particularly true in countries where the state sought to centralize authority and impose a uniform system of governance across the country, as in the case of Guinea, Tanzania, and Mozambique—each of which banned or abolished traditional authorities altogether at various points in their histories (Baldwin 2020). There has, however, been a recent resurgence in the influence of traditional authorities in Africa (Englebert 2002b; Englebert 2002a). Today, institutional plurality, or what Sklar (1993) refers to as "mixed government," is the norm across most of the continent.[2]

One question that has sparked considerable debate in the literature is why these traditional authorities persist in spite of the consolidation trends that we observe over the past three decades (Baldwin 2015).[3] In this chapter, I hope provide a more nuanced perspective on this question. I am less interested in *why* traditional authorities remain an important force in African politics, and more interested in *where* these authorities still hold sway. In particular, I ask whether individuals that reside "outside" of the formal state have a different valuation of traditional authorities than those who reside "inside" of the state. I argue that this valuation is a function of their visibility. In unincorporated regions, where state infrastructure is sparse and formal state capacity is limited, individuals should view African traditional leaders more positively than they do elsewhere. In these areas, traditional authorities fill the "governance gap" left by incomplete consolidation (Koelble and LiPuma 2011)—they execute important social functions such as dispute resolution and service provision, and they often serve as intermediaries between citizens and the formal state (Williams

2. According to Sklar (1993, 86–87): "As in previous (both ancient and modern) epochs of mixed government, African polities today are governed by unified sovereign authorities. However, there are also two separate dimensions of governmental authority, as there were in medieval Europe. These back-to-back domains of authority are readily identifiable as the realm of state sovereignty and the realm of traditional government; both systems effectively govern the same communities of citizen-subject. Although dualistic systems of political authority can be found in other parts of the world, their establishment by combinations of custom and law in Africa is more comprehensive and systematic than elsewhere."

3. The literature clusters around three main explanations. The first is popular demand: Tribal chiefs, headmen, and other traditional institutions are broadly seen as legitimate authorities (Logan 2013), or they provide functions beneficial to the local populace (Baldwin 2015). The second explanation revolves around state weakness—Africa's often dysfunctional states leave a governance deficit filled by historically important institutions with centuries of experience in dispute resolution and public goods provision (LiPuma and Koelble 2009; Dionne 2017, Ch. 6). The third is that they continue to exist because governments want them to, either as symbionts or as agents of the state to which governments who lack sufficient capacity can delegate certain aspects of governance (de Kadt and Larreguy 2018).

2010). In state-incorporated areas, by contrast, formal authorities are more visible to general public. Residents of state-controlled territory are more likely to interact with formal authorities, and basic services are more likely to be provided by agents of the local or national government. The reduced visibility of traditional authorities and the increased salience of formal authorities in these regions should therefore attenuate individuals' valuations of traditional leaders.

It is important to note that this chapter does not directly engage with the question of whether informal and formal authorities are complements or substitutes (seeHenn (2022) for a recent example). As van der Windt et al. (2019) point out, complementarity and substitution cannot be derived from simple correlations between individuals' support for traditional leaders and their support for the state. Instead, van der Windt and his coauthors argue that the appropriate estimand is not a correlation, but rather a type of constant elasticity of substitution between the various types of authorities that exist in a region. Unfortunately, the survey data I employ in this chapter are unsuitable for the type of analysis that van der Windt et al. (2019) envision. Nevertheless, it is still possible to learn a great deal about individual attitudes towards traditional and formal authorities, and how they vary in response to an individual's social and political geography.

The chapter proceeds as follows. The subsequent section outlines my theory and hypotheses. Section 5.3 provides a descriptive analysis of where traditional authorities are still seen as relevant and influential by the local populace. Section 5.4 provides correlational tests of the hypotheses detailed in Section 5.2, as well as a quasi-experimental analysis that seeks to identify the effect of state-incorporation on support for traditional authorities. A final section summarizes my results.

5.2 Theoretical Expectations

Where in Africa is support for traditional authorities concentrated? I argue that support for traditional authorities is derived primarily from their salience to the lives of the individuals over which they govern. In regions of state that are removed from centers of government power, traditional authorities take on increased visibility in the local community. In the absence of the state, individuals will turn to these traditional authorities for services ranging security and dispute resolution to short-term loans and agricultural insurance. This suggests the following hypothesis:

Hypothesis 1: *The perceived legitimacy and influence of traditional authorities will be stronger in unincorporated regions of the state than in unincorporated regions.*

In state-incorporated areas, by contrast, most governance functions are executed by elected officials conducting constituent services, or by members of the formal bureaucracy—the government employees such as postal workers and police officers that we encounter in our daily lives. In state-incorporated areas, then, the central government effectively substitutes for the services provided by traditional authorities in unincorporated areas, thereby increasing the salience of state authorities to residents of state-incorporated areas. This suggests Hypothesis 2:

Hypothesis 2: *The perceived legitimacy of state authorities will be stronger in incorporated regions of the state than in incorporated regions.*

To test Hypothesis 1, I create two outcome variables based on cross-sectional data collected in the most recent round of the AfroBarometer Survey (Round 8). The first outcome, the *TA Perceptions Index*, combines eight survey items into a single composite index, similar to those developed by Logan (2009) and Henn (2022). These items, detailed in Table 5.1, include several perceptual questions that gauge attitudes such as the respondent's level of trust in a traditional authority and the ideal amount of influence a traditional authority should hold in a local community. The index also includes a single behavioral measure—the frequency with which the respondent reports contact with a traditional authority in the past year. Responses are recoded such that higher values indicate more positive perceptions of traditional leaders (See Appendix for recoding procedures). The second outcome—the *TA Influence Index*—is composed of a battery of questions unique to AfroBarometer Round 8, which ask respondents how much influence traditional authorities hold in certain policy domains, including governing the local community, allocating land, influencing voting behavior, and resolving disputes. To test Hypothesis 2, I create two additional indices based on components of the *TA Perceptions Index*. The *LC Perceptions Index* includes the first five items in the *TA Perceptions Index*, asked about local councilors, and the the *MP Perceptions Index* includes the same five items, asked about the respondent's member of parliament (MP) or representative in the national legislature. The remaining three items of the *TA Perceptions Index* are not included in the formal authorities indices, as respondents were not asked evaluate local councilors or MPs on these characteristics.

Table 5.1. Summary Statistics for Outcome Indices and their Components

Variable	n	μ	σ	Min	Max
Traditional Authorities Perception Index					
Contact with Traditional Authority, Past Year	42,800	0.833	1.149	0	3
Traditional Authority Listens To People Like Me	40,978	1.311	1.125	0	3
Trust in Traditional Authority	41,081	1.871	1.095	0	3
Traditional Authority Does Not Engage in Corruption	39,263	1.999	0.850	0	3
Approve of Traditional Authority Performance	37,483	1.786	0.920	0	3
Ideal Traditional Authority Influence in Local Community	40,931	2.650	1.157	0	4
Traditional Authority Serves Community Interests	41,391	0.567	0.496	0	1
Traditional Authority Should Advise on Voting	41,785	0.289	0.453	0	1
TA Perceptions Index	43,721	52.074	21.020	0	100
Traditional Authorities Influence Index					
Traditional Authority Influence: Governing Local Community	40,815	1.825	1.084	0	3
Traditional Authority Influence: Allocating Land	40,696	1.646	1.164	0	3
Traditional Authority Influence: Voting Behavior	40,480	1.262	1.150	0	3
Traditional Authority Influence: Dispute Resolution	41,222	2.069	1.062	0	3
TA Influence Index	41,667	56.719	28.112	0	100
Local Councilor Perceptions Index					
Contact with Local Councilor, Past Year	42,493	0.490	0.903	0	3
Local Councilor Listens To People Like Me	42,199	0.760	0.899	0	3
Trust in Local Councilor	41,769	1.397	1.083	0	3
Local Councilor Does Not Engage in Corruption	39,695	1.665	0.841	0	3
Approve of Local Councilor Performance	39,515	1.391	0.922	0	3
LC Perceptions Index	43,855	37.192	19.857	0	100
Parliament Perceptions Index					
Contact with MP, Past Year	45,917	0.222	0.639	0	3
MP Listens To People Like Me	44,770	0.571	0.806	0	3
Trust in MP	44,516	1.334	1.094	0	3
MP Does Not Engage in Corruption	41,168	1.578	0.876	0	3
Approve of MP Performance	43,069	1.294	0.918	0	3
MP Perceptions Index	46,218	32.440	18.142	0	100

All four indices are constructed by summing the respondent's numerical responses (e.g., point value) for each survey item and dividing this sum by the total points possible (based the maximum value of each survey item and the total number of items the respondent completed) to obtain a single score that can range from 0 to 100, where higher values indicate more favorable perceptions, or, in the case of the *TA Influence Index*, greater levels of influence. Principal component analysis suggests that the components of each of these indices represent more than one dimension of support for a given authority (see Appendix). Nevertheless, all results presented in the main text replicate when substituting the first principal component of each index in place of the index itself.[4]

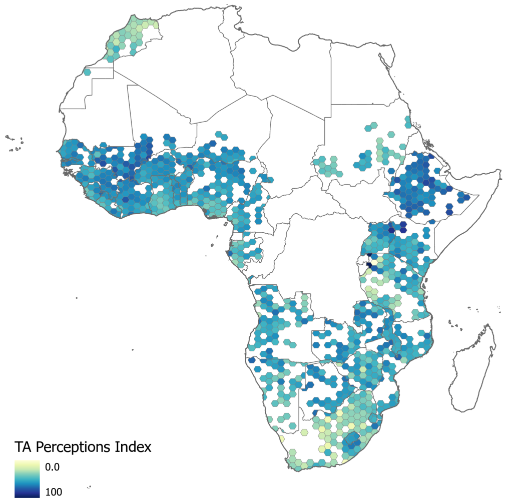
5.3 Spatial Distribution of Support for Traditional Authorities

Because the GPS coordinates of each enumeration area are reported by AfroBarometer, it is possible to explore spatial patterns of support for traditional and formal authorities across Africa.[5] In Figure 5.1, I plot the average index value for each of the four outcome indices within a 10,000 km^2 hexagonal grid cell.

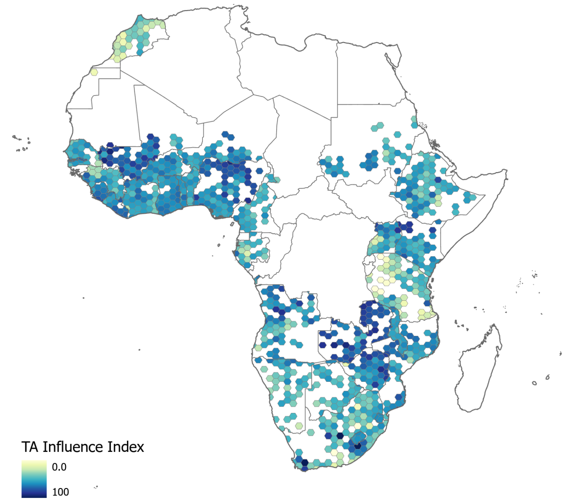
Visual inspection of the maps in Figure 5.1 reveal a great deal of spatial clustering; respondents that have relatively positive perceptions of a given authority are likely to live near respondents with similar perceptions of that authority. The same is true of respondents with comparatively negative perceptions. I test whether or not this spatial clustering is statistically significant by calculating Moran's I for each index. Moran's I is a widely-used measure of spatial autocorrelation. Values range from -1 to 1, where negative values represent dispersion and positive values represent clustering. Results in Table 5.2 indicate that all four indices tend to cluster in space, which suggests that support for both types of authorities is geographically determined. Above, I argue that the relevant geographical factor driving these patterns is the incorporation status of a given region. In the following section, I test this proposition empirically.

4. Results presented in the main text turn out to be more conservative than the estimates that use the first principal component as an outcome variable.

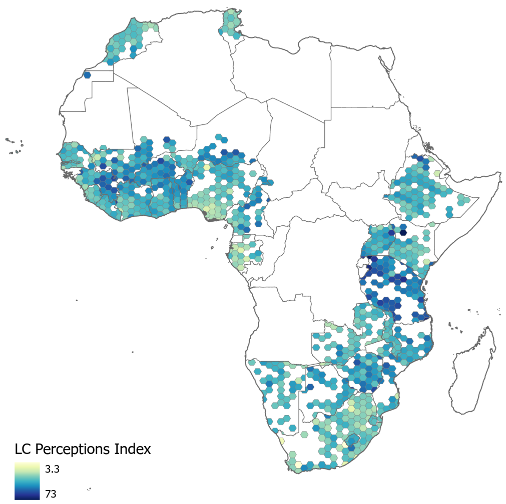
5. To protect respondents' privacy, AfroBarometer introduces small random variations into GPS coordinates. Deviations are generally within ±10 kilometers, though these deviations are of sufficient magnitude in two instances to push the location of an enumeration area across an international border. These minor deviations do not affect the analyses presented in this chapter.



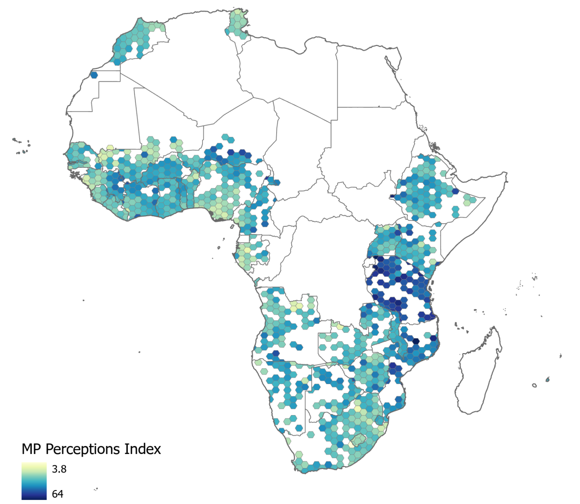
(a) TA Perceptions Index



(b) TA Influence Index



(c) LC Perceptions Index



(d) MP Perceptions Index

Figure 5.1. Spatial distribution of the four main outcome variables, averaged within a 10,000 km^2 hexagonal grid cell. Even at this relatively low level of resolution, there are obvious patterns of spatial clustering among respondents with high index values and those with low index values.

Table 5.2. Spatial Autocorrelation By Index

Index	Moran's I	z Score	p Value
TA Perceptions Index	0.437	104.791	0.01
TA Influence Index	0.320	76.341	0.01
LC Perceptions Index	0.290	73.867	0.01
MP Perceptions Index	0.379	94.842	0.01

5.4 Perceptions, Influence, and State Incorporation

Is there any systematic difference in how the public views traditional authorities in state-incorporated areas versus unincorporated areas? In this section, I outline the results of two separate analyses, each of which show that traditional authorities are perceived more positively by, and retain more influence over, residents of unincorporated territory. Contrary to expectations, however, I also find disconfirmatory evidence for Hypothesis 2: Residents of state-incorporated areas tend have more *negative* perceptions of state-affiliated authorities than residents of unincorporated areas.

5.4.1 Traditional Authorities and State Incorporation

As a preliminary test of Hypothesis 1, I estimate the following OLS model:

$$y_i = \alpha + \beta(\text{State})_i + \gamma Z_i + \epsilon_i, \tag{5.1}$$

where y is either the respondent's *TA Perceptions Index* or their *TA Influence Index* value, State is a binary "treatment" variable that takes a value of 1 if control point density (CPD—the measure of territorial control detailed in Chapter 2) at the respondent's location is greater than the mean control point density for the respondent's home country, and Z is vector of demographic controls (see Table 5.7 in the Appendix for summary statistics). I estimate Model 5.1 using sampling weights; standard errors are adjusted to account for AfroBarometer's complex design.

Figure 5.2 plots the results of Model 5.1 (the full set of results are available in Table 5.10 in the Appendix). The coefficient estimate for the State variable is negative and significant for both outcomes. These findings confirm Hypothesis 1—respondents' perceptions of traditional authorities, and the influence that these authorities wield in their respective locales, is higher in unincorporated areas than in state-incorporated areas. To illustrate these differences, Figure 5.3 plots the predicted values of both the *TA Perceptions Index* and the *TA Influence Index* in incorporated and unincorporated regions. The average marginal effect of treatment on the *TA Perceptions Index* is -2.984, which represents a roughly 3 percentage point decline in respondents' valuation of traditional authorities when moving from an unincorporated area to a state-incorporated area. The effect of state-incorporation is less pronounced for the *TA Influence Index*. The average marginal effect is

-1.017; this difference is significant at the 90% confidence level ($p = 0.055$). The comparatively weak effect of state-incorporation on the *TA Influence Index* is not necessarily surprising. While Hypothesis 1 anticipates that traditional authorities will have greater influence over unincorporated populations, domains such as local governance, land allocation, and dispute resolution (three components of the index) increasingly fall within the purview of formal governments, even in Africa's hinterlands. It is not unreasonable, then, to hypothesize that the gap between traditional authorities' influence in unincorporated and incorporated areas has narrowed over the past two decades, as the state has taken on a more expansive role in the lives of many citizens. Unfortunately, it is not possible to test this proposition empirically, as this specific battery has not be asked in previous iterations of AfroBarometer.

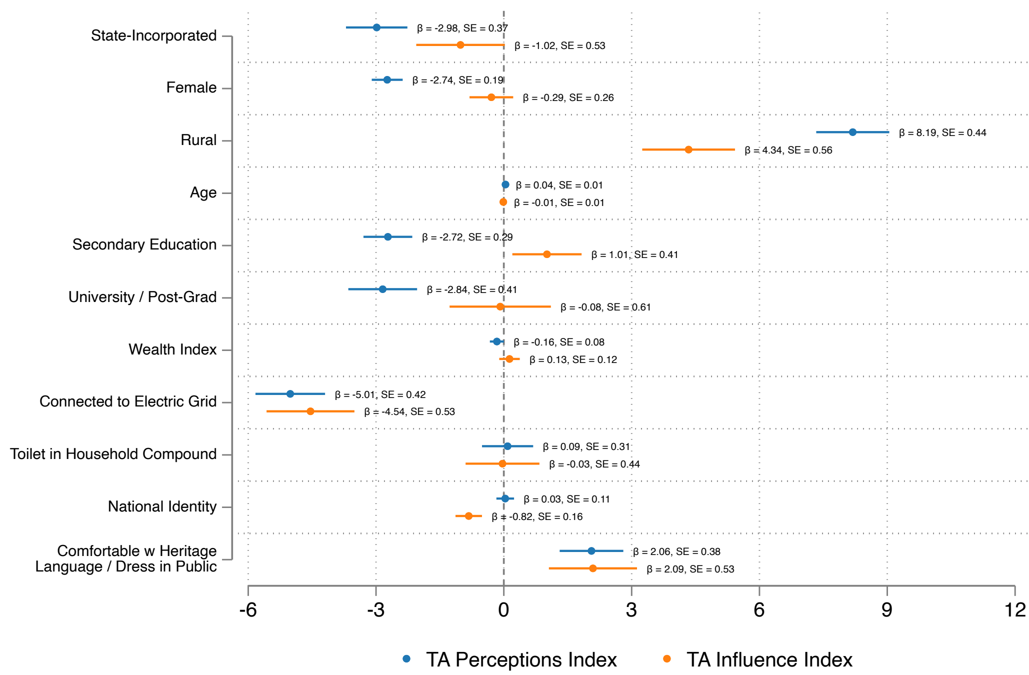


Figure 5.2. Results of Model 5.1 showing a significant negative correlation between state-incorporation and the two informal authorities outcome variables. Marker labels contain coefficient estimates and standard errors.

The remainder of the results from Model 5.1 are also broadly consistent with theoretical expectations. Female respondents are less likely to report favorable perceptions of traditional authorities than male respondents. This is likely driven by patriarchal structure of certain traditional

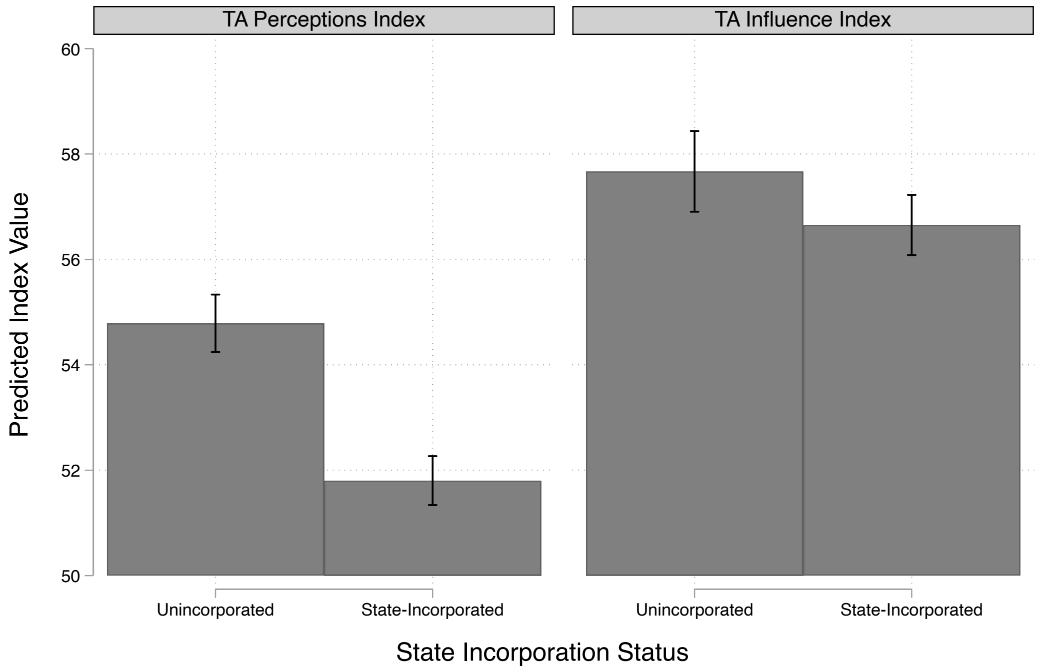


Figure 5.3. Predicted *TA Perceptions Index* and *TA Influence Index* values and 95% confidence intervals in state-incorporated and unincorporated territory.

societies in Africa and the discriminatory attitudes they espouse, which have contributed to host of negative outcomes for women, including high rates of intimate partner violence and widespread exclusion from politics (Jayachandran 2015, Alangea et al. 2018, Sikweyiya et al. 2020, Robinson and Gottlieb 2021). The same is true of highly educated respondents; the negative correlation between educational attainment and perceptions of traditional authorities mirrors findings from Logan (2013, 369). Although neither the respondent's age nor wealth have any significant effect on either of the two outcomes, I do uncover a significant negative correlation between access to the electrical grid and the two outcome variables, as well as a significant positive correlation between the respondent's comfort with traditional language and clothing and the two outcome variables. The single best predictor of both indices, however, is the binary rural-urban classification. Rural respondents are significantly more likely to report positive perceptions of traditional authorities, and traditional authorities are significantly more likely to influence rural communities than urban communities. This finding lends some empirical support to arguments made by Mamdani (1996) that traditional governance in Africa is a primarily rural phenomenon.[6]

5.4.2 Regression Discontinuity Results

Hypothesis 1 implies that "treated" respondents—those whose control point density (CPD) scores exceed the country mean—should view traditional authorities more negatively than "untreated" respondents, or those whose CPD scores are below the country mean. Although CPD scores are not randomly assigned, Henn (2022) demonstrates the feasibility of estimating the causal effect of state-incorporation on respondents' perceptions of traditional authorities by using a geographic regression discontinuity design (RDD). In this analysis, I employ a similar design to confirm the results from Section 5.4.1, and to take some preliminary steps towards causal identification.

The RDD requires a continuous running variable and a known cutoff, used to discriminate between treated and untreated observations, to estimate a local average treatment effect (LATE) in the vicinity of the cutoff. Identification relies on three core assumptions. The first is that treatment is at least partially determined by the running variable, and that in the absence of treatment, $E[y_i^0|x_i = x_0]$ and $E[y_i^1|x_i = x_0]$ are continuous functions of x across the cutoff (i.e., without treatment, we would not observe the discontinuity). Second, respondents are unable to manipulate the running variable with any real precision, to effectively self-select into treatment or control. While it is certainly possible that individuals may "vote with their feet" and relocate within their home country to neighborhoods with either higher or lower levels of state control depending on their preferences, settlement patterns in many African countries tend to be sticky (Laver 1976, Herbst 1990). Residence is more likely to be determined by economic considerations, climate factors, and social ties rather than proximity to the state (Garcia et al. 2015, Wesolowski et al. 2015). The final assumption is that predetermined or exogenous covariates should be balanced at the cutoff. To test this assumption, I estimate the effect of the discontinuity on a number of plausibly exogenous covariates. Results in Table 5.3 suggests that there are very few significant differences between respondents on one side of the cutoff and respondents on the other.

This RDD setup uses (log) CPD as the running variable, and the country mean as the cutoff,

6. It is worth noting that a similar analysis by Logan (2013) using AfroBarometer Round 4 data shows no statistically significant correlation between the rural-urban indicator and the respondent's desire to expand the role of traditional authorities.

Table 5.3. Balance on Exogenous Covariates

AfroBarometer Item	Variable	RD Effect	Robust p
Q101	Female	-0.00044	0.977
Q14	Age	1.43760	0.022**
Q97	Education	0.05416	0.076*
Q3	Overall Direction of the Country	-0.00405	0.827
Q4A	Economic Condition of the Country	-0.01967	0.631
Q4B	Present Living Conditions	-0.04543	0.303
Q14	Freeness and Fairness of Last Election	0.07306	0.237
Q36	Extent of Democracy in Country	0.13532	0.211
Q47B	Fairness of High Taxes for the Rich to Help Poor	0.07320	0.207
Q59F	Effects of Social Media on Society	-0.06809	0.252
Q55E	Gets News from Social Media	-0.00392	0.956
Q65C	Government Borrows too Much from China	-0.00019	0.994

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

c. For ease of interpretation, I transform CPD to a z score, such that $c = 0$. My baseline specification is:

$$y_i = \alpha + \beta_1 \tau_i + \beta_2 x_i + \beta_3 (x_i \cdot \tau_i) + \gamma Z_i + \epsilon_i, \tag{5.2}$$

where y is either the *TA Perceptions Index* or the *TA Influence Index*, x is the centered running variable, and τ is a binary treatment indicator that takes a value of 1 if $x_i \geq 0$ and 0 otherwise. The coefficient of interest is β_1, which represents the treatment effect at the cutoff. Although the RDD assumes that covariates are balanced at the cutoff, I do include a vector of demographic controls, Z; these are the same set of controls used in Model 5.4.1 (see Table 5.7). Figures 5.4 and 5.5 show the discontinuity for the full sample. For estimation purposes, I restrict the sample to respondents within $\frac{1}{10}$ of a standard deviation from the cutoff. Figure 5.12 in the Appendix provides RD estimates for the *TA Perceptions Index* using a range of bandwidths to confirm that results are not driven by the choice of bandwidth. As a robustness check, I also calculate the RD effect using a non-parametric specification:

$$y_i = m(x_i) + \epsilon_i, \tag{5.3}$$

where $m(x_0) = E[y_i | x_i = x_0]$, and $x_0 \in x_i$. I estimate $m(x_i)$ using a kernel-weighted local polynomial (quadratic) model (Gelman and Imbens 2019, Calonico et al. 2014). This model does not include any controls, though results are robust to their inclusion (see Table 5.12 in the Appendix).

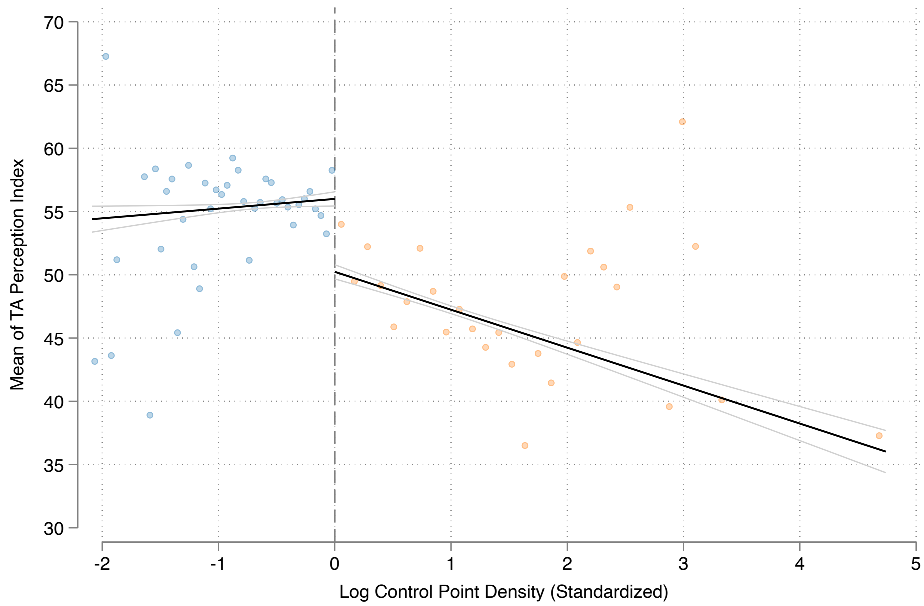


Figure 5.4. Binned scatter plot of log control point density (standardized) and the *TA Perceptions Index* using the full sample. Global least squares estimates (without controls) and 95% confidence intervals shown.

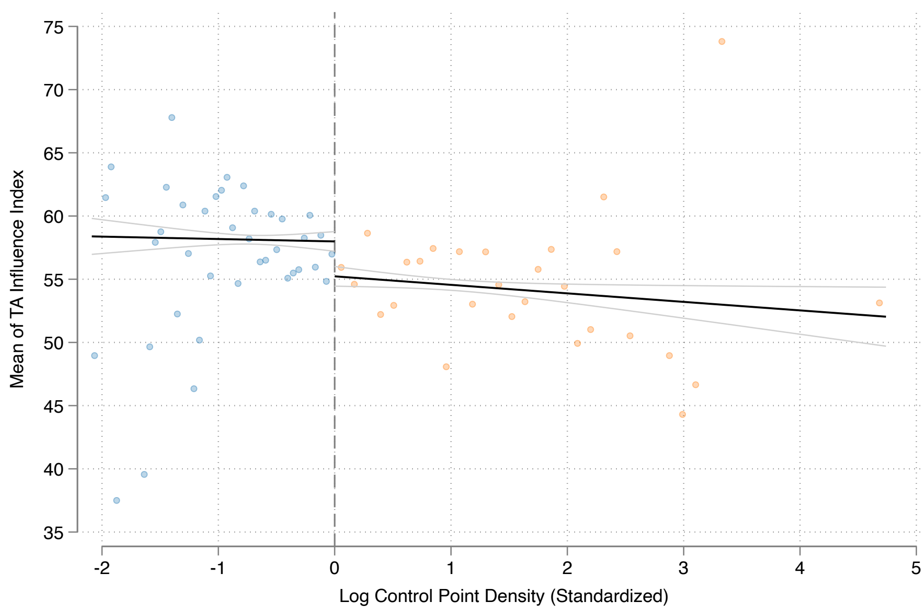


Figure 5.5. Binned scatter plot of log control point density (standardized) and the *TA Influence Index* using the full sample. Global least squares estimates (without controls) and 95% confidence intervals shown.

Table 5.4 provides the full set of results from Models 5.2 and 5.3. Columns (1) through (3) estimate the RD effect on the *TA Perceptions Index*, and the Columns (4) through (6) estimate the RD effect on the *TA Influence Index*. Results are consistent with OLS estimates in Figure 5.2. There is a significant, negative treatment effect of about 5.4% for the *TA Perceptions Index* with the full OLS specification. The magnitude of the effect is smaller using the local polynomial model, but remains significant at the 90% confidence level. These results are a strong indication that state-incorporation does, in fact, dampen respondents' perceptions of traditional authorities. I do not uncover a similar effect on the influence of traditional authorities. Although the RD point estimates are negative, they do not attain statistical significance in any specification. As I note in Section 5.4.1, I attribute this null effect to the increasingly expansive role of the state in contemporary African politics. While traditional authorities still engender respect from the local populace, their governance functions may be crowded out by formal institutions such as local councils, state courts, and land registries.

5.4.3 Formal Authorities and State Incorporation

The analyses thus far have focused on the perceptions and influence of traditional authorities in incorporated and unincorporated regions. In this section, I evaluate respondents' perceptions of *formal* authorities, or those associated with the state or local government. If the theory outlined in Section 5.2 is correct—that individuals will view the most salient authority in their region in a more positive light—we should expect to see that the perceived legitimacy of state authorities will be stronger in incorporated regions of the state than in unincorporated regions (i.e., Hypothesis 2). To test this hypothesis, I reestimate Model 5.1 using the *LC Perceptions Index* and the *MP Perceptions Index* as the outcome variables.

Results are visualized in Figure 5.6; Table 5.13 in the Appendix contains the full set of results. Coefficient estimates for the state-incorporation variable are negative and significant in both models, this is the opposite of what is predicted by Hypothesis 2. In fact, the magnitudes of the differences in index values between incorporated and unincorporated areas are on par with those I estimate for the two traditional authorities indices above. Figure 5.7 plots the predicted value of each state authority index by state incorporation status. The average marginal effect of state-incorporation on respondents' perceptions of local councilors is -2.166. The average marginal effect of state-

Table 5.4. RDD Results

	TA Perceptions Index			*TA Influence Index*		
	OLS		Loc. Poly.	OLS		Loc. Poly.
	(1)	(2)	(3)	(4)	(5)	(6)
RD Estimate	-6.141***	-5.360***	-1.836*	-2.873	-1.225	-1.538
	(1.641)	(1.655)	(1.003)	(2.274)	(2.338)	(1.195)
Constant	58.68***	55.00***		58.26***	56.17***	
	(1.120)	(2.457)		(1.554)	(3.471)	
Observations	2,304	2,092	43,453	2,230	2,036	41,415
Controls	No	Yes	No	No	Yes	No
Bandwidth	0.1	0.1		0.1	0.1	
R^2	0.007	0.058		0.002	0.021	
Kernel Type			Triangular			Triangular
Robust 95% CI			[-4.301 ; -.079]			[-4.013 ; 1.31]
Conventional p			0.0672			0.198
Robust p			0.0420			0.320
Order Loc. Poly. (p)			2			2
Order Bias (q)			3			3
BW Loc. Poly. (h)			0.534			0.792
BW Bias (b)			0.860			1.046

Standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

incorporation for MPs is slightly smaller at -1.501, though this difference remains significant at the 99% confidence level.

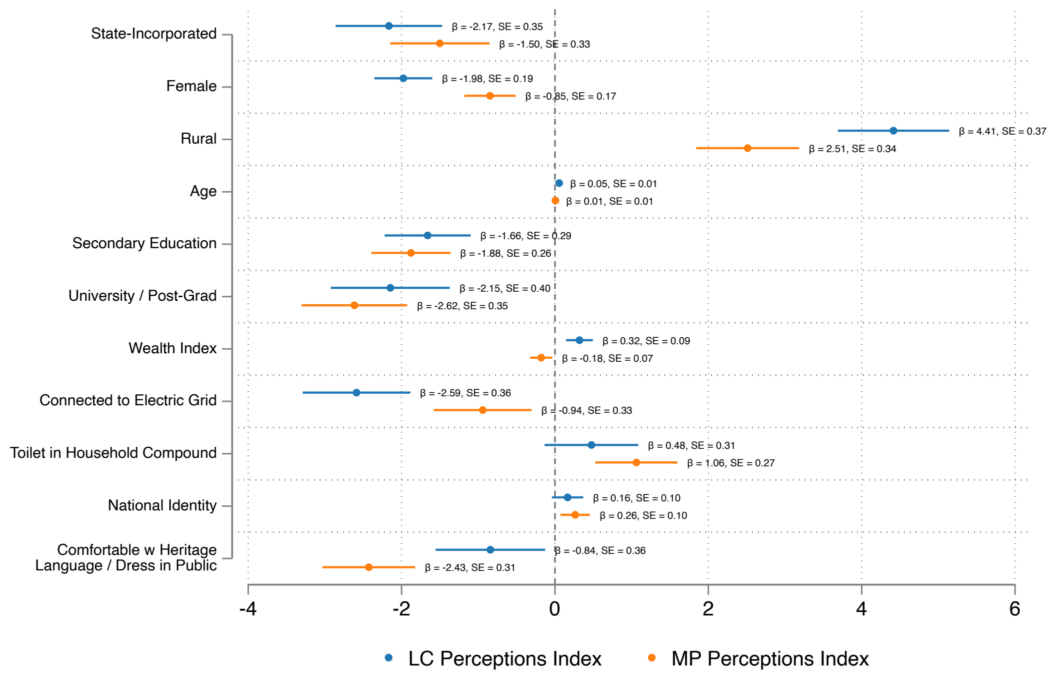


Figure 5.6. Results of Model 5.1 showing a significant negative correlation between state-incorporation and the two formal authorities outcome variables. Marker labels contain coefficient estimates and standard errors.

Although this analysis presents evidence inconsistent with Hypothesis 2, my findings do mirror results from Logan (2009, 112), who uncovers a positive correlation between support for traditional leaders and support for elected leaders using data from AfroBarometer Rounds 1 and 2.[7] Logan's analysis, however, does not disaggregate by state-incorporation status.

Two other sets of results from these models are difficult to reconcile with theoretical priors. First, rural respondents are significantly more likely to hold positive perceptions of their local councilors and MPs than urban respondents. One potential explanation for this result is that these positive perceptions are a function of credit claiming. In rural areas, where infrastructure is sparse and public goods provision is comparatively weak, residents may be more likely to attribute any facilities and services that do exist to their elected representatives. These representatives, in turn, are

7. Although our indices differ slightly based on available survey items, I estimate the same positive correlation between the *TA Perceptions Index* and the two formal authorities indices in Table 5.6. Correlations range from 28% for MPs to to 36% for local councilors.

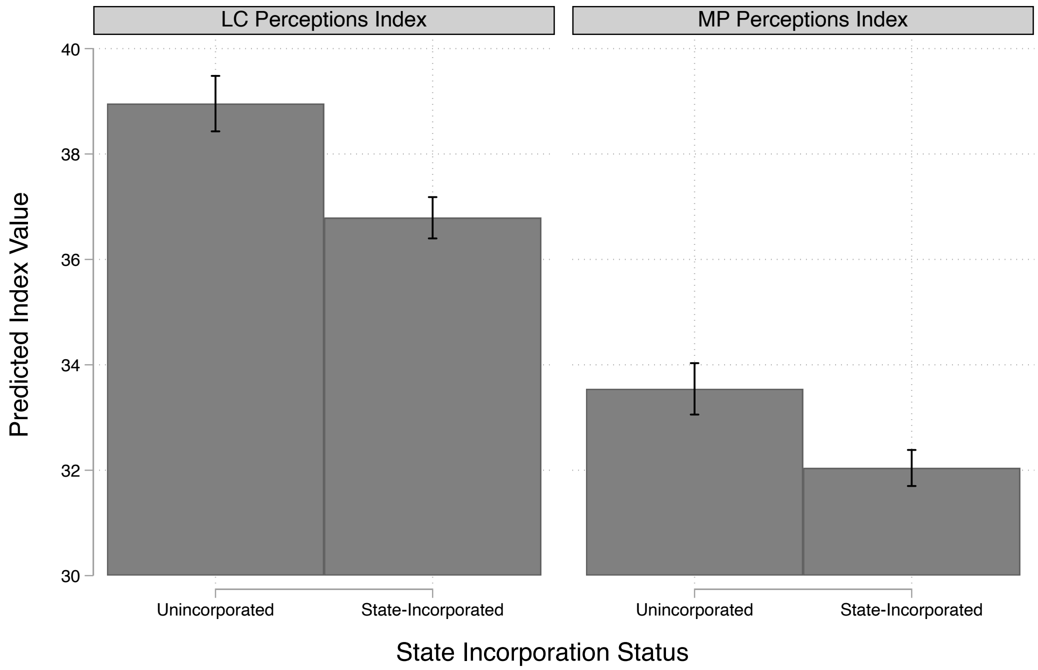


Figure 5.7. Predicted *LC Perceptions Index* and *MP Influence Index* values and 95% confidence intervals in state-incorporated and unincorporated territory.

happy to claim credit for programmatic benefits (Koter 2016). Alternatively, this result may be driven by clientelism. Rural AfroBarometer respondents are 3.82% more likely than urban respondents to report being offered money or a gift in exchange for a vote in the most recent election.[8] These non-programmatic benefits may buy elected officials sustained support outside of the election cycle. The second surprising result is the significant negative correlation between educational attainment and the two formal authorities indices. It is possible that highly educated respondents are more savvy observers of politics, and thus more perceptive of government corruption than respondents with less education. This, in turn, attenuates their valuations of formal authorities. AfroBarometer data provide some support for this explanation: 6.29% of respondents with a post-secondary education report that their local councilor does not engage in corruption, compared to 17.75% of respondents with a primary education or less—this is a difference of almost 11.5 percentage points.[9] A similar difference exists between high-education and low-education respondents in their respective

8. A t-test indicates that this difference is statistically significant at the 99% confidence level; Mean Difference = 0.038, SE = 0.004.
9. Note that this corruption variable is a component of all three (TA, LC, MP) perceptions indices.

assessments of corruption among both MPs ($d = 11.39\%$) and traditional authorities ($d = 18.03\%$).

The remaining results are all consistent with expectations. Although the estimated coefficients are fairly small, respondents who identify more with their national identity have more positive perceptions of formal authorities than those who identify with their tribal or ethnic group. Conversely, respondents who report being comfortable with their heritage language and dress in public spaces have more negative perceptions of formal authorities. It is worth noting that the coefficients on these two variables in this analysis have the opposite sign as those estimated in Section 5.4.1, which suggests that *identity* is an important factor in determining support for one type of authority over another.

5.5 Discussion

Results in Section 5.4 show that support for both traditional and formal authorities is significantly lower in state-incorporated regions of the continent than in unincorporated regions. It is difficult to ascertain from available AfroBarometer data why these results are inconsistent with Hypothesis 2, though it is possible that the *quality* of governance offered by state authorities is low compared to that offered by traditional authorities, even in state-incorporated areas. Many of the governments in the AfroBarometer sample are plagued by corruption, graft, and other performance issues, so informal authorities may be viewed in comparatively positive light, even in areas where most public goods are provisioned by the state.

5.6 Appendix

Primary Outcome Variables

The following eight questions from AfroBarometer Round 8 comprise the *TA Perceptions Index*. Five of the eight variables (marked with an asterisk) comprise the *LC Perceptions Index* and the *MP Perceptions Index*. Recoding rules are provided in the bullet points. The binary recode only applies to the *TA Perceptions Index*.

1. *Contact with {TA, LC, MP}, Past Year**

   {Q12D, Q12A, Q12B}: During the past year, how often have you contacted any of the following persons about some important problem or to give them your views: A traditional leader / A local government councilor / A Deputy in the National Assembly?

   (a) Recode
   - Substantive response values retained
   - All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

   (b) Binary recode
   - {0 Never, 1 Only once} $\Longrightarrow$ 0 Infrequent Contact
   - {2 A few times, 3 Often} $\Longrightarrow$ 1 Frequent Contact
   - All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

2. *{TA, LC, MP} Listens To People Like Me**

   {Q38C, Q38B, Q38A}: How much of the time do you think the following try their best to listen to what people like you have to say: Traditional leaders / Municipal or communal councilors / Deputies of the National Assembly?

   (a) Recode
   - Substantive response values retained
   - All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

   (b) Binary recode
   - { 0 Never, 1 Only sometimes} $\Longrightarrow$ 0 Seldom Listens
   - {2 Often, 3 Always} $\Longrightarrow$ 1 Often Listens
   - All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

3. *Trust in {TA, LC, MP}**

   {Q41K, Q41D, Q41B}: How much do you trust each of the following, or haven't you heard enough about them to say: Traditional leaders / Your commune council / National Assembly?

(a) Recode

- Substantive response values retained
- All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

(b) Binary recode

- {0 Not at all, 1 Just a little} $\Longrightarrow$ 0 Low Trust
- {2 Somewhat, 3 A lot} $\Longrightarrow$ 1 High Trust
- All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

4. *{TA, LC, MP} Does Not Engage in Corruption**

{Q42H, Q42D, Q42B}: How many of the following people do you think are involved in corruption, or haven't you heard enough about them to say: Traditional leaders / Local Government Councilor / MP or National Assembly Rep?

(a) Recode

- 0 None $\Longrightarrow$ 3
- 1 Some of them $\Longrightarrow$ 2
- 2 Most of them $\Longrightarrow$ 1
- 3 All of them $\Longrightarrow$ 0
- All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

(b) Binary recode

- 0 None $\Longrightarrow$ 1 No TA Corruption
- {1 Some of them, 2 Most of them, 3 All of them} $\Longrightarrow$ 0 TA Corruption
- All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

5. *Approve of {TA, LC, MP} Performance**

{Q51D, Q51C, Q51B}: Do you approve or disapprove of the way the following people have performed their jobs over the past twelve months, or haven't you heard enough about them to say: Your traditional leader / Local Government Councilor / MP or National Assembly Rep?

(a) Recode

- 1 Strongly disapprove $\Longrightarrow$ 0 Strongly Disapprove
- 2 Disapprove $\Longrightarrow$ 1 Disapprove
- 3 Approve $\Longrightarrow$ 2 Approve
- 4 Strongly approve $\Longrightarrow$ 3 Strongly Approve
- All other responses {7 Not applicable, 8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

(b) Binary recode

- {1 Strongly disapprove, 2 Disapprove} $\Longrightarrow$ 0 Disapprove
- {3 Approve, 4 Strongly approve} $\Longrightarrow$ 1 Approve

- All other responses {7 Not applicable, 8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

6. *Ideal TA Influence in Local Community*

   Q87E: Do you think that the amount of influence traditional leaders have in governing your local community should increase, stay the same, or decrease?

   (a) Recode
      - 1 Decrease a lot $\Longrightarrow$ 0 Decrease A Lot
      - 2 Decrease somewhat $\Longrightarrow$ 1 Decrease Somewhat
      - 3 Stay the same $\Longrightarrow$ 2 Stay the Same
      - 4 Increase somewhat $\Longrightarrow$ 3 Increase Somewhat
      - 5 Increase a lot $\Longrightarrow$ 4 Increase A Lot
      - All other responses {7 Not applicable, 8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

   (b) Binary recode
      - {1 Decrease a lot, 2 Decrease somewhat} $\Longrightarrow$ 0 Decrease
      - {3 Stay the same, 4 Increase somewhat, 5 Increase a lot} $\Longrightarrow$ 1 Same or Increase
      - All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

7. *TA Serve Community Interests*

   Q88: Which of these statements is closest to your own opinion?
   Statement 1: Traditional leaders mostly look out for what is best for the people in their communities.
   Statement 2: Traditional leaders mostly serve the interests of politicians and government officials.
   Statement 3: Traditional leaders mostly look out for their own personal interests.

   (a) Recode same as binary

   (b) Binary recode
      - {0 Do not agree with any of these statements, 1 Statement 3, 2 Statement 2} $\Longrightarrow$ 0 Serve Own / Gov't Interests
      - 3 Statement 1 $\Longrightarrow$ 1 Serve Community Interests
      - All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

8. *TA Should Advise on Voting*

   Q89A: Which of the following statements is closest to your view? Choose Statement 1 or Statement 2.
   Statement 1: Traditional leaders have a better grasp of political issues than ordinary people; they should give their people advice about how to vote.
   Statement 2: Traditional leaders should stay out of politics and leave people to make their own decisions about how to vote.

(a) Recode same as binary

(b) Binary recode

- {1 Agree very strongly with Statement 1, 2 Agree with Statement 1} $\Longrightarrow$ 1 TA Should Advise Voting
- {3 Agree with Statement 2, 4 Agree very strongly with Statement 2, 5 Agree with neither} $\Longrightarrow$ 0 TA Should Stay Out of Politics
- All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

Table 5.5. Traditional Authority TA Perceptions Index (Using Binary Component Variables)

Variable	n	μ	σ	Min	Max
Traditional Authorities Perception Index (Binary Recode)					
Contact with Traditional Authority, Past Year	42,800	0.3897	0.4877	0	1
Traditional Authority Listens To People Like Me	40,978	0.4286	0.4949	0	1
Trust in Traditional Authority	41,081	0.6357	0.4812	0	1
Traditional Authority Does Not Engage in Corruption	39,263	0.2893	0.4535	0	1
Approve of Traditional Authority Performance	37,483	0.6928	0.4614	0	1
Ideal Traditional Authority Influence in Local Community	40,931	0.8535	0.3536	0	1
Traditional Authority Serves Community Interests	41,391	0.5666	0.4956	0	1
Traditional Authority Should Advise on Voting	41,785	0.2892	0.4534	0	1
TA Perceptions Index (Binary Recode)	43,721	50.3117	25.1241	0	100

The *Traditional Authorities TA Influence Index* is created from the following four questions from AfroBarometer Round 8.

1. *TA Influence: Governing Local Community*

   Q87A: Now let's talk about traditional leaders and their role in politics and government in this country. How much influence do traditional leaders currently have in each of the following areas: Governing your local community?

2. *TA Influence: Allocating Land*

   Q87B: How much influence do traditional leaders currently have in each of the following areas: Allocating land?

3. *TA Influence: Voting Behavior*

   Q87C: How much influence do traditional leaders currently have in each of the following areas: Influencing how people in their communities vote?

4. *TA Influence: Dispute Resolution*

   Q87D: How much influence do traditional leaders currently have in each of the following areas: Solving local disputes?

   All four variables are recoded as follows:

- 1 None $\Longrightarrow$ 0 None

- 2 A small amount $\Longrightarrow$ 1 Little
- 3 Some $\Longrightarrow$ 2 Some
- 4 A lot $\Longrightarrow$ 3 A Lot
- All other responses {8 Refused, 9 Don't know/Haven't heard, -1 Missing} coded as missing

Table 5.6. Correlations Between Indices

	TA Perceptions	TA Influence	LC Perceptions	MP Perceptions
TA Perceptions	1.000			
TA Influence	0.277	1.000		
LC Perceptions	0.362	0.077	1.000	
MP Perceptions	0.281	0.036	0.612	1.000

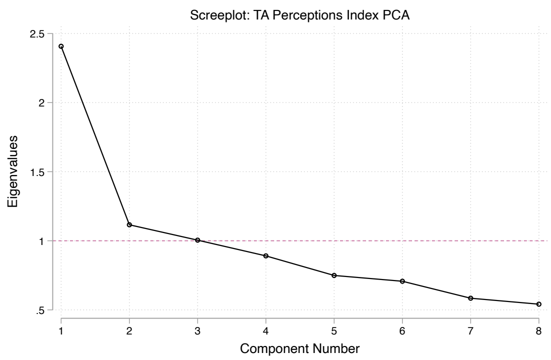
Summary Statistics

Table 5.7. Summary Statistics for Covariates

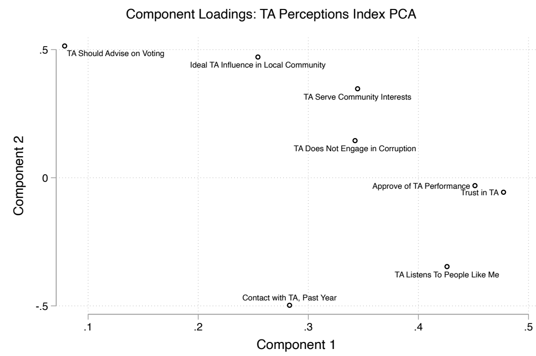
Variable	n	μ	σ	Min	Max	Percent
State Incorporation Status	46,560	0.6160	0.4864	0	1	
Female	48,084	0.4999	0.5000	0	1	
Rural	47,364	0.5553	0.4969	0	1	
Age	48,069	37.0628	14.7818	18	102	
Connected to Electric Grid	47,959	0.5710	0.4949	0	1	
Toilet in Household Compound	47,994	0.6940	0.4608	0	1	
National Identity	42,458	0.4856	1.1875	-2	2	
Comfort w Heritage Language / Dress in Public	46,327	0.8664	0.3402	0	1	
Wealth Index	48,084	3.8166	2.0161	0	7	
Household Owns Radio	47,804	0.7515	0.4321	0	1	
Household Owns TV	47,552	0.6086	0.4881	0	1	
Household Owns Vehicle / Motorbike	47,412	0.3864	0.4869	0	1	
Household Owns Computer	47,289	0.2616	0.4395	0	1	
Household Owns Bank Account	47,254	0.5096	0.4999	0	1	
Household Owns Mobile Phone	47,794	0.9065	0.2911	0	1	
Mobile Phone has Internet Access	39,388	0.5218	0.4995	0	1	
Education (Ordinal)						
≤ Primary	22,978					47.96%
≤ Secondary Education	16,943					35.36%
≤ University / Post-Grad	7,988					16.67%

Note: Wealth Index is the sum of the seven binary variables indicating household ownership of a given item. National Identity (Q82B) is recoded as negative if respondent identifies more strongly with ethnic identity, positive if respondent identifies with national identity, and 0 if the respondent is indifferent.

Principal Component Analysis

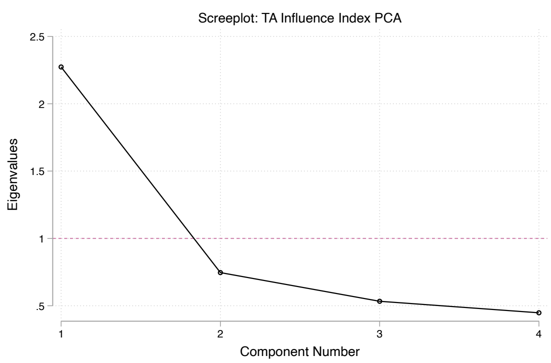


(a) Scree Plot

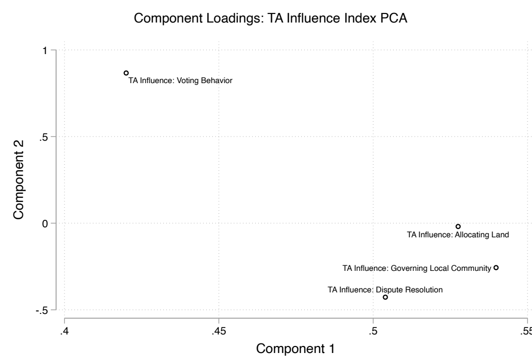


(b) Loadings Plot

Figure 5.8. Principal component analysis of the *TA Perceptions Index*.

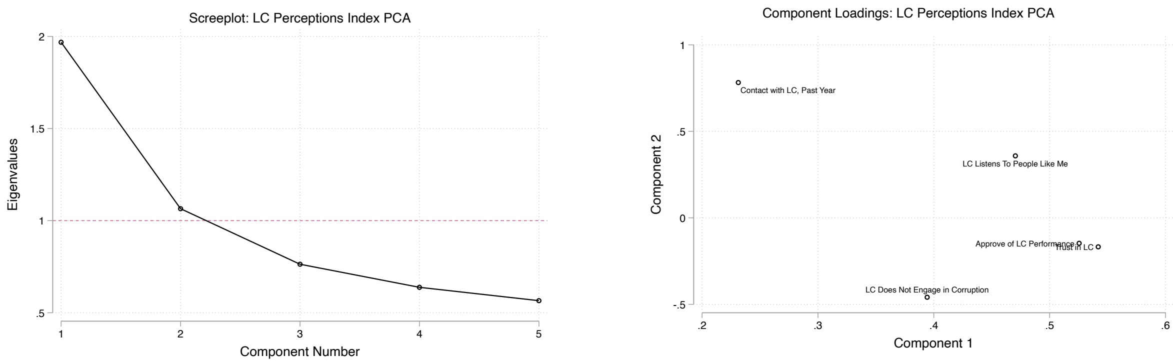


(a) Scree Plot



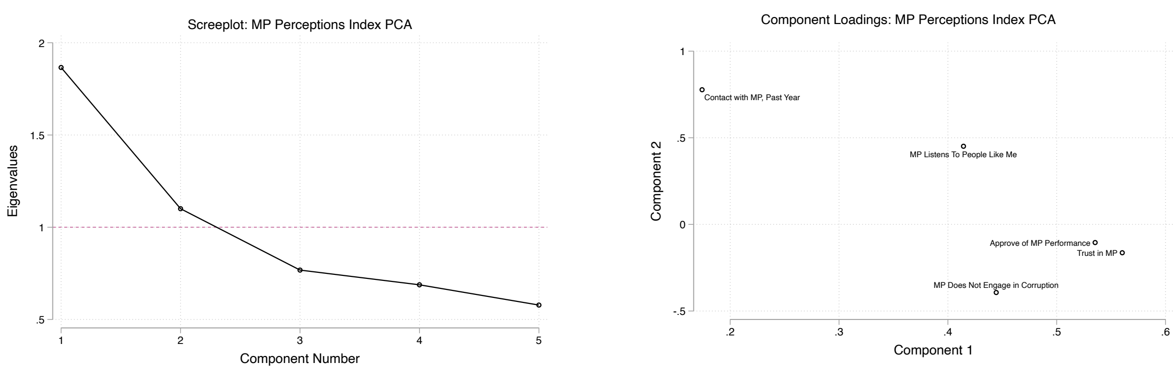
(b) Loadings Plot

Figure 5.9. Principal component analysis of the *TA Influence Index*.



(a) Scree Plot **(b)** Loadings Plot

Figure 5.10. Principal component analysis of the *LC Influence Index*.



(a) Scree Plot **(b)** Loadings Plot

Figure 5.11. Principal component analysis of the *LC Influence Index*.

Table 5.8. Bivariate OLS Using First Component as Outcome

	TA Perceptions	TA Influence	LC Perceptions	MP Perceptions
State-Incorporated	-0.441***	-0.221***	-0.297***	-0.246***
	(0.0294)	(0.0258)	(0.0255)	(0.0244)
Constant	0.271***	0.132***	0.195***	0.155***
	(0.0209)	(0.0194)	(0.0198)	(0.0186)
Observations	33,089	39,042	34,540	37,110
R^2	0.020	0.005	0.010	0.008

Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.9. Full OLS Specification Using First Component as Outcome

	TA Perceptions	TA Influence	LC Perceptions	MP Perceptions
State-Incorporated	-0.212***	-0.0533*	-0.129***	-0.0997***
	(0.0294)	(0.0290)	(0.0275)	(0.0267)
Female	-0.161***	-0.0175	-0.0818***	-0.0205
	(0.0161)	(0.0142)	(0.0151)	(0.0138)
Rural	0.544***	0.255***	0.299***	0.187***
	(0.0346)	(0.0308)	(0.0289)	(0.0280)
Age	0.00342***	-0.000727	0.00252***	-0.000841
	(0.000700)	(0.000671)	(0.000656)	(0.000624)
Secondary Education	-0.250***	0.0629***	-0.172***	-0.210***
	(0.0243)	(0.0231)	(0.0221)	(0.0217)
University / Post-Grad	-0.239***	0.00658	-0.206***	-0.274***
	(0.0343)	(0.0331)	(0.0305)	(0.0286)
Wealth Index	-0.00950	0.00763	0.00942	-0.0257***
	(0.00742)	(0.00696)	(0.00698)	(0.00605)
Connected to Electric Grid	-0.364***	-0.257***	-0.178***	-0.0688**
	(0.0318)	(0.0285)	(0.0279)	(0.0268)
Toilet in Household Compound	0.0894***	-0.00724	0.0158	0.0776***
	(0.0252)	(0.0243)	(0.0239)	(0.0221)
National Identity	0.00130	-0.0427***	0.0111	0.0121
	(0.00894)	(0.00871)	(0.00807)	(0.00804)
Comfort w Heritage Lang. / Dress in Public	0.0785**	0.117***	-0.0420	-0.166***
	(0.0306)	(0.0289)	(0.0283)	(0.0251)
Constant	0.00354	-0.0600	0.0351	0.334***
	(0.0612)	(0.0564)	(0.0545)	(0.0518)
Observations	29,270	33,957	31,228	33,068
R^2	0.108	0.024	0.041	0.032

Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

OLS Results: Traditional Authorities and State Incorporation

Table 5.10 provides the full set of results for Model 5.1. Columns (1) and (2) shows the effect of state incorporation on perceptions of traditional authorities, columns (3) and (4) shows the effect of state incorporation on traditional authorities' influence in the local community.

Table 5.10. Effect of State Incorporation on Attitudes Toward Traditional Authorities

	TA Perceptions Index		*TA Influence Index*	
	(1)	(2)	(3)	(4)
State-Incorporated	-7.374***	-2.984***	-3.994***	-1.017*
	(0.373)	(0.367)	(0.469)	(0.530)
Female		-2.737***		-0.293
		(0.185)		(0.262)
Rural		8.190***		4.337***
		(0.437)		(0.555)
Age		0.0391***		-0.0124
		(0.00837)		(0.0124)
Secondary Education		-2.722***		1.012**
		(0.292)		(0.415)
University / Post-Grad		-2.843***		-0.0843
		(0.412)		(0.606)
Wealth Index		-0.163*		0.133
		(0.0839)		(0.123)
Connected to Electric Grid		-5.013***		-4.538***
		(0.417)		(0.526)
Toilet in Household Compound		0.0881		-0.0311
		(0.306)		(0.442)
National Identity		0.0324		-0.824***
		(0.106)		(0.159)
Comfort w Heritage Language / Dress in Public		2.057***		2.091***
		(0.381)		(0.527)
Constant	56.74***	52.80***	59.11***	55.84***
	(0.261)	(0.745)	(0.355)	(1.037)
Observations	43,453	37,496	41,415	35,896
R^2	0.029	0.129	0.005	0.021

Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 5.11 provides the results of Model 5.1 using the alternative coding of the *TA Perceptions Index*. Column (1) provides the bivariate correlation and Column (2) includes all controls. Results are consistent with those in the first two columns of Table 5.10.

Table 5.11. Effect of State Incorporation on Perceptions of Traditional Authorities

	TA Perceptions Index	
	(1)	(2)
State-Incorporated	-9.034***	-3.860***
	(0.438)	(0.434)
Female		-2.742***
		(0.221)
Rural		9.007***
		(0.515)
Age		0.0291***
		(0.0104)
Secondary Education		-3.709***
		(0.354)
University / Post-Grad		-3.735***
		(0.493)
Wealth Index		-0.411***
		(0.0990)
Connected to Electric Grid		-5.816***
		(0.469)
Toilet in Household Compound		0.137
		(0.364)
National Identity		-0.0543
		(0.125)
Comfortable w Heritage Language / Dress in Public		1.755***
		(0.442)
Constant	56.06***	53.62***
	(0.309)	(0.869)
Observations	43,453	37,496
R^2	0.031	0.128

Note: Outcome variable is alternate (binary recode) index.
Standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

Supplementary RDD Results

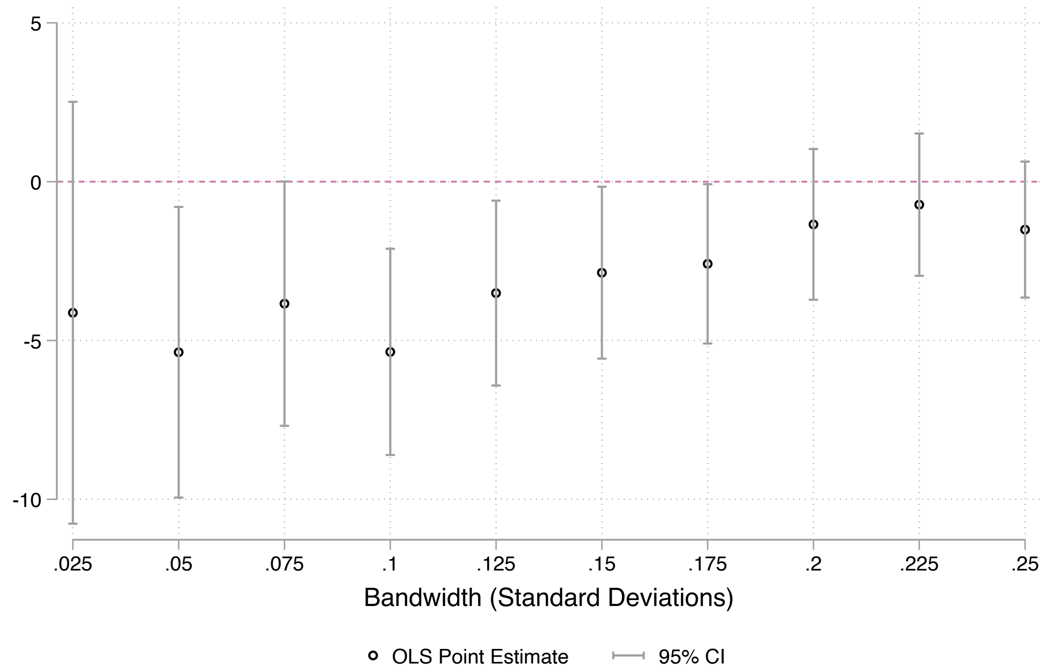


Figure 5.12. RD estimates for the *TA Perceptions Index* at different bandwidths.

Table 5.12. Local Polynomial Models with Controls

	TA Perceptions Index	TA Influence Index
	(1)	(2)
RD Estimate	-2.300**	-1.239
	(1.017)	(1.205)
Observations	37,496	35,896
Controls	Yes	Yes
Kernel Type	Triangular	Triangular
Robust 95% CI	[-4.776 ; -.505]	[-4.11 ; 1.263]
Conventional p	0.0237	0.304
Robust p	0.0154	0.299
Order Local Polynomial (p)	2	2
Order Bias (q)	3	3
BW Local Polynomial (h)	0.581	0.873
BW Bias (b)	0.960	1.169

Standard errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

OLS Results: Formal Authorities and State Incorporation

Table 5.13. Effect of State Incorporation on Attitudes Toward Formal Authorities

	LC Perceptions Index		*MP Perceptions Index*	
	(1)	(2)	(3)	(4)
State-Incorporated	-4.170***	-2.166***	-3.149***	-1.501***
	(0.320)	(0.354)	(0.300)	(0.330)
Female		-1.977***		-0.848***
		(0.193)		(0.171)
Rural		4.415***		2.513***
		(0.370)		(0.342)
Age		0.0548***		0.00553
		(0.00831)		(0.00768)
Secondary Education		-1.660***		-1.878***
		(0.286)		(0.263)
University / Post-Grad		-2.147***		-2.615***
		(0.396)		(0.352)
Wealth Index		0.320***		-0.179**
		(0.0891)		(0.0743)
Connected to Electric Grid		-2.588***		-0.942***
		(0.358)		(0.325)
Toilet in Household Compound		0.476		1.062***
		(0.312)		(0.273)
National Identity		0.165		0.263***
		(0.104)		(0.0981)
Comfort w Heritage Language / Dress in Public		-0.842**		-2.427***
		(0.364)		(0.309)
Constant	40.04***	36.78***	34.52***	35.82***
	(0.250)	(0.697)	(0.232)	(0.637)
Observations	42,347	37,067	44,694	39,029
R^2	0.010	0.040	0.007	0.024

Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

5.7 Works Cited

Alangea, Deda Ogum, Adolphina Addoley Addo-Lartey, Yandisa Sikweyiya, Esnat Dorothy Chirwa, Dorcas Coker-Appiah, Rachel Jewkes, and Richard Mawuena Kofi Adanu. 2018. "Prevalence and Risk Factors of Intimate Partner Violence among Women in Four Districts of the Central Region of Ghana: Baseline Findings from a Cluster Randomised Controlled Trial." *PLOS ONE* 13 (7): e0200874.

Baldwin, Kate. 2015. *The Paradox of Traditional Chiefs in Democratic Africa.* New York: Cambridge University Press.

———. 2020. "Chiefs, Democracy, and Development in Contemporary Africa." *Current History* 119 (817): 163–168.

Bates, Robert H., and Steven A. Block. 2013. "Revisiting African Agriculture: Institutional Change and Productivity Growth." *The Journal of Politics* 75 (2): 372–384.

Boone, Catherine. 1997. "'Empirical Statehood' and Reconfigurations of Political Order," edited by Leonardo Villalon and Philip Huxtable, 129–142. Boulder, CO, USA: Lynne Rienner Press.

Bratton, Michael, and Nicholas van de Walle. 1997. *Democratic Experiments in Africa: Regime Transitions in Comparative Perspective.* Cambridge University Press.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik. 2014. "Robust Data-Driven Inference in the Regression-Discontinuity Design." *The Stata Journal* 14 (4): 909–946.

de Kadt, Daniel, and Horacio A. Larreguy. 2018. "Agents of the Regime? Traditional Leaders and Electoral Politics in South Africa." *The Journal of Politics* 80 (2): 382–399.

Dionne, Kim Yi. 2017. *Doomed Interventions: The Failure of Global Responses to AIDS in Africa.* Cambridge: Cambridge University Press.

Englebert, Pierre. 2002a. "Born-Again Buganda or the Limits of Traditional Resurgence in Africa." *The Journal of Modern African Studies* 40 (3): 345–368.

———. 2002b. "Patterns and Theories of Traditional Resurgence in Tropical Africa." *Mondes en Développement* (Louvain-la-Neuve) 118 (2): 51–64.

Fosu, Augustin. 2018. *Governance and Development in Africa: A Review Essay (Working Paper 298).* Working Paper Series. Abidjan, Côte d'Ivoire: African Development Bank.

Garcia, Andres J., Deepa K. Pindolia, Kenneth K. Lopiano, and Andrew J. Tatem. 2015. "Modeling Internal Migration Flows in Sub-Saharan Africa Using Census Microdata." *Migration Studies* 3 (1): 89–110.

Gelman, Andrew, and Guido Imbens. 2019. "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics* 37 (3): 447–456.

Henn, Soeren J. 2022. "Complements or Substitutes? How Institutional Arrangements Bind Traditional Authorities and the State in Africa." *American Political Science Review,* 1–20.

Herbst, Jeffrey. 1990. "Migration, the Politics of Protest, and State Consolidation in Africa." *African Affairs* 89 (355): 183–203. JSTOR: 722241.

Jayachandran, Seema. 2015. “The Roots of Gender Inequality in Developing Countries.” *Annual Review of Economics* 7 (1): 63–88.

Koelble, Thomas A., and Edward LiPuma. 2011. “Traditional Leaders and the Culture of Governance in South Africa.” *Governance* 24 (1): 5–29.

Koter, Dominika. 2016. *Beyond Ethnic Politics in Africa.* Cambridge University Press.

Kpundeh, Sahr, and Brian Levy. 2004. *Building State Capacity in Africa.* World Bank Institute Development Studies. The World Bank.

Kyed, Helene Maria, and Lars Buur. 2007. “Introduction: Traditional Authority and Democratization in Africa.” In *State Recognition and Democratization in Sub-Saharan Africa: A New Dawn for Traditional Authorities?,* edited by Lars Buur and Helene Maria Kyed, 1–28. Palgrave Studies in Governance, Security, and Development. New York: Palgrave Macmillan US.

Laver, Michael. 1976. “‘Exit, Voice, and Loyalty’ Revisited: The Strategic Production and Consumption of Public and Private Goods.” *British Journal of Political Science* 6 (4): 463–482.

LiPuma, Edward, and Thomas A. Koelble. 2009. “Deliberative Democracy and the Politics of Traditional Leadership in South Africa: A Case of Despotic Domination or Democratic Deliberation?” *Journal of Contemporary African Studies* 27 (2): 201–223.

Logan, Carolyn. 2009. “Selected Chiefs, Elected Councillors and Hybrid Democrats: Popular Perspectives on the Co-Existence of Democracy and Traditional Authority*.” *The Journal of Modern African Studies* 47 (1): 101–128.

———. 2013. “The Roots of Resilience: Exploring Popular Support for African Traditional Authorities.” *African Affairs* 112 (448): 353–376.

Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism.* First Edition. Princeton, N.J: Princeton University Press.

Mosley, Paul, and John Weeks. 1993. “Has Recovery Begun? “Africa’s Adjustment in the 1980s” Revisited.” *World Development* 21 (10): 1583–1606.

Robinson, Amanda Lea, and Jessica Gottlieb. 2021. “How to Close the Gender Gap in Political Participation: Lessons from Matrilineal Societies in Africa.” *British Journal of Political Science* 51 (1): 68–92.

Sikweyiya, Yandisa, Adolphina Addoley Addo-Lartey, Deda Ogum Alangea, Phyllis Dako-Gyeke, Esnat D. Chirwa, Dorcas Coker-Appiah, Richard M. K. Adanu, and Rachel Jewkes. 2020. “Patriarchy and Gender-Inequitable Attitudes as Drivers of Intimate Partner Violence against Women in the Central Region of Ghana.” *BMC Public Health* 20 (1): 682.

Sklar, Richard L. 1993. “The African Frontier for Political Science.” In *Africa and the Disciplines: The Contributions of Research in Africa to the Social Sciences and Humanities,* edited by Robert H. Bates, V. Y. Mudimbe, and Jean F. O’Barr. University of Chicago Press.

van der Windt, Peter, Macartan Humphreys, Lily Medina, Jeffrey F. Timmons, and Maarten Voors. 2019. “Citizen Attitudes Toward Traditional and State Authorities: Substitutes or Complements?” *Comparative Political Studies* 52 (12): 1810–1840.

Von Trotha, Trutz. 1996. “From Administrative to Civil Chieftaincy: Some Problems and Prospects of African Chieftaincy.” *The Journal of Legal Pluralism and Unofficial Law* 28 (37-38): 79–107.

Wesolowski, Amy, Wendy Prudhomme O’Meara, Nathan Eagle, Andrew J. Tatem, and Caroline O. Buckee. 2015. “Evaluating Spatial Interaction Models for Regional Mobility in Sub-Saharan Africa.” *PLOS Computational Biology* 11 (7): e1004267.

Williams, J. Michael. 2010. *Chieftaincy, the State, and Democracy: Political Legitimacy in Post-Apartheid South Africa*. Indiana University Press.

Chapter 6

Public Health Outcomes and the State

Abstract

Scholars are increasingly interested in the quality of governance and public goods provision by states, and in particular, whether and why outcomes vary *within* states. In this chapter, I focus on one important type of public good—public health—and assess whether living in state-controlled territory has any discernible effect on the incidence of common endemic diseases, such as malaria and HIV. Consistent with previous research, I find that malaria rates are significantly lower in state-incorporated regions of Africa. Paradoxically, I find that the opposite is true of HIV. This finding contradicts the few studies that evaluate the relationship between state-capacity and HIV. To explain this finding, I propose and test the theory that anti-HIV stigma attenuates the efficacy of public health messaging and other interventions in state-incorporated areas. Because people fear that association with HIV, or with groups at high risk of HIV, will reveal information about their private behaviors to the state, they will avoid contact with the public health system, resulting in lower levels of HIV-related knowledge. Less knowledgable individuals in these incorporated areas are more likely to engage in high-risk behavior, resulting in higher transmission and incidence rates of HIV. My results suggest that this explanation is plausible, though further research based on non-observational data is warranted.

6.1 Introduction

Over the last few years of the COVID-19 pandemic, we observed a great deal of variation in the success that various national governments achieved in reducing the transmission and mortality rates of the virus. The pandemic underscored the pivotal role of the state in mitigating the negative effects of public health emergencies and other disasters, as public health outcomes often depend on the capacity of the state to deal with these types of shocks (Serikbayeva et al. 2021). Indeed, a handful of recent studies find that higher capacity states tended to be more proactive in their pandemic response, and more successful in moderating the disease burden of COVID-19 (Bosancianu et al. 2020; Bollyky et al. 2022; Christensen and Lægreid 2020; Yen et al. 2022). These findings jive with the conventional wisdom that high-capacity states are better equipped to manage endemic disease than less-developed or less-capacious states (Gizelis 2009; Majeed and Gillani 2017). Higher levels of state capacity allow countries to detect and monitor disease outbreaks, target public health programming toward vulnerable demographics, and implement effective policy interventions. Such efforts, in turn, are thought to result in lower rates of disease transmission, reduced infant mortality, and higher life expectancy (D'Arcy and Nistotskaya 2017; Hanson and Sigman 2021; Holmberg and Rothstein 2011).

In this chapter, I explore how *subnational variation* in state capacity influences public health outcomes across the African continent. My analysis focuses on malaria and HIV, as these diseases represent the two most significant long-term epidemics affecting African countries. In line with previous literature, I hypothesize that disease rates will tend to be lower in state-incorporated areas:

Hypothesis 1: *The incidence of endemic diseases, such as malaria or HIV, will be lower in state-incorporated areas than it is in areas of limited statehood.*

State-incorporated areas are regions of the state with a high degree of government penetration; they are characterized by high levels of infrastructure density and administrative presence. In such areas, we expect states to be better able to respond to ongoing epidemics or emergent outbreaks of disease. Agents of the state based in these regions are closer to the problems they are meant to address, and thus encounter fewer logistical barriers in coordinating public health responses, implementing policy (e.g., vector eradication in the case of malaria), regulating behavior (e.g., enforcing mask mandates and physical distancing in the early days of the COVID-19 pandemic), and providing educational resources. Additionally, these high-capacity areas are typically associated with increased health-related public goods provision, such as education, sanitation and water treatment, food safety standards, and transportation infrastructure (D'Arcy and Nistotskaya 2017; Dittmar and Meisenzahl

2017).

Hypothesis 1 is consistent with findings from Boussalis et al. (2012), who conduct one of the few existing studies that evaluates the effects of state capacity on disease incidence at the subnational level. Using methods analogous to those outlined in Section 6.3 below, these authors show that malaria rates are significantly lower in high-capacity Indian states, as measured by relative political extraction (the ratio of actual tax effort to predicted tax effort).[1] This hypothesis is also broadly consistent with work done by Price-Smith et al. (2004) and Gizelis (2009). Using country-level (rather than subnational) data and a pooled time-series cross sectional model, Gizelis finds that high-capacity states are more effective in mitigating HIV transmission than low-capacity states. In what follows, I find evidence to support Hypothesis 1 in the case of malaria. Similar to Boussalis et al. (2012), my results show that the incidence of malaria is significantly lower in state-incorporated territory than it is in unincorporated territory. Surprisingly, however, I find that the opposite is true of HIV. Contrary to expectations, rates of HIV are consistently higher in state-incorporated territory. This relationship persists throughout the entire 18-year study period for which high-quality subnational data is available.

In sections 6.4 and 6.5, I explore this puzzling finding in more detail. I argue that these divergent epidemic dynamics may be the result of anti-HIV stigma and discrimination, which remain rampant throughout Africa. Unlike other endemic diseases, HIV spreads primarily through sexual contact or through intravenous drug use, both of which are sensitive subjects in many cultures. This is also a disease that tends to infect already marginalized populations, including women, commercial sex workers, men who have sex with men, and transgender individuals—all groups whose status or lifestyles make them susceptible to harassment or sanctioning by authorities. In state-incorporated territory, where the threat and consequences of official sanctioning are particularly acute, individuals may attempt to avoid any association with HIV, or with one of these high-risk groups, and in so doing, shy way from public health messaging that encourages risk-reducing behavior, such as safer sex practices and the use of clean needles. My analyses show that this avoidance of the public health system, motivated by a fear of revealing private behaviors to the state, is one plausible explanation of the higher rates of HIV that we observe in state consolidated territory.

1. See Swaminathan and Thomas (2012).

The empirics in this chapter confirm some of our basic intuitions about the relationship between state capacity and public health, while at the same time establishing a counterintuitive finding in the case of HIV, which suggests that this relationship is much more nuanced than previously thought. The chapter takes some preliminary steps to develop and test a possible explanation for this for this finding. Substantively, this research contributes to the literature on welfare-enhancing public goods provision. In recent years, political scientists have renewed their interest in the effects of living in state-consolidated versus non-consolidated territory, and whether or not the state is beneficial for social and economic development. Much of this debate is structured in terms of state versus non-state (e.g., traditional authorities, rebel groups, NGOs) provision (Lee et al. 2014; Post et al. 2017; Börzel and Risse 2021; Carlitz and Lust 2021). Due to data limitations, I focus primarily on dynamics inside state-controlled territory, and blackbox some of the health-related interventions implemented by non-state actors.[2] My findings do, however, support the theory that the state does have a positive impact on public health, not just in reducing the rates of common diseases like malaria, but also in marginally increasing levels of awareness of—and voluntary testing for—stigmatized diseases like HIV. Finally, this chapter contributes to the epidemiological literature, suggesting ways to tailor interventions for malaria- and HIV-vulnerable populations based on political geography.

6.2 Malaria and HIV-AIDS in Africa

Malaria and HIV-AIDS are among the most widespread communicable diseases in Africa. The combined burden of these two diseases is enormous; the World Health Organization's Global Disease Observatory estimates that HIV-AIDS kills more Africans than any other illness (122 deaths per 100k population per year). Malaria comes in third on this list (62 deaths per 100k), just behind diarrheal disease (67 deaths per 100k). These diseases pose a significant economic burden as well. Economists estimate that malaria is responsible for a "growth penalty" of between 1.1 and 1.3% of GDP per year in Africa (Andrade et al. 2022), while HIV-AIDS reduces growth by 2 to 4% per year on average in African countries (Nketiah-Amponsah et al. 2019). These estimates do not account for personal spending on prophylaxis, which can be substantial. Chima et al. (2003) estimate that

2. Examples include criminal organizations enforcing COVID-19 mask mandates and lockdowns in El Salvador and Brazil during the early pandemic, and religious organizations fostering ART compliance in Ethiopia (Kumsa and Tucho 2019).

monthly per capita expenditures on malaria prevention reach $0.41 USD ($1.88 per household) in Malawi and $3.88 ($26 per household) in Cameroon. Pre-Exposure Prophylaxis (PrEP, discussed in Section 6.2.2), a drug used to prevent HIV transmission, can cost between $394 to $760 per year in Africa (based on Zambia estimates from Hendrickson et al. (2022))—a price unaffordable for many Africans.[3] Given the immense human and economic costs associated with these diseases, states have significant incentive to minimize their spread.

6.2.1 Malaria

Malaria is a mosquito-borne infectious disease caused by parasitic organisms from the *Plasmodium* genus. In human populations, malaria is caused by a half dozen different species of parasite, though most severe infections result from *P. falciparum*, while milder infections result from *P. vivax* and other *Plasmodium* species. These parasites are present in the *Anopheles* genus of mosquitos, which represent the primary vector of malaria transmission to human populations. Because these mosquitos require standing water to reproduce, and their survival and lifecycle is temperature dependent, their geographic range is well established and widely used by epidemiologists as a baseline measure of malaria suitability (Christiansen-Jucht et al. 2014; CDC 2020; Villena et al. 2022).

While most high-income countries have eradicated the disease, malaria remains endemic throughout tropical Africa. An estimated 95% of global malaria cases occur in sub-Saharan Africa, and and 96% of all malaria-related deaths occur on the continent (WHO 2022). The disease primarily affects children under the age of five (who represent 80% of malaria-related deaths in Africa), pregnant women, and individuals with weakened immune systems (WHO 2022). The high prevalence of malaria in Africa is due to a number of factors, including poor sanitation, inadequate access to healthcare, and a lack of mosquito control measures. While certain drugs are available for chemoprophylaxis, most prevention measures involve vector control, including insecticide-treated nets, indoor residual spraying, and larviciding. There is currently one approved malaria vaccine (RTS,S/AS01 / Mosquirix), which has been shown to reduce the rate of deadly severe malaria in children by 30% (Duffy 2022). Pilot rollouts of the the vaccine are currently underway in Ghana, Kenya and Malawi.

3. Despite the widespread adoption of PrEP in wealthy countries, a meta-analysis by Case et al. (2019) finds that PrEP is *not* a cost-effective method of HIV prevention in Africa given its unaffordability. Even in the United States, the cost of PrEP without insurance ranges from $22,000 to $30,000 USD per year.

Effective treatments for malaria do exist, though access to these drugs is limited on the continent, and drug efficacy is becoming an increasingly concerning problem as the parasites adapt resistance to common antimalarial drugs (WHO 2020).

6.2.2 HIV-AIDS

Africa is one of the regions in the world most heavily affected by the HIV-AIDS epidemic. According to the Joint United Nations Programme on HIV-AIDS (UNAIDS), 71.25% of global HIV-AIDS cases are in sub-Saharan Africa (UNAIDS 2022). The region accounts for nearly two-thirds of all new HIV infections and three-quarters of all AIDS-related deaths (Kharsany and Karim 2016).

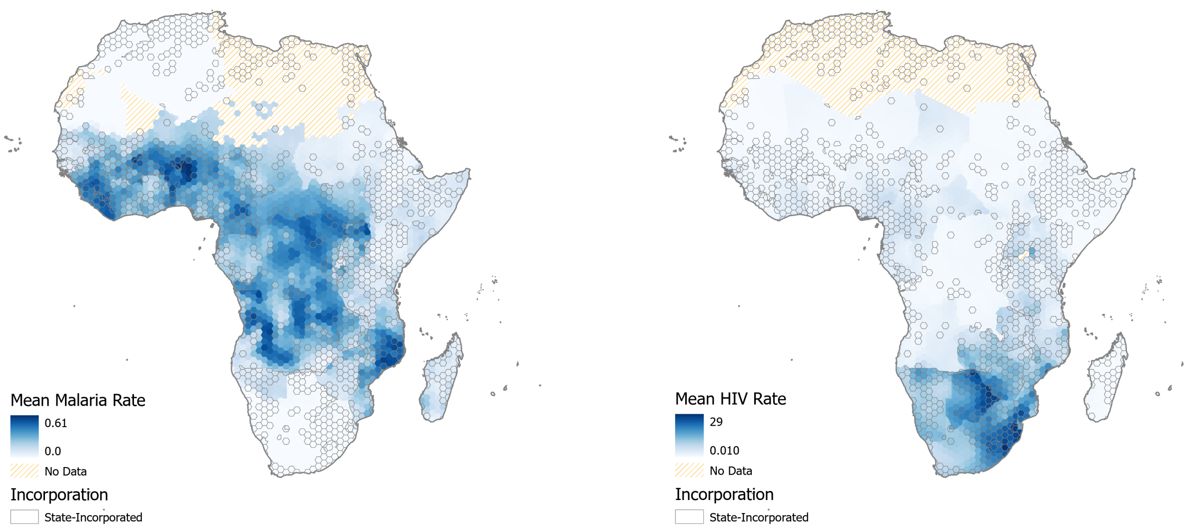
The virus is primarily spread through sexual contact, sharing of needles, and from mother to child during pregnancy, childbirth, or breastfeeding (vertical transmission). The HIV epidemic in Africa has been fueled by a range of factors, including poverty, gender inequality, and limited access to healthcare and prevention services. In many African countries, stigma and discrimination against people living with HIV remain major barriers to prevention, treatment, and care. However, there has been significant progress in recent years to address the HIV epidemic in Africa. Access to highly active antiretroviral therapy (HAART)—a combination of drugs that prevents the virus from making copies of itself in the body and reduces viral load—has increased significantly (UNAIDS 2013). There have also been successful prevention initiatives such as the distribution of condoms, Pre-Exposure Prophylaxis (PrEP), and voluntary medical male circumcision. Despite these efforts, however, the HIV epidemic in Africa remains a significant public health challenge that requires continued attention and investment.

6.3 The Effects of State Incorporation on Disease Incidence

To evaluate Hypothesis 1, I focus on two outcome variables. The first is the incidence rate of the Malaria-causing parasite *P. falciparum*, which is estimated globally for the years 2000 to 2020 at a 5×5 kilometer resolution by the Malaria Atlas Project (MAP).[4] This rate is defined as the proportion of children 2 to 10 years of age showing detectable *P. falciparum* parasite; this is a standard measure of malaria prevalence used by epidemiologists. The second outcome variable is the incidence rate

4. Although *P. vivax* is present in tropical Africa, its incidence is quite rare (generally less than 10 cases per thousand) and its geographic extent is limited to the Horn of Africa and the southern regions of Sudan.

of HIV. Subnational HIV data come from the Institute for Health Metrics and Evaluation (IHME), which estimates the prevalence of HIV among adults 15 to 49 years of age in each area of a 5×5 kilometer grid cell for the years 2000 through 2017.[5] Both of these sources use high-quality sentinel surveillance data to generate spatiotemporal estimates of incidence rates, which alleviates some of the measurement concerns associated with data collected by national and regional governments. Indeed, the MAP and IHME estimates were each developed, in part, to account for subnational spatial heterogeneity in disease prevalence and to guide policy making and prevention efforts in the absence of reliable national or local data. While all seroprevalence estimates involve some degree of uncertainty, these estimates represent what is arguably the closest approximation of the "true" extent of malaria and HIV on the continent in the 2000s and 2010s.



(a) Malaria Incidence Rates (2017) **(b)** HIV Incidence Rates (2017)

Figure 6.1. Spatial distribution of the two primary outcome variables for the year 2017. Areas coded as state-incorporated are outlined in gray. Note that the unit of analysis is the 10,000 km^2 hexagonal grid cell; individual pixels from the input rasters are averaged within each cell to produce a single cell-year specific measure.

Hypothesis 1 also requires a measure of state incorporation, which I derive from the control point density measure detailed in Chapter 4. I create a binary "treatment" variable that takes the value 1 if mean control point density in a given grid cell is above the median control point density for the entire country, and 0 otherwise. For the main analyses, the outcome variables, treatment variable,

5. See Dwyer-Lindgren et al. (2019) for a description of these data.

and all covariates are mean aggregated to the 10,000 km^2 hexagonal grid cell shown in Figure 6.1. In the Supplementary Information, I replicate all analyses using using a more granular 1000 km^2 grid cell, though results are consistent across different levels of aggregation.

I estimate the effect of state incorporation on disease rates using a standard repeated measures mixed model with a factorial interaction term:

$$\text{Incidence}_{ij} = \alpha + \beta_1 \text{State}_{ij} + \beta_2 \text{Year}_{ij} + \beta_3 (\text{State} \times \text{Year})_{ij} + \gamma X_{ij} + u_j + \epsilon_{ij}, \qquad (6.1)$$

where Incidence is the rate of either malaria or HIV in year i for cell j, State is the treatment variable, which is interacted with the time variable Year, X is a vector of disease-specific control variables measured at the cell-year, u_j is cell-specific random effect, and ϵ is a within-cell error term with an autoregressive (AR-1) structure to account for serial correlation in the time series. β_1 and β_2 are the main effects of the treatment variable State (i.e., the mean difference between treated and control cells) and the variable Year (i.e., a time trend), respectively, and β_3 is the interaction coefficient. Hypothesis 1 suggests that $\beta_1 < 0$ in both the malaria and HIV models.

The specification in Equation 6.1 includes several controls that influence the baseline rates of malaria and HIV in a given locale. Because malaria is a predominantly rural disease that tends to affect less-affluent individuals, while HIV is generally understood to be the opposite—an urban disease that affects wealthier or more cosmopolitan communities, I control for both urban territory and wealth in each model.[6] My measure of urban territory is a binary variable that takes a value 1 if at least half of the grid cell contains urban territory, and 0 otherwise. I include the cell-specific average of nighttime light emissions as a proxy for wealth. In the malaria model, I include a measure of malaria suitability (i.e., vector habitat availability) from a hydro-climatic model developed by Smith et al. (2020).[7] The HIV model controls for the rate of male circumcision, which is one of the

6. The reality is a bit more nuanced. While malaria transmission rates do tend to be higher in rural areas, urban transmission rates—particularly rates in periurban areas and in urban centers situated in wet savannas and forest zones—have increased over the past two decades, due in large part to the rapid urbanization of the African continent and the adaptation of various anopheline species (i.e., the mosquitos that carry the malaria-causing Plasmodium parasites) to urban aquatic habitats (e.g., blocked water drains, potholes, tires, etc.) (Robert et al. 2003; Machault et al. 2012). HIV, on the other hand, is generally an urban disease, though it tends to be highly-correlated with poverty rates, as poverty is associated with high-risk behaviors such as commercial sex work and and intravenous drug use (Cohen et al. 1997; Stockemer and Lamontagne 2007).

7. The basic suitability model dates back over a century to the work of Sir Ronald Ross, who won the Nobel Prize in 1902 for his discovery of the malaria parasite. Ross developed a simple mathematical model to explain the relationship between

strongest available predictors of aggregate HIV transmission rates (Szabo and Short 2000). I do not control for cell population or population density, as both outcomes are normalized as relative rates rather than raw counts.

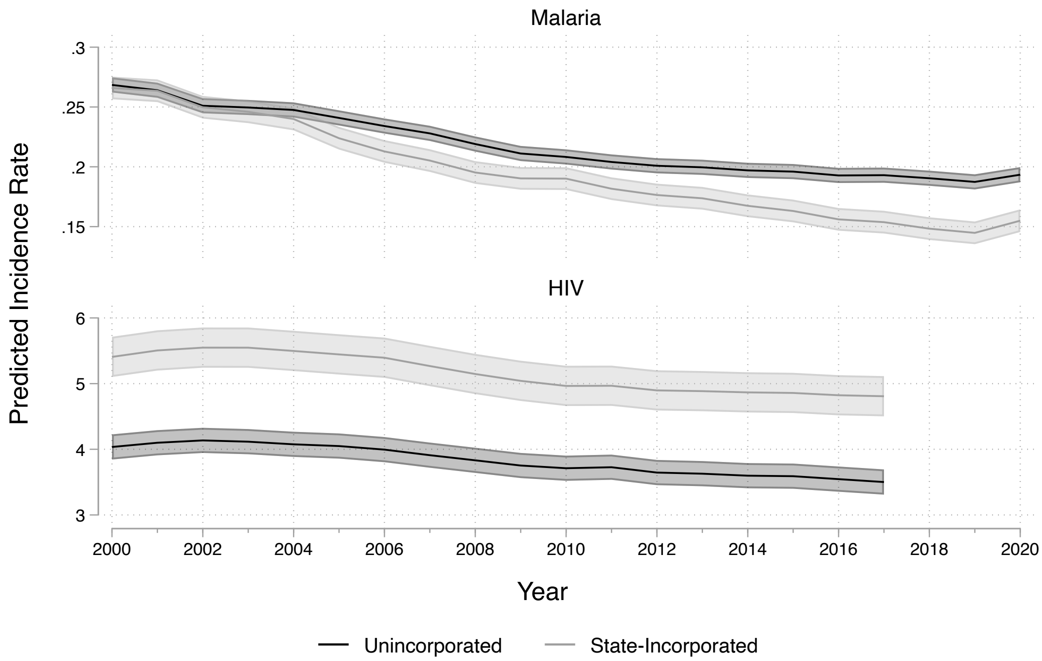


Figure 6.2. Predicted disease incidence rates and 95% confidence intervals in state-incorporated and unincorporated territory by disease and year.

The full set of results are available in Table 6.5 in the Appendix. The substantive effect of the State × Year interaction is visualized in Figure 6.2. Here, we see mixed support for Hypothesis 1. In the malaria model, the coefficient for the State variable is negative, though not significant at conventional levels. However, the average marginal effect of treatment is -0.023, which is statistically significant $\left(\chi^2_{(1)} = 14.94\right)$—malaria rates are roughly a tenth of a standard deviation lower, on average, in state-incorporated areas than they are in unincorporated areas. Coefficients on the interaction term β_3 are negative and significant for the years 2004 through 2020, indicating that, in each of these years, the malaria rate is significantly lower in state-incorporated territory than unincorporated

mosquito populations and the incidence of malaria in humans (Ross 1916; Mandal et al. 2011). Because mosquitos thrive in temperatures between 16 and 34°C and require standing water to breed, nearly all modern suitability models are based on temperature and rainfall data (60-80 mm of rainfall per month is the standard proxy for anopheline breeding habitat). The Smith et al. (2020) suitability measure used in these analyses also incorporates hydrological data to determine whether ground surface water is available for larval habitats.

territory, and that this difference is increasing over time; I calculate a marginal effect of -0.003 in the year 2000, and a much larger effect of -0.038 in 2020. This finding, coupled with the decreasing time trend in both malaria curves, likely reflects the efficacy of scaled-up prevention, diagnosis, and treatment interventions in reducing malaria burden, particularly in state-consolidated areas, over the past 15 years (O'Meara et al. 2010). The results in Table 6.5 also show that the malaria incidence rate is positively and significantly correlated with urban territory (see Footnote 6) and malaria suitability, and negatively and significantly correlated with wealth, as proxied by nighttime light emissions.

Rates of HIV, on the other hand, are consistently higher in state-incorporated territory. In the HIV model, the coefficient for the State variable is positive and statistically significant, and the average marginal effect is 1.33, which represents a quarter of a standard deviation increase in the HIV incidence rate between state-incorporated and unincorporated territory over the entire study period. This is a statistically significant difference $\left(\chi^2_{(1)} = 47.66\right)$. The HIV model also returns positive coefficients on urban territory and nighttime lights, though the nighttime lights coefficient is not significant at conventional levels. We also get a negative and significant coefficient on male circumcision, which is consistent with the epidemiological literature. Overall, these results do *not* support Hypothesis 1—contrary to expectations, the incidence of HIV is significantly higher in state-incorporated areas.

6.4 Anti-HIV Stigma and State Incorporation

Why should we expect divergent rates of malaria and HIV in state-incorporated territory? One possible explanation for the counterintuitive findings in Section 6.3 is that these inflated rates of HIV are the result of anti-HIV stigma, which is pervasive in many African countries (Teshale and Tesema 2022). In Table 6.1, for example, I summarize the results of several HIV-related opinion items from a combined African sample of the Demographic and Health Survey (DHS). These indicators reflect a fairly high degree of prejudice against people living with HIV (PLHIV). While only 35% of respondents would feel ashamed if a family member was HIV-positive, nearly 60% would want this fact to remain a secret. Tellingly, 82% of respondents who would *not* be ashamed of an HIV-positive family member would still want their family member's status to remain a secret. Roughly 70% of the sample feel that HIV-positive people are the subject of gossip or viewed negatively by others.

Table 6.1. DHS Indicators of HIV Stigma

DHS Survey Item	n	Affirmative Responses
Would want HIV infection in family to remain a secret	288501	59.12%
Would be ashamed if family member had HIV	264593	36.61%
Willing to care for a relative who has AIDS*	252847	82.79%
Would buy vegetables from a vendor with HIV-AIDS*	536609	59.13%
Should child with HIV be allowed to attend school with other children?*	277661	70.00%
Should a female teacher with HIV who is not sick continue teaching?*	242763	68.81%
People talk badly about people living with HIV	259101	71.09%
People living with HIV lose respect of others	258074	67.88%
People hesitate to take an HIV test for fear of others' reaction if positive	256574	82.89%
People hesitate to take an HIV test because of fear / stigma / discrimination	6539	35.08%

Note: Table excludes "Don't Know" and missing responses from tabulation.
*Questions for which a negative response connotes stigma are tagged with an asterisk.

The fact is that Malaria and HIV are qualitatively different in terms of their epidemiology. They vary not only in their modes of transmission, but also in the populations they tend to affect. While malaria is, in some ways, indiscriminate, in that nearly everyone living in a transmission zone is potentially susceptible to infection, HIV is largely spread through "taboo" behaviors, and disproportionately affects marginalized or criminalized groups. HIV is therefore viewed disparagingly by both governments and the broader society. Fear of this intense stigma can influence the ways in which at-risk individuals consume information. In particular, it may cause people to deliberately disengage from the public health system—especially programs and messaging concerning HIV—in order to avoid any association with the disease, or to avoid calling attention to their to their occupation, lifestyle, or behavioral choices. Accessing information about HIV or other stigmatized diseases may be construed as tacit acknowledgement that an individual engages in sordid—and perhaps illegal—behaviors such as sex work, same-sex or extramarital relations, or intravenous drug use. These dynamics are widely recognized by public health practitioners. According to the U.S. Department of Health and Human Services, for example, HIV stigma can prevent people from learning their status, and discourage those at risk from seeking out HIV prevention tools and testing, and from talking openly with their sex partners about safer sex options (HIV.gov 2022). This leads to a vicious circle, in which those most at risk of infection are the same people least likely to educate themselves about prevention, thus increasing their chances of contracting the disease (Parker et al. 2002).

Stigma is particularly problematic in state-incorporated areas. On one hand, the consequences of exposing one's private behaviors are more severe. The same capacity for harm reduction, which states leverage for public good in the case of malaria, can also be used to target discrimination and violence. By definition, the coercive capacity of the state is concentrated in state-consolidated territory. Those living "inside" the state are thus more susceptible to a range of government actions, including harassment and abuse by security personnel, extortion, arrest, legal prosecution, deportation, and even death—all of which incentivize discretion and avoidance of the public health system (Davis 2017).[8] On the other hand, the probability that an innocuous interaction with the public health system results in discovery is elevated in state-incorporated territory. Even at private clinics, the mere proximity to the state raises the possibility that a person seeking HIV-related resources or support will be tagged as a sexual deviant or drug user and become a target of authorities.[9] The of severity of possible repercussions, coupled with the increased risks of exposure in state-incorporated areas may engender alienation from important health-related resources and promote riskier behavior.

To briefly summarize: I argue that one mechanism that may drive higher HIV rates in state-incorporated territory is stigma. Stigma can lead individuals to disengage from the public health system, particularly in countries in which affiliation with HIV, or with demographics that are highly-susceptible to HIV, may result in prosecution or other forms of formal sanctioning and harassment by governmental authorities. This disengagement, in turn, leads to lower levels of HIV-related knowledge and prophylaxis, which increases HIV transmission and incidence rates. This suggests the following hypothesis:

Hypothesis 2: *We should observe higher rates of HIV in state-incorporated territory, particularly in countries with high levels of institutionalized anti-HIV stigma or discrimination.*

8. Within Africa, Egypt and Sudan both allow for the deportation HIV-positive non-nationals; a number of other countries in the Middle East and Central Asia have similar policies on the books (UNAIDS-NCPI 2022).

9. Evidence from Thailand, for example, suggests that concerns about confidentiality and fear of being reported to authorities dissuades people from seeking out HIV-related information and treatment (Churcher 2013). Similarly, studies conducted in France, the U.S., and Denmark show that undocumented immigrants (another criminalized demographic) avoid the healthcare system entirely until health issues become critical because of their concerns of being reported to authorities (Hacker et al. 2015).

To test this hypothesis, I add a measure of HIV stigmatization to the interaction term in Equation 6.1:

$$\text{Incidence}_{ij} = \alpha + \beta_a \text{State}_{ij} + \beta_b \text{Year}_{ij} + \beta_c \text{Stigma}_{ij} + \beta_{ab}(\text{State} \times \text{Year})_{ij} + \beta_{ac}(\text{State} \times \text{Stigma})_{ij}$$
$$+ \beta_{bc}(\text{Year} \times \text{Stigma})_{ij} + \beta_{abc}(\text{State} \times \text{Stigma} \times \text{Year})_{ij} + \gamma X_{ij} + u_j + \epsilon_{ij} \quad (6.2)$$

The Stigma variable in Equation 6.2 is a composite measure of *official state discrimination* against HIV-positive or HIV-vulnerable individuals, based on an index proposed by Lyons et al. (2022). The index is the sum of ten binary indicators from the Global AIDS Monitoring (GAM) National Commitments and Policy Instrument (NCPI) published by UNAIDS, which are detailed in Table 6.2. Higher values represent greater levels of institutionalized stigma and discrimination. β_{ac} is the coefficient of interest; an estimate greater than zero indicates a positive relationship between HIV rates and stigma in state-incorporated areas. The vector X includes the same HIV-specific controls that I use in Equation 6.1: urban territory, nighttime light emissions, and rates of male circumcision.

Table 6.2. UNAIDS-NCPI Indicators of HIV Stigma & Discrimination

UNAIDS-NCPI Item	n	% of Countries
No Formal Legal Protections for People Living with HIV (PLHIV)	54	11.11%
Laws Criminalizing Transmission of, Non-Disclosure of, Exposure to HIV	54	83.33%
Anti-LGBT Laws - Morality	54	40.74%
Anti-LGBT Laws - General	54	7.41%
Anti-LGBT Laws - Promotion / Propaganda	54	7.41%
Arrest or Prosecution for Consensual Same Sex Intercourse	54	29.63%
Arrest or Prosecution of Drug Users	54	68.52%
Arrest or Prosecution of PLHIV for Transmission of / Non-Disclosure of / Exposure to HIV	54	1.85%
Arrest or Prosecution for Vertical Transmission	54	1.85%
Arrest or Prosecution for Commercial Sex Work	54	48.15%
Arrest or Prosecution of Transgender Individuals	54	9.26%
Restricted Entry / Stay / Residence of PLHIV	54	5.56%
HIV Test Required for at Least One Type of Entry Visa	54	7.41%

Note: Values listed as "unknown" or "undetermined" are coded as 0.

Table 6.6 in the Appendix provides estimates from Model 6.2. Consistent with Hypothesis 2, β_{ac} is positive and significant, indicating that HIV rates do tend to increase in state-incorporated areas as the level of official state stigma increases. Figure 6.3 depicts this relationship by plotting predicted HIV rates in state-incorporated and unincorporated territory across all levels of official state stigma.

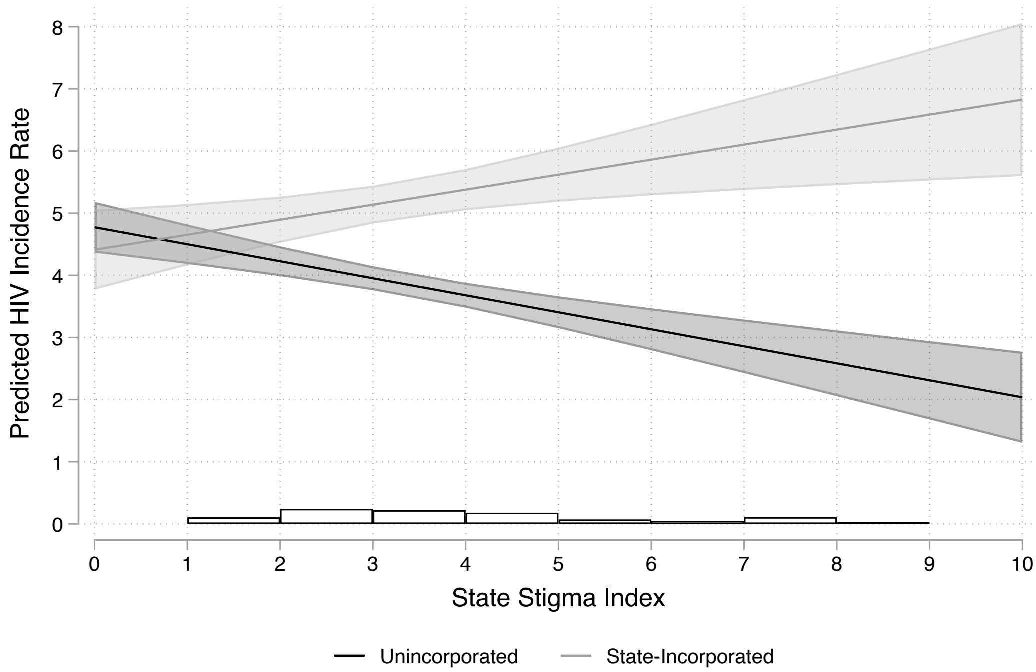


Figure 6.3. Predicted HIV incidence rates and 95% confidence intervals.

One unexpected finding from this analysis is that HIV rates and stigma exhibit a statistically significant inverse relationship (captured by a negative estimate of β_c) in unincorporated territory—HIV rates decrease in these areas as the level of state stigma increases. Although I conducted this analysis with no strong priors about how this relationship would play out in unincorporated territory, my baseline expectation was either a null result or a parallel upward trend. The effect of stigma on individual behavior should be exclusive to ($\beta_c = 0$), or at least more pronounced ($0 < \beta_c < \beta_{ac}$) in state-incorporated areas where the threat of official sanctioning is most acute.

One possible interpretation of this finding is that the stigma mechanism functions differently in incorporated regions than it does in unincorporated regions. As I argue above, those living inside of the state may be hesitant to interact with the public health system prior to HIV infection or testing, in order to minimize the possibility of exposing their private behavior (e.g., same-sex intercourse, intravenous drug use) to authorities. Individuals outside of the state, by contrast, may be more interested in concealing any possible *effects* of their private behavior from their immediate social circle. While consumption of HIV-related information may go unnoticed by both state authorities and the broader community in unincorporated areas, HIV is, in many ways, a highly-conspicuous disease. Untreated HIV infection results in visible wasting and recurrent secondary infections (e.g., pneumonia, thrush, shingles), and treatment requires a daily regimen of costly antiretroviral therapy—the expense (manifest in reduced household spending) and logistics (e.g., regular trips to a pharmacy or dispensary) of which are difficult to conceal from friends, family, and neighbors, especially in close-knit communities.

Given the high degree of social stigma attached to HIV in many African countries, it is reasonable to expect individuals—particularly those in unincorporated areas where reputation and other social pressures are more salient than the fear of sanctioning by government agents—to take active steps to reduce their risk of contracting HIV and other sexually transmitted infections, as testing positive for these diseases may reveal embarrassing information about an individual's private behavior. Using geocoded data on condom use provided by IHME, I am able to empirically test this proposition. I re-estimate Model 6.2, swapping the HIV incidence rate with a new outcome measure: the percentage of the population in a given cell that reports using a condom during their last sexual encounter. In this specification, the vector X includes controls for urban territory and nighttime lights, though I replace the male circumcision variable with the cell-specific marriage rate, as marriage is a more relevant consideration in the decision to use a condom than circumcision. I also control for the level of anti-HIV prejudice (a 10 point composite index based on the variables in Table 6.1), and the HIV incidence rate (as condom use tends to increase as the prevalence of HIV in a community increases).[10]

Full results of this specification are available in Table 6.7 in the Appendix. Figure 6.4 shows the effect of stigma by year and state incorporation status. Official stigma has a very clear positive effect on condom usage in unincorporated areas; a one point increase in official stigma increases the rate of condom usage by almost three-quarters of a percentage point on average ($\beta_c = 0.74$). The substantive (i.e., average conditional) effect is marginally lower at 0.72, though highly significant. This effect, however, disappears in state-incorporated territory, where I estimate an non-significant substantive effect of 0.19. In line with existing research, rates of condom use decrease where the marriage rate is high, and increase in wealthy and urban areas, and in regions where HIV prevalence is high. Condom

10. The HIV prejudice / social stigma variable was created from geocoded DHS data. Each respondent in the dataset was assigned a composite index value based on their responses to the DHS survey questions listed in Table 6.1. These data were then kriged (a form of spatial interpolation) to generate values for grid cells (the unit of analysis) that did not contain any DHS respondents.

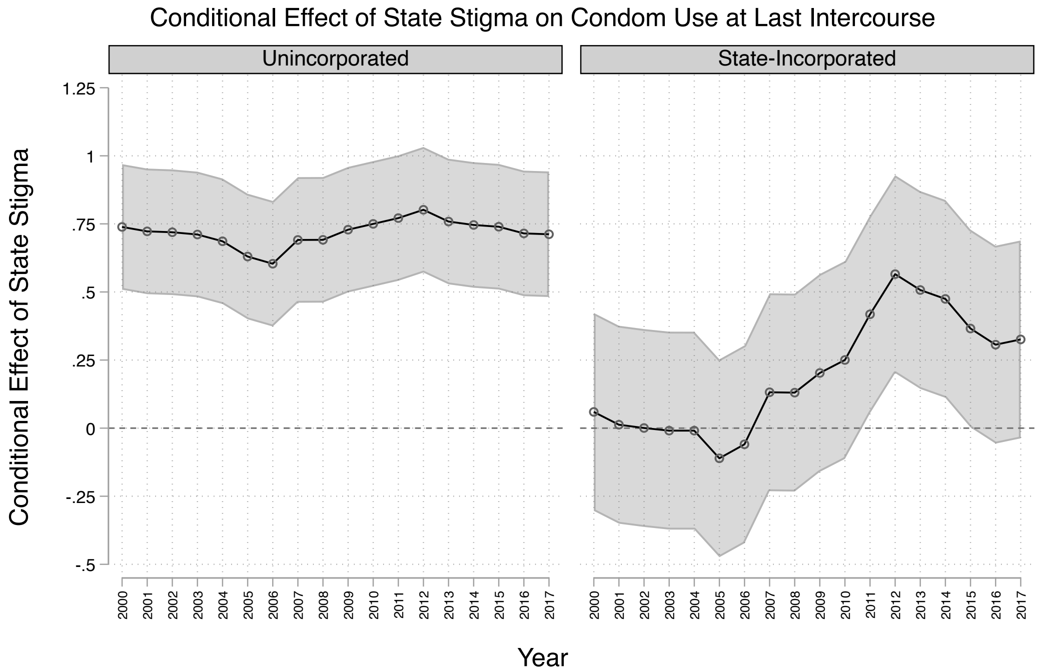


Figure 6.4. Conditional effects of state stigma and 95% confidence intervals, by year and state incorporation status.

use also tends to decrease as the level of anti-HIV prejudice in the community increases. This result is consistent with work by Peretti-Watel et al. (2007), who identify a strong negative relationship between anti–HIV discrimination in a person's social environment and condom use. Link and Phelan (2002) and Skinner and Mfecane (2004) posit that these high-risk sexual behaviors may result from a desire to appear HIV-negative during sexual encounters, as condom use may be construed as evidence of HIV infection.

We need to be careful in how we interpret these results, as condom use is a blunt metric of HIV risk mitigation. Aside from HIV prevention, condoms are used for family planning purposes, the prevention of other sexually transmitted infections, and as a matter of personal preference. Also, HIV transmission occurs in a variety of ways, including vertically (i.e., mother to child) and through needle reuse and sharing among persons who inject drugs (PWID). These results should therefore not be construed as direct evidence that anti-HIV stigma is motivating residents of unincorporated territory in Africa to take active steps to reduce their exposure to HIV. These results are, however, consistent with the story outlined above: If HIV infection reveals private information about an person's behavior to their immediate community, it is reasonable to take precautionary measures against infection. This fear of exposure outweighs the more immediate goal of appearing HIV-negative to sexual partners. This set of results is also consistent with the findings from Model 6.2, which show a significant reduction in the HIV incidence rate in unincorporated areas as official stigma increases. It is possible that rates of HIV transmission are reduced by more widespread condom use in areas with low state presence.

6.5 Anti-HIV Stigma and Avoidance of the Public Health System

In Section 6.3, I show that HIV rates are higher on average in state-incorporated territory than they are in unincorporated territory. This finding runs contrary to prevailing assumptions about the relationship between state capacity and public health—high-capacity states, or high-capacity areas within states, should exhibit lower rates of disease incidence, morbidity, and mortality than lower capacity states (Majeed and Gillani 2017; Knutsen and Kolvani 2022). In Section 6.4, I argue that one possible explanation for this finding is institutionalized anti-HIV stigma. In high capacity areas, stigma serves to alienate at-risk individuals from HIV-related resources, thus reducing basic knowledge about the disease, which increases high-risk behavior and thus transmission rates. Analyses in Section 6.4 seem to confirm at least part of this story—Model 6.2 uncovers a statistically significant positive correlation between stigma and the HIV incidence rate, and the subsequent analysis of condom use suggests that residents of state-incorporated territory do, in fact, engage in riskier sexual behaviors as the intensity of both official and community-based stigma increase.

If stigma is leading residents of state-incorporated territory to avoid the public health system, there are a couple of observable implications that might serve to further confirm or disconfirm this story. First, avoidance should result in lower levels of knowledge about HIV transmission and infection:

Hypothesis 3: *Individuals living in high-stigma state-incorporated territory will be less informed about HIV than residents of other territorial configurations.*

Although this information mechanism is explicit in the explanation I offer in Section 6.4, Hypothesis 3 is in many ways counterintuitive. The underlying assumption in much of the existing literature is that high-capacity states are better able to manage disease transmission by providing education and resources, and encouraging the adoption of mitigation behaviors such as condom use and needle exchange. Stigma, however, may attenuate these efforts. Skinner and Mfecane (2004), for example, argue that in the South African context, anti-HIV stigma impacts how people receive educational inputs. The taboos surrounding sex, gender identity, and vice promulgated by religion, traditional culture, and social mores make health education difficult under favorable circumstances; the addition of information about a stigmatized disease such as HIV-AIDS only complicates educational campaigns and may actually accelerate disengagement.

A second observable implication of this stigma-based explanation of HIV dynamics is that we should observe less frequent rates of HIV testing in these same high-stigma high-capacity areas where we observe lower levels of HIV-related knowledge:

Hypothesis 4: *Rates of HIV testing will be lower in high-stigma state-incorporated territory than it is elsewhere.*

Stigma is a well-documented barrier to voluntary HIV testing (Mahajan et al. 2008; Herek et al. 2003; Obermeyer and Osborn 2007; Pool et al. 2001; Kalichman and Simbayi 2003).[11] Nearly all of these behavioral epidemiology studies report confidentiality concerns and fear of how one's peers would react to a positive diagnosis as the primary barriers that prevent individuals from testing. It is worth noting that both of these factors are broadly congruent with the theory I outline above: concerns about confidentiality may indicate a fear of having one's private behaviors revealed to the state, while the anticipation of social shunning upon testing positive may be driven by the same fear of ex post exposure that I argue contributes to higher rates of condom usage in unincorporated territory.

To test these two hypotheses, I turn to cross-sectional data from the Demographic and Health Survey (DHS) on HIV-related knowledge and behavior. I use geocoded data from the most recent DHS men's and women's surveys available by country (generally DHS-V, DHS-VI, and DHS-7). The combined data are based on surveys conducted between 2003 and 2018; the full dataset includes 688,397 respondents across 33 African countries, though I lose about half of the sample due to listwise deletion. For Hypothesis 3, I assemble an index of HIV-related knowledge based on ten separate yes or no questions that gauge the respondent's knowledge of HIV transmission and morbidity. The

11. Ferree et al. (2021) find that similar dynamics play out in the case of COVID-19 testing in Malawi.

constituent variables are listed in Table 6.3. These variables are recoded such that accurate answers take on a value of 1, and inaccurate answers a value of 0.[12] The final item in the table, "Knows of a place to get an HIV test," is coded as 1 if the respondent answered affirmatively. The percentage of accurate responses for the total sample are listed in the last column of Table 6.3. Because not all items are included in each DHS country sample, the knowledge index is calculated as the overall percentage of correct responses:

$$\text{Knowledge Index} = \left(\frac{\text{Number of Accurate Responses}}{\text{Total Number of Items Answered}}\right) \times 100$$

On average, each respondent answered 8.89 questions ($\sigma = 1.38$). The mean accuracy rate is 73.73% ($n = 590618$, $\sigma = 23.14$). To assess testing behaviors for Hypothesis 4, I combine two individual survey items—a binary variable indicating whether or not the respondent has ever had an HIV test, and a continuous variable recording the number of months since the respondent's most recent HIV test—into a single binary variable that takes the value 1 if a respondent tested for HIV in the past two years, and 0 otherwise. I use GIS data on the location of each cluster in the DHS sample to assign individual respondents to a geographic location; these locations are then used to code the state incorporation variable.[13]

Table 6.3. Constituent Variables of HIV Knowledge Index

DHS Survey Item	Accuracy	n	Accurate Responses
A healthy-looking person can have HIV-AIDS	True	578021	79.24%
Contract HIV from saliva	False	283702	48.01%
Contract HIV from witchcraft or supernatural means	False	530134	78.49%
Contract HIV from mosquito bites	False	562224	66.04%
Contract HIV from sharing food with HIV-positive person	False	571734	77.96%
Always using a condom reduces the chance of contracting HIV	True	571760	77.22%
Drugs available to reduce HIV transmission to baby during pregnancy	True	485463	75.77%
HIV can be transmitted to child during delivery	True	585558	75.37%
HIV can be transmitted to child during breastfeeding	True	585552	75.98%
Knows of a place to get an HIV test	—	502250	78.33%

12. Responses of "don't know" / "not sure" / "it depends" are coded 0. All other responses, such as "I am HIV Positive," are coded as missing.

13. Although DHS uses a randomized displacement to mask the true location of respondents for privacy purposes, this error (2 km max in urban areas and 5 km max in rural areas) is much smaller than the 10,000 km^2 grid cell used as the the unit of analysis; the treatment variable should therefore have a high degree of fidelity in all but a few edge cases.

I test Hypotheses 3 and 4 using the following OLS/LPM specification with a factorial interaction term:

$$y_i = \alpha + \beta_1 \text{State}_i + \beta_2 \text{Stigma}_i + \beta_3 (\text{State} \times \text{Stigma})_i + \gamma X_i + \epsilon_i, \tag{6.3}$$

where the outcome, y_i, is either the knowledge index or the binary testing variable. The two treatment variables, State and Stigma, are the same variables used in each of the preceding analyses, though for ease of interpretation, I recode the Stigma variable as binary (values below the mean are coded 0, and 1 otherwise). X is a vector of controls that includes basic demographic characteristics such as the respondent's age, gender, marital status, education level, literacy, household wealth, total number of sexual partners in the past year (excluding the respondent's spouse), as well as a binary indicator of whether or not the respondent has any children (as many Africans encounter HIV-AIDS information and testing as a routine component of antenatal care). I also include a control for urban territory, the average HIV incidence rate in the respondent's grid cell, and a measure of HIV-related prejudice (see Table 6.1; this is the same variable used as a control in the analysis of condom use). All models are estimated using sampling weights, and standard errors are adjusted to account for the complex sampling protocols (cluster sampling of PSUs) employed by DHS.

Table 6.10 in the Appendix presents the full results of Model 6.3. Columns (1) and (2) detail parameter estimates for the knowledge index and binary testing outcome respectively. Column (3) provides estimates of a logit specification for the binary testing outcome, which I include as a robustness check. In order to better visualize the interaction effect, I plot predicted knowledge index values and predicted testing rates in Figures 6.5 and 6.6, respectively. Results are broadly consistent with Hypothesis 3. In Figure 6.5, we see that high levels of stigma are associated with a statistically significant reduction (1.86 percentage points on average) in HIV-related knowledge, though this reduction is more pronounced in unincorporated territory (-4.94 points) than it is incorporated territory (-0.88 points). These differential effects may be an indication that public health information is, in fact, diffusing more easily in state-incorporated territory; this would seem to confirm the conventional wisdom that education at least partially mediates reduced rates of communicable disease in high-capacity areas.

The predictive margins plotted in Figure 6.6 do not support Hypothesis 4. Results indicate

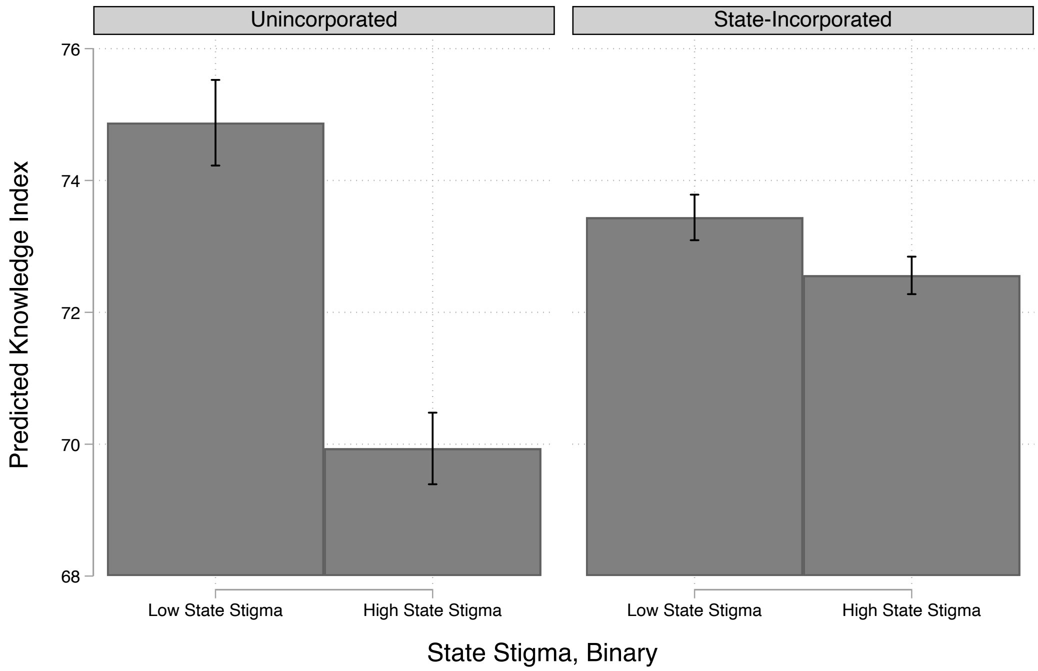


Figure 6.5. Predicted HIV knowledge by state incorporation status and stigma level.

a significantly higher rate of HIV testing in high-stigma state-incorporated areas as compared to low-stigma state-incorporated areas (an increase of 2.5 percentage points on average). This effect is reversed in unincorporated areas, where I estimate a 6.9 percentage point reduction in HIV testing in high-stigma areas. It is difficult to speculate on what is driving these divergent effects without more data, though it is possible that the negative effect of state stigma in unincorporated territory is related to this fear of social shunning or ostracism if a test comes out positive—the same fear I propose is driving the higher rates of condom usage unincorporated territory. The positive effect of stigma in state-controlled areas is more puzzling, especially because it is inconsistent with findings from previous studies of voluntary testing (though all of these studies use either a personal- or community-based measures of stigma, rather than a measure of institutional discrimination).

6.6 Discussion

The results presented in this paper paint a nuanced picture of the relationship between subnational state capacity and public health outcomes. On one hand, I find that that the hypothesized

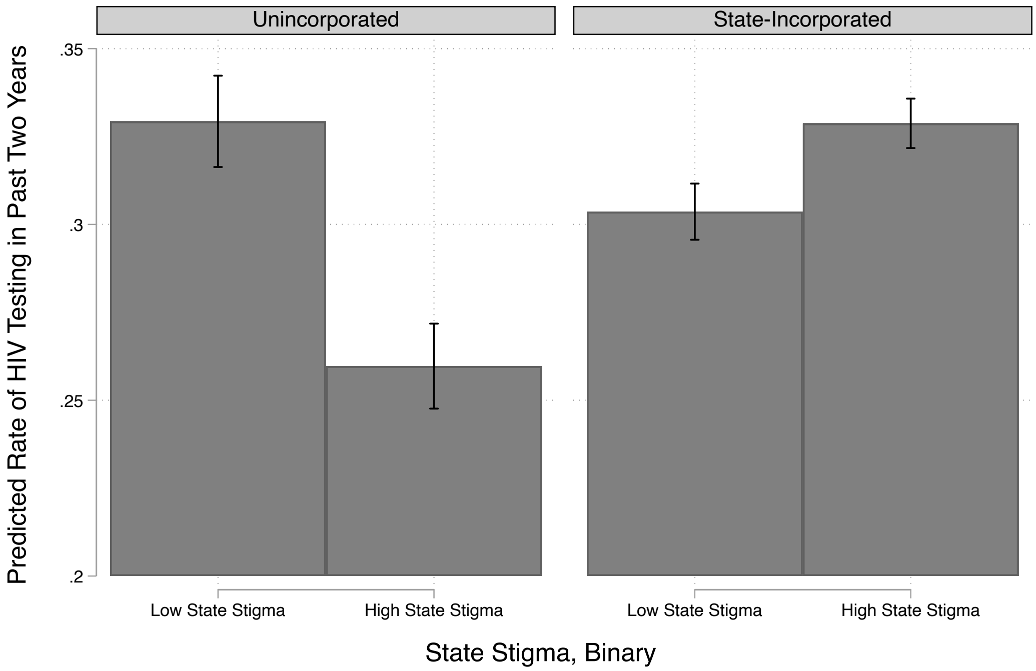


Figure 6.6. Predicted probability of an HIV test in the past 24 months (LPM estimates), by state incorporation status and stigma level.

inverse relationship between state capacity and disease rates holds, at least in the case of malaria. On the other hand, I find that the opposite relationship exists in the case of HIV. Because these two diseases are qualitatively different in terms of their epidemiology—one (HIV) engendering a high degree of stigma and discrimination that the other (malaria) does not, I posit that this stigma may be responsible for the increased prevalence of HIV in state-incorporated areas.

In Section 6.4, I argue that stigma may alter individuals' risk calculus in state-incorporated areas. Because the behaviors that tend to result in HIV infection—commercial sex work, same-sex relations, and intravenous drug use—tend to be criminalized, and because the probability that the state is able to detect and sanction these behaviors is higher in state-incorporated territory, residents of these areas will isolate themselves from the institutions responsible for public health promotion out of fear of revealing their ostracized behavior to the state. Due largely to the limitations of observational data, and the shortage of subnational data at the continent scale, I am only able to conduct indirect tests of this theory, though results of these tests provide suggestive evidence that the theory is plausible. I show, for example, that HIV-related knowledge and condom use rates tend to decrease in state-incorporated areas as stigma increases. These are findings we would expect if residents of these areas are insulated from public health messaging, as the theory stipulates. These findings are also consistent with the higher rates of HIV incidence I uncover in Section 6.3.

Admittedly, I also uncover findings that are more difficult to reconcile with this theory of stigma-driven avoidance. First, the negative effect of stigma on HIV-related knowledge is more pronounced in unincorporated areas than in state areas. And second, we see a divergent effect of state stigma on HIV testing in state and non-state areas. It is difficult to conjecture what factors may precipitate these results. Further research is needed to provide more direct tests of the theory, and to rule out alternative explanations (e.g., self-selection of risk-acceptant individuals into state-incorporated areas).

Overall, results in this chapter indicate that states, NGOs, and other advocates need to improve geographic-based targeting of disease prevention measures. For conventional diseases such as malaria, and perhaps influenza, cholera, tuberculosis as well, states must find ways to extend their mitigation programs beyond state-controlled territory. This obviously poses logistical challenges, though developments in information and communications technology may prove beneficial in these efforts. A recent quasi-experiment in Ghana, for example, finds that pushing malaria-related SMS messages encouraging preventative measures is associated with a 20.6% decrease in malaria rates in children in treated areas (Mohammed et al. 2019). For stigmatized diseases such as HIV and other sexually transmitted infections—though perhaps also mental health disorders and leprosy—states must take into consideration the ways in which stigma may affect their public health messaging, and perhaps take steps to improve medical confidentiality and eliminate discriminatory laws that target HIV risk groups.

6.7 Appendix

Table 6.4. Summary Statistics for Grid Cell Analyses

Variable	n	μ	σ	Min	Max
Mean HIV Rate	61,020	4.233	5.600	0.004	39.289
Mean Malaria Rate	74,298	0.210	0.187	0.000	0.653
Mean Condom Usage	61,020	11.881	12.645	0.346	79.366
Incorporation Status: State-Incorporated	85,767	0.310	0.463	0.000	1.000
Urban	85,749	0.463	0.499	0.000	1.000
Mean Nighttime Light Emissions	85,749	0.153	1.304	0.000	69.134
Mean Male Circumcision Rate	61,020	78.255	30.458	0.747	99.960
State Stigma	84,771	3.598	1.957	1.000	9.000
Mean Married / Partnered	61,020	61.595	11.784	14.147	92.246
Mean Community Stigma / HIV Prejudice	77,166	4.807	1.544	0.906	8.895
Malaria Suitability	84,605	3.881	3.878	0.000	12.000

Table 6.5. Effects of State Incorporation on Malaria & HIV Rates

	Malaria Rates		HIV Rates	
	(1)		(2)	
	Estimate	Std. Error	Estimate	Std. Error
Incorporation Status: State-Incorporated	-0.00253	(0.00609)	1.372***	(0.194)
2001	-0.00444***	(0.000556)	0.0643***	(0.00755)
2002	-0.0174***	(0.000783)	0.0999***	(0.0107)
2003	-0.0189***	(0.000955)	0.0805***	(0.0131)
2004	-0.0209***	(0.00110)	0.0402**	(0.0151)
2005	-0.0276***	(0.00122)	0.0139	(0.0168)
2006	-0.0344***	(0.00133)	-0.0405*	(0.0184)
2007	-0.0405***	(0.00144)	-0.125***	(0.0199)
2008	-0.0494***	(0.00153)	-0.203***	(0.0212)
2009	-0.0573***	(0.00161)	-0.283***	(0.0225)
2010	-0.0602***	(0.00170)	-0.324***	(0.0237)
2011	-0.0644***	(0.00177)	-0.308***	(0.0248)
2012	-0.0676***	(0.00184)	-0.390***	(0.0259)
2013	-0.0689***	(0.00191)	-0.407***	(0.0270)
2014	-0.0714***	(0.00197)	-0.437***	(0.0280)
2015	-0.0724***	(0.00204)	-0.445***	(0.0290)
2016	-0.0757***	(0.00209)	-0.490***	(0.0299)
2017	-0.0754***	(0.00215)	-0.534***	(0.0308)
2018	-0.0780***	(0.00220)		
2019	-0.0811***	(0.00226)		
2020	-0.0750***	(0.00231)		
State-Incorporated × 2001	0.00207*	(0.000974)	0.0338*	(0.0137)
State-Incorporated × 2002	0.00125	(0.00137)	0.0412*	(0.0194)
State-Incorporated × 2003	-0.00108	(0.00167)	0.0599*	(0.0237)
State-Incorporated × 2004	-0.00507**	(0.00192)	0.0504	(0.0274)
State-Incorporated × 2005	-0.0144***	(0.00214)	0.0239	(0.0306)
State-Incorporated × 2006	-0.0187***	(0.00234)	0.0278	(0.0335)
State-Incorporated × 2007	-0.0202***	(0.00252)	-0.0141	(0.0361)
State-Incorporated × 2008	-0.0213***	(0.00268)	-0.0573	(0.0386)
State-Incorporated × 2009	-0.0183***	(0.00283)	-0.0822*	(0.0409)
State-Incorporated × 2010	-0.0156***	(0.00297)	-0.118**	(0.0431)
State-Incorporated × 2011	-0.0198***	(0.00310)	-0.131**	(0.0451)
State-Incorporated × 2012	-0.0219***	(0.00323)	-0.120*	(0.0471)
State-Incorporated × 2013	-0.0234***	(0.00335)	-0.115*	(0.0490)
State-Incorporated × 2014	-0.0271***	(0.00346)	-0.103*	(0.0508)
State-Incorporated × 2015	-0.0304***	(0.00357)	-0.104*	(0.0526)
State-Incorporated × 2016	-0.0342***	(0.00367)	-0.0943	(0.0543)
State-Incorporated × 2017	-0.0367***	(0.00377)	-0.0650	(0.0559)
State-Incorporated × 2018	-0.0396***	(0.00386)		
State-Incorporated × 2019	-0.0401***	(0.00395)		
State-Incorporated × 2020	-0.0359***	(0.00404)		
Urban	0.0252***	(0.00541)	0.868***	(0.176)
Nighttime Light Emissions	-0.00535**	(0.00180)	0.192	(0.148)
Male Circumcision Rate			-0.0453***	(0.00132)
Malaria Suitability	0.0253***	(0.000616)		
Constant	0.145***	(0.00433)	7.162***	(0.151)
Random Effects (SD)				
Cell Intercepts	0.00000233***	(0.000000797)	0.0000421***	(0.0000152)
Residuals: AR(1)				
Residual	0.151***	(0.00155)	4.551***	(0.0583)
Rho	0.9839***	(0.00033)	0.9967***	(0.0000862)
N	73773		61020	
AIC	67916.74		-311677.30	
BIC	68295.53		-311235.30	

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 6.6. Effects of State Incorporation and Official Stigma on HIV Rates

	HIV Rates				
	Estimate	Std. Error		Estimate	Std. Error
Incorporation Status: State-Inc.	-1.280***	(0.364)	2001 × Stigma	-0.00362	(0.00466)
2001	0.0789***	(0.0179)	2002 × Stigma	-0.00562	(0.00658)
2002	0.123***	(0.0253)	2003 × Stigma	-0.00828	(0.00805)
2003	0.109***	(0.0309)	2004 × Stigma	-0.00949	(0.00928)
2004	0.0794*	(0.0357)	2005 × Stigma	0.0108	(0.0104)
2005	-0.0223	(0.0398)	2006 × Stigma	0.0103	(0.0113)
2006	-0.0765	(0.0436)	2007 × Stigma	0.0134	(0.0122)
2007	-0.174***	(0.0470)	2008 × Stigma	0.0144	(0.0131)
2008	-0.257***	(0.0502)	2009 × Stigma	0.0121	(0.0139)
2009	-0.332***	(0.0532)	2010 × Stigma	0.0174	(0.0146)
2010	-0.394***	(0.0560)	2011 × Stigma	0.00573	(0.0153)
2011	-0.340***	(0.0587)	2012 × Stigma	0.00831	(0.0160)
2012	-0.432***	(0.0613)	2013 × Stigma	-0.00194	(0.0166)
2013	-0.415***	(0.0637)	2014 × Stigma	-0.00352	(0.0172)
2014	-0.439***	(0.0661)	2015 × Stigma	-0.00649	(0.0178)
2015	-0.436***	(0.0683)	2016 × Stigma	0.0109	(0.0183)
2016	-0.537***	(0.0705)	2017 × Stigma	0.0122	(0.0189)
2017	-0.584***	(0.0726)	State-Inc. × 2001 × Stigma	-0.0103	(0.00870)
State-Inc. × 2001	0.0617	(0.0318)	State-Inc. × 2002 × Stigma	-0.0314*	(0.0123)
State-Inc. × 2002	0.137**	(0.0449)	State-Inc. × 2003 × Stigma	-0.0603***	(0.0150)
State-Inc. × 2003	0.248***	(0.0550)	State-Inc. × 2004 × Stigma	-0.0840***	(0.0174)
State-Inc. × 2004	0.303***	(0.0634)	State-Inc. × 2005 × Stigma	-0.109***	(0.0194)
State-Inc. × 2005	0.368***	(0.0708)	State-Inc. × 2006 × Stigma	-0.123***	(0.0212)
State-Inc. × 2006	0.415***	(0.0775)	State-Inc. × 2007 × Stigma	-0.151***	(0.0229)
State-Inc. × 2007	0.464***	(0.0837)	State-Inc. × 2008 × Stigma	-0.196***	(0.0244)
State-Inc. × 2008	0.569***	(0.0893)	State-Inc. × 2009 × Stigma	-0.248***	(0.0259)
State-Inc. × 2009	0.711***	(0.0947)	State-Inc. × 2010 × Stigma	-0.288***	(0.0273)
State-Inc. × 2010	0.804***	(0.0997)	State-Inc. × 2011 × Stigma	-0.304***	(0.0286)
State-Inc. × 2011	0.832***	(0.104)	State-Inc. × 2012 × Stigma	-0.330***	(0.0298)
State-Inc. × 2012	0.925***	(0.109)	State-Inc. × 2013 × Stigma	-0.340***	(0.0310)
State-Inc. × 2013	0.955***	(0.113)	State-Inc. × 2014 × Stigma	-0.358***	(0.0322)
State-Inc. × 2014	1.019***	(0.117)	State-Inc. × 2015 × Stigma	-0.359***	(0.0333)
State-Inc. × 2015	1.016***	(0.121)	State-Inc. × 2016 × Stigma	-0.374***	(0.0343)
State-Inc. × 2016	1.075***	(0.125)	State-Inc. × 2017 × Stigma	-0.370***	(0.0353)
State-Inc. × 2017	1.092***	(0.129)	Urban	1.017***	(0.162)
Stigma	-0.209***	(0.0522)	Nighttime Light Emissions	0.626**	(0.239)
State-Inc. × Stigma	0.664***	(0.0972)	Male Circumcision Rate	-0.0381***	(0.00140)
			Community Stigma / HIV Prejudice	-1.572***	(0.0512)
Continued in next 3 columns ⟶			Constant	14.92***	(0.313)
			Random Effects (SD)		
			Cell Intercepts	2.98e-09***	(1.20e-09)
			Residuals: AR(1)		
			Residual	4.017***	(0.0540)
			Rho	0.995***	(0.0001)
			N	55260	
			AIC	58811.72	
			BIC	59516.39	

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 6.7. Effects of State Incorporation and Official Stigma on Condom Use at Last Intercourse

	Condom Use				
	Estimate	Std. Error		Estimate	Std. Error
Incorporation Status: State-Inc.	1.738*	(0.815)	2001 × Stigma	-0.0103	(0.0419)
2001	0.172	(0.161)	2002 × Stigma	-0.00842	(0.0419)
2002	0.323*	(0.161)	2003 × Stigma	-0.0109	(0.0420)
2003	0.536***	(0.161)	2004 × Stigma	-0.0348	(0.0420)
2004	0.821***	(0.161)	2005 × Stigma	-0.0877*	(0.0420)
2005	1.290***	(0.161)	2006 × Stigma	-0.109**	(0.0420)
2006	1.734***	(0.161)	2007 × Stigma	-0.0163	(0.0420)
2007	1.593***	(0.161)	2008 × Stigma	-0.00608	(0.0420)
2008	1.880***	(0.161)	2009 × Stigma	0.0401	(0.0420)
2009	1.736***	(0.161)	2010 × Stigma	0.0730	(0.0421)
2010	1.688***	(0.161)	2011 × Stigma	0.0835*	(0.0420)
2011	1.832***	(0.161)	2012 × Stigma	0.110**	(0.0420)
2012	1.989***	(0.161)	2013 × Stigma	0.0733	(0.0420)
2013	2.329***	(0.161)	2014 × Stigma	0.0673	(0.0420)
2014	2.411***	(0.161)	2015 × Stigma	0.0648	(0.0421)
2015	2.333***	(0.161)	2016 × Stigma	0.0420	(0.0421)
2016	2.237***	(0.161)	2017 × Stigma	0.0361	(0.0421)
2017	2.263***	(0.161)	State-Inc. × 2001 × Stigma	-0.0271	(0.0784)
State-Inc. × 2001	0.146	(0.287)	State-Inc. × 2002 × Stigma	-0.0231	(0.0784)
State-Inc. × 2002	0.197	(0.287)	State-Inc. × 2003 × Stigma	-0.0150	(0.0784)
State-Inc. × 2003	0.368	(0.287)	State-Inc. × 2004 × Stigma	0.0127	(0.0784)
State-Inc. × 2004	0.494	(0.287)	State-Inc. × 2005 × Stigma	-0.0360	(0.0784)
State-Inc. × 2005	0.832**	(0.287)	State-Inc. × 2006 × Stigma	0.0551	(0.0784)
State-Inc. × 2006	0.712*	(0.287)	State-Inc. × 2007 × Stigma	0.154*	(0.0784)
State-Inc. × 2007	0.740**	(0.287)	State-Inc. × 2008 × Stigma	0.160*	(0.0785)
State-Inc. × 2008	1.022***	(0.287)	State-Inc. × 2009 × Stigma	0.186*	(0.0785)
State-Inc. × 2009	0.876**	(0.287)	State-Inc. × 2010 × Stigma	0.205**	(0.0785)
State-Inc. × 2010	0.965***	(0.287)	State-Inc. × 2011 × Stigma	0.339***	(0.0785)
State-Inc. × 2011	0.718*	(0.287)	State-Inc. × 2012 × Stigma	0.459***	(0.0785)
State-Inc. × 2012	0.491	(0.287)	State-Inc. × 2013 × Stigma	0.453***	(0.0785)
State-Inc. × 2013	0.723*	(0.287)	State-Inc. × 2014 × Stigma	0.440***	(0.0785)
State-Inc. × 2014	0.941**	(0.287)	State-Inc. × 2015 × Stigma	0.329***	(0.0785)
State-Inc. × 2015	1.209***	(0.287)	State-Inc. × 2016 × Stigma	0.298***	(0.0785)
State-Inc. × 2016	1.133***	(0.287)	State-Inc. × 2017 × Stigma	0.319***	(0.0785)
State-Inc. × 2017	0.977***	(0.287)	Urban	1.110**	(0.357)
Stigma	0.637***	(0.117)	Nighttime Light Emissions	2.311***	(0.530)
State-Inc. × Stigma	-0.651**	(0.218)	Married / Partnered	-0.142***	(0.00348)
			HIV Rate	0.343***	(0.0101)
			Community Stigma / HIV Prejudice	-2.974***	(0.111)
Continued in next 3 columns ⟶			Constant	28.79***	(0.728)
			Random Effects (SD)		
			Cell Intercepts	8.718***	(0.117)
			Residual	2.294***	(0.00712)
			N	55260	
			AIC	265839.90	
			BIC	266544.60	

Standard errors in parentheses.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 6.8. Summary Statistics for DHS Analyses — Continuous & Binary Variables

Variable	n	μ	σ	Min	Max
HIV Knowledge Index	590,618	73.733	23.149	0.000	100.000
Testing, Past Two Years	523,839	0.366	0.482	0.000	1.000
Incorporation Status: State-Incorporated	663,909	0.737	0.440	0.000	1.000
State Stigma: High State Stigma	663,909	0.498	0.500	0.000	1.000
Age	688,397	29.501	10.413	15.000	64.000
Male	688,397	0.295	0.456	0.000	1.000
Partnered or Married	683,246	0.350	0.477	0.000	1.000
Years of Education	688,173	5.769	4.750	0.000	27.000
$\geq$ 1 Sexual Partners Other than Spouse	502,754	0.033	0.177	0.000	1.000
One or More Children	681,775	0.689	0.463	0.000	1.000
Urban	688,397	0.374	0.484	0.000	1.000
Mean HIV Rate in Grid Cell	625,709	4.608	5.885	0.062	28.930
Community Stigma / HIV Prejudice	560,928	4.561	2.968	0.000	10.000

Table 6.9. Summary Statistics for DHS Analyses — Ordinal Variables

Variable	n	Percent
Literate		
Illiterate	251,171	36.81
Able to Read Part of a Sentence	68,874	10.09
Able to Read Full Sentence	362,369	53.1
Wealth Quintile		
1 (Bottom)	134,081	19.48
2	130,666	18.98
3	132,972	19.32
4	135,976	19.75
5 (Top)	154,702	22.47

Table 6.10. Effects of State Incorporation and Official Stigma on HIV Knowledge and Testing Rates

	Knowledge Index	Testing, Past 2 Years	Testing, Past 2 Years
		LPM	Logit (Odds Ratio)
	(1)	(2)	(3)
Incorporation Status: State-Incorporated	-1.436***	-0.0257***	0.873***
	(0.374)	(0.00716)	(0.0315)
State Stigma: High State Stigma	-4.940***	-0.0696***	0.693***
	(0.407)	(0.00807)	(0.0290)
State-Incorporated × High State Stigma	4.061***	0.0947***	1.622***
	(0.474)	(0.00951)	(0.0796)
Age	0.0817***	-0.00478***	0.976***
	(0.00513)	(0.000124)	(0.000615)
Male	-1.860***	-0.0705***	0.698***
	(0.125)	(0.00267)	(0.0103)
Partnered or Married	-0.914***	-0.0618***	0.742***
	(0.0926)	(0.00216)	(0.00885)
Years of Education	0.865***	0.0204***	1.104***
	(0.0164)	(0.000422)	(0.00229)
Literate : Able to Read Part of Sentence	3.063***	0.0319***	1.175***
	(0.172)	(0.00385)	(0.0230)
Literate: Able to Read Full Sentence	6.287***	0.0197***	1.116***
	(0.164)	(0.00354)	(0.0199)
Wealth Index (Quintile) = 2	1.251***	0.00206	1.012
	(0.194)	(0.00357)	(0.0189)
Wealth Index (Quintile) = 3	2.112***	0.00291	1.019
	(0.215)	(0.00392)	(0.0207)
Wealth Index (Quintile) = 4	3.203***	0.00163	1.009
	(0.226)	(0.00450)	(0.0234)
Wealth Index (Quintile) = 5	3.873***	-0.00674	0.968
	(0.246)	(0.00532)	(0.0261)
≥1 Sexual Partners Other than Spouse	-0.466**	0.0295***	1.177***
	(0.209)	(0.00586)	(0.0349)
One or More Children	4.589***	0.299***	4.474***
	(0.121)	(0.00314)	(0.0744)
Urban	-0.568***	-0.00249	0.983
	(0.164)	(0.00401)	(0.0197)
Mean HIV Rate in Grid Cell	0.304***	0.00998***	1.048***
	(0.00913)	(0.000299)	(0.00150)
Community Stigma / HIV Prejudice	-0.180***	-0.00197***	0.990***
	(0.00192)	(4.46e-05)	(0.000229)
Constant	68.85***	0.283***	0.360***
	(0.371)	(0.00717)	(0.0130)
Observations	398,465	344,642	344,642
R^2	0.240	0.155	

Standard errors in parentheses.
***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

6.8 Works Cited

Andrade, Mônica V., Kenya Noronha, Bernardo P. C. Diniz, Gilvan Guedes, Lucas R. Carvalho, Valéria A. Silva, Júlia A. Calazans, André S. Santos, Daniel N. Silva, and Marcia C. Castro. 2022. "The Economic Burden of Malaria: A Systematic Review." *Malaria Journal* 21 (1): 283.

Bollyky, Thomas J, Erin N Hulland, Ryan M Barber, James K Collins, Samantha Kiernan, Mark Moses, David M Pigott, Robert C Reiner Jr, Reed J D Sorensen, Cristiana Abbafati, Christopher Adolph, Adrien Allorant, Joanne O Amlag, Aleksandr Y Aravkin, Bree Bang-Jensen, Austin Carter, Rachel Castellano, Emma Castro, Suman Chakrabarti, Emily Combs, Xiaochen Dai, William James Dangel, Carolyn Dapper, Amanda Deen, Bruce B Duncan, Lucas Earl, Megan Erickson, Samuel B Ewald, Tatiana Fedosseeva, Alize J Ferrari, Abraham D Flaxman, Nancy Fullman, Emmanuela Gakidou, Bayan Galal, John Gallagher, John R Giles, Gaorui Guo, Jiawei He, Monika Helak, Bethany M Huntley, Bulat Idrisov, Casey Johanns, Kate E LeGrand, Ian D Letourneau, Akiaja Lindstrom, Emily Linebarger, Paulo A Lotufo, Rafael Lozano, Beatrice Magistro, Deborah Carvalho Malta, Johan Månsson, Ana M Mantilla Herrera, Fatima Marinho, Alemnesh H Mirkuzie, Ali H Mokdad, Lorenzo Monasta, Paulami Naik, Shuhei Nomura, James Kevin O'Halloran, Christopher M Odell, Latera Tesfaye Olana, Samuel M Ostroff, Maja Pasovic, Valeria Maria de Azeredo Passos, Louise Penberthy, Grace Reinke, Damian Francesco Santomauro, Maria Inês Schmidt, Aleksei Sholokhov, Emma Spurlock, Christopher E Troeger, Elena Varavikova, Anh T Vo, Theo Vos, Rebecca Walcott, Ally Walker, Simon D Wigley, Charles Shey Wiysonge, Nahom Alemseged Worku, Yifan Wu, Sarah Wulf Hanson, Peng Zheng, Simon I Hay, Christopher J L Murray, and Joseph L Dieleman. 2022. "Pandemic Preparedness and COVID-19: An Exploratory Analysis of Infection and Fatality Rates, and Contextual Factors Associated with Preparedness in 177 Countries, from Jan 1, 2020, to Sept 30, 2021." *The Lancet* 399 (10334): 1489–1512.

Börzel, Tanja A., and Thomas Risse. 2021. *Effective Governance Under Anarchy: Institutions, Legitimacy, and Social Trust in Areas of Limited Statehood.* Cambridge University Press.

Bosancianu, Constantin Manuel, Kim Yi Dionne, Hanno Hilbig, Macartan Humphreys, Sampada Kc, Nils Lieber, and Alex Scacco. 2020. *Political and Social Correlates of Covid-19 Mortality,* ub3zd.

Boussalis, Constantine, Hal T. Nelson, and Siddharth Swaminathan. 2012. "Towards Comprehensive Malaria Planning: The Effect of Government Capacity, Health Policy, and Land Use Variables on Malaria Incidence in India." *Social Science & Medicine* 75 (7): 1213–1221.

Carlitz, Ruth, and Ellen Lust. 2021. "Governance Beyond the State: Social Institutions and Service Delivery." In *The Oxford Handbook of the Quality of Government,* edited by Andreas Bågenholm, Monika Bauhr, Marcia Grimes, and Bo Rothstein, 0. Oxford University Press.

Case, Kelsey K, Gabriela B Gomez, and Timothy B Hallett. 2019. "The Impact, Cost and Cost-effectiveness of Oral Pre-exposure Prophylaxis in sub-Saharan Africa: A Scoping Review of Modelling Contributions and Way Forward." *Journal of the International AIDS Society* 22 (9): e25390.

CDC. 2020. *Anopheles Mosquitoes.* https://www.cdc.gov/malaria/about/biology/index.html.

Chima, Reginald Ikechukwu, Catherine A Goodman, and Anne Mills. 2003. "The Economic Impact of Malaria in Africa: A Critical Review of the Evidence." *Health Policy* 63 (1): 17–36.

Christensen, Tom, and Per Lægreid. 2020. "Balancing Governance Capacity and Legitimacy: How the Norwegian Government Handled the COVID-19 Crisis as a High Performer." *Public Administration Review* 80 (5): 774–779.

Christiansen-Jucht, Céline, Paul E. Parham, Adam Saddler, Jacob C. Koella, and María-Gloria Basáñez. 2014. "Temperature during Larval Development and Adult Maintenance Influences the Survival of Anopheles Gambiae s.s." *Parasites & Vectors* 7 (1): 489.

Churcher, Sian. 2013. "Stigma Related to HIV and AIDS as a Barrier to Accessing Health Care in Thailand: A Review of Recent Literature." *WHO South-East Asia Journal of Public Health* 2 (1): 12–22.

Cohen, Mitchell, Gustavo Guizzardi, Dominique Hausser, and Luc Van Campenhoudt, eds. 1997. *Sexual Interactions and HIV Risk: New Conceptual Perspectives in European Research.* London: Taylor & Francis.

D'Arcy, Michelle, and Marina Nistotskaya. 2017. "State First, Then Democracy: Using Cadastral Records to Explain Governmental Performance in Public Goods Provision." *Governance* 30 (2): 193–209.

Davis, Sara L. M. 2017. "The Uncounted: Politics of Data and Visibility in Global Health." *The International Journal of Human Rights* 21 (8): 1144–1163.

Dittmar, Jeremiah, and Ralf Meisenzahl. 2017. *State Capacity and Public Goods: Institutional Change, Human Capital, and Growth in Historic Germany.* CEPR Discussion Paper No. DP12037, 2968378, Rochester, NY.

Duffy, Patrick E. 2022. "Making a Good Malaria Vaccine Better." *Trends in Parasitology* 38 (1): 9–10.

Dwyer-Lindgren, Laura, Michael A. Cork, Amber Sligar, Krista M. Steuben, Kate F. Wilson, Naomi R. Provost, Benjamin K. Mayala, John D. VanderHeide, Michael L. Collison, Jason B. Hall, Molly H. Biehl, Austin Carter, Tahvi Frank, Dirk Douwes-Schultz, Roy Burstein, Daniel C. Casey, Aniruddha Deshpande, Lucas Earl, Charbel El Bcheraoui, Tamer H. Farag, Nathaniel J. Henry, Damaris Kinyoki, Laurie B. Marczak, Molly R. Nixon, Aaron Osgood-Zimmerman, David Pigott, Robert C. Reiner, Jennifer M. Ross, Lauren E. Schaeffer, David L. Smith, Nicole Davis Weaver, Kirsten E. Wiens, Jeffrey W. Eaton, Jessica E. Justman, Alex Opio, Benn Sartorius, Frank Tanser, Njeri Wabiri, Peter Piot, Christopher J. L. Murray, and Simon I. Hay. 2019. "Mapping HIV Prevalence in Sub-Saharan Africa between 2000 and 2017." *Nature* 570 (7760): 189–193.

Ferree, Karen E., Adam S. Harris, Boniface Dulani, Kristen Kao, Ellen Lust, and Erica Metheney. 2021. "Stigma, Trust, and Procedural Integrity: Covid-19 Testing in Malawi." *World Development* 141:105351.

Gizelis, Theodora-Ismene. 2009. "Wealth Alone Does Not Buy Health: Political Capacity, Democracy, and the Spread of AIDS." *Political Geography* 28 (2): 121–131.

Hacker, Karen, Maria Anies, Barbara L Folb, and Leah Zallman. 2015. "Barriers to Health Care for Undocumented Immigrants: A Literature Review." *Risk Management and Healthcare Policy* 8:175–183.

Hanson, Jonathan K., and Rachel Sigman. 2021. "Leviathan's Latent Dimensions: Measuring State Capacity for Comparative Political Research." *The Journal of Politics* 83 (4): 1495–1510.

Hendrickson, Cheryl, Lawrence C. Long, Craig van Rensburg, Cassidy W. Claassen, Mwansa Njelesani, Crispin Moyo, Lloyd Mulenga, Heidi O'Bra, Colin A. Russell, and Brooke E. Nichols. 2022. "The Early-Stage Comprehensive Costs of Routine PrEP Implementation and Scale-up in Zambia." *PLOS Global Public Health* 2 (11): e0001246.

Herek, Gregory M., John P. Capitanio, and Keith F. Widaman. 2003. "Stigma, Social Risk, and Health Policy: Public Attitudes toward HIV Surveillance Policies and the Social Construction of Illness." *Health Psychology* (US) 22:533–540.

Holmberg, Sören, and Bo Rothstein. 2011. "Dying of Corruption." *Health Economics, Policy and Law* 6 (4): 529–547.

Kalichman, S. C., and L. C. Simbayi. 2003. "HIV Testing Attitudes, AIDS Stigma, and Voluntary HIV Counselling and Testing in a Black Township in Cape Town, South Africa." *Sexually Transmitted Infections* 79 (6): 442–447.

Kharsany, Ayesha B.M., and Quarraisha A. Karim. 2016. "HIV Infection and AIDS in Sub-Saharan Africa: Current Status, Challenges and Opportunities." *The Open AIDS Journal* 10:34–48.

Knutsen, Carl Henrik, and Palina Kolvani. 2022. *Fighting the Disease or Manipulating the Data? Democracy, State Capacity, and the COVID-19 Pandemic.* SSRN Scholarly Paper, 4019437, Rochester, NY.

Kumsa, Diribe Makonene, and Gudina Terefe Tucho. 2019. "The Impact of Formal and Informal Institutions on ART Drug Adherence." *Journal of the International Association of Providers of AIDS Care* 18:2325958219845419.

Lee, Melissa M., Gregor Walter-Drop, and John Wiesel. 2014. "Taking the State (Back) Out? Statehood and the Delivery of Collective Goods." *Governance* 27 (4): 635–654.

Link, Bruce G., and Jo C. Phelan. 2002. "McKeown and the Idea That Social Conditions Are Fundamental Causes of Disease." *American Journal of Public Health* 92 (5): 730–732.

Lyons, Carrie, Victoria Bendaud, Christine Bourey, Taavi Erkkola, Ishwarya Ravichandran, Omar Syarif, Anne Stangl, Judy Chang, Laura Ferguson, Laura Nyblade, Joseph Amon, Alexandrina Iovita, Eglė Janušonytė, Pim Looze, Laurel Sprague, Keith Sabin, UNAIDS Task Team, Stefan Baral, and Sarah M. Murray. 2022. "Global Assessment of Existing HIV and Key Population Stigma Indicators: A Data Mapping Exercise to Inform Country-Level Stigma Measurement." *PLOS Medicine* 19 (2): e1003914.

Machault, Vanessa, Cécile Vignolles, Frédéric Pagès, Libasse Gadiaga, Yves M. Tourre, Abdoulaye Gaye, Cheikh Sokhna, Jean-François Trape, Jean-Pierre Lacaux, and Christophe Rogier. 2012. "Risk Mapping of Anopheles Gambiae s.l. Densities Using Remotely-Sensed Environmental and Meteorological Data in an Urban Area: Dakar, Senegal." *PLOS ONE* 7 (11): e50674.

Mahajan, Anish P., Jennifer N. Sayles, Vishal A. Patel, Robert H. Remien, Daniel Ortiz, Greg Szekeres, and Thomas J. Coates. 2008. "Stigma in the HIV/AIDS Epidemic: A Review of the Literature and Recommendations for the Way Forward." *AIDS* 22 (Suppl 2): S67–S79.

Majeed, Muhammad Tariq, and Seemab Gillani. 2017. "State Capacity and Health Outcomes: An Empirical Analysis." *Pakistan Journal of Commerce and Social Sciences (PJCSS)* 11 (2): 671–697.

Mandal, Sandip, Ram Rup Sarkar, and Somdatta Sinha. 2011. “Mathematical Models of Malaria - a Review.” *Malaria Journal* 10 (1): 202.

Mohammed, Aliyu, Princess Ruhama Acheampong, Easmon Otupiri, Francis Adjei Osei, Roderick Larson-Reindorf, and Ellis Owusu-Dabo. 2019. “Mobile Phone Short Message Service (SMS) as a Malaria Control Tool: A Quasi-Experimental Study.” *BMC Public Health* 19 (1): 1193.

Nketiah-Amponsah, Edward, Mohammed Abubakari, and Priscilla Twumasi Baffour. 2019. “Effect of HIV/AIDS on Economic Growth in Sub-Saharan Africa: Recent Evidence.” *International Advances in Economic Research* 25 (4): 469–480.

O’Meara, Wendy Prudhomme, Judith Nekesa Mangeni, Rick Steketee, and Brian Greenwood. 2010. “Changes in the Burden of Malaria in Sub-Saharan Africa.” *The Lancet Infectious Diseases* 10 (8): 545–555.

Obermeyer, Carla Makhlouf, and Michelle Osborn. 2007. “The Utilization of Testing and Counseling for HIV: A Review of the Social and Behavioral Evidence.” *American Journal of Public Health* 97 (10): 1762–1774.

Parker, Richard, Peter Aggleton, Kathy Attawell, Julie Pulerwitz, and Lisanne Brown. 2002. *HIV/AIDS-related Stigma and Discrimination: A Conceptual Framework and an Agenda for Action.* Horizons Report. Washington, DC: Population Council.

Peretti-Watel, Patrick, Bruno Spire, Yolande Obadia, Jean-Paul Moatti, and for the VESPA Group. 2007. “Discrimination against HIV-Infected People and the Spread of HIV: Some Evidence from France.” *PLOS ONE* 2 (5): e411.

Pool, R., S. Nyanzi, and J. A. G. Whitworth. 2001. “Attitudes to Voluntary Counselling and Testing for HIV among Pregnant Women in Rural South-West Uganda.” *AIDS Care* 13 (5): 605–615.

Post, Alison E., Vivian Bronsoler, and Lana Salman. 2017. “Hybrid Regimes for Local Public Goods Provision: A Framework for Analysis.” *Perspectives on Politics* 15 (4): 952–966.

Price-Smith, Andrew, Steven Tauber, and Anand Bhat. 2004. “State Capacity and HIV Incidence Reduction in the Developing World: Preliminary Empirical Evidence.” *Seton Hall Journal of Diplomacy and International Relations* 5:149.

Robert, Vincent, Kate Macintyre, Joseph Keating, Jean-François Trape, Jean-Bernard Duchemin, Mcwilson Warren, and John C. Beier. 2003. “Malaria Transmission in Urban Sub-Saharan Africa.” *American Journal of Tropical Medicine and Hygiene* 68 (2): 169–176.

Ross, Ronald. 1916. “An Application of the Theory of Probabilities to the Study of a Priori Pathometry.—Part I.” *Proceedings of the Royal Society of London,* Series A, 92 (638): 204–230.

Serikbayeva, Balzhan, Kanat Abdulla, and Yessengali Oskenbayev. 2021. “State Capacity in Responding to COVID-19.” *International Journal of Public Administration* 44 (11-12): 920–930.

Skinner, D., and S. Mfecane. 2004. “Stigma, Discrimination and the Implications for People Living with HIV / AIDS in South Africa : Original Article.” *SAHARA : Journal of Social Aspects of HIV / AIDS Research Alliance* 1 (3): 157–164.

Smith, M. W., T. Willis, L. Alfieri, W. H. M. James, M. A. Trigg, D. Yamazaki, A. J. Hardy, B. Bisselink, A. De Roo, M. G. Macklin, and C. J. Thomas. 2020. “Incorporating Hydrology into Climate Suitability

Models Changes Projections of Malaria Transmission in Africa." *Nature Communications* 11 (1): 4353.

Stockemer, Daniel, and Bernadette Lamontagne. 2007. "HIV/AIDS in Africa: Explaining the Differences in HIV Prevalence Rates." *Contemporary Politics* 13 (4): 365–378.

Swaminathan, Siddharth, and John Thomas. 2012. "The Politics of Births in India." In *The Performance of Nations,* 245–263. Plymouth, UK: Rowman & Littlefield Publishers.

Szabo, Robert, and Roger V. Short. 2000. "How Does Male Circumcision Protect against HIV Infection?" *BMJ* 320 (7249): 1592–1594.

Teshale, Achamyeleh Birhanu, and Getayeneh Antehunegn Tesema. 2022. "Discriminatory Attitude towards People Living with HIV/AIDS and Its Associated Factors among Adult Population in 15 Sub-Saharan African Nations." *PLOS ONE* 17 (2): e0261978.

UNAIDS, Joint United Nations Programme on HIV/AIDS. 2013. *Access to Antiretroviral Therapy in Africa: Status Report on Progress towards the 2015 Targets.* Technical report. Geneva.

———. 2022. *In Danger: UNAIDS Global AIDS Update 2022.* Technical report. Geneva.

Villena, Oswaldo C., Sadie J. Ryan, Courtney C. Murdock, and Leah R. Johnson. 2022. "Temperature Impacts the Environmental Suitability for Malaria Transmission by Anopheles Gambiae and Anopheles Stephensi." *Ecology* 103 (8): e3685.

WHO, World Health Organization. 2020. *Report on Antimalarial Drug Efficacy, Resistance and Response: 10 Years of Surveillance (2010-2019).* Technical report. Geneva.

———. 2022. *World Malaria Report 2022.* World Health Organization.

Yen, Wei-Ting, Li-Yin Liu, Eunji Won, and Testriono. 2022. "The Imperative of State Capacity in Public Health Crisis: Asia's Early COVID-19 Policy Responses." *Governance* 35 (3): 777–798.

Chapter 7

Conclusion

This dissertation grapples with two core questions in the subfield of comparative politics: First, over which territories within a state's de jure borders do central governments exercise de facto control? And second, is life different for those individuals living inside of state-controlled territories compared to those living outside of the state? I examine these questions in the African context—a region that is notorious for incomplete state consolidation (Jackson and Rosberg 1982), and in which scholars have documented subnational geographic disparities in outcomes ranging from economic development to public health (Iddawela et al. 2021; Yourkavitch et al. 2018).

In the first half of the dissertation, I introduce a new measure of territorial control based on publicly available GIS data on the location of official state-related infrastructure across the African continent. Mapping this measure allows us to visualize both within-country and across-country variation in state control throughout the African continent. Although I argue that that this is a valid measure of control, further research should seek to "ground truth" this measure by (for instance) validating the extent of state-incorporated and unincorporated areas with the assessments of area specialists and other experts, or with large-n survey data from African residents who are best equipped to report their own incorporation status.

Descriptive and correlational analyses of this measure in Chapters 2 and 3 confirm much of the received wisdom on state building in Africa, much of which stems from the seminal work of Herbst (2014). In Chapter 3, for example, I show that control tends to correlate with population density and economic production, though I do uncover one unexpected finding: control seems to be positively associated with rugged terrain. This finding is inconsistent with much of the extant civil war literature (e.g., Fearon and Laitin 2003), though it does echo results from Buhaug and Rød (2006) and Buhaug and Lujala (2005) that indicate a negative correlation between conflict onset and rugged or mountainous terrain.[1] I explain this contradictory finding by pointing to the human geography of the African continent. Acemoglu et al. (2001) note that human settlements in Africa tend to be at higher elevations due to the continent's harsh climates and the threat of malaria-carrying mosquitos, and most of the continent's transportation infrastructure tends to be located at lower elevations along the coasts. We should therefore expect control to be concentrated in these high- and low-elevation regions. In Figure 3.2, I show this to be generally true.

In Chapter 4, I extend the analyses from Chapter 3 to account for the spatial dimensions of territorial control. Chapter 4 sketches out a theory of state consolidation in which leaders attempt to grow the state into areas that are economically or strategically salient to the central government. Unfortunately, the data available for this dissertation do not allow me to test this theory directly. This is, in fact, the most significant limitation for the analyses conducted in this dissertation—it is not possible to assess whether the state whether African states have asserted control over particular regions because of their economic or strategic value, or whether these areas have become more valuable because they enjoy the benefits of state control. This is clearly one important direction for further research. Because GIS data on state infrastructure has become easier to acquire since the beginning of this project, it will be possible in the future to assemble a time-series measure of territorial control which might be leveraged for causal studies of state expansion.

Despite the limitations of the cross-sectional territorial control measure I employ throughout this dissertation, in Chapter 4, I am able to corroborate some of the naïve correlations uncovered in Chapter 3: geographic patterns of territorial control across Africa are highly correlated with population and economic activity. This chapter also uncovers various new findings that have not been widely articulated in the literature. The first is that control tends to cluster in space. In other words, areas with high levels of state control are not randomly distributed throughout a country's territory. While I am careful to avoid causal claims, this pattern is consistent with the theory I propose in Chapter 4, that states will attempt to expand the geographic scope of their control through a process of accretion, incorporating previously uncontrolled territory proximate to areas of state control before

1. The civil war literature (e.g., Fearon and Laitin 2003) posits a *positive* relationship between rugged terrain and conflict, as rugged terrain is more difficult for the center to control, and thus represents a "softer" target for rebel groups and other belligerents.

more distant territories. This finding deserves closer attention in future research.

The second novel finding that Chapter 4 underscores is the uneven expansion of African states throughout their de jure borders. Using a series of random forest models, I am able to approximate the "ecological niche" of the Africa state—areas which are broadly amenable to state control. I find, however, that many of the areas that fall within this ecological niche are completely devoid of state infrastructure. This is puzzling, as states are generally understood to maximize control. This pattern raises questions about the state's cost constraints, which may limit its expansion into these favorable areas, but also questions about non-physical, or technological modes of radiating control into surrounding territory.

The second half of the dissertation engages with the question of whether life differs for those who live "inside" versus "outside" of the state. The answer, suggested by Chapters 5 and 6, seems to be a nuanced yes. In Chapter 5, I show that there are differences in how residents *perceive* traditional authorities in incorporated and unincorporated areas—people that live in regions of high state control tend to have more negative perceptions of traditional authorities than those living in regions with low state control. This is not necessarily true, however, when it comes to the actual *influence* of traditional authorities. There is no apparent difference in the degree to which traditional authorities exercise governance functions inside and outside of the state. This raises certain questions about whether states may collaborate with or co-opt traditional authorities within state-controlled areas in order to reduce the costs governance. Perhaps the most surprising finding from Chapter 5 is that residents of state-incorporated territory have more negative perceptions of their local councillors and members of parliament than their counterparts in unincorporated territory; one avenue for future research is to probe the reasons for this particular disparity.

Chapter 7 provides another clear example of how life differs in areas under state control. In the context of public health, we see very clear differences in the rates of common endemic diseases such as malaria and HIV. HIV is much more prevalent in incorporated areas, while the opposite is true in the case of malaria. These divergent patterns of disease prevalence are surprising, as the literature suggests that areas of high state capacity should exhibit lower rates of disease (Serikbayeva et al. 2021). I posit that these higher rates of HIV in state-incorporated areas may be the result of anti-HIV stigma. Stigma may reduce an individual's propensity to seek out HIV-related information or prophylactics, and this effect may be amplified in state-controlled, areas where the fear of discovery of and potential prosecution for stigmatized high-risk behaviors (e.g., commercial sex work, intravenous drug use) are acute. Preliminary tests suggest some support for this stigma hypothesis, though future research is certainly warranted to explore the exact mechanism at play.

Overall, this dissertation proposes a new subnational measure of territorial control, and uses this measure to address the question “Where is the state in Africa?” This measure allows me to map spatial variation in territorial control across Africa with a level of granularity that is uncommon in the existing literature. I show that this measure correlates with common proxies of state capacity and control, and demonstrate the utility of this measure in assessing disparities in governance and public health outcomes within states. Further development of this measure (e.g., expanding to a time series) opens up exciting possibilities for future research, including a more rigorous exploration of the expansion and contraction of the African state.

7.1 Works Cited

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review* 91 (5): 1369–1401.

Buhaug, Halvard, and Päivi Lujala. 2005. “Accounting for Scale: Measuring Geography in Quantitative Studies of Civil War.” *Political Geography* 24 (4): 399–418.

Buhaug, Halvard, and Jan Ketil Rød. 2006. “Local Determinants of African Civil Wars, 1970–2001.” *Political Geography* 25 (3): 315–335.

Fearon, James D., and David D. Laitin. 2003. “Ethnicity, Insurgency, and Civil War.” *American Political Science Review* 97 (1): 75–90.

Herbst, Jeffrey. 2014. *States and Power in Africa: Comparative Lessons in Authority and Control - Second Edition.* Princeton University Press.

Iddawela, Yohan, Neil Lee, and Andrés Rodríguez-Pose. 2021. “Quality of Sub-national Government and Regional Development in Africa.” *The Journal of Development Studies* 57 (8): 1282–1302.

Jackson, Robert H., and Carl G. Rosberg. 1982. “Why Africa’s Weak States Persist: The Empirical and the Juridical in Statehood.” *World Politics* 35 (1): 1–24.

Serikbayeva, Balzhan, Kanat Abdulla, and Yessengali Oskenbayev. 2021. “State Capacity in Responding to COVID-19.” *International Journal of Public Administration* 44 (11-12): 920–930.

Yourkavitch, Jennifer, Clara Burgert-Brucker, Shireen Assaf, and Stephen Delgado. 2018. “Using Geographical Analysis to Identify Child Health Inequality in Sub-Saharan Africa.” *PLOS ONE* 13 (8): e0201870.