2021

Birth registration and educational access in Sub-Saharan Africa: The case for an explanatory spatial research design

Thomas Edward Griffin
University of Vermont

BIRTH REGISTRATION AND EDUCATIONAL ACCESS IN SUB-SAHARAN AFRICA: THE CASE FOR AN EXPLANATORY SPATIAL RESEARCH DESIGN

A Dissertation Presented

by

Thomas E. Griffin

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Educational Leadership and Policy Studies

May, 2021

Defense Date: March 19, 2021
Dissertation Examination Committee:

Bernice R. Garnett, Sc.D., Advisor
Kelsey Gleason, Sc.D., Chairperson
Jason C. Garvey, Ph.D.
Jessica S. Strolin-Goltzman, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

In 2019, the United Nations Children's Fund (UNICEF) estimated the global number of children under the age of 5 without birth registrations at 166 million, with the largest share being present in Sub-Saharan Africa. As the author witnessed firsthand while working in Cameroon, the lack of birth registration documentation (i.e. birth certificates) precluded students from progressing from primary to secondary education. Struck by this example of social exclusion, the purpose of this study was to examine the extent to which birth registration acted as a barrier to educational access in primary and secondary education systems elsewhere across Sub-Saharan Africa. An interdisciplinary conceptual framework revealed a gap in academic literature with only a few studies having explored the relationship between birth registration and access to education in a regional context. This study filled such gap by advancing an innovative *explanatory spatial* mixed methods research design to analyze secondary data from UNICEF and the United Nations Educational, Scientific, and Cultural Organization (UNESCO). This unique design consisted of an initial quantitative multiple regression analysis followed by a spatial autocorrelation analysis, using geographic information systems (GIS), to explain the geography of the initial results.

Results from this pragmatic research approach, outlined in a journal-article dissertation format, were intended to be made useful for researchers and policymakers alike. Noteworthy for the former audience, the quantitative strand found that while birth registration was not a significant predictor of access to education at any level of schooling, there were significant effects of gross domestic product (GDP) per capita and rurality on educational access (Article #1). For the latter readers, choropleth maps of birth registration revealed some areas of neighboring countries with similar levels of low registration despite the absence of statistical clustering. However, access to education demonstrated statistically significant cluster patterns ($p<0.05$) at the primary and lower-secondary levels, offering organizations like UNICEF and UNESCO noteworthy findings that could better inform policy interventions (Article #2). Finally, the author integrated both data strands using a multivariate cluster analysis in the ArcGIS platform, providing a compelling argument for the use of spatial mixed methods in educational policy research (Article #3).

DEDICATION

I dedicate this dissertation to my Cameroonian family: M. Denis, M. Boubakary, Mme. Fatimat, and most importantly, the children of Garoua.

Je dédie ce thèse à ma famille Camerounaise: M. Denis, M. Boubakary, Mme. Fatimat, et surtout, les enfants de Garoua.

ACKNOWLEDGEMENTS

This dissertation was just one component of a five-year long doctoral journey. There are so many wonderful people that have supported me along the way with whom I would like to share this accomplishment. After all, this celebration is far more a reflection of that support network than any individual effort on my part and I would like to offer you all my most sincere gratitude.

For their unwavering and endless support from Little Rock, I first thank my parents, whom I love so dearly, Dr. Robert Griffin and Mrs. Ann Griffin. Likewise, for making Franklin and Amherst my go-to places of R&R throughout this process, I have to thank my New Hampshire parents, Mr. Eric Hahn and Mrs. Jennifer Hahn.

For bringing me into the UVM team and encouraging me throughout this phase of my lifelong learning adventure, I would like to thank Dean Cynthia Belliveau. Furthermore, I thank my colleagues and great friends at UVM Continuing & Distance Education (CDE) for their positive energy and camaraderie along the way.

For running the show in Bravo Company, 403rd Civil Affairs Battalion, and being my much-needed voice of reason, I have to acknowledge the Army's most outstanding Non-Commissioned Officer, First Sergeant Dennis Clark. Thank you, battle.

For their wisdom and ability to get the best out of me, I would like to thank my dissertation committee, and, in particular, my remarkable advisor, Dr. Bernice Garnett.

I want to express the highest degree of appreciation for my dear friend and colleague, Associate Dean Beth Taylor-Nolan. I seriously could not have endured these last five years without your side-by-side empathy and friendship. I look forward to celebrating our accomplishments with Guinness, Franklin, and Charlie this Summer.

Finally, to the person I look up to most in life and the leader whom I aspire to become, I thank my amazing wife and best friend, Captain Sarah Griffin. I am truly blessed to share a life with someone that gives me so much strength, energy, love, and the perfect balance of tough-love and compassion. She has endured the long nights, the lost weekends, the deployment, and so much more in support of this goal and I will be forever grateful. I love you, Sally.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1. Setting

Imagine for a moment that you are a 13-year old student in Cameroon's austere Far North Region looking to escape an area plagued with violence from the Boko Haram conflict. You believe that an education will be your ticket to prosperity, so you work tirelessly during primary school and are one of the highest achievers in your class. Then, as you try to sit for the primary school exit exam, or the Certificat d'Études Primaires, you are sent home because you do not possess a birth certificate (Civil Registration Centre for Development, 2011; Kindzeka, 2017). Your opportunity to continue formal schooling evaporates. Sadly, vulnerable populations beyond the conflict-affected region of Northern Cameroon also face scenarios like this. Consider the case of an interviewee in Crush and Tawodzera's (2014) study, who was unable to enroll her daughter in school after fleeing to South Africa from Zimbabwe. Like many other asylum seekers, she had been unable to secure a birth certificate in Zimbabwe prior to her departure from the country's crisis conditions. Then, recognizing her daughter's diminishing pathway to further education without such document, she admitted that she could not afford to return for one even if she was willing to risk her safety to do so. In both cases, the cost of bureaucratic exclusion from the education system was the future of a vulnerable child.

1.2. Problem Statement

According to the United Nations Children's Fund (UNICEF), an estimated 237 million children under age 5 do not have a birth certificate worldwide and one in four are not registered (Selim, 2019). It is widely recognized that a registered birth and the

possession of corresponding birth certificate documentation establishes the basis for an individual's officially recorded identity, which then connects them to a number of legal rights and social benefits (Appell, 2014; Dunning et al., 2014; Hunter, 2019). Pelowski et al. (2016) argue that these processes impose major challenges in developing countries, where low registration levels have been attributed to a lack of administrative infrastructure and a variety of sociocultural factors. It is problematic when access to basic services, ranging from education to healthcare, is contingent upon birth registration and/or the possession of a birth certificate (Brewer et al., 2015). This is especially the case when the state fails to provide a functioning civil registration system or deliberately excludes some of its inhabitants from acquiring their rightful legal identity, as was the case in many African countries throughout the 20th century (International Institute for Vital Registration and Statistics, 1988). In other words, birth registration systems emplaced by colonizers decades ago still act as a barrier to legal identity formation and its subsequent benefits (Johnston et al., 2000).

Globally, birth certificates often act as mechanisms for social exclusion, denying children the legal documentation necessary to access state benefits and services (Chereni, 2017). With their births unregistered by the state, millions of "missing" children remain "hidden in plain sight," or bureaucratically invisible from state institutions (Ajmera & Fields, 2016; Sumner, 2015). Education is one sector of society where exclusion, via missing birth certificates, is particularly problematic, as it can prevent children from entering school systems, progressing between levels, or from exiting with a credential (Hunter, 2019). This was evident in the Dominican Republic, where Corbacho et al. (2012) found that the absence of birth registrations translated to fewer years of

educational attainment. It has been unclear the degree to which cases like this and the introductory examples from Cameroon and South Africa are reproduced at scale. Without this understanding, one could reasonably assume that the lack of birth certificates could be a factor in the United Nations Educational, Scientific, and Cultural Organization (UNESCO)'s estimation of over 260 million children that were out of school in 2016 (United Nations Educational, Scientific and Cultural Organization, 2018). Consequently, disenfranchised youth, excluded from society en masse, could plausibly offer a recipe for social unrest or violent conflict, particularly if concentrated in geographical areas already strained by weak governance, post-colonial tension, or large degrees of human migration.

1.3. Purpose Statement

The purpose of this study was to examine the extent to which birth registration hinders educational access in primary and secondary education systems across Sub-Saharan Africa. This region is home to a number of conflict-affected fragile states, or those impacted by conflict, income disparity, political instability, and/or on-going state fragility (Dryden-Peterson, 2009). These countries' education systems are susceptible to a range of institutional barriers to quality schooling, from individual- and group-level characteristics, like nationality (Buckland, 2011) or gender (Lloyd et al., 2005) to systemic inequalities in curriculum and pedagogy (Dryden-Peterson, 2009). Throughout the available scholarly literature, an indicator of birth registration is a notable absence from this list of barriers to educational access – an omission I intended to inspect through empirical research. Modifying one of Creswell and Plano Clark's (2018) mixed methods research designs, I employed an *explanatory spatial* study design. In this technique, I

analyzed quantitative data first before explaining the results with a cross-country spatial analysis. In the first phase, I employed quantitative data, obtained through the UNICEF and UNESCO data warehouses, to measure the relationship between birth registration and educational access. In the second phase, I investigated each variable independently, using a spatial analysis method, to determine if there were any geographic patterns in their distribution across the study area.

I was particularly interested in the process of addressing these variables through the employment of both statistical and spatial relationships. While quantitative methods can illuminate the strength of relationships between variables, spatiality is a tremendously valuable approach for understanding if there are relationships in how they are distributed (McGrew Jr. et al., 2014; Starr, 2003). Additionally, a regional analysis can sometimes uncover spatial relationships better than a global sample because large-N studies often ignore spatial variations (Østby et al., 2009). As such, my study uses Sub-Saharan Africa as the sample region to explain the factors connecting birth registration and educational access. Comparatively, Africa stands out among other regions in both of these variables of interest within this study. Sub-Saharan Africa accounts for nearly 60% all children under the age of 5 whose births are not registered, which is higher than any other region (United Nations Children's Fund, 2019e). The region also lags behind others in literacy rates and net attendance rate of primary school (Roser & Ortiz-Ospina, 2016). Given that spatial dependence of social phenomena, or clustering, has been shown to be strongly present in Africa, the spatial context could enable researchers, practitioners, and policymakers to better understand how the study's variables pertain to inter- and intra-state distributions (O'loughlin & Anselin, 1991).

1.4. Significance

This study contributed to academic literature from both a topical and a methodological standpoint. On the topics of birth registration, legal identity formation, and education, it builds on the only two studies examining these interconnected relationships in depth. Corbacho et al. (2012) revealed the impact of birth registration on education attainment in the Dominican Republic and Apland et al. (2014) conducted a mixed methods analysis of birth registration and access to services (education and health) across cases in Sierra Leone, Vietnam, Kenya, and India. Influenced by these quantitative analyses, my study expanded the geographical scope to explore the possibility of regional trends linking these variables. Even similar studies focusing on birth certificate rates and social exclusion in Africa arose primarily from South Africa (Hlatshwayo & Vally, 2014; Jewkes & Wood, 1998; Porteus et al., 2000) or Nigeria (Akomolafe et al., 2019; Anaduaka, 2020; Makinde et al., 2016). I sought to be the first to examine the intersection of legal identity and education by combining the quantitative approaches found within many of these studies with a geographic information systems (GIS) analysis technique. In doing so, I added a rigorous cross-continental analysis to the slowly developing body of interdisciplinary literature advancing spatial mixed methods in empirical research (Frels et al., 2011; Harbers & Ingram, 2017).

The stakes of this research are high. Inequalities in the provision of education and barriers to accessing the education system are both considered to be factors fueling conflict or perpetuating it in conflict-affected fragile states (Burde et al., 2017; Dryden-Peterson, 2009). Across the world, examples of a complex relationship between education and conflict are aplenty (Brock, 2011; Davies, 2003; King, 2011). Østby et al.

(2019) identified 42 quantitative studies between 1996 and 2016 that tested the relationship between various educational factors, as independent variables, and political violence or conflict measures, as dependent variables. Many of these studies analyzed global samples (Elbakidze & Jin, 2015; Hegre et al., 2013; Lange, 2012). The problem with this macro-level approach to investigating the education-conflict relationship is that it prioritizes generalizable findings over sub-regional conclusions, which are likely full of nuance. Additionally, only one of the studies (Barakat & Urdal, 2009) focused on an aggregate educational exclusion measure (low educational attainment) and none of the studies examined individual barriers to educational access, such as birth registration rates. Uncovering these unique differences could enable organizations, like UNICEF, to take more targeted actions when addressing the problem of conflict-inducing social inequalities.

I will deem this dissertation a success if it eventually leads, directly or indirectly, to even just one more child with a birth certificate or a secondary school diploma in hand. As part of the United Nations' Sustainable Development Goals 4 and 16, the intergovernmental organization is striving to ensure inclusive/equitable quality education and provide legal identities, including birth registrations, for all by 2030, respectively (United Nations Children's Fund, 2017). As of 2019, UNICEF estimates that 166 *million* children under age 5 have never been recorded (United Nations Children's Fund, 2019e). If my study can illuminate the extent of this problem, spatially, then it might enable organizations like UNICEF to re-examine their policies of resource distribution. Additionally, this study has the potential to benefit policymakers and researchers simultaneously. As my conceptual framework in Chapter 2 suggests, there are many

direct and indirect consequences of social exclusion. The scholarly implications of this research will extend to disciplines across the social sciences, including fields like postcolonial studies or conflict studies. Investigating the conditions accompanying social exclusion is a key to reducing the growing chasm between the 'haves' and 'have nots' across Sub-Saharan Africa (Tikly, 2001).

1.5. Dissertation Outline

Addressing the problems of under-registration and educational exclusion in Sub-Saharan Africa required a well-articulated research approach in pursuit of clear and noteworthy findings. This dissertation achieved these objectives and the following chapters will guide readers through them, from a comprehensive literature review through a summative review of its key findings. Chapter 2 and Chapter 3 will provide the context for this study by situating it among prior research and describing its unique explanatory spatial research design, respectively. In Chapter 2, I trace the theoretical and empirical underpinnings of my study using a conceptual framework and highlight gaps in scholarly literature on my topic. Specifically, I shed light on the small body of literature covering the relationship between birth registration and educational access as well as the need for innovative spatial designs in social science research. In Chapter 3, I provide a comprehensive explanation of the explanatory spatial research design, including the six-step process that I employed to analyze secondary data from UNICEF and UNESCO with multiple regression and spatial autocorrelation tests. In doing so, Chapter 3 concludes by giving readers confidence that the results were gained in a transparent, rigorous, valid, and reliable manner.

The results of this study are spread over three independent journal articles in Chapters 4, 5, and 6. Each journal article represents a different strand in the sequential research design – the quantitative analysis, spatial analysis, and mixed methods integration components, respectively. From a written flow perspective, I find it important to note that these articles were designed to be standalone products, as I intend to submit them individually for publication in academic journals at a later date. As such, you may notice during the mixed methods article that Griffin (2021a) refers to Chapter 4 and Griffin (2021b) refers to Chapter 5 of this dissertation. Although these two former chapters are presented in a traditional research format (e.g. overview/literature review, findings, discussion, etc.), the mixed methods article (i.e. Chapter 6) was written with specific formatting for *The Journal of Mixed Methods Research* to advance the idea of an explanatory spatial mixed methods design (Fetters & Freshwater, 2015). Together, these articles produce a cohesive argument for the use of explanatory spatial designs while highlighting birth registration and its relationship with educational access in Sub-Saharan Africa as a case study.

CHAPTER 2: LITERATURE REVIEW

2.1. Introduction

This chapter is designed to critically examine and identify gaps in the available scholarly literature on the interrelated topics of birth registration, legal identity, access to education, and the reproduction of social phenomena over space (Pan, 2016). I synthesize existing literature and theory to show how birth registration processes can be problematized in post-colonial Sub-Saharan Africa. I also rely on theoretical constructs to demonstrate the consequences of such ongoing reproduction of social processes, where individuals are excluded from state services at scale. Both the problem and the consequences have sufficient theoretical backing to help readers understand the context behind my study. The gap between them, however, is the limited body of empirical research testing relationships among variables of legal identity (e.g. birth registration) and state services (e.g. access to education). This gap will become apparent in the following section, which shows the rationale behind my conceptual framework for the study. In the subsequent sections, I will present literature that highlights the relationships between (a) birth registration and legal identity, (b) legal identity and access to state services, and (c) access to state services and various democratic or economic outcomes. This chapter will conclude by examining the spatial nature of these relationships with prior research that shows how social phenomena can become geographically distributed. Combined, these sections will lay the foundation for the next chapter, where I will tie these literature-supported relationships to the methodological decisions for this dissertation.

2.2. Conceptual Framework

Researchers often employ conceptual frameworks to help critically examine each component idea and/or theory that they will test throughout a study (Maxwell, 2013). Conceptual frameworks are pragmatic "navigational devices" because they guide the researcher through complex human behaviors to credible, theory-based results (Evans et al., 2011). They can be employed in narrative or graphical form; either way, they present the presumed relationships between the key variables or constructs in the study (Miles et al., 2013). Using the narrative format to conceptualize the relationship between globalization and education in the postcolonial world, Tikly (2001) was able to focus specifically on the management of exclusion in Sub-Saharan Africa. This framework highly influences my study, although I find it more useful to employ a graphical framework to untangle a complex string of micro-relationships between birth registration, legal identities, and educational access.

Tikly (2001) provided a foundational example for framing themes together across disciplines to uncover the economic, political, and cultural dimensions of exclusion in education systems throughout Sub-Saharan Africa. Primarily driven by the forces of globalization, his essay highlighted a general commonality in the experiences among postcolonial, low-income countries. These experiences exposed two conditions that drove my interest on the topic of educational access in the region. First, he argued that the highly-selective and elitist colonial education structures, opportunities, and forms of schooling have been resistant to change over time. These education systems have remained exclusive over time, and, shaped by globalization, have crystalized the colonial order of limited opportunities for social mobility. This is why I hypothesized that colonial

educational or social structures, like birth registration, could be spatially clustered along post-colonial boundaries. Second, he described how out-of-school youth, often in urban settings, have displayed a despondency with their education systems. As a result, he argued for widened access across all educational levels if these colonial exclusion structures are to change. Thus, not only is Tikly's framework useful from a theoretical standpoint, but he also recognized the need to practically assess the lingering effects of the colonial era.

The conceptual framework for this study (see Figure 2.1) adopts a complex societal-systems approach to account for the complexity, interdependence, and integration of prior research and theory (Marshall & Cole, 2014). Inspired by Gourevitch (2008), I link together multiple theories across a range of academic disciplines to produce a fuller account of the issue at hand. By doing so, a Western researcher, like myself, can construct a meaningful and possibly catalytic study, while avoiding the misrepresentation and exploitation of the countries whom I hope to serve (Andreotti, 2011; Vanner, 2015). Drawing from postcolonial theory in education and sociology, I address the significance of birth certificates in individuals' identity formation processes. Drawing from a structural functionalism theory in international relations, I highlight societal functions of education and pose the possible implications of this study. Drawing from human geography, I review the concept of spatial distribution to show how social phenomena can be replicated at scale. Together, the relationships described across this interdisciplinary conceptual framework reveal a gap in academic literature with only a few studies having come close to exploring the relationship between identity formation processes and access to education in a regional context.

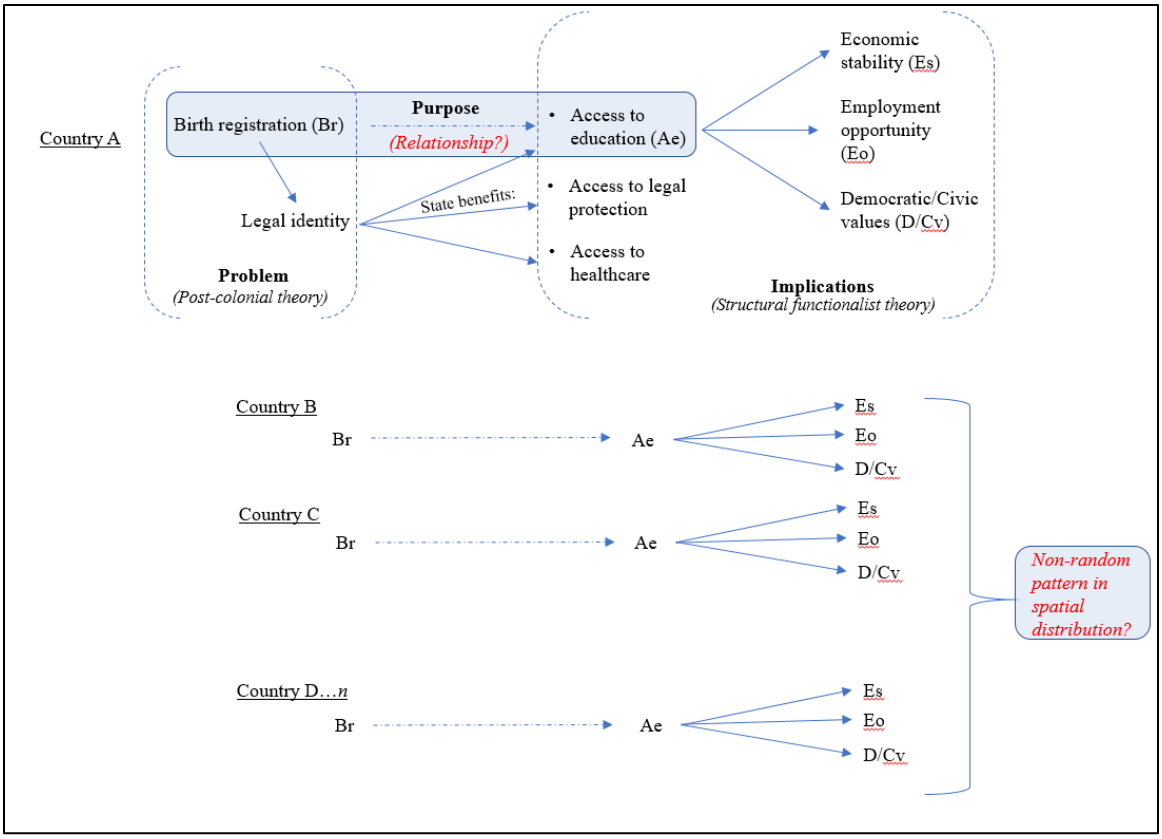

Figure 2.1 Conceptual Framework

2.3. Birth Registration and Identity Formation

An individual's first interaction with the state is often the official recording of their birth (Bequele, 2005). The lasting benefits of this crucial moment range from the ability to travel in and outside of one's home country (Hunter, 2019) to accessing basic public services, like the healthcare system (Bennouna et al., 2016). In addition to being excluded from state structures, the millions of people across the world without birth certificates could also be facing damaging psycho-social effects (United Nations Children's Fund, 2013c). According to Hegel's model of recognition, people have a fundamental need for recognition and their identity is dependent upon recognition by the "other" (Greenhill, 2008; Honneth, 1996). Employing the critical lens of postcolonialism,

the following sections examine how birth registration systems emplaced by colonizers still act as a barrier to legal identity formation processes. This section of the literature review will cover the first relationship in my conceptual framework, which is between birth registration and legal identity formation. Statistical evidence will convey the scope of birth registration as a global and regional problem while postcolonial theory will frame the theoretical concerns of such staggering systemic failures.

2.3.1. Birth Registration

Birth registration, the key independent variable of this study, is the starting point from which I will trace a linear sequence of events in one's lifetime throughout my conceptual framework. Birth registration is an official and permanent process that establishes an individual's existence based upon the legal requirements of his or her country (Selim, 2019; United Nations Children's Fund, 2019e). According to the United Nations Convention on the Rights of the Child (CRC), it is a fundamental human right and it provides the foundation for other rights of the child, ranging from the right to a name and nationality to the right to be free from economic and social exploitation (Todres, 2003; United Nations Children's Fund, 2005). An implicitly related human right is the issuance of a birth certificate, which is the outcome of birth registration processes and is often conducted by a government authority, like a civil registrar (Gerber et al., 2011; Selim, 2019; United Nations Children's Fund, 2019e). This study focuses primarily on the birth registration process because the standard indicator for country-level birth-related statistics is the percentage of children under age 5 with a birth certificate *or* whose birth was registered by civil authorities (United Nations Children's Fund, 2019e).

In its ongoing monitoring of Sustainable Development Goal 16.9 (United Nations Children's Fund, 2018a), the United Nations has produced the most comprehensive, global birth registration figures to date. Reporting on this goal, two key documents from the United Nations Children's Fund (UNICEF) enable researchers, policymakers, and other audiences to examine trends in birth registration data throughout the most recent decade, the 2010s. In its 2013 report, *Every child's birth right: Inequities and trends in birth registration,* the UN reported nearly 230 million children under age five that did not have their births registered by civil authorities. In its 2019 report, *Birth registration for every child by 2020: Are we on track?,* the global number of children without birth registrations improved, as the figure reduced to 166 million. Despite governments' successful efforts to influence this trend by improving civil registries and reducing cost barriers, some regions still lag behind (United Nations Children's Fund, 2019e). Presenting these trends by region, UNICEF showed that the areas with the lowest average percentages of children with birth registrations in 2013 were Eastern/Southern Africa (38%), South Asia (39%), and West/Central Africa (47%) (United Nations Children's Fund, 2013c). In 2019, these same regions witnessed the lowest birth registration values, although South Asia rose sharply to 70% while West/Central Africa and Eastern/Southern Africa only increased to 51% and 40%, respectively (United Nations Children's Fund, 2019e). The fact that sub-Saharan Africa fell so far behind other regions makes it an interesting case region for research and there could be a few factors behind this trend.

Scholarly literature covering the topic of global birth registrations is not tremendously robust, unlike literature surrounding the other UN Sustainable

Development Goals (e.g. climate action, income inequality, etc.). In fact, there are few analyses beyond the aforementioned UNICEF reports that even provide macro-level studies of birth registration trends (Bhatia et al., 2017, 2019). Publications on the birth registration topic generally fall into one of three broad areas. First, there is a body of literature seeking to improve the quality and accessibility of civil registration systems' data globally (Mikkelsen et al., 2015; Mills et al., 2017; Suthar et al., 2019; World Bank & World Health Organization, 2014). In 2007, *The Lancet* journal published a series of articles calling for improvements to vital statistics processes, including births, birth registrations, deaths, and causes of death (AbouZahr et al., 2007; Hill et al., 2007; Mahapatra et al., 2007). As the UNICEF data above suggests, however, over one hundred million children are still "invisible," without a trace of any legal record (Setel et al., 2007). After comparing birth registration data since 1980 to estimates of fertility, Phillips et al. (2018) suggest there are also still gaps in the quality, timeliness, and availability of birth registration systems' data, especially in low- and middle-income countries. Even the UNICEF reports (2013b, 2019) suggest that household surveys are the key source for data monitoring due to unreliable administrative records across countries. I will discuss some of these gaps and my plans to mitigate such data-related limitations in the next chapter.

The second category of literature focuses on the barriers to birth registration and the provision of birth certificates. Despite varying birth registration rates by world region, examples of barriers can be found across all corners of the world. In Latin America, Sub-Saharan Africa, and in Indonesia, research points to a combination of high fiscal costs of procurement (relative to each country) and the opportunity costs of far travel distances to

civil registrars as the greatest barriers to birth registration (Duff et al., 2016; Hunter & Brill, 2016; Issa & Mgaya, 2018; United Nations Children's Fund, 2017). Examples of quality issues (e.g. completeness, process bottlenecks, etc.) in civil registration service provision range from Indonesia to New York State (Bennouna et al., 2016; Melnik et al., 2015). In Brazil, China, and Latin America, demographic characteristics of marginalized families, ranging from indigenous peoples and ethnic minorities to low-income populations, correlated with low percentages of birth registration (Hunter & Sugiyama, 2018; Li et al., 2010; Mackenzie, 2008). In Ghana, these percentages were lowest in children who were born at home (due to rurality), whose mothers were young (15-19 years old), and whose families were in the poorest wealth quintiles (Dake & Fuseini, 2018). Given this myriad of challenges in receiving birth certificates, it might not be a surprise, then, that those left behind by these barriers are faced with a number of lifelong consequences.

The third and final broad category of literature covers the relationship between birth registration, legal identity, and access to various social services. This perspective focuses on the output of the birth registration process, or legal identity, which subsequently affords individuals with rights and entitlements according to their country's national laws (Apland et al., 2014; Mackenzie, 2008). This category also emphasizes the significant negative effects on those who are socially excluded by their lack of birth registration (Harbitz & Tamargo, 2009; Jeong et al., 2018). Some publications specifically refer to birth registration as, "a passport to protection," because it is the first step in legal recognition, which is the essential mechanism that secures their other human rights (Cappa et al., 2014; United Nations Children's Fund, 2013b). My dissertation is

intended to reside within this greater body of studies by focusing on the consequences of low birth registration instead of the barriers one faces during the acquisition process. Later in this literature review, I will review the specific publications associated with birth registration's associated benefits, like healthcare and education. Taking a step back, however, it is critical to review the concept of legal identity, because that is the specific legal *means* to guaranteeing and protecting an individual's human rights (Office of the United Nations High Commissioner for Human Rights, 2014).

2.3.2. Legal Identity

Possessing a birth certificate provides individuals with a means to claim state benefits, but it is their legal identity that actually confers such rights (Apland et al., 2014). That is not to diminish the role of the birth certificate, as it is a "breeder" document that is often used for other forms of identification, like driver's licenses or passports (Office of the United Nations High Commissioner for Human Rights, 2014; United Nations Children's Fund, 2013b). This relationship is illustrated in my conceptual framework by an arrow pointing from birth registration to legal identity. Legal identity formation is critical for empowerment in the modern world by acting as a mechanism for socio-economic mobility, personal development, and sense of belonging (Hunter, 2019). In many developing countries, access to legal identities is not equal and the imposition of identity documentation for social services purposes can even create a dangerous demand market that excludes marginalized populations (Vandenabeele & Lao, 2007). In fact, evidence from Bolivia, Ecuador, and Guatemala suggests that the lack of legal identities was an aggregating factor for social, economic, and political exclusion (Harbitz & Tamargo, 2009). In a powerful statement illuminating this point, Dunning et al. (2014)

argued, "…it is no accident that those lacking birth registration and legal identity are typically the most vulnerable people in the poorest countries" (p. 3). Through a critical post-colonial theory lens, one can see how the legacies of colonization in Africa have made this problem no accident.

Postcolonial theory is a lens through which one can critically examine the impact of the West's chaotic and traumatic historical attempts to transform other societies in a vain of modernity (Amoko, 2013). Furthermore, today's political economies of formerly colonized states are examples where systemic inequalities in the present are a reflection of colonial rule in the past. This is worrisome for the integrity of birth registration processes as, "…the conferral of and benefits from various instruments of legal identity are contingent upon the political economy of countries…" (Ladner et al., 2014). Over 20 years ago, UNICEF acknowledged the leftover structures of colonial governments and their lasting effects on civil registration systems, which did not even register indigenous populations in some Sub-Saharan countries (United Nations Children's Fund, 1998). Therefore, decolonization processes, which focus on the long-term divestment of colonial power, must take a critical look at these type of government instruments (L. T. Smith, 2012). Postcolonial theory, especially when applied in education research, creates the conditions for non-coercive relationships with the excluded 'Other' populations of the Global South (Andreotti, 2011; Gandhi, 1998). Combine the theory with a Hegelian (1977) concept of identity and recognition of/by the other and it becomes conceivable that the legal identities granted through birth registration are still coercive legacies of colonial bureaucracy.

In theory, legal recognition is based upon a minimum commonality of people, placing humans on a level playing field across legal systems that administer their universal interests equally (Douzinas, 2002). A postcolonial lens tells otherwise. From this perspective, one can see how the history of civil registration and legal personhood is rooted in systems emplaced by those in power, who then used such power to construct or withhold the identities of others (S. Q. Jensen, 2011; Szreter & Breckenridge, 2012). In this sense, non-recognition, often a characteristic of colonial subjects (International Institute for Vital Registration and Statistics, 1988), represents a form of harm that has uncompromising consequences in a postcolonial aftermath (Villet, 2011). These consequences can range from the denial of social services (e.g. public education) – the key problem within the scope of my study – to damaging psychological effects, like diminished self-worth (Hopkins & Blackwood, 2011). Skovdal and Andreouli (2011) describe identity as, "…socio-psychological process of positioning oneself in relation to others," which is consistent with Hegel's intertwined concepts of self-consciousness and human need for recognition (Hegel, 1977). This was a major influence behind my study, as I pondered if there are groups of people, arrayed across the formerly colonized states of Sub-Saharan Africa, that are today powerless in their identity positioning process.

In addition to recognizing lasting systems of power found in identity processes, postcolonial theory offers researchers with a lens for other critical analyses. One of those is literary criticism, which builds on Edward Said's work on false images of the Orient (Hamadi, 2014; Said, 1978). Said argued that certain texts, especially in academia and government, are given certain authority and legitimacy and should therefore be critically examined through discourse analysis (Andreotti, 2011; Said, 1978). Applied here, the

very UNICEF (2019) report that I have relied upon to cite global birth registration figures also contains a fascinating first sentence when considered from a postcolonial lens. It states, “Society first acknowledges a child’s existence and identity through birth registration” (p. 6). In placing such a weight to birth registration as the guarantor of identity, even UNICEF fails to recognize other powerful forms of identity, like that granted by being a member of a family (Douzinas, 2002; Hegel, 1910). This critique should not diminish UNICEF’s tireless work in the field. Rather, it is to make the point that even such an entity can unintentionally convey hegemonic and Eurocentric values through discourse. If that is possible, then so is the idea that birth registration in Africa, which originated with colonial administrations, can be fostering marginalization and vulnerability at scale (Bequele, 2005).

2.4. Legal Identity and Access to State Services

The second relationship depicted in my conceptual framework represents key contextual variables for this study, such as the connection of legal identity with access to state services. As Hunter (2019) argues, proof of one’s legal identity is an essential mechanism for social inclusion, meaning we should expect an inverse relationship between increases in birth registrations and decreases in marginality. This is the case because the absence of a legal identity is an impediment to an individual’s capacity to access state benefits (Bhabha, 2011). In fact, providing legal identity documentation (Sustainable Development Goal 16.9) is often a required stepping stone to at least 10 other SDGs, giving those left behind with a number of legal, social, and/or economic disadvantages (Dahan & Gelb, 2015a). Bhabha (2011) argues that people in those

marginalized situations are suffering from a form of twenty-first-century statelessness, where they are bureaucratically precluded from accessing their legitimate claims to services. Birth registration, as one measure of legal identity, has been described as a child's key to unlocking these state benefits, which can ultimately make their lives safer and more successful (Baron, 2012). In this dissertation, I am seeking to understand how many millions of doors to opportunity remain locked as a result of Sub-Saharan Africa's post-colonial education systems.

Before proceeding with this research aim, it should be noted that empirical research – whether in the existing studies outlined below or in my study – relies upon measurement and observation (Galvan & Galvan, 2017). Unfortunately, legal identity is a complex social process that is often hard to measure with existing data indicators (Dahan & Gelb, 2015a). In fact, Oppenheim and Powell (2015) argue that many national and international actors (e.g. UNICEF, USAID, etc.) are still trying to determine how to accurately and effectively measure legal identity even despite the acknowledgement that it is such a critical human right that avails social benefits to individuals. Birth registration stands out as the most common indicator that these actors can use to evaluate progress against goals of universal legal identity within countries (Dahan & Gelb, 2015a; Oppenheim & Powell, 2015). For this reason, I created a shaded box around birth registration in my conceptual framework. This is to show that birth registration is both a step in the legal identity process and an indicator that will be used to represent legal identity in its relationship (or lack thereof) with access to state services, such as education. In other words, while it is legal identity that confers one's right to benefits, it

is the relationship between birth registration and access to education that will enable me to test the strength of such conferral across my study.

As mentioned earlier, my study will be situated within the small body of empirical studies covering various social outcomes resulting from birth registration. In 2009, Harbitz and Boekle-Giuffrida (2009) posited that legal identity documentation was a seldom-discussed subject among scholarly literature. Nearly a decade later, their argument stands. Most of the academic studies pertaining to birth registration outcomes have emphasized various health factors. At the systems level, Phillips et al. (2015) found that improvements in civil registration and vital systems, which includes birth registration, coincided with improved health outcomes in terms of healthy life expectancy, maternal mortality ratio, and child mortality risk. Other studies showed a negative relationship between birth registration and undernutrition, as unregistered children had lower height/weight z-scores and lower scores on the Early Childhood Development Index (Comandini et al., 2016; Jeong et al., 2018). In another study, a lack of birth certificates also correlated with fewer vaccines among children without such documentation (Brito et al., 2017). Even qualitative evidence from indigenous populations in Northern Ontario showed how unregistered births disproportionally affected health outcomes of marginalized populations in Canada (Sanders & Burnett, 2019). It is important to provide this context to show that birth registration is often part of an integrated governance and legal system that protects and promotes children's rights (Apland et al., 2014). Therefore, if there is evidence that birth registration is correlated with the various health services described above, then access to education could presumably also be impacted by the same system(s) of state services.

2.4.1. Birth Registration and Access to Education

There is a gap within academic literature in the relationships between birth registration and educational variables, with very few studies investigating them in detail. The menu of sources outlining the various healthcare rights associated with birth registration does not extend in a similar manner when it comes to education. In fact, Corbacho et al. (2012) and Apland et al. (2014) offer the only noteworthy studies that lay the foundation for this subject altogether. The former was the first paper to identify the missing research about the lack of legal identity and its impact on schooling outcomes. The latter then followed suit with a detailed study on the relationship between birth registration and social services, including mixed methods analyses covering access to education and access to health services. Both of these studies, groundbreaking as they were, came with limitations that have yet to be addressed in this small field of scholarly inquiry.

Corbacho et al.'s (2012) study uncovered a negative effect of the lack of birth registration on educational attainment, but it was too limited in scope to necessarily apply to contexts beyond the Dominican Republic. With a limited scope, one cannot make inferences as to the replicability of this phenomenon in other countries or regions beyond the Caribbean. Their study justifies the need for additional research elsewhere, as their multiple models yielded noteworthy and robust results about the effect of birth certificates on total years of educational attainment. Studies in other contexts may be valuable in assessing the validity of their claims about the causality in this relationship. An additional limitation in their study is that readers cannot clearly identify the precise point at which birth certificates become a barrier to attainment. In other words, their

study was primarily *outcome*-based, emphasizing the fewer total grades of completed education among 7-18 year old students without birth certificates than those with them. Their emphasis on outcomes can be contrasted with their findings on *access*, where they did not find any evidence that a lack of birth certificates affects students' initial entrance to the country's education system. As this dissertation's opening setting in Chapter 1 suggests, there are points in one's educational journey besides initial school entrance where birth registration can become a barrier (Kindzeka, 2017). I therefore argue that it is critical to focus on educational access at each level of education (e.g. primary- and secondary-education) to critically examine how birth registration might act as a barrier to education.

Apland et al. (2014) built upon Corbacho et al.'s (2012) work with quantitative analyses exploring the relationship between birth registration and access to education in Vietnam, Kenya, Sierra Leone, and India. Like Corbacho et al., they identified a significant lack of empirical research on this subject and they were interested in how birth registration benefits or hinders children's ability to access crucial state services. They certainly advance the research on this topic, offering rigorous mixed methods studies within the four aforementioned countries. While they do examine comparisons in birth certificate requirements across each country's educational systems, the study reads more like four individual analyses than one comparative analysis. The multi-country study is a step in the right direction for examining this issue on a grander scale, as they found that students with birth registrations were more likely to be attending formal, age-appropriate, education than those without. Little research has revisited this topic since these two studies were published, so my dissertation aims to both build on their foundation within

the field and become the first to offer a true comparative assessment across countries within one region.

There is clearly a gap in the literature linking birth registration and access to education together, meaning that the few existing studies, by Corbacho et al. (2012) and Apland et al. (2014), will warrant extra scrutiny. The preceding paragraphs highlighted the findings and limitations of each study, leaving one final gap found within their respective methodologies. Corbacho et al. (2012) strictly used a quantitative, econometric approach when examining birth registration and education in the Dominican Republic. Although this may serve as a useful example for an in-depth, intra-country, study, it is too deep in analysis for my aim of examining these relationships across a broader region. Conversely, Apland et al. (2014) used a mixed methods research design (i.e. quantitative and qualitative analyses) in their study of similar variables across four countries in different corners of the globe. In this case, the study is too broad, as my focus on Sub-Saharan Africa as a region is a direct result of the staggering statistics presented earlier in this chapter. In the latter case, however, Apland et al. (2014), provided a compelling use of mixed methods research by emphasizing its benefits in terms of informing advocacy and programming interventions. They do so by using the *why* and *how* associated with qualitative research to provide further insight into their quantitative findings. In my study, I replace the qualitative component with a spatial analysis to achieve a similar intent. Without such empirical analysis, we are left only with theory to infer the consequences of birth registration as a barrier to such vital social services.

2.5. Access to Education and the Promise of Structural Functions

The relationship between birth registration and access to education binds together two parenthetical associations within my conceptual framework. The first, discussed above, describes the problem(s) associated with legal identity formation in a post-colonial context. The second, presented below, depicts the implications should these problems manifest at scale within and across countries. In this section, I argue that the existence of barriers to individuals' identity formation and progressive levels of education could preclude individuals from a number of societal benefits. This argument is rooted in theoretical consequences stemming from breakdowns among key social structures upon which individuals depend for guaranteeing their needs. As Wendt (1987) famously argues in his proposition of the agent-structure problem in international relations theory, individual human agents and social structures are mutually interdependent entities. Regardless of one's philosophical stance on the purpose of education – for example, as a function of democracy (Beane, 1998) or as a function of economic growth (Hanushek & Wößmann, 2010) – it is one of society's key social structures. It is thought to deliver a promise, or certain outcomes, to better individuals and bond diverse groups with a common good (Boman, 2006). If it fails to deliver on such a promise, by a number of ways described throughout this section, then it could lead to greater instances of conflict or societal instability. Applying a structural functionalist lens to this issue, which assumes that society's institutions or structures lie in an integrated and equilibrium state, I will examine the theoretical consequences when such a state is disrupted by failing to fulfill its promise(s) (Avruch & Black, 1987; Hallpike, 1973; Morrow, 1978).

Formal education systems are theoretically associated with human capital formation, as individuals accumulate skills and knowledge through progressive levels of schooling. Becker (1993) goes so far as to claim that education is the most important investment in human capital, citing earnings data by education level even after controlling for numerous socioeconomic factors. It is no surprise, then, that the spread of modern public mass education systems in the developing world throughout the postcolonial era was intended to improve human capital at scale (Saleh, 2016; Sokoloff & Engerman, 2000). The institution of education became a clear gateway to better economic futures for citizens and countries. At an individual level, the perception of returns on investment (ROI) and increased human capital often drives parents' decisions to invest in their children's quality education (R. Jensen, 2010). On a community scale, it is the aggregate demand for increased future earnings that pressures employers to invest in their labor supply's human capital improvement (Kosack, 2009). As human capital begins to grow across its citizens, a country will presumably become more economically modernized. Therefore, the economic stability of a country and the employment opportunities for its citizens appear linked with access to their education systems.

Access to education and its subsequent human capital accumulation is both a vehicle to modernity and an indicator of democracy. Investment in human capital is credited with shifts from agrarian to industrial societies, as workers transform their knowledge and skills into higher yields of production (Rostow, 1990; Schultz, 1961). As an upward mobility device, human capital then results in increases of real earnings per worker (Saleh, 2016; Schultz, 1961). According to modernization theorists, like Lipset (1959), higher levels of education correlate with higher degrees of industrialization and

greater democracy. Similarly, Croke et al. (2016) revealed a negative relationship between education and political participation in authoritarian regimes, further strengthening the argument that education and democracy are wed together. This should not mislead readers into thinking that there is a sequential relationship between modernity and democracy. Rather, the framework implies that, in theory, modern democratic societies use education and deliberate investments in human capital to strengthen their democratic fabric simultaneously.

Governments often make conscious efforts to further shape their constituents' values through the employment of education (Ansell & Lindvall, 2013). First, education can promote ideologies to malleable young audiences. Cantoni et al. (2017) suggested a correlation between states' design of education content and political ideology exists, citing the subsequent reduction of social friction and establishment of social norms. Likewise, in a study of public goods provision in Kenya, Miguel (2004) described the public school curriculum as, "…an aggressively employed nation-building tool" (p. 335). Second, education can ultimately lead to greater political participation among students and parents. In Mali, for example, parents who sent their children to public schools were more likely to have voted in elections due to more frequent interactions with agents of the state, like teachers or administrators (Bleck, 2013). This effect is likely due to increased individual agency, as deliberate improvements in training and knowledge have been shown to influence political participation (Gourevitch, 2008). As states witness returns on their educational investments in the form of stronger and more participative constituencies, further support to the education system will presumably accelerate the relationships described throughout this conceptual framework and corresponding

literature review. The key point is to show that a breakdown at any of the relational lines in this framework can have serious implications within a single country, or, likely worse, across multiple countries or regions.

2.6. Mapping Spatial Phenomena

The previous three sections of this chapter were all concerned with the scholarly literature pointing to a sequence of relationships *within* countries, showing how birth registration, legal identity, state services/benefits, and various civic outcomes might all be interconnected in Country A, for example. Acting together, these sections uncovered a gap in the literature, where it is unclear if birth registration, acting as a function of legal identity, correlates with access to education, which is one of a few state services often requiring proof of identity. As mentioned in Chapter 1, the purpose of my study is to understand if birth registration acts as a barrier for children across Sub-Saharan Africa in accessing education. Research question #1 aims to explore if the relationship exists over a study area by testing the relationship between birth registration and access to education across countries A through *n* (i.e. the number of countries of Sub-Saharan Africa). Research question #2 is best represented by Countries A through *n* in my conceptual framework, where I will explore the spatial distribution of these variables, separately, across the same countries. The purpose of the latter is to identify any non-random distributions or clusters of variables. The relative spacing of countries B and C in the conceptual framework, as compared to D…*n,* was a purposeful display to show how variables can be clustered in areas. This final section of Chapter 2 will highlight literature suggesting that spatial distribution analyses are important in uncovering spatial

relationships. In it, I argue that Sub-Saharan Africa makes for an appropriate case study for clustering of social phenomena given its history of artificial colonial borders drawn over pre-existing cultures.

From a theoretical perspective, geographic space is a key concept in human geography as it allows us to represent events, places, and social or environmental phenomena on a flattened map (R. J. Johnston et al., 2000). I will discuss some of the theoretical assumptions behind space and spatiality in Chapter 3, but, for now, I find it helpful to present readers with a basic overview of the term. Agnew (2011) asserted that the concept of space, as it is conceived of in modern times, rose to academic prominence in the seventeenth century. Since then, he argues, space has produced geographic knowledge by allowing people to assign locations and meaning that are reflective of the "spatial imprint" of physical, social, or economic processes. Likewise, Bailey and Gatrell (1995) believed that a host of practitioners beyond geographers (e.g. epidemiologists, planners, sociologists, etc.) face problems of spatial nature. An even stronger "spatial turn" has occurred in the social sciences over recent decades, as computer mapping software has enabled researchers, "…to reason that everything happens in space, but because *where* things happen is critical to knowing *how* and *why* they happen" (Warf & Arias, 2008, p. 1). Additionally, Arias (2010) argued that a geographical conception of space is, "…a means to integrate diverse sources of information and to understand how broad social processes play out unevenly in different locations" (p. 29). This final point underscores the essence of my study, as I wanted to understand if/how birth registration and educational access intersect with one another across Sub-Saharan Africa.

From a methodological perspective, geographers are often interested in how social phenomena transcend geographical spaces. While this topic will be covered at length in Chapter 3, literature about spatial autocorrelation suggests that it is worth addressing my particular variables of interest from this perspective. According to McGrew Jr. et al. (2014), part of the geographic research process is to explore the distribution of variables across adjacent spaces to discover any significant patterns in the values of each spatial attribute. This stems from research suggesting that many ecological phenomena have been shown to demonstrate patches and/or gradients in their observed values across space (Legendre, 1993). Spatial autocorrelation, which applies a statistical technique to examine correlations across geographic areas, has been growing in popularity since the 1970s, as researchers sought to test hypotheses about spatial relationships (Getis, 2008). Some authors identified the benefit of spatial autocorrelation as the ability to determine if high values are found near places of other high values or if low values tend to move in concert with other low values (Bolstad, 2016). This technique has already been employed in Africa alone, ranging from research about the spatial distribution of non-farm enterprise activity in rural Africa (Owoo & Naudé, 2014) to child health inequalities in sub-Saharan Africa (Yourkavitch et al., 2018). In the latter case, the distribution of health indicators across small geographic units, or clusters, aided Yourkavitch et al. (2018) in identifying areas of low health coverage that could then better inform resource allocation. From this perspective, my study assumed that it was conceivable that birth registration and educational access are similar social science variables that could uncover actionable instances of spatial autocorrelation.

The inquiry into social phenomena in their spatial form(s) often stems from geographers' interest in border studies and the way these lines shape human behavior. Africa, as a continent, offers an interesting case study for border studies because the artificially drawn borders conceived by colonial powers have been shown to cross indigenous spaces, leaving behind a legacy of social problems in the post-colonial era (W. F. S. Miles, 2008, 2014; Rudincova, 2015; Touval, 1966). While these lines have often been described as arbitrary and artificial, Herbst (1989) argued that borders are always artificial because states are not natural and are rather socially constructed by those with political interests. As such, some believe that decolonization processes will never be complete until the current borders are re-drawn in a manner that is not drawn by and for the former colonial powers (Ramutsindela, 2010). Until this occurs, modern African borders will likely continue to cut across linguistic or ethnic geographies (Ali et al., 2019; Nettle, 1996). From a postcolonial perspective, this has damaging social effects on identities and communities, as it precludes a certain type of pre-colonial African sense of unity and togetherness that would have been shared among socio-cultural groups had it not been for colonization (Müller, 2020; Touval, 1967). Even colonial education systems, which spread a common structure of schooling, pedagogies, and forms of knowledge throughout Sub-Saharan Africa, have provided a shared basis for postcolonial reform efforts (Tikly, 2001). Thus, any social variable should have grounds for critical examination to help understand if commonalities in their attributes are tied, at least in part, to geography.

The institution of education is no stranger to the lasting effects of colonial borders in Africa, making it a compelling dependent variable in my study. One example of this is

the French and British colonial legacies in the Cameroonian education system, where the country's partition along colonial linguistic lines has continued to shape parallel education systems within the country today (Dupraz, 2019). Another example of this is Huillery's (2009) research, which showed that modern educational outcomes have been determined in French-speaking West Africa by public investments during colonial times. In other words, the areas with early investments in their education systems essentially received a head start in persisting investments over time. Given that education systems in Africa are largely similar to colonial education systems (Shizha, 2005), it is certainly possible that the same social processes emplaced in education systems decades ago could still be in place. This is not necessarily an argument in favor of radical, systemic change, but rather an acknowledgement that Luke (2011) was justified in cautiously approaching the idea of moving educational innovations from one cultural context or nation to another. This literature suggests that characteristics from educational systems might cluster in similar socio-cultural contexts. Likewise, returning to the colonial birth registration practices described earlier in this chapter, attitudes towards civil registration systems may tend to follow colonial lines. These are the theoretical assumptions that was tested using my study's innovative methodology.

2.7. Conclusion

After highlighting the millions of children in Sub-Saharan Africa without birth certificates in Chapter 1, this chapter was intended to shed light on the processes by which that problem occurs and generates certain consequences. In this chapter, I reviewed some of the barriers inhibiting the birth registration process in developing

countries (Duff et al., 2016; Hunter & Brill, 2016; Issa & Mgaya, 2018; United Nations Children's Fund, 2017) and the significant negative effects of those excluded from social benefits as a result (Harbitz & Tamargo, 2009; Jeong et al., 2018). Through a post-colonial lens, it became evident that these lasting conditions likely stemmed from deliberate systems of identity recognition that favored those in power during colonial-era Africa (International Institute for Vital Registration and Statistics, 1988; S. Q. Jensen, 2011; Szreter & Breckenridge, 2012; Villet, 2011). This deep-rooted social exclusion persists as access to state services, from healthcare to education, is dependent on proof of one's legal identity (Hunter, 2019). Because legal identity itself is difficult to measure, I turned to birth registration as an indicator of access to state services, such as public education systems (Dahan & Gelb, 2015b; Oppenheim & Powell, 2015). The scant body of research on this relationship, with notable exceptions in Corbacho et al. (2012) and Apland et al. (2014), clearly uncovered a need for further research. After all, as I concluded in the preceding sections, academic literature points to dire consequences if education systems fail to deliver on their social promises (Avruch & Black, 1987; Boman, 2006), especially considering the degree to which social phenomena can transcend geographic spaces.

The final component of an effective literature review is to reiterate any gaps found within the relevant literature that help the researcher identify areas for future study (Pan, 2016). In this dissertation, I aimed to fill a void in the limited research testing the relationship between birth registration and access to education. I sought to strike a balance between the depth found in Corbacho et al.'s (2012) study in a Dominican Republic context and the breadth found in Apland et al.'s (2014) mixed methods study

spanning four countries in different geographic regions. Methodologically, the latter study influenced my decision to use complementary research methods to more thoroughly investigate this relationship, although I substituted a spatial method in place of a qualitative one. With an emphasis on Sub-Saharan Africa, I built on Tikly's (2001) theoretical framework of educational exclusion in a post-colonial construct by testing out such assertions, empirically. In taking such an interdisciplinary approach, my study makes unique contributions to social justice literature across academia, from human geography to international education. Finally, the fact that an estimated 237 million children under age 5 worldwide do not possess a birth certificate implies that one of the largest gaps might be found in a lack of effective interventions in birth registration. Motivated by a sense of pragmatism and purpose, my explanatory spatial mixed methods study was an attempt to both understand the *what* and the *where* of this significant birth registration shortfall (Warf & Arias, 2008).

CHAPTER 3: METHODS

3.1. Chapter Overview

In this chapter, I lay out the methodological plan for this dissertation. Lying at the intersection of philosophy, strategic inquiry, and specific methods, the mixed methods research design will fill the gaps in literature addressed in the previous chapter (Creswell, 2009). This chapter follows Creswell's (2009) interconnection framework by covering these three foundational components of a research design. First, I share the story of how I discovered my research topic firsthand in Cameroon and provide insight into my philosophic worldviews that help contextualize this research. Next, I discuss how these experiences and paradigms translated into actionable research questions that naturally lend themselves to a mixed methods strategy. I rely heavily on Creswell and Plano Clark's (2018) explanatory sequential research design, but will modify it with an innovative spatial turn, called an "explanatory spatial" design. Then, I provide a detailed, six-step, blueprint for employing quantitative and spatial analytical techniques to better understand the relationship between birth registration and educational access across Sub-Saharan Africa. Finally, I conclude by reviewing the study's limitations to set a transparent and rigorous stage for the study's three main journal articles (Chapters 4-6).

3.2. Researcher Positionality

Slightly over one year prior to this writing, I was participating in what would be the first of many engaging and personally fruitful conversations with a district-level primary school inspector (e.g. superintendent) in Northern Cameroon. Any outside observer would not have bet on this relationship to endure, however, as the conversation

would cease with me admitting an inability to aid in his top priority for the district. This was unexpected because I had been trained for this exact situation – to build rapport with cultural and political leaders – when becoming an Army Civil Affairs Officer a few years prior (Joint Chiefs of Staff, 2018). So, I set off to develop a human connection by quickly scanning for anything in his office that could help strike a natural conversation. The soccer trophies from three of his schools' recent championships were the perfect entrée into sharing my same passion for the game. While that bought me a few jovial minutes, it quickly became apparent that his focus for discussion that day would be the only thing in his office larger than the trophies – the stack of his students' applications for retroactive birth certificates.

The school inspector painted a sad picture of the countless young adolescents within the city who were about to finish primary school (and their formal educations altogether) unless they could acquire a birth certificate as their ticket to secondary school. Unfortunately, the cost for shepherding their applications through the formal processes was simply too expensive for most families to shoulder without any government assistance. On the heels of a national election, the risk of being perceived as foreign intervention in a civil registration and identity process (i.e. birth registration) was too high for my team to participate in any sort of problem-solving capacity. Although we ended up improving school infrastructure and providing school supplies to the district over the next nine months, the fact that the problem went unsolved still continues to trouble me. Like the civil-military efforts before me in Cameroon, I am driven to help avert the scores of out-of-school youth from becoming vulnerable to recruitment by the neighboring violent extremist group, Boko Haram (Kimmons, 2017). In the following

sections, I will highlight some philosophical underpinnings of these vulnerabilities that drive my interest in birth registration and access to education. These underpinnings, or philosophical assumptions, will then set the foundation for the methodological choices made throughout the remainder of this chapter.

3.3. Philosophical Assumptions

Outlining the theoretical lens(-es) in a study is a critical step in bridging the gap between paradigmatic worldviews and methodological approaches (Creswell & Plano Clark, 2018; Crotty, 1998). Researchers should explicitly state their philosophical assumptions and set of beliefs, or worldviews, to be clear about how they expect to gain knowledge from their study (Creswell & Plano Clark, 2018). How the researcher conceives and implements their ideas provides the critical foundation that guides their inquiry (Nastasi et al., 2010). The combination of my firsthand experience witnessing the educational exclusion problem in Cameroon and my underlying philosophical perspectives provide readers with full transparency of my positionality as a researcher. For instance, my military deployment provided the international background to the problem, directing my attention to the concept of birth certificates as a variable in a social exclusion equation. However, the preceding years of working to combat exclusive admissions policies in an American postsecondary context helped ground my passion for the subject of educational access. Through a concurrent phase of academic study, I found that new paradigms were emerging to re-examine the historical inequalities in education and the traditional barriers to learning (Ilon & Lee, 2014). Therefore, I now find it

especially critical to be forthright in how I approach these problems in a global setting when I innately apply a Western conception of systems and space (L. T. Smith, 2012).

Given my dual identity as a citizen-soldier and an educator, it seems fitting that I apply the perspectives and values from multiple lenses to approach this dissertation. I naturally rely on traits from each when encountering real-world problems requiring thorough analysis and creative solutions. It therefore makes sense that I adopt a paradigmatic pluralism approach, which is one of nine core characteristics of mixed methods research and combines two or more worldviews (Tashakkori & Teddlie, 2010). A dialectical perspective towards paradigmatic pluralism is perhaps the most appropriate outlook on my position as the researcher in this study. With dialectical pluralism, researchers embrace certain intellectual tensions and contradictory ideas while simultaneously promoting knowledge that is both universalistic theoretical and locally practical (Greene & Hall, 2010; R. B. Johnson, 2017; Onwuegbuzie & Frels, 2013). Creswell and Plano Clark (2018) argue that although these contradictory ideas cannot be reconciled, honoring them with multiple worldviews can generate unique insights into the world. With a particular emphasis on the concept of space in this study, I embrace the intersecting tensions between a pragmatic epistemology on one hand and a postpositivist ontology on the other hand.

Pragmatism is an ideal worldview to serve as the foundation for my study for two epistemological reasons. First, with a commitment to produce socially useful knowledge, pragmatism is a problem-centered perspective that focuses on the consequences of research through multiple methods of data collection (Creswell & Plano Clark, 2018; Yvonne Feilzer, 2010). As stated above, my goal in this study is to uncover a potential

relationship between birth registration and access to education, but it is the implications of the possible findings that make this study so intriguing. The literature presented at the tail end of the previous chapter pointed towards serious consequences of social exclusion and failed education systems across some of the world's most vulnerable populations. Second, I believe that exploring this topic over a spatial frame of reference (i.e. Sub-Saharan Africa) is a function of practicality because it binds the research problem at hand to a digestible yet specific study area. In that sense, I apply an "epistemology of the grid" to this study, which is, "…a procedure for locating and segmenting a complex, relational, and dynamic social reality" (Dixon & Jones, 1998, p. 251). The explanatory power of spatial analysis, as Dixon and Jones (1998) argue, is rooted in the grid epistemology because it facilitates the surveillance of sociospatial structures over flattened maps represented by a grid of latitude and longitude.

Like Hall (2013) and Maxwell (2011), I am not convinced that there is one best worldview for mixed methods research. The dialectic irony here is that my ontological position on the nature of reality (i.e. that there is one *singular* reality) is more consistent with a *pluralistic* postpositivist worldview than a pragmatic one (Creswell & Plano Clark, 2018). In other words, I argue that it takes a collection of worldviews examining one single reality to address complex, societal research problems. Again, I will refer to spatial concepts to illuminate my ontological position here. In discrete space, objects, like countries or rivers, are found at precise points on Earth (Ballas et al., 2017). Confirming these points with a GPS, for example, would suggest that they exist at a single and verifiable set of geographical grid coordinates. By applying their own worldviews, people may disagree over the meaning associated with that location, which is referred to as

"place" in human geography (Tuan, 1979). Regardless of those varying places, the objects themselves still reside within a single spatial reality (i.e. location). My challenge in this study is to ensure that reality as it is represented in my spatial analysis model conveys the reality as it truly is (Steinberg & Steinberg, 2015). The aim of my pragmatic and postpositivist study will be to understand how certain socio-structural problems manifest over space. As the preceding literature review showed, there are places that can already serve as empirical and theoretical benchmarks.

Based on my philosophic worldview (i.e. pragmatic and postpositivist) and the strategy of inquiry driving my research question (i.e. explaining the extent of a problem), a mixed methods research design is the most fitting for this study (Creswell & Plano Clark, 2018). There are a few characteristics about this approach that are important to recognize before discussing key research decisions, ranging from data collection to analysis techniques. First, it is an observational study, intended on alerting the academic community to a phenomenon by assessing specific associations between variables (L. M. Sullivan, 2018). Second, it can be characterized as *descriptive* non-experimental research (as opposed to *predictive* non-experimental research) because the purpose is to describe and document the characteristics of the phenomenon rather than predict future conditions or infer causality (Johnson, 2001). Finally, it is an implicitly spatial study, where the observations represent geographic features (i.e. countries) but the locations themselves are not under scrutiny (McGrew Jr. et al., 2014). Together, these features allow for a thorough examination of birth registration and educational access variables across a large, cross-national, study area.

3.4. Research Questions

Mixed methods research questions, according to Plano Clark and Badiee (2010), generally mirror the personal worldviews of the researcher. My study is no different from their claim, as I approached the quantitative strand of my study from a pragmatic perspective and my spatial strand from the postpositivist worldview described above. The former sought to determine which, if any, dependent educational variables are most closely associated, or correlated, with the independent variable of birth registration. The latter aimed to uncover the ways in which our socio-spatial world of borders and post-colonial constructs affect the variables included in such analysis. Tashakkori and Creswell (2007) argued that the two strands of a mixed methods study are deserved of their own research questions along with a third explicitly formulated mixed methods question. My study will follow this direction with the research questions outlined below and in Figure 3.1. The first question follows the characteristics of a traditional quantitative research question, which investigates the relationships between variables of interest (Plano Clark & Badiee, 2010). The spatial question represents an inquiry into the geographic values of location, extent, distribution, and spatial association of the key variables (Board, 1984; Nyerges, 1991). The final overarching question is a combined mixed methods question that questions how the methodology helped explain the content, or the phenomena at the center of this study (Creswell & Plano Clark, 2018). These questions, paired with corresponding hypotheses, provide a comprehensive inquiry into the statistical and spatial relationships associated with birth registration and educational access in Sub-Saharan Africa.

<table>
<tr><th>Research Questions</th><th>Data Strand</th><th>Hypotheses</th><th>Sample</th><th>Independent Variable</th><th>Dependent Variables</th><th>Test</th></tr>
<tr><td rowspan="3">RQ 1: Is there a relationship between birth registration and educational access across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?</td><td rowspan="3">Quantitative</td><td rowspan="3">H0: no statistically significant relationship exists (β=0) between birth registration and educational access

HA: there is a statistically significant relationship (β≠0) between birth registration and educational access</td><td rowspan="5">All countries in Sub-Saharan Africa meeting data screening requirements (n= 40)</td><td>Percentage of children under age 5 whose births are registered</td><td>Educational Access – Primary School (Composite Variable)</td><td rowspan="3">Multiple Regression Analysis</td></tr>
<tr><td>GDP per Capita</td><td>Educational Access – Lower-Secondary School (Composite Variable)</td></tr>
<tr><td>Rural Population %</td><td>Educational Access – Upper-Secondary School (Composite Variable)</td></tr>
<tr><td>RQ 2: Is there a non-random pattern in the spatial distribution of birth registration, GDP per capita, rurality, and educational access across countries in Sub-Saharan Africa since 2010?</td><td>Spatial</td><td>H0: attribute values of birth registration, GDP per capita, rurality, and educational access are randomly distributed among countries in the study area

HA: attribute values of birth registration, GDP per capita, rurality, and educational access are not randomly distributed among countries in the study area</td><td colspan="2">• Birth registration rate
• Educational access – primary school
• Educational access – lower secondary school
• Educational access – upper secondary school
• GDP per capita
• Rural population %</td><td>Moran's I Index</td></tr>
<tr><td>RQ 3: To what extent does the spatial distribution within birth registration, other socioeconomic or geographic variables, and educational access help explain the relationship between them in a cross-national study across countries in Sub-Saharan Africa since 2010?</td><td>Mixed Methods</td><td>N/A</td><td colspan="2">Varies depending on pragmatic need (See Chapter 6)</td><td>Multivariate Cluster Analysis</td></tr>
</table>

Figure 3.1 Research Questions and Data Strand Matrix

In an effort to clearly and comprehensively address the quantitative and spatial research questions, they are each be closely tied with research hypotheses. While the research design will be explained in the subsequent section, it is important to describe how it is derived precisely from the research questions through a method known as hypothesis testing. In many cases, this process involves a null hypothesis, representing no difference among variables' means, and its converse form, the alternative hypothesis (McGrew Jr. et al., 2014; L. M. Sullivan, 2018). In some cases, the null hypothesis is driven by the researcher's estimated guess about the relationships between variables (Steinberg & Steinberg, 2015). The former case, which is used in my study, is a non-directional hypothesis and the latter case is a directional hypothesis, where the researcher believes there will be a positive or negative association among variables (McGrew Jr. et al., 2014). The purpose of a hypothesis test is to determine a threshold, or critical value,

for the point when the researcher can statistically accept the null hypothesis and reject the alternative hypothesis, or vice versa (L. M. Sullivan, 2018). In the case of my first two research questions and hypotheses below, if the null hypotheses (H_0) are rejected, indicating a relationship between variables, then we accept the alternative hypotheses (H_A) and the answer to the research question is "yes." The following list outlines all research questions for this study.

3.4.1. Situating the Research Questions

Before listing each of the research questions, I find it helpful to orient readers to the timeframe and variables that are listed consistently throughout them. First, each question covers a time period of 2010 to present. From a practical standpoint, 2010 allows for the inclusion of multiple rounds of UNICEF's MICS survey, which should provide sufficient data coverage for ensuring that a complete sample of countries could be used for my cross-national analysis (United Nations Children's Fund, 2013a). I will discuss the role of time in more detail at the beginning of the next section. Second, I am using a composite of three indicators – net attendance rates, school completion rates, and out-of-school rates – as a proxy for educational access. Additionally, the first research question contains three sub-questions, with the educational level in brackets, to clearly show how they will be tested separately to show differences in access by level. Like the timeframe, each of these variables will be covered in more detail during the data collection section of this chapter. The research questions are listed first, however, to show the audience how my research decisions flow logically from detailed research questions and hypotheses to a specific mixed methods study design.

3.4.2. List of Research Questions

Research Question #1 (Quantitative). *Is there a relationship between birth registration and educational access across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?*

H_0: no statistically significant relationship exists ($\beta=0$) between birth registration and educational access

H_A: there is a statistically significant relationship ($\beta\neq0$) between birth registration and educational access

Sub-Question #1a. *Is there a relationship between birth registration and educational access [at the primary school level] across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?*

Sub-Question #1b. *Is there a relationship between birth registration and educational access [at the lower-secondary school level] across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?*

Sub-Question #1c. *Is there a relationship between birth registration and educational access [at the upper-secondary school level] across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?*

Research Question #2 (Spatial). *Is there a non-random pattern in the spatial distribution of birth registration, GDP per capita, rurality, and educational access across countries in Sub-Saharan Africa since 2010?*

H_0: attribute values of birth registration, GDP per capita, rurality, and educational access are randomly distributed among countries in the study area

H_A: attribute values of birth registration, GDP per capita, rurality, and educational access are not randomly distributed among countries in the study area

Research Question #3 (Mixed). *To what extent does the spatial distribution **within** birth registration, other socioeconomic/geographic variables, and educational access help explain the relationship **between** them in a cross-national study across countries in Sub-Saharan Africa since 2010?*

3.5. Research Design

After presenting the research questions above, I find it critical to clarify that this study was not designed to be longitudinal. For example, I did not collect birth registration data on a sample of 5-year olds and then re-survey them 5-10 years later to determine its effect on their educational outcomes. Instead, I was seeking a relationship, or association, between the countries that struggle with birth registration rates and the countries with lower educational access. Another way of framing the study is to ask, "are the countries with low birth registration also the same countries where educational access is lower?" In essence, this is a restated way of addressing Research Question #3. Based on my personal experience in Cameroon, it is very possible for students of primary school age to retroactively receive birth certificates provided they follow an expensive and often convoluted administrative process. Therefore, I found it believable that the countries with bureaucratic barriers in place for one government institution (i.e. civil registration) could be the same countries with barriers in place in another government institution (i.e. public education). This clarification is critical for readers to ensure they are not approaching my research design from a longitudinal perspective but rather a relational one.

In this dissertation, I employed a mixed methods research design to examine the relationship between birth registration and access to education in Sub-Saharan Africa. My research design was typology-based, meaning that I adapted a foundational mixed methods typology – an explanatory sequential design – to fit this study's research questions (Creswell & Plano Clark, 2018). I called my research design an "explanatory spatial" model in an effort to preserve this core typology while reflecting the need for a methodological variant (Creswell & Plano Clark, 2006). In the traditional explanatory sequential design, according to Creswell and Plano Clark (2018), a researcher begins with a quantitative data analysis before moving sequentially to a qualitative strand to explain those initial results with more depth. In my design, I also began with a quantitative strand but I followed it with a spatial analysis using selected variables from the quantitative analysis. Creswell and Plano Clark also asserted that mixed methods researchers should include a notation system to clearly convey the flow of data analyses. The following notation best symbolizes my study:

$$\text{QUAN}_{\text{relationship}} \rightarrow \text{spatial}_{\text{distribution}}$$

In this notation arrangement, the capitalization of "QUAN" denotes the priority of the quantitative strand while the directional arrow and "spatial" notation represents the deductive and supplemental approach to explaining the initial results (Creswell, 2010; Morse, 1991). Since this was an untested research design, prior to my dissertation, I included additional terms in the subscript of the notation to clearly present the purpose for each strand. In other words, the purpose of my study was to test for a *relationship* between key variables in the quantitative strand before explaining those results by testing

for their spatial *distribution.* I believe that there are two important reasons why this research design has the potential to be unique and effective.

First, due to the complex nature of the research problem at hand, a need existed to obtain more complete and corroborated results than with only one strand (i.e. quantitative or spatial) alone (Creswell & Plano Clark, 2018). This is an issue and study area requiring analytical breadth, as multiple research methods can provide deeper understanding and enhanced descriptions of the variables of interest (R. B. Johnson et al., 2007). Given the socio-spatial theme depicted in my conceptual framework and supporting literature, mixed methods research offered a fitting typology, as it is inclusive of the many ways that humans can make sense out of the social world (Greene, 2007). Additionally, this need for a completeness of results stems from a place of pragmatism, as the aim of this study was to generate attention towards the countless children across Sub-Saharan Africa that are excluded from their education system for lack of birth documentation (Creswell & Plano Clark, 2018). After all, my resource-limited team was unable to directly aid the students of Northern Cameroon with this issue but perhaps this study could uncover how many more places are in similar unfortunate situations. Combining multiple strands of data could enable a humanitarian or development organization, like UNICEF or USAID, respectively, to not only understand the extent to which birth registration acts as an educational barrier in Sub-Saharan Africa (i.e. research question #1), but also to determine *where* to allocate resources most effectively (i.e. research question #2).

Second, there was a need for a more deliberate variant in which a spatial component becomes integrated with a quantitative strand in mixed methods research.

McGrew Jr. et al. (2014) argued that looking at spatial variation and distribution is a logical next step to statistical data analysis techniques. Despite this seemingly obvious next step, from a human geographer's perspective (i.e. always wanting to know *where),* spatial analyses were notably absent from Creswell and Plano Clark's (2018) detailed review of mixed methods typologies that crossed academic disciplines and many adaptations to core research designs (pp. 54-57). There has been evidence of and arguments for the use of geographic information systems (GIS), a spatial analysis technique, in a mixed methods setting (Brown et al., 2017; Giordano & Cole, 2018; Preston & Wilson, 2014; Warshawsky, 2014; Yeager & Steiger, 2013). A substantial amount of literature also existed that contained the concept of "qualitative GIS" as a mixed method (Cope & Elwood, 2009; Jung & Elwood, 2010; Shay et al., 2016; Sinha et al., 2017). What I proposed, however, was not simply another example of a spatial mixed methods study in practice. Rather, I was highlighting a need to assess the utility of a distinctly sequential employment of multiple data strands (e.g. quantitative and spatial). The aim here was to establish a credible framework for a future 'explanatory spatial' mixed methods design. This assessment will contribute to Yoon & Lubienski's (2017) call for a mixed methods geospatial research agenda that supports education policy analysis. Specifically, they recognized that mixed-methods GIS can allow researchers to re-conceptualize identities as spatially constituted and help them understand spatial marginalization in educational policies – the precise purpose of my dissertation.

I use a mixed methods procedural diagram, adopted from Ivankova et al. (2006), to clearly delineate the steps required to perform my study. Like a conceptual framework,

the procedural diagram (see Figure 3.2) serves as a road map for the remainder of this chapter.

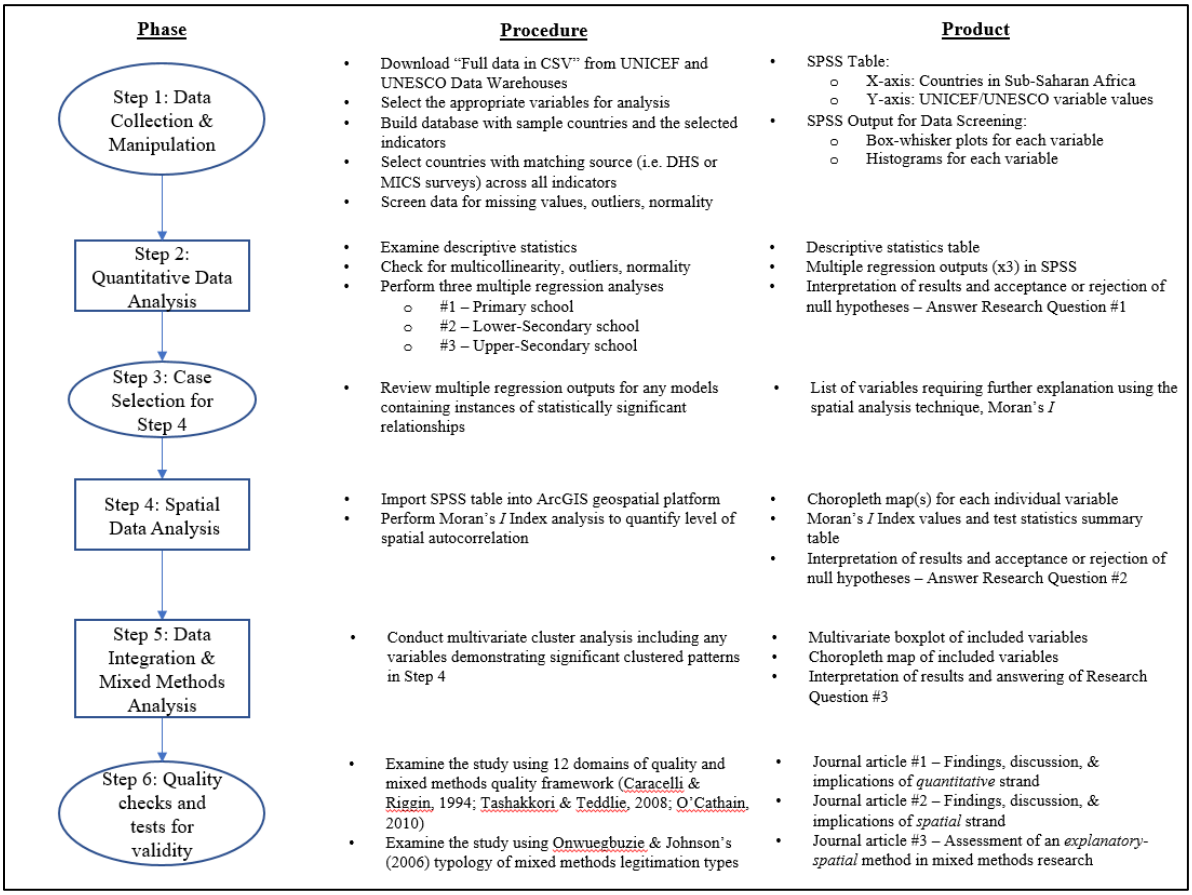

Figure 3.2 Procedural Diagram

Each stage will be described in detail as I will highlight the data collection and analysis techniques along with the intended outcome of each step. Beginning with data collection, I will describe my process for building a custom secondary dataset using the United Nations Children's Fund (UNICEF) and United Nations Educational, Scientific and Cultural Organization (UNESCO) data warehouses. From there, I will test my quantitative hypotheses associated with Research Question #1 with three multiple regression by quantifying the association between variables of birth registration and

educational access (L. M. Sullivan, 2018). The results from this test will then dictate the variable selection that will be used in the subsequent spatial analysis. The screening and decision-making criteria will be presented thoroughly below. I will then perform a Global Moran's *I* test on each of the selected variables to test for non-random spatial clustering of their attribute values across Sub-Saharan Africa (Research Question #2). Finally, I will integrate both data strands into a final graphical format to provide an interpretation of extent. Revisiting Figure 2 from above, one can see how these two tests will cover the two specific research questions while the integration of data will address the overarching Research Question #3.

3.6. Data Collection (Step 1)

This study relied on publicly available secondary data to test for relationships between birth registration and educational access (see Appendix A for institutional permission to use such data for dissertation purposes). The secondary data, which was collected by the UNICEF and UNESCO, came in spatially-aggregated data sets, where the data represented national values for each respective indicator (McGrew Jr. et al., 2014). This allows for a cross-national analysis, in which country-specific particularities can be differentiated from social-structural regularities across countries within a specified region or, in this case, Sub-Saharan Africa (Kohn, 1987). Secondary data is becoming more prevalent in academic research, as technology enables large amounts of data to be made readily accessible to researchers for empirical analyses (M. P. Johnston, 2017). It is especially useful in situations where time and financial resources are constraints (McGrew Jr. et al., 2014). For example, it would be ideal, yet logistically impossible, to

have replicated the opening anecdote of Chapter 1 across all countries in Sub-Saharan Africa. Instead, secondary data made it possible to understand if the birth registration issue in Northern Cameroon was precluding youths across the region from accessing progressive levels of formal education. The following sub-sections address the key roles of UNICEF and UNESCO, which lead in providing international data as well as preserving quality and consistency considerations of secondary data sources (Atkinson & Brandolini, 2001; Murray & Newby, 2012).

3.6.1. UNICEF Data Warehouse

In this study, I will built a dataset for analysis partly using the openly-accessible secondary data acquired through the UNICEF Data Warehouse (United Nations Children's Fund, 2020a). UNICEF is a well-established entity in terms of worldwide data curation and has even been designated as a custodian or co-custodian of 17 Sustainable Development Goals by an Inter-Agency & Expert Group on SDG Indicators (United Nations Children's Fund, 2018a). In fact, UNICEF is the custodian agency for official birth registration reporting, making it the ideal entity for this study, as birth registration is the key independent variable that will be used for analysis. The UNICEF global database, or data warehouse, includes data from over 197 countries, with an estimated coverage of at least 80% of the developing world, according to Murray and Newby (2012). It includes many topical indicators, ranging from child protection to education to immunization, and covers a wide range of demographic groups, like newborns, school-age children and women of reproductive age. This variety in available data points allows users, like researchers and policymakers, to conveniently access data across their geographic regions or content areas of interest.

Assembling and maintaining a robust global database requires UNICEF to aggregate data from multiple sources. This helps fulfill their commitments to improving the quality of data systems, leaving no country behind in their data coverage, and developing international norms and standards for data collection (United Nations Children's Fund, 2018b). However, adolescent and socially marginalized populations, including out-of-school children, are often left on the margins of global data collection efforts (Azzopardi et al., 2019; Patton et al., 2012). Macro-level research studies therefore need to combine data sources sometimes to improve the coverage across indicators or variables. This was the case in a recent study on adolescent health and wellbeing, globally, where researchers developed a dataset using sources from UNICEF, the Institute for Health Metrics and Evaluation (IMHE) and the International Labour Organization (Azzopardi et al., 2019). The UNICEF data warehouse contains some of these data sources; in the case of child protection (i.e. birth registration) and education, they draw data primarily from their own Multiple Indicator Cluster Surveys (MICS) or the U.S. Agency for International Development's (USAID) Demographic and Health Surveys (DHS) Program.

Upon an initial scan of the UNICEF Data Warehouse, it became apparent that the data provided for most of my countries of interest were either sourced through MICS or DHS surveys. These two types of household surveys are closely aligned with one another in structure and measurement approach, serving as the primary data source in most developing countries for children-related health and well-being indicators (Hancioglu & Arnold, 2013). As such, they are practical sources for researchers and the use of them in a paired manner has been well-represented within academic literature (Amouzou et al.,

2017; Ayede et al., 2018; Gray et al., 2013; Larsen et al., 2017). Conducted every 3-5 years in low- and lower-middle income countries, these household surveys and interviews create a representative sample of each country under inquiry (Ayede et al., 2018). They rely on cluster sampling, which divides the population into mutually exclusive and exhaustive categories (i.e. clusters) before randomly sampling individuals (or families) within each cluster (McGrew Jr. et al., 2014). The data from each cluster is then combined to produce a total sample that is representative of the country's greater population. The MICS and DHS surveys were therefore the most appropriate way to access a comprehensive dataset that could be used in my study, which rested on the ability to compare birth registration and education indicators in a macro, cross-national, approach. The following two paragraphs will provide a brief overview of each survey program, respectively.

Multiple Indicator Cluster Surveys. Covering 116 countries since its inception in the 1990s, the Multiple Indicator Cluster Surveys (MICS) program, managed by UNICEF, is one of the largest household survey programs focused on children and women (Khan & Hancioglu, 2019). Khan and Hancioglu's recent publication provides perhaps the most simple, yet thorough, overview of the program and is a useful resource should readers find an interest in a more detailed explanation of the program. To briefly summarize, however, MICS is a household survey program, led by UNICEF-trained field teams and enumerators. The program, which uses a multistage sample design with clusters of 20-25 households, is designed to be representative of the populations from which these clusters are drawn. The MICS field teams conduct face-to-face interviews with household respondents of individual men and women aged 15-49, who self-report

current and retrospective data on themselves. These interviews cover a wide variety of UNICEF's standard, globally recommended data indicators, like Water, Sanitation and Hygiene (WASH), maternal and newborn health, and child development/nutrition. Data on children under 5 years of age (one questionnaire) and ages 5-17 (another questionnaire) comes from their mothers or caregivers if their mothers do not live in the household or are deceased. The extensively-trained MICS program enumerators upload questionnaire data, obtained using tablets during fieldwork with the households, on a daily basis to a cloud server for quality assurance by UNICEF's survey managers. Given the longevity and scale of the program, I determined this to be a trustworthy data source for my analysis.

Demographic and Health Surveys. Like MICS, USAID's Demographic and Health Surveys (DHS) program uses standardized questionnaires and procedures to facilitate comparable analysis across countries (U.S. Agency for International Development, 2018). According to the *Guide to DHS Statistics* (2018), this program was established in 1984 and has been implemented in over 90 countries, most of which (but not all) are recipients of USAID assistance. This program shares many similarities with the MICS program and, while this brief summary may be redundant to the previous paragraph, it is useful to show the consistency in protocol across these survey programs. For example, the DHS program's main objective, like MICS, is to collect information on basic demographic and health indicators across households, ranging from the nutritional status of mothers and children to family planning to attitudes/behaviors towards HIV/AIDS (U.S. Agency for International Development, 2012). The program also seeks to build a nationally representative sample using existing sampling frames drawn from a

country's most recent census or list of communities containing a measure of population size. The DHS field-based data collection teams also receive detailed training and an in-depth interviewer's manual to ensure consistency and data quality as they implement the Household Questionnaire and the Woman's Questionnaire throughout each cluster of households (U.S. Agency for International Development, 2020).

Similar to MICS, respondents in the DHS program's Household Questionnaire should be any knowledgeable person age 15 or older living in the household (U.S. Agency for International Development, 2018). These individuals report information about the characteristics of each person in the household, including their age, sex, education, and relationship to the head of household. This also helps identify women of reproductive age (15-49) in the household, whom are then interviewed as well to provide information on 11 standard survey sections, covering topics like pregnancy and postnatal care, child immunization, and other health considerations. All of these results are updated in a database and then aggregated, generally within 12 months since initiation in a country, to produce a key findings reports for each country. Since this data, along with the MICS results, feed into the UNICEF data warehouse, researchers are easily able to export cross-country data to conduct research on a variety of variables, as shown below.

3.6.2. UNESCO Institute for Statistics

UNICEF's sister United Nations agency, UNESCO, housed a similar database that contained variables that I could export for use in my custom dataset for this dissertation. The database, called the UNESCO Institute for Statistics (UIS), serves as an official source of education, science, culture, and communication data that is effective for comparative, transformational, and development purposes (United Nations Educational,

Scientific and Cultural Organization, 2020c). It contained national-level socio-economic (GDP per capita) and geographic/demographic (rural population %) indicators that I used to hold constant, or control for, during my multiple regression analyses. These indicators were not available in the UNICEF data warehouse, making the UIS system a viable alternative source despite the additional step of needing to merge the data together with my exported UNICEF data. Like UNICEF, the UNESCO database contains data that was originally sourced through household surveys while also including information obtained through population censuses, learning assessments, financial estimates, and administrative data (United Nations Educational, Scientific and Cultural Organization, 2020d). In fact, the two variables that I exported from the UIS database, GDP per capita and rurality, were originally sourced through UNESCO's economic estimates and household surveys, respectively. These two variables were then added to my database, for a total of six indicators that I analyzed throughout this study.

3.6.3. Key Independent Variable: Birth Registration

According to UNICEF (2020b), the standard indicator for birth registration in DHS and MICS surveys is the percentage of children under age 5 (0 to 59 months) with a birth certificate or whose birth was reported as registered with civil authorities at the time of survey. This indicator is calculated by the following equation (United Nations Children's Fund, 2019c):

*Birth registration % under 5 years old = (the number of children with a birth certificate or registration with civil authorities / the total number of children in the population) * 100*

I used this as the key independent variable in my quantitative data analysis because it is arguably the most well-established global indicator related to birth registration. The DHS survey has been collecting birth registration in more than 90 countries since 1993 and the MICS survey has collected birth registration since 1999 in about 50 low- and middle-income countries worldwide (United Nations Children's Fund, 2020b). Additionally, UNICEF reviews all birth registration data points for consistency, overall data quality, proper methodology and sampling, national representativeness, and plausibility prior to entering them into their global database on birth registration. Although there are some limitations of the data itself, the UNICEF database enabled me to include birth registration percentages for most countries throughout Sub-Saharan Africa.

3.6.4. Controls: GDP per Capita and Rurality

In a multiple regression analysis, it can be useful to hold some factors constant, or control for them, to help move toward an understanding of the true causal effect of the key explanatory variable of interest (Arkes, 2019). In this case, my task was to select additional explanatory variables that might explain at least a portion of the variance in educational access while avoiding the effects of confounding, or multicollinearity (Pallant, 2020). Confounding occurs when there is a distortion on your outcomes that is created by other independent variables beyond your key explanatory variable, which, in this case, was birth registration (Sullivan, 2018). As such, I chose to include a common measure of country-level socioeconomic level, or gross domestic product (GDP) per capita, and a measure of geographic/demographic composition, or the percentage of the population that is considered to be rural. Both of these variables have been used in prior research related to educational access (Checchi, 2003) and to the lack of birth certificates

(Brito et al., 2017). After importing both variables from the UNESCO UIS database, I included them in a correlation matrix, alongside birth registration, to ensure that there were no strong correlations between any of the three explanatory variables (i.e. a sign of confounding). Finally, it is important to note that while additional explanatory variables might have been useful, the small sample size in my study (n<40) meant that three independent variables would be the maximum number that I could use effectively (Pallant, 2020; Tabachnick & Fidell, 2007).

GDP per capita. According to the Organisation for Economic Co-operation and Development (2014), GDP per capita is the most widely used comparative indicator of economic standing and is often used for comparing the sizes of economies across countries. They calculate it in the following way:

GDP per capita = Sum of a country's marketed goods and services produced within the national boundaries (in U.S. Dollars) / the country's total population

Rural population percentage. The United Nations (2018) maintains a measurement of rural populations by counting or estimating the de facto population living in areas classified as rural. Rurality is generally considered a combination of population density and the geographic size of a locality (United Nations Statistics Division, 2017). The UN has recognized an urban-rural dichotomy, where qualities of life, access to services, and levels of development all can be impacted by these sociodemographic geographies. Recognizing the impact that this dichotomy might have on educational access, I used the rural population percentage in my study, which the UN (2018) calculates as:

*Rural population % = (Rural population / total population) * 100*

3.6.5. Dependent Variable: Access to Education

The outcome variable for my study, access to education, is more complex than birth registration as there are many ways that it can be measured. I originally thought that access to education would only be a reflection of student enrollments at the education level of interest (e.g. primary-education, secondary-education, etc.). Some studies have in fact used gross enrollment rates at various education levels as indicators of access (Checchi, 2003; Donou-Adonsou, 2019). In doing so, however, this approach often only takes into account those who have enrolled when, in reality, I am more concerned with those students who are *left out* of their country's education system. It became apparent through literature that investigating access to education is best done by using proxies with a combination of variables, ranging from children who had never been to school to students who completed some but not all levels of age-appropriate education to those who had dropped out of school (Lewin & Sabates, 2011). On this complex outlook on access to education, Lewin (2009) writes:

> First, access to education needs to be broadly conceived if it is to be used as a focus for domestic policy and to targets which shape policy and practice. Simple definitions of enrolment registration need to be accompanied by a broader vision which includes entry and progression at an appropriate age, regular attendance, satisfactory achievement, appropriate chances to progress to post-primary, and more equitable distribution of opportunity. (p. 171)

I combined attendance, completion, and out-of-school rates into one composite dependent variable in my study. Not only were these indicators easily accessible through the UNICEF data warehouse, but there was clearly a need to include such variables based

on Lewin's (2009) trailblazing study on access to education in Sub-Saharan Africa. This will help us consider all aspects of the access equation, as it will both shed light on attendance (i.e. those students enrolled in school) but will also follow his recognition of exclusion as an important converse of access:

> The largest numbers of those excluded from education are those who enrol in Grade 1 but fail to complete their schooling. Of the 113 million children of primary school age in SSA over 32 million children remain outside primary schools. More than double this number fails to participate through to the end of secondary schooling. Only in some fragile states where civil unrest, war and other forms of conflict have destroyed infrastructure and disrupted education systems, is it likely that a majority of those out of school are those who have never attended school. (Lewin, 2009, p. 157)

The following three sub-sections will briefly show how the three variables are calculated consistently across the MICS and DHS surveys. In each case, the variable is available at the primary, lower secondary, and upper secondary education levels. According to the International Standard Classification of Education (ISCED), primary education generally covers ages 6-11, or 1st through 6th grades in United States terms. Secondary education covers ages 11 or 12 through 18 or 19, or 7th through 9th grades for lower secondary and 10th through 12th grades, or the first three years of vocational education, for upper secondary education (National Center for Education Statistics, n.d.).

Net attendance ratio. As I mentioned above, the net attendance ratio, like enrollment, shows the percentage of students attending the respective level of education within a country. I argue that this is in an indicator of how accessible a country's

education system is because it directly covers the degree to which students are includes or enrolled. This indicator is calculated by the following equation (United Nations Children's Fund, 2019d):

*Net attendance ratio = (the number of children attending primary or secondary school who are of official primary or secondary age / the total number of official primary or secondary school age children) * 100*

According to the DHS program (U.S. Agency for International Development, 2018), countries' national Ministries of Education and the UNESCO Institute for Statistics databases are the key sources for establishing the specific age range and the month of each academic year start point for each country. It should also be noted that the denominator in this equation and those that follow is based on the household survey respondents' self-reported answers about their children's ages. Without such clarification, we would be in a paradoxical situation where it would be difficult to discern the total number of school-aged children in the population without the birth documentation which can be so notably absent in this region.

Out-of-school rate. Serving as the converse to attendance, I found it important to include out-of-school rates as a second method of measuring access to education. According to UNESCO (2020), higher out-of-school rates in countries can indicate a need to improve access to education as it can help identify excluded populations. This indicator captures children and adolescents who are not enrolled in primary nor secondary levels of education. According to UNICEF (2016), "…in data collected with DHS or MICS a child is considered to be *in school* if she or he attended school *at any time* during the academic reference year provided in the questionnaire" (p. 21). This

indicator is calculated by the following equation (United Nations Children's Fund, 2020c):

*Out-of-school rate = (the number of children **who are not** attending primary or secondary school who are of official primary or secondary age / the total number of official primary or secondary school age children) * 100*

The administrative sections of surveys like MICS or the DHS program provide the important data for this indicator because they can directly identify households where children are out of school (United Nations Educational, Scientific and Cultural Organization, 2020b). This makes sense, as national-level Ministries of Education (see above) may only contain enrollment information and not necessarily be able to track those who are not enrolled in school.

Completion rate. Although it may be intuitive to think of access to education in terms of enrollment, or inputs to education, I also considered how outcome data might be an indicator as well. For example, if a country displays a high percentage of primary school graduates but a low percentage of lower- or upper-secondary school, then there might be reason to investigate reasons why students were precluded from finishing. This is consistent with the opening vignette of Chapter 1, as Cameroonian students cannot sit for exit exams without a birth certificate and are therefore barred entry to the subsequent education level. In addition to access, however, UNESCO (2020a) notes that a low completion rate could be indicative of certain educational quality indicators, such as high dropouts/repetition or late completion. For this reason, I did not want to consider completion rates as a standalone proxy for educational access. Rather, it was used to capture any trends and possible barriers associated with each of the educational levels.

This indicator is calculated by the following equation (United Nations Educational, Scientific and Cultural Organization, 2020a):

*Completion rate = (the number of persons in the relevant age group who have completed the last grade of the given level of education / the total population of the same age group) * 100*

It should be noted that relevant age group is defined as, "a cohort of children or young people aged 3-5 years above the intended age for the last grade of each level of education" (United Nations Educational, Scientific and Cultural Organization, 2020a). In other words, this includes all children who have completed a level of education (e.g. primary school) by the time they are 3-5 years older than the official age of entry into the last grade of that level of education.

Composite Variable. One tactic for combining multiple variables representing a particular construct into a single variable is the use of a composite variable. This should be done after critically analyzing the theoretical composition of the construct and selecting variables that logically represent its meaning (Song et al., 2013). As the preceding paragraphs have shown, net attendance, out-of-school, and completion rates all maintained sufficient theoretical grounds for inclusion in my conception of the term, "educational access." A composite variable has been used before to study access in Sub-Saharan Africa, making it a viable technique in international educational policy research (Spaull & Taylor, 2015). Creating the composite variable only required two steps. First, I needed to invert the out-of-school variable so that it would follow a consistent scale with the other two variables, where higher values represented better access. I accomplished this by using the SPSS compute function and subtracting all out-of-school values for each

country from 100. Second, I used Song et al.'s (2013) simple averaging technique to compute the average of net attendance, completion rate, and the inverted out-of-school rate for each country. I repeated this process at the primary-school, lower-secondary, and upper-secondary school levels, thereby creating three separate educational access variables for use in my quantitative and spatial analyses.

3.6.6. Data Screening

The procedure for building and screening my dataset using the secondary data from the UNICEF and UNESCO data warehouses required a few steps. First, I exported the full datasets from the respective warehouses in comma-separated values (CSV) format and convert them to .xlsx (Microsoft Excel) files. I then selected each of the variables listed above (e.g. birth registration, completion rate, etc.) and deleted all other unnecessary variables (i.e. those not pertaining to my study) from the dataset. These variables served as the columns, or fields, in my final data table. Similarly, I deleted all countries except for those found within Sub-Saharan Africa. These countries represented the rows in my final data table. Once these fields were established, I scanned the dataset to determine which country-level values would remain for my data analyses. My selection criteria included the most recent MICS or DHS survey since 2010 within each country. For example, if Ghana, contained birth registration data from the MICS6 survey in 2018 and the MICS5 survey in 2013, then I would only include the former values in my table. Once the dataset was established, I imported it into the SPSS statistical software and proceeded to screen the data for suitability purposes.

There are a series of data cleaning issues that must generally be resolved between the data collection and data analysis phases of a study (Tabachnick & Fidell, 2007).

Although I selected a research design based on secondary data to minimize these pre-analysis data issues, it is worth mentioning them briefly here. First, there is the topic of missing data. It is typically important to determine if data are missing completely at random to understand if the omissions are effects of unobserved factors beyond the other variables in the study (Cox et al., 2014). In my dataset, I believed that it was likely the case that any missing country values were not random and they were, in fact, due to conditions associated with countries under severe internal conflict (e.g. Somalia, Libya, etc.). In these cases, I removed the missing countries from the analysis. Second, I used box-whisker plots to identify outliers, or extreme values, in the dataset (L. M. Sullivan, 2018). In Chapter 4, you will read about the case of Gabon, which I needed to remove because of the necessity to identify, evaluate, and describe any values that are more than two standard deviations from the mean, as was the case here (Tabachnick & Fidell, 2007). Similarly, I examined the variables for any noteworthy examples of skewness and kurtosis using the histogram feature in SPSS but did not find any in need of further action. Since my sample size was relatively small (i.e. less than 50), I planned to identify and describe any non-normal measures of relative dispersion, but, again, this proved to not be needed (McGrew Jr. et al., 2014). Since there were no major data quality issues, I was then able to proceed to the first of two data analysis phases.

3.7. Quantitative Data Analysis (Step 2)

The analysis of my quantitative data strand consisted of two steps. First, I used a numerical procedure and graphical technique, known as descriptive statistics, to organize my data and describe the key characteristics of each observed variable (Fisher &

Marshall, 2009). Descriptive statistics are used to summarize the basic characteristics of the dataset, but they will not offer any inferential capabilities, like the regression analysis outlined below (Bernhardt, 2013). For my ratio-scale variables, I analyzed the three general classes of descriptive statistics. According to Larson (2006), these include measures of central tendency (e.g. mean, median, and mode), dispersion (e.g. standard deviation and interquartile range), and shape (e.g. skewness and kurtosis). I performed these analyses using the Analyze→Descriptive Statistics→Frequencies options in the SPSS statistical software package. As an output to this step, I built a single table in that included the descriptive statistics for each variable (see Table 4.2). While the primary purpose for this step was to examine my main variables of interest for any abnormalities that could affect the regression analysis, examining the variables in this table also provides value as a standalone product (discussed further in Chapter 4). With so little joint research on birth registration and educational access, these descriptive statistics may be an important first step in how much or how little the phenomena differ across Sub-Saharan Africa.

In the second step of the quantitative data analysis phase, I used inferential statistics, through three multiple regression analyses, to examine the relationships between explanatory variables (e.g. birth registration, GDP per capita, and rurality) and outcome variables, or educational access. In a geographic study, the purpose of inferential statistics is to use a sample of spatially observed data to test hypotheses about a statistical population, or the total set of information under investigation (McGrew Jr. et al., 2014). In the case of my study, I used all countries in Sub-Saharan Africa for which birth registration and educational access data was available to then make inferences about

the relationships among variables across the region. Since the *X* and *Y* variables in my study were random variables (i.e. collected independently of one another) it made sense to employ a quantitative analysis to understand the degree of relationship between them (Howell, 2010). Additionally, through regression, one can investigate both how a set of factors (i.e. called explanatory or predictor variables) explains an outcome (i.e. the dependent, or response, variable) and how that outcome is affected by each individual predictor (Arkes, 2019). The following steps trace my methodological process for both conducting the multiple regression analyses and testing the null hypothesis that there was no statistically significant relationship between birth registration and educational access at three levels of education.

Ensure requirements and assumptions are met. According to McGrew Jr. et al. (2014), there are five requirements for conducting a multiple regression analysis (p. 271). First, the variables must be measured by an interval or ratio scale. All of the variables in my dataset followed the latter (e.g. birth registration %, rurality, access, etc.). Second, the test assumes that there is a linear relationship between the *X* and *Y* variables. As my literature review suggested, my assumption was that there would be a positive linear association where educational access would be higher in countries with higher birth registration rates. Third, the variables are assumed to be normally distributed in a bivariate manner. I tested this assumption during the descriptive statistics step described above when using a histogram to examine the shape of the distribution (i.e. the third class of descriptive statistics). The final two requirements, both related to residuals, calls for the variance of residuals to be equal for all values of the outcome variable and for residual errors to be independent of one another. The latter case was a specified output in

the SPSS software when performing the regression analyses and confirmed to not be an issue in any of my three models. Only after assessing each of these five considerations could I then move on to the three multiple regression analyses.

Analyze the multiple regression equation. When aiming to estimate the effects of birth registration on educational access, it was important to remove the influence of other variables (i.e. country-level wealth and geography) by holding them constant (Arkes, 2019). It does this by allowing researchers to evaluate each independent variable for the amount of total variation they explain in the dependent variable (McGrew Jr. et al., 2014). Therefore, the multiple regression equation, according to McGrew Jr. et al. (2014, p. 270) can be depicted as follows:

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_n X_n$$

In this equation: Y = the outcome variable
a = Y-intercept
β = regression coefficient for each predictor variable
X = the values for each predictor variable

In this study's equation, Y is the access to education composite variable, X_1 represents birth registration, X_2 represents GDP per capita, and X_3 represents the percentage of population that is considered rural for each country. This equation was applied at the primary-school, lower-secondary school, and upper-secondary school levels for three total multiple regression analyses.

Select and compute the test statistic. Arkes (2019) provides a test statistic equation, *t,* for the multiple regression analysis, which I employed in my study. A test statistic provides a value that can summarize the information in your dataset (Sullivan, 2018). This value is then compared against a critical value that determines whether or not

we can accept or reject the null hypothesis. There are a few types of test statistics, like the *z*, *F*, or x^2 (Howell, 2010; Sullivan, 2018). The *t* distribution that I used, which is appropriate for small samples, is similar in shape to a normal distribution but it allows for a greater margin of error due to the smaller sample size (Sullivan, 2018). It is calculated for a regression analysis using the equation outlined by Arkes (2019, p. 19):

$$t = \frac{coefficient\ estimate\ (\beta)}{standard\ error\ (SE(\beta))}$$

Interpret results and determine whether to reject or accept the null hypothesis. After selecting the appropriate test statistic, McGrew Jr. et al. (2014) suggest that researchers then select the level of significance for that test. In my study, I selected a significance level, or α, of .05 because it is a commonly accepted threshold in hypothesis testing. An α of .05 represents a null hypothesis rejection region of 5% of the area under the *t* distribution curve (Howell, 2010; Sullivan 2018). Furthermore, because I did not necessarily have a logical basis to assume a directional relationship between variables (i.e. positive or negative), I selected a two-tailed alternate hypothesis (McGrew Jr. et al., 2014). This meant that my rejection region would be the upper 2.5% *and* the lower 2.5% of the area under the distribution curve at a .05 significance level. 95% of the area under the distribution curve, then, is the non-rejection area, or the area in which we can accept the null hypothesis if the test statistic falls within that region.

I used the *t* value from the previous step to determine whether to reject or accept the null hypothesis. This can be done using a critical value table for the *t* distribution in the appendices of many statistical textbooks (Howell, 2010; McGrew Jr. at al., 2014; Sullivan, 2018). I would use the degrees of freedom, or *n*-1, at .05 two-sided test α, to

look up the critical value for *t.* For example, if my sample size was 35, then my degrees of freedom would be 34 and the critical value would be 2.032. Therefore, according to my decision rule, I would reject the null hypothesis if my calculated *t* value is less than -2.032 or greater than 2.032. Using those results, I would accept the null hypothesis if *t* fell between these two values, which encompasses 95% of the area under the *t* distribution curve. Alternatively, as was the case here, most statistical software packages can perform this calculation and facilitate this process in a much quicker manner. As a reminder, accepting the null hypothesis would mean that the coefficient estimate, β, was equal to zero for each explanatory variable. A rejected null hypothesis would mean that there the coefficient estimate(s) were significantly different from zero (Arkes, 2019). between them. Finally, the SPSS statistical software calculate the *p*-value, or the exact significance level associated with the test (Arkes, 2019; McGrew Jr. et al., 2014). This would enable me to account for Type I error to understand the probability of rejecting the null hypothesis when it is in fact true. As a final output, I create a table showing the coefficient estimates, β, and *p* values for each of the analyzed variables (see Table 4.3). This was a concise way of presenting all of the quantitative results and allowing for a much easier interpretation than if each test had been presented separately (American Psychological Association, 2019).

3.8. Variable Selection (Step 3)

The primary intent behind an explanatory sequential research design is to use a second, qualitative data phase to help explain specific results found during the initial quantitative phase (Creswell & Plano Clark, 2018). In my explanatory spatial design, I

followed this same idea while replacing the qualitative data strand with a spatial data analysis in Step 4. A key step linking the two phases of data analysis (i.e. Step 3) is to determine which quantitative results are in need of additional clarification or explanation, according to Creswell and Plano Clark. This can be done by further examining any significant results, non-significant results, outliers, group differences, or a combination of each. When selecting these criteria for additional insight using a second data strand, the researcher may also need to refine their research questions and re-design the second data phase with potentially new participants. This can pose challenges to approval processes, such as institutional review board (IRB) reviews, as the second phase cannot be fully specified in advance without knowing the quantitative results first. In my case, I did not face such hurdles in the IRB process because I did not need to alter my data collection requirements nor did I modify my spatial analysis techniques or research questions. I simply used the quantitative results to drive the list of variables in need of additional analysis.

The variables that I would select for the subsequent data analysis stage would depend on the acceptance or rejection of the null hypotheses in the preceding quantitative phase. The input to this step, the multiple regression result table, helped me quickly identify which variables would be most suitable for further analysis. The plan was to select any of the regression models in Step 2 where I rejected the null hypotheses and the respective dependent variable(s) displayed a statistically significant relationship with any of the predictor variables across Sub-Saharan Africa. I would then examine these selected variables in the spatial autocorrelation tests to determine if they also showed any instances of non-random spatial distributions, or clustering. The purpose of this

examination would be a pragmatic one, as it could help international organizations identify the areas where possible interventions in birth registration could also impact access to education. Chapter 6 discusses the importance of this case selection process in greater detail as it is a key component of my explanatory spatial variant and offers researchers several advantageous options for subsequent explanation of initial results.

3.9. Spatial Analysis (Step 4)

The spatial data analysis strand followed a similar fashion as the quantitative strand in which the inferential examination was preceded by a descriptive analysis. Steinberg and Steinberg (2015) argue that descriptive mapping provides audiences with a baseline geographic understanding of how each of the independent and dependent variables manifest themselves across space (i.e. Sub-Saharan Africa). This is important because my study was intended primarily for policy, non-governmental and/or educational audiences as opposed to geographers or those with a deep understanding of spatial analyses. Classifying spatial data into visual categories is one way of enhancing readers' understanding of complex spatial patterns (McGrew Jr. et al., 2014). Choropleth mapping is a method by which geographers do this, assigning color shades to polygon map features (e.g. countries) based on their attribute values (Ballas et al., 2017). This is especially useful for quantitative continuous variables with a wide variance, as was the case with birth registration and educational access rates across Sub-Saharan Africa.

As a result of the selections made in in Step 3, I built an individual choropleth map, using the ArcGIS geospatial mapping software, for each of the study's six variables. Multiple individual maps may appear excessive, but it enabled me to quickly discover

unique spatial patterns across the continent. Like SPSS, ArcGIS makes it convenient for users to import one dataset before performing repetitive analyses or processes across variables. I used an equal intervals classification scheme to assign counties a color shade based on which one of five intervals (i.e. range of values) contained their particular variables fell into (Ballas et al., 2017). This method ensured that I was classifying variables consistently across each of the separate maps and making unique values become more visually identifiable. Like the descriptive statistics portion of my quantitative analysis, this step offered me the flexibility to use the maps as a standalone feature for discussion.

In the second phase of my spatial analysis, I tested for spatial autocorrelation within each of the variables selected after the quantitative correlation analysis. Unlike a quantitative correlation, however, spatial autocorrelation is an examination within each variable and does not test for relationships *between* variables (Getis, 2007). Getis (2007) argues that a "spatial twist" on correlation arose from a recognition that many data observations, especially in close proximity to each other, were not necessarily independent (p. 492). In fact, Moran (1948) argued that the geographical distribution of phenomena should be investigated to question whether the presence of some quality in one area made it more or less likely in a neighboring area. The Moran's *I* Index, initiated by his 1948 work, would later become popular among geographers as a method of incorporating the numeric values of areas to quantify the level of spatial autocorrelation among them (McGrew Jr. et al., 2014). This method would enable them to determine if these area attribute values were clustered versus dispersed across a study area and/or if autocorrelations were strong in one region versus weak in another (Bolstad, 2016). This

method is specifically designed to take into account the manifestation of social conditions across adjacent locations (McGrew Jr. et al., 2014). As I mentioned in the literature review, I argue that this type of spatial technique is critical for examining if colonial-era social processes/structures have continued to yield similar results among neighboring countries that share pre- or post-colonial experiences. I used Moran's *I* index to address this inquiry through the following steps.

Ensure requirements and assumptions are met. Similar to the multiple regression analysis, the Moran's *I* index contained a few requirements that needed to be met before proceeding to the analysis step (McGrew Jr. et al., 2014, pp. 230-231). First, there had to be a minimum of 30 geographic features. Even after screening my dataset for countries without adequate data sources, my sample size exceeded this threshold. Second, the attribute values (i.e. the specific numeric values for each variable within each country) must be measured on an interval/ratio scale. I used the same dataset in this test as in the quantitative data strand described above, thereby ensuring that this condition was met. Finally, the same assumption of normality described in my quantitative data screening applies here, as the test is suited for datasets that are assumed to contain a normal distribution. Again, had any of my variables failed to meet this assumption in the quantitative screening then they would also have been excluded from this spatial analysis.

Calculate Moran's I Index. There are a few important terms (see italicized text) incorporated into the Moran's *I* index equation presented below that should first be defined for audiences perhaps unfamiliar with spatial terminology. An *area* is defined in my study as a Sub-Saharan African country. In geographic terms, an area is a polygon feature that represents objects large enough to have boundaries, like countries or tracts of

land (Law & Collins, 2015). Adjacent polygons, or areas, are often *joined* together by a common edge, or a national border (Bolstad, 2016). Two countries that are joined together are known as a *contiguous pair,* or each unique pair of neighboring locations throughout the study area (McGrew Jr. et al., 2014). Additionally, each area generally contains some sort of *attribute value,* which is non-spatial information about a geographic feature, in this case each of the variable values (e.g. birth registration rate, educational access, etc.) corresponding with each country in the dataset (Law & Collins, 2015). I took particular interest in the *variance*, or the degree of variability measured by squared deviation from the mean (Sullivan, 2018). Each of these italicized terms can be found in the Moran's *I* index equation as outlined by McGrew Jr. et al. (2014, p.229):

$$I = \frac{(\text{number of areas})(\text{sum of cross} - \text{products for all continuous pairs } (i,j)}{(\text{number of joins})(\text{variance of the area attribute values})}$$

$$I = \frac{n\sum(x_i - \bar{x})(x_j - \bar{x})}{J\sum(x - \bar{x})^2}$$

In this equation: n = number of areas
$\bar{x}$ = the mean of all attribute values
x_i and x_j = the values of all contiguous pairs
J = the number of joins

Similar to the quantitative data analysis, I relied on a statistical software package to complete the spatial strand of my study. The platform, called ArcGIS, contained the functionality to calculate the Moran's *I* index equation presented above. It is one of the many tools found within the "Geoprocessing" wizard. Geoprocessing is a method by which the GIS system performs an operation to manipulate an input dataset into a new output dataset (Law & Collins, 2015). Specifically, the Moran's *I* option is found within Spatial Statistics Tools → Analyzing Patterns → Spatial Autocorrelation (Global

Moran's *I*). After inputting my dataset into this tool, it returned the Moran's *I* value, the variance, as well as the z-score and p-value, which I will discuss in the next step (Esri, n.db). Again, like SPSS, ArcGIS produces its output results in a customizable results window, which will be presented in Chapter 5.

Select and compute the test statistic. The next step after calculating the Moran's *I* index was to convert it into a test statistic that could be used to determine statistical significance. In this case, McGrew Jr. et al. (2014) guides researchers to convert *I* to a *Z*-score. I therefore used their equation to compute the *Z*-score:

$$Z = \frac{I - E_I}{\sqrt{Var_I}}$$

In this equation: I = Moran's *I* index calculated for each variable in step b.
E_I = -1/(*n*-1)
Var_i = the variance under the assumption of normality. This was calculated by the equation:

$$Var_I = \frac{n^2 J + 3J^2 - n\sum L^2}{J^2(n^2 - 1)}$$

In this equation: n = number of areas (i.e. countries)
J = the number of joins
$\sum L$ = the sum of the number of joins for each individual area (i.e. the total number of countries bordering other countries)

A *Z*-score, or standard score, shows how far the value is from the mean of the distribution (McGrew Jr. et al., 2014). For example, a *Z*-score of exactly 1 would indicate a test statistic of exactly 1 standard deviation above the mean for that particular variable. To understand the probability of this score occurring, or the *p*-value, one can look up the *Z*-score using a table of normal values found in many quantitative textbooks (Howell, 2010; McGrew Jr. et al., 2014; Sullivan, 2018). Alternatively, the ArcGIS software contains the

functionality for this calculation in the same spatial statistics process described above. As such, the latter would become a quality assurance test against the former. Again, this follows a similar process to the quantitative hypothesis test process described throughout Step 2.

Interpret results and determine whether to reject or accept the null hypothesis. The spatial analysis strand required two steps to explain its results. First, it was important to interpret the *I* value because it signals the direction of the correlation and the value indicates the strength of the correlation. With Moran's index, the *I* value of +1 would be a perfectly clustered display of spatial autocorrelation and the *I* value of -1 would be a perfectly dispersed, or anticorrelated value (Bolstad, 2016). In other words, positive values represent autocorrelation, or clustering, and negative values represent dispersion. Again, as with a correlation analysis, it is possible to interpret the degree of correlation using this value on a standalone basis; however, it is the test statistic value (*t*-value or *Z*-score) that determines the level of statistical significance associated with this clustering or randomness (McGrew Jr. et al., 2014).

Second, and the final piece of this data strand, was to determine whether or not to reject the null hypothesis. I used a significance level, or α, of .05 to remain consistent with the statistical probability employed in my regression analysis. If the *p*-value was less than .05 then I would reject the null hypothesis (that attribute values are randomly distributed across the study area) and accept the alternate hypothesis (that attribute values are not randomly distributed). If I rejected the null hypothesis and the Moran's *I* value is positive, then I would be able to confidently say that the values for that particular variable were clustered. If I rejected the null hypothesis and the Moran's *I* was negative, then

there were significant results to show that the attribute values were not randomly distributed and there was some factor that may be influencing their non-random dispersion (McGrew Jr. et al., 2014). Either result certainly warrants future research on this subject, as discussed in Chapter 5. To aid in this interpretation, I created a table that listed all relevant variables, their Moran's *I* value, their *Z*-score, and their *p*-value (see Table 5.1). As I mentioned in the literature review, I was particularly interested to see if there were certain variables that displayed statistically significant evidence of clustering, perhaps along pre- and/or post-colonial national boundaries. This results table of each variable, combined with their associated choropleth maps, provided a baseline analysis for this inquiry.

3.10. Data Integration (Step 5)

Although there were effective standalone components in the quantitative and spatial strands in Step 2 and Step 4, the power of mixed methods research comes when integrating the two types of data (Fielding, 2012). Mixing is an important procedure as it allows the generalizable findings of quantitative research to be connected with a second (often qualitative but in this case spatial) research method, which helps readers better understand complex social phenomena (Zhang & Creswell, 2013). However, integration processes have often been the least discussed procedures in scholarly literature, meaning that mixed methods researchers today should pay special attention to the critical nature of clearly articulating the need for mixing. Integration of traditional mixed methods data sources can be challenging because quantitative and qualitative research have been described as poles on a multidimensional continuum, each with their own distinct

components (Bazeley, 2010). Part of why I am advocating for a new, explanatory spatial, research method is that the spatial analysis is already close in nature to quantitative research, allowing for a smoother integration in practice. In my study, the integration process allows for knowledge creation that is greater than simply the sum of its quantitative and spatial parts.

The integration process in my study consisted of a multivariate cluster analysis technique to effectively merge the quantitative strand, which allowed one to identify statistical relationships between variables, and the spatial strand, which helped illuminate the spatial distributions within the variables. ArcGIS Pro's Multivariate Clustering tool was employed here because it can determine the most natural clusters in the dataset, where the features within each cluster are as similar as possible while also maximizing differences between the clusters themselves (Esri, n.da). As such, it supported Bazeley's (2010) argument that GIS is a particularly powerful analytic and communication tool for the computer-assisted integration of data sources. The goal here was to understand if any variables clustered in concert with one another over the specified geographical space (i.e. Sub-Saharan Africa. This technique has been used before and will naturally fall within a larger body of literature highlighting the efficacy of multivariate methods in geospatial analyses (Long et al., 2010; Marquinez et al., 2003; Shaban et al., 2010; Shan et al., 2013; Srivastava et al., 2012).

Returning to my pragmatic worldview presented in Chapter 2, the cases that were included in the multivariate cluster analysis were intended to yield the greatest opportunity for action or future research. That is, they were the relationships that were most likely to uncover any likely barriers for accessing education and to be

geographically clustered in a manner that could help narrow the scope for policymakers and/or development organizations. In doing so, this technique was designed to answer Research Question #3, about the extent to which birth registration acts as a barrier to accessing education in Sub-Sahara Africa. Chapter 6 will present three useful ways in which different combinations of variables could be integrated in a GIS-based multivariate clustering analysis. Again, this is the power of mixed methods research. Significant quantitative results only show *what* phenomena are occurring but follow-on spatial methods will address a key logical next question of *where* they are occurring.

3.11. Review for Limitations and Validity (Step 6)

I find it critical to interpret the results of my study, from Steps 2 through 5, within the context of the innovative research design and its supporting theoretical constructs. I also want to be fully transparent as this explanatory spatial mixed methods research design may warrant additional scrutiny from academic communities. Recognizing these essential considerations in a mixed methods study, Tashakkori and Teddlie (2008) offer a useful framework, consisting of Caracelli and Riggin's (1994) domains of quality, for assessing the quality of mixed methods research. Their second domain, “Design Quality,” will be especially relevant to my study as my research design comes with some limitations that may affect its reliability. Reliability is defined in this case as the consistency over time and/or the stability of measurement techniques over a variety of conditions (Drost, 2011, p. 108). Following each of my data analysis stages, I examined the methods and results for any unforeseen reliability limitations within my study.

3.11.1. Reliability

From a design quality and reliability perspective, I will point out that this study, like most research, was performed in a time- and resource-constrained environment. In an ideal setting, my study would include comprehensive data samples from every country in Sub-Saharan Africa during the same timeframe for an optimal cross-national comparison. That was simply not feasible as consistency in international data collection processes is rarely perfect from country to county (McGrew Jr. et al., 2014). As I mentioned in the data collection section above, however, UNICEF and UNESCO are very clear and transparent in establishing strict protocol for their data enumerators. I therefore decided to only include cases where their true data sources (i.e. MICS, DHS, etc.) were specified, even knowing that this could have resulted in some missing countries from the analysis. Additionally, these strict guidelines make it impractical to conduct the same surveys across every country in the same year. As such, my database included cases across Sub-Saharan African countries ranging from 2010-2020 to align my timeframe with the most recent rounds of MICS and DHS surveys. This ensured that I was only using data from the most recent data sources in each country to minimize any potential error due to lack of consistency in survey procedure. In other words, both design quality and reliability considerations were actually addressed early in this study by making key decisions in how I assembled my table of secondary data.

3.11.2. Validity

On the topic of validity, Domain 4, “Interpretive Rigor,” in Tashakkori and Teddlie's (2008) mixed methods quality framework considers the degree to which the study’s conclusions are drawn from credible and trustworthy findings. This is similar to

the concept of internal validity, or the validity of the research itself and, subsequently, the researcher's ability to draw correct inferences from their data (Creswell, 2009; Drost, 2011). In mixed methods research, it is critical to address potential threats to validity, which are those that impact correct inferences as well as accurate assessments from the integrated data (Creswell & Plano Clark, 2018). Creswell and Plano Clark note that one threat to internal validity in a mixed methods study comes from keeping results from the data strands separately. This was mitigated in my study by clearly articulating the differing options for mixed methods integration in Step 5 (described above) and again in a more thorough review during Chapter 6. These sections show how I managed to investigate both the quantitative and spatial data strands in an integrated manner. Another threat to validity, they describe, especially in a sequential design, is failing to identify the most important results from the first strand to explain with the second strand. This threat is why I have outlined clear case selection protocol. In doing so, readers should be able to infer that the multivariate maps presented in Chapter 6 are the result of the most noteworthy and closely associated variables.

The final domain of quality that is worth mentioning as a key consideration in my mixed methods study is the fifth domain, "Inference Transferability" (Tashakkori & Teddlie, 2008). This is equivalent to the concept of external validity in a traditional quantitative study (O'Cathain, 2010). External validity is typically an examination of the threat posed by generalizing and transferring inferences across contexts (Drost, 2011). In my study, I will again be clear that my intent was not necessarily to follow a traditional research model of using the study to make inferences across other global contexts. In fact, the reason I included a spatial component is because I believe that socio-spatial context,

particularly geographic clustering, is paramount for examining the relationships among social phenomena. This study originated from a place of post-colonial inquiry and I would find it inappropriate to transfer the results from this study to other situations that did not suffer from such colonial legacies as the African continent. As such, my strategy for addressing political legitimation, another term for mixed methods validity, was to ensure that I maintain a neutral perspective throughout the research process that considers how stakeholders and policymakers can apply this research towards the generation of actionable solutions (Onwuegbuzie & Johnson, 2006). As the last step in this process, I used Tashakkori & Teddlie's (2008) mixed methods quality framework and Onwuegbuzie & Johnson's (2006) typology of mixed methods legitimation types to perform one final screening for limitations. As such, limitations will be discussed in each of the dissertation’s following three journal articles (i.e. Chapters 4-6).

CHAPTER 4: MEASURING THE RELATIONSHIP BETWEEN BIRTH REGISTRATION AND EDUCATIONAL ACCESS IN SUB-SAHARAN AFRICA (ARTICLE #1)

4.1. Introduction

As the world becomes increasingly connected, whether through social processes or intertwined financial systems, access to a quality education has often become a prerequisite for opportunity. Sustained access to education, according to Little and Lewin (2011), is essential for poverty reduction and inclusion, making it a key component of development policy. Despite improvements in global progress towards the United Nations' Sustainable Development Goal (#4) of ensuring inclusive and equitable lifelong learning opportunities for all, millions of young people remain left behind in the pursuit for quality education (United Nations Economic and Social Council, 2020). The "access challenge" remains greatest in Sub-Saharan Africa, where current (2019) enrollment of young people in pre-primary, primary, and secondary education is several times less than the number it would need to achieve universal enrollment goals by the year 2030 (United Nations Children's Fund, 2019a). Recognizing this trend, nearly 40% of educational research in Africa from 2010 to 2018 addressed the thematic areas of "access to education" or "equitable, inclusive education" (Mitchell et al., 2020). Evidence of education disparities by gender, socioeconomic status, and geography shows that Sub-Saharan Africa (SSA) remains as a region warranting ongoing research into inclusive visions of access (Kuépié et al., 2015; Lewin, 2009; Majgaard & Mingat, 2012).

This article seeks to fill a void in education literature by addressing an under-researched relationship between birth registration and access to education. Theoretically, there are grounds to believe that birth registration could be added to the aforementioned

list of disparities, where those children lacking birth certificates become left behind. Millions of children across the world lacking official records of their birth are considered "missing" or "invisible" from state institutions without such critical identity documents (Ajmera & Fields, 2016; Sumner, 2015). The figure of children under the age of 5 without a registered birth – a key United Nations indicator for the monitoring of child protection – was estimated to be at 166 million as of 2019 (United Nations Children's Fund, 2019e). Although there have been some instances where birth registration precluded children from furthering their education beyond primary school, including the author's personal observations in Cameroon, it has largely been absent from substantive academic research (Civil Registration Centre for Development, 2011; Crush & Tawodzera, 2014; Kindzeka, 2017). Empirically, only Corbacho et al. (2012) and Apland et al. (2014) offer rigorous glimpses into birth registration processes from a systematic educational exclusion perspective. This relationship has therefore yet to be tested in a Sub-Saharan Africa context, which is a region that lags behind other world regions in both birth registration levels and educational outcomes (Roser & Ortiz-Ospina, 2016; United Nations Children's Fund, 2019e).

Sub-Saharan Africa will serve as the area of focus for this quantitative study, as it is unclear whether the countries with low birth registration levels are also the same countries where access to education is limited. This study sought to answer the question: *Is there a relationship between birth registration and educational access across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?* A multiple linear regression model will help describe the relationship between birth registration and access to education in Sub-Saharan Africa from 2010-2020. The

methods section will both outline this empirical strategy and address the manner in which this dataset was created and analyzed using openly accessible secondary data from the United Nations Children's Fund (UNICEF) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO). In the results section, descriptive statistics will illuminate important trends in access by educational level while the regression models will demonstrate an insignificant statistical relationship between the key variables. The discussion section will situate these results within the scant body of literature on the topic, highlighting limitations in international data accessibility. Finally, intended for international/comparative education researchers and intergovernmental policy experts, the article will conclude with research and policy recommendations articulating the need for a spatial understanding of social and educational exclusion.

4.2. Background

A number of barriers exist at the primary-, secondary-, or tertiary-level education levels across Sub-Saharan Africa, as access is far from being an equally shared commodity, especially in a post-colonial context (Tikly, 2001). In fact, proxy measures of access have allowed researchers to investigate any barriers to it across Sub-Saharan Africa (Lewin & Sabates, 2011). In arguably one of the most detailed analyses on this topic over the past decade, a *World Bank* report provided evidence of several inequities among various populations, like girls or financially disadvantaged, that became widened with each higher level of education beyond primary school (Majgaard & Mingat, 2012). This was consistent with a *Save the Children* report about conflicted affected fragile states, where poverty, urban/rural setting, gender, disability status, age, language, and displacement were all identified as individual- or group-level barriers affecting access to

education (Dryden-Peterson, 2009). These fragile states, which can describe many of the conditions in Sub-Saharan Africa, "…are countries that are impacted by conflict, income disparity, weak governance, and/or inequality in resource allocation" (p. 21). Even sub-regional and country-level studies have indicated disparities in their education systems along these similar lines (Buckland, 2011; Kuépié et al., 2015). Relatively absent from this seemingly endless list of barriers, however, is birth registration status, which is interesting since birth certificates are often used as "passports" to key social services, such as education or health care (Cappa et al., 2014).

Lacking a birth certificate or growing up without having had your birth registered is one unfortunate way that individuals can be systematically excluded from the many social benefits of being a country's citizen. Individuals in these cases are effectively "stateless," and are deprived of accessing their legitimate claims to social services (Bhabha, 2011). In fact, without such legal identity documentation, individuals have been shown to suffer significant effects from their social exclusion, ranging from ongoing conditions of poverty (Harbitz & Tamargo, 2009) to detrimental impacts to their human development (Jeong et al., 2018). This is especially devastating as there have been countless examples where, due to high costs of procurement or travel distances to civil registrars, low levels of birth registration were tied to disparities among several vulnerable populations, like indigenous populations, ethnic minorities, and those in the poorest wealth quintiles (Dake & Fuseini, 2018; Hunter & Brill, 2016; Hunter & Sugiyama, 2018; Issa & Mgaya, 2018; Mackenzie, 2008). One noteworthy example highlighting the detrimental effects of low birth registration is that it has displayed negative relationships with undernutrition in children and lower child vaccination levels

(Brito et al., 2017; Comandini et al., 2016). If birth registration status is able to affect such an essential human right as access to healthcare, then it is understandable to assume that it could also impact one's access to an education as well.

When examining these patterns of exclusion from a postcolonial lens, it becomes clear that some of these social processes may not be an accident. In fact, non-recognition was a characteristic of colonized populations, as civil registration and legal identity processes were rooted in systems created by and for those in power during the colonial era (S. Q. Jensen, 2011; Szreter & Breckenridge, 2012; Villet, 2011). A postcolonial lens in educational research allows us to see the perspective of the excluded 'Other' populations in the Global South to understand the consequences of such colonial legacies (Andreotti, 2011; Gandhi, 1998). The colonial era left a wave of marginalizing effects on identities and communities, as the artificially drawn colonial boundaries have literally shaped socio-cultural constructs crossing linguistic and ethnic geographies (Ali et al., 2019; Müller, 2020; Touval, 1967). Civil registration systems, which often maintain a country's birth registration processes, is one institution that continues to witness logistical and administrative challenges in Africa (Bequele, 2005; Bhatia et al., 2019). Many in this region, especially, still live in areas with weak systems of civil registration (Mikkelsen et al., 2015; Mills et al., 2017; Suthar et al., 2019). Thus, if these systems are tied to access with vital services, then the degree to which they act as a barrier to vulnerable populations is certainly deserved of scholarly attention.

The underlying theoretical argument driving this study was that some populations in Sub-Saharan Africa might be confined to similar geographies of exclusion, where they are both lacking birth registration and an access to education. Not only do civil

registration systems have histories shaped by colonial powers, but the same might also be exhibited in African countries' education systems. In some cases, education systems in Africa are structurally similar to their colonial-era versions (Shizha, 2005) and, in other cases, the colonial-era investments in education systems (particularly in French-speaking West Africa) have ties to modern educational outcomes (Huillery, 2009). Cameroon represents a clear example of the lasting effects of colonization, as its colonial partition among the French and British has resulted in parallel educations systems within the country well into the 21st century (Dupraz, 2019). In terms of educational access, then, this study sought to understand if country-specific situations, like that in Cameroon, witnessed similar characteristics as other countries with similar colonial experiences. The idea was to expose certain social structures, conditions, or barriers, like birth registration, that were actually more regularly manifesting themselves across the region than previously known.

4.3. Methods

4.3.1. Data Collection

This study relied upon a custom dataset developed by combining secondary data from two United Nations entities – the United Nations Children's Fund (UNICEF) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO). UNICEF is the custodian agency for birth registration and other child protection indicators, whereas UNESCO's Institute for Statistics (UIS) is the official source for internationally-comparable education data (United Nations Children's Fund, 2020a; United Nations Educational, Scientific and Cultural Organization, 2020c). Covering nearly 200 countries and most of the developing world, according to Murray and Newby

(2012), these UN data sources enable researchers to effectively conduct cross-national analyses. With Sub-Saharan Africa as the specified region, a cross-national analysis aims to differentiate country-specific peculiarities from social-structural regularities (Kohn, 1987). In other words, this data may help determine if birth registration acts as a barrier to accessing education in a systemic manner throughout the region or not.

Merging indicators from multiple data structures required a few screening decisions to ensure that any subsequent analyses would be conducted using complete, clean, and consistent data across the region. This process began by exporting birth registration and education data from UNICEF for the 49 countries of Sub-Saharan Africa. A quick visual scan helped screen out countries with missing data across multiple variables. Table 4.1 displays the omitted countries and the final countries included in the regression analysis. The second step was to ensure that the dataset included only the most recent data points for each country during the timeframe of 2010 to 2020. For reliability purposes, all of the included UNICEF data needed to be sourced by the UNICEF Multiple Indicator Cluster Surveys (MICS) or the U.S. Agency for International Development's (USAID) Demographic and Health Surveys (DHS) Program and the year needed to be consistent (+/- one year) across variables for each country. These programs are similarly structured, nationally-representative, household survey programs that are conducted on multi-year frequencies in most low- and lower-middle income countries (Ayede et al., 2018; Hancioglu & Arnold, 2013). This meant that some countries' most recent data dated back to 2010 as they have not had a more recent set of surveys across each indicator. Finally, a VLOOKUP function was used to combine socioeconomic and demographic indicators for each country (e.g. gross domestic product per capita and rural

population %, respectively) from the UIS database with the UNICEF data. A final missingness screening was then conducted and the complete dataset ultimately consisted of *n*=40 countries across Sub-Saharan Africa for the data analysis phase.

Table 4.1: List of countries in Sub-Saharan Africa by inclusion or omission in this study

Included Countries (*n*=40)	**Omitted Countries (*n*=9)**
Angola, Benin, Burkina Faso, Burundi, Cote D'Ivoire, Cameroon, Central African Republic, Chad, Comoros, Congo, Democratic Republic of the Congo, Eswatini, Ethiopia, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Sierra Leone, South Africa, South Sudan, Togo, Uganda, United Republic of Tanzania, Zambia, Zimbabwe	Botswana, Cabo Verde, Djibouti, Equatorial Guinea, Eritrea, Gabon, Mauritius, Seychelles, Somalia

Access to Education. Effective and congruent research guiding educational policy hinges on at least some degree of commonality in the definition of access. In an exploration of policy options to improve the delivery of basic education in developing countries, Chimombo (2005) recognized that access could be defined in terms of opportunity costs, gender, cultural attitudes, quality and effectiveness. In that sense, access to education could appropriately be conceived as a combination of *inputs*, like enrollment or attendance, and *outcomes*, like retention or completion. Lewin (2009) advances a broad vision for access to education that includes access to acceptable staffing and learning materials, progression through educational levels to successful age-appropriate completion, consistent attendance, and access to post-primary opportunities. This wide understanding of access provides the subject with a strong theoretical foundation. A narrowed scope, however, offers perhaps a more suitable empirical

approach as evidenced by research strategies implementing composite measures of educational access (Spaull & Taylor, 2015). As such, access to education here will henceforth be considered a composite of those attending school (i.e. an input measure), those not attending school (i.e. an indicator of exclusion), and those who have completed school (i.e. an outcome measure). This article will therefore investigate factors that could be acting as barriers to or enablers of a multifaceted conception of access.

Access to education was the outcome variable in this study and was analyzed at the primary, lower-secondary, and upper-secondary school levels. This enables readers to understand how access and its relationship with socioeconomic and demographic conditions can vary by becoming more restrictive with each successive level of education. This indicator was created using a composite variable of net attendance, completion, and out-of-school rates for each country in the dataset. This is the first study to operationalize access in this manner but it was influenced by Spaull and Taylor's (2015) composite measures of educational quantity and quality. Simple averaging is a commonly used approach to create composite variables and was used here to reduce the number of separate dependent variables by using one score for access (Song et al., 2013). The first input into this average score was net attendance, which is arguably the most straightforward indicator for access as it directly assesses the degree to which students are enrolled in school compared to the total population of school-aged children (United Nations Children's Fund, 2019a). Completion rates, on the other hand, are useful when considering access from an educational quality perspective, as the student life cycle and can account for high incidence of drop out, low retention, or late completion (Spaull & Taylor, 2015; United Nations Children's Fund, 2020a). Finally, and perhaps most

importantly, an out-of-school rate (using an inverted score) was used to address those *excluded* from their education systems for entire academic years (UNICEF 2016). Together, these variables are not only better suited for a quantitative analysis free of multiple testing concerns, but they also could provide a useful framework when considering a comprehensive idea of access in the future.

Birth Registration. Birth registration was the key predictor variable at the center of this study, as the intent was to better understand how it related to educational access. Unlike access to education, which required manual data manipulation to create a composite variable, birth registration was a standalone variable in both the DHS and MICS surveys. It was useful for this study as it is arguably the most well-established global indicator related to birth registration. According to UNICEF (United Nations Children's Fund, 2019e), the standard indicator for birth registration is the percentage of children under age 5 (0 to 59 months) with a birth certificate or whose birth was reported as registered with civil authorities at the time of survey. This variable is calculated by dividing the total number of children in the surveyed population by the number of the number of children with a birth certificate or registration with civil authorities. Therefore, countries with higher values likely have more well-functioning civil registration and vital statistics (CRVS) systems (Mikkelsen et al., 2015). This study was built on a hypothesis that the countries with stronger CRVS systems (i.e. a government institution) could presumably also have stronger public education systems.

Controls. The model selection in a regression analysis is critically dependent upon an optimal selection of control variables that limits omitted-variables bias (Arkes, 2019). Although birth registration is the key explanatory variable in this study, additional

hypothesized barriers were needed to control for the diverse demographic and socioeconomic conditions across Sub-Saharan Africa. There are differing opinions on the optimal number of additional predictor variables with some creating specific formulas based on sample size (Oliker, 1978) while others argue for as small a number of variables as possible (Mahmoud et al., 2015). In general, Stevens (1996) recommends at least 15 cases per predictor variable, which means that 3 predictor variables for this study, where n=40, is sufficient. In this study, two UIS-provided indicators, gross domestic product (GDP) per capita and rural population percentage, were used in addition to birth registration as possible explanatory measures of access to education. Not only are these indicators supported by literature, but they also did not correlate highly with one another in a correlation matrix, reducing the concern of multicollinearity issues. The former, intended to be a proxy for socioeconomic status of each country, has been linked with educational coverage (another synonym for access) at the post-primary education levels (Majgaard & Mingat, 2012). Similarly, research into the latter suggests that there may be something about rurality that affects access to education, particularly in developing countries (Schafft & Jackson, 2010; Vasconcellos, 1997). One purpose of this article's regression model was to understand if these conditions have remained true following the second decade of the 21st century.

4.3.2. Data Analysis

A multiple regression model was used to demonstrate the strength and direction of the relationships between one key explanatory variable (birth registration), two control variables (GDP per capita and rural population), and one dependent variable (access to education) for n=40 countries across Sub-Saharan Africa between 2010 and 2020 (Arkes,

2019). In doing so, the regression model sought to identify determinants/correlates across these relationships (Wang et al., 2017). A multiple regression model accomplishes this by allowing researchers to evaluate each independent variable for the amount of total variation they explain in the dependent variable (McGrew Jr. et al., 2014). The multiple regression equation, according to McGrew Jr. et al. (2014) can be depicted as follows:

$$Y = a + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_n X_n$$

In this equation, *Y* is the dependent, or outcome variable, *a* is the Y-intercept, *β* is the regression coefficient for each predictor variable and *X* is the value for each predictor variable (p. 270). In this study's equation, *Y* is the access to education composite variable, X_1 represents birth registration, X_2 represents GDP per capita, and X_3 represents the percentage of population that is considered rural for each country. Discussed in further detail below, this equation was applied at the primary-school, lower-secondary school, and upper-secondary school levels for three total calculations.

The objectives from this multiple regression were to examine and potentially uncover noteworthy results using both descriptive and inferential statistics. From a descriptive perspective, the means and standard deviations, two measures of central tendency, were important indicators for describing the location and dispersion, respectively, of each variable across the countries in Sub-Saharan Africa (L. M. Sullivan, 2018). Doing so allowed one to make descriptive observations about trends in accessibility and birth registration. On the inferential side, a hypothesis test was used to determine if the coefficient estimates, β, for each variable were statistically significant (Arkes, 2019). In this situation, the null hypothesis, H_0, is that β=0, while the alternate hypothesis, H_1 is that β≠0. The *p*-value, or threshold for rejecting the null hypothesis, was

set at a 5% significance level, or $p<0.05$. Hypothesis testing was critical for understanding if any relationships between explanatory variables – birth registration, GDP per capita, and rurality – and access to education were due to chance. Therefore, descriptive and inferential statistics possess the ability to provide key information about these variables in their own right while also combining with one another to further explain such relationships.

Using the SPSS statistical software, the multiple regression analysis followed Pallant's (2020) guide for performing the tests and interpreting their outputs. The first step was to check for multicollinearity, which was not apparent in this study as there were no bivariate correlations greater than .7 nor were there any issues with tolerance or variance inflation factor (VIF). Next, an examination of the residual boxplots yielded no major deviations from normality nor abnormal, systematic, patterns in the residuals. A subsequent screening for outliers, however, uncovered one data point with Mahalanobis and Cook's distances that were far too high when compared to the rest of the dataset. As such, the country of Gabon was removed and the regression analyses was run for a second time. The analysis then required an evaluation of both the regression model itself and each independent variable, individually, to assess statistical significance and size of the coefficient estimates, β. Finally, the outputs were used to produce regression equations that allow researchers to examine if birth registration affects access to education across Sub-Saharan African, after controlling for GDP and rurality. The next section outlines each of these results followed by a discussion that situates them within the context of international education and development literature.

4.4. Results

4.4.1. Descriptive Statistics

The descriptive statistics (see Table 4.2) indicate that education becomes more inaccessible with each higher level beyond primary school. In the access to education composite variable, the unit of measurement is percentage points (i.e. a range from 0 to 100), where the higher number represents greater access. Access to education shows a declining trend of mean values with each educational level, as the mean at the primary school level represents a 71.0% accessibility value, compared to 47.8% and 31.5% at the lower- and upper-secondary education levels, respectively. Similarly, the standard deviations follow the same trend, ranging from 16.1% for primary schools, to 15.1% for lower-secondary schools, to 12.6% at the upper-secondary level. Birth registration and rural population, also measured by percentages, displayed similar mean values of 56.9% and 61.2% respectively.

Another way to describe the dataset is by directing attention to the large degree of intra-region inequality, or variability, associated with socioeconomic/demographic indicators. The standard deviation in rural population (SD=15.3) is nearly half as small as birth registration (SD=26.4), however. GDP per capita, measured in U.S. dollars, showed a mean of $1,533 and a standard deviation of $1,319. In other words, there was tremendous variability here, as the standard deviation was almost as great as the mean itself. The range for GDP per capita, from $275 (Burkina Faso) to $6,132 (South Africa), had a skewed right distribution of the variable, illuminating this trend of inequality. Finally, it is worth noting that all other variables were normally distributed.

Table 4.2: Descriptive statistics

Variable (n=40)	*M*	*SD*
Access to Education – Primary	71.01	16.12
Access to Education – Lower Secondary	47.76	15.06
Access to Education – Upper Secondary	31.51	12.64
Birth Registration	56.92	26.44
GDP per Capita	1,532.63	1,319.18
Rural Population	61.23	15.31

Note. Unit for all variables except for GDP per capita (US$) are percentages (%)

4.4.2. Multiple Regression Results

Although the statistical significance of individual predictor variables is more commonly reported than the overall significance of the model, the latter is noteworthy here (Arkes, 2019). Table 4.3 contains the results of this study's multivariate regression analyses. When considered together, the three predictor variables – birth registration, GDP per capita, and rural population – explained a portion of the variance in educational access in a statistically significant ($p < 0.05$) manner across all three multiple regression models – primary school, lower-secondary, and upper-secondary – in Sub-Saharan Africa. The adjusted R^2 indicator was more appropriate for this study than the R^2 value because of the small sample size, where n=40 (Pallant, 2020). The three explanatory variables explained 13.3% of the variance in educational access at the primary-school level ($p < 0.05$). They also explained 37.6% of the variance in access to lower-secondary school ($p < 0.001$) and 35.1% of the variance in access to upper-secondary school ($p < 0.001$). In other words, the three predictor variables account for over two times the amount of variance in access to education in secondary school than compared to primary school.

Table 4.3: Results of each multiple regression analysis

	Access: Primary	Access: Lower-Secondary	Access: Upper-Secondary
(Constant)	68.946	54.575	40.551
	(14.948)	(11.844)	(10.147)
Birth registration	0.044	0.008	-0.055
	(0.098)	(0.077)	(0.066)
GDP per capita	0.004**	0.006**	0.005**
	(0.002)	(0.002)	(0.001)
Rural population	-0.117	-0.258*	-0.217*
	(0.181)	(0.144)	(0.123)
R-Squared	0.200	0.424	0.400
Adjusted R-squared	0.133	0.376	0.351
No. observations	40	40	40

Note. Parenthetical values represent the standard errors. Asterisks, *, **, indicate significance at the 90%, and 95% levels, respectively.

Multivariate regression analyses are useful because they can uncover which predictor variables in a model are making statistically significant unique contributions to the equation (Pallant, 2020). In this case, the coefficient estimate (β), for the key explanatory variable, birth registration, was not significant in its relationship with access to education at any education level. Furthermore, the coefficient estimates for rural population % were not significant in any model at a 5% significance but they were significant at a 10% level at the lower- and upper-secondary school levels ($p = 0.081$ and $p = 0.085$, respectively). The coefficient estimates for GDP per capita, on the other hand, were statistically significant in all three models. Therefore, the null hypotheses could only be rejected for GDP per capita in this study as there was statistically significant evidence that its coefficient estimates were different from 0 (Arkes, 2019). Holding other variables constant, then, countries with higher GDP per capita can be expected to also

display greater access to education. Finally, the coefficient estimates for all variables in each model were very small in absolute terms ($\beta < 0.3$). For example, a one percent increase in birth registration would only increase primary school access by 0.044%. Interestingly, the coefficient estimates were largest (ranging from -.12 to -.26) in the rural population variable while also carrying negative values. This meant that for each percentage point increase in rural population, access to education was expected to decrease.

The final component of the multiple regression analysis was to insert the results into updated regression equations for access to education at each level of schooling. Modifying the model from McGrew Jr. et al. (2014), the three regression equations can be expressed as:

$$Y_{primary} = 68.9 + 0.044_{birth} + 0.004_{GDPcap} - 0.117_{rural}$$
$$Y_{lower} = 54.6 + 0.008_{birth} + 0.006_{GDPcap} - 0.258_{rural}$$
$$Y_{upper} = 40.6 - 0.055_{birth} + 0.005_{GDPcap} - 0.217_{rural}$$

In these models, *Y* is the outcome variable at the three levels of education (marked by italicized subscript), followed by the Y-intercept and the regression coefficient for each explanatory variable. Although these models are unique to Sub-Saharan Africa between 2010 and 2020, they offer a useful side-by-side comparison to identify trends by education level. Not only does the decreasing Y-intercept become apparent by successive education levels, but so too does the consistently small β values for GDP per capita and the negative sign for rural coefficient estimates. Considered together, the descriptive statistics and these multivariate results offer unique contributions to the educational policy literature, as shown below.

4.5. Discussion

In some countries, access to various state services and social protections is dependent upon the possession of a birth certificate, which is especially problematic in developing countries, where weak civil registration systems have failed to provide millions of individuals with their rightful legal identities (Dahan & Gelb, 2015a; Hunter & Brill, 2016). The purpose of this article was to examine the accessibility of one of these services, public education, in the context of birth registration in Sub-Saharan Africa between 2010 and 2020. Holding geographic and socioeconomic factors constant, this article advanced a multiple regression model that allowed for comparisons at successive levels of education – primary, lower-secondary, and upper-secondary. While the model did reveal significant and nearly-significant evidence that GDP per capita and rurality, respectively, affect access to education, the key explanatory variable of birth registration failed to yield statistically significant evidence worthy of rejecting a null hypothesis. Despite the lack of a unique statistical contribution, birth registration nonetheless remained a component of a regression model that explained 30% - 40% of the variance in educational access at multiple levels. This consideration, when paired with descriptive statistics, certainly illuminates how access to public education becomes more exclusive beyond primary school. The rest of this section will situate these key findings among the relevant literature, setting the stage for a conclusion that offers clear research and policy recommendations.

4.5.1. Birth Registration

The fact that birth registration did not have a statistically significant impact on access to education was not consistent with results from the few other studies with similar

aims. For example, Apland et al. (2014) found that participants' possession of a birth certificate made them more likely to have attended age-appropriate education and less likely to discontinue their education in India, Kenya, and Sierra Leone. Similarly, Corbacho et al. (2012) found significant evidence in the Dominican Republic that children without birth certificates had lower educational attainment than those with them and were especially likely to be excluded during the transition between primary and secondary school beyond grade 8. In fact, the model presented in the study on Sub-Saharan Africa was built on this assumption that birth registration might inhibit access to education beyond the primary-school level. The assumption resulted from the author's personal experience in Cameroon and a finding by Corbacho et al. (2012) that initial access to the educational system (i.e. entry to primary school) was not correlated with birth registration. This means that the transition from primary to secondary school in Sub-Saharan Africa was not impacted by birth registration as expected based on prior research.

Despite findings that were contradictory with prior research, this study makes a unique contribution to international education and development literature. First, the scope of the study, with a regional emphasis on Sub-Saharan Africa, was much wider than those of Apland et al. (2014) or Corbacho et al. (2012). Ambitious efforts are necessary, however, to understand the extent to which birth registration can hinder educational access beyond the limited examples in India, Kenya, Sierra Leone, and the Dominican Republic. After all, the combination of academic research in these examples as well as news stories of this problem occurring in Cameroon, Zimbabwe, and Vietnam suggest that there is a need to address it within their geography-specific contexts (Crush &

Tawodzera, 2014; Kindzeka, 2017; Lam, 2001). Again, there was justified reason to assume that geography mattered when it came to institutional processes (e.g. civil registration and education systems) because many education systems in Sub-Saharan Africa have remained structurally similar to their colonial-era forms (Shizha, 2005) and because these systems have generated common structures of schooling and pedagogy (Tikly, 2001). In that case, perhaps a sub-regional (e.g. West Africa, East Africa, etc.) might be a more appropriate unit for a follow-on analysis. There still could be a relationship between birth registration and access to education in one of these smaller units that became overshadowed by the quantitative analysis that included too many countries in dissimilar areas.

The fact that birth registration did not offer a statistically significant contribution to the variance in educational access does not mean that we can dismiss a relationship between these variables altogether. After all, some authors have recently argued that finding statistical significance should not necessarily be the goal of regression analyses as there are other meaningful considerations of inquiry (Arkes, 2019; McShane et al., 2019). Recognizing the absence of significance in this study, as well as the utility of alternative methods for examining relationships between variables, scatterplots were employed here to visually inspect how birth registration related to access at each educational level. As expected (based on the regression analyses), none of the scatterplots demonstrated strong or even moderate correlations, a measure of effect size, between variables (G. M. Sullivan & Feinn, 2012). However, the scatterplot at the primary school level revealed a fascinating cluster of data points representing countries that neighbor one another, geographically (see Figure 4.1).

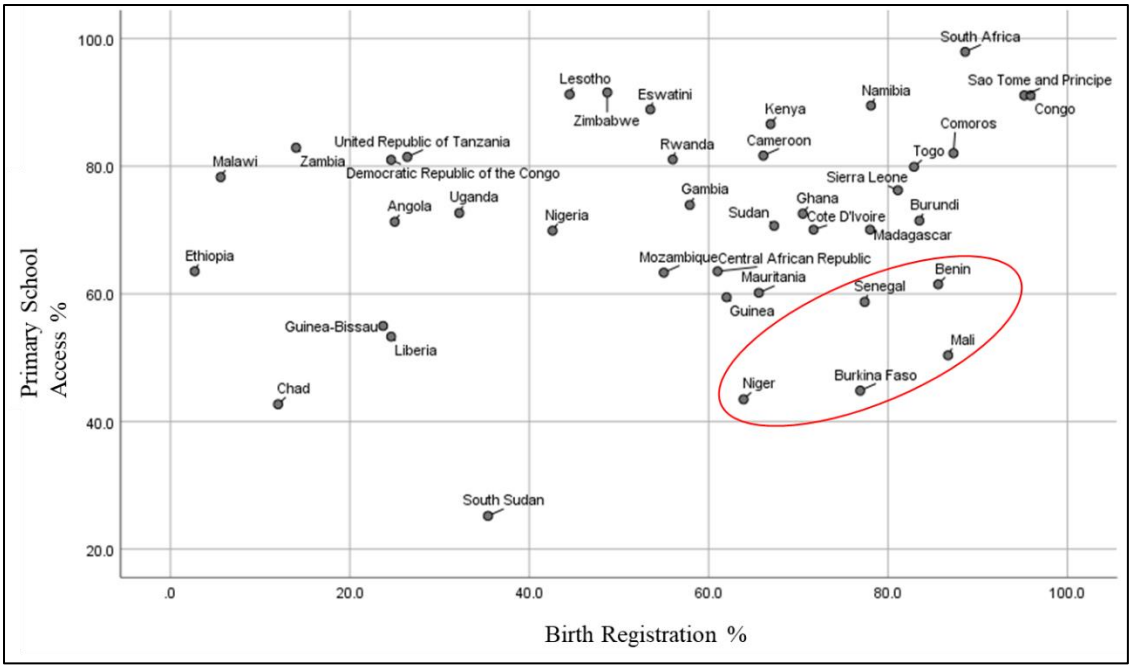

Figure 4.1. Scatterplot of Birth Registration and Access to Education at the Primary-School Level

The encircled countries of Benin, Senegal, Mali, Burkina Faso, and Niger, are all West African countries that lie in one contiguous pattern, spatially. They also appear to demonstrate similar conditions where birth registration is relatively high (compared to the data set) yet access to primary education is relatively low. This creates an interesting case for a multivariate spatial analysis, where one can examine how the variables relate to one another over a geographic space. In turn, this might help us understand why there are some clusters of countries, like those encircled in Figure 4.1, that have similar circumstances pertaining to these variables of interest. In other words, is there something about the human geography of this area that results in lower educational access despite similar birth registration levels as other countries in the region? This question may not have arisen without an initial quantitative analysis and it is beyond the scope of this paper, though it is covered in a subsequent and related spatial article (Griffin 2021c).

4.5.2 Explanatory Variables

The statistically significant finding that GDP per capita was positively associated with access to education aligned with previous research and further strengthens that body of literature, both generally and in Africa, specifically. A few key global studies, with over 80 countries each, provided empirical evidence that the economic characteristics of countries both limited access to secondary education and overall years of schooling (Checchi, 2003; Thomas et al., 2001). While Checchi's study used GDP per capita as a control for the degree of development, Thomas et al.'s paper focused on the variable as the key explanatory variable. In the latter case, increases in per capita GDP were positively related to the labor force's average years of schooling, an indicator that country-level wealth does relate to educational access. The findings presented in this study on Sub-Saharan Africa are noteworthy as GDP per capita demonstrated a statistically significant relationship, albeit a very small one, with access to education at the primary- and secondary-education levels.

In an African context, more recent studies have demonstrated similar relationships between socioeconomic status, levels of development and school attendance/quality in Nigeria and Botswana (Kazeem et al., 2010; Mbulawa & Mehta, 2016). Interestingly, Mbulawa & Mehta's (2016) paper showed a unidirectional causality between economic development and access to education in Botswana, whereas Oketch (2006) found economic growth (measured by GDP per capita growth) to have a two-way causal flow with investment in education. This is an important consideration because it provides ample reason to suggest that educational investments may also need to be included in any future studies about access to education in Africa. The main question, here, is to

understand if access to education and educational investment have any degree of meaningful relationship. If so, the study on birth registration and educational access might be worth revisiting with educational or other social service expenditures additional controls depending on the accessibility of quality data. This will be discussed further in the recommendations for research below.

In light of the findings surrounding GDP per capita, it is also not surprising that rurality affected access as those from poorer and more rural households have been shown to have severely limited access to education (Majgaard & Mingat, 2012). This urban/rural divide in access to education may not be a simple function of geography (e.g. distance/travel time to primary- or secondary-school), however, as Kazeem et al. (2010) found the gap to shrink after controlling for household wealth. From a similar critical lens towards the urban/rural geographies, we should be careful from concluding that rurality affected access to education at any particular education level despite the fact that some of the rural population values were significant at the 10% threshold, as described in the previous section. There have been some discrepancies in prior research, for example, as some have argued that the urban/rural gap in educational access is greater in pre-tertiary (i.e. secondary) education (Anlimachie & Avoada, 2020; Lewin, 2009), while others found that access was widest among the youngest groups of rural children (Kuépié et al., 2015). A deeper investigation into geography is beyond the scope of this paper, although it is clear from a brief scan of literature that gender gaps in educational access should go hand-in-hand with those examining social conditions of rural women across Sub-Saharan Africa, in particular (Ingutia et al., 2020; Porter, 2011). Furthermore, the educational experience found within rural African communities is a social justice issue and that,

"…all endeavors to social injustices with regard to rural education should be characterized by difference, attributable to diversity in rurality" (Hlalele, 2012, p. 117).

4.5.3. Limitations

From a statistical power standpoint, this study was limited by a relatively small sample size for a multiple regression analysis (n=40). Although Jenkins and Quintana-Ascencio (2020) argue that regression analyses can be used at $n \geq 25$, the small sample limited the number of independent variables that could appropriately be used, as described in the methods section above. Therefore, this study could only control for one socioeconomic (e.g. GDP per capita) and one demographic (e.g. rural population %) variable. This left out a variety of possible explanatory variables that may be more appropriate for the contexts specific to Sub-Saharan Africa. For example, the study by Apland et al. (2014) noted that mothers' education levels, access to better living conditions in the household, and gender of the student all mattered when presenting statistically significant relationships between birth certificate holders and those attending age-appropriate education. However, a trade-off occurs here because their study was limited by the fact that it was not nationally representative of the countries' populations for its three-country study whereas this paper was able to look comparatively across a region (e.g. Sub-Saharan Africa). Given that large disparities in birth registration data exist globally, these decisions will likely continue to plague researchers focused on low- and middle-income countries (Phillips et al., 2018).

An additional limitation in this study is the omission of 9 countries out of 49 total countries in Sub-Saharan Africa. As mentioned above, Botswana, Cabo Verde, Djibouti, Equatorial Guinea, Eritrea, Mauritius, Seychelles, and Somalia did not meet the screening

criteria and were subsequently excluded from the regression analyses. While some of these are island nations (Cabo Verde, Mauritius, and Seychelles), others (notably Djibouti, Eritrea, and Somalia) have experienced a fair degree of armed conflict or social unrest over the last decade (De Waal, 2015). It is plausible, then, that the omission of these countries might have had at least a minor effect on the regression result, especially since there has been substantial evidence of conflict acting as a barrier to educational access (Dryden-Peterson, 2009; Justino, 2014; A. Smith, 2014). Furthermore, the conflict-affected countries omitted from this study due to a lack of data are clustered in the Horn of Africa. This in itself may warrant additional attention from a human geography perspective as there could be omitted variable bias due to a range of geographic, demographic, or other social conditions. Again, the solution to such a limitation may be reducing the scope of a study to a sub-regional level. However, as Apland et al. (2014) reminded readers, these early studies on birth registration and educational access are better served as indicators of possible trends than as concrete results transferrable to other geographical contexts.

One final component of this study that could be interpreted as a limitation is the operationalization of access to education using a composite variable. Araya-Ajoy et al (2018) recognize a certain trade-off occurs when using composite variables. On one hand, composite measures can sometimes compound sampling or measurement errors (i.e. from aggregating separate components), but, on the other hand, single composite measures can be easier to estimate and/or interpret. According to Allen (2017), the key to clearly operationalizing a variable is to support the conceptualization of such concept with a strong theoretical foundation. The study presented in this article might face criticism,

then, because it is the first of its kind that operationalizes educational access using the specific combination of attendance, out-of-school, and completion indicators. Therefore, it has been untested in empirical research until this point, meaning that there could be some credible concerns of reliability.

Although these concerns with composite variables are understandable, it can be argued that they are unnecessary. First, empirically, Spaull & Taylor (2015) set a precedent for the use of composite educational measures in Africa and demonstrated how combining indicators actually helped overcome earlier limitations where studies focused primarily on educational quality *or* quantity. Second, conceptually, some have maintained that it is critical for educational access to be defined in terms of both inputs, like enrollment, or outputs, like completion outcomes (Lewin, 2009; Little & Lewin, 2011). In this model, a truly inclusive and accessible education system is one that is highly attended, results in few out-of-school children, and sees high quantities of completions. Thus, a comprehensive conception of access, or a composite measure, is not only easier to analyze in a statistical analysis, like Araya-Ajoy et al. (2018) mentioned, but it also is fundamentally a more inclusive model of access.

4.6. Conclusion

This study was the first of its kind to examine the relationship between birth registration and educational access at a regional level, using Sub-Saharan Africa as the study area. The purpose was to build on previously cited studies that found statistically significant relationships between these variables of interest across a few selected countries. Although this paper did not find significant results with regard to the effect of birth registration on educational access, the fact that it yielded contradictory findings to

those studies shows that research has barely scratched the surface on this topic. Instead, this paper yielded important results elsewhere, highlighting the impacts of GDP per capita and the urban/rural divide in terms of educational access. Both sets of findings result in opportunities for future researchers and practitioners alike in the field to help address the large inequities in educational access and the tens of millions of unregistered children across the region. This paper concludes by offering recommendations for each in the remainder of this section.

International and/or comparative education researchers could benefit from both a change in methodology and in geographic scope to guide further research into the relationship between birth registration and access to education. The reason for this claim is due to the contradictory findings found throughout the limited research on this subject to date (see section 4.2 above). Even within studies (see Apland et al.), there have been contradictions between quantitative and qualitative findings, with the latter yielding less evidence of an existent relationship than the statistical results demonstrated. Additionally, the presence of documented instances where birth registration acted as a mechanism for exclusion suggests that further research is warranted (Crush & Tawodzera, 2014; Kindzeka, 2017; Sumner, 2015). As such, one recommendation for Sub-Saharan Africa would be to widen the methodological framework to include qualitative or even spatial analysis techniques. The purpose here would be to uncover clusters within the region, whether within a single country or in a cluster of neighboring countries, where a possible relationship may exist. A starting place might be in/around the places of these documented instances, like Cameroon (Kindzeka, 2017) or Zimbabwe (Crush & Tawodzera, 2014). In either case, the methods should continue to view access as a

combination of school enrollment, quality, and completion at multiple educational levels to help understand not just *if*, but also *where* and *how*¸ barriers to education manifest themselves throughout the region.

Should this further research uncover clusters of disparities in educational access, it would help key intergovernmental development organizations, like UNICEF or UNESCO, plan assistance initiatives with better geographical precision. This is important as any educational reform efforts rely on effective strategies that recognize these context-specific and diverse needs that are more inclusive of all learners (Chimombo, 2005). As an example, pursuits of universal educational access provision in the global South, and in Africa, specifically, may be more aptly suited for non-formal basic education programmes (Little & Lewin, 2011). Alternatively, this further research may reinforce the personalized stories cited above, which raised issues with educational access in the transitions to secondary education levels. These are consistent with this study's descriptive trends in diminishing access by successive education levels and its inferential findings related to GDP per capita and rurality. In sum, any development programs targeting educational access must also recognize these other socioeconomic and demographic variables at hand. In these cases, governments and/or aid organizations should simultaneously consider interventions targeting poverty reduction or strategies addressing the reasons (i.e. child labor) why adolescents remain out of school (Justino, 2014). In other words, there is much work that can be done to address both the shortfalls in birth registration and in educational access across Sub-Saharan Africa in tandem with one another, whether or not a statistical relationship exists among the two.

CHAPTER 5: EXAMINING THE GEOGRAPHIES OF EDUCATIONAL ACCESS AND BIRTH REGISTRATION IN SUB-SAHARAN AFRICA (ARTICLE #2)

5.1. Introduction

Spatial distributions of social or ecological phenomena are important for understanding if patches and/or gradients in their observed values exist over a given geographical space (Legendre, 1993). In one poignant example, Tikly (2001) proposed a conceptual framework of exclusion in Sub-Saharan Africa, offering readers with foundational insight into the general commonality among postcolonial, low-income countries. This showed how education systems in these areas continued to solidify colonial-era social structures over time, thereby perpetuating limited opportunities for social mobility. This is not uncommon to the African continent, where an unfortunate legacy of lingering social problems has been, at least in part, due to geography, as artificially drawn borders cut through linguistic or ethnic populations (Ali et al., 2019; W. F. S. Miles, 2014; Nettle, 1996; Rudincova, 2015; Touval, 1966). It is not surprising, then, that evidence of spatial dependencies, or clusters, between various social variables can be found in Africa (O'loughlin & Anselin, 1991). A key question in social science or geographical research is whether social variables, like birth registration levels or educational indicators, are consistent throughout the region or if they are, in fact, legacies of a manufactured human geography.

This study will build upon the few scholarly publications highlighting access to and exclusion from key social services, like public education, in Sub-Saharan Africa. Tikly (2001) laid the theoretical foundation for this topic in a theoretical framework that reviewed social exclusion in the context of a rapidly globalizing world at the turn of the

millennium. Shifting to an empirical viewpoint, Majgaard and Mingat (2012) displayed the first robust, comparative analysis of disparities in the education systems in the region, finding accessibility disparities by socioeconomic status, gender, rurality. That study, exhaustive as it was, failed to include any mention of birth registration as one of the many factors that may shape a student's entry into various educational levels (e.g. primary, secondary, etc.). This is a notable omission, especially since Apland et al. (2014) and Corbacho et al. (2012) conducted in-depth statistical analyses demonstrating significant relationships between birth registration levels and access to education in the few countries they examined. Those studies left a gap in the literature with none having explored this relationship on a regional level. Finally, a recent study by Ahmed et al. (2019) is the only to use GIS tools to map the extent of spatial inequality in access to learning in an African context (Ethiopia). For educational research, that study provided a compelling case for the use of spatial methods in understanding whether educational access is equitably distributed. Again, with an emphasis solely on one country, an opportunity was missed for cross-country comparative analysis.

Seeking to fill these gaps in both content and methodology, the study will unfold in the following manner. First, a brief background section will cover an overview of the problem that is social exclusion in Sub-Saharan Africa while simultaneously setting the methodological foundation with the review of an influential spatial study driving this research. Second, the methods section will discuss the spatial research design by showing how it is the natural evolution beyond Griffin's (2021a) quantitative analysis on the same topic. Special attention will be given to describe the specific geographic research methods, choropleth mapping and spatial autocorrelation, since this study is intended for

educational policy audiences, with whom these concepts might be foreign. The results section will cover both the descriptive and inferential findings associated with such methods, allowing readers to clearly visualize trends in educational access and birth registration across Sub-Saharan Africa. The discussion section then reminds the audience of the unique contributions of this study to international education research. Finally, the article concludes with recommendations for further policy and research, notably recognizing the need for better integration of geographic and quantitative analyses in cross-national, comparative, research in the social sciences.

5.2. Background

This article was intended to address a problem of social exclusion in Sub-Saharan Africa, where inequalities in educational access continue to disenfranchise vulnerable youth. A range of systemic barriers to quality schooling, ranging from nationality status to gender, have persisted across this region, leaving individuals behind in their pursuits of better opportunity through education (Buckland, 2011; Dryden-Peterson, 2009; Lloyd et al., 2005). Not only does this region trail most, if not all, other world regions in terms of literacy and net attendance rates, but there were an estimated 260 million children out of school in Sub-Saharan Africa as recent as 2016 (Roser & Ortiz-Ospina, 2016; United Nations Educational, Scientific and Cultural Organization, 2018). The stakes to persisting inequalities are high, as barriers to the provision of education in concentrated geographic areas have been considered as factors contributing to and/or exacerbating civil instability across several countries in the Global South (Brock, 2011; Burde et al., 2017; Dryden-Peterson, 2009; King, 2011; Østby et al., 2019). Østby (2019) reviewed over 40 studies

examining the relationship between education and civil conflict, yet none of the studies examined individual barriers to educational access, such as birth registration. This is problematic, as Hunter (2019) showed how the lack of a birth certificate can prevent children from entering school systems, progressing between levels, or from exiting with a credential.

When one is excluded from their rightful claim to social services due to a missing birth certificate, they are essentially left behind at the starting line. Other metaphors for these circumstances see the individuals as, "missing," or "hidden in plain sight," both of which make the point that children without birth documentation are essentially invisible from state institutions (Ajmera & Fields, 2016; Sumner, 2015). These critical documents, which provide the basis for one's legal identity, therefore act as mechanisms of social exclusion, precluding children (whom are already vulnerable in Sub-Saharan Africa) from accessing key social services, like education or healthcare (Brewer et al., 2015; Chereni, 2017; Hunter, 2019). This is especially worrisome as many of the civil registration systems in Sub-Saharan Africa were emplaced by colonizers decades ago and still create conditions of exclusion and marginalization today (International Institute for Vital Registration and Statistics, 1988; R. J. Johnston et al., 2000). Consequently, Sub-Saharan Africa is the world region with the highest number of unregistered children under the age of 5, at an estimated 166 million children (United Nations Children's Fund, 2019e). This places the region at a tremendous risk of failing to meet the United Nations' Sustainable Development Goals 4 and 16, for ensuring inclusive/equitable quality education and provide legal identities, including birth registrations, for all by 2030, respectively (United Nations Children's Fund, 2017).

Given that Sub-Saharan Africa possesses several of the world's lowest country-level values in both birth registration and educational indicators, there was reason to believe that geographical pockets within the region where these variables clustered together (Roser & Ortiz-Ospina, 2016; United Nations Children's Fund, 2019e). On one hand, there were individual instances where birth registration acted as a barrier to furthering one's education beyond the primary level across opposing corners of the region, from Cameroon to Zimbabwe (Crush & Tawodzera, 2014; Kindzeka, 2017). On the other hand, a recent quantitative study found that no such relationship of statistical significance existed between birth registration and educational access across Sub-Saharan Africa over the last decade (Griffin, 2021a). That study, however, offered only a macro, regional-level, statistical analysis and, furthermore, its relatively small sample size (n=40) could have potentially masked trends found within the variables themselves. A natural next step in geographic and mixed methods research is to employ alternative tools that can further explore the spatial distribution of such variables and/or explain these initial quantitative results, respectively (Creswell & Plano Clark, 2018; McGrew Jr. et al., 2014). After all, the spatial dependence of social phenomena, or clustering of them, has been shown to be strongly present in Africa (O'loughlin & Anselin, 1991). This article therefore sought to answer the question: *Is there is a non-random pattern in the spatial distribution of the educational access and demographic variables presented in Griffin's (2021a) multiple regression analyses?*

A valuable predecessor to this article was the work of A. Y. Ahmed et al. (2019), as they explored the spatial inequality of access to meaningful learning in Ethiopia. Their work was influential to this study for two key reasons. First, they used geographic

information systems (GIS) to show how access to substantive learning was not equitably distributed within the Amhara region of Ethiopia. As mentioned in section 5.1, they recognized that social phenomena could have vast geographical variations within a certain study area. The study presented below will expand the scope to include an investigation of educational access beyond Ethiopia (e.g., to the countries of Sub-Saharan Africa), but it is critical to credit A.Y. Ahmed et al. with proving that examining this topic, spatially, was a rigorous and effective model. Second, like Spaull and Taylor (2015), they operationalized access to education using a composite measure from multiple indicators that covered educational access and educational quality. In fact, they argued, "…the rationale to conceptualise and operationalise a combined measure of educational access and quality is logical in the context of expanding education systems" (p. 577). Although the study below operationalized access slightly differently – as a composite of net attendance, out-of-school rates, and completion rates – the key here is that both studies were conceptualized using an expansive vision of access that recognized and incorporated the complexity of such a topic. Doing so allows readers to gain a deeper understanding of the way that access to education is distributed, spatially.

5.3. Methods

The research design for this study could be described as a cross-national area pattern analysis, a phrase that joins two key elements. The first element, the cross-national piece, is an analysis technique that enables researchers to examine an entire study region to understand if observed trends are due to country-specific peculiarities or social-structural regularities across the area (Kohn, 1987). This element essentially

defines the geographic scope of the study and units of analysis, which in this case is countries within Sub-Saharan Africa. As the term implies, an area pattern analysis is appropriate for studying practical problems in geography where patterns may emerge from the attribute values assigned to neighboring geographic areas (e.g. towns, districts, countries, etc.) (McGrew Jr. et al., 2014). A benefit of area pattern analyses is that it can be descriptive and/or inferential, allowing researchers to uncover spatial trends in multiple ways (Savigny & Wijeyaratne, 1995). For example, a researcher can employ a choropleth map to describe the dataset, much like a histogram describes the data in a quantitative analyses, according to Savigny and Wijeyaratne. Likewise, similar to a Pearson's correlation test in a quantitative study, a Moran's *I* test can be used to infer relationships within the spatial dataset and even offers similar correlation values, ranging from -1 to 1 as an indicator of relational strength (Bolstad, 2016). The following section outlines the ways in which both techniques were used to examine the geographies of educational access and demographic indicators across the countries of Sub-Saharan Africa.

5.3.1. Data

This study drew from secondary datasets assembled by two mission-similar entities, the United Nations Children's Fund (UNICEF) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO). These are two of the world's leading agencies for officially recorded data about birth registration, child protection, and internationally-comparable education data (United Nations Children's Fund, 2020a; United Nations Educational, Scientific and Cultural Organization, 2020c). Both datasets actually follow rigorous quality control processes to aggregate data sourced

primarily through UNICEF Multiple Indicator Cluster Surveys (MICS) or the U.S. Agency for International Development's (USAID) Demographic and Health Surveys (DHS) Program, two nationally-representative household survey programs spanning most low- and lower-middle income countries (Ayede et al., 2018; Hancioglu & Arnold, 2013). Since this article was intended to be a spatial follow-up examination of Griffin's (2021a) quantitative analysis, it is recommended to see that preceding article for a detailed description of the secondary dataset itself.

In summary, however, Griffin's study focused on six variables in a multiple regression model across an *n* of 40 countries in Sub-Saharan Africa between 2010 and 2020. The key outcome variables were access to education at the primary-school, lower-secondary, and upper-secondary education levels. In each of these three variables, access to education was defined as a composite of net attendance, out-of-school, and completion rates, allowing for a thorough and comprehensive perspective of access. The key explanatory variable was birth registration, measured by the percentage of children under age 5 (0 to 59 months) with a birth certificate or whose birth was reported as registered with civil authorities at the time of survey. Further explanatory variables representing socioeconomic status and demographic characteristics were gross domestic product (GDP) per capita and the percentage of each country's population that was defined as rural. The purpose of the latter three explanatory variables in Griffin's study was to understand if they could explain the variance in access to education across the successive education levels in Sub-Saharan Africa.

Based on the results of Griffin's (2021) three multiple regression analyses, all six variables were included in this subsequent spatial analysis presented here. There, birth

registration, GDP per capita, and rurality, when considered together, significantly explained roughly one-third of the variance in access to education at each schooling level. As McGrew Jr et al. (2014) argued, a natural next step in the research process, from a geographer's perspective, is to critically examine the spatial distribution of the variables themselves to highlight any unique clustering or non-random dispersion. Similarly, although birth registration did not have a statistically significant unique contribution to the explanation of educational access, the models themselves warranted a deeper empirical investigation. According to Arkes (2019) and McShane et al. (2019), however, there are other effective ways of analyzing a research problem beyond simply hunting for statistical significance between variables. This is one key benefit of sequential mixed methods, where one strand of data is analyzed and interpreted to then determine how best to further make sense of those findings with subsequent research (Creswell & Plano Clark, 2018). Furthermore, the decision to include more variables was one of pragmatism, hopeful that it may provide future social science researchers with a general baseline understanding of how these social conditions manifest themselves across the region.

5.3.2. Analysis

Choropleth mapping. One of the key drivers of this study was a need to visualize the variables of interest (birth registration, access to education, etc.) across Sub-Saharan Africa to facilitate an easier pursuit of pattern recognition or clustering. The ArcGIS platform was best suited for this project since it is a geographic information system (GIS) that allows users to identify characteristics of geographic places and perform sophisticated spatial analyses (Law & Collins, 2015). The first step in this

study's analysis phase was to assess the six key variables, descriptively, by presenting them on maps in which their values were divided into classes and depicted across geographical units, or countries, across the study region (Savigny & Wijeyaratne, 1995). By classifying data this way, McGrew Jr. et al. (2014) argue, "…communication is enhanced, detailed spatial information is better understood, and complex spatial patterns are represented more clearly" (p. 27). One of the most common ways of doing this is through the use of a choropleth map, where land areas, like countries, are shaded according to continuous or categorical data values (Ballas et al., 2017).

Ballas et al. (2017) point out that creating thematic maps and geovisualizations, like choropleth maps, can be accomplished in many ways, requiring the researcher or cartographer to make key choices in the classification scheme and visualization technique to best fit their objectives. The former is arguably the most important decision, as the number of classes, or groups of similar values, will dictate how much variability there is on the map. This study was specifically interested in the potential clustering of educational access, lending itself to an equal intervals scheme. Here, the range of values within each variable was subdivided into five equal classes to determine which countries, if any, were clustered around a specific class or if they were randomly dispersed across the five classes. Much like the common statistical significance level of 0.05 in quantitative analyses, the five-class model represents a common number of choropleth classes for cartographers interested in mapping geographic patterns (McGrew Jr. et al., 2014). An example of this model would be if the values for a particular variable in the dataset ranged from 20 to 80 (i.e. a range of 60), then each of the five classes would represent 12 percentage points. Each class is then assigned a color or shade, based on the

visualization technique, so as to group the geographic units (in this case, countries in Sub-Saharan Africa) according to their attribute values.

The second decision in creating the choropleth maps for this study was more of a stylistic one, which is a reminder of the art that is cartography. In Figures 5.1 through 5.6, readers can see how the educational access variables (see Figures 5.1-5.3) follow a blue gradient scheme, while the demographic variables (see Figures 5.4-5.6) follow a green gradient. This was a stylistic decision intended to especially serve readers coming from the preceding article (see Griffin, 2021a), where the now-blue maps represent outcome variables, and the now-green maps represent the explanatory ones. Furthermore, the contrast differences, where a darker shade in some maps corresponds to a higher value for a variable or a lower value in other maps, were designed so that important elements would clearly stand out to viewers, a critical component to thematic mapping (Ballas et al., 2017). Again, referencing Griffin's (2021) prior study, this was important to depict, visually, since *higher* levels of rurality but *lower* values of birth registration were assumed to relate with educational access. It is therefore critical to examine the scale of each map prior to one's interpretation. With these maps produced, they could serve both as a baseline for inferential spatial analysis and as a standalone product for depicting the spatial distribution of key educational and/or demographic variables of interest.

Global Moran's I. If the descriptive mapping is most useful for showing where certain variables are distributed throughout Sub-Saharan, then the inferential mapping is best served for determining, statistically, if the magnitudes of their values vary from place to place (McGrew Jr. et al., 2014). This is known as spatial autocorrelation, which is a method for understanding the dispersion or clustering *within* single variables, which

should not be confused with a traditional Pearson's correlation that tests bivariate relationships *between* variables (Getis, 2007). Spatial autocorrelation became popularized in geographic research following Moran's (1948) questioning as to whether the values of a variable in one area made it more or less likely to see similar values of that same variable in a neighboring area. Tikly's (2001) work, summarized earlier in this article, is a reminder that some social processes, particularly in education, may generate similar conditions across neighboring countries sharing colonial-era experiences. Therefore, a spatial autocorrelation test was an appropriate analysis for understanding if there was any statistical evidence backing this idea across educational access, birth registration, and demographic indicators. Furthermore, the dataset for this study lent itself to a Global Moran's *I* test, one of the most common tests of spatial autocorrelation, because it had over 30 geographic features (i.e. countries), it was on a ratio scale, and the data adhered to assumptions of normality (McGrew Jr. et al., 2014).

Unlike the choropleth mapping, which required a few key stylistic decisions, the Global Moran's *I* test was performed using a single process in the ArcGIS software program. This platform allows researchers to use a variety of geoprocessing tools, manipulating input data into some form of new, often more meaningful, spatial output (Law & Collins, 2015). In this case, the input data was the same dataset described above, with six key variables of interest covering the countries of Sub-Saharan Africa. The analysis itself is a mathematical equation that produces the autocorrelation value, *I*, as a quotient of the areas in the dataset (i.e. countries) and their attribute values divided by the number of spatial joins (i.e. # of contiguous borders between countries) and their variance (Bolstad, 2016). This value depicts the degree to which variables' attribute values are

dispersed (*I* values approaching -1), random (*I* values close to 0) or clustered (*I* values approaching 1) The output, gained by selecting the 'Global Moran's' option in the ArcGIS Analysis toolbox, generates an HTML report file that contains the Moran's *I* value, its *Z*-score, and the *p*-value for the analysis. Table 5.1 summarizes these values for all six variables of interest. The results, described in further detail below, were then assessed using the same significance threshold as Griffin's (2021a) study, $p<0.05$. In short, however, any *p*-values less than 0.05 resulted in a rejected null hypothesis, meaning that the variable's attribute values were not randomly distributed throughout the study area.

Table 5.1: Moran's *I* and *p*-values for spatial autocorrelation tests of six variables

Variable	Moran's *I* value	*p*-value
Access to Education – Primary	0.38	<0.01
Access to Education – Lower Secondary	0.26	<0.05
Access to Education – Upper Secondary	0.11	0.23
Birth Registration	0.13	0.16
GDP per Capita	0.12	0.13
Rural Population	0.31	<0.01

5.4. Results

5.4.1. Access to Education

The three green-shaded choropleth maps representing access to education across Sub-Saharan Africa offer several key findings with regard both to the geographical levels (i.e. regional- or country-levels) and the educational levels (i.e. primary, lower-, and upper-secondary). At the regional level, a comparative look across the maps shows that access to education is much higher, generally, across Sub-Saharan Africa at the primary-school level than at the secondary levels. For convenience, Table 5.2 displays the number of countries in each equal class of access values across the three education levels. In this

table, all but six countries at the primary-school level are in the middle to high classes of educational access, representing 68.84% or greater on the composite access scores. Conversely, all but six to seven countries at the secondary-school level fall within the middle to low educational access classes, representing lower than 61.09% or 43.09% composite access scores at the lower-secondary or upper-secondary levels, respectively. At the country level, South Africa stood out as the only country in the highest access class at each level of education. South Sudan, on the other hand, ranked as the only country lying in the lowest class of access at each educational level. Although these tabular, cross-map findings are valuable, examining the individual choropleth maps alongside their respective Moran's *I* results uncover even more detailed trends in educational access.

Table 4.2: Number of countries by choropleth class for each educational level

Variable	Primary	Lower Secondary	Upper Secondary
Lowest access class	1	6	5
Low-middle	5	13	14
Middle access class	8	11	12
Middle-high	15	6	4
Highest access class	8	1	2

In Sub-Saharan Africa, access to education at the primary school level demonstrated a statistically significant clustered pattern ($p<0.001$). This meant that the null hypothesis could be rejected, that the spatial distribution of values were clustered, with a moderate *I* value of 0.38, and that the result was likely not due to chance. Turning to the choropleth map, three clusters became visibly apparent. Five countries, colored yellow, clustered at the southern edge of Africa, representing some of the highest values of educational access at the primary level. Moving north, the next ostensible cluster was a

group of eight countries with high primary school access in Central Africa, marked by the lightest shade of green. Although South Sudan stuck out as the lone country in the lowest class of primary school access, the next lowest class, displayed in a dark forest green shade, was clustered in a horizontal pattern across four countries at the northernmost point in Sub-Saharan Africa. This is a region known as the Sahel, which is a ~4,000km arc-like belt of land below the Sahara Desert that stretches horizontally across the continent (Suleiman, 2017).

The lower secondary school level also demonstrated a significant spatial autocorrelation ($p<0.05$). Additionally, it too had a moderate Moran's *I* value, at 0.25. Interestingly, despite the smaller *I* value than primary education level, there were actually four noticeable clusters of values when looking at its choropleth map. The most noticeable cluster is the large swath of the Sahel that is depicted in the darkest shade of green, where five contiguous countries possess the lowest access levels in the entire region. There are then two clusters of the second lowest access class (dark forest green) on opposite sides of the region, with seven countries on the western coast and four countries on the eastern coast. Interestingly, all four countries within the latter cluster (Uganda, Rwanda, Burundi, and Tanzania) were some of the few countries that decreased in access by two choropleth classes from the primary school level. Other countries witnessing this phenomenon included Zimbabwe, Mozambique, and Cote d'Ivoire. In other words, these seven countries were in the second *highest* class of educational access at the primary school level but then fell to the second *lowest* class at the lower-secondary level. This is noteworthy in the context of educational transitions and access, which will be covered in the discussion section of this article. Finally, the last cluster consisted of

five countries in the central part of Sub-Saharan Africa with moderate (middle class) levels of educational access.

Unlike the first two levels, access to education at the upper-secondary school level did not display any significant clustering across Sub-Saharan Africa. The choropleth map supports this assertion with a much more dispersed set of values across the region. There is still a single cluster of six countries in western Africa that appear to be clustered; however, the region as a whole did not demonstrate significant autocorrelation. Notable other findings from this education level resulted from comparing this map to the lower-secondary map. First, two pairs of neighboring countries (Burkina Faso & Niger and Central African Republic & South Sudan) remained as countries in the lowest class of access. Second, Gabon and Tanzania were the only two countries in Sub-Saharan Africa that witnessed a lower degree of access at each successive level of access. In fact, Tanzania joined the four aforementioned countries as the five countries in Sub-Saharan Africa with the lowest access values at the upper-secondary level. After reviewing the next set of variables, these findings will be discussed in the context of existing educational research to further investigate access throughout the entire region.

5.4.2. Birth Registration

Birth registration, which was the key explanatory variable influencing this study, appeared to show no significant evidence of spatial autocorrelation across Sub-Saharan Africa. Its choropleth map illuminated this dispersion as the values were fairly evenly distributed throughout the region. In fact, the highest two classes of birth registration each contained 10 countries and the third and fourth classes contained six and seven countries, respectively. Only four countries stood out with very low levels of birth

registration. Similar to upper-secondary access, however, the birth registration map did display a few instances where there were neighboring countries with similar values despite the lack of statistical clustering. First, there is a wide vertical cluster of countries in the second lowest class of birth registration (dark blue) in central Africa. Second, there is a thin horizontal, contiguous, layer of countries in the northwestern part of the region that all fall in the second highest class of birth registration. Finally, the range of attribute values varies dramatically (see the five classes in the map's legend), showing just how different birth registration processes may be functioning across the region.

5.4.3. Socioeconomic and Demographic Variables

The remaining findings allow readers to understand how key explanatory variables in Griffin's (2021a) study, used to control for socioeconomic status (GDP per capita) and geographic setting of each country (rurality), were spatially distributed. GDP per capita was statistically correlated with access to education in that study, making it intriguing to examine spatially, especially now that access to education has been presented above. Based on the Moran's analysis, it was not found to demonstrate spatial autocorrelation. Its choropleth map looks almost like one massive cluster, however, as 30 of the 37 countries fell in the lowest class of GDP per capita. On one hand, this highlights the tremendous poverty levels across the region as well as the vast inequalities in income distribution. The map effectively brings attention to the low absolute values of GDP per capita. On the other hand, this also shows one downside of choropleth mapping as it was not possible to tell if there were pockets of clusters to be found throughout the region.

The rural populations of countries in Sub-Saharan Africa were significantly clustered at a $p<0.05$ level. Additionally, the Moran's I value of 0.31 was moderate. The

most noteworthy finding, made possible by the choropleth map, was that the most sparsely populated countries were also mostly landlocked. Conversely, the least rural countries were those found on the southern and western coasts of the continent. This should not be surprising given that many large African cities (e.g. Libreville, Lagos, Dakar, etc.) are all found along the coastline. It is also noteworthy to compare this map side-by-side to that of the educational access maps, particularly at the two secondary levels. Doing so will illuminate some key trends in that the countries with the least educational access are also some of the most rural. This is another benefit as the statistical correlations were not significant at the 0.05 threshold between these variables in Griffin's earlier study, yet the maps still demonstrate at least some degree of existent relationship. With this in mind, the following discussion section will cover some of the topical and methodological implications of such findings.

5.5. Discussion

The purpose of this study was to examine the spatial distribution of educational access across Sub-Saharan Africa as well as some of its possible barriers, like birth registration, country-level wealth, and rurality. One of the benefits of mixed methods research, according to Creswell and Plano Clark (2018) is that it enables researchers to harness the strengths of multiple modalities in an attempt to further explain results and/or to generate new knowledge on a topic. In this case, it was unsettling that Griffin's (2021a) quantitative analysis yielded no statistical relationship between birth registration and educational access. This was due to the fact that empirical evidence has shown how birth registration can act as a barrier to accessing social services, like education (Crush & Tawodzera, 2014; Kindzeka, 2017) or healthcare (Bennouna et al., 2016; Brito et al.,

2017). In that sense, a spatial analysis was presumed to be helpful as a method for understanding if any geographical clusters of each variable were similarly distributed, despite an overall lack of statistical significance between them. The section presented below discusses how this article was effective dispelling the notion of birth registration as a region-wide barrier to education while also uncovering some unexpected and unique results with regard to the spatial distribution of educational access in Sub-Saharan Africa.

5.5.1. Birth registration: Case closed? Not quite.

The most interesting finding stemming from this spatial inquiry into birth registration, specifically, was that it did not display any instances of country-level clustering that were similar in spatial distribution with any of the other variables (i.e. the other five choropleth maps). On one hand, this should not be surprising since Griffin's (2021a) quantitative analysis revealed that birth registration was not significantly correlated with any of the study's other variables. On the other hand, however, it would have been noteworthy had any similar patterns between variables emerged, similar to the way in which rural population and access to secondary education yielded similar examples of clustering among landlocked countries' values despite not being significantly related to one another in Griffin's prior study. Additionally, it would have made sense if any mirroring instances of clustering had occurred given Tikly's (2001) theoretical framework of shared pre- or post-colonial experience in Sub-Saharan Africa. Furthermore, it was unexpected that none of the explanatory variables displayed similar clustering because Bhatia et al.'s (2017) found that wealth and urban/rural inequalities in birth certificate coverage tend to persist in most low- and middle-income countries, many of which are found in Sub-Saharan Africa.

This study's choropleth map for birth registration may not be particularly noteworthy on a standalone basis, but it is intriguing when situated amongst one of UNICEF's most recent (2019) status reports on the progress of birth registration indicators worldwide. In this report, a choropleth map for birth registration in Africa demonstrated a similarly dispersed pattern of values across the continent (p. 17). Interestingly, the report also presented an identical map but substituted birth *certificate* rates instead of birth registration (p. 19). The map then becomes substantially clustered, with the lowest rates of birth certificates in central and eastern Africa. In fact, the report suggested that in many of these countries, more than a quarter of registered children lack a birth certificate due to a variety of reasons, like expensive fees or broken acquisition processes for obtaining these crucial documents. In fact, the report showed how, on average, 40% of children in Eastern and Southern Africa are registered but only 22% have birth certificates.

These UNICEF statistics showing the stark contrasts in birth registration and birth certificate rates suggest that perhaps it is too early to dismiss a potential relationship between birth-related civil registration processes and access to state services. Even an international law review suggests that the human right to birth registration may not actually extend to include the issuance of a birth certificate, despite an implicit linkage between the two (Gerber et al., 2011). The presence of ostensible clusters of low-certificate countries and the mired standards of birth-related civil registration processes is therefore certainly worth untangling. Assuming one can acquire cross-national birth *certificate* data, a logical next step would be to replicate this study to again scan for spatial clustering alongside possibly related variables, like access to education. This

would be especially meaningful because, as (Hunter, 2019) asserts, there is divided academic literature as to whether identity documents (e.g. birth certificates) preclude those in the Global South from accessing state services or if their deprivation of services is a function of broader socioeconomic inequalities.

5.5.2. Toward a geography of educational access in Sub-Saharan Africa

This article's findings pertaining to educational access in Sub-Saharan Africa underscore the value of spatial analyses as an analytical technique. Rather than discussing all of the findings presented above, this section will cover those examples that were made particularly visible from this methodology. In the first example, the descriptive mapping (e.g. choropleth mapping) of educational access complemented the descriptive statistics found in Griffin's (2021a) quantitative analysis. In Griffin's preceding study, the mean value of educational access decreased from the primary school to the secondary, highlighting the declining access with successive education levels. The three choropleth maps here effectively built on this by showing *where* these decreases occurred, which became quite apparent by examining the distribution of values across each of the maps' five classes. This matters because two key World Bank reports in 2008 were dedicated to examining the critical factors affecting educational transitions in the region (The World Bank, 2008), calling for better strategies to address an often-overlooked need for secondary educational development (Verspoor, 2008). The findings presented in this article show that little has changed over a decade later, suggesting that perhaps more serious attention be given to Fredriksen and Fossberg's (2014) recommendations for substantial investments, augmented with external aid support, to secondary education systems in the region.

In the next example, the spatial autocorrelation displayed at the primary level and its associated choropleth map shed light on low educational access found across the Sahel region. Despite global and even regional progress toward Sustainable Development Goal 4.1 (i.e. universal primary/secondary education by 2030), clearly the Sahel sub-region, which covers nine countries in a horizontal band from Mauritania in the West extending to Sudan in the east, is being left behind (Sissoko et al., 2011; United Nations Children's Fund, 2018b). The spatial analysis presented in this article was effective in showing the significant cluster of Sahelian countries with low primary school access, but this is just a first step in actually addressing the complex issue at hand. For example, several recent UN press releases have shown how armed conflict has forced dramatic increases in school closures throughout the Sahel (A. Ahmed, 2020; United Nations Children's Fund, 2019b; United Nations Office for the Coordination of Humanitarian Affairs, 2018). Conflict is likely just one of many confounding issues leading to low access. Griffin's (2021a) quantitative findings and Eizenga's (2019) policy paper are reminders that poverty and rurality, both of which are clearly present in the Sahel, can play a critical role in determining one's likelihood of getting a formal education. After all, Shekar et al. (2016) assert, "Problems with education access, especially in rural areas, begin in primary school, so the basis for increasing people's educational opportunities is compromised when they are very young" (p. 25).

The final example highlighting the benefits of this spatial methodology is the cluster of four East African countries (Tanzania, Uganda, Rwanda, and Burundi) that visibly fell by two classes of educational access from the primary- to the lower-secondary-school levels. This finding is noteworthy as it certainly supports a body of

literature covering the inequities in secondary school access in these East African countries. In Uganda, Tanzania, and Kenya, for example, low school enrollment was shown to stem from the opportunity costs of attending school or perceived low returns in the labor market while the increasing abolishment of school fees has presented greater opportunities for children to continue learning (M. Oketch & Rolleston, 2007). Research from Williams et al. (2015) supports this idea in neighboring Rwanda, where even in a fee-free setting, "hidden costs" of schooling (e.g. exam fees, parent-teacher association dues, etc.) made secondary school inaccessible to the poorest children. In Uganda, specifically, implementing universal secondary education policies did not improve access for boys, perhaps due to some of these similar losses in opportunity costs of school, yet there were optimistic findings about the improvement in girls' attendance (Asankha & Takashi, 2011; Zuze & Leibbrandt, 2011). Likewise, in Tanzania, poor school infrastructure and inadequate transportation to school could be added to this list of barriers to secondary school access (Human Rights Watch, 2017). The point here is that using spatial analyses to identify clusters of areas facing similar issues, combined with a subsequent look into the possible causes of these issues among the clusters' component parts, can be an effective model for building comprehensive and targeted sub-regional action plans.

5.5.3. Limitations

As shown above, this article revealed a number of credible and significant findings, yet it managed to do so while working with a limited secondary data across key variables. Birth registration data, for example, remains limited by many developing countries' weak civil registration systems (Bhatia et al., 2017; Mikkelsen et al., 2015;

Phillips et al., 2018). It is important to keep in mind that the source data behind the UNICEF and UNESCO databases – the MICS and DHS survey programs – are used to produce nationally-representative samples for each country, which is then made available on the respective agency's data warehouse. Consequently, there is a need for more accessible disaggregated data across the countries involved in the MICS and DHS survey programs, which include most of the countries in Sub-Saharan Africa (Gray et al., 2013). This data could then give researchers a look into the subnational spatial distributions of social phenomena, as shown through key examples like access to meaningful learning in Ethiopia (A. Y. Ahmed et al., 2019) or child health inequalities in Sub-Saharan Africa, broadly (Yourkavitch et al., 2018). Since this article was the first of its kind to examine these key variables of interest across Sub-Saharan Africa broadly, there is certainly opportunity for subsequent research at the subnational level.

5.6. Conclusion

This study offered an innovative spatial follow-up to Griffin's (2021a) quantitative analysis examining the relationship between birth registration and educational access in Sub-Saharan Africa. Employing the Global Moran's *I* test of spatial autocorrelation allowed for a deeper assessment of six individual variables to understand if their values were distributed across the region in any sort of significant pattern(s). Although birth *registration* did not demonstrate significant clustering throughout the region, it opened the door for future research into the distribution of birth *certificate* rates, especially when compared to recent UNICEF reports that suggest sub-regional inequities exist, especially in Southern and Eastern Africa (United Nations Children's Fund, 2019e). This could prove useful when situated within Tikly's (2001) framework of social

exclusion in Sub-Saharan Africa or compared to Hunter's (2019) investigation of identity documentation in the Global South. This study's implications are actually far greater, however, in terms of advancing a new geography of educational access that would complement key prior research efforts, notably Lewin (2009) and Majgaard & Mingat (2012). Comparing choropleth maps of educational access at three levels (primary, lower-secondary, and upper-secondary) showed where transitions between levels resulted in noticeably lower values of accessibility. Access to education at the primary and lower-secondary levels, specifically, then revealed significant spatial autocorrelation, or clustering, of low values, especially in the Sahel and East Africa regions, respectively. These findings, enabled through spatial analyses, therefore simultaneously provided significant contributions to academic research and development policies.

After producing new knowledge at the intersection of international education and human geography, this article also opens a door to further inquiry and/or policy development. It reaped the benefits of mixed methods research, as it stems from a place of pragmatism, emphasizing tangible recommendations for either researcher or practitioner audiences (Creswell & Plano Clark, 2018). The spatial component, specifically, builds on Griffin's (2021a) earlier study by revealing certain geographical areas that could be further analyzed to understand the underlying socio-spatial context(s). In fact, essentially any clusters of countries found on the choropleth maps, particularly in those with significant evidence of spatial autocorrelation, can be used to drive future research agendas. In the East African secondary access cluster described above, for example, there is an opportunity to investigate the social processes or structures creating such similar values in neighboring countries. Thus, if this study was critical for

identifying *where* certain variables clustered together in Sub-Saharan Africa, then additional research could prove invaluable for understanding *why* these social processes manifest themselves, spatially, across the region.

Finally, the choropleth maps themselves could allow policymakers or development agencies to easily identify outlier countries (like South Sudan, for example) where immediate aid might be best appropriated. Although this article only summarized/discussed the key findings, it is the aim of the researcher that fellow practitioners use the spatial information to continue the quest to reduce socio-structural barriers to identity processes (e.g. birth registration) or social services (e.g. public education) in some of these already vulnerable regions. This action-oriented purpose is similar to the aforementioned research by Ahmed et al. (2019), who used spatial techniques to direct policy attention toward educational inequalities in Ethiopia. Following their model (while expanding the scope to a larger geographical area), the article presented here may allow development agencies to prioritize certain countries or regions when creating interventions that target the Sustainable Development Goals 4 and/or 16, which call for universal access to quality education or legal identities, respectively, for all by 2030 (United Nations Children's Fund, 2017). Thus, it is not too lofty a goal to presume that the spatial analyses presented here may literally become maps to better futures for all.

5.7. Figures

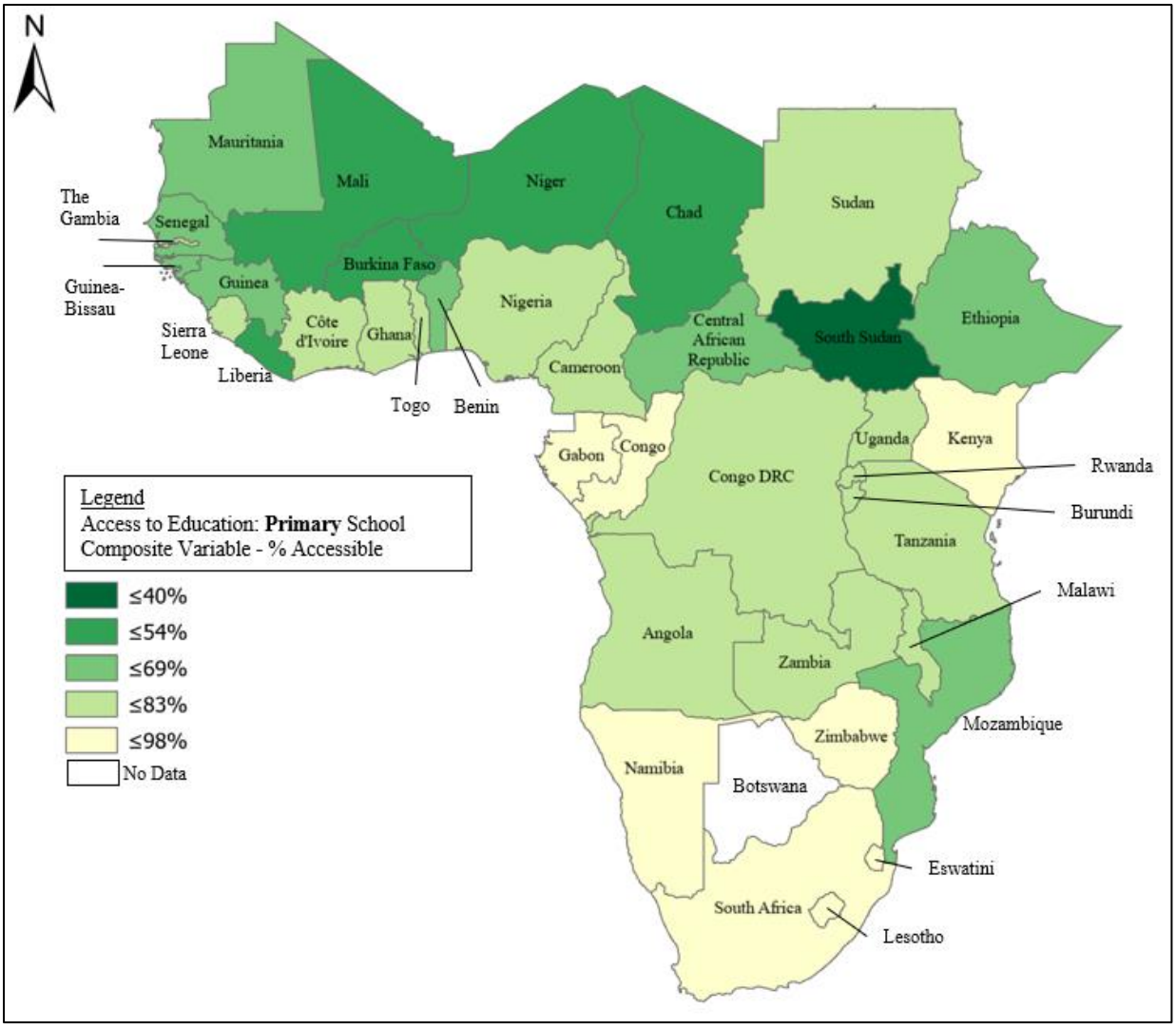

Figure 5.1 Choropleth map of access to education at the primary school level in Sub-Saharan Africa

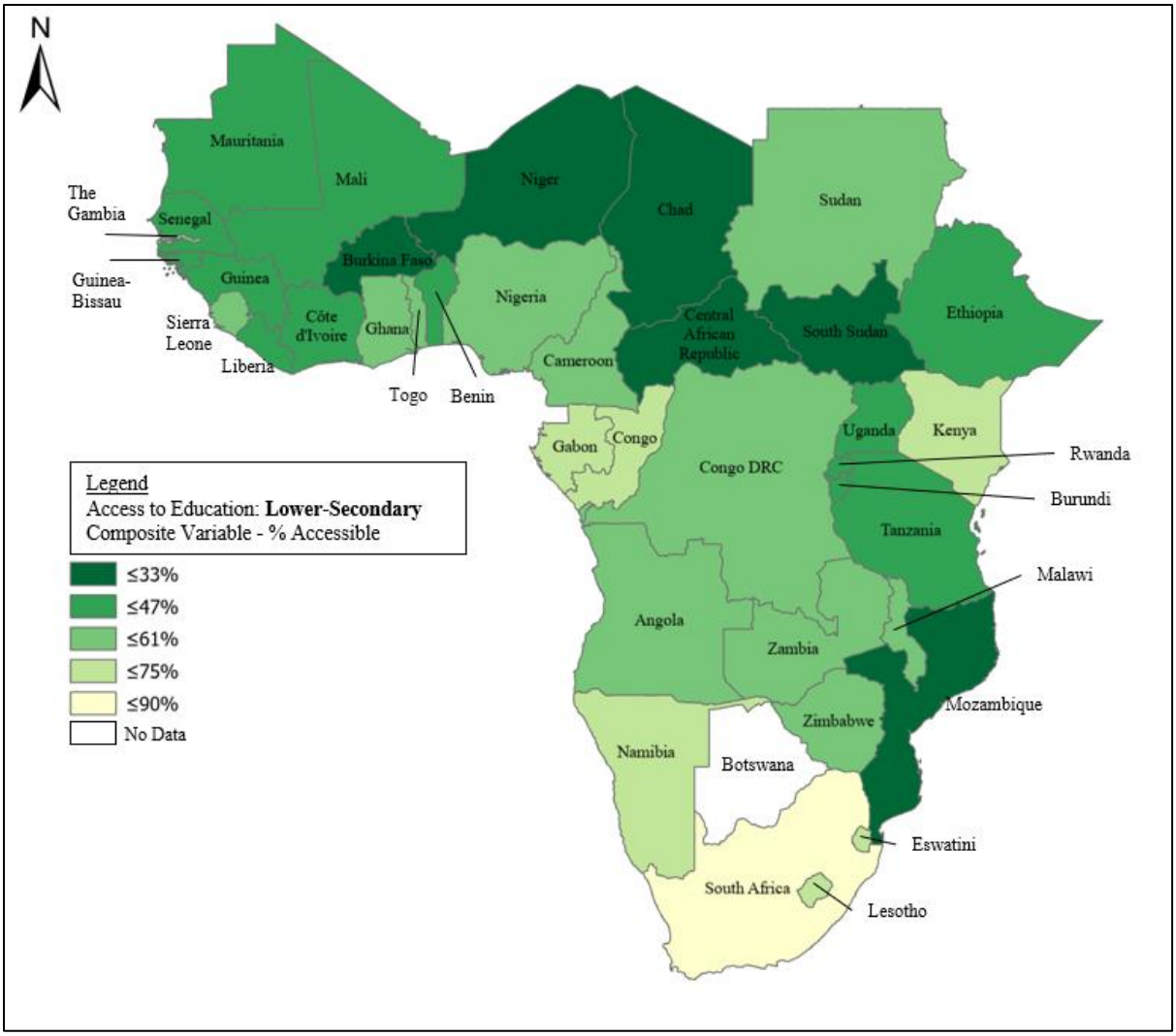

Figure 5.2 Choropleth map of access to the lower-secondary education level in Sub-Saharan Africa

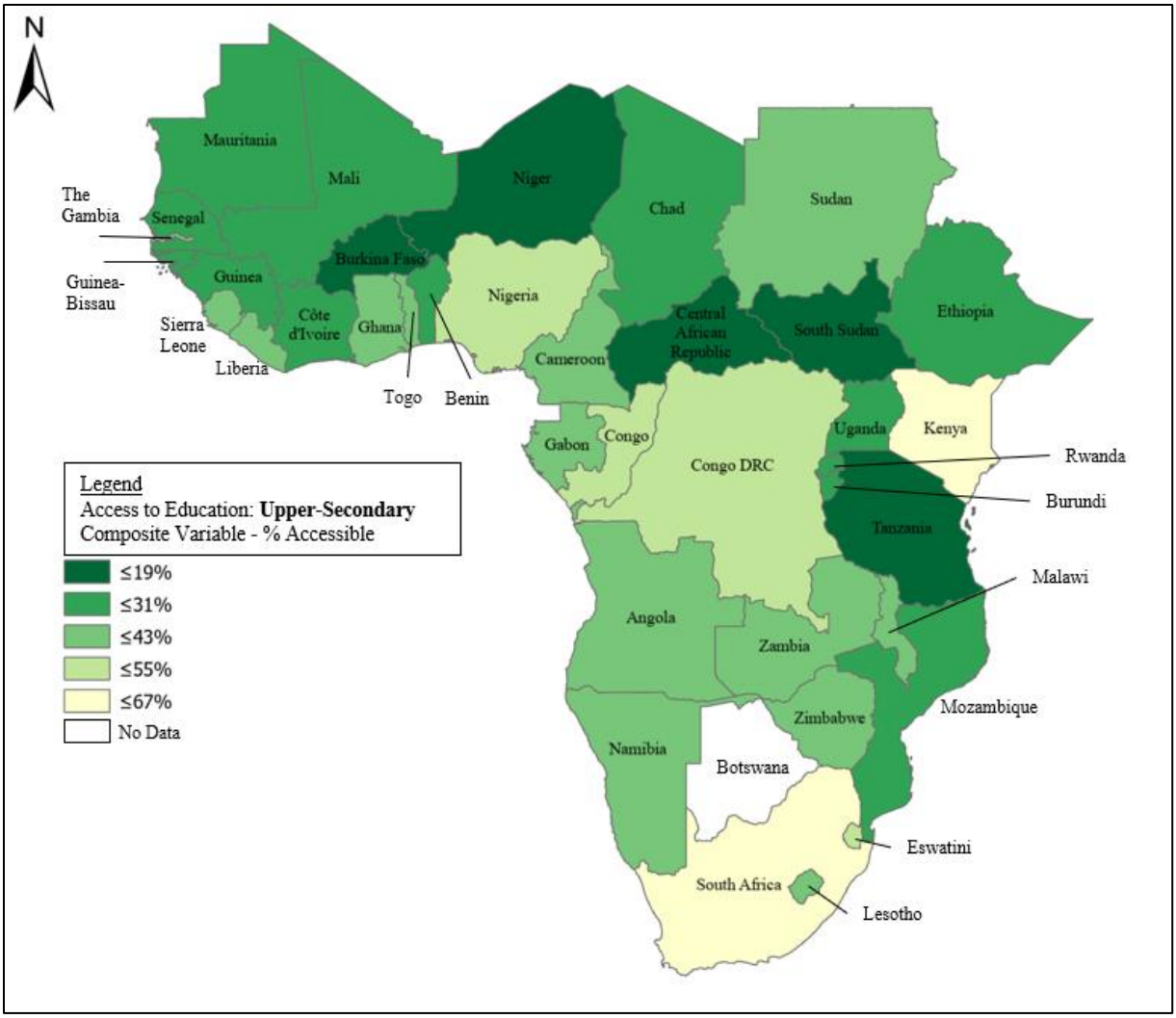

Figure 5.3 Choropleth map of access to the upper-secondary education level in Sub-Saharan Africa

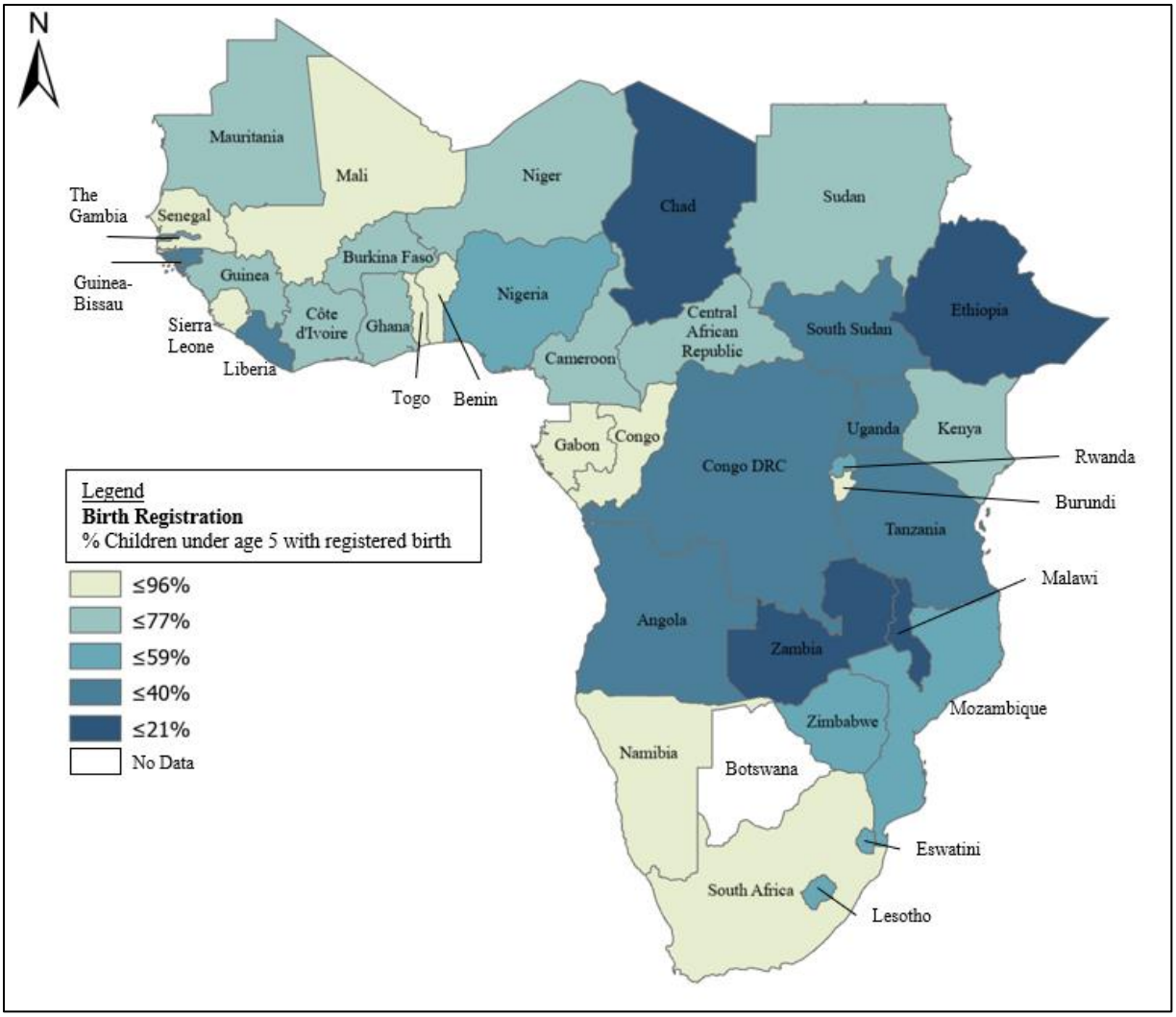

Figure 5.4 Choropleth map of birth registration rates across Sub-Saharan Africa

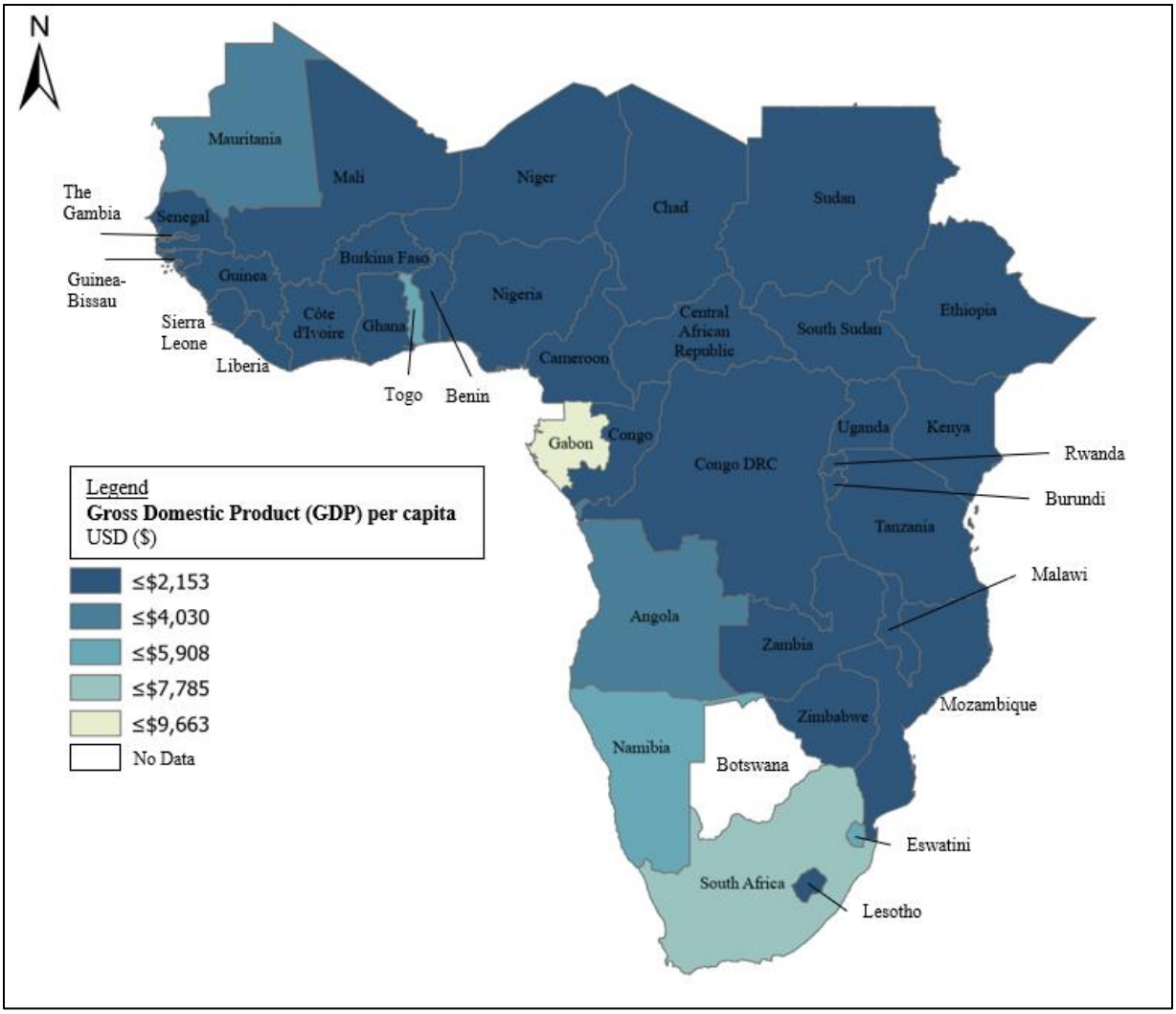

Figure 5.5 Choropleth map of gross domestic product (GDP) per capita across Sub-Saharan Africa

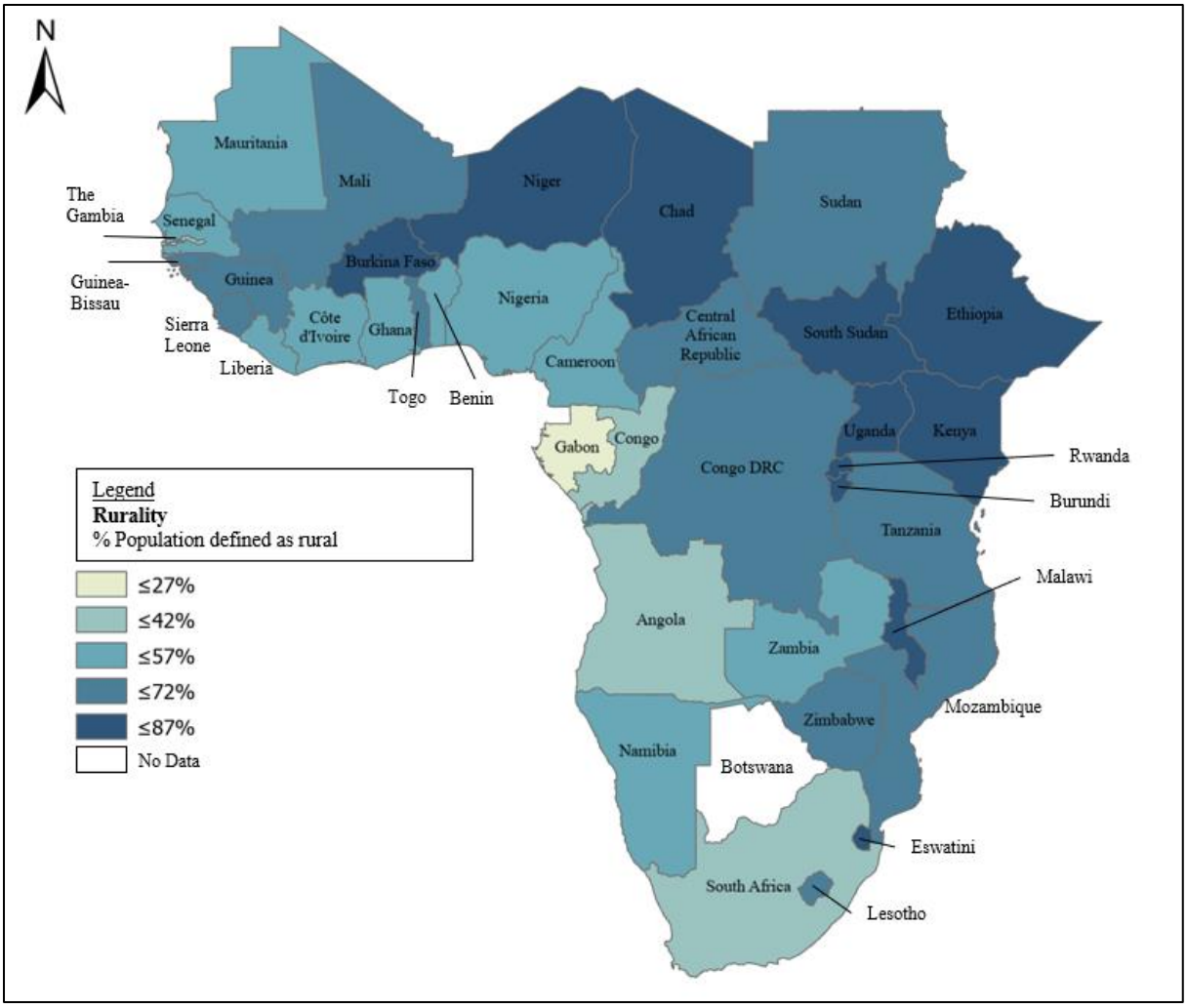

Figure 5.6 Choropleth map of rural population percentages across Sub-Saharan Africa

CHAPTER 6: WHERE'S THE WHERE? THE CASE FOR AN EXPLANATORY SPATIAL RESEARCH DESIGN (ARTICLE #3)

6.1. Background

Nearly two decades ago, Johnson and Onwuegbuzie (2004) argued in one of the field's seminal works that the time had come for mixed methods research to be taken seriously as a paradigm alongside its elder siblings, quantitative and qualitative research. That paper stemmed from an apparent newfound need for enhanced empirical integration despite Maxwell's (2016) later claim that the combination of quantitative and qualitative research in the natural and social sciences had been occurring for centuries or generations, respectively. History aside, the aim of this paper is to recognize that a similar point is in the present, where the time has come to formally welcome spatial analysis into the mixed methods family. In doing so, this article advances an emerging opportunity for a mixed methods variant, called an *explanatory spatial* design, that crystallizes a place for spatial data strands in mixed methods research (Creswell & Plano Clark, 2006). A specific variant like this will then aid future researchers with a pragmatic structure that can build on prior studies where spatial analyses have been used successfully within mixed methods frameworks to increase understanding of the ever-complex world (Teddlie & Tashakkori, 2006; Yoon et al., 2018). Ironically, in his effort to "map" the developing landscape of mixed methods, Creswell's (2010) notable omission of spatial mention leaves geographers excluded from this increasingly valued field of mixed methods research (McKim, 2017).

The contributions of geographers' prior efforts in combining their craft with other methods should not go unnoticed. Examples of relevant, spatially-driven, empirical

studies, each with dozens of citations, range from using mixed methods to address socio-spatial differences in urban gardening (McClintock et al., 2016; Preston & Wilson, 2014) or combining methods to identify disparities among low-income populations with regard to transportation services or healthy food options (Shannon, 2015; Shay et al., 2016). Recognizing this growing utility and employment of spatial analysis tools, geographic information systems (GIS), and spatial data in applied geographic settings, some authors have sought to make cases for mixed-methods geospatial approaches (Harbers & Ingram, 2017; Yeager & Steiger, 2013; Yoon & Lubienski, 2017). Furthermore, Brown et al. (2017) and Jones (2017) offer some of the only comprehensive critical examinations of spatial analyses within existing mixed methods frameworks, emphasizing validity and spatial representations, respectively. These two articles are key examples where the authors placed the greatest priority on a methodological critique. Therefore, they were influential in developing this paper, which aims first to advance a specific mixed methods variant while secondarily contributing to an emerging geography of educational policy (Yoon et al., 2018).

Spatial methods have already begun to carve a place for themselves in mixed methods research in two distinct ways. The first is called "qualitative GIS," and could perhaps be considered a mixed methods variant in its own right considering that there have already been entire texts advancing it into the fold of mixed method literature (Cope & Elwood, 2009). As the term suggests, qualitative GIS enables researchers to incorporate the accounts of individuals' lived experiences into innovative mapping approaches in an attempt to solve complex, sociospatial, problems (Jung & Elwood, 2010; Yoon & Lubienski, 2017). A desire to understand communities' perceptions of

space, as well as their interactions with or imaginations of it, prompted social scientists to combine some of the context-rich strengths of qualitative research with the cartographic approaches for visualizing the lived environment (Baur et al., 2014; Rucks-Ahidiana & Bierbaum, 2015). In fact, these were several of the theoretical forces behind participatory GIS, which, after emerging in the 1990s, likely set the foundation for what later became qualitative GIS (Elwood, 2006). Participatory GIS, like its later methodological cousin qualitative GIS, became an effective technique for "democratizing" GIS because it enabled participants to have control over how their experiences with place would be geographically represented, thereby better informing social processes (Abbot et al., 1998; Dunn, 2007). As the case study presented later in this article will show, an explanatory spatial design can accomplish this same objective of enhancing social inquiry and policy.

The second instance where geographic techniques have found a place within existing mixed methods frameworks, especially in recent years, is in visualizing the results of mixed methods studies. Dickinson (2010) argued that graphics in mixed methods research, including a specific eye toward maps, can allow audiences to gain inferences through the visual itself or through the graphics' ability to integrate sources. Creswell and Plano Clark (2018) emphasized the latter when covering advanced techniques for integrating data strands in a mixed methods study. One of the benefits of qualitative GIS, coincidentally, is its ability to incorporate narrative, visual, and sometimes quantitative considerations into comprehensive, yet digestible, visualizations (Preston & Wilson, 2014). This often means that innovate ways of examining complex relationships between spatial phenomena can be possible, offering researchers with exciting new approaches to spatial problem-solving (Yeager & Steiger, 2013). Expanding

beyond just its visual benefits, Gregory and Healey (2007) remind us that the strength of geographic information systems lies in their ability to facilitate the creation and analysis of spatial databases. Building on this aspect of GIS literature, a twofold outlook on GIS as both a visualization and analysis tool is an important component of the proposed explanatory spatial design.

Given these two examples' newfound footholds into mainstream mixed methods literature, it is time for spatial considerations to form more concrete and rigorous typologies or variants. This article will advance the need for a specific *explanatory spatial* variant to bond together quantitative and spatial analyses in the pursuit of better sociospatial clarity. It will make the case for such variant by reviewing a case study where this method was creatively applied to assess the relationship between birth registration and educational access across Sub-Saharan Africa (Griffin, 2021a; 2021b). The argumentative strategy here is to break down each of the component parts of this study to show their unique contribution(s) to this design, thereby unfolding in the following manner. First, an overview of the research design will provide a general framework for understanding its specific linear protocol. Next, it will cover the dichotomous philosophical underpinnings of such pragmatic research. This will naturally lead into the rationale behind this variant's specified research questions, giving readers an idea of precisely why the "where" is an important objective of inquiry. Subsequently, it will explain how clear case selection criteria must be emplaced to create a synergistic and logical flow between the quantitative and spatial analyses. Finally, the article makes a compelling case for the implementation of multivariate cluster analysis as a technique to integrate the data strands, tying the whole study into one cohesive product. The article

will conclude by situating this variant as a peer among other mixed methods designs while being forthright in its possible limitations.

6.2. Research design: Embracing the sum of its parts

The proposed explanatory spatial design was influenced by one of the core mixed methods approaches, called an explanatory sequential design (Creswell & Plano Clark, 2018). With nearly identical aims and methodological sequences, the explanatory spatial variant substitutes a geospatial approach in place of a qualitative one in the latter portion of the study. It is helpful, then, to begin with an overview of the explanatory sequential design before advancing the idea of an innovative spatial variant. In sequence, according to Creswell and Plano Clark, the explanatory sequential design requires researchers to analyze a quantitative dataset in the first strand, determine which results need further explanation, analyze a second stage using qualitative data, and interpret the integrated results. The intent of this design, they argue, is to use the latter qualitative analysis to explain the mechanisms or processes behind the initial quantitative results. They also assert, "This design lends itself to emergent approaches in which the second phase can be designed based on what is learned from the initial quantitative phase" (p. 81). Geographic information systems (GIS) platforms possess dozens of spatial analysis tools, making them a natural fit as an emergent means for explaining initial quantitative results.

An explanatory spatial research design is simply a spatial twist of the explanatory sequential mixed methods typology, much like the way in which spatial autocorrelation became a "spatial twist" of a traditional correlation analysis (Getis, 2007). The remainder of this article will use a case study to illustrate how this explanatory spatial research

design can be effectively implemented. In the case study, Griffin (2021a) first used a quantitative multiple regression analysis to examine the relationship between access to education in Sub-Saharan Africa and birth registration rates, while controlling for country-level wealth (e.g. GDP per capita) and geography (e.g. % of population as rural). Based on those results, Griffin (2021b) then selected which variables would be included in a follow-up spatial autocorrelation analysis to understand their distribution throughout the region. The purpose was to further explain the quantitative results by showing *where* there were clusters of low educational access and any significant explanatory variables within the region. The following notation system summarizes these steps:

$$QUAN_{relationship} \rightarrow spatial_{distribution}$$

It can be interpreted by using the capital "QUAN" letters to give methodological priority to the initial quantitative results while the directional arrow and lower case "spatial" represents the deductive means by which it is used to explain initial results (Creswell, 2010; Morse, 1991).

An explanatory spatial research design is powerful because its main parts – a quantitative strand and a spatial strand – are almost as beneficial on a standalone basis as they are when integrated (the latter to be discussed later). Multiple regression, for example, is highly effective at explaining which variables in a set of factors explain an outcome, even when holding constant certain predictor variables (Arkes, 2019). In Griffin's (2021a) case, readers are able to see exactly how birth registration, GDP per capita, and rurality make unique contributions to the variance in educational access (or not). Likewise, the Global Moran's *I* analysis, a test of spatial autocorrelation, grants researchers with key insights into the effect of a variable's values in one area on its

values in neighboring areas, therefore providing a measure of clustering or dispersion (McGrew Jr. et al., 2014). Again, in Griffin's (2021b) case, this test was conducted on all six variables of interest, allowing readers to quickly compare the degree to which variables are geographically distributed in similar patterns. Although it is out of the scope here to explain all of Griffin's individual results, Table 6.1 summarizes the key findings for both data strands to show just how many key points can be found within each strand.

Table 6.1: Key findings of each strand in Griffin's quantitative (2021a) and spatial studies (2021b)

Data Strand	Key Findings
Quantitative (Griffin, 2021a)	
(Descriptive)	1. Education becomes more inaccessible with each successive level of schooling, from primary through upper-secondary education. 2. GDP per capita demonstrates the greatest degree of inequality among all six variables, ranging from $275 (Burkina Faso) to $6,132 (South Africa).
(Inferential)	1. Together, the three predictor variables - birth registration, GDP per capita, and rural population - explained 13.3% of the variance in access to the primary school ($p < 0.05$), 37.6% at the lower-secondary level ($p < 0.001$) and 35.1% at the upper-secondary school ($p < 0.001$). 2. The key explanatory variable, birth registration, was not significant in its relationship with access to education at any education level. 3. The coefficient estimates for rural population % was not significant in any model at a 5% significance level but it was significant at a 10% level at the lower- and upper-secondary school levels ($p = 0.081$ and $p = 0.085$, respectively). 4. The coefficient estimates for GDP per capita were statistically significant in all three models.
Spatial (Griffin, 2021b)	
(Descriptive)	1. Access to education is much higher, generally, across Sub-Saharan Africa at the primary-school level than at the secondary levels. 2. Birth registration displayed a few instances where there were groups of neighboring countries with similar values despite a lack of statistical clustering. 3. Choropleth map of GDP per capita highlighted the tremendous poverty levels across the region as well as the vast inequalities in income distribution.
(Inferential)	1. Access to education demonstrated statistically significant clustered patterns ($p<0.05$) at the primary and lower-secondary levels. 2. Birth registration, which was the key explanatory variable influencing this study, showed no significant evidence of spatial autocorrelation across Sub-Saharan Africa. 3. The rural populations of countries in Sub-Saharan Africa were significantly clustered at a $p<0.05$ level.

As Table 6.1 makes clear, the quantitative and spatial strands both yield results that are descriptive and inferential. In essence, this produces four distinct categories of findings. The benefit of this feature is that it gives researchers multiple ways of examining the issue at hand. Additionally, should some of these categories fail to generate any noteworthy findings, this improves the chances that at least one category will be of value to the overall objective of the study. For example, in both inferential components Griffin's case study, birth registration – the key explanatory variable of interest – did not significantly contribute to the variance in educational access in the multiple regression analysis nor did it demonstrate any significant instances of spatial autocorrelation. This was surprising given the detailed literature review and theoretical grounds for suggesting that either of those outcomes should not have been the case. Examining the choropleth map for birth registration rates in Sub-Saharan Africa, or the descriptive component of the spatial strand, uncovered a consistent pattern with a similar UNICEF map of birth *registration* while deviating considerably from the spatial distribution of birth *certificate* rates. Therefore, the explanatory spatial design proved valuable as the two inferential components nearly closed the door on the relationship between birth-related legal identity and access to education, while the descriptive mapping opened the door to a clear pathway for future research.

Another key feature of an explanatory spatial design is the practical value of explaining quantitative findings with subsequent geospatial techniques. Determining where and how the social phenomena of interest manifest over space is a logical next step in geographic research processes and is especially valuable for identifying certain spatial processes (McGrew Jr. et al., 2014). In fact, this type of geographic discovery is quickly

becoming much more useful across the social sciences as its applications are plentiful across academic disciplines, such as healthcare, business/marketing, education, and environmental justice, to name a few (Ballas et al., 2017). In that sense, Griffin's spatial findings (2021b) illuminated both individual outlier countries, where UNICEF or UNESCO could apply essential development aid to improve access to education, and clusters of countries, where researchers might need to pursue additional insight into their accessibility issues. Interestingly, though a previous argument praised the standalone value of each data strand, consider them here without one another. Absent the initial quantitative results, the spatial findings likely would not have brought attention to the effects of certain explanatory variables, such as rurality, on access to education. This would leave the aforementioned humanitarian organizations without a comprehensive outlook on the situation at hand for a particular geography. Conversely, without the spatial findings, interested parties would need to scour a raw data table to determine which countries are in most need of better policies or more in-depth research.

The sum of the quantitative and spatial parts in an explanatory spatial research design is clearly a comprehensive account of the statistical relationships and geographic distributions behind complex social phenomena. The intent and strength behind this design is for the latter to further explain the former in a logical, straightforward manner (Creswell & Plano Clark, 2018). In pursuit of such an objective, the strands will likely complement one another while offering valuable insight in their own right. Later in this article, the advancement of an explanatory spatial variant will be driven home by demonstrating how the mixed methods integration will make the study greater than the sum of these parts. In the interim, however, this article asserts that the philosophical

perspectives, research questions, and case selection criteria are all intricate factors that must be evaluated when considering the employment of an explanatory spatial design.

6.3. Philosophical foundations: Giving space a place

When it comes to the philosophical underpinnings of an explanatory spatial variant, it is best to once again follow Creswell and Plano Clark's (2018) direction and consider using multiple philosophical assumptions to fit the differing strands of data in a sequential research design. Although methodological or philosophical purists might point out the inherent tensions or contradictions in such an approach, others argue that dialectical pluralism builds a paradigmatic setting that is both richly theoretical and logically practical (Greene & Hall, 2010; R. B. Johnson, 2017; Tashakkori & Teddlie, 2010). Unpacking this philosophical jargon could in and of itself lead to entire discussions outside of this article's scope. To avoid such endeavor, this sub-section recognizes that adding a spatial twist to the explanatory sequential research design is best suited for a pluralistic approach based on the practical and just nature of spatiality in social science research. Human geographers, particularly with the increasing popularity of quantitatively oriented spatial analysis tools, have resorted to adopting pluralistic approaches as they assess the nature of spatial realities and knowledge production (Poon, 2005). These emerging spatial constructs have also brought more attention to the power of GIS in producing conceptions of space that have a solid place among mixed methods literature (Giordano & Cole, 2018). Thus, an explanatory spatial design enables researchers to draw benefits of pragmatist and postpositivist philosophical assumptions, as shown in the corresponding theoretical framework for this design (Figure 6.1).

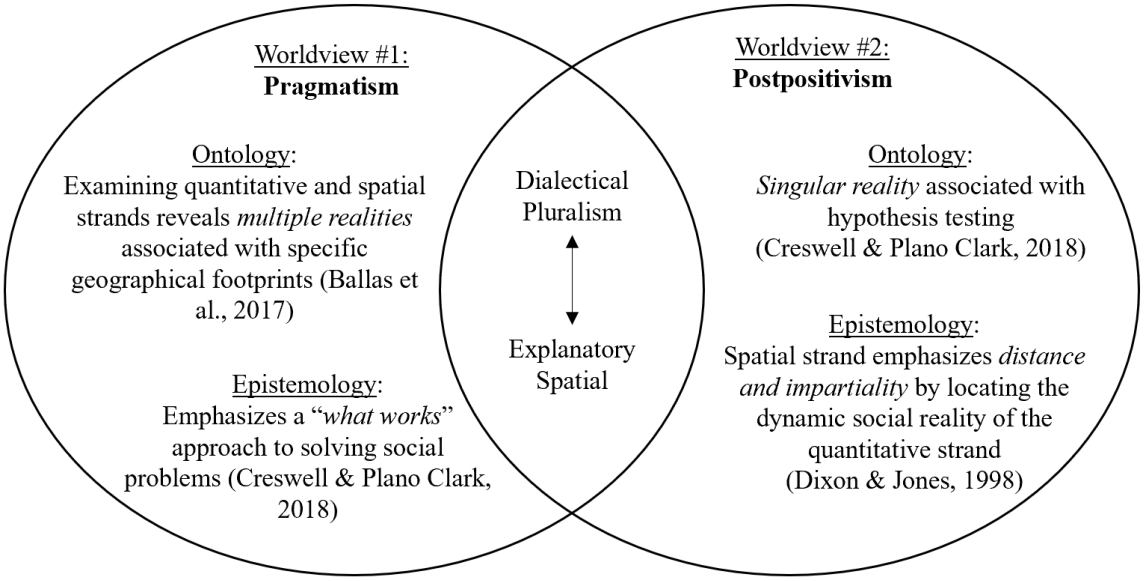

Figure 6.1 A theoretical framework highlighting the dialectical pluralism associated with an explanatory spatial research design

Pragmatism is an ideal worldview to incorporate spatiality because it enables researchers to embrace a problem-centered perspective in producing socially useful knowledge (Creswell & Plano Clark, 2018; Yvonne Feilzer, 2010). Space, in human geography and the social sciences, is inclusive of theoretical, conceptual, and practical applications for understanding the lived experiences of those in the world around us (Ballas et al., 2017; Dixon & Jones, 1998; Tuan, 1979). This philosophical assumption led to Griffin's case study in Sub-Saharan Africa, as his motivation for undertaking the research was a practical need to understand a problem that he witnessed firsthand in Cameroon, where birth registration acted as a barrier to secondary education. Unexpectedly, his quantitative analysis did not yield any significant relationship between birth registration and educational access beyond Cameroon. According to Mertens et al. (2010), however, quantitative data does not always reflect observed phenomena on the

ground, often forcing international development organizations to adopt mixed methods for deeper analyses as a matter of practice. Conversely, Ballas et al. (2017) also argued that although objects may not always require a perfect representation in GIS, they still do have a true geographical footprint in terms of location, area, and shape. Therefore, because quantitative and spatial strands can reveal multiple realities, from an ontological perspective, pragmatism was an ideal philosophical worldview for determining which is best suited for the research question(s) at hand (Creswell & Plano Clark, 2018).

An explanatory spatial research design is also fitting for epistemological pluralism, as researchers focus on "what works" to solve a problem (e.g. pragmatism) while maintaining a postpositivist emphasis on distance and impartiality (Creswell & Plano Clark, 2018). Distance is arguably the most common spatial characteristic as many research problems across the social sciences seek an understanding of how nearby phenomena relate to one another (Ballas et al., 2017). This can be seen in a popular discussion of spatial theory, or "grid epistemology," where Dixon and Jones (1998) argued that spatial considerations are necessary for, "…locating and segmenting a complex, relational, and dynamic social reality" (p. 251). In Griffin's (2021) case study, his approach for understanding the dynamic social reality was a cross-national one, allowing for an analytical differentiation between country-specific particularities and socio-structural regularities across Sub-Saharan Africa (Kohn, 1987). Accomplishing this task required datasets from United Nations sources that had established rigorous quality standards for including consistent household survey information in their databases (United Nations Children's Fund, 2020a; United Nations Educational, Scientific and Cultural Organization, 2020c). In other words, the researcher critically relied on

impartiality as a practical mechanism for ensuring that he could gain a trustworthy look into the complex research questions of interest (Creswell & Plano Clark, 2018). So, though not all cases will need the specific combination of pragmatism and postpositivism, the key point here is that an explanatory spatial design benefits tremendously from a paradigmatic pluralism that offers an effective combination of flexibility and rigor.

6.4. Research questions: Where is the where?

Mixed methods research questions, according to Plano Clark and Badiee (2010), are often generated from the researcher(s)' worldviews and are influenced by their communities of practice. This makes the topic of research questions a natural next step in evaluating the pragmatic nature of an explanatory spatial research design. However, just as Gregory et al. (2015) recognized a need to define appropriate research questions in the rapidly developing research field of spatial humanities, there is a need here to understand how research questions can best lead to meaningful pursuits of spatial inquiry. After all, nowhere in Creswell and Plano Clark's (2018) detailed review of mixed methods research questions do they recognize the need to ask "where" in explaining the intersections of mixed methods data strands. This is noteworthy, as Getis (2008) showed several ways in which spatial analysis techniques (specifically, spatial autocorrelation) can address various research questions, ranging from understanding temporal autocorrelation to the explicit examination of spatial effects. He recognized that questioning "spatial relations," or the relationships among geographic units, was a key concept that could aid researchers in the natural and social sciences in understanding

distance-related effects like cost, effort, friction, and decay effect. To that end, Griffin's case study provides one example of how spatiality can be included in a set of mixed methods research questions to effectively lead readers into understanding the process of explaining initial quantitative results with subsequent spatial analyses.

Interestingly, none of Griffin's research questions directly ask "where" certain phenomena exist, yet all three of them contain an inherently spatial component (see Table 6.2 for a list of his research questions). As Creswell and Plano Clark (2018) recommended, he included one research question for each data strand (i.e. quantitative and spatial) and a third explicitly spatial question. The first question asked about the relationship between birth registration and educational access across the countries of Sub-Saharan Africa. Although this is a quantitative question, the defined study area (e.g. Sub-Saharan Africa) gives readers critical insight into the geographic unit of interest. Thus, a natural next step, from a human geographer's perspective, is to understand how the spatial distribution of these variables are represented throughout the study area to determine which areas might explain the relationship(s) found in the first question. The second question, then, asked if there was a non-random distribution of these variables in the region, which alluded to a null hypothesis where there the variables of interest are randomly dispersed throughout the area. This question therefore helps researchers examine where there might be statistically significant clusters of countries for each variable. Finally, the mixed methods question aims to pair the data strands by addressing the extent to which the latter spatial results help to explain the initial quantitative results. So, while the word, "where," is not literally necessary in any question, the underlying assumption is that had a significant relationship existed between birth registration and

access to education, then practitioners, like UNICEF, could have quickly understand which clusters of countries might be contributing to such a result. Furthermore, the research questions will inevitably change from study to study, but those questions guiding Griffin's case study offer valuable insight into the multi-purposed questions that an explanatory spatial research design can aim to address.

Table 6.2: Research questions for each strand in Griffin's (2021) explanatory spatial research design

Data Strand	Research Question
Quantitative	Is there a relationship between birth registration and educational access across the countries in Sub-Saharan Africa since 2010, while controlling for GDP per capita and rurality?
Spatial	Is there a non-random pattern in the spatial distribution of birth registration, GDP per capita, rurality, and educational access across countries in Sub-Saharan Africa since 2010?
Middle Methods	To what extent does the spatial distribution within birth registration, other socioeconomic or geographic variables, and educational access help explain the relationship between them in a cross-national study across countries in Sub-Saharan Africa since 2010?

6.5. Case selection: Choosing with purpose

Another individual component of an explanatory spatial research design worthy of its own review is the case selection process that occurs between the first and second data analysis stages. Koivu and Hinze (2017) argue that case selection is a key component of mixed methods research, yet they also acknowledged that researchers need to be more transparent in the techniques or strategies they choose to bridge this gap between data strands. Likewise, Ivankova et al. (2006), while focusing specifically on an explanatory sequential design, recognized the significance of case selection between stages. However, they asserted that there was little evidence of established guidelines for researchers to

follow when attempting to proceed with such methodological decisions. There were even fewer case selection processes when incorporating spatial tools into mixed methods studies until Ingram and Harbers (2020) advanced two strategies for taking spatial considerations into the case selection decisions. This section therefore aims to build on their detailed framework by sharing a case selection process in an explanatory spatial design. It will deviate from the preceding sections, which emphasized the need for the design itself, by reviewing Griffin's (2021) case selection criteria as a model for implementing a clearly articulated roadmap that explains the transitions from quantitative analysis to spatial analysis.

Choosing which cases from the initial quantitative strand to include in the subsequent (in this case, spatial) strand can occur in a few ways. In some cases, not all confirmatory, explanatory, or exploratory analysis criteria meant for the second data strand can be specified in advance (Nastasi et al., 2010). In these cases, a key purpose of the sequential design is to understand what the initial results are indicating, which will then enable the researcher to more specifically design the secondary data strand. Although this can sometimes have administrative challenges when it comes to institutional review processes (Creswell and Plano Clark, 2018), mixed methods researchers often face logistical constraints that make this choice the best option (Koivu & Hinze, 2017). In other cases, researchers determine which results will be explained in the latter analysis stage by deciding ahead of time how significant results, non-significant results, outliers, and/or group differences will be used to refine or guide the sequential stage (Creswell and Plano Clark, 2018). Seawright (2016) argued that methodological literature often focused on researchers' tendency to select cases based on the

characteristics of dependent variables and instead advanced case selection strategies based on the extreme values found within a study's key independent variable(s). Emphasizing the former, Ingram and Harbers (2020) argued that identifying geographic clustering of outcomes enables one to understand if the their outcomes defy other regional spatial patterns.

These two strategies influenced Griffin's (2021) case study in the sense that he placed an emphasis of case selection based on which regression models contained significant results across its independent variables while also examining the dependent variables using a follow-up spatial autocorrelation (see Figure 6.2). Statistical significance need not necessarily always be of key importance in a study (Arkes, 2019), though it made logical sense to include in Griffin's model. In the quantitative analysis, he employed three multiple regression models with birth registration, rurality, and GDP per capita as predictor variables and access to education composite measures at three educational levels as the dependent variables. Any model consisting of statistically significant relationships among any variables would result in all variables from that model being included in the subsequent spatial analysis (see criteria #1). The spatial autocorrelation tests were intended to identify significant clusters within each variable to help explain how the social phenomena occurred over space. Another case selection process occurred after this step as well, as any variables demonstrating statistically significant clusters would then be included in the final integration phase (see criteria #2). The purpose here was to determine if any of these variables clustered in concert with one another in hopes of uncovering a deeper understanding of the relationships both within and between variables.

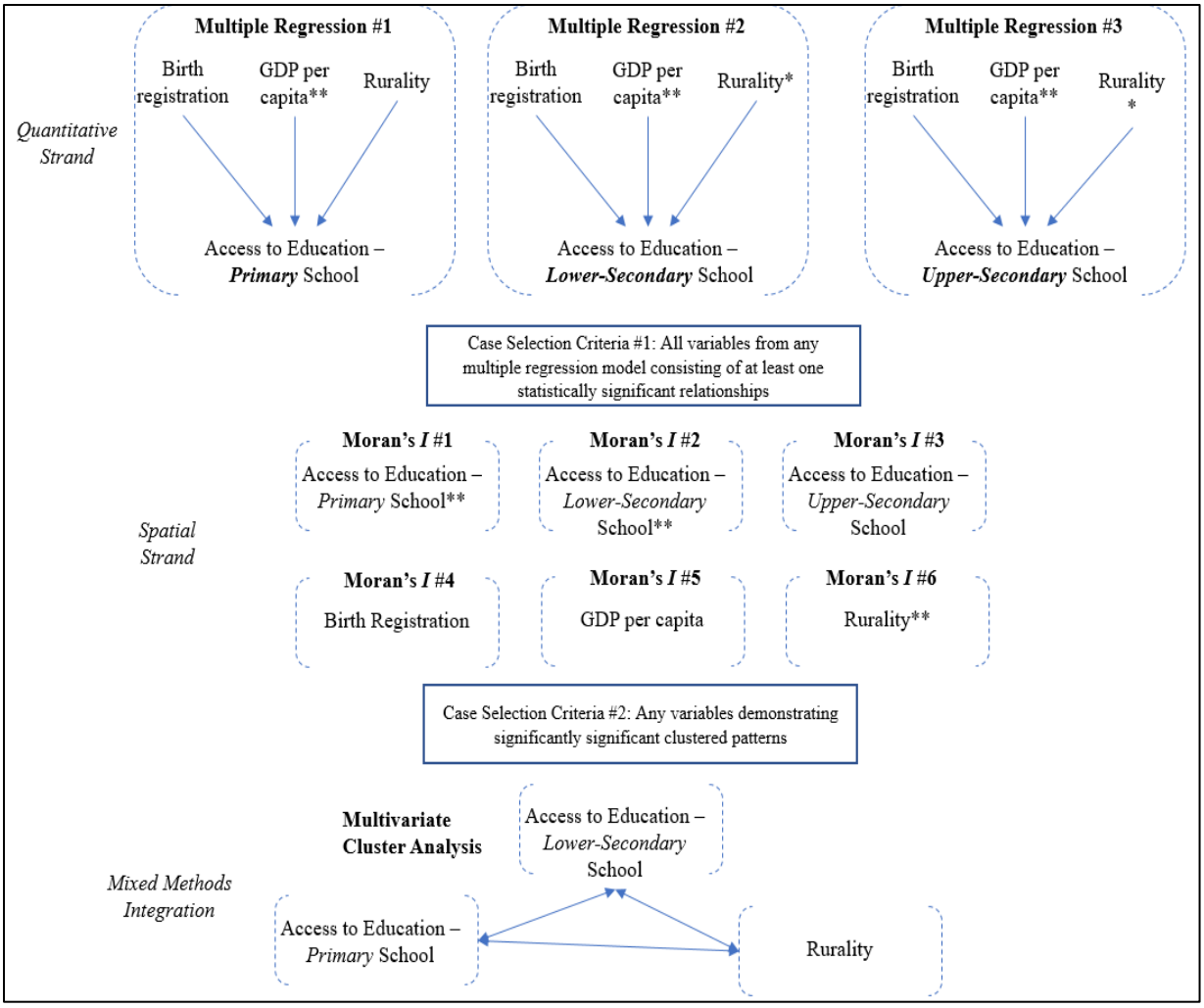

Figure 6.2 Case selection process for Griffin's (2021) explanatory spatial research design

6.6. Integration: Greater than the sum of parts

As the previous sections have shown, an explanatory spatial research variant contains many component parts, each with their own value added to the overall strength of such a design. The culminating point of this article is that the final phase in this design, particularly the integration of quantitative and spatial factors, will provide unquestionable value to a researcher's inquiry into sociospatial phenomena. In one of the early editions of the *Journal of Mixed Methods Research*, O'Cathain et al. (2007, p. 60) recognized the

"considerable potential" of integration during the analysis stages of mixed methods studies after reviewing how few researchers had covered the topic up to that point. Although there is considerable opportunity for mixed methods integration with the rapid acceleration of computer technologies (Bazeley, 2010), there has still been disagreement over the utility of these efforts. On one hand, critics assert that integration is leading to "diffraction," where the study becomes interrupted or splintered (Uprichard & Dawney, 2019). On the other, these technologies, like "geo-referencing," are said to contribute to "richness," or enhanced analytic density, according to (Fielding, 2012). An explanatory spatial design certainly leans toward the latter, as it "achieves integration" by merging carefully selected variables (i.e. through case selection processes described above) using a GIS platform (Fetters et al., 2013).

One method for integrating data in an explanatory spatial research design is through a multivariate cluster analysis. Whereas the quantitative strand allows one to identify statistical relationships between variables and the spatial strand helps illuminate the spatial distributions within the variables, a multivariate cluster analysis can effectively merge the two. A cluster analysis, according to Shaban et al. (2010), "…seeks to classify homogeneous subgroups of cases in a population by minimizing the *Within-group* and maximizing *Between-group* variations" (p. 1786). One goal here might be to understand if any variables cluster in concert with one another over a certain geographical space. Multivariate clustering using GIS has been useful in some examples of prior research, where scientists needed to understand landscape patterns (Long et al., 2010), solve geologic problems (Parks, 1966), and examine water pollution variations (Shaban et al., 2010). Furthermore, these are situated within a larger body of literature

highlighting the efficacy of multivariate methods in geospatial analyses (Marquinez et al., 2003; Shan et al., 2013; Srivastava et al., 2012). There is therefore sufficient evidence here that multivariate spatial analysis allows researchers to effectively analyze, visualize, and interpret the spatial relationships within and among multiple variables. Griffin's case study is an example of the potential for this technique when applied to an explanatory spatial design.

This article offers three useful ways in which the variables could be integrated in a GIS-based multivariate clustering analysis based on the datasets described in Griffin's quantitative and spatial articles (2021a and 2021b, respectively). ArcGIS Pro's Multivariate Clustering tool was employed here because it determines the most natural clusters in the dataset, where the features within each cluster are as similar as possible while also maximizing differences between the clusters themselves (Esri, n.da). It does so by using a k-means algorithm, which separates the data into *k* clusters based on minimizing the distances between the sum of squared errors and the mean of each cluster (Eghtesadifard et al., 2020; Jain, 2010). It is important to note that there are many ways by which the researcher can define the *k* number of clusters for use in their multivariate analysis (Chiang & Mirkin, 2010; Pelleg & Moore, 2000). The key point, though, is that it is up to the researcher to input the number of clusters to tell the GIS platform how to run its analysis. In this case, each of the three models that will be described below were broken into five clusters. The purpose of this decision was to ensure methodological consistency among each model presented here as well as the choropleth maps presented in Griffin's (2021b) spatial analysis article. In theory, this should allow readers to more easily identify trends across the differing clustered approaches.

The first way that a multivariate cluster analysis could be useful for Griffin's (2021) study would be to continue with the natural evolution of its sequential design based on the selection criteria outlined in the previous section. In this example, the selected variables would be those that demonstrated statistically significant instances of spatial autocorrelation – access to primary school, access to lower secondary school, and rural population percentage (see Figure 6.3). As mentioned earlier, the purpose here would be to determine if these variables relate to one another in any sort of visibly unique spatial pattern. In other words, it would be interesting to see if the values among these variables cluster with one another in the same way as they clustered within themselves. As the multivariate cluster box-plot and corresponding map show, there are two obvious inverse relationships when it comes to access to education at both the primary and lower secondary levels (see blue line #1 and purple line #5). In these cases, three countries with the highest average levels of access to education are also countries with the least rural populations, and, conversely, four of the Sahelian countries with the least accessible education systems are also the most rural. Interestingly, the three middle clusters on the box-plot (lines #2 – 4) do not demonstrate the same patterns of inverse relationships as the other two. This is perhaps why there was not a statistically significant relationship among these variables during the initial quantitative data analysis strand despite the fact that there are some seemingly related instances of educational access and rurality.

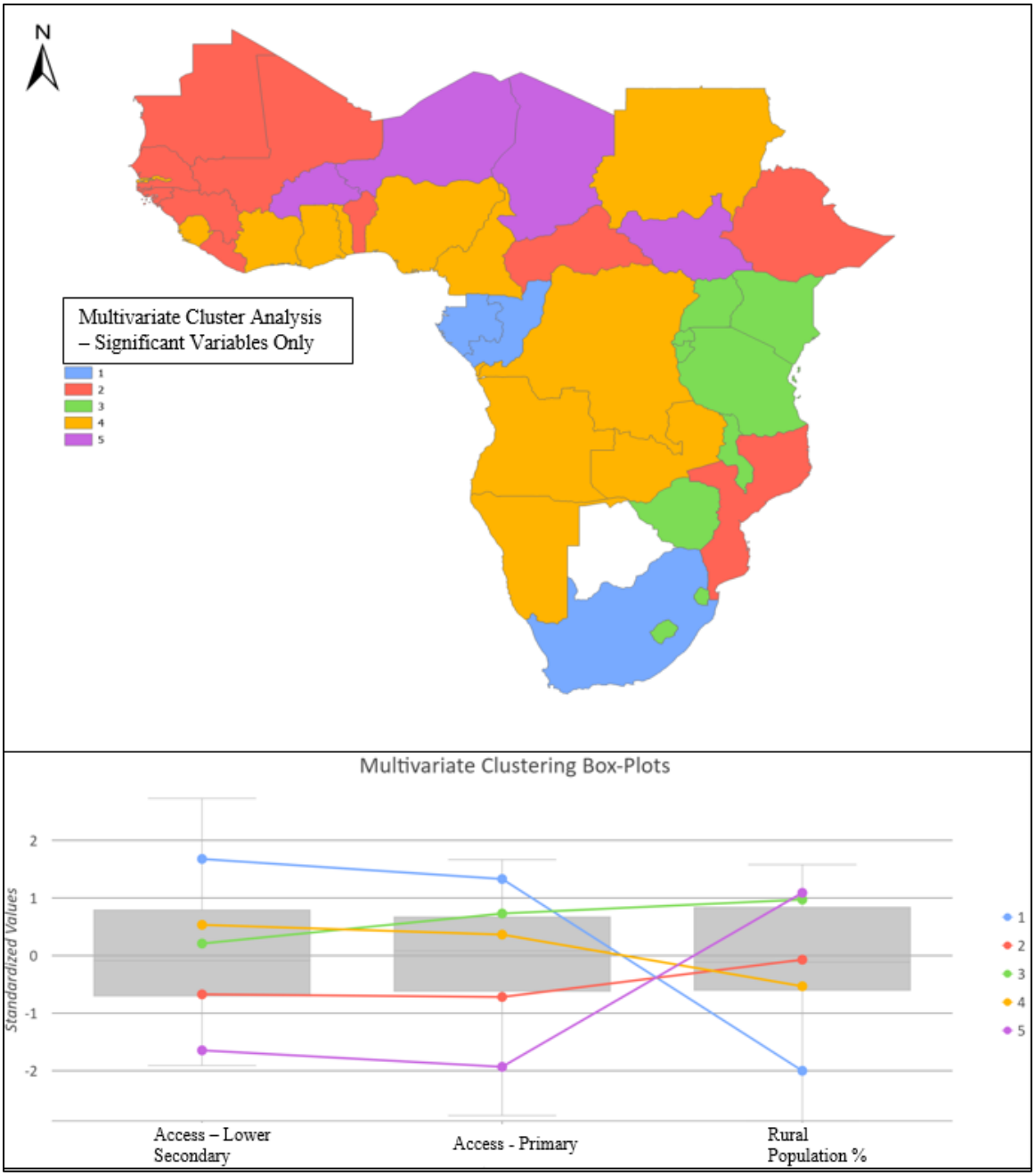

Figure 6.3 Multivariate cluster analysis #1 with choropleth map and associated box-plot. Analysis included only the variables meeting case selection criteria following Griffin's (2021b) spatial analysis.

The second option in which a multivariate cluster analysis could be useful in this case would be if the researcher subjectively chose the variables based upon the key themes of their study. In Griffin's case, he was most interested in the impact of birth registration on educational access at the primary, lower-secondary, and upper-secondary level and therefore selected these four variables (see Figure 6.4). This option uncovered two noteworthy, clustered relationships among the four variables. First, the four countries with above-average access at each educational level (see the red line #2) also displayed significantly higher birth registration levels than any other cluster. If not careful, one might conclude, then, that birth registration has a positive relationship with educational access based on this trend alone. This leads us to the second noteworthy relationship, depicted in green (line # 3), where the cluster of Central African countries with well below-average birth registration levels were actually some of the countries demonstrating average levels of educational access within this dataset. This is useful information, then, as it helps to confirm the other findings in Griffin's mixed methods study, where the quantitative strand found no significant relationship between birth registration and educational access (2021a) nor did these variables demonstrate any significant spatial clustering in concert with one another (2021b). Although there is some methodological tension over the employment of confirmatory motivations in mixed methods studies, it was helpful here to triangulate the findings from the quantitative and spatial strands (Small, 2011). Therefore, if the first option was useful in an explanatory way, then this option is perhaps best suited as a confirmatory one.

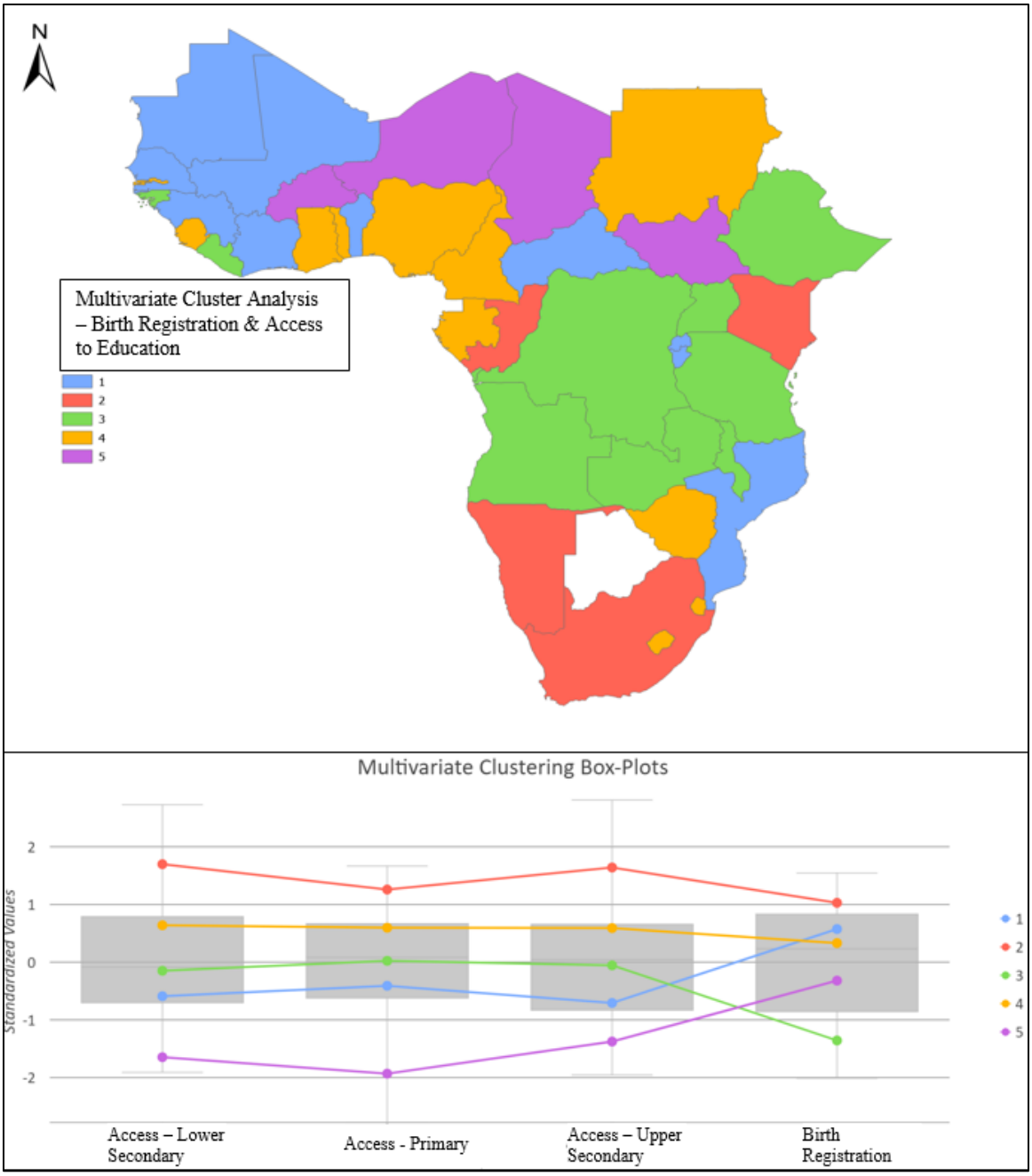

Figure 6.4 Multivariate cluster analysis #2 with choropleth map and associated box-plot. Analysis included selected variables of interest in Griffin's (2021a; 2021b) explanatory spatial study.

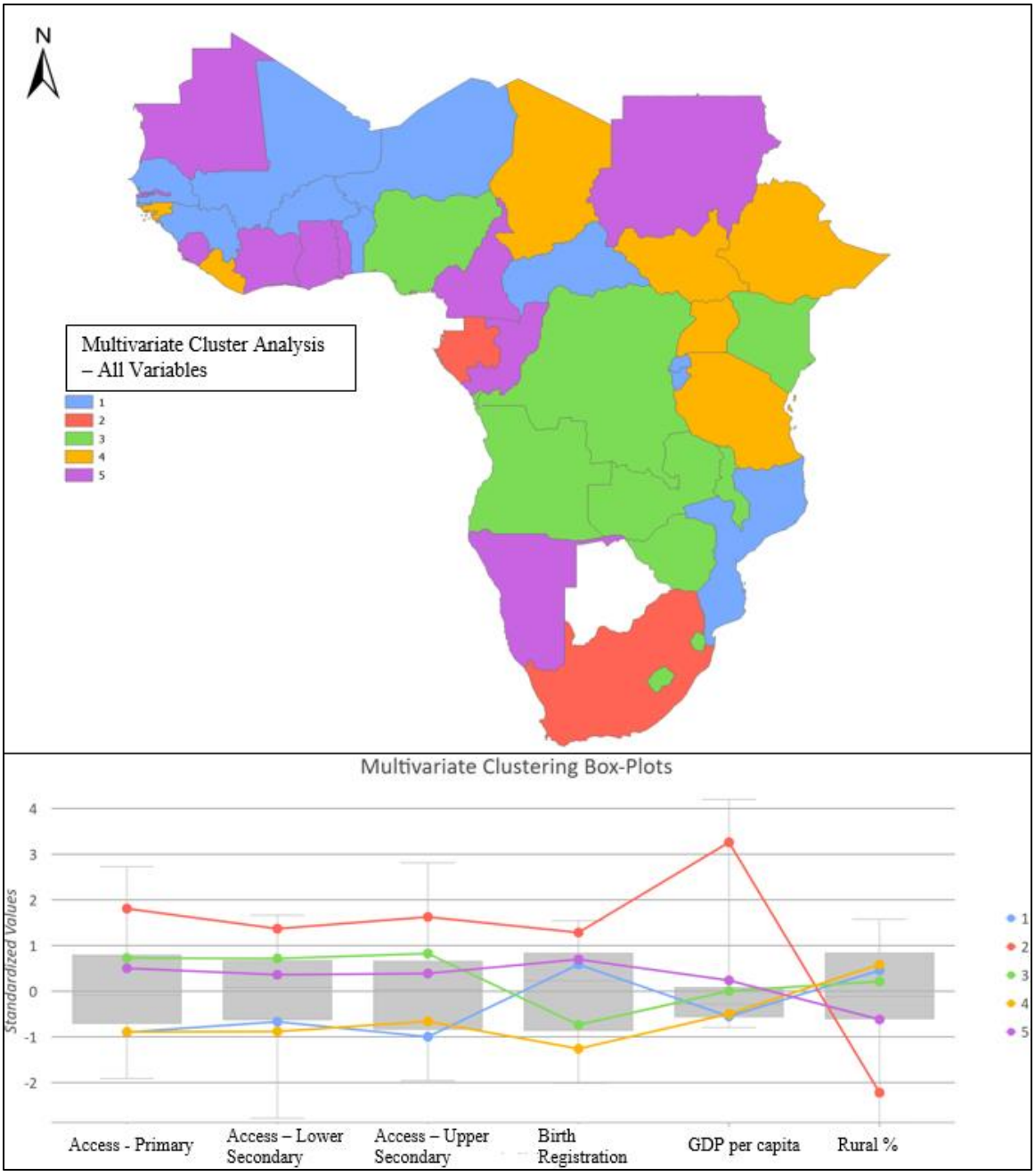

Figure 6.5 Multivariate cluster analysis #6 with choropleth map and associated box-plot. Analysis included all variables in Griffin's (2021a; 2021b) explanatory spatial study.

The third and final option that shows the benefit of multivariate cluster analysis in a comprehensive manner, where the researcher might include all variables included in the study. The purpose here would be to understand how all of the variables might cluster

together over the space within the study area. In this case, Griffin could include the four variables described in the preceding paragraph plus his other two explanatory variables, GDP per capita and rurality, for a total of six variables (see Figure 6.5). When looking at the variables this way, one could say that a trend, in general, is that countries with higher educational access have higher birth registration rates, are wealthier countries, and are demographically less rural (and vice versa). This appears to be the case for the red, orange, and purple lines on the boxplot (#2, #4, and #5, respectively). The blue and green countries therefore offer noteworthy findings as they deviate from these trends when it comes to birth registration. From a pragmatic standpoint, then, these offer opportunities for further research, as a comparative analysis among these countries might be beneficial for understanding underlying phenomena contributing to these trends.

The intent of integration is central to a mixed methods study and often differs depending on the research design or mixed methods typology (Creswell & Plano Clark, 2018). In an explanatory spatial design, integration is well suited for a cluster analysis because, according to McGrew Jr. et al. (2014), such an approach is always descriptive in nature. This offers researchers maximum flexibility for addressing their research problem(s) and the integration is dependent upon the study's guiding purpose. For example, returning to the focus on research questions from the preceding section, Griffin's (2021) mixed methods research question asked, "To what extent does the spatial distribution within birth registration and educational access variables help explain the relationship between them?" The short answer is that it depends. It is likely dependent upon the needs of the individual conducting the research. Look at the three aforementioned cluster approaches to illuminate this argument. A methodologist might be

interested in the first model to understand if the sequential flow of the design yielded any unexpectedly different results from an approach where each strand was considered on a standalone basis. The second model might be more suited for a researcher strictly interested in the spatial relationships among key variables, or in this case, birth registration and educational access. The third model might be best for an international organization, like UNICEF, as they try to determine practical solutions to educational barriers. Moreover, the bottom line is that this integrative step proves to be much more valuable than simply the sum of a quantitative strand or spatial strand in an explanatory spatial design.

An explanatory spatial design with an integration that is greater than the sum of its strands is one where clear meta-inferences are presented. A meta-inference is essentially an integrated conclusion, theoeretical statement, narrative, or other type of understanding that is gleaned from the inferences among its component strands (Tashakkori & Teddlie, 2008; Venkatesh et al., 2013). Meta-inferences can be affected by a number of design issues, ranging from the sequencing of phases to the paradigmatic setting, requiring the mixed methods researcher to be very clear in articulating these conclusive meta statements (Onwuegbuzie & Johnson, 2006). To clearly articulate this in an explanatory spatial setting, the article here proposes a meta-inference template in a similar fashion as Creswell and Plano Clark's (2018) example purpose statements, where the audience can fill in the underlined prompts. As such, a researcher could adopt a meta-inference statement, such as:

The [phenomenon, relationship, etc.] between [these variables] was observed to be located [in this spatial pattern] within [the units of this geographical space]. Therefore,

[these stakeholders] could effectively [take this action] as result of this explanatory spatial research.

This statement is powerful as a meta-inference for two reasons. First, it clearly draws the inferences from both strands in a sort of "what-where" manner. The quantitative strand shows what was observed (i.e. a relationship between variables) and the spatial strand showed where it transpired (i.e. the geographic distribution). Second, this reiterates the pragmatic nature of an explanatory spatial design, as the research is intended to provide researchers and/or policymakers with tangible and actionable direction.

6.7. Conclusion

The purpose of this article was to advance the need for an *explanatory spatial* mixed methods design. Nastasi et al. (2010) suggest that updating typologies or common mixed methods designs is a positive sign of innovation within the field and that doing so will help inform future methods, pedagogy, and communication. In that sense, this article was partly an argument in favor of a new variant and partly a roadmap for implementing it. Deconstructing its component parts allowed for a thorough case to be made in terms of its benefits to researchers (see Table 6.3). The article traced Griffin's study, the first specifically named explanatory spatial design, to show how key decisions were made within each step of the design to help audiences better understand the relationships between birth registration, demographics, and educational access across Sub-Saharan Africa. It showed how spatial methodologies can be especially useful for explaining initial quantitative results, both theoretically and empirically. It also demonstrated how deliberate research questions may force the researcher to be creative and deliberate in

selecting cases as they proceed with each step of the sequential design. Finally, this article placed an emphasis on the integration technique, also seeking to advance the use of multivariate cluster analysis as a beneficial and pragmatic approach for tying the research together into a meaningful whole.

Table 6.3: Benefits of each design component of an explanatory spatial research design

Component	Benefit to Researcher
Research Design	1. Allows researchers to explain the initial quantitative results by inspecting *where* there might be clusters of each variable within a study area 2. The study's main parts – a quantitative strand and a spatial strand – are almost as beneficial on a standalone basis as they are when integrated 3. Both data strands can yield noteworthy results that are descriptive and inferential, essentially producing four distinct categories of findings
Philosophical Pluralism	1. Enables researchers to reap the benefits of both pragmatist and postpositivist philosophical worldviews and assumptions 2. Pluralism encourages researchers to focus on "what works" to solve a problem (e.g. pragmatism) while maintaining a postpositivist emphasis on distance and impartiality (Creswell & Plano Clark, 2018)
Research Questions	Adding a research question about space (i.e. asking 'where?') fills a gap in literature mixed methods research questions (Creswell & Plano Clark, 2018; Gregory et al., 2015)
Case Selection	Clarity and specificity in the case selection process can help minimize threats to validity in a sequential research design (Creswell & Plano Clark, 2018), which is a particular concern found within geographic studies
Integration	1. Well suited for a cluster analysis because, according to McGrew Jr. et al. (2014), such an approach is always descriptive in nature, offering researchers maximum flexibility for addressing their research problem(s) 2. A multivariate cluster analysis can be tailored to fit the study's guiding purpose and therefore could be confirmatory, explanatory, or exploratory 3. Meta-inference template allows researchers to wed inferences from both strands with practically-oriented direction.

As you can see, the explanatory spatial design is certainly loaded with pragmatic and postpositivist benefits, but, like any novel idea, should be broached with a word of caution. Until the variant is tested in practice, its limitations will remain veiled. In the interim, a brief look into two existing validity concerns with mixed methods research can potentially set future researchers up for a successful implementation of such a design. First, Creswell and Plano Clark (2018), argue that validity can be threatened in an explanatory sequential design (i.e. the typology most similar to that presented in this paper) by failing to identify key quantitative results in need of further explanation, failing to explain surprising results found in the subsequent analysis strand, and failing to connect the two strands with one another. Each of these issues were addressed earlier in this article, particularly in the sections on case selection and spatial integration, and could offer researchers guidance in connecting quantitative and spatial strands together. Second, researchers should be careful of forming meta-inferences based on their explanatory spatial designs (Onwuegbuzie & Johnson, 2006), as the design is primarily intended to understand the sociospatial context(s) and nuances of the specific study area itself. In fact, Griffin (2021a; 2021b) avoided the need to exhaust any of Onwuegbuzie and Johnson's (2006) legitimation types related to meta-inference by simply clearly stating that his purpose was to understand barriers to educational access in Sub-Saharan Africa, specifically. It is important to make this distinction, as many researchers without familiarity of spatial techniques may be left grasping for generalizability, or external validity, especially given its deep ties to quantitative research (Polit & Beck, 2010).

Although researchers can address threats to mixed methods validity through case selection, multivariate integration, and naming the study's specific context(s), there may

still be validity or reliability limitations due to the geospatial nature of the design. Researchers have argued that validity can be a challenge when investigating geographic problems, often due to difficulties in expressing complex spatial concepts with existing technologies or inconsistencies in international data quality (Brown et al., 2017; McGrew Jr. et al., 2014). Others have attempted to make GIS techniques more accessible or usable in response to some of these concerns (Ballatore et al., 2020). There is an opportunity, then, to begin to address these validity considerations by integrating them within existing mixed methods validation frameworks, like that presented in Dellinger and Leech's (2007, p. 322) guide for addressing validity in traditional qualitative and quantitative settings. Here, they present a framework inclusive of construct validation elements for each mixed methods data strand, including categories for design quality, interpretive rigor, legitimation, and measurement-related reliability. Therefore, one logical next step in the advancement of an explanatory spatial design would be to create a similar validation framework inclusive of spatial techniques once they have been assessed in empirical settings beyond Griffin's (2021) lone example.

With a well-planned implementation strategy that surpasses the sum of its parts while accounting for its limitations, an explanatory spatial research design has the potential to make several contributions to mixed methods research. First, and foremost, it is a new variant that offers researchers with a unique way of expanding their methodological tool kits to solve complex sociospatial problems. Second, it contributes to the growing lists of mixed methods designs, made by and for educational and social science researchers, which has a notable omission of any spatial reference (Creswell & Plano Clark, 2018; Teddlie & Tashakkori, 2006). Third, it builds on the work of

Qualitative GIS by highlighting its applicability to solving real-world social issues, thereby continuing to highlight the utility of geographic considerations in mainstream mixed methods literature (Brown et al., 2017; Cope & Elwood, 2009; Jung & Elwood, 2010; Shay et al., 2016). Finally, from a philosophical perspective, the explanatory spatial design, at least in Griffin's (2021) case, supported literature advocating for a dialectical pluralism, suggesting that pursuing methods that overcome the inherent tensions between worldviews, like pragmatism and postpositivism, can actually generate rich new knowledge (R. B. Johnson, 2017; R. B. Johnson & Onwuegbuzie, 2004; Onwuegbuzie & Frels, 2013). After all, the social world around us, from our communities to the way we communicate to our societal structures, is full of inherently spatial interactions (Ballas et al., 2017). Therefore, an explanatory spatial design is one way, until proven ineffective, to conceptualize and begin to understand these ever-complex sociospatial relationships.

BIBLIOGRAPHY

Abbot, J., Chambers, R., Dunn, C., Harris, T., Merode, E., Porter, G., Townsend, J., & Weiner, D. (1998). Participatory GIS: Opportunity or oxymoron. *PLA Notes.*, *33*, 27–33.

AbouZahr, C., Adjei, S., & Kanchanachitra, C. (2007). From data to policy: Good practices and cautionary tales. *The Lancet*, *369*(9566), 1039–1046. https://doi.org/10.1016/S0140-6736(07)60463-2

Agnew, J. A. (2011). Space and place. In J. A. Agnew & D. N. Livingstone (Eds.), *The SAGE handbook of geographical knowledge* (pp. 316–330). SAGE Publications.

Ahmed, A. (2020, August 27). *Schools caught up in armed conflict sweeping across the Sahel*. UNHCR. https://www.unhcr.org/news/stories/2020/8/5f4782404/schools-caught-armed-conflict-sweeping-across-sahel.html

Ahmed, A. Y., Miller, V. W., Gebremeskel, H. H., & Ebessa, A. D. (2019). Mapping inequality in access to meaningful learning in secondary education in Ethiopia: Implications for sustainable development. *Educational Studies*, *45*(5), 554–581. https://doi.org/10.1080/03055698.2018.1509777

Ajmera, M., & Fields, G. A. (2016). *Invisible children: Reimagining international development at the grassroots*. Palgrave Macmillan.

Akomolafe, A. A., Fadiji, F. A., & Gayawan, E. (2019). Evaluation of geographical variation in live-birth registrations using Bayesian method. *Communications in Statistics - Theory and Methods*, 1–14. https://doi.org/10.1080/03610926.2019.1625921

Ali, M., Fjeldstad, O.-H., Jiang, B., & Shifa, A. B. (2019). Colonial legacy, state-building and the salience of ethnicity in Sub-Saharan Africa. *The Economic Journal*, *129*(619), 1048–1081. https://doi.org/10.1111/ecoj.12595

Allen, M. (2017). *The SAGE encyclopedia of communication research methods*. SAGE Publications, Inc. https://doi.org/10.4135/9781483381411

American Psychological Association. (2019). *Publication manual of the American Psychological Association* (7th ed.). American Psychological Association.

Amoko, A. (2013). Race and postcoloniality. In S. Malpas & P. Wake (Eds.), *The Routledge companion to critical and cultural theory* (2nd ed., pp. 131–143). Routledge.

Amouzou, A., Mehra, V., Carvajal–Aguirre, L., Khan, S. M., Sitrin, D., & Vaz, L. M. (2017). Measuring postnatal care contacts for mothers and newborns: An analysis of data from the MICS and DHS surveys. *Journal of Global Health*, *7*(2). https://doi.org/10.7189/jogh.07.020502

Anaduaka, U. (2020). *An empirical analysis of the socioeconomic and demographic characteristics influencing birth certification in Nigeria*. https://assets.researchsquare.com/files/rs-18158/v3_stamped.pdf

Andreotti, V. (2011). *Actionable postcolonial theory in education*. Palgrave Macmillan.

Anlimachie, M. A., & Avoada, C. (2020). Socio-economic impact of closing the rural-urban gap in pre-tertiary education in Ghana: Context and strategies. *International Journal of Educational Development*, *77*, 1–12.

Ansell, B., & Lindvall, J. (2013). The political origins of primary education systems: Ideology, institutions, and interdenominational conflict in an era of nation-building. *The American Political Science Review*, *107*(3), 505–522.

Apland, K., Blitz, B. K., Calabia, D., Fielder, M., Hamilton, C., Indika, N., Lakshman, R., Lynch, M., & Yarrow, E. (2014). *Birth registration and children's rights: A complex story*. Plan International. http://eprints.mdx.ac.uk/17346/

Appell, A. R. (2014). Certifying identity. *Capital University Law Review*, *42*(361), 1–42.

Araya-Ajoy, Y. G., Bolstad, G. H., Brommer, J., Careau, V., Dingemanse, N. J., & Wright, J. (2018). Demographic measures of an individual's "pace of life": Fecundity rate, lifespan, generation time, or a composite variable? *Behavioral Ecology and Sociobiology*, *72*(75), 1–14.

Arias, S. (2010). Rethinking space: An outsider's view of the spatial turn. *GeoJournal*, *75*(1), 29–41.

Arkes, J. (2019). *Regression analysis: A practical introduction*. Routledge.

Asankha, P., & Takashi, Y. (2011). Impacts of universal secondary education policy on secondary school enrollments in Uganda. *Journal of Accounting, Finance and Economics*, *1*(1), 16–30.

Atkinson, A. B., & Brandolini, A. (2001). Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD Countries as a case study. *Journal of Economic Literature*, *39*(3), 771–799. https://doi.org/10.1257/jel.39.3.771

Avruch, K., & Black, P. W. (1987). A generic theory of conflict resolution: A critique. *Negotiation Journal*, *3*(1), 87–96. https://doi.org/10.1007/BF00999036

Ayede, A. I., Kirolos, A., Fowobaje, K. R., Williams, L. J., Bakare, A. A., Oyewole, O. B., Olorunfemi, O. B., Kuna, O., Iwuala, N. T., Oguntoye, A., Kusoro, S. O., Okunlola, M. E., Qazi, S. A., Nair, H., Falade, A. G., & Campbell, H. (2018). A prospective validation study in South-West Nigeria on caregiver report of childhood pneumonia and antibiotic treatment using Demographic and Health Survey (DHS) and Multiple Indicator Cluster Survey (MICS) questions. *Journal of Global Health*, *8*(2), 1–11. https://doi.org/10.7189/jogh.08-020806

Azzopardi, P. S., Hearps, S. J. C., Francis, K. L., Kennedy, E. C., Mokdad, A. H., Kassebaum, N. J., Lim, S., Irvine, C. M. S., Vos, T., Brown, A. D., Dogra, S., Kinner, S. A., Kaoma, N. S., Naguib, M., Reavley, N. J., Requejo, J., Santelli, J. S., Sawyer, S. M., Skirbekk, V., … Patton, G. C. (2019). Progress in adolescent health and wellbeing: Tracking 12 headline indicators for 195 countries and territories, 1990–2016. *The Lancet*, *393*(10176), 1101–1118. https://doi.org/10.1016/S0140-6736(18)32427-9

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*. Routledge.

Ballas, D., Clarke, G., Franklin, R. S., & Newing, A. (2017). *GIS and the social sciences: Theory and applications*. Routledge.

Ballatore, A., McClintock, W., Goldberg, G., & Kuhn, W. (2020). Towards a usability scale for participatory GIS. In P. Kyriakidis, D. Hadjimitsis, D. Skarlatos, & A. Mansourian (Eds.), *Geospatial technologies for local and regional development* (pp. 327–348). Springer International Publishing. https://doi.org/10.1007/978-3-030-14745-7_18

Barakat, B., & Urdal, H. (2009). *Breaking the waves? Does education mediate the relationship between youth bulges and political violence?* The World Bank. https://doi.org/10.1596/1813-9450-5114

Baron, A. (2012, November 9). *At United Nations Headquarters in New York, UNICEF Executive Director Anthony Lake and other high-level panelists discuss the importance of birth registration.* ReliefWeb. https://reliefweb.int/report/world/united-nations-headquarters-new-york-unicef-executive-director-anthony-lake-and-other

Baur, N., Hering, L., Raschke, A. L., & Thierbach, C. (2014). Theory and methods in spatial analysis. Towards integrating qualitative, quantitative and cartographic approaches in the social sciences and humanities. *Historical Social Research / Historische Sozialforschung*, *39*(2), 7–50.

Bazeley, P. (2010). Computer-Assisted Integration of Mixed Methods Data Sources and Analyses. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 431–468). SAGE Publications.

Beane, J. A. (1998). Reclaiming a democratic purpose for education. *Educational Leadership*, *56*(2), 8–11.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). University of Chicago Press.

Bennouna, C., Feldman, B., Usman, R., Adiputra, R., Kusumaningrum, S., & Stark, L. (2016). Using the three delays model to examine civil registration barriers in Indonesia. *PLoS One*, *11*(12), 1–15. https://doi.org/10.1371/journal.pone.0168405

Bequele, A. (2005). *Universal birth registration: The challenge in Africa*. https://static1.squarespace.com/static/5519047ce4b0d9aaa8c82e69/t/5551e3dfe4b0ac09774e6170/1431430111352/universal_birth_registration.pdf

Bernhardt, V. (2013). *Data analysis for continuous school improvement*. Routledge.

Bhabha, J. (2011). From citizen to migrant: The scope of child statelessness in the twenty-first century. In J. Bhabha (Ed.), *Children without a state: A global human rights challenge* (pp. 1–42). MIT University Press.

Bhatia, A., Ferreira, L. Z., Barros, A. J. D., & Victora, C. G. (2017). Who and where are the uncounted children? Inequalities in birth certificate coverage among children under five years in 94 countries using nationally representative household surveys. *International Journal for Equity in Health*, *16*(1), 1–11. https://doi.org/10.1186/s12939-017-0635-6

Bhatia, A., Krieger, N., Beckfield, J., Barros, A. J. D., & Victora, C. (2019). Are inequities decreasing? Birth registration for children under five in low-income and middle-income countries, 1999–2016. *BMJ Global Health*, *4*(6), 1–13. https://doi.org/10.1136/bmjgh-2019-001926

Bleck, J. (2013). Do Francophone and Islamic schooling communities participate differently? Disaggregating parents' political behaviour in Mali. *The Journal of Modern African Studies*, *51*(3), 377–408.

Board, C. (1984). Higher order map-using tasks: Geographical lessons in danger of being forgotten. *Cartographica: The International Journal for Geographic Information*

and Geovisualization, *21*(1), 85–97. https://doi.org/10.3138/0348-1643-2553-4578

Bolstad, P. (2016). *GIS fundamentals: A first text on geographic information systems* (5th ed.). Eider (PressMinnesota).

Boman, Y. (2006). The struggle between conflicting beliefs: On the promise of education. *Journal of Curriculum Studies*, *38*(5), 545–568. https://doi.org/10.1080/00220270600670783

Brewer, M., Menzies, N., & Schott, J. (2015). *Identification systems don't always serve the bottom 40%.* The World Bank. http://www.governancenow.com/files/947040BRI00PUB0opment0Issue080FINAL.pdf

Brito, S., Corbacho, A., & Osorio, R. (2017). Does birth under-registration reduce childhood immunization? Evidence from the Dominican Republic. *Health Economics Review*, *7*(1), 1–18. https://doi.org/10.1186/s13561-017-0149-3

Brock, C. (2011). *Education as a global concern*. Bloomsbury Publishing.

Brown, G., Strickland-Munro, J., Kobryn, H., & Moore, S. A. (2017). Mixed methods participatory GIS: An evaluation of the validity of qualitative and quantitative mapping methods. *Applied Geography*, *79*, 153–166. https://doi.org/10.1016/j.apgeog.2016.12.015

Buckland, S. (2011). From policy to practice: The challenges to educational access for non-nationals in South Africa. *International Journal of Educational Development*, *31*(4), 367–373. https://doi.org/10.1016/j.ijedudev.2011.01.005

Burde, D., Kapit, A., Wahl, R. L., Guven, O., & Skarpeteig, M. I. (2017). Education in emergencies: A review of theory and research. *Review of Educational Research*, *87*(3), 619–658. https://doi.org/10.3102/0034654316671594

Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N., & Zhang, Y. J. (2017). Curriculum and ideology. *Journal of Political Economy*, *125*(2), 338–392. https://doi.org/10.1086/690951

Cappa, C., Gregson, K., Wardlaw, T., & Bissell, S. (2014). Birth registration: A child's passport to protection. *The Lancet Global Health*, *2*(2), e67–e68. https://doi.org/10.1016/S2214-109X(13)70180-3

Caracelli, V. J., & Riggin, L. J. C. (1994). Mixed-method evaluation: Developing quality criteria through concept mapping: Mixed-Method collaboration. *Evaluation Practice*, *15*(2), 139–152. https://doi.org/10.1016/0886-1633(94)90005-1

Checchi, D. (2003). Inequality in incomes and access to education: A cross-country analysis (1960–95). *LABOUR*, *17*(2), 153–201. https://doi.org/10.1111/1467-9914.00226

Chereni, A. (2017). Researching the Dynamics of Birth Registration and Social Exclusion for Child Rights Advocacy: The Unique Role of Qualitative Research. *Forum: Qualitative Social Research*, *18*(1), 1–28.

Chiang, M. M.-T., & Mirkin, B. (2010). Intelligent choice of the number of clusters in k -means clustering: An experimental study with different cluster spreads. *Journal of Classification*, *27*(1), 3–40. https://doi.org/10.1007/s00357-010-9049-5

Chimombo, J. P. G. (2005). Issues in basic education in developing countries: An exploration of policy options for improved delivery. *Journal of International Cooperation in Education, 8*(1), 129–152.

Civil Registration Centre for Development. (2011). *Civil registration support in Cameroon* (pp. 1–77). UNICEF. https://www.unicef.org/evaldatabase/files/2011_Cameroon_-_Civil_Registration_Support_in_Cameroon.final.pdf

Comandini, O., Cabras, S., & Marini, E. (2016). Birth registration and child undernutrition in sub-Saharan Africa. *Public Health Nutrition, 19*(10), 1757–1767. https://doi.org/10.1017/S136898001500333X

Cope, M., & Elwood, S. (Eds.). (2009). *Qualitative GIS: A mixed methods approach*. SAGE Publications.

Corbacho, A., Brito, S., & Rivas, R. O. (2012). *Birth Registration and the impact on educational attainment* (IDB Working Paper Series No. IDB-WP-345). Inter-American Development Bank. https://publications.iadb.org/publications/english/document/Birth-Registration-and-the-Impact-on-Educational-Attainment.pdf

Cox, B. E., McIntosh, K. L., Reason, R. D., & Terenzini, P. T. (2014). Working with missing data in higher education research: A primer in real-world example. *The Review of Higher Education, 37*(3), 377–402.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). SAGE Publications.

Creswell, J. W. (2010). Mapping the developing landscape of mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 45–68). SAGE Publications.

Creswell, J. W., & Plano Clark, V. L. (2006). *Designing and conducting mixed methods research*. SAGE Publications.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Croke, K., Grossman, G., Larreguy, H. A., & Marshall, J. (2016). Deliberate disengagement: How education can decrease political participation in electoral authoritarian regimes. *American Political Science Review, 110*(3), 579–600. https://doi.org/10.1017/S0003055416000253

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. SAGE Publications.

Crush, J., & Tawodzera, G. (2014). Exclusion and discrimination: Zimbabwean migrant children and South African schools. *Journal of International Migration and Integration, 15*(4), 677–693. https://doi.org/10.1007/s12134-013-0283-7

Dahan, M., & Gelb, A. (2015a). *The identity target in the post-2015 development agenda* (Note 19). World Bank Group. http://pubdocs.worldbank.org/en/553511442506682828/1603399-TransportICT-Newsletter-Note-19.pdf

Dahan, M., & Gelb, A. (2015). *The criterion problem: Measuring the legal identity target in the post-2015 agenda*. World Bank Blogs. https://blogs.worldbank.org/digital-development/criterion-problem-measuring-legal-identity-target-post-2015-agenda

Dake, F. A. A., & Fuseini, K. (2018). Registered or unregistered? Levels and differentials in registration and certification of births in Ghana. *BMC International Health and Human Rights*, *18*(1), 1–25. https://doi.org/10.1186/s12914-018-0163-5

Davies, L. (2003). *Education and conflict: Complexity and chaos*. Routledge.

De Waal, A. (2015). *The real politics of the Horn of Africa: Money, war and the business of power*. Polity Press.

Dellinger, A. B., & Leech, N. L. (2007). Toward a unified validation framework in mixed methods research. *Journal of Mixed Methods Research*, *1*(4), 309–332. https://doi.org/10.1177/1558689807306147

Dickinson, W. D. (2010). Visual displays for mixed methods findings. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 469–504). SAGE Publications.

Dixon, D. P., & Jones, J. P. (1998). My dinner with Derrida, or spatial analysis and poststructuralism do lunch. *Environment and Planning A: Economy and Space*, *30*(2), 247–260. https://doi.org/10.1068/a300247

Donou-Adonsou, F. (2019). Technology, education, and economic growth in Sub-Saharan Africa. *Telecommunications Policy*, *43*(4), 353–360. https://doi.org/10.1016/j.telpol.2018.08.005

Douzinas, C. (2002). Identity, recognition, rights or what can Hegel teach us about human rights? *Journal of Law and Society*, *29*(3), 379–405. https://doi.org/10.1111/1467-6478.00225

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives*, *38*(1), 105–124.

Dryden-Peterson, S. (2009). *Barriers to accessing primary education in conflict-affected fragile states*. Save the Children. https://wcfia.harvard.edu/files/wcfia/files/2942.pdf

Duff, P., Kusumaningrum, S., & Stark, L. (2016). Barriers to birth registration in Indonesia. *The Lancet Global Health*, *4*(4), e234–e235. https://doi.org/10.1016/S2214-109X(15)00321-6

Dunn, C. E. (2007). Participatory GIS — a people's GIS? *Progress in Human Geography*, *31*(5), 616–637. https://doi.org/10.1177/0309132507081493

Dunning, C., Gelb, A., & Raghavan, S. (2014). *Birth registration, legal identity, and the post-2015 agenda* (CGD Policy Paper 046). Center for Global Development. https://www.files.ethz.ch/isn/183719/birth-registration-legal-identity.pdf

Dupraz, Y. (2019). French and British colonial legacies in education: Evidence from the partition of Cameroon. *The Journal of Economic History*, *79*(3), 628–668. https://doi.org/10.1017/S0022050719000299

Eghtesadifard, M., Afkhami, P., & Bazyar, A. (2020). An integrated approach to the selection of municipal solid waste landfills through GIS, K-Means and multi-criteria decision analysis. *Environmental Research*, *185*, 1–16.

Eizenga, D. (2019). *Long term trends across security and development in the Sahel* (West African Papers No. 25). Sahel Research Group, University of Florida. https://doi.org/10.1787/630477ee-en

Elbakidze, L., & Jin, Y. H. (2015). Are economic development and education improvement associated with participation in transnational terrorism? *Risk Analysis*, *35*(8), 1520–1535. https://doi.org/10.1111/risa.12378

Elwood, S. (2006). Critical issues in participatory GIS: Deconstructions, Reconstructions, and new research directions. *Transactions in GIS*, *10*(5), 693–708. https://doi.org/10.1111/j.1467-9671.2006.01023.x

Esri. (n.da). *How multivariate clustering works*. https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-multivariate-clustering-works.htm

Esri. (n.db). *How spatial autocorrelation (Global Moran's I) works*. https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm

Evans, B. C., Coon, D. W., & Ume, E. (2011). Use of theoretical frameworks as a pragmatic guide for mixed methods studies: A methodological necessity? *Journal of Mixed Methods Research*, *5*(4), 276–292. https://doi.org/10.1177/1558689811412972

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health Services Research*, *48*(6), 2134–2156. https://doi.org/10.1111/1475-6773.12117

Fetters, M. D., & Freshwater, D. (2015). Publishing a methodological mixed methods research article. *Journal of Mixed Methods Research*, *9*(3), 203–213. https://doi.org/10.1177/1558689815594687

Fielding, N. G. (2012). Triangulation and mixed methods designs: Data integration with new research technologies. *Journal of Mixed Methods Research*, *6*(2), 124–136. https://doi.org/10.1177/1558689812437101

Fisher, M. J., & Marshall, A. P. (2009). Understanding descriptive statistics. *Australian Critical Care*, *22*(2), 93–97. https://doi.org/10.1016/j.aucc.2008.11.003

Fredriksen, B., & Fossberg, C. H. (2014). The case for investing in secondary education in sub-Saharan Africa (SSA): Challenges and opportunities. *International Review of Education*, *60*(2), 235–259. https://doi.org/10.1007/s11159-014-9407-3

Frels, J. G., Frels, R. K., & Onwuegbuzie, A. J. (2011). Geographic information systems: A mixed methods spatial approach in business and management research and beyond. *International Journal of Multiple Research Approaches*, *5*(3), 367–386. https://doi.org/10.5172/mra.2011.5.3.367

Galvan, J. L., & Galvan, M. C. (2017). *Writing literature reviews: A guide for students of the social and behavioral sciences* (7th ed.). Routledge.

Gandhi, L. (1998). *Postcolonial theory: A critical introduction*. Columbia University Press.

Gerber, P., Gargett, A., & Castan, M. (2011). Does the right to birth registration include a right to a birth certificate? *Netherlands Quarterly of Human Rights*, *29*(4), 434–459. https://doi.org/10.1177/016934411102900403

Getis, A. (2007). Reflections on spatial autocorrelation. *Regional Science and Urban Economics*, *37*(4), 491–496. https://doi.org/10.1016/j.regsciurbeco.2007.04.005

Getis, A. (2008). A history of the concept of spatial autocorrelation: A geographer's perspective. *Geographical Analysis*, *40*(3), 297–309. https://doi.org/10.1111/j.1538-4632.2008.00727.x

Giordano, A., & Cole, T. (2018). The limits of GIS: Towards a GIS of place. *Transactions in GIS*, *22*(3), 664–676. https://doi.org/10.1111/tgis.12342

Gourevitch, P. (2008). The role of politics in economic development. *Annual Review of Political Science*, *11*(1), 137–159. https://doi.org/10.1146/annurev.polisci.11.053006.185507

Gray, N., Azzopardi, P., Kennedy, E., Willersdorf, E., & Creati, M. (2013). Improving adolescent reproductive health in Asia and the Pacific: Do we have the data? A review of DHS and MICS surveys in nine countries. *Asia Pacific Journal of Public Health*, *25*(2), 134–144. https://doi.org/10.1177/1010539511417423

Greene, J. C. (2007). *Mixed methods in social inquiry*. Jossey-Bass.

Greene, J. C., & Hall, J. N. (2010). Dialectics and Pragmatism: Being of Consequence. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 119–144). SAGE Publications.

Greenhill, B. (2008). Recognition and collective identity formation in international politics. *European Journal of International Relations*, *14*(2), 343–368. https://doi.org/10.1177/1354066108089246

Gregory, I., Donaldson, C., Murrieta-Flores, P., & Rayson, P. (2015). Geoparsing, GIS, and textual analysis: Current developments in spatial humanities research. *International Journal of Humanities and Arts Computing*, *9*(1), 1–14. https://doi.org/10.3366/ijhac.2015.0135

Gregory, I. N., & Healey, R. G. (2007). Historical GIS: Structuring, mapping and analysing geographies of the past. *Progress in Human Geography*, *31*(5), 638–653. https://doi.org/10.1177/0309132507081495

Griffin, T.E. (2021a). *Measuring the relationship between birth registration and educational access in Sub-Saharan Africa* [Unpublished doctoral dissertation, Article #1]. University of Vermont.

Griffin, T.E. (2021b). *Examining the geographies of educational access and birth registration in Sub-Saharan Africa* [Unpublished doctoral dissertation, Article #2]. University of Vermont.

Hall, R. (2013). Mixed methods: In search of a paradigm. In T. Le & Q. Le (Eds.), *Conducting Research in a Changing and Challenging World* (pp. 71–78). Nova Science Publishers.

Hallpike, C. R. (1973). Functionalist interpretations of primitive warfare. *Man*, *8*(3), 451–470. https://doi.org/10.2307/2800321

Hamadi, L. (2014). Edward Said: The postcolonial theory and the literature of decolonization. *European Scientific Journal*, 39–46.

Hancioglu, A., & Arnold, F. (2013). Measuring coverage in MNCH: Tracking progress in health for women and children using DHS and MICS household surveys. *PLoS Medicine*, *10*(5), 1–8. https://doi.org/10.1371/journal.pmed.1001391

Hanushek, E. A., & Wößmann, L. (2010). Education and economic growth. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 2, pp. 245–252). Elsevier.

Harbers, I., & Ingram, M. C. (2017). Geo-nested analysis: Mixed-methods research with spatially dependent data. *Political Analysis*, *25*(3), 289–307. https://doi.org/10.1017/pan.2017.4

Harbitz, M., & Boekle-Giuffrida, B. (2009). *Democratic governance, citizenship, and legal identity: linking theoretical discussion and operational reality* [Working Paper]. Inter-American Development Bank. https://publications.iadb.org/publications/english/document/Democratic-Governance-Citizenship-and-Legal-Identity-Linking-Theoretical-Discussion-and-Operational-Reality.pdf

Harbitz, M., & Tamargo, M. del C. (2009). *The significance of legal identity in situations of poverty and social exclusion: The link between gender, ethnicity, and legal identity* [Technical Note]. Inter-American Development Bank. https://publications.iadb.org/publications/english/document/The-Significance-of-Legal-Identity-in-Situations-of-Poverty-and-Social-Exclusion-The-Link-between-Gender-Ethnicity-and-Legal-Identity.pdf

Hegel, G. W. F. (1910). *The phenomenology of mind* (J.B. Baillie, Trans.). The Macmillan Company. (Original work published 1807)

Hegel, G. W. F. (1977). *Phenomenology of Spirit* (A.V. Miller, Trans.). Oxford University Press. (Original work published 1807)

Hegre, H., Karlsen, J., Nygård, H. M., Strand, H., & Urdal, H. (2013). Predicting armed conflict, 2010–2050. *International Studies Quarterly*, *57*(2), 250–270. https://doi.org/10.1111/isqu.12007

Herbst, J. (1989). The creation and matintenance of national boundaries in Africa. *International Organization*, *43*(4), 673–692. https://doi.org/10.1017/S0020818300034482

Hill, K., Lopez, A. D., Shibuya, K., & Jha, P. (2007). Interim measures for meeting needs for health sector data: Births, deaths, and causes of death. *The Lancet*, *370*(9600), 1726–1735. https://doi.org/10.1016/S0140-6736(07)61309-9

Hlalele, D. (2012). Social justice and rural education in South Africa. *Perspectives in Education*, *30*(1), 111–118.

Hlatshwayo, M., & Vally, S. (2014). Violence, resilience and solidarity: The right to education for child migrants in South Africa. *School Psychology International*, *35*(3), 266–279. https://doi.org/10.1177/0143034313511004

Honneth, A. (1996). *The struggle for recognition: The moral grammar of social conflicts* (J. Anderson, Trans.). MIT University Press; Polity Press.

Hopkins, N., & Blackwood, L. (2011). Everyday citizenship: Identity and recognition. *Journal of Community & Applied Social Psychology*, *21*(3), 215–227. https://doi.org/10.1002/casp.1088

Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Wadsworth, Cengage Learning.

Huillery, E. (2009). History matters: The long-term impact of colonial public investments in French West Africa. *American Economic Journal: Applied Economics*, *1*(2), 176–215. https://doi.org/10.1257/app.1.2.176

Human Rights Watch. (2017). *"I had a dream to finish school": Barriers to secondary education in Tanzania*. Human Rights Watch. https://www.hrw.org/report/2017/02/14/i-had-dream-finish-school/barriers-secondary-education-tanzania

Hunter, W. (2019). Identity documents, welfare enhancement, and group empowerment in the Global South. *The Journal of Development Studies*, *55*(3), 366–383. https://doi.org/10.1080/00220388.2018.1451637

Hunter, W., & Brill, R. (2016). "Documents, please": Advances in social protection and birth certification in the developing world. *World Politics*, *68*(2), 191–228. https://doi.org/10.1017/S0043887115000465

Hunter, W., & Sugiyama, N. B. (2018). Making the newest citizens: Achieving universal birth registration in Contemporary Brazil. *The Journal of Development Studies*, *54*(3), 397–412. https://doi.org/10.1080/00220388.2017.1316378

Ilon, L., & Lee, J. (2014). Changing views of inequality and implications for leadership and learning. In A. H. Normore & J. S. Brooks (Eds.), *Educational leadership for ethics and social justice: Views from the social sciences* (pp. 113–134). Information Age Publishing.

Ingram, M. C., & Harbers, I. (2020). Spatial tools for case selection: Using LISA statistics to design mixed-methods research. *Political Science Research and Methods*, *8*, 747–763.

Ingutia, R., Rezitis, A. N., & Sumelius, J. (2020). Child poverty, status of rural women and education in sub Saharan Africa. *Children and Youth Services Review*, *111*, 1–8.

International Institute for Vital Registration and Statistics. (1988). *Organization and status of civil registration in Africa and recommendations for improvement* (Technical Papers: Number 31). International Institute for Vital Registration and Statistics. https://www.cdc.gov/nchs/data/isp/031_Organization_and_Status_of_Civil_Regist_in_Africa_and_Recommendation_for_Improvement.pdf

Issa, I. H., & Mgaya, N. R. (2018). Factors for low birth registration in Mbeya District: A case of Utengule Usongwe Ward in Mbeya District. *International Journal of Innovative Science and Research Technology*, *3*(6), 292–296.

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods*, *18*(1), 3–20. https://doi.org/10.1177/1525822X05282260

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666.

Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PLOS ONE*, *15*(2), 1–15. https://doi.org/10.1371/journal.pone.0229345

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548. https://doi.org/10.1162/qjec.2010.125.2.515

Jensen, S. Q. (2011). Othering, identity formation and agency. *Qualitative Studies*, *2*(2), 63–78. https://doi.org/10.7146/qs.v2i2.5510

Jeong, J., Bhatia, A., & Fink, G. (2018). Associations between birth registration and early child growth and development: Evidence from 31 low- and middle-income countries. *BMC Public Health*, *18*(673), 1–8. https://doi.org/10.1186/s12889-018-5598-z

Jewkes, R., & Wood, K. (1998). Competing discourses of vital registration and personhood: Perspectives from rural South Africa. *Social Science & Medicine, 46*(8), 1043–1056. https://doi.org/10.1016/S0277-9536(97)10036-3

Johnson, B. (2001). Toward a new classification of nonexperimental quantitative research. *Educational Researcher, 30*(2), 3–13. https://doi.org/10.3102/0013189X030002003

Johnson, R. B. (2017). Dialectical pluralism: A metaparadigm whose time has come. *Journal of Mixed Methods Research, 11*(2), 156–173. https://doi.org/10.1177/1558689815607692

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26. https://doi.org/10.3102/0013189X033007014

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112–133. https://doi.org/10.1177/1558689806298224

Johnston, M. P. (2017). Secondary data analysis: A method of which the time has come. *Qualitative and Quantitative Methods in Libraries, 3*(3), 619–626.

Johnston, R. J., Gregory, D., Pratt, G., & Watts, M. (Eds.). (2000). *The dictionary of human geography* (4th ed.). Wiley-Blackwell.

Joint Chiefs of Staff. (2018). *Civil-military operations (JP 3-57)*. United States Department of Defense. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_57.pdf

Jones, K. (2017). Using a theory of practice to clarify epistemological challenges in mixed methods research: An example of theorizing, modeling, and mapping changing west African seed systems. *Journal of Mixed Methods Research, 11*(3), 355–373. https://doi.org/10.1177/1558689815614960

Jung, J.-K., & Elwood, S. (2010). Extending the qualitative capabilities of GIS: Computer-aided qualitative GIS. *Transactions in GIS, 14*(1), 63–87. https://doi.org/10.1111/j.1467-9671.2009.01182.x

Justino, P. (2014). *Barriers to education in conflict-affected countries and policy opportunities*. UNESCO Institute for Statistics. https://www.allinschool.org/media/1941/file/Paper-OOSCI-Barriers-to-Education-in-Conflict-Affected-Countries-2014-en.pdf

Kazeem, A., Jensen, L., & Stokes, C. S. (2010). School attendance in Nigeria: Understanding the impact and intersection of gender, urban-rural residence, and socioeconomic status. *Comparative Education Review, 54*(2), 295–319.

Khan, S., & Hancioglu, A. (2019). Multiple indicator cluster surveys: Delivering robust data on children and women across the globe. *Studies in Family Planning, 50*(3), 279–286. https://doi.org/10.1111/sifp.12103

Kimmons, S. (2017, May 31). *Using education as a weapon, Soldiers help Cameroon deter Boko Haram*. United States Army. https://www.army.mil/article/186462/using_education_as_a_weapon_soldiers_help_cameroon_deter_boko_haram

Kindzeka, M. E. (2017, November 15). *Cameroon, UN tackle huge birth registration backlog in North*. Voice of America. https://www.voanews.com/africa/cameroon-un-tackle-huge-birth-registration-backlog-north

King, E. (2011). The Multiple Relationships between Education and Conflict: Reflections of Rwandan Teachers and Students. In K. Mundy & S. Dryden-Peterson (Eds.), *Educating children in conflict zones: A tribute to Jackie Kirk* (pp. 137–151). Teachers College Press.

Kohn, M. L. (1987). Cross-national research as an analytic strategy: American Sociological Association, 1987 Presidential Address. *American Sociological Review*, *52*(6), 713–731. https://doi.org/10.2307/2095831

Koivu, K. L., & Hinze, A. M. (2017). Cases of convenience? The divergence of theory from practice in case selection in qualitative and mixed-methods research. *PS: Political Science & Politics*, *50*(04), 1023–1027. https://doi.org/10.1017/S1049096517001214

Kosack, S. (2009). Realising education for all: Defining and using the political will to invest in primary education. *Comparative Education*, *45*(4), 495–523. https://doi.org/10.1080/03050060903391586

Kuépié, M., Shapiro, D., & Tenikue, M. (2015). Access to schooling and staying in school in selected Sub-Saharan African countries. *African Development Review*, *27*(4), 403–414. https://doi.org/10.1111/1467-8268.12156

Ladner, D., Jensen, E. G., & Saunders, S. E. (2014). A critical assessment of legal identity: What it promises and what it delivers. *Hague Journal on the Rule of Law*, *6*(1), 47–74. https://doi.org/10.1017/S1876404513000043

Lam, T. D. T. (2001, December 28). *POPULATION-VIETNAM: Govt steps up birth certificate campaign.* Inter Press Service News Agency. http://www.ipsnews.net/2001/12/population-vietnam-govt-steps-up-birth-certificate-campaign/

Lange, M. (2012). *Educations in ethnic violence: identity, educational bubbles, and resource mobilization*. Cambridge University Press.

Larsen, D. A., Grisham, T., Slawsky, E., & Narine, L. (2017). An individual-level meta-analysis assessing the impact of community-level sanitation access on child stunting, anemia, and diarrhea: Evidence from DHS and MICS surveys. *PLOS Neglected Tropical Diseases*, *11*(6), 1–13. https://doi.org/10.1371/journal.pntd.0005591

Larson, M. G. (2006). Descriptive statistics and graphical displays. *Circulation*, *114*(1), 76–81. https://doi.org/10.1161/CIRCULATIONAHA.105.584474

Law, M., & Collins, A. (2015). *Getting to know ArcGIS* (4th ed). Esri Press.

Legendre, P. (1993). Spatial autocorrelation: Trouble or new paradigm? *Ecology, 74*(6), 1659–1673. https://doi.org/10.2307/1939924

Lewin, K. M. (2009). Access to education in sub-Saharan Africa: Patterns, problems and possibilities. *Comparative Education*, *45*(2), 151–174. https://doi.org/10.1080/03050060902920518

Lewin, K. M., & Sabates, R. (2011). *Changing patterns of access to education in Anglophone and Francophone countries in Sub Saharan Africa: Is education for all pro-poor?* (CREATE Pathways to Access Research Monograph No. 52).

University of Sussex Centre for International Education. https://files.eric.ed.gov/fulltext/ED519555.pdf
Li, S., Zhang, Y., & Feldman, M. W. (2010). Birth registration in China: Practices, problems and policies. *Population Research and Policy Review*, *29*(3), 297–317. https://doi.org/10.1007/s11113-009-9141-x
Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, *53*(01), 69–105. https://doi.org/10.2307/1951731
Little, A. W., & Lewin, K. M. (2011). The policies, politics and progress of access to basic education. *Journal of Education Policy*, *26*(4), 477–482. https://doi.org/10.1080/02680939.2011.555004
Lloyd, C. B., Mete, C., & Sathar, Z. A. (2005). The effect of gender differences in primary school access, type, and quality on the decision to enroll in rural Pakistan. *Economic Development and Cultural Change*, *53*(3), 685–710. https://doi.org/10.1086/427042
Long, J., Nelson, T., & Wulder, M. (2010). Regionalization of landscape pattern indices using multivariate cluster analysis. *Enivornmental Management*, *46*(1), 134–142.
Luke, A. (2011). Generalizing across borders: Policy and the limits of educational science. *Educational Researcher*, *40*(8), 367–377. https://doi.org/10.3102/0013189X11424314
Mackenzie, Y. (2008). The campaign for universal birth registration in Latin America: Ensuring all Latin American children's inherent right to life and survival by first guaranteeing their right to a legal identity. *Georgia Journal of International and Comparative Law*, *37*, 519–554.
Mahapatra, P., Shibuya, K., Lopez, A. D., Coullare, F., Notzon, F. C., Rao, C., & Szreter, S. (2007). Civil registration systems and vital statistics: Successes and missed opportunities. *The Lancet*, *370*(9599), 1653–1663. https://doi.org/10.1016/S0140-6736(07)61308-7
Mahmoud, M. A., Saad, A. E. N., & Shaer, R. E. (2015). Phase II multiple linear regression profile with small sample sizes. *Quality and Reliability Engineering International*, *31*(5), 851–861. https://doi.org/10.1002/qre.1644
Majgaard, K., & Mingat, A. (2012). *Education in Sub-Saharan Africa: A comparative analysis* (World Bank Study 70603). The World Bank. http://documents1.worldbank.org/curated/en/892631468003571777/pdf/709790PUB0EPI0067926B09780821388891.pdf
Makinde, O. A., Olapeju, B., Ogbuoji, O., & Babalola, S. (2016). Trends in the completeness of birth registration in Nigeria: 2002–2010. *Demographic Research*, *35*(12), 315–338.
Marquinez, J., Lastra, J., & Garcia, P. (2003). Estimation models for precipitation in mountainous regions: The use of GIS and multivariate analysis. *Journal of Hydrology*, *270*, 1–11.
Marshall, M. G., & Cole, B. R. (2014). *The problem of "state failure" and the complex societal-systems approach*. Center for Systemic Peace. http://www.systemicpeace.org/assets/handbookofpoliticalconflictmarshallandcole5nov13.pdf

Maxwell, J. A. (2011). Paradigms or toolkits? Philosophical and methodological positions as heuristics for mixed methods research. *Mid-Western Educational Researcher*, *24*(2), 27–30.
Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*. SAGE Publications.
Maxwell, J. A. (2016). Expanding the history and range of mixed methods research. *Journal of Mixed Methods Research*, *10*(1), 12–27. https://doi.org/10.1177/1558689815571132
Mbulawa, S., & Mehta, P. (2016). The effect of access and quality of education on economic development in Botswana. *International Journal of Business, Economics and Management*, *3*(11), 144–159.
McClintock, N., Mahmoudi, D., Simpson, M., & Santos, J. P. (2016). Socio-spatial differentiation in the Sustainable City: A mixed-methods assessment of residential gardens in metropolitan Portland, Oregon, USA. *Landscape and Urban Planning*, *148*, 1–16. https://doi.org/10.1016/j.landurbplan.2015.12.008
McGrew Jr., J. C., Lembo Jr., A. J., & Monroe, C. B. (2014). *An introduction to statistical problem solving in geography* (3rd ed.). Waveland Press.
McKim, C. A. (2017). The value of mixed methods research: A mixed methods study. *Journal of Mixed Methods Research*, *11*(2), 202–222. https://doi.org/10.1177/1558689815607096
McShane, B. B., Gal, D., Gelman, A., Robert, C., & Tackett, J. L. (2019). Abandon statistical significance. *The American Statistician*, *73*(51), 235–245.
Melnik, T. A., Guldal, C. G., Schoen, L. D., Alicandro, J., & Henfield, P. (2015). Barriers in accurate and complete birth registration in New York State. *Maternal and Child Health Journal*, *19*(9), 1943–1948. https://doi.org/10.1007/s10995-015-1711-4
Mertens, D. M., Bledsoe, K. L., Sullivan, M., & Wilson, A. (2010). Utilization of mixed methods for transformative purposes. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 193–214). SAGE Publications.
Miguel, E. (2004). Tribe or nation? Nation building and public goods in Kenya versus Tanzania. *World Politics*, *56*(3), 327–362.
Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., de Savigny, D., Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: Monitoring data quality and progress. *The Lancet*, *386*(10001), 1395–1406. https://doi.org/10.1016/S0140-6736(15)60171-4
Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). SAGE Publications.
Miles, W. F. S. (2008). Legacies of Anglo-French colonial borders: A West African and Southeast Asian comparison. *Journal of Borderlands Studies*, *23*(2), 83–102. https://doi.org/10.1080/08865655.2008.9695699
Miles, W. F. S. (2014). *Scars of partition: Postcolonial legacies in French and British borderlands*. University of Nebraska Press.

Mills, S., Abouzahr, C., Kim, J., Rassekh, B. M., & Sarpong, D. (2017). *Civil registration and vital statistics for monitoring the sustainable development goals*. The World Bank. https://doi.org/10.1596/27533

Mitchell, R., Rose, P., & Asare, S. (2020). Education research in Sub-Saharan Africa: Quality, visibility, and agendas. *Comparative Education Review, 64*(3), 363–383.

Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological), 10*(2), 243–251.

Morrow, W. E. (1978). *Chains of thought: Philosophical essays in South African education*. Southern Book Publishers.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research, 40*(2), 120–123.

Müller, T. R. (2020). Colonial borders and hybrid identities: Lessons from the case of Eritrea. *Borderlands, 19*(1), 147–173.

Murray, C., & Newby, H. (2012). Data resource profile: United Nations Children's Fund (UNICEF). *International Journal of Epidemiology, 41*(6), 1595–1601. https://doi.org/10.1093/ije/dys185

Nastasi, B. K., Hitchcock, J. H., & Brown, L. M. (2010). An inclusive framework for conceptualizing mixed methods design typologies: Moving toward fully integrated synergistic research models. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 305–338). SAGE Publications.

National Center for Education Statistics. (n.d.). *Education indicators: An international perspective*. https://nces.ed.gov/Pubs/eiip/eiip1s01.asp

Nettle, D. (1996). Language diversity in West Africa: An ecological approach. *Journal of Anthropological Archaeology, 15*(4), 403–438. https://doi.org/10.1006/jaar.1996.0015

Nyerges, T. L. (1991). Analytical map use. *Cartography and Geographic Information Systems, 18*(1), 11–22. https://doi.org/10.1559/152304091783805635

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 531–556). SAGE Publications.

O'Cathain, A., Murphy, E., & Nicholl, J. (2007). Integration and publications as indicators of "yield" from mixed methods studies. *Journal of Mixed Methods Research, 1*(2), 147–163. https://doi.org/10.1177/1558689806299094

Office of the United Nations High Commissioner for Human Rights. (2014). *Birth registration and the right of everyone to recognition everywhere as a person before the law* (Annual Report A/HRC/27/22). United Nations. https://www.refworld.org/docid/53ff324e4.html

Oketch, M. O. (2006). Determinants of human capital formation and economic growth of African countries. *Economics of Education Review, 25*(1), 554–564.

Oketch, M., & Rolleston, C. (2007). Policies on free primary and secondary education in East Africa: Retrospect and prospect. *Review of Research in Education, 31*(1), 131–158. https://doi.org/10.3102/0091732X07300046131

Oliker, V. I. (1978). On the relationship between the sample size and the number of variables in a linear regression model. *Communications in Statistics - Theory and Methods, 7*(6), 509–516. https://doi.org/10.1080/03610927808827645

O'loughlin, J., & Anselin, L. (1991). Bringing geography back to the study of international relations: Spatial dependence and regional context in Africa, 1966–1978. *International Interactions*, *17*(1), 29–61. https://doi.org/10.1080/03050629108434769

Onwuegbuzie, A. J., & Frels, R. K. (2013). Introduction: Toward a new research philosophy for addressing social justice issues: Critical dialectical pluralism 1.0. *International Journal of Multiple Research Approaches*, *7*(1), 9–26. https://doi.org/10.5172/mra.2013.7.1.9

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, *13*(1), 48–63.

Oppenheim, B., & Powell, B. M. (2015). *Legal identity in the 2030 Agenda for Sustainable Development: Lessons from Kibera, Kenya* [Policy Paper]. Open Society Justice Initiative. https://www.justiceinitiative.org/uploads/0a6472de-a975-4a3b-b3ad-2b979891d645/legal-identity-2030-agenda-lessons-kibera-kenya-2051216.pdf

Organisation for Economic Co-operation and Development. (2014). *Society at a glance: Asia/Pacific 2014*. OECD. https://www.oecd-ilibrary.org/social-issues-migration-health/society-at-a-glance-asia-pacific-2014_9789264220553-en

Østby, G., Nordås, R., & Rød, J. K. (2009). Regional inequalities and civil conflict in Sub-Saharan Africa. *International Studies Quarterly*, *53*(2), 301–324. https://doi.org/10.1111/j.1468-2478.2009.00535.x

Østby, G., Urdal, H., & Dupuy, K. (2019). Does education lead to pacification? A systematic review of statistical studies on education and political violence. *Review of Educational Research*, *89*(1), 46–92. https://doi.org/10.3102/0034654318800236

Owoo, N., & Naudé, W. (2014). *Non-farm enterprise productivity and spatial autocorrelation in rural Africa: Evidence from Ethiopia and Nigeria* (IZA Discussion Paper No. 8295). Institute for the Study of Labor (IZA). https://papers.ssrn.com/abstract=2465330

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM statistics* (7th ed.). Open University Press.

Pan, M. L. (2016). *Preparing literature reviews: Qualitative and quantitative approaches* (5th ed.). Routledge.

Parks, J. M. (1966). Cluster analysis applied to multivariate geologic problems. *The Journal of Geology*, *74*(5), 703–715.

Patton, G. C., Coffey, C., Cappa, C., Currie, D., Riley, L., Gore, F., Degenhardt, L., Richardson, D., Astone, N., Sangowawa, A. O., Mokdad, A., & Ferguson, J. (2012). Health of the world's adolescents: A synthesis of internationally comparable data. *The Lancet*, *379*(9826), 1665–1675. https://doi.org/10.1016/S0140-6736(12)60203-7

Pelleg, D., & Moore, A. W. (2000). X-means: Extending k-means with efficient estimation of the number of clusters. *ICML*, *1*, 727–734.

Pelowski, M., Wamai, R. G., Wangombe, J., Nyakundi, H., Oduwo, G. O., Ngugi, B. K., & Ogembo, J. G. (2016). How would children register their own births? Insights from a survey of students regarding birth registration knowledge and policy suggestions in Kenya. *PLOS ONE, 11*(3), e0149925. https://doi.org/10.1371/journal.pone.0149925

Phillips, D. E., AbouZahr, C., Lopez, A. D., Mikkelsen, L., de Savigny, D., Lozano, R., Wilmoth, J., & Setel, P. W. (2015). Are well functioning civil registration and vital statistics systems associated with better health outcomes? *The Lancet, 386*(10001), 1386–1394. https://doi.org/10.1016/S0140-6736(15)60172-6

Phillips, D. E., Adair, T., & Lopez, A. D. (2018). How useful are registered birth statistics for health and social policy? A global systematic assessment of the availability and quality of birth registration data. *Population Health Metrics*, *16*(21), 1–13. https://doi.org/10.1186/s12963-018-0180-6

Plano Clark, V. L., & Badiee, M. (2010). Research questions in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (2nd ed., pp. 275–304). SAGE Publications.

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies*, *47*(11), 1451–1458. https://doi.org/10.1016/j.ijnurstu.2010.06.004

Poon, J. P. H. (2005). Quantitative methods: Not positively positivist. *Progress in Human Geography*, *29*(6), 766–772. https://doi.org/10.1191/0309132505ph583pr

Porter, G. (2011). 'I think a woman who travels a lot is befriending other men and that's why she travels': Mobility constraints and their implications for rural women and girls in sub-Saharan Africa. *Gender, Place and Culture*, *18*(1), 65–81.

Porteus, K., Clacherty, G., Mdiya, L., Pelo, J., Matsai, K., Qwabe, S., & Donald, D. (2000). 'Out of school' children in South Africa: An analysis of causes in a group of marginalised, urban 7- to 15-year-olds. *Support for Learning*, *15*(1), 8–12. https://doi.org/10.1111/1467-9604.00135

Preston, B., & Wilson, M. W. (2014). Practicing GIS as mixed method: Affordances and limitations in an urban gardening study. *Annals of the Association of American Geographers*, *104*(3), 510–529. https://doi.org/10.1080/00045608.2014.892325

Ramutsindela, M. (2010). Africa's borders in colonial and postcolonial contexts. In P. Motzafi-Haller & M. Feige (Eds.), *HAGAR studies in culture, polity and identities* (Vol. 9, pp. 13–29). BGU Print Unit, Israel.

Roser, M., & Ortiz-Ospina, E. (2016). *Global education*. Our World in Data. https://ourworldindata.org/global-education#

Rostow, W. W. (1990). *The stages of economic growth: A non-communist manifesto*. Cambridge University Press.

Rucks-Ahidiana, Z., & Bierbaum, A. H. (2015). Qualitative spaces: Integrating spatial analysis for a mixed methods approach. *International Journal of Qualitative Methods*, *14*(2), 92–103. https://doi.org/10.1177/160940691501400208

Rudincova, K. (2015). When colonial borders still matter: The emergence of South Sudan. *Journal of African History, Politics and Society*, *1*(1), 89–115.

Said, E. (1978). *Orientalism*. Pantheon Books.

Saleh, M. (2016). Public mass modern education, religion, and human capital in twentieth-century Egypt. *The Journal of Economic History*, *76*(3), 697–735. https://doi.org/10.1017/S0022050716000796

Sanders, C., & Burnett, K. (2019). A case study in personal identification and social determinants of health: Unregistered births among indigenous people in Northern Ontario. *International Journal of Environmental Research and Public Health*, *16*(4), 567–576. https://doi.org/10.3390/ijerph16040567

Savigny, D. D., & Wijeyaratne, P. (1995). *GIS for health and the environment: Proceedings of an international workshop held in Colombo, Sri Lanka, 5-10 September 1994*. IDRC. https://idl-bnc-idrc.dspacedirect.org/handle/10625/14959

Schafft, K. A., & Jackson, A. Y. (2010). *Rural education for the twenty-first century: Identity, place, and community in a globalizing world*. Penn State Press.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, *51*(1), 1–17.

Seawright, J. (2016). The case for selecting cases that are deviant or extreme on the independent variable. *Sociological Methods & Research*, *45*(3), 493–525. https://doi.org/10.1177/0049124116643556

Selim, L. (2019, December 19). *What is birth registration and why does it matter?* UNICEF. https://www.unicef.org/stories/what-birth-registration-and-why-does-it-matter

Setel, P. W., Macfarlane, S. B., Szreter, S., Mikkelsen, L., Jha, P., Stout, S., & AbouZahr, C. (2007). A scandal of invisibility: Making everyone count by counting everyone. *The Lancet*, *370*(9598), 1569–1577. https://doi.org/10.1016/S0140-6736(07)61307-5

Shaban, M., Urban, B., El Saadi, A., & Faisal, M. (2010). Detection and mapping of water pollution variation in the Nile Delta using multivariate clustering and GIS techniques. *Journal of Environmental Management*, *91*(8), 1785–1793.

Shan, Y., Tysklind, M., Hao, F., Ouyang, W., Chen, S., & Lin, C. (2013). Identification of sources of heavy metals in agricultural soils using multivariate analysis and GIS. *Journal of Soils and Sediments*, *13*(4), 720–729.

Shannon, J. (2015). Rethinking food deserts using mixed-methods GIS. *Cityscape*, *17*(1), 85–96.

Shay, E., Combs, T. S., Findley, D., Kolosna, C., Madeley, M., & Salvesen, D. (2016). Identifying transportation disadvantage: Mixed-methods analysis combining GIS mapping with qualitative data. *Transport Policy*, *48*, 129–138. https://doi.org/10.1016/j.tranpol.2016.03.002

Shekar, M., Yazbeck, A., Hasan, R., & Bakilana, A. (2016). *Population and development in the Sahel: Policy choices to catalyze a demographic dividend* [Discussion Paper]. The World Bank. https://doi.org/10.1596/25293

Shizha, E. (2005). Reclaiming our memories: The education dilemma in postcolonial African school curricula. In A. A. Abdi & A. Cleghorn (Eds.), *Issues in African education: Sociological perspectives* (pp. 65–83). Palgrave Macmillan. https://doi.org/10.1057/9781403977199_4

Sinha, G., Smucker, T. A., Lovell, E. J., Velempini, K., Miller, S. A., Weiner, D., & Wangui, E. E. (2017). The pedagogical benefits of participatory GIS for

geographic education. *Journal of Geography*, *116*(4), 165–179. https://doi.org/10.1080/00221341.2016.1215488
Sissoko, K., Keulen, H. van, Verhagen, J., Tekken, V., & Battaglini, A. (2011). Agriculture, livelihoods and climate change in the West African Sahel. *Regional Environmental Change*, *11*(1), 119–125. https://doi.org/10.1007/s10113-010-0164-y
Skovdal, M., & Andreouli, E. (2011). Using identity and recognition as a framework to understand and promote the resilience of caregiving children in Western Kenya. *Journal of Social Policy*, *40*(3), 613–630. https://doi.org/10.1017/S0047279410000693
Small, M. L. (2011). How to conduct a mixed methods study: Recent trends in a rapidly growing literature. *Annual Review of Sociology*, *37*(1), 57–86. https://doi.org/10.1146/annurev.soc.012809.102657
Smith, A. (2014). Contemporary challenges for education in conflict affected countries. *Journal of International and Comparative Education*, *3*(1), 113–125.
Smith, L. T. (2012). *Decolonizing methodologies: Research and indigenous peoples* (2nd ed.). Zed Books.
Sokoloff, K. L., & Engerman, S. L. (2000). History lessons: Institutions, factor endowments, and paths of development in the new world. *Journal of Economic Perspectives*, *14*(3), 217–232. https://doi.org/10.1257/jep.14.3.217
Song, M.-K., Lin, F.-C., Ward, S. E., & Fine, J. P. (2013). Composite variables: When and how. *Nursing Research*, *62*(1), 45–49. https://doi.org/10.1097/NNR.0b013e3182741948
Spaull, N., & Taylor, S. (2015). Access to what? Creating a composite measure of educational quantity and educational quality for 11 African Countries. *Comparative Education Review*, *59*(1), 133–165. https://doi.org/10.1086/679295
Srivastava, P. K., Han, D., Gupta, M., & Mukherjee, S. (2012). Integrated framework for monitoring groundwater pollution using a geographical information system and multivariate analysis. *Hydrological Sciences Journal*, *57*(7), 1453–1472.
Starr, H. (2003). The power of place and the future of spatial analysis in the study of conflict. *Conflict Management and Peace Science*, *20*(1), 1–20. https://doi.org/10.1177/073889420302000101
Steinberg, S. L., & Steinberg, S. J. (2015). *GIS research methods: Incorporating spatial perspectives*. Esri Press.
Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Lawrence Erlbaum.
Suleiman, M. D. (2017, February 27). *Sahel region, Africa*. The Conversation. http://theconversation.com/sahel-region-africa-72569
Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1
Sullivan, L. M. (2018). *Essentials of biostatistics in public health* (3rd ed.). Jones & Bartlett Learning.
Sumner, C. (2015). *Indonesia's missing millions: Erasing discrimination in birth certification in indonesia* (CGD Policy Paper 064). Center for Global

Development. https://www.cgdev.org/sites/default/files/CGD-Policy-Paper-64-Sumner-Missing-Millions.pdf
Suthar, A. B., Khalifa, A., Yin, S., Wenz, K., Fat, D. M., Mills, S. L., Nichols, E., AbouZahr, C., & Mrkic, S. (2019). Evaluation of approaches to strengthen civil registration and vital statistics systems: A systematic review and synthesis of policies in 25 countries. *PLoS Medicine*, *16*(9), 1–21. https://doi.org/10.1371/journal.pmed.1002929
Szreter, S., & Breckenridge, K. (2012). Recognition and registration: The infrastructure of personhood in world history. *Proceedings of the British Academy*, *182*, 1–36.
Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed). Pearson.
Tashakkori, A., & Creswell, J. W. (2007). Editorial: Exploring the nature of research questions in mixed methods research. *Journal of Mixed Methods Research*, *1*(3), 207–211. https://doi.org/10.1177/1558689807302814
Tashakkori, A., & Teddlie, C. (2008). Quality of inferences in mixed methods research: Calling for an integrative framework. In M. M. Bergman (Ed.), *Advances in mixed methods research: Theories and applications* (pp. 101–119). SAGE Publications.
Tashakkori, A., & Teddlie, C. (2010). *Sage handbook of mixed methods in social & behavioral research* (2nd ed.). SAGE Publications.
Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, *13*(1), 12–28.
The World Bank. (2008). *Transitions in secondary education in Sub-Saharan Africa: Equity and efficiency issues*. The World Bank. https://doi.org/10.1596/978-0-8213-7342-2
Thomas, V., Wang, Y., & Fan, X. (2001). *Measuring education inequality: Gini coefficients of education* (Policy Research Working Paper 2525). The World Bank. https://openknowledge.worldbank.org/handle/10986/19738
Tikly, L. (2001). Globalisation and education in the postcolonial world: Towards a conceptual framework. *Comparative Education*, *37*(2), 151–171. https://doi.org/10.1080/03050060124481
Todres, J. (2003). Birth registration: An essential first step toward ensuring the rights of all children. *Human Rights Brief*, *10*(3), 32–35.
Touval, S. (1966). Treaties, borders, and the partition of Africa. *The Journal of African History*, *7*(2), 279–293.
Touval, S. (1967). The organization of African unity and African borders. *International Organization*, *21*(1), 102–127.
Tuan, Y.-F. (1979). Space and place: Humanistic perspective. In S. Gale & G. Olsson (Eds.), *Philosophy in geography* (pp. 387–427). Springer Netherlands. https://doi.org/10.1007/978-94-009-9394-5_19
United Nations Children's Fund. (1998). UNICEF on deficient birth registration in developing countries. *Population and Development Review*, *24*(3), 659–664. https://doi.org/10.2307/2808179
United Nations Children's Fund. (2005). *The 'rights' start to life: A statistical analysis of birth registration*. UNICEF. https://www.unicef.org/publications/index_25248.html

United Nations Children's Fund. (2013a). *MICS5 indicators and definitions*. UNICEF. http://mics.unicef.org/tools?round=mics5
United Nations Children's Fund. (2013b). *A passport to protection: A guide to birth registration programming*. UNICEF. https://www.refworld.org/pdfid/52b2e2bd4.pdf
United Nations Children's Fund. (2013c). *Every child's birth right: Inequities and trends in birth registration*. UNICEF. https://www.unicef.org/publications/index_71514.html
United Nations Children's Fund. (2016). *Monitoring education participation: Framework for monitoring Children and Adolescents who are out of school or at risk of dropping out* (UNICEF Series on Education Participation and Dropout Prevention - Volume 1). UNICEF. https://www.unicef.org/eca/reports/monitoring-education-participation
United Nations Children's Fund. (2017). *A snapshot of civil registration in Sub-Saharan Africa*. UNICEF. https://data.unicef.org/resources/snapshot-civil-registration-sub-saharan-africa/
United Nations Children's Fund. (2018a). *SDG global indicators related to children* (UNICEF Briefing Note Series). UNICEF. https://data.unicef.org/wp-content/uploads/2018/05/SDG-Briefing-Notes-web-1.pdf
United Nations Children's Fund. (2018b). *Progress for every child in the SDG era*. UNICEF. https://data.unicef.org/resources/progress-for-every-child-2018/
United Nations Children's Fund. (2019a). *Every child learns: UNICEF education strategy 2019-2030*. UNICEF. https://www.unicef.org/media/59856/file/UNICEF-education-strategy-2019-2030.pdf
United Nations Children's Fund. (2019b, February 27). *School closures in the Sahel double in the last two years due to growing insecurity – UNICEF*. https://www.unicef.org/press-releases/school-closures-sahel-double-last-two-years-due-growing-insecurity-unicef
United Nations Children's Fund. (2019c). *Metadata*. https://unstats.un.org/sdgs/metadata/files/Metadata-16-09-01.pdf
United Nations Children's Fund. (2019d). *Education overview*. UNICEF. https://data.unicef.org/topic/education/overview/
United Nations Children's Fund. (2019e). *Birth Registration for every child by 2030: Are we on track?* UNICEF. https://data.unicef.org/resources/birth-registration-for-every-child-by-2030/
United Nations Children's Fund. (2020a). *UNICEF data warehouse* [Data set]. UNICEF. https://data.unicef.org/dv_index/
United Nations Children's Fund. (2020b). *Birth registration*. UNICEF. https://data.unicef.org/topic/child-protection/birth-registration/
United Nations Children's Fund. (2020c). *MICS6 Indicators and Definitions*. UNICEF. http://mics.unicef.org/tools
United Nations Department of Economic and Social Affairs. (2018). *Glossary of demographic terms*. https://population.un.org/wup/General/GlossaryDemographicTerms.aspx

United Nations Economic and Social Council. (2020). *Progress towards the Sustainable Development Goals* (E/2020/57). United Nations. https://undocs.org/en/E/2020/57

United Nations Educational, Scientific and Cultural Organization. (2018). *One in five children, adolescents and youth is out of school*. UNESCO Institute for Statistics. https://unstats.un.org/sdgs/metadata/files/Metadata-16-09-01.pdf

United Nations Educational, Scientific and Cultural Organization. (2020a). *4.1.4—Completion rate (primary, lower secondary, upper secondary)*. UNESCO Institute for Statistics. http://tcg.uis.unesco.org/4-1-4-completion-rate-primary-lower-secondary-upper-secondary/

United Nations Educational, Scientific and Cultural Organization. (2020b). *4.1.5—Out-of-school rate (primary, lower secondary, upper secondary)*. UNESCO Institute for Statistics. http://tcg.uis.unesco.org/4-1-5-out-of-school-rate-primary-lower-secondary-upper-secondary/

United Nations Educational, Scientific and Cultural Organization. (2020c). *UNESCO Institute for Statistics* [Data set]. UNESCO. http://data.uis.unesco.org/Index.aspx

United Nations Educational, Scientific and Cultural Organization. (2020d). *Background information on education statistics in the UIS database*. UNESCO. http://data.uis.unesco.org/ModalHelp/OECD/background-information-education-statistics-uis-database-2020-en.pdf

United Nations Office for the Coordination of Humanitarian Affairs. (2018, July 11). *The Sahel crisis deepens—Education under attack in Burkina Faso*. OCHA. https://www.unocha.org/story/sahel-crisis-deepens-education-under-attack-burkina-faso

United Nations Statistics Division. (2017). *Population density and urbanization*. United Nations. https://unstats.un.org/unsd/demographic/sconcerns/densurb/densurbmethods.htm

Uprichard, E., & Dawney, L. (2019). Data diffraction: Challenging data integration in mixed methods research. *Journal of Mixed Methods Research*, *13*(1), 19–32. https://doi.org/10.1177/1558689816674650

U.S. Agency for International Development. (2012). *Demographic and Health Survey: Sampling and house listing manual*. USAID. https://www.dhsprogram.com/pubs/pdf/DHSM4/DHS6_Sampling_Manual_Sept2012_DHSM4.pdf

U.S. Agency for International Development. (2018). *Guide to DHS statistics* [DHS-7 (version 2)]. USAID. https://www.dhsprogram.com/pubs/pdf/DHSG1/Guide_to_DHS_Statistics_DHS-7_v2.pdf

U.S. Agency for International Development. (2020). *Demographic and Health Survey: Interviewer's manual*. USAID. https://www.dhsprogram.com/pubs/pdf/DHSM1/DHS8-Interviewer's-Manual-EN-25Sep2020-DHSM1.pdf

Vandenabeele, C., & Lao, C. V. (2007). *Legal identity for inclusive development*. Asian Development Bank. https://www.think-asia.org/bitstream/handle/11540/4860/Legal%20Identity%20for%20Inclusive%20Development%20%282007%29.pdf?sequence=1

Vanner, C. (2015). Positionality at the center: Constructing an epistemological and methodological approach for a western feminist doctoral candidate conducting research in the postcolonial. *International Journal of Qualitative Methods*, *14*(4), 1–12. https://doi.org/10.1177/1609406915618094

Vasconcellos, E. (1997). Rural transport and access to education in developing countries: Policy issues. *Journal of Transport Geography*, *5*(2), 127–136. https://doi.org/10.1016/S0966-6923(96)00075-0

Venkatesh, V., Brown, S. A., & Bala, H. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. *MIS Quarterly*, *37*(1), 21–54.

Verspoor, A. M. (2008). *At the crossroads: Choices for secondary education in Sub-Saharan Africa*. The World Bank. https://doi.org/10.1596/978-0-8213-7113-8

Villet, C. (2011). Hegel and Fanon on the question of mutual recognition: A comparative analysis. *The Journal of Pan African Studies*, *4*(7), 39–51.

Wang, X., Luo, R., Zhang, L., & Rozelle, S. (2017). The education gap of China's migrant children and rural counterparts. *The Journal of Development Studies*, *53*(11), 1865–1881.

Warf, B., & Arias, S. (2008). Introduction: The reinsertion of space in the humanities and social sciences. In B. Warf & S. Arias (Eds.), *The spatial turn: Interdisciplinary perspectives* (pp. 1–10). Routledge.

Warshawsky, D. (2014). The potential for mixed methods: Results from the field in urban South Africa. *The Professional Geographer*, *66*(1), 160–168. https://doi.org/10.1080/00330124.2013.768141

Wendt, A. E. (1987). The agent-structure problem in international relations theory. *International Organization*, *41*(3), 335–370.

Williams, T. P., Abbott, P., & Mupenzi, A. (2015). 'Education at our school is not free': The hidden costs of fee-free schooling in Rwanda. *Compare: A Journal of Comparative and International Education*, *45*(6), 931–952. https://doi.org/10.1080/03057925.2014.938611

World Bank & World Health Organization. (2014). *Global civil registration and vital statistics: A scaling up investment plan 2015-2024*. World Bank Group. https://openknowledge.worldbank.org/handle/10986/18962

Yeager, C. D., & Steiger, T. (2013). Applied geography in a digital age: The case for mixed methods. *Applied Geography*, *39*, 1–4. https://doi.org/10.1016/j.apgeog.2012.12.001

Yoon, E.-S., Gulson, K., & Lubienski, C. (2018). A brief history of the geography of education policy: Ongoing conversations and generative tensions. *AERA Open*, *4*(4), 1–9. https://doi.org/10.1177/2332858418820940

Yoon, E.-S., & Lubienski, C. (2017). Thinking critically in space: Toward a mixed-methods geospatial approach to education policy analysis. *Educational Researcher*, *47*(1), 53–61. https://doi.org/10.3102/0013189X17737284

Yourkavitch, J., Burgert-Brucker, C., Assaf, S., & Delgado, S. (2018). Using geographical analysis to identify child health inequality in sub-Saharan Africa. *PLOS ONE*, *13*(8), 1–23. https://doi.org/10.1371/journal.pone.0201870

Yvonne Feilzer, M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, *4*(1), 6–16. https://doi.org/10.1177/1558689809349691

Zhang, W., & Creswell, J. (2013). The use of "mixing" procedure of mixed methods in health services research. *Medical Care*, *51*(8), e51–e57. https://doi.org/10.1097/MLR.0b013e31824642fd

Zuze, T. L., & Leibbrandt, M. (2011). Free education and social inequality in Ugandan primary schools: A step backward or a step in the right direction? *International Journal of Educational Development*, *31*(2), 169–178. https://doi.org/10.1016/j.ijedudev.2010.06.013

APPENDIX A: INSTITUTIONAL REVIEW BOARD APPROVAL

The UNIVERSITY of VERMONT
Committees on Human Subjects
Serving the University of Vermont
and the UVM Medical Center

RESEARCH PROTECTIONS OFFICE
213 Waterman Building
85 South Prospect Street
Burlington, Vermont 05405
(802) 656-5040
www.uvm.edu/irb/

Exemption Certification - Modification

To:	Thomas Griffin
From:	Gale Weld, Research Review Analyst, CIP
Approved Date:	December 17, 2020
Study#:	CHRBSS (Behavioral): STUDY00001338
Study Title:	Birth registration and access to education in Sub-Saharan Africa: An explanatory spatial analysis
Submission ID:	MOD00006835
Sponsor:	Internal Funding
Finalized Documents:	Exemption #4_v3 ;

Thank you for submitting the proposed modification to this exempt project. After review, it has been determined that the project still qualifies for exemption as indicated below under Section 45 CFR 46.104.

While the project is exempt from IRB review, it is required that researchers follow all human subject protection regulations and notify the IRB of any problems that arise during the conduct of the project.

Exemption Category: (4)(i) The identifiable private information or identifiable biospecimens are publicly available
Exempt Criteria 45 CFR 46.104(d)(4)(i) The identifiable private information or identifiable bio specimens are publicly available.

Consent/HIPAA/Waiver Determinations:
•Not Applicable

This determination applies only to the activities described in this IRB submission and will no longer apply should any changes be made. If changes are necessary, please submit a modification for consideration of a continued exemption.