Population Density and Governance in Africa

Pranish Desai (852380)

Supervisor(s):

Professor Rod Alence

A Research Report submitted in partial fulfillment of the requirements for the degree of Master of Arts in the field of e-Science

in the

School of Social Sciences
University of the Witwatersrand, Johannesburg

July 2021

Declaration

I, Pranish Desai (852380), declare that this research report is my own, unaided work. It is being submitted for the degree of Master of Arts in the field of e-Science at the University of the Witwatersrand, Johannesburg. It has not been submitted for any degree or examination at any other university.

Pranish Desai (852380)

July 2021

This page is intentionally left blank.

Acknowledgements

I would like to thank Professor Rod Alence for supervising me, and for his valuable input in carrying out this investigation, particularly in terms of the research methodology and his advice on how to revise the thesis.

Thank you to my reader, Mr Michael O'Donovan who offered useful insight which helped in the geospatial analysis components of the study.

Thank you to the team at Afrobarometer for granting me access to their geographically referenced data.

I would also like to thank my source of funding for the degree, the DSI-NICIS National e-Science Postgraduate Teaching and Training Platform (NEPTTP). Opinions expressed and conclusions arrived at, are my own and are not necessarily to be attributed to the NEPTTP.

I am also grateful to the academic and administrative staff at the University of the Witwatersrand who are involved with the NEPTTP. I would like to specifically thank Dr Helen Robertson and Ms Casey Sparkes.

Finally, I would like to express my profound gratitude to my parents, Pragnesh and Nisha Desai for their support throughout my schooling and university years.

Contents

List of Figures

List of Tables

List of Abbreviations

GPS	**G**lobal **P**ositioning **S**ystem
GPW	**G**ridded **P**opulation of the **W**orld
IQI	**I**nfrastructure **Q**uality **I**ndex
ITI	**I**nstitutional **T**rust **I**ndex
LME	**L**inear **M**ixed-**E**ffects
PSU	**P**rimary **S**ampling **U**nit
RLI	**R**ule of **L**aw **I**ndex
SGI	**S**atisfaction with **G**overnment **I**ndex
SSA	**S**ub-**S**aharan **A**frica
UN	**U**nited **N**ations
WBG	**W**orld **B**ank **G**roup

Chapter 1

Introduction

1.1 Introduction

Over the past half-century, much of Africa has experienced unprecedented rates of population growth and urbanisation (UN Habitat, 2016). However, a general consensus which exists amongst scholars about the African continent is that it has historically been sparsely populated, with low population densities particularly common inland (Badenhorst, 1951; Green, 2012a; Griffiths, 1995). Yet, up until the latter half of the 20th century obtaining reliable population estimates across the continent which could confirm this belief was a difficult task for three main reasons. First, many African societies lacked the resources capable of sustaining a full census, second the technical capacity to obtain accurate population measures was absent, and third there was a lack of consensus regarding whether national borders or tribal borders were the appropriate unit to use when considering the "area" dimension of population size and population density (United Nations, 1950). Once more reliable population data was available, this assumption of a sparsely populated continent found empirical support with several studies finding that substantial population clusters located in Africa were rare (Akyeampong, Bates, Nunn, & Robinson, 2014; Bates, 1987; Vengroff, 1976).

In recent years, some of this research into the population density patterns of Africa also sought to link lower densities across the continent with some of the continent's historical governance struggles with issues such as state size, governance capacity, economic development, inequality and conflict (Green, 2012a; Herbst, 2014;

Stephan & Tedrow, 1974). This largely qualitative body of research is closely associated with research into how the rural-urban divide has affected the political economy of Sub-Saharan Africa. The main scholarly argument regarding the effect which population density has on development and governance in Africa posits that through pre-colonial, colonial, and post-colonial times, rulers in African societies have had to grapple with the costs of governing over sparsely populated societies (Herbst, 2014). As a result, African societies – regardless of who was ruling them – have often been plagued by a weak state apparatus that was unable to efficiently perform essential functions such as tax collection, public goods distribution, sustaining state institutions and ensuring border security (Herbst, 2014; Robinson, 2002).

The emergence of more reliable population data for Africa has also coincided with rapid population growth and urbanisation in much of the continent – a decades-long process which has had further significant demographic, governance and developmental implications for the continent. To note one example of this rapid change, between 1965 and 2015, Africa's total population tripled, and the period also saw an average annual urbanisation rate of approximately 4% (UN Habitat, 2016). However, few studies into the density-governance relationship have investigated whether these demographic shifts have resulted in any changes to the theorised density-governance relationship within the African context. Specifically, there has been little research which uses contemporary data and primarily quantitative and spatial methods to determine whether low density areas across the African continent continue to experience specific governance and developmental issues relative to higher density areas in the continent. It is from this void in the existing research concerning the density-governance debate that this report presents the results of an investigation which sought to ascertain how the density-governance relationship operates within the contemporary African setting.

The investigation used 43,108 geographically referenced data points reflecting 2015 data from 27 countries[1] in Sub-Saharan Africa to determine whether population density patterns have a substantive effect on four dimensions of governance: infrastructure quality, institutional trust, the rule of law, and satisfaction with government. The remainder of this chapter provides more context regarding the purpose

[1] Appendix A contains the list of countries in the sample.

and significance of this cross-sectional study, the research question and research hypotheses, as well as the limitations and ethical considerations which affect the scope of the study.

1.2 Purpose of the Study

The purpose of the study is to determine whether variations in population density patterns across the African continent have a tangible effect in determining governance outcomes – particularly at the subnational level. Given that the existing literature on how the density-governance relationship operates in Africa is primarily qualitative, the intention is for this study's findings to help "fill" this gap within the current body of knowledge on this topic. To do this, four indicators of governance quality were created using data collected during Round 6 (2014–15) of the pan-African survey Afrobarometer (Afrobarometer, 2020). This data was then spatially linked to raster population data for 2015 compiled by the University of Southampton's WorldPop Project (University of Southampton WorldPop datasets, 2020) to test whether variations in population density across 43,108 data points located at the subnational level have an effect on infrastructure quality, institutional trust, the rule of law and satisfaction with government at those same data points. This merged dataset was used within statistical modelling techniques including multivariate linear regression and mixed-effects modelling to test the hypotheses which this chapter outlines.

1.3 Significance of the Study

The study's findings add to the existing knowledge about how population density patterns affect governance in Africa. The existing literature generally argues that areas with lower population densities are linked with governance struggles, while areas with higher population densities do not experience governance problems to the same extent (Bates, 1987; Herbst, 2014). The findings of this primarily quantitative study either solidify, or question the validity of this existing body of knowledge. If results reported in this study do show that areas in Africa which have lower population densities are more prone to certain governance problems such as lower infrastructure quality, or lower levels of institutional trust, then this lends further credibility to the existing beliefs about the density-governance relationship. By contrast, results which show no substantive link between population density and governance quality would call into question some of the assumptions of the existing literature with regard to contemporary issues of African governance.

Perhaps the most complicated potential outcome of this study are results which show that lower density areas may actually perform better on certain governance issues when compared to higher density areas, even when controlling for other factors. This is because such a result would confirm the belief that population density does influence governance outcomes across the continent – but suggest that the effect may be in the opposite direction to the theorised one. In this scenario, the fact that the data sources used in this study are publicly available, and the fact that the methods used in this study are replicable means that this study can help inform how future studies consider the density-governance relationship in the African context. Lastly, given the rapid demographic changes which the African continent is continually experiencing (UN Habitat, 2016), the results of the study can potentially help policymaking by identifying some of the demographic and geographic patterns which can result in different intrastate levels of infrastructure quality, institutional trust, or satisfaction with government.

1.4 Research Question

The intention is for the study's findings to help answer the following research questions:

1. To what extent do variations in population density have a causal influence on variations in specific governance indicators at the subnational level in Sub-Saharan Africa?

   Where the specific governance indicators are:

   (a) Infrastructure Quality Index (IQI)

   (b) Institutional Trust Index (ITI)

   (c) Rule of Law Index (RLI)

(d) Satisfaction with Government Index (SGI)

2. Does this influence differ depending on whether a governance indicator measures more "objective" aspects of governance quality, or if the indicator measures the "subjective" perceptions citizens have of governance quality within their society?

The first question asks for clarity on the nature of the relationship between variations in population density and variations in governance outcomes at the subnational level in Sub-Saharan Africa. It inquires whether population density actually has a meaningful influence on the four aspects of governance included in this study, as well as the direction and extent of this effect. The second question queries whether the effect which variations in population density have is different in cases where more "objective" aspects of governance are measured, when compared to more "subjective" aspects of governance. Since the data source for the governance indicators used in this study was the pan-African survey Afrobarometer (Afrobarometer, 2020), the more "subjective" measures of government performance are contingent on the beliefs citizens hold about the quality of this performance. The more subjective measures included in this study are institutional trust, satisfaction with government, and the rule of law. On the other hand, the indicator measuring quality of infrastructure is a more "objective" measure since it depends on whether or not the Afrobarometer interviewer found evidence of specific infrastructural facilities such as an electric grid, piped water and paved roads at a specific sample location (Isbell, 2017).[2]

1.5 Research Hypotheses

Based off the arguments of the existing research on this topic, one expects that there is a positive relationship between population density and specific aspects of governance in Africa. Specifically, the expectation is areas with lower densities are linked to adverse governance outcomes, while the opposite applies for the effect of higher density areas (Herbst, 2014). However, in framing the null and alternate hypotheses, the key determinant of whether the first null hypothesis will be

[2]Chapter 3 which outlines the research methodology used in this study contains more information about the governance indicators mentioned in this sub-section.

rejected depends on the strength and significance of the effect – as opposed to the expected direction of this effect. The set 2 hypotheses address the "directional" question by outlining the baseline (null) expectation that there will be no directional difference in how variations in subnational population densities influence different governance outcomes. Chapter 3 elaborates on the way in which the results of the investigation will influence whether these null hypotheses will be rejected in favour of these alternate hypotheses.

The following sets of null and alternate hypotheses inform this study:

Set 1 Hypotheses:

H_0: Controlling for other potentially causal factors, variations in population density at the subnational level have no meaningful causal effect on variations in specific governance indicators throughout Sub-Saharan Africa.

H_a: Controlling for other potentially causal factors, variations in population density at the subnational level have a meaningful causal effect on variations in specific governance indicators throughout Sub-Saharan Africa.

Where the specific governance indicators are:

1. Infrastructure Quality Index (IQI)
2. Institutional Trust Index (ITI)
3. Rule of Law Index (RLI)
4. Satisfaction with Government Index (SGI)

Set 2 Hypotheses:

H_0: There is no directional difference when comparing the effect of instances where subnational variations in population density influences "objective" indicators of governance such as infrastructure quality, and instances where subnational variations in population density influences "subjective" indicators of governance such as institutional trust, the rule of law and satisfaction with government.

H_a: There is a directional difference when comparing the effect of instances where subnational variations in population density influences "objective" indicators of governance such as infrastructure quality, and instances where subnational variations in population density influences "subjective" indicators of governance such as institutional trust, the rule of law and satisfaction with government.

In this second set of hypotheses, the null hypothesis expects that regardless of whether a governance quality indicator is "subjective" or "objective" the effect which variations in population density have on variations in these indicators is directionally identical. That is, if there is a positive causal link – higher densities are linked with better governance outcomes – between population density and the objective governance indicator of infrastructure quality, then we would also expect the effect which population density has on the "subjective" governance indicators to be directionally positive. By contrast, the alternate hypothesis considers the situation where variations in population density do have a different directional impact on variations in governance quality depending on whether the governance quality indicator is an "objective" or a "subjective" measure.

1.6 Limitations

The most prominent limitation of this study relates to the nature of the sample used in this thesis. In the main statistical models, this study encompasses 43,108 data points collected from 27 countries in Sub-Saharan Africa with the data primarily measuring governance quality and population densities as they existed in the year 2015. The 27 countries included in the study – see Appendix A – are all the non-island Sub-Saharan countries included in Round 6 of the Afrobarometer survey. The interviews for Round 6 of the Afrobarometer survey were conducted during 2014 and 2015 (Isbell, 2017). The implications of this limitation are twofold. First, the study is cross-sectional and does not encompass multiple years/decades in its analysis of how variation in population densities affects governance. Consequently, the results of this investigation primarily reflect how the density-governance relationship operated within the sample countries during 2015 – as opposed to how the density-governance relationship has operated over any lengthier period. Second, since the study comprises of data from 27 countries in Sub-Saharan Africa, its findings are not necessarily translatable to other parts of the world – and to a lesser extent, even the non-sample countries within Sub-Saharan Africa. While some of the general patterns found in this study may mimic those found in other parts of the continent or other parts of the world, it is beyond the scope of this study to account for the extent of these potentially similar patterns.

Beyond this main limitation, the other key limitations of the study pertain to how governance quality is measured in this study at the subnational level. As previously stated, governance quality is measured using four indexes: the Infrastructure Quality Index (IQI), the Institutional Trust Index (ITI), the Rule of Law Index (RLI) and the Satisfaction with Government Index (SGI). The limitation of measuring governance quality at the subnational level is that this thesis is more reliant on the more "subjective" perceptions of governance quality which citizens have than more "objective" measures of governance quality such as the nationally coded World Bank's Worldwide Governance Indicators (Kaufmann, Kraay, & Mastruzzi, 2011; World Bank, 2020c). This issue is grappled with in two ways. First, the identification of four distinct measures of governance quality ensures a focus on specific elements of governance as opposed to a ubiquitous concept of governance. Therefore, the findings of this study are pertinent to how population density affects these four aspects of governance. Second, to counteract the issue of "subjectivity" in these indicators, the control variables selected in this study focus on the demographic aspects of the data points such as the age, gender, level of educational attainment and employment status of Afrobarometer survey respondents.

1.7 Ethical Considerations

I re-iterate my commitment to academic best practices with regard to concerns of academic honesty, transparency and data usage. In particular, I strive to adhere to all ethical standards stipulated by the University of the Witwatersrand's School of Social Sciences. All data used in this study is secondarily sourced – that is, taken from existing databases, most of which are publicly available in some form. There was no formal ethical clearance required to conduct this research, save a School of Social Sciences Ethics Committee waiver. The waiver number for this thesis is: WINTR2020/07/02.

1.8 Report Structure

This report consists of six chapters. Subsequent to this introductory chapter which outlined the purpose and scope of the study, Chapter 2 provides a theoretical background to this study. This background is in the form of a literature review which focusses on how population patterns in Africa have historically been accounted for. The second chapter also includes a recounting of various studies which have analysed the links between population density patterns, development, and governance in the African context. Chapter 3 outlines and justifies the research design and methodology used in this study, with emphasis on the sample used in the study, the data sources used in the study, and the variables used to test the hypotheses this chapter has outlined. Chapter 4 presents the main results of the study while Chapter 5 discusses these results and the implications of these results with respect to both the research hypotheses, and the existing research on this topic. Finally, Chapter 6 concludes the study by re-iterating the key findings of the study and offering some suggestions for future studies into the effect which population density patterns have on governance outcomes in Africa.

Chapter 2

Theoretical Background

2.1 Introduction

This chapter provides a literature review encompassing two important topics pertinent to this study. The first section of the review considers studies on the measuring of African population patterns, with a focus on the challenges which have influenced the way in which population estimates for the continent have historically been made. The second section will focus on research which has sought to link Africa's population density patterns to issues of development and governance.

2.2 Studying Africa's Population Patterns

As the introduction section of this report mentioned, the ability to adequately obtain measures for population densities within Africa had historically been problematic for three reasons. These three reasons included many African societies lacking the resources which would allow for the ability to conduct a full census across vast areas. There was also a weak technical capacity across the continent which limited the reliability of any population estimates which were collected (Badenhorst, 1951). The final issue was a disagreement among scholars over the appropriate geographical unit with reference to which a census should obtain population count and population density estimates. The essential disagreement was between those who posited that the territorial borders in Africa acknowledged by the European colonial powers ought determine the areal aspect of density, and those who suggested that given the ethnic diversity of the whole continent, tribal borders were the superior proxy for area (Vengroff, 1976).

One prominent example of the juxtaposition of these problems is the United Nations Demographic Yearbook of 1949 (United Nations, 1950). The report asserted that out of an estimated continental population of 198 million people, only 44% of this estimated population had experienced a census of even a "fair" degree of reliability. Consequently, the report classified the reliability of its own African population and density estimates to be "poor" (United Nations, 1950). Nevertheless, the report estimated an extremely low continental population density figure of 6.5 persons per km^2 for 1949 (United Nations, 1950)[1]. The issue of unreliable contemporaneous estimates for population in Africa existed until deep into the 20th century. Several studies from this period noted that numerous newly independent African states had a lack of resources capable of financing and sustaining a full census especially in rural areas (Badenhorst, 1951; McEvedy & Jones, 1978; Stephan & Tedrow, 1974).

However, because the disagreement over the principal form of political territory continued to persist, some of these studies resorted to calculating density figures based off both national and tribal boundaries (Bates, 1987; Stephan & Tedrow, 1974; Vengroff, 1976). These studies noted that density levels were generally low regardless of the form of political territory used. One example of this was in a study done by Vengroff (1976) who found that 67 tribal territories out of a sample of 101 territories had an average density of less than 60 people per square mile. Similarly, out of a sample of 33 countries, 21 had an average density of less than 60 people per square mile (Vengroff, 1976).

These issues with obtaining accurate concurrent population estimates have also complicated attempts to make accurate projections of the historical sizes of African societal populations – especially for the period until 1950. The debate surrounding how to best estimate the historic size of African populations and understand

[1] As a means of comparison, the global average density in 1949 was 18 persons per km^2, and of all world regions, only Oceania recorded a lower average population density than Africa at that time (United Nations, 1950).

the continent's lower historical densities is an ongoing one. This debate has significant implications for understanding fields such as Africa's economic history and political evolution. Although there are differences among notable scholars studying Africa's historical populations, there is consensus among these scholars that the issue of low densities was one exacerbated by historical processes such as the transatlantic slave trade which particularly affected societies in West Africa[2], and the Indian Ocean slave trade which especially distorted population patterns in East Africa[3] (Akyeampong et al., 2014; Manning, 2015; McEvedy & Jones, 1978). In fact, scholars posit that these historical processes had such an effect that Africa's total population size likely stagnated between 1790 and 1890 (Akyeampong et al., 2014).

Some researchers find that the presence of European colonization in an area was inversely correlated with the population densities of those areas (Angeles & Neanidis, 2015). Such impactful historical processes also create challenges for the estimation of pre-colonial African populations and as a result, scholarly work on this topic has attempted to estimate these population numbers with reference to population size and growth patterns from other regions of the world. More recent developments in this debate have seen scholars use historical population size and growth rate figures adapted from South Asian countries such as India (Manning, 2015) and Indonesia (Frankema & Jerven, 2014) to create population databases for African societies for the period until 1950. Although such research and databases do not feature prominently in this report on how the population density-governance relationship in Africa has operated in more recent years, the research does create options for further quantitative analysis into the historic effect of population densities on governance and development in Africa.

In recent years there have been several developments which have allowed for accurate contemporaneous population count and population density estimates throughout Africa. At a basic level, most national governments across the continent now have the resources, access to modern transportation systems and technical capacities to conduct a full census encompassing both urban and rural areas. National

[2]The transatlantic slave trade peaked in the 1790s (Akyeampong et al., 2014).

[3]This is mainly in reference to the wave of the Indian Ocean slave trade which peaked in the 1850s (Akyeampong et al., 2014).

governments having these resources and this capacity has also meant that the debate over whether to use national or tribal borders as the proxy for the areal dimension of population counts or aggregate population densities has settled in favour of national borders. Furthermore, the emergence of satellite imaging has been of substantial benefit in helping to understand large-scale and small-scale human settlements throughout the world (Lloyd et al., 2019; Lloyd, Sorichetta, & Tatem, 2017).

The development of satellite imaging has also allowed for the emergence of relatively reliable proxies for human settlements such as road density data (World Bank, 2020a) and night-time lights (World Bank, 2020b). This study benefitted immensely from the both the improvement of census infrastructure across the continent, as well as the emergence of modern satellite imaging techniques. This is evident in the fact that the two main data sources which are used in this study to measure population density are the University of Southampton's WorldPop Project (University of Southampton WorldPop datasets, 2020), and the Gridded Population of the World (GPW) dataset compiled by NASA's Socioeconomic Data and Applications Center (Socioeconomic Data and Applications Center, 2020).

2.3 Population Density, Development and Governance in Africa

Over the last fifty years a growing body of scholarship has emerged which has linked the sparsely populated nature of African societies to specific governance and developmental struggles experienced by these societies. Among the struggles the research has identified as being partially linked to low population densities are issues of state size, governance capacity, conflict, inequality and economic development. This research all falls within the purview of the broader scholarship on African governance which as Alence (2017) notes can be characterized as "broadly institutionalist in its attention to the organization of political life". What is meant by this is that studies of African governance have emphasized how governance institutions are treated – whether by focussing on the effect social and political actors have on institutional formation and institutional constraints, or by stressing how cultural antecedents influence institutions (Alence, 2017).

Earlier studies which focussed on the link between population density and governance have usually linked Africa's low population densities with an inferior capacity for state formation, state consolidation and resource provision (Bates, 1987; Stephan & Tedrow, 1974; Vengroff, 1976). Later studies such as those by Herbst (2014), Gennaioli and Rainer (2007) and Green (2012a) have also linked Africa's low historical densities to governance issues including governance capacity and colonial era public goods provisions. Some other scholars including Cooper (1996), Mamdani (1996) and Boone (2003) use a case study approach in looking at the related issue of how the geographical distribution of human settlements within African societies can cause friction between how urban regions are governed, and how rural regions are governed.

The most comprehensive rationale forwarded in favour of arguing the population density-governance link exists comes from the book *States and Power in Africa* by Jeffrey Herbst (Herbst, 2014). Herbst's essential argument is that low population densities have historically meant that regardless of who the rulers of a specific society in Africa were, it has always been a costly endeavour – in terms of the resources and effort required – for states to maintain control over these sparsely populated societies (Herbst, 2014). Consequently, African societies including those from the pre-colonial, colonial and post-colonial periods were often plagued by a weak state apparatus unable to adequately perform essential functions such as tax collection (Frankema, 2011; Kasara, 2007), efficient public goods distribution, sustaining state institutions (Herbst, 2014) and ensuring border security (Englebert, Tarango, & Carter, 2002). Thus, in Herbst's view, the struggles with state formation and state consolidation throughout African history have their roots in the sparsely populated nature of the continent's landscape (Herbst, 2014).

Other authors have studied the role population density patterns play in exacerbating conflict in Africa. Whereas Raleigh and Hegre (2009) find that conflicts in Central Africa are more prevalent in areas where multiple populations cluster locally, Green (2012b) argues that Africa's historically low densities contributed to higher poverty levels, unequal property rights and ethnically fractionalized populations within post-colonial African states. These problems were further aggravated by African societies' rapid population growth since the mid–20th century, thereby increasing the likelihood of conflicts based off land disputes (Green, 2012b).

Other studies researching the influence of population density patterns in Africa have noted how these patterns have affected economic development (Hopkins, 2009; Van de Walle, 2009). Whereas scholars focussing on governance have argued that low densities meant that areas further away from societal capitals were harder to effectively govern, other authors have argued that several land abundant-low density African societies have faced developmental struggles because of small, sparsely distributed labour forces. Several studies have used historical population density and urbanization estimates as a proxy for income levels from pre-industrial eras because of the belief that dense populations could only exist if a region had a certain level of prosperity (Acemoglu, Johnson, & Robinson, 2002; Nunn & Puga, 2012). These studies have argued that even accounting for less reliable data, Africa had higher urban densities relative to other world regions such as the Americas and Oceania – thereby indicating a greater level of pre-colonial prosperity up until 1500. These authors argue that the "reversal" towards lower population densities, and by extension lower prosperity in regions such as Sub-Saharan Africa is largely attributable to the legacy of European colonization (Acemoglu et al., 2002). However aspects of this argument have been criticized by others including Bandyopadhyay and Green (2012) and Austin (2008b) who utilize alternative measures of population density in 1500 and find that the relationship between pre-modern population density and modern income levels is not robust when considering the intervening colonial period (Bandyopadhyay & Green, 2012).

Whilst acknowledging the impact of colonization as a whole on Africa's population density patterns and economic development, other scholars have studied the specific impact of how the slave trade radically transformed Africa's population density patterns and prospects for economic development in much of the 18th and 19th centuries (Akyeampong et al., 2014; Austin, 2008a; Bates, 1987). The crux of these arguments is that the slave trade had the substantial effect of deepening the divide between the abundant landscape of much of the continent, and the labour shortages directly caused by the slave trade. These authors argue that without the slave trade, this disparity would have been less severe – as was the case in other colonized parts of the world such as South Asia and Central America – and therefore many regions in Africa would have been better placed to continue their pre-colonial economic development trajectories (Austin, 2008a).

In recent years there have been studies which have debated the supposed links between population density patterns and issues of inequality and economic development in post-colonial Africa (Fosu, Bates, & Hoeffler, 2006; Min, Gaba, Sarr, & Agalassou, 2013; Mveyange, 2015). While Mveyange (2015) finds that a combination of urbanisation, lights per capita and population density patterns helps with explaining regional inequality in Africa, Fosu et al. (2006) argue that the example of states such as Botswana offer instruction for how individual African states can economically develop in spite of low population densities in rural areas. Furthermore, Gosh, Anderson, Elvidge, and Sutton (2013) discover that night-time satellite imagery can be a substitute measure for poverty at the provincial/state/regional level in Africa. In accounting for the disparity in economic development between higher density urban areas and lower density rural areas, some research has suggested that an element of reverse causality may exist in that while population density is linked with better infrastructure partly because higher densities improve the incentives for building and maintaining infrastructure, it is also the case that these superior facilities in turn attract migrants from lower density rural areas (Bates, 1987). Thus, while the main causal link sees rises in population density increase the likelihood of better infrastructure, there is a smaller causal link which in turn sees this better infrastructure increase population sizes – and in turn densities – within urban areas.

A common thread which links these studies into how population densities in Africa influence governance, economic development and the incidence of conflict is that the vast majority of them take a primarily qualitative approach to studying the problem. Furthermore, most of these studies have placed greater emphasis on the historical aspects of the relationship between population density and governance in Africa. The effect of this approach to understanding the density-governance relationship is that there is a "gap" in the literature in that there have been few studies which have relied primarily on quantitative and spatial methods to understand how this relationship has operated in more recent times. The aim is for this study's findings to help contribute to the closing of this "gap" since it is a primarily quantitative study which seeks to determine whether existing subnational variations in population density have a meaningful effect on specific governance outcomes in Sub-Saharan Africa.

2.4 Conclusion

As can be seen from this literature review, the study of population count and population density patterns in Africa has historically faced substantial challenges. With improvements in the reliability of aggregate population data for the continent, there has been a growing body of scholarship which has sought to understand the historical influence which population density patterns – particularly low densities – have had on historical governance and developmental outcomes across the African continent. However, these studies have generally been qualitative in nature and also tend to focus on the historical aspects of the population density-governance relationship. In short, there have been very few quantitatively based studies which have sought to determine whether a density-governance link currently exists in Africa. This is a relevant concern because of the rapid population growth and rapid urbanisation the continent has experienced over the past half-century (UN Habitat, 2016). The intent is for this study's findings to help "fill" this void since these findings have been developed using a combination of mapping and statistical analysis tools to determine whether variations in population density patterns influence variations in specific subnational governance outcomes within the contemporary African setting.

Chapter 3

Research Methodology

3.1 Introduction

This chapter discusses the research methodology used in the attempt to understand how variations in population density affect specific governance outcomes at the subnational level in Sub-Saharan Africa. The chapter commences by providing guidance on the research design of the study before moving on to discuss the sample encompassed by the study. Thereafter, the chapter focusses on the data sources, sampling procedure and data preparation methods employed within the study. The subsequent three sections provide a comprehensive overview of the variables used in this study. Finally, this chapter provides an outline of how the hypotheses set out in Chapter 1 are tested and evaluated.

3.2 Research Design

3.2.1 Overall Research Design

The research design incorporates hypothesis-led methods, sampling methods and statistical methods. The hypotheses presented in Chapter 1 informed how this methodology was selected as a means of testing and evaluating the falsifiability of the null hypotheses. Sampling methods are relevant to the study in two ways. First, due to issues of data availability, the study encompasses a sample of 27 Sub-Saharan

African countries to determine how the population density-governance relationship operates within the contemporary context. Second, because Afrobarometer data itself consists of nationally representative samples of citizens, the 43,108 data points analysed in this study represent a sample of co-ordinate locations at which the subnational relationship between population density and governance is evaluated. The remainder of this section discusses the main statistical methods used to evaluate the validity of the null hypotheses. The main dataset comprising of 43,108 data points – reflecting data for 2015 – was used to construct multivariate linear models, and mixed-effects models as the main determinant of the validity of the hypotheses. However, a smaller dataset of 4,930 data points – also reflecting 2015 data – was used as a robustness check on the main results of the study.

3.2.2 Multivariate Linear Model

The multivariate linear model is one of the cornerstones of statistical modelling and it is appropriate to use in this study because the four governance outcomes which are dependent variables are all measured on a continuous scale ranging from 0 to 10. These four dependent variables are: the Infrastructure Quality Index (IQI), the Institutional Trust Index (ITI), the Rule of Law Index (RLI) and the Satisfaction with Government Index (SGI). The model tests for the extent of variation within a dependent variable which is caused by a predictor variable (Fox & Weisberg, 2018). In the multivariate linear models used in this study, the key predictor variable is the population density of a data point, while the key dependent variable is a specific governance outcome. Within the multivariate linear models, the study also includes four control variables measured at the level of the individual data point. These four control variables record the age of an Afrobarometer Round 6 respondent, whether they have graduated from high school or not, whether they are male/female and whether they held some employment at the time of the Afrobarometer survey[1].

The debate surrounding the appropriate number of control variables to use in a quantitative study within the social sciences is an ongoing one. Whereas some scholars endorse including as many control variables as a model allows for so that the testing of a relationship between two variables is as statistically exhaustive as

[1] Latter sections of this chapter provide an overview of the dependent and predictor variables used in this study.

possible, others favour using a limited number of carefully identified control variables so that it is easier to discern the precise effect which an independent variable has on a dependent variable given the presence of control variables (Achen, 2005). For the purposes of this study, the latter approach was adopted as the study associates greater risk with including variables which may conceal the relationships between the carefully identified variables, than with potentially excluding several variables which may or may not have an influence on both population density patterns and on governance outcomes at the subnational level. Thus, many potentially relevant control variables measured at the national level such as national income levels, the presence of conflict in a region, or whether an individual is situated in a landlocked country do not feature in the study since they detract from the study's specific purpose of more clearly distinguishing the direction, size and significance of the effect which subnational variations in population density have on subnational variations in governance outcomes. The study deals with the potential risks which arise from the deliberate narrowing of the number of control variables through other means such as mixed-effects models which distinguishes only by the country in which an individual resides, tests for nonlinearities in the relationship and robustness check models which measure the density-governance relationship at a slightly "higher" unit of analysis than the individual level[2].

3.2.3 Mixed-Effects Model

In addition to the multivariate linear model, a mixed-effects model was also used to study the relationship between population density and governance outcomes. The key difference between the multivariate linear model and the mixed-effects model is that the mixed-effects model controls for the country from which a data point originates at the group level. Whereas the independent, dependent and other control variables are measured at the level of the 43,108 subnational data points, the control variable for country is measured at a "higher", grouping variable level since it also considers whether the variation – if any – caused by the country where a data point is located is substantive. A key advantage of the mixed-effects model compared to the classical multivariate linear model is that it accounts for non-independence amongst the data points (Fox & Weisberg, 2018). In more simple

[2]The rest of this Chapter elaborates on these various models.

terms, the model accounts for the possibility that there may be more similarities between data points from the same country when compared to data points from different countries. These models also consider the potential for population density having a nonlinear effect on governance outcomes by including orthogonal polynomial terms within the model. The group level variable takes into account the random intercepts by Afrobarometer Primary Sampling Units (PSUs) – of which there are 4,930 in this study – which are nested within the 27-country sample.

3.2.4 Robustness Check procedure

As part of evaluating the validity of the research hypotheses, the robustness check models include the shifting of the key data point across which variation in population density and variation in governance outcomes are measured. For the main multivariate linear and mixed-effects models, the study measures variation across 43,108 data points representing 43,108 Afrobarometer Round 6 survey respondents from the 27-country sample. By contrast, the robustness check models measure variation across the 4,930 data points representing the 4,930 unique Afrobarometer PSUs located within the 27-country sample. As a result, the robustness check models measure variation in a more aggregated way than the main models do. For both the main models and the robustness check models, the data reflects population densities and governance scores for the year 2015.

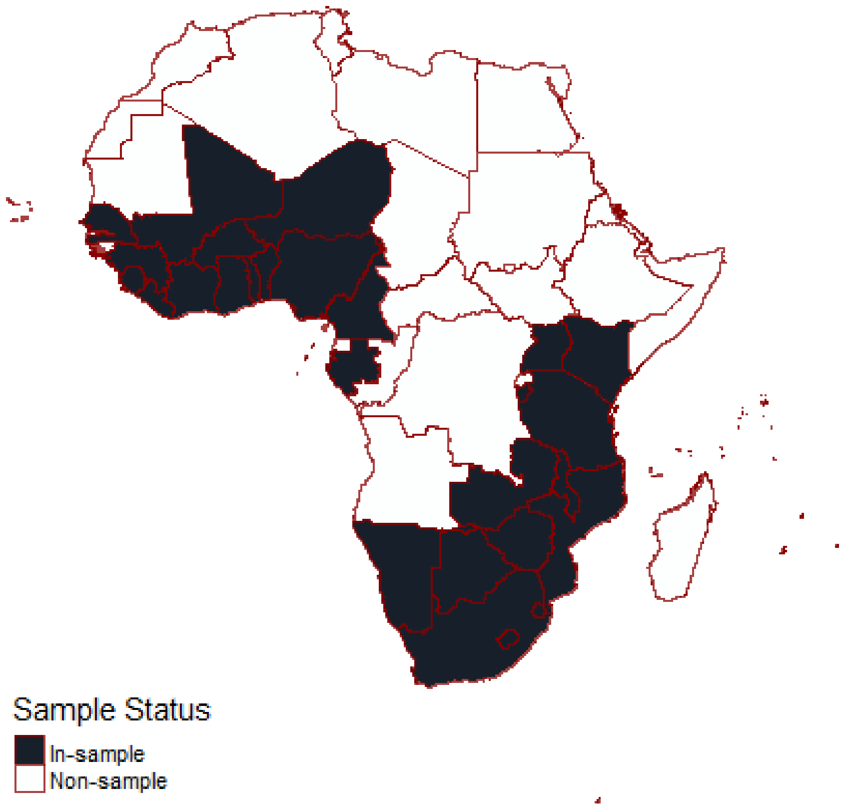
3.3 Investigation Sample

In total, there are 27 countries in Sub-Saharan Africa included in this study. These countries are: Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cote d'Ivoire, Eswatini, Gabon, Ghana, Guinea, Kenya, Lesotho, Liberia, Malawi, Mali, Mozambique, Namibia, Niger, Nigeria, Senegal, Sierra Leone, South Africa, Tanzania, Togo, Uganda, Zambia and Zimbabwe. Figure 3.1 maps the countries included within this study's sample[3]. In order to be selected as a sample country for this study, a country needed to be a non-island country in Sub-Saharan Africa where a Round 6 Afrobarometer survey was conducted. In line with the existing literature, North

[3]The shapefile used to create the African national borders which are the base of this map was obtained from the ICPAC Geoportal (ICPAC Geoportal, 2020).

African and island African nations were excluded from the study[4] on the basis of their different historical evolutions when compared to mainland Sub-Saharan African countries (Herbst, 2014).

FIGURE 3.1: Map of Sample



The year 2015 was selected as the sample year of the investigation as it is the most recent year for which there is geographically referenced data for both the independent variable of population density and geographically referenced versions of the dependent variables which are created using data from Round 6 of the Afrobarometer survey. The study is not a longitudinal one and therefore does not analyse how the population density-governance relationship changes over time.

[4]Round 6 Afrobarometer countries which are not included in this study are: Algeria, Cabo Verde, Egypt, Madagascar, Mauritius, Morocco, Sao Tome and Principe, Sudan and Tunisia.

3.4 Data Sources

The data used to calculate the four subnational governance indicators, as well as four of the control variables used in the study are all taken from the geocoded results of Round 6 of the pan-African survey Afrobarometer (Afrobarometer, 2020). The data in Round 6 of the Afrobarometer survey encompassed 36 African countries with interviews being conducted during 2014 and 2015 (Isbell, 2017). The number of survey respondents per country in the Afrobarometer survey is usually either 1200 or 2400, and the samples of the national populations which were surveyed were constructed using a design framework which was representative, stratified and clustered (Isbell, 2017). The data has a +/- 3% margin of error at the 95% confidence level. The geocoded aspect of the Afrobarometer data links a select number of interviewees[5] to the co-ordinates of a specific Afrobarometer Primary Sampling Unit (PSU). The Primary Sampling Unit within the Afrobarometer survey is the key geographically referenced subnational unit with reference to which survey respondents are selected and interviewed (Afrobarometer, 2020). In total, there are 43,108 Afrobarometer respondents located in 4,930 PSUs from 27 countries in Sub-Saharan Africa included in this study[6]. The governance quality and control variables data used within the robustness check models is also from Afrobarometer Round 6. The key difference is that whereas the main models measure variation across 43,108 Afrobarometer respondents, the robustness check models measure variation across 4,930 Afrobarometer PSUs.

The data used to measure the independent variable of population density is taken from the University of Southampton's WorldPop Project (University of Southampton WorldPop datasets, 2020). Specifically, the data source for the population densities of the 27 sample countries is the WorldPop Project's UN adjusted 1km resolution raster database for the year 2015 (University of Southampton WorldPop datasets, 2020). The raster data measures the number of people per km^2 at a resolution of 30 arc – which is approximately equivalent to 1km at the equator. The data is also adjusted to match the population estimates of the United Nations' Population Division within its Department of Economic and Social Affairs (United Nations

[5]It is most often the case that 8 interview respondents comprise a PSU. However, for South Africa there are usually 4 respondents per PSU. In some rare cases, the number of respondents per PSU is neither 4 nor 8 (Isbell, 2017).

[6]See Appendix A for country-level information about this sample.

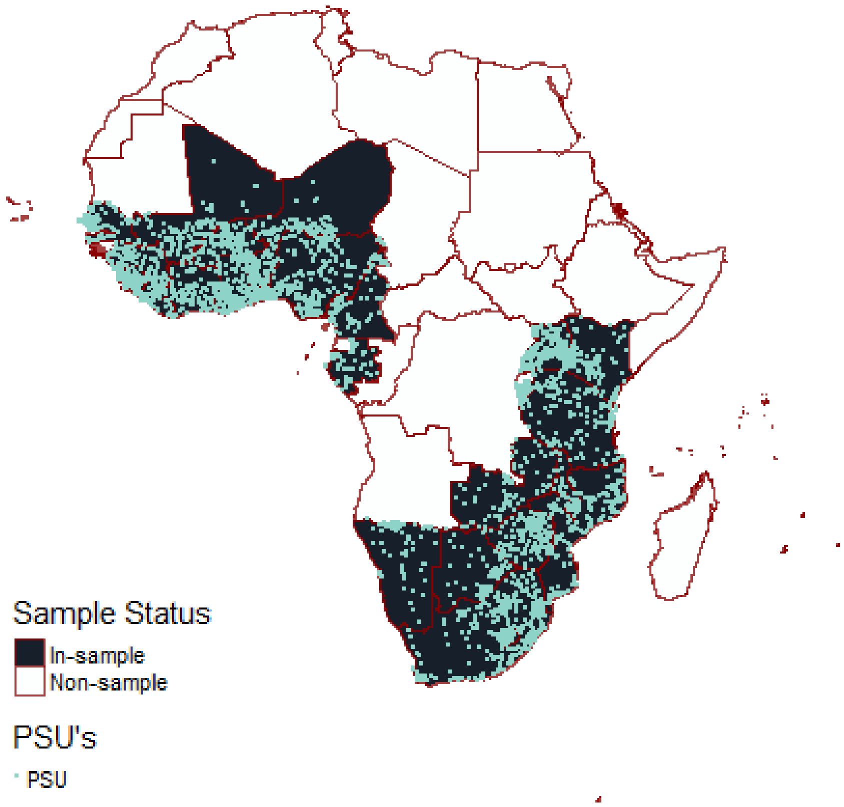
Department of Economic and Social Affairs Population Division, 2020). Given the extremely high resolution of this raster population density data, a merged raster containing population densities for all unique longitude-latitude grid cells for the 27-country sample has the population densities of nearly 15 million such grid cells [7]. However, since there were only 43,108 Afrobarometer respondents located in 4,930 PSUs included in the study, the actual number of grid cells containing population densities which were matched with the location of Afrobarometer PSUs on the basis of identical co-ordinate schemes was also 4,930 density cells.

3.5 Sampling Procedure

For the four dependent variables and the control variables, the responses of 43,108 Afrobarometer Round 6 respondents from 27 countries were geographically referenced at the level of the Afrobarometer Primary Sampling Unit (PSU). Consequently, the responses of these 43,108 respondents were geographically referenced at the locations of 4,930 Afrobarometer PSUs by using their Geographical Positioning System (GPS) co-ordinates. Figure 3.2 plots the locations of the PSUs on top of the base sample map used in figure 3.1 (Afrobarometer, 2020). The locations of the 4,930 PSUs within the 27 sample countries are represented by luminescent green dots. Appendix A lists the number of Afrobarometer respondents and Afrobarometer PSUs per country.

[7] For more information on this aggregate population density data refer to Appendix B.

FIGURE 3.2: Map of Round 6 Afrobarometer PSUs



3.6 Data Preparation

The data preparation, statistical modelling, data visualization and map creation components of this study were primarily conducted using the statistical computing language of R within the development environment of RStudio[8]. The main data preparation tasks were the "cleaning" of the Afrobarometer and WorldPop Project datasets, and the subsequent joining of these datasets. While "cleaning" the Afrobarometer Round 6 dataset, the survey questions which were not used to create

[8]The following R packages were explicitly used during the course of the study: *Simple Features for R* package (Pebesma, 2020), *dplyr* package (Wickham, François, Henry, & Müller, 2020), *raster* package (Hijmans, 2020), *stargazer* package (Hlavac, 2020), *car* package (Fox & Weisberg, 2020a), *visreg* package (Breheny, 2020), *lme4* package (Bolker, 2020), *influence.ME* package (Nieuwenhuis, Pelzer, & te Grotenhuis, 2020), *tmap* package (Tennekes, 2020), *ggplot2* package (Wickham & Pedersen, 2020), *caTools* package (Tuszynski & Dietze, 2020), *effects* package (Fox & Weisberg, 2020b) and the *sjPlot* package (Lüdecke, 2020).

either the governance indicators or any of the control variables were removed. All non-sample countries were also removed from the dataset.

For the population density data, individual raster files with 2015 population density data for the sample countries were downloaded from the WorldPop Project database (University of Southampton WorldPop datasets, 2020). The grid cells which did not have population density data were removed because missing data for the independent variable at specific co-ordinates would be of no use in answering the research question. The individual raster files of the sample countries were then merged into one large raster file containing the population densities in km^2 at 14,743,996 co-ordinate pairs[9].

Since both the Afrobarometer dataset, and the population density data were geographically referenced, it was possible to merge these two datasets by pairing the co-ordinate locations of Afrobarometer PSUs with the corresponding population densities at those co-ordinates. Where population density grid cells had no matching co-ordinate pair amongst the PSUs, these grid cells were "dropped" from the dataset. During this pairing process there were some occasions when a PSU's co-ordinates matched with more than one population density grid cell[10]. In these instances, the multiple population density cells which had "paired" with one PSU were averaged to get a single population density for that location. In the end, there was a dataset consisting of 43,108 data points geographically referenced at the level of 4,930 PSUs. This final dataset was used when conducting statistical analysis, when drawing maps, and when visualizing results.

3.7 Dependent Variables: Governance Quality Indicators

3.7.1 General Overview

As Chapter 1 stated, this study measures subnational governance quality using four indicators: the Infrastructure Quality Index (IQI), the Institutional Trust Index (ITI),

[9]See Appendix B for more information about this merged dataset.

[10]This usually occured when a PSU's co-ordinates fell precisely on the border of more than one raster cell.

the Rule of Law Index (RLI) and the Satisfaction with Government Index (SGI). Each indicator was created using data from Round 6 of the pan-African survey Afrobarometer (Afrobarometer, 2020). Each indicator was rescaled so that governance quality would be measured on a continuous scale ranging from 0–10. This section provides a comprehensive overview of how each indicator was created for the purposes of this study. Appendix C includes maps of the governance indicator scores each PSU records[11].

3.7.2 Infrastructure Quality Index (IQI)

The first dependent variable used in the study is the Infrastructure Quality Index (IQI) and the index measures the extent to which respondents within a PSU have access to a specific set of infrastructural facilities or not. The index was created using the answers to ten questions from the Round 6 Afrobarometer survey. These ten questions were answered by an Afrobarometer interviewer in conjunction with a field supervisor. The responses to each question were binary coded with an answer of 0 indicating that the facility was not present, while an answer of 1 indicated that the facility was present (Isbell, 2017). The most common general form of these ten questions was framed as follows: *"Are the following facilities present in the primary sampling unit/enumeration area, or within easy walking distance: [insert facility type]?"* (Isbell, 2017). Table 3.1 summarises the question number and corresponding facility name/type.

TABLE 3.1: Summary of Infrastructure Quality Index Questions

R6 Afrobarometer Question Number	Infrastructural Facility Name/Type
Q EA_SVC_A	Electricity grid in the PSU
Q EA_SVC_B	Piped water system in the PSU
Q EA_SVC_C	Sewage system in the PSU
Q EA_SVC_D	Cell phone service in the PSU
Q EA_FAC_A	Post office in the PSU
Q EA_FAC_B	School in the PSU
Q EA_FAC_C	Police station in the PSU
Q EA_FAC_D	Health Clinic in the PSU
Q BEA_FAC_F	Bank in the PSU
Q EA_ROAD_A	Tarred/paved road

[11] These mapped scores reflect the PSU-level governance outcome scores taken from the dataset used in the robustness check models of this study.

Since there are 10 questions, the affirmative responses to each question were aggregated. So if an Afrobarometer interviewer found that a person in a PSU had access to every facility, then the IQI score for that person was 10, while if the person had access to none of these facilities, then the IQI score would be 0. Thus, higher scores on the IQI indicate superior access to infrastructural quality. In comparison to the other governance quality indicators, the IQI is more "objective" since it is based on whether the Afrobarometer interviewer and field supervisor found the presence of a specific infrastructural facility within a PSU. By contrast, the ITI, the RLI and the SGI were all based on the "subjective" responses given by survey participants. Non-substantive data points – such as those coded as "Don't know" or as "missing" – were recoded as NA and were excluded from the aggregation process. For the robustness check models, the IQIs of respondents were averaged with reference to the number of respondents within a PSU.

3.7.3 Institutional Trust Index (ITI)

The second governance indicator is the Institutional Trust Index (ITI) which seeks to measure the extent to which Afrobarometer respondents trust key governance institutions, key political institutions, and key institutions within civil society. This indicator is also measured on a continuous scale with the maximum possible score on the ITI being 10. The indicator is adapted from Question 52 of the Round 6 Afrobarometer survey. The general form of this question is as follows: *"How much do you trust each of the following, or haven't you heard enough about them to say: [insert institution]?"* (Isbell, 2017). Table 3.2 provides a summary of the institutions included within this index.

TABLE 3.2: Summary of Institutional Trust Index Questions

R6 Afrobarometer Question Number	Institution Name/Type
Q 52A	President/Prime Minister
Q 52B	Parliament/National Assembly
Q 52C	National Electoral Commission
Q 52D	Tax Department
Q 52E	Local Government Council
Q 52F	Ruling Party
Q 52G	Opposition Party
Q 52H	Police
Q 52I	Army
Q 52J	Courts
Q 52K	Traditional Leaders
Q 52L	Religious Leaders

In order to operationalize these Afrobarometer questions into a 0–10 continuous measure of institutional trust, the aggregate responses to these questions were rescaled. Respondents to these questions were given four substantive answer options which were coded on a 0–3 categorical scale with higher scored responses indicating the respondent had greater trust in the specific institution. The allowed responses were: "0 = Not at all, 1 = Just a little, 2 = Somewhat, 3 = A lot" (Isbell, 2017). Non-substantive responses such as "Don't knows", refused to answer or missing answers were excluded from the averaging process for this variable. Before rescaling the responses to these questions, the responses each respondent gave to the questions were averaged so that each respondent had an average trust score ranging from 0–3. To rescale according to a 0–10 scale, each respondent's average trust score was multiplied by 3.333. So, if a respondent claimed to have "a lot" of trust in every single institution, then their overall personal trust score was 3 * 3.333 which equals 9.999, which is then rounded up to 10 – the highest possible score on the ITI. The respondent level ITI scores were used in the study's main models, while the averaged PSU-level ITI scores were used in the robustness check models.

The reasoning behind including traditional and religious leaders within the ITI is that within the African context, religious leaders and traditional leaders are among the most pivotal actors in civil society. These leaders often act as conduits in facilitating exchanges between citizens and their governments – and also in facilitating exchanges among ruling elites (Sanny & Asiamah, 2020). Even though governance quality is measured at the subnational level in this study, national-level institutions such as presidents/prime ministers, parliaments and national election commissions were included in the ITI because these institutions have the ability to influence more localized aspects of governance through actions like directing government expenditures for local projects, ensuring local stability, and assisting local leaders in times of crisis.

3.7.4 Rule of Law Index (RLI)

The third governance indicator used is the Rule of Law Index (RLI), and it captures the extent to which Afrobarometer respondents believe that the rule of law exists within their society. This index is also continuously measured up until 10 and it includes components relating to whether Afrobarometer respondents believe there is equal treatment under the law, and the extent to which respondents perceive executive branch leaders to respect the legal authority vested in other branches of government such as courts and parliaments. Higher scores on this indicator highlight increasingly positive perceptions of the rule of law within a society. As table 3.3 lists, the indicator was calculated using responses to Questions 45C, 45E, 51B, 51C and 51D of the Round 6 Afrobarometer survey. The general form of these questions is: *"In your opinion, how often, in this country: [insert violation of the rule of law]?"* (Isbell, 2017).

TABLE 3.3: Summary of Rule of Law Index Questions

R6 Afrobarometer Question Number	Violation of the Rule of Law
Q 45C	President ignores courts and laws
Q 45E	President ignores parliament
Q 51B	Unequal treatment under the law
Q 51C	Officials go unpunished after law-breaking
Q 51D	Ordinary citizens go unpunished after law-breaking

In order to rescale the RLI onto a 0–10 scale with higher RLI scores indicating a greater perceived level of the rule of law, the categorical responses were recoded so that this was possible. This is because within the original categorical options available to Afrobarometer respondents, "higher" responses indicated a greater perception that violations of the rule of law were occurring (Isbell, 2017). Table

3.4 summarises this recode summary. Once the responses were recoded, the same averaging and rescaling process which was used to calculate the Institutional Trust Index was followed.

TABLE 3.4: Rule of Law Index Recode Summary

Response Label	Original Scale Value	Recoded Scale Value
Never	0	3
Rarely	1	2
Often	2	1
Always	3	0

3.7.5 Satisfaction with Government Index (SGI)

The final governance indicator is the Satisfaction with Government Index (SGI) and it measures the degree of satisfaction Afrobarometer respondents have in terms of how their government is dealing with various policy action areas. As with the other governance indicators, the SGI is also measured on a continuous scale which maxes out at 10. The indicator is calculated using responses to Question 66 of the Afrobarometer Round 6 survey, the general form of which is: *"How well or badly would you say the current government is handling the following matters, or haven't you heard enough to say: [insert policy action area]?"* (Isbell, 2017). Table 3.5 summarises the policy action areas included in calculating this variable.

TABLE 3.5: Summary of Satisfaction with Government Index Questions

R6 Afrobarometer Question Number	Policy Action Area
Q 66A	Managing the economy
Q 66B	Improving the living standards of the poor
Q 66C	Creating jobs
Q 66D	Keeping prices down
Q 66E	Narrowing income gaps
Q 66F	Reducing crime
Q 66G	Improving basic health services
Q 66H	Addressing educational needs
Q 66I	Providing water and sanitation services
Q 66J	Ensuring everyone has enough to eat
Q 66K	Fighting corruption in government
Q 66L	Maintaining roads and bridges
Q 66M	Providing reliable electric supply

To ensure uniformity among the "subjective" measures of governance quality, the responses pertaining to how Afrobarometer respondents thought the government was handling various policy action areas were recoded, rescaled and then averaged. This is because as table 3.6 shows, the responses to these questions were originally coded on a 1:4 categorical scale as opposed to a 0:3 scale (Isbell, 2017). After the responses were recoded, they were then averaged and rescaled by respondent to calculate an individual's satisfaction score. This averaging and rescaling occurred in an identical manner to the ITI and the RLI.

TABLE 3.6: Satisfaction with Government Index Recode Summary

Response Label	Original Scale Value	Recoded Scale Value
Very Badly	1	0
Fairly Badly	2	1
Fairly Well	3	2
Very Well	4	3

3.8 Predictor Variables: Population Density and Control Variables

3.8.1 Independent Variable: Population Density

The independent variable used in this study is population density which is a measurement of the average number of persons within an areal unit. As previously stated, the main data source this study used to measure population density is the University of Southampton's WorldPop Project (University of Southampton WorldPop datasets, 2020) with the data reflecting the densities of geographically referenced co-ordinate grid cells for 2015 in persons per km^2. The WorldPop database is used to measure the population densities of each Afrobarometer PSU within the multivariate linear and mixed-effects models.

For both the main models, and the robustness check models a logarithmic transformation was used on the data for population density. This is in line with previous studies on the effects which population density patterns have on several governance and developmental issues (Alesina & Spolaore, 2005; Craig, 1984). The reasoning behind applying a logarithmic transformation to population density data within statistical models is that it is more important to consider ratio differences than absolute differences when comparing different densities. One example of the importance of ratio differences versus absolute differences is when considering two sets of population densities. While set one compares the density of a location with 100 people per km^2 and a location with 200 people per km^2, set two compares a location with 10,000 people per km^2 and a location with 10,100 people per km^2. Although the absolute difference between densities in either set is 100 people per km^2, the percentage difference between densities in set one is 200% while for set two it is just under 1%. Therefore, the advantage of using a logarithmic transformation on population density data within statistical models is that the transformed version of the variable follows a normal distribution thereby ensuring that extreme values do not disproportionately influence the model. Within the mixed-effects models, polynomial terms of population density were included to test whether any nonlinearities are present in the relationship between population density and governance outcomes.

Figure 3.3 displays the population density category of each PSU location, while table 3.7 provides a frequency summary of the number of PSU data points which fall within each population density category.

FIGURE 3.3: Map depicting 2015 Population Densities at 4,930 Afrobarometer PSUs in km^2

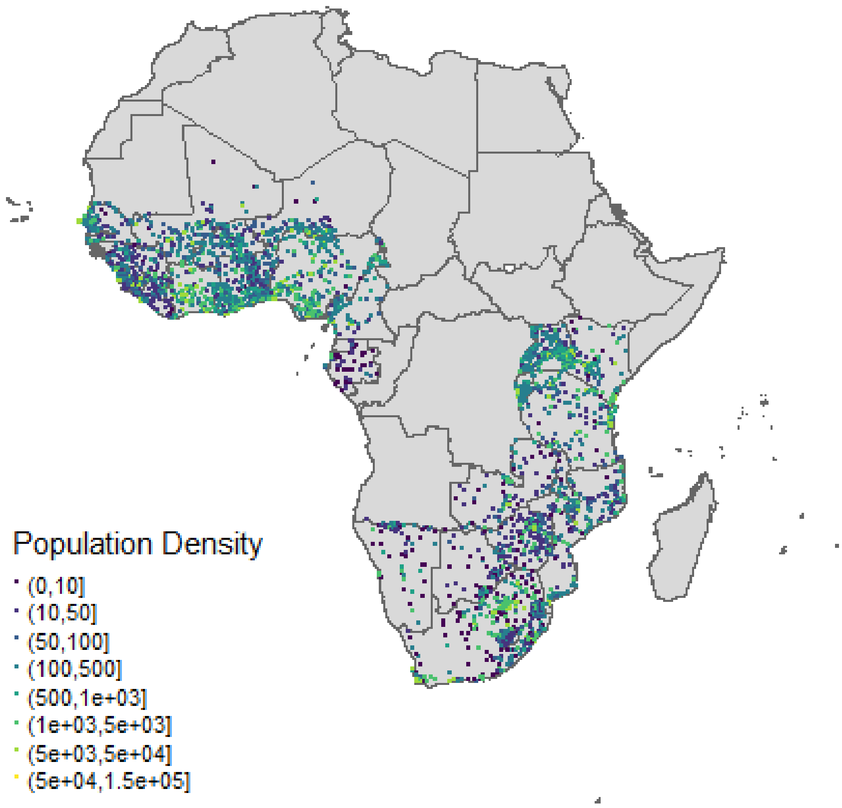


TABLE 3.7: Number of Data Points per 2015 Population Density Category at 4,930 Afrobarometer PSUs

Population Density Category (persons per km^2)	Number of Points per Category
0–10	290
10–50	710
50–100	599
100–500	1,363
500–1,000	457
1,000–5,000	897
5,000–50,000	612
50,000–150,000	2

3.8.2 Control Variable 1: Education

The first control variable used in the study is a measure of educational attainment. This control variable was selected due to research showing links between education levels and governance outcomes such as institutional trust and approval of government performance[12] (Drakos et al., 2016).

This variable is adapted from the responses to Question 97 of the Round 6 Afrobarometer survey which asks respondents: *"What is your highest level of education?"* (Isbell, 2017). Meaningful responses to this question were classified according to a ten-point scale ranging from 0–9. If a respondent received no formal schooling, only informal schooling, had only attended/completed primary school, or had only attended but not completed high school then they were classified as having not graduated from high school. By contrast, if a respondent had completed high school, had received some post-secondary education, had attended/completed university, or had received a post-graduate degree, then the respondent was classified as having graduated from high school. Respondents who did not know, refused to answer or had a missing answer were excluded from the calculation of each PSU's education ratio (Isbell, 2017). In total, 12,702 respondents had completed high school, while 30,406 respondents had not.

For the robustness check models, educational attainment was measured on a continuous scale ranging from 0 to 1. Educational attainment scores represented the proportion of respondents within each Afrobarometer PSU who had completed high school. The more respondents who graduated from high school within a PSU, the higher the score for that PSU. For instance, if out of an 8 person PSU, 3 people had completed high school, then the education ratio for that PSU was 0.375.

3.8.3 Control Variable 2: Employment

The second control variable is a measure of whether or not an Afrobarometer respondent held some level of employment at the time of the Round 6 Afrobarometer

[12] In the European context, higher levels of education were determined to be correlated with higher levels of trust (Drakos, Kallandranis, & Karidis, 2016). However, given the much lower rates of educational attainment in Africa, it may not necessarily be the case that the direction of influence would remain precisely the same in Africa as it is in Europe when assessing the education-trust relationship.

survey in their country. This variable was also selected because of the links found between it and factors such as institutional trust and approval of government performance[13] (Drakos et al., 2016). The variable was calculated from the responses to Question 95 of the Afrobarometer survey which asked: *"Do you have a job that pays a cash income? If yes, is it full-time or part-time? If no, are you presently looking for a job?"* (Isbell, 2017). Substantive responses to this question were originally recorded on a four-point scale ranging from 0 to 3. Regardless of whether they were actively seeking employment, respondents who held no employment at the time of the survey were reclassified as having no employment. Respondents who had part-time or full-time employment were categorized as having some employment. Once again respondents with non-substantive responses to this question were excluded from the calculation of the employment ratio for each data point. Of the 43,108 respondents in this study, 15,893 held some employment at the time of the Afrobarometer survey, while the remaining 27,215 did not hold any employment. As with educational attainment, the robustness check models measured the level of employment within a PSU on a continuous 0–1 scale. A higher employment ratio indicates a greater proportion of respondents held some level of employment.

3.8.4 Control Variable 3: Gender

The third control variable is the gender of the Afrobarometer respondent. The variable was determined using the answers by Afrobarometer survey interviewers to Question 101 which asked the interviewer to record the gender of the person who they had just interviewed (Isbell, 2017). The sample contained 21,388 male respondents and 21,720 female respondents. The robustness check models include a gender ratio variable also measured on a continuous 0–1 scale with higher scores indicating a greater proportion of male respondents within a data point. A score of 0.5 within an 8 person PSU means that there were 4 males interviewed in that PSU.

[13]The inclusion of the control variables of education and employment status within the same model may raise concerns about multicollinearity – whereby predictor variables are heavily correlated thereby distorting the model. These concerns are addressed in the next chapter.

3.8.5 Control Variable 4: Age

The fourth control variable is the age of an Afrobarometer respondent. The variable was based off the answer to Question 1 of the survey (Isbell, 2017). 42,837 respondents gave an age to their interviewer and the range of ages was from 18 years old to 105 years old. The control variable for age included in the robustness check models recorded the average age of the people located in a PSU. Based off an analysis of the distribution of this variable using the *car* package in R (Fox & Weisberg, 2020a), this variable was transformed so that each respondent's age was subject to the -0.33 power. So, the age variable in the regression models is in fact $Age^{-0.33}$.

3.8.6 Control Variable 5: Country

The final control variable of country is only included within the mixed-effects models. The variable simply denotes the country within which each Afrobarometer respondent is located. Operationally, the variable is used to clarify whether there is a country-level effect (grouping variable), which influences how the population density-governance outcome relationship operates amongst the 27-country sample. In the mixed-effects models, potential variation within PSUs within a country was controlled for by nesting the country-level effect by PSU using the term Country/PSU.

3.9 Summary of Variables and Data Sources

TABLE 3.8: Summary of Variables, Models and Data Sources

Variable Name	**Variable Type**	**Models used in**	Data Source
Infrastructure Quality Index (IQI)	Dependent Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Institutional Trust Index (ITI)	Dependent Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Rule of Law Index (RLI)	Dependent Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Satisfaction with Government Index (SGI)	Dependent Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Education	Control Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Employment	Control Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Gender	Control Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Age	Control Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(Afrobarometer, 2020)
Country	Control Variable	Mixed-Effects (nested by PSU), Robustness Check	(Afrobarometer, 2020)
Population Density (PSU)	Independent Variable	Multivariate Linear, Mixed-Effects, Robustness Check	(University of Southampton WorldPop datasets, 2020)

3.10 Procedure for Testing Hypotheses

Chapter 1 set out two null hypotheses and their validity is evaluated on the basis of the results of the statistical models. The first null hypothesis presupposed that variations in population density have no meaningful causal effect on variations in the four governance indicators at the subnational level in Sub-Saharan Africa. The second null hypothesis assumed that there is no directional difference when one compares the effect which population density has on the "objective" indicator of governance quality (the IQI), and the effect it has on "subjective" indicators of governance quality (the ITI, RLI and SGI).

The multivariate linear models are used to determine what the direction, size and significance of the effect population density has on governance outcomes at the subnational level is. The results of the mixed-effects models add the component consideration of whether the results obtained in the classic multivariate model hold even when country-level effects are controlled for. With both models, close attention is paid to the extent and significance of these relationships. This is done by looking at the size of coefficients and observing whether this effect is statistically robust at the 90% ($p<0.1$), 95% ($p<0.05$) and 99% ($p<0.01$) levels. If the effect is robust, then the first null hypothesis will be rejected. The results of the robustness check models used within the study will also determine whether this relationship is meaningful even when the key unit of analysis is shifted from the Afrobarometer respondent to the level of the Afrobarometer PSU.

In terms of evaluating the second null hypothesis, attention is paid to whether the independent variable of population density has an identical directional effect on each governance quality indicator – regardless of whether the indicator is "objective" (IQI), or "subjective" (ITI, RLI, SGI) in nature. Simply put, the second null hypothesis outlines the expectation of a positive directional effect of population density on infrastructure quality, to correspond to a positive directional effect of population density on institutional trust, the rule of law and satisfaction with government. If there is a different directional effect depending on whether a governance quality indicator is "subjective" or "objective", then this will result in the rejection of the second null hypothesis in favour of the alternate hypothesis.

3.11 Conclusion

This chapter has outlined the research methodology of this study of the subnational relationship between population density and four governance indicators: the Infrastructure Quality Index (IQI), the Institutional Trust Index (ITI), the Rule of Law Index (RLI), and the Satisfaction with Government Index (SGI). The chapter began with a discussion of the research design used in the study with emphasis on the regression specifications of the regression models employed to test the hypotheses. Thereafter the chapter discussed the investigation's sample, the data sources used, the sampling procedure of the study and the way in which the data was prepared for the purposes of the study. Following this, the chapter provided a comprehensive overview of how the variables used in this study were created, and outlined what they were intended to measure. Lastly, the chapter signaled how the research hypotheses are evaluated.

Chapter 4

Results and Analysis

4.1 Introduction

This chapter presents the results of the study into the nature of the relationship which exists between population density and governance outcomes at the subnational level in Sub-Saharan Africa. The chapter begins by providing a set of summary statistics reflecting the 43,108 data points from 2015 which comprise the main sample of this study. Thereafter, the paper briefly discusses the bivariate relationships which exist between the variables included in this study. The next section presents and analyses the main results of the study with a focus on the results of the multivariate linear and mixed-effects models used within the study. This section also focusses on the potential for nonlinear effects within the density-governance relationship, and reports how the relationship varies by country for each governance outcome as derived from the mixed-effects models. The fifth section of this chapter presents the results of the robustness check models. The chapter concludes with a summary of the main findings.

4.2 Summary Statistics

Table 4.1 provides the summary statistics for each variable across the 43,108 data points which represent the Round 6 Afrobarometer survey respondents from 27 countries included in this study. As the table shows, most of the control variables in the study are measured on binary 0–1 scale. The median respondent lives in a rural area, did not graduate from high school, and did not hold any form of employment at the time of the survey. There are slightly more females than males in the sample, and the median person is in their mid-thirties.

TABLE 4.1: Summary Statistics of each variable (2015)

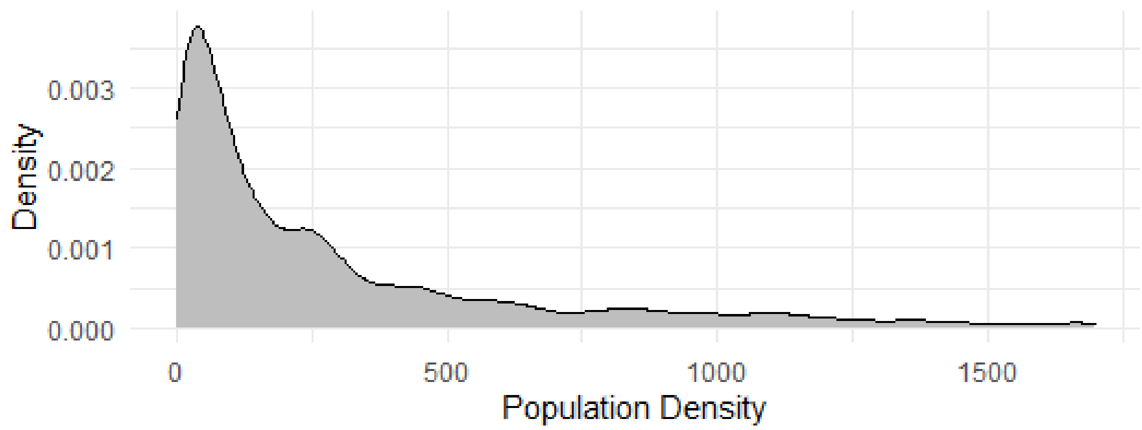
Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Urban	43,108	0.38	0.49	0	0	0	1	1
Some Employment	43,108	0.37	0.48	0	0	0	1	1
HS Graduate	43,108	0.29	0.46	0	0	0	1	1
Male	43,108	0.50	0.50	0	0	0	1	1
Infrastructure Quality Index (IQI)	43,108	4.98	2.61	0	3	5	7	10
Institutional Trust Index (ITI)	43,108	5.59	2.53	0.00	3.89	5.83	7.50	10.00
Rule of Law Index (RLI)	43,108	5.87	2.28	0	4.4	6	7.3	10
Satisfaction with Government Index (SGI)	43,108	3.79	2.20	0.00	2.31	3.85	5.38	10.00
Age	42,837	37.02	14.49	18.00	26.00	34.00	45.00	105.00
Population Density (km^2)	43,108	2,072.44	4,674.66	1.00	72.83	282.19	1,681.49	81,184.61
Population Density (log)	43,108	5.76	2.18	0.00	4.29	5.64	7.43	11.30

The four governance outcomes which are the dependent variables in this study are all measured on a 0–10 continuous scale. Of the three "subjective" measures of governance quality, it is evident that across the sample, the worst performing indicator is the Satisfaction with Government Index (SGI) which has a mean of 3.79 and a median of 3.85. By contrast, the means and medians for both the Institutional Trust Index (ITI), and the Rule of Law Index (RLI) are above the indicators' scale's midpoint of 5. Perhaps paradoxically, these summary statistics indicate that across the 27-country sample, citizens are much more likely to trust governmental institutions such as the president/prime minister, parliament, the police and the courts than be satisfied with the actual job their government is doing in various policy action areas such as improving living standards, improving educational services, fighting corruption and reducing crime. For all three "subjective" measures of governance quality, it is clear that the distribution of outcomes for these variables is quite even with extremely small differences between median and mean values

The distribution of scores for the more "objective" measure of governance quality – the Infrastructure Quality Index (IQI) is even given that the median IQI score of 5 barely exceeds the mean IQI of 4.98. The general distribution of scores for this variable suggest that infrastructure quality across the sample is not as low as satisfaction with government is, but also not as high as the average performances on the institutional trust and rule of law variables. It is also noteworthy from these mean and median scores that apart from the SGI, most scores on the IQI, the RLI and the ITI are neither extremely strong, nor extremely weak. This suggests middling governance outcomes across the 27-country sample for the year 2015.

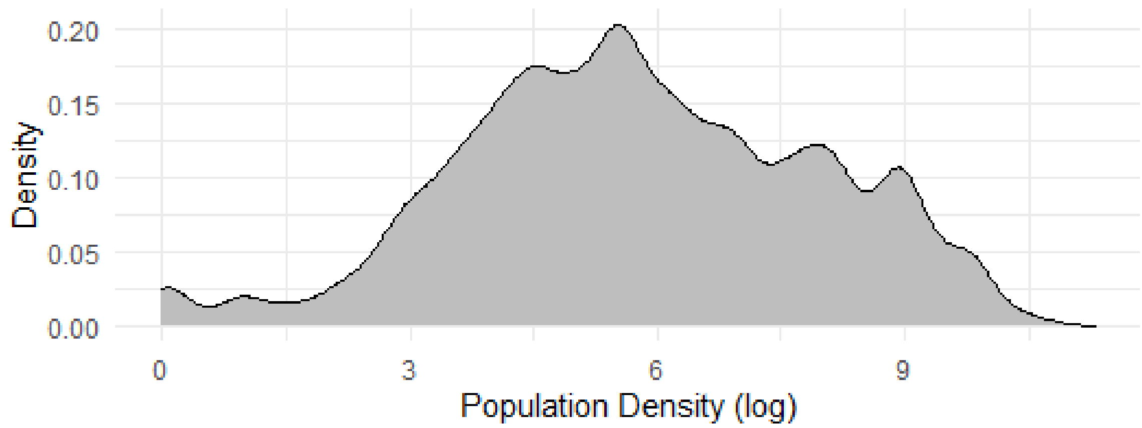
Regarding the independent variable of population density, the distribution of the untransformed variable has a distinct "right-skew" with a mean over seven times greater than the median. The uneven distribution of this variable is also apparent in the very high standard deviation which indicates that in its untransformed state, the population density variable is being heavily impacted by a few data points with very high population densities. In fact, within the entire 43,108 person sample, only 15 survey respondents – all of whom are located in Nairobi, Kenya – live in a population grid cell with a population density in excess of 50,000 people per km^2. This uneven distribution validates the decision to apply a logarithmic transformation to the population density variable. As can been seen in table 4.1 and figures 4.1 and 4.2, the transformed version of the variable follows a much more "normal" distribution which minimizes the risk that extreme values for this variable disproportionately impact the results of the regression models.

FIGURE 4.1: Density distribution plot for Population Density (km^2)



It should be noted that the density plot for the untransformed variable in figure 4.1 has a limit on the x-axis which means that the figure only displays the distribution of the population densities of the 75% smallest values for population density. It is easiest to understand the implications of these distribution plots with reference to the rural-urban divide. If one uses the definition of a rural area which is provided by the United Nations Food and Agriculture Organization, then the differentiation point between an urban area and a rural area occurs at the 300 persons per km^2 mark (Food and Agriculture Organization, 2018). In the plot for density (log), this urban-rural distinction occurs at approximately log(5.7). Considering these boundaries, the plot for the distribution of the untransformed population density variable (figure 4.1) is different in nature for rural areas and urban areas with the rural portion of the plot predominantly unimodal. The urban portion of the plot (> 300 persons per km^2) is more multimodal and has a wide distribution[1]. By contrast, there is a more "normal" distribution in the log-transformed version of population density (figure 4.2). The density distribution plot for this variable is multimodal with a sharp increase in the slope gradient from approximately log(1.5) – population density of 5 persons per km^2 – and an intermittent decline in the slope gradient starting once the rural-urban threshold of log(5.7) – population density of 300 persons per km^2 – is crossed.

FIGURE 4.2: Density distribution plot for Population Density (log)



[1]Especially when one considers that densities of over 1,700 persons per km^2 are excluded from the plot.

4.3 Bivariate Analysis

TABLE 4.2: Correlation between variables (2015)

	Urban	Some Employment	HS	Male	IQI	RLI	ITI	SGI	Age	Density	Density (log)
Urban	1	0.13	0.28	-0.001	0.56	-0.12	-0.21	-0.05	-0.09	0.45	0.52
Some Employment	0.13	1	0.18	0.10	0.12	-0.01	-0.05	-0.03	0.004	0.08	0.13
HS Graduate	0.28	0.18	1	0.09	0.28	-0.07	-0.17	-0.02	-0.17	0.15	0.19
Male	-0.001	0.10	0.09	1	0.002	-0.01	-0.002	-0.01	0.10	-0.0002	-0.001
IQI	0.56	0.12	0.28	0.002	1	-0.06	-0.17	-0.01	-0.07	0.28	0.36
RLI	-0.12	-0.01	-0.07	-0.01	-0.06	1	0.31	0.27	0.03	-0.08	-0.10
ITI	-0.21	-0.05	-0.17	-0.002	-0.17	0.31	1	0.42	0.10	-0.13	-0.16
SGI	-0.05	-0.03	-0.02	-0.01	-0.01	0.27	0.42	1	-0.01	-0.03	-0.04
Age	-0.09	0.004	-0.17	0.10	-0.07	0.03	0.10	-0.01	1	-0.05	-0.07
Population Density (km^2)	0.45	0.08	0.15	-0.0002	0.28	-0.08	-0.13	-0.03	-0.05	1	0.64
Population Density (log)	0.52	0.13	0.19	-0.001	0.36	-0.10	-0.16	-0.04	-0.07	0.64	1

The first component of the bivariate analysis in this study is an analysis of the results displayed in table 4.2 which show the correlations between the variables used in this study. One of the most noteworthy aspects of these results is that the "objective" measure of government performance – the IQI – is negatively correlated with the ITI, the RLI and the SGI. This indicates that better levels of infrastructural quality are linked with lower levels of institutional trust, less satisfaction with government, and more scepticism that the rule of law exists within a society. This finding would potentially violate the assumption of a directionally uniform relationship between population density and each of the outcome variables set out in Chapter 1's research hypotheses.

This somewhat unexpected result in terms of the "directional" aspect of the relationship between the governance indicators themselves is mimicked in the unexpected directions of the coefficients representing the bivariate relationships between the independent variable of population density and the four governance indicators. In both the original and transformed versions of population density, the independent variable is negatively correlated with the subjective measures of governance quality, namely the ITI, RLI and SGI. However, the correlation between population density and the IQI is directionally positive and therefore in line with the expectations outlined in the existing research on the topic (Green, 2012a; Herbst, 2014). As the literature review chapter of this thesis mentioned, some scholars have raised the possibility that reverse causality may exist in the relationship between higher density urban areas and infrastructure quality in Africa in that while superior infrastructure may be centered around urban areas in order to meet the needs of larger density areas, it may also be the case that this superior infrastructure in turn increases population size and density in the longer run as more migrants are attracted from lower density rural areas (Bates, 1987). Since this study takes a cross-sectional approach, the results of the bivariate and multivariate models cannot by themselves fully comment on the extent of the reverse causality within the year 2015 itself[2], however, this study agrees with the assumption that the main causal link – especially in the short term – is of population density on infrastructure quality.

It is also worthwhile to study the degree of correlation among the various predictor

[2]To do so would require a more longitudinal dataset, something which is mentioned as a suitable avenue for future research in Chapter 6.

variables used in the study. This is because a high degree of correlation between predictor variables could indicate multicollinearity, whereby the standard errors within the multivariate linear and mixed-effects models could be inflated. Inflated standard errors would hint at structural flaws with the model (Fox & Weisberg, 2018). The correlation table indicates that population density in both the original and transformed form is not heavily correlated with the age, gender, education or employment control variables. The correlation table also includes a variable recording whether a person lives in a rural or urban location. As one would expect, there is a substantial correlation of 0.45 between population density and urban, and 0.52 between population density (log) and urban. This is expected given the traditionally close link between higher population densities and urban locations (Day & Day, 1973). Owing to this high correlation, the urban variable is excluded from the main regression models in order to minimize the risk of including theoretically similar measures within the same model. This is a relevant concern because in recent years classifications of urban and rural areas have been done with reference to the corresponding population densities of those areas (Dijkstra & Poelman, 2014). Appendix D includes summary results of regression models which include this urban variable.

Table 4.3 presents the results of the bivariate regression models between the log-transformed version of population density, and each of the four governance indicators. In each of the four bivariate models, population density is a meaningful predictor of variations in governance outcomes at the subnational level. Population density is a significant predictor at the 90%, 95% and 99% levels, with the size and direction of this effect varying by governance indicator. The impact of this finding is that it increases the likelihood – albeit subject to further confirmation – that population density is a meaningful predictor of governance outcomes. As a result, it is possible that the set 1 null hypothesis which posits that no meaningful relationship exists between variables does not hold.

On the issue of the direction of the predictive relationships, the results of the bivariate linear models once again report the unexpected discovery found in the correlation results. Although higher population densities are positively linked with better infrastructure quality, higher population densities are also linked with weaker scores on the ITI, the RLI and the SGI. While the strength of these directional effects differs by dependent variable, this result nevertheless increases the possibility that in the African context, population density affects "objective" measures of governance quality, and "subjective" measures of governance quality in different ways.

TABLE 4.3: Summary of Bivariate Linear Regression Models (2015)

	Dependent variable:			
	IQI	ITI	RLI	SGI
	(1)	(2)	(3)	(4)
log(Population Density)	0.433***	−0.182***	−0.106***	−0.040***
	(0.005)	(0.006)	(0.005)	(0.005)
Constant	2.486***	6.635***	6.486***	4.019***
	(0.033)	(0.034)	(0.031)	(0.030)
Observations	43,108	43,108	43,108	43,108
R^2	0.130	0.025	0.010	0.002
Adjusted R^2	0.130	0.025	0.010	0.002
Residual Std. Error	2.437	2.498	2.266	2.200
F Statistic	6,458.922***	1,092.166***	451.397***	67.670***
Note:				*p<0.1; **p<0.05; ***p<0.01

In terms of the strength of this directional effect, population density is strongest as a positive predictor of better infrastructure quality – coefficient of 0.433, and weakest as a negative predictor of satisfaction with government (-0.040). Although the coefficients for each of the bivariate density-governance relationships are robust at the 99% level, the size of these coefficients are themselves not overwhelming. Based on the results of these bivariate models it could at best be argued that population density is a moderately strong positive predictor of the IQI, but a weakly negative predictor of the ITI, RLI and SGI. This can also be seen in the very low R^2 and Adjusted R^2 scores which range from 0.002 to 0.130. The meaning of these low scores is that the bivariate models are not accounting for much variation in the respective dependent variables. However, as the next section shows, this aggregate explanatory effect of the models increases with the addition of control variables.

4.4 Main Results

This section presents the main results of the study. The section starts with the results of the multivariate linear regression models which include the four demographic control variables measuring educational attainment, employment status, gender and age. The section then reports the results of the mixed-effects models which remove the assumption of "non-independence" between data points by controlling for the country wherein an Afrobarometer PSU is located. The section also includes an analysis determining whether there are any nonlinearities present in the relationship between population density and governance outcomes in Sub-Saharan Africa. After this, the section briefly reports on the country-level random effects for each of the four density-governance relationships this study encompasses.

4.4.1 Main Results: Multivariate Linear Models

Table 4.4 displays the results of the multivariate linear models used to clarify the relationship between population density and governance outcomes at the subnational level in Sub-Saharan Africa. The reason for the smaller number of observations included in the multivariate models – 42,837 – is that 271 respondents did not give their age. One noteworthy aspect of these models which include control variables is that it is uniformly the case that the introduction of control variables has decreased the strength – though not necessarily the associated level of statistical significance – of the predictive effect which variations in population density have on variations in governance outcomes when compared to the coefficients reported in table 4.3.

TABLE 4.4: Summary of Multivariate Linear Regression Models (2015)

	Dependent variable:			
	IQI	ITI	RLI	SGI
	(1)	(2)	(3)	(4)
log(Population Density)	0.377***	−0.147***	−0.098***	−0.037***
	(0.005)	(0.006)	(0.005)	(0.005)
HS Graduate	1.194***	−0.730***	−0.272***	−0.038
	(0.026)	(0.027)	(0.025)	(0.024)
Some Employment	0.232***	−0.090***	0.081***	−0.097***
	(0.024)	(0.025)	(0.023)	(0.023)
Male	−0.092***	0.025	−0.041*	−0.016
	(0.023)	(0.024)	(0.022)	(0.021)
$Age^{-0.33}$	−0.459***	1.300***	0.168*	−0.361***
	(0.102)	(0.106)	(0.097)	(0.094)
Constant	3.377***	3.962***	6.156***	4.809***
	(0.215)	(0.224)	(0.205)	(0.199)
Observations	42,837	42,837	42,837	42,837
R^2	0.178	0.048	0.014	0.003
Adjusted R^2	0.178	0.048	0.014	0.002
Residual Std. Error	2.369	2.465	2.260	2.197
F Statistic	1,851.011***	434.454***	118.282***	22.247***
Note:				*p<0.1; **p<0.05; ***p<0.01

Once again, the “strongest” coefficient is within the model where population density is accounting for variations in infrastructure quality (0.377), while the “weakest” coefficient is in the model where population density is a predictor of satisfaction with government (-0.037). For all four governance indicators, population density is a meaningful predictor at the 90%, 95% and the 99% levels indicating that in most cases, the set 1 null hypothesis which presupposes no meaningful relationship is likely to fall away.

Regarding the question of direction, the results reported in table 4.4 once again point to a violation of the assumption that the effect which population density has on governance quality indicators is uniform regardless of whether the dependent variable is an “objective” or “subjective” measurement of government performance. While the directional effect of population density on the IQI remains positive, the directional effect of population density on the ITI, the RLI and the SGI is still negative. Therefore, these results once again suggest the potential that within the context of Sub-Saharan Africa, the question of whether population density has a different directional effect on different governance quality indicators is a legitimate one.

It is also useful to observe the relationships which the various control variables have with the governance indicators. As one would expect there is a very strong positive relationship between higher educational attainment and the IQI, and a reasonably strong positive link between employment and the IQI. However, both of these control variables also have a negative effect on the ITI and the SGI. This suggests that respondents who have graduated from high school, and respondents who hold some employment are much more sceptical of institutions in their society, and less satisfied with their government than their counterparts who have not completed high school, and who are not employed. Regarding the gender ratio, female respondents are more likely to have access to superior infrastructural quality and more likely to believe that the rule of law exists in society relative to their male counterparts. Older respondents are more likely to trust in institutions and have confidence that the rule of law exists than their younger counterparts. However, older respondents are also more dissatisfied with their government, and have inferior access to infrastructural facilities.

One final consideration based off the results of these models is that the overall model is much stronger at predicting the IQI than it is at predicting the ITI, the

RLI or the SGI. This is a surprising outcome because as Chapter 1 mentioned, the control variables were mainly chosen as a way to account for the subjective nature of several of the governance indicators. At 0.014 and 0.003 respectively, the R^2 values for the RLI and SGI are notably low. What this indicates is that particularly at the subnational level, the multivariate linear models are not accounting for large portions of the variations in perceptions of the rule of law, or satisfaction with government. The remainder of this section reports on whether other factors such as group level predictors and nonlinearities influence how the density-governance relationship operates in contemporary Sub-Saharan Africa.

4.4.2 Main Results: Mixed-Effects Models

The key difference between the multivariate linear model and the mixed-effects model is that the latter model nullifies the assumption of "non-independence" amongst data points (Fox & Weisberg, 2018). As it relates to this specific study, whereas the classical regression model assumes that the country in which an Afrobarometer respondent is located is of no relevance, the mixed-effects model actually controls for the possibility that respondents from within the same country may share more characteristics than respondents from different countries. Therefore, the results reported in table 4.5 introduce the control variable of country at the grouping level of the model. This country variable is also nested by the PSU with which an Afrobarometer respondent is linked. The models reported in this section are all varying-intercept only models since the term Country/PSU is the only one measured at the group level. However, Appendix D also includes the results of varying-slope models wherein variations in population density are also measured at the group level.

TABLE 4.5: Summary of Mixed-Effects Regression Models (2015)

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
log(Population Density)	0.217***	−0.125***	−0.070***	−0.028***
	(0.007)	(0.008)	(0.007)	(0.007)
HS Graduate	0.124***	−0.368***	−0.117***	−0.007
	(0.009)	(0.027)	(0.026)	(0.025)
Some Employment	0.020**	−0.027	−0.047**	−0.064***
	(0.008)	(0.025)	(0.024)	(0.023)
Male	−0.010	−0.005	−0.045**	−0.015
	(0.006)	(0.021)	(0.020)	(0.019)
$Age^{-0.33}$	−0.027	1.077***	0.300***	−0.149*
	(0.030)	(0.099)	(0.093)	(0.088)
Constant	3.800***	4.199***	5.731***	4.316***
	(0.224)	(0.273)	(0.242)	(0.229)
Observations	42,837	42,837	42,837	42,837
Log Likelihood	−53,046.280	−96,031.500	−93,208.040	−91,321.300
Akaike Inf. Crit.	106,110.600	192,081.000	186,434.100	182,660.600
Bayesian Inf. Crit.	106,188.500	192,159.000	186,512.100	182,738.600
Note:				*p<0.1; **p<0.05; ***p<0.01

For all the governance indicators, the introduction of the country nested by PSU variable at the grouping level weakens the coefficient reflecting the log-transformed population density's causal influence on these indicators. Population density remains strongest as a positive predictor of infrastructure quality (coefficient of 0.217), and weakest as a negative predictor of the SGI (coefficient of -0.028). Population density continues to be a negative predictor of all three "subjective" measures of governance quality, even with the introduction of the group level control variable. The fact that the coefficients for population density is significant at the 90%, 95% and 99% confidence levels in all four models indicates that population density is a meaningful predictor of subnational governance outcomes in Sub-Saharan Africa in the contemporary setting.

The results of the mixed-effects models calculated across 27 Sub-Saharan African countries confirm that both null hypotheses outlined by the study in Chapter 1 are false and can be rejected in favour of the alternate hypotheses. The first null hypothesis presupposed that no meaningful relationship existed between variations in population density and variations in specific governance quality indicators at the subnational level in Sub-Saharan Africa. However, as the results presented in this chapter have shown, in all cases, population density is a statistically meaningful predictor of governance quality as measured in terms of infrastructure quality, institutional trust, perceptions of the rule of law and satisfaction with government. This relationship holds at the 90%, 95% and 99% levels regardless of whether the country of origin of an Afrobarometer PSU is controlled for or not. It is therefore clear that variations in population density at the subnational level do have a meaningful effect on variations in these four governance quality indicators at the subnational level in Sub-Saharan Africa.

The second null hypothesis held that there was directional uniformity in terms of the effect which population density has on governance quality regardless of whether the governance quality indicator under scrutiny is measured in more "objective" terms (the IQI), or more "subjective" terms (ITI, RLI and SGI). This null hypothesis does not hold because it is clear from the results presented in this chapter that while population density is a positive predictor of infrastructure quality, it is a negative predictor of institutional trust, perceptions of the rule of law and satisfaction with government. Thus, higher population densities at the subnational level are meaningfully linked with better infrastructure, but are also meaningfully linked with lower levels of institutional trust, lower perceptions of the existence of the rule of law, and lower satisfaction with government. While the strength of population density as a predictor variable is muted in terms of the severity of the effect – as seen in weakly positive and weakly negative coefficients – distinctive and statistically meaningful patterns do exist when studying the subnational relationship which exists between population density and governance outcomes in Sub-Saharan Africa. Later in this chapter, the study reports the results of the robustness check models which measured variation at the level of the 4,930 PSUs in the 27 sample countries – as opposed to the main models which measured variations across the 43,108/42,837 individual level Afrobarometer respondents.

4.4.3 Nonlinearities

In addition to identifying that there is a linear relationship between population density and the four governance indicators in the context of Sub-Saharan Africa, it is also useful to assess whether there are any nonlinear effects which are present in these relationships. More specifically, does the effect of population density on governance at the subnational level differ between higher density urban areas and lower density rural areas? To assess whether nonlinearities are present in this relationship, orthogonal polynomial terms for the predictor variable of population density up to the second order were included in the mixed-effects models[3]. The orthogonal polynomial terms are transformed so that the linear regressor of population density and the quadratic regressor of population density-squared are no longer correlated with one another (Fox & Weisberg, 2018). Table 4.6 reports a summary of the results of the regression models which include these polynomial terms.

[3] Third-order polynomials were excluded because the addition of them to the models was non-substantive in terms of them showing whether or not a meaningful nonlinear component exists within the density-governance relationship.

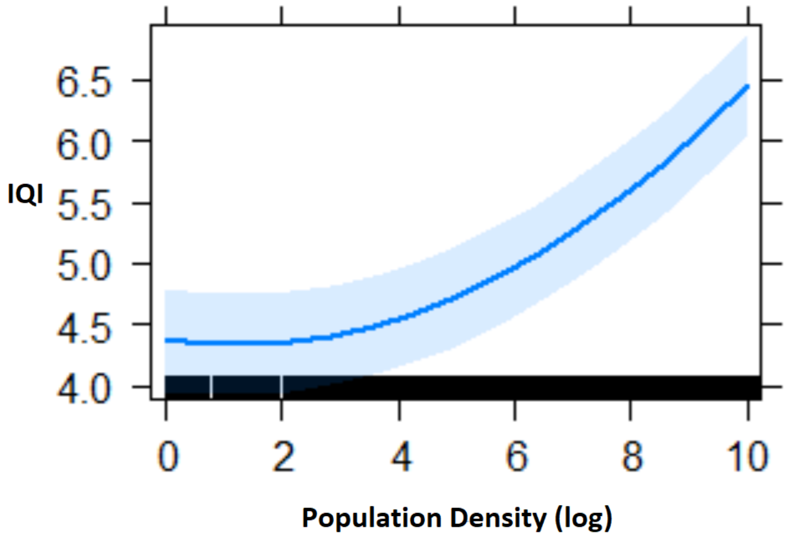
TABLE 4.6: Summary of Mixed-Effects models including polynomial terms (2015)

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
poly(log(Population Density), 1)	109.257***	−57.183***	−31.950***	−12.966***
	(3.060)	(3.464)	(3.008)	(3.306)
poly(log(Population Density), 2)	35.185***	−20.078***	−5.088*	−3.030
	(2.689)	(3.385)	(2.942)	(3.226)
HS Graduate	0.123***	−0.361***	−0.115***	−0.006
	(0.009)	(0.028)	(0.026)	(0.025)
Some Employment	0.019**	−0.027	−0.047**	−0.064***
	(0.008)	(0.025)	(0.024)	(0.023)
Male	−0.010	−0.005	−0.046**	−0.015
	(0.006)	(0.021)	(0.020)	(0.019)
$Age^{-0.33}$	−0.025	1.071***	0.298***	−0.150*
	(0.030)	(0.099)	(0.093)	(0.088)
Constant	5.047***	3.490***	5.329***	4.153***
	(0.216)	(0.268)	(0.239)	(0.224)
Observations	42,837	42,837	42,837	42,837
Log Likelihood	−52,953.100	−96,005.720	−93,198.440	−91,312.650
Akaike Inf. Crit.	105,926.200	192,031.400	186,416.900	182,645.300
Bayesian Inf. Crit.	106,012.900	192,118.100	186,503.500	182,732.000
Note:				*p<0.1; **p<0.05; ***p<0.01

The most noticeable difference between the results reported in table 4.6, and those reported in table 4.5 is that the coefficients and standard errors of the orthogonal polynomial terms for population density are much larger than the coefficients and standard errors for population density within any of the linear models. This large difference in coefficients is entirely the consequence of the orthogonal polynomial terms being uncorrelated with one another. Taken in isolation, the size of each population density coefficient is not as important as understanding how the different coefficients relate to one another. In this regard, the polynomial-inclusive models once again demonstrate that population density is strongest as a positive predictor of infrastructure quality, and weakest as a negative predictor of satisfaction with government. In the models where the IQI and ITI are the dependent variables, both population density terms are robust predictors of governance quality at the 90%, 95% and 99% levels. This indicates that there is both a significant linear component, and a significant nonlinear component to the density-infrastructure and density-trust relationships. When the RLI is the dependent variable, population density is a significant first-order (linear) predictor at all three levels of significance, but at the second-order (nonlinear), this effect is only significant at the 90% level. Where satisfaction with government is concerned, population density is a meaningful – albeit weak – linear predictor of outcomes. However, there is no meaningful nonlinear component to the density-satisfaction relationship across the sample.

Figures 4.3 to 4.6 depict the linear and nonlinear components of the effect which population density has on each of the four governance quality indicators. As the results reported in table 4.6 indicated, figures 4.3 and 4.4 show that the nonlinear components are strongest when considering the density-infrastructure and density-trust relationships. For the IQI, the nonlinear effect – as evidenced by more acute curvature – is strongest between lower log-form densities ranging from log(0) to log(4). In real-world terms, this range encompasses rural areas with densities between 0 persons per km^2, and densities of approximately 55 persons per km^2. Above this range, the relationship is much more linear indicating that at higher densities, the effect of population density on infrastructure quality is more clearly a positive one with higher densities being linked with better infrastructure. In practical terms, the increase shown in figure 4.3 between 4.5 and 6 on the IQI scale between log(4) and log(10) population densities – which corresponds to the difference between densities of 55 and 22,000 persons per km^2 – means that people living in higher density urban areas with densities of over 20,000 are expected to have access to around 1.5 more infrastructural facilities (see table 3.1 for a summary of facilities) relative to their counterparts living in lower density rural areas.

FIGURE 4.3: Nonlinearities Effect Plot: IQI



As Figure 4.4 shows, the nonlinear component when the ITI is the dependent variable is also clearer at lower densities, but at log-form densities greater than log(6)[4], this relationship becomes a largely linear, negative one. That is, respondents living in very high density urban areas are less inclined to trust institutions by about 1 scoring unit on the ITI relative to respondents who live in rural areas or less dense urban areas. Figures 4.5 and 4.6 show why the nonlinear explanatory components in the models including the RLI and SGI are weaker since the relationship which population density has with these two indicators is a slightly negative, predominantly linear one throughout.

[4]The approximate equivalent of log(6) is 400 persons per km^2.

FIGURE 4.4: Nonlinearities Effect Plot: ITI

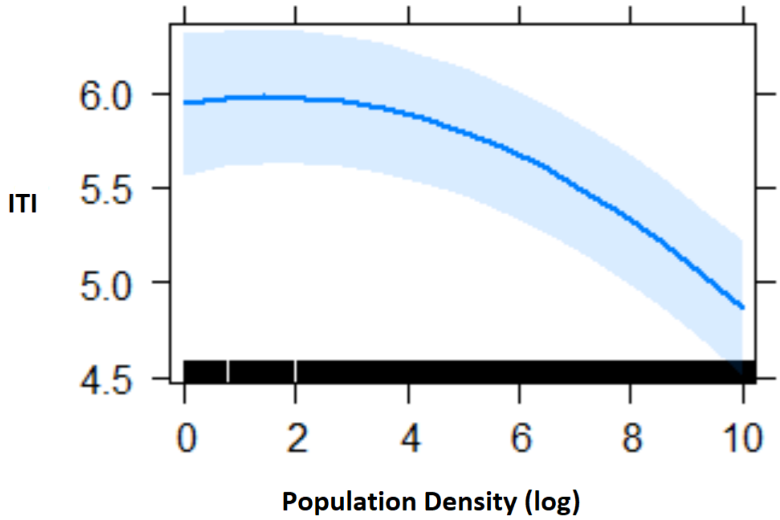


FIGURE 4.5: Nonlinearities Effect Plot: RLI

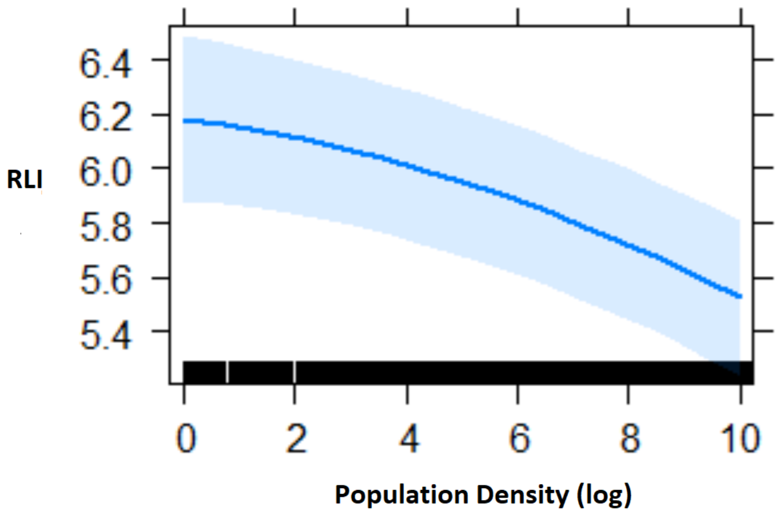
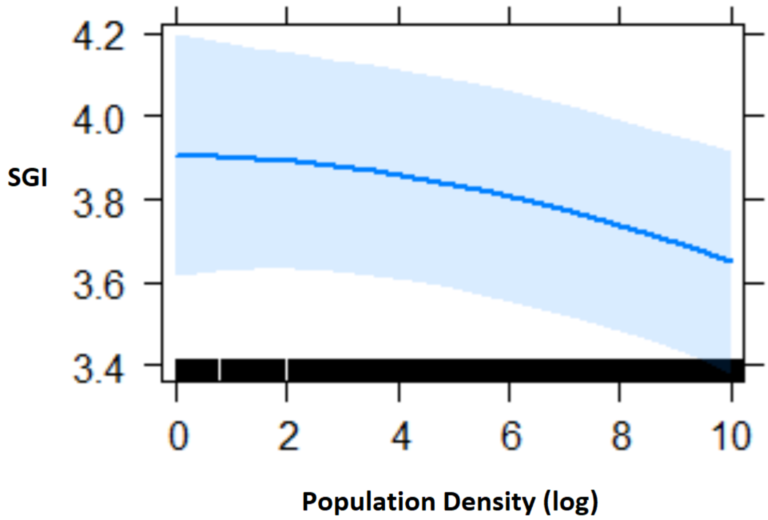


FIGURE 4.6: Nonlinearities Effect Plot: SGI



4.4.4 Country-Level Effects

It is also worthwhile to consider how the country-level random effects operate within each mixed-effects model. The introduction of the country control variable which is nested by PSU means that while there is still a "common slope" to the regression line, the point of intersection for the regression line now differs based on country. Since there are 27 countries in the sample, there would be 27 points of country-level intersection. Figures 4.7 through 4.10 report the variation in random effects by country for each governance indicator. It should be noted that since this study does not control for other country level factors such as income levels, economic strength and geographic characteristics, this section is limited to reporting which countries have the most notable divergences from the mean of all countries' intercepts – and does not provide a comprehensive understanding of why this is the case for specific countries. As such, this brief section is more descriptive than analytical.

Figure 4.7 depicts the random effects estimates for each country as per the results of the polynomial inclusive mixed-effect model results reported in table 4.6. Since the random effects scores reflect deviations from the corresponding fixed effect coefficient, the mean of the country scores displayed in this figure is 0. The individual country scores reflect the amount by which a country's actual score on the

IQI, differs from their expected score on the IQI given the model's component predictor variables. Consequently, countries where governance scores most exceed their expected scores have the largest positive intercepts, and the countries where governance quality scores fall most short of their expected levels have the smallest negative intercepts. For the IQI, the countries which most exceed their expected scores are Cameroon (2.68) and Botswana (2.54). The countries which fall short of their expected scores by the highest margins are Mozambique (-1.60) and Namibia (-1.47). Apart from these four countries, only two other countries – Gabon (1.04) and Uganda (-1.26) – fall outside the range of -1 to 1 when considering the difference between their expected performance and actual performance on the IQI.

FIGURE 4.7: Random Effects by Country for IQI

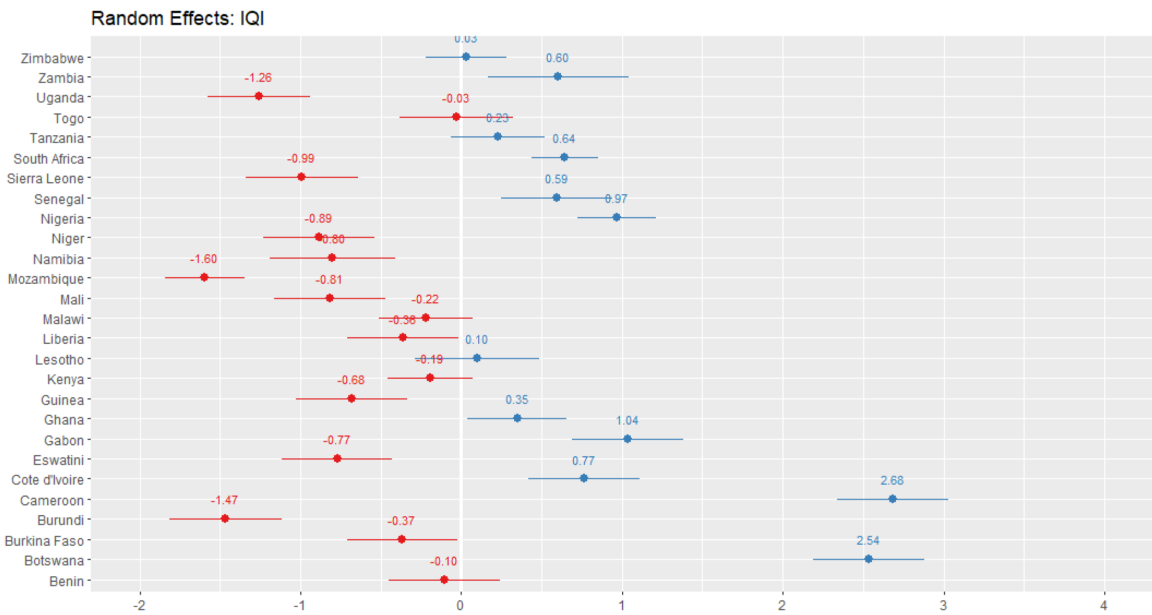


Figure 4.8 reports the random effects coefficients by country for the indicator measuring institutional trust across the 43,108 data points. Three countries do more than 1 unit better than they would have been expected to given their performances on the predictor variables and they are: Niger (1.73), Burundi (1.50) and Namibia (1.12). By contrast, four in-sample countries do more than 1 unit worse than they would be expected to on the ITI and they are: Gabon (-1.71), Liberia (-1.54), Nigeria (-1.37) and Togo (-1.03).

FIGURE 4.8: Random Effects by Country for ITI

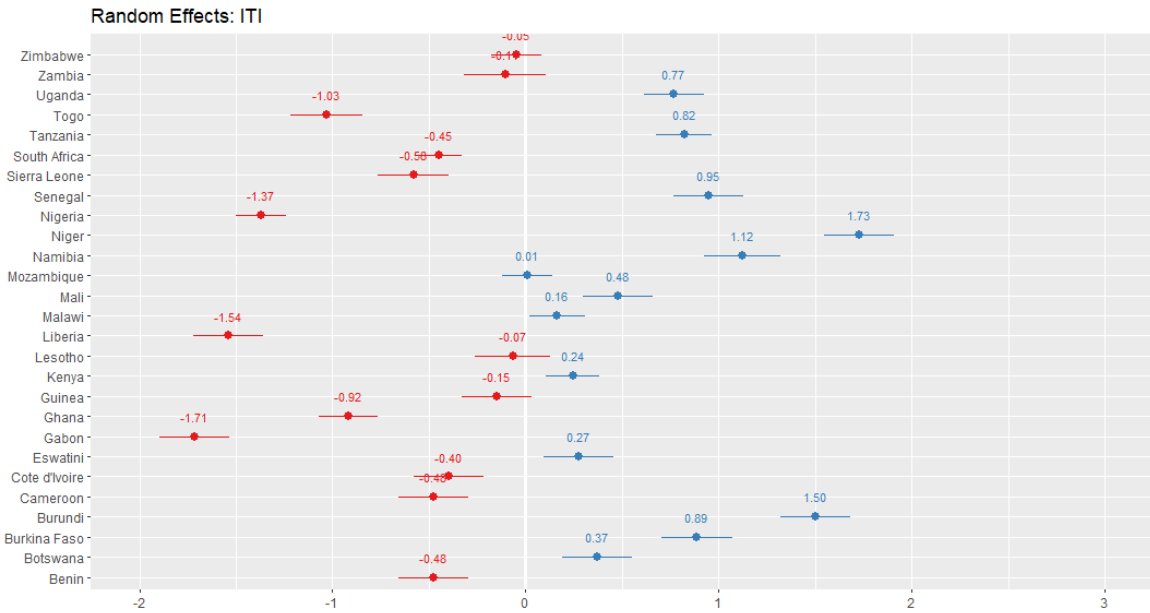


Figure 4.9 depicts how random effects vary by 27 countries in Sub-Saharan Africa for the Rule of Law Index (RLI) which measures the extent to which respondents believe that the rule of law exists in their society. The RLI scores within 23 of the 27 countries included in the sample largely conform to what their expected scores would be since they fall within the -1 to 1 range. The four exceptions to this are: Namibia (1.68), Botswana (1.37), South Africa (-1.39) and Gabon (-1.74). While Namibia and Botswana do much better than expected on the RLI, South Africa and Gabon do much worse.

FIGURE 4.9: Random Effects by Country for RLI

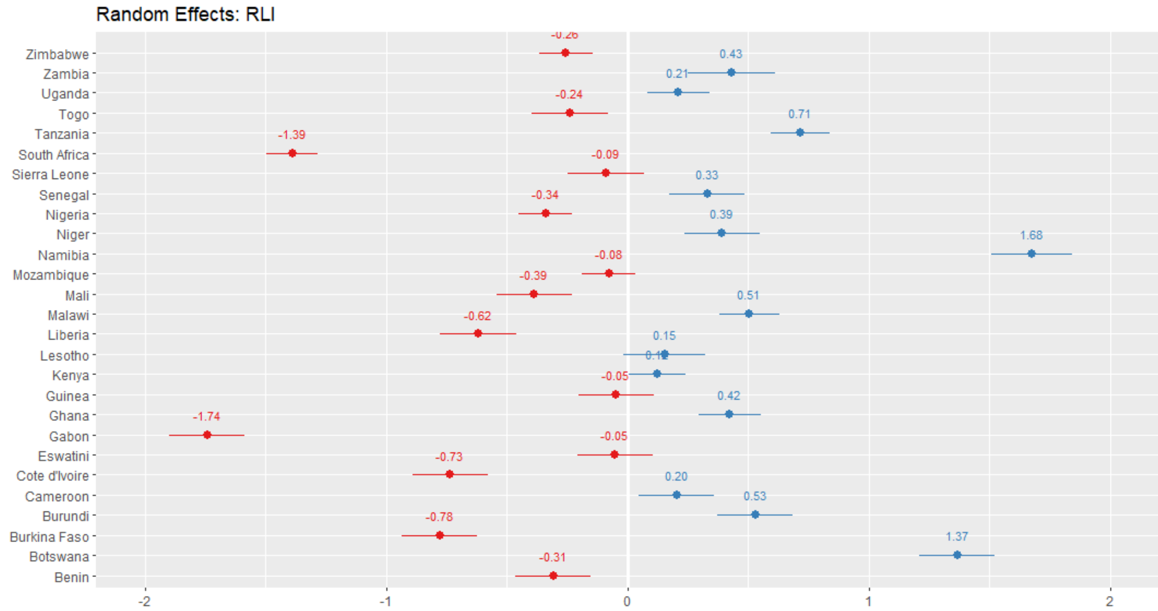
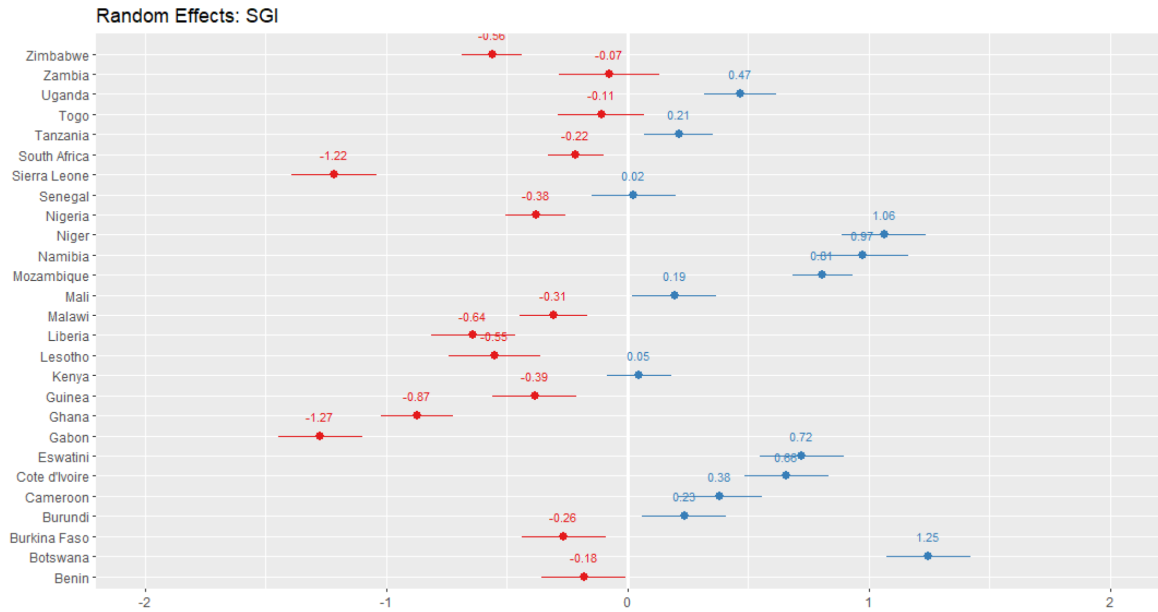


FIGURE 4.10: Random Effects by Country for SGI



The final figure showing the random effects by country is 4.10 which shows how random-effects differ by country when the dependent variable is the Satisfaction with Government Index (SGI). As the summary statistics section of this Chapter showed, the SGI was by far the indicator which indicated the lowest quality of governance, with the mean SGI score across all data points being only 3.79. This indicator is also the one in which there is the smallest difference between the highest random-effect score and the lowest random-effect coefficient indicating that even when country-level effects are considered, respondents across the 27-country sample are generally dissatisfied with how their governments are dealing with various policy action areas. The countries which most exceed expectations within this model are Botswana (1.25) and Niger (1.06). On the other hand, the countries who fall most short in relation to their expected performance on the SGI are Gabon (-1.27) and Sierra Leone (-1.22).

Across the four indicators, Botswana is the country where data points most consistently exceed their expected scores on various governance indicators. In fact, the only indicator where Botswana does not do at least 1 unit better than expected is on the measure of institutional trust – where the country's data points still exceed expectations by 0.37 units. Given that Botswana is known as a low population density country which has better levels of economic development, political stability and human development (Fosu et al., 2006), it is unsurprising to see it consistently outperform expectations within the country-level random effects component of the mixed-effects models. A more interesting case study may be Namibia. Namibia performs worse than expected on the more "objective" measure of government performance, the IQI (-0.80). However, data points located in Namibia also perform much better than expected on the "subjective" indicators of governance quality: the ITI (1.12), the RLI (1.68) and the SGI (0.97). Given that the models' results found population density to be positively related to better infrastructure, but negatively related to scores on the three subjective measures of governance quality, then Namibia as a low population density country in Sub-Saharan Africa fits this archetype very well. The country which most often falls short of expectations across the four governance variables is Gabon which does worse than expected in all four mixed-effects models.

4.5 Robustness Check Models

The main regression models used within this study all measured the subnational population density-governance relationship across 43,108 individual level data points which reflected 43,108 survey respondents from the Round 6 Afrobarometer survey. These main models found that while higher population densities in Sub-Saharan Africa were linked to better infrastructure quality, higher densities were also linked with lower levels of institutional trust, less satisfaction with government, and greater scepticism that the rule of law exists within a society. The vast majority of these findings were found to be significant at the 90%, 95% and 99% levels of confidence.

As a robustness check on these main results, the study now reports on the results of the robustness check models which changed the unit of analysis across which variations in population density and governance outcomes are measured. Within the robustness check models, the prime unit of analysis is the PSU with which a Round 6 Afrobarometer respondent is linked. Across the 27-country sample, there are 4,930 such data points. Table 4.7 provides the summary statistics for study's predictor and outcome levels at these 4,930 data points.

TABLE 4.7: Variable Summary Statistics for Robustness Check sample of 4,930 Afrobarometer PSUs (2015)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Education Ratio	4,930	0.31	0.28	0	0.1	0.2	0.5	1
Employment Ratio	4,930	0.36	0.27	0	0.1	0.4	0.5	1
Gender Ratio	4,930	0.50	0.06	0.00	0.50	0.50	0.50	1.00
Urban Ratio	4,930	0.40	0.48	0	0	0	1	1
PSU Average Age	4,930	37.07	6.69	21.12	32.25	36.38	41.12	67.75
IQI	4,930	5.05	2.58	0	3	5	7	10
ITI	4,930	5.54	1.51	0.35	4.53	5.61	6.60	9.64
RLI	4,930	5.78	1.24	1.17	4.98	5.83	6.62	10.00
SGI	4,930	3.79	1.26	0.03	2.89	3.79	4.68	9.00
Population Density (km^2)	4,930	2,067.59	4,624.26	1.00	67.23	270.42	1,664.02	54,695.13
Population Density (log)	4,930	5.71	2.23	0.00	4.21	5.60	7.42	10.91

It is notable that when compared to the summary statistics measured at the individual level (see table 4.1) the summary scores for the governance indicators at the PSU level are quite similar – especially when looking at the means and medians of the four governance indicators. Once again scores tend to be lowest on the satisfaction with government indicator, and highest on the measures for institutional trust and the rule of law. At 5, the median for the IQI across the 4,930 PSUs is identical to the median for the IQI across the 43,108 individual respondents. Perhaps the biggest difference between the different units of measurement is that the total range of scores across the governance indicators is smaller when considering the dataset with fewer units. The clearest sign of this is that the minimum and maximum values in the individual level dataset were 0 and 10 for each governance indicator, but at the PSU level, this is only the case for the IQI.

Table 4.8 reports the results of mixed-effects models which use the robustness check dataset comprising 4,930 data points to test the subnational link between population density and measures of governance quality across 27 countries in Sub-Saharan Africa. The direct comparison of these robustness check results is with the mixed-effects models displayed in table 4.5. The reason why the main models are still preferable to this robustness check model is that the robustness check model relies on the "artificial" PSU-level aggregation of the various variables. By contrast, the main models measure variability in governance outcomes across the original Afrobarometer unit of analysis – the individual respondent.

TABLE 4.8: Summary of Mixed-Effects Robustness Check Regression Models calculated across 4,930 Afrobarometer PSUs (2015)

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
log(Population Density)	0.271***	−0.108***	−0.064***	−0.039***
	(0.015)	(0.008)	(0.007)	(0.008)
Education Ratio	3.117***	−0.881***	−0.444***	0.070
	(0.149)	(0.083)	(0.071)	(0.079)
Employment Ratio	0.652***	−0.191**	0.015	−0.039
	(0.139)	(0.077)	(0.066)	(0.073)
Gender Ratio	−0.344	−0.197	−0.120	−0.375
	(0.496)	(0.275)	(0.236)	(0.262)
$Age^{-0.33}$	−1.232**	−0.398	−0.741**	−1.622***
	(0.616)	(0.342)	(0.293)	(0.326)
Constant	5.166***	7.472***	7.981***	7.628***
	(1.339)	(0.756)	(0.647)	(0.714)
Observations	4,930	4,930	4,930	4,930
Log Likelihood	−10,625.810	−7,735.659	−6,972.576	−7,496.502
Akaike Inf. Crit.	21,267.620	15,487.320	13,961.150	15,009.000
Bayesian Inf. Crit.	21,319.640	15,539.340	14,013.180	15,061.030
Note:				*p<0.1; **p<0.05; ***p<0.01

Using the smaller sample which measures the variables at a "higher" geographical/administrative unit has strengthened the causal influence of population density on both infrastructure quality, and satisfaction with government, but has weakened the predictive power of population density on institutional trust and the rule of law. As was the case in all previous models regardless of sample size or control variables, population density is a meaningful positive predictor of the IQI, and a meaningful negative predictor of the ITI, RLI and SGI at the 90%, 95% and 99% levels. It is also worth noting that higher rates of educational attainment within a PSU is an extremely strong predictor of infrastructure quality with a coefficient of 3.117.

The main implication of the results of the robustness check models is that they confirm that the effect which population density has on subnational measures of governance quality is meaningful in contemporary Sub-Saharan Africa. These results provide further evidence that neither the set 1 null hypothesis nor the set 2 null hypothesis can stand. This is because population density is a meaningful predictor of four governance quality indicators, but this effect is directionally different depending on whether a governance indicator is "objective" like the IQI, or "subjective" like the ITI, RLI and SGI.

4.6 Conclusion

This chapter presented the results of the investigation into the subnational relationship which exists between population density and governance indicators measuring infrastructure quality, institutional trust, perceptions of the rule of law and satisfaction with government. The results indicate that in accordance with the expected relationship outlined in the existing literature, higher population densities are meaningfully linked to better infrastructure quality at the individual level – even when controlling for other factors. However, contrary to expectations, higher population densities are also meaningfully linked to lower levels of institutional trust, greater scepticism that the rule of law exists within a society, and less satisfaction with government performance. The chapter also reported that there are some nonlinear components present in the density-IQI, density-ITI and density-RLI relationships. These nonlinear components are more distinct in rural areas which have lower population densities ranging from 0 to 300 persons per km^2, but beyond this range the linear component is much more prominent. The results of the robustness check models indicate that these findings hold even when one switches the main unit of analysis from the individual level to the Afrobarometer PSU level. The next chapter includes a discussion on the implications of these findings.

Chapter 5

Discussion and Implications

5.1 Introduction

This chapter considers the results of the regression models which were reported in Chapter 4 with respect to the broader questions which motivated this study. The chapter begins with a summary assessment of the standing of the research questions and research hypotheses which Chapter 1 outlined. Thereafter, the chapter includes a discussion which contrasts the ways in which population density affects “objective” and “subjective” measures of governance quality at the subnational level in Sub-Saharan Africa. The fifth section of the chapter considers the findings of this study and their implications for the existing body of scholarship on the topic of how the population density-governance operates in Africa. Finally, the chapter briefly discusses the implications of the study for policymaking in Sub-Saharan Africa.

5.2 Summary of Research Question and Research Hypothesis Evaluation

Chapter 1 outlined two research questions the study’s findings were meant to help answer. The first research question asked: "To what extent do variations in population density have a causal influence on variations in specific governance indicators at the subnational level in Sub-Saharan Africa?" The second research question was: "Does this influence differ depending on whether a governance indicator measures more "objective" aspects of governance quality, or if the indicator measures the "subjective" perceptions citizens have of governance quality within their society?" In order to have a reference point by which the study's findings could help answer the research question, two sets of hypotheses were formulated. The first null hypothesis assumed that population density had no meaningful causal influence on the four governance indicators. The second hypothesis presupposed that regardless of whether this effect was statistically meaningful or not, the directional effect which population density had on the four governance indicators would be uniform.

Based off the results of the regression models reported in Chapter 4, the first null hypothesis was rejected in favour of the alternate hypothesis. The results of the multivariate linear regression models showed that population density was a statistically meaningful predictor of infrastructure quality, the rule of law, institutional trust and satisfaction with government across 43,108 data points located in 27 Sub-Saharan African countries. Furthermore, the results of the mixed-effects models – which controlled for the country in which Afrobarometer respondents were located – also showed that population density was a meaningful predictor of all four governance quality indicators. The strength of population density's effect on these governance indicators was not overwhelming since at no point did the coefficients for population density exceed +/- 0.4 within linear models which controlled for other factors. The results of the models which controlled for nonlinear effects indicated that there is a clear nonlinear component to the density-ITI and density-IQI relationships – particularly at lower densities. By contrast, the nonlinear components within the density-RLI and density-SGI relationships are less distinct. Altogether, the patterns visible across the results of several models suggested that the null hypothesis of no meaningful relationship existing between population density and governance could not hold. The results of the robustness check models which measured variation across the 4,930 PSUs within the 27-country sample indicated that this relationship is robust even when one switches the unit across which one measures variations in population density and governance outcomes.

The set 2 null hypothesis was also rejected in favour of the alternate hypothesis. This is because the results indicated that population density affects "objective" indicators of governance quality in a different way than it affects "subjective" measures of governance quality. The study included four governance quality indicators, three of which were subjective measures since they were created using the answers respondents to the Round 6 Afrobarometer survey gave regarding their own perceptions of the performance of their governments. These three subjective measures measured institutional trust, perceptions of the rule of law, and satisfaction with government. The "objective" measure of governance measured infrastructure quality on the basis of whether or not an Afrobarometer interviewer and their field supervisor found the "objective" presence of a set of specific infrastructural facilities within an Afrobarometer PSU. Across bivariate, classical multivariate and mixed-effects models, population density was a positive predictor of infrastructure quality, but a negative predictor of institutional trust, the rule of law and satisfaction with government. More specifically, higher population densities were linked with better infrastructure, but also linked with lower levels of institutional trust, greater scepticism that the rule of law existed within a society, and lower levels of satisfaction with government. Since most of these results were robust at the 90%, 95% and 99% levels – regardless of sampling unit – the set 2 null hypothesis was rejected in favour of the alternate hypothesis. The next two sections include a discussion of what factors may explain this disjunct in the relationship population density has with objective measures of governance quality, and subjective measures of governance quality in Sub-Saharan Africa.

5.3 Population Density and Objective Measures of Governance

As previously stated, this study's findings indicated that in most cases, higher population densities were meaningfully linked with better outcomes in terms of infrastructure quality at the subnational level in Sub-Saharan Africa. In the context of this investigation, infrastructure quality was the sole governance quality indicator which did not rely on the subjective perceptions which citizens have of their own governments. One challenge which this study faced is that there are few "objective" measures of governance quality which are measurable at the subnational level. This is because most existing measures of governance quality are measured at the national level (Kaufmann et al., 2011). Moreover, Afrobarometer which was the main data source for this study is a survey which mainly samples the perspective of citizens. As a result, the questions within the Round 6 survey which were not answered by citizens, mainly pertained to infrastructure. The answers to these questions were captured within the Infrastructure Quality Index (IQI) which determined whether or not an Afrobarometer sampling location contained a specific set of infrastructural facilities (see table 3.1).

As one would intuitively suspect, superior performances on the IQI were heavily linked with whether or not a sampling location was located in a higher density, urban area. Urban regions, particularly in Sub-Saharan Africa are known to have better access to facilities such as piped water, paved roads, cellular services and schools (Mveyange, 2015). Higher levels of educational attainment were also a powerful predictor of infrastructure quality. Since population density is positively correlated with both education and urbanisation (see table 4.2), it is not surprising that population density is also a positive predictor of infrastructure quality. This finding is also in-line with the expected relationship based on the existing scholarship on the population density-governance quality relationship in Africa. This scholarship largely argues that a common thread throughout African history is that rulers located in high density urban locales have often struggled to govern over the mainly rural and low density interior parts of their societies (Herbst, 2014). What was surprising in this study's findings was the disparity between objective conditions at the subnational level, and the actual perceptions of citizens living in the same areas. In short, why were better levels of infrastructure quality negatively linked with institutional trust, perceptions of the rule of law and satisfaction with government? The next section briefly examines this dynamic.

5.4 Population Density and Subjective Measures of Governance

This study investigated the relationship between population density and three subjective measures of governance quality calculated at the level of 43,108 Afrobarometer Round 6 respondents. These three subjective indicators were: the Institutional Trust Index (ITI), the Rule of Law Index (RLI) and the Satisfaction with Government Index (SGI). Contrary to expectations, the study found that population density had a directionally negative causal influence on these three variables. That is, higher densities at the subnational level were linked to lower levels of institutional trust, more cynicism that the rule of law existed within a society, and lower levels of satisfaction with government.

One potential explanation for this is found in the predictive value of the control variables included in the study. In particular, higher levels of educational attainment were linked with higher population densities. Yet, at the same time, greater educational attainment was a strong negative predictor of institutional trust and the rule of law. This indicates that in Sub-Saharan Africa, Afrobarometer survey respondents who have graduated high school and live in high density urban areas are more likely to mistrust governance institutions and be sceptical that the rule of law is maintained in their society than their counterparts who did not complete high school and live in low density rural parts of the country. This is despite the fact that these respondents who live in high density urban areas have access to superior infrastructure quality than those living in low density rural areas.

A plausible interpretation of this disparity is the possibility that respondents who have more formal education, and who live in high density urban areas may have higher expectations of both their local and national governments than their counterparts in low density rural areas. Furthermore, since respondents living in high density urban areas are much more likely to live in the same area as government officials, this close proximity to the operational aspects of governance may increase their scepticism towards local and national governance institutions, and their scepticism of the presence of the rule of law within their society. The causal links between educational attainment, population density, and satisfaction with government, while frequently negative are far less distinct in severity (see tables 4.2, 4.4 and 4.5). This is likely because scores on the SGI were generally very low regardless of sampling unit (mean of 3.79 on a 0–10 scale) and thus the predictor variables in this study had less variability within which to discern patterns.

These findings of lower perceptions of governance quality being present in high population density, more urbanised parts of countries somewhat mimic studies where social indicators rather than governance indicators are the outcome variable. More specifically, there are some studies which measure how variations in population density influences concepts such as "quality of life" (Fassio & Rollero, 2013; Greyling & Rossouw, 2017) and "perceived quality of neighbourhood" (Walton, Murray, & Thomas, 2008). These studies also tend to rely on survey-based indicators and often find that higher population densities are linked to lower perceptions of neighbourhood conditions or economic development – regardless of the objective conditions in these areas (Kanbur & Venables, 2005; Walton et al., 2008). Some of these studies use case study countries which are included in this study such as South Africa, and find robust relationships exist between higher population densities and higher crime rates, lower adult literacy rates and higher HIV prevalence rates over an approximately 20 year period (Greyling & Rossouw, 2017). Beyond the difference between measuring aspects of governance quality, and measuring social indicators, another difference between this study and these other studies is that this study uses a much larger sample of data points, and a much larger sample of countries. Nevertheless, the findings of this study do suggest that subjective measures of governance quality behave similarly to subjective social quality indicators when one is testing for the effect which variations in subnational population density have on variations in the way citizens perceive the quality of their own lives, and the quality of governance in their society.

5.5 Implications for Existing Body of Scholarship

In many respects, the findings of this primarily quantitative study lend further credence to the largely qualitative existing body of scholarship on the topic of the population density-governance relationship in Africa. This study sought to “fill” a void in the existing literature by using quantitative analysis and spatial analysis methods to determine whether subnational variations in population density have a meaningful effect on four governance indicators. The existing research on this topic has linked Africa’s historically low population densities to issues of state size (Vengroff, 1976), state capacity (Herbst, 2014), border security (Englebert et al., 2002), the prevalence of conflicts (Raleigh & Hegre, 2009) and issues of economic development (Van de Walle, 2009).

Using a dataset comprised of 43,108 data points reflecting 2015 data, this study found support for the existing literature’s thesis that Africa’s population density patterns do influence governance outcomes across the continent. This is particularly the case when one considers how variations in population density influence variations in infrastructural quality even when controlling for other factors such as educational attainment and national-level effects. The literature suggests that Africa's historically low population densities have meant that rulers through the pre-colonial, colonial and post-colonial periods have struggled to fully exert authority over sparsely populated lands (Herbst, 2014). In accordance with this argument, this study found that there continues to be subnational effect whereby higher population densities in Sub-Saharan Africa are meaningfully linked with better infrastructure quality, while the opposite applies for low density areas. Thus, when it comes to the subnational relationship between population density and an objective measure of governance quality such as infrastructure quality, this study's findings solidify this existing literature.

The results of the statistical models where subjective measures of governance quality were the outcome variable have more complex implications for the existing scholarship on the population density – governance relationship in Africa. These results found that contrary to expectations, lower population densities at the subnational level are meaningfully linked to higher levels of institutional trust, greater confidence that the rule of law exists within a society, and greater satisfaction with government. These results appear to be similar to more recent scholarship on how variations in population densities influence the perceptions of citizens regarding the state of "social quality" in their own locales. There is often a disjunct between these perceptions held by citizens in higher density urban areas – which tend to be more cynical – when measured against the actual living conditions in these areas which are often superior to those experienced by their less dissatisfied counterparts who reside in lower density rural areas (Walton et al., 2008). Consequently, the implication of this study's results wherein subjective measures of governance were considered is that the existing scholarship may be correct that higher population densities in Sub-Saharan Africa are linked to some better material conditions in those areas, but the perceptions of the citizens themselves may not share those realities.

5.6 Implications for Policymaking

It is useful to consider some of the implications of the results of this study for policymaking across the continent. Since this study had a large sample of 43,108 data points spread across 27 Sub-Saharan countries, these implications are by nature generic. The results of the study revealed that there continues to be a disparity in the infrastructural facilities which citizens across Africa have depending on whether they live in higher density urban areas, or lower density rural areas. While these differences are understandable given the histories of many countries in Sub-Saharan Africa, the fact that there were still stark disparities as recently as 2015 highlights the continued need for national governments, and local governments to continue to work to improve living conditions within low density rural areas. Given the already low levels of satisfaction with government across the continent, such policy actions are necessary if governments across the continent are to maintain political stability and increase institutional confidence within their respective societies. Practical ways of dealing with this disparity will by nature have to reconcile with the many resource constraints which national and local governments across the continent continue to face.

The implications for policymaking of the study's findings on how population density patterns influence subjective perceptions of governance quality are not as readily clear. This study found that higher population densities were meaningfully linked with lower levels of institutional trust, and greater scepticism of the presence of the rule of law within a society. To a lesser degree, higher population densities were also linked with slightly lower levels of satisfaction with government. It is essential for policymakers to grapple with the disjunct which exists amongst citizens who tend to have more formal education and live in high density urban areas, yet who display less institutional trust or confidence in the existence of the rule of law in their society – despite these citizens having access to better infrastructural facilities than their counterparts with less formal education who live in low density rural areas.

While this finding is broadly in-line with research from other parts of the world which suggest people living in high density areas are more sceptical of societal quality (Walton et al., 2008), this disjunct is particularly concerning in the African context because of the rapid demographic changes which much of Sub-Saharan Africa has experienced over the last half-century. Given that these changes which include rapid population growth and rapid urbanisation are projected to continue for much of the 21st century (UN Habitat, 2016), it is imperative that policymakers strive to build institutional trust, work on maintaining the rule of law, and improve the level of satisfaction which their citizens have in their ability to deal with essential policy action areas like reducing poverty, reducing inequality, increasing employment opportunities and reducing crime. If governments at the national and local level are unable to adequately grapple with these problems, then there is an increased likelihood that more African countries will continue to be trapped in a rising tide of citizen dissatisfaction coupled with ongoing governance struggles.

5.7 Conclusion

This chapter discussed the implications of the results reported in Chapter 4 of this study's investigation into the subnational relationship between population density and governance quality in Sub-Saharan Africa. The chapter began by evaluating the standing of the research questions and research hypotheses in light of these results. Thereafter, the chapter considered why population density had contrasting directional influences on objective measures of governance quality such as infrastructure quality, and subjective measures of governance quality such as institutional trust, perceptions of the rule of law and satisfaction with government. The chapter then deliberated on the implications of these findings for the broader scholarship on how the population density-governance relationship operates in Africa. Lastly, the chapter briefly discussed some of the implications of the study's findings for policymaking in Sub-Saharan Africa.

Chapter 6

Conclusion

6.1 Conclusion

This study sought to determine whether variations in population densities measured at the subnational level have a meaningful effect on variations in four measures of governance quality: infrastructure quality, institutional trust, perceptions of the rule of law and satisfaction with government. The relationship was measured across 43,108 data points located in 27 countries in Sub-Saharan Africa using data from 2015. These data points represented 43,108 survey respondents included within Round 6 of the pan-African survey Afrobarometer. The study found that higher population densities at the subnational level are linked with better infrastructure quality, but are also linked with lower levels of institutional trust, a greater scepticism that the rule of law exists within a society, and lower levels of satisfaction with government.

This report began by providing a background to the investigation. The first chapter also outlined the purpose and significance of the study. Chapter 1 then listed two research questions and two sets of hypotheses. The opening chapter also discussed the limitations of the study and commented on the ethical considerations pertinent to the study.

The second chapter provided a theoretical background to the study in the form of a literature review. This chapter focussed on existing scholarship which has studied Africa's population patterns, and also studies which have sought to link Africa's historically low population densities to issues of governance and economic development. Chapter 2 noted that this existing literature is primarily qualitative in nature, and focusses more on historical demographic and governance processes on the continent as opposed to having a more contemporary approach. The intention was for the study's findings to help fill this "gap" in the existing literature by using statistical analysis and spatial analysis methods applied to recent data to determine whether there is an ongoing link between population density patterns and governance quality at the subnational level in Sub-Saharan Africa.

Chapter 3 gave a comprehensive overview of the research methodology used in this study. The chapter provided details about the multivariate linear, and mixed-effects models which were used to clarify the relationship between population density and governance in Sub-Saharan Africa. The third chapter also gave information about the investigation's sample, the data sources used to measure the variables and the sampling procedure of the study. Chapter 3 also identified how the four indicators used in this study to measure governance quality at the subnational level were calculated. These four indicators were: the Infrastructure Quality Index (IQI), the Institutional Trust Index (ITI), the Rule of Law Index (RLI), and the Satisfaction with Government Index (SGI). This chapter also provided a rationale for why a logarithmic transformation was applied to the measure for population density, and gave guidance on the selection of the control variables included within the regression models.

Chapter 4 presented the results of the investigation. The results indicated that higher population densities at the subnational level are linked with superior infrastructural quality, but are also linked with slightly lower levels of institutional trust, slightly more scepticism that the rule of law exists within a society, and slightly lower levels of satisfaction with government. The study found that population density is a statistically meaningful predictor of the IQI, the ITI, the RLI and the SGI at the 90%, 95% and 99% levels of confidence in both the multivariate linear models and the mixed-effects models. The study's findings also suggested that there are nonlinear components to the density-IQI, density-ITI and density-RLI relationships, with this nonlinear component more prominent at lower population densities. As a test for robustness the study measured for variations in population density and governance quality using a smaller dataset comprising 4,930 Afrobarometer Primary Sampling Units within the 27-country sample. The results of these robustness tests suggested that the causal effect which population density has on measures of governance quality meaningfully holds regardless of the unit of variation across which one measures this relationship.

Chapter 5 considered the implications of these results for the standing of the null hypotheses, the standing of the existing body of scholarship on the density-governance relationship in Africa and the implications for policymaking on the continent. Based off the findings of the study, both null hypotheses were rejected at the 90%, 95% and 99% confidence levels. The first null hypothesis was rejected since it posited that no meaningful relationship existed between population density and governance quality at the subnational level in Sub-Saharan Africa. The second null hypothesis was rejected because contrary to expectations, population density affects objective measures of governance quality such as infrastructure quality in a different way than population density affects subjective measures of governance quality like institutional trust, perceptions about the rule of law and satisfaction with government. In terms of accounting for this unexpected outcome, these results suggest that population density's effect on subjective measures of governance quality is similar to the effect found in other studies based off social surveys of representative samples of citizens (Walton et al., 2008). Regarding the existing body of scholarship on the density-governance relationship in Africa, this study's findings generally solidified the idea that population density does have a meaningful role to play in accounting for some of the variations in governance quality between high-density urban areas and low-density rural areas. As such, it appears that in addition to plausible arguments that low population densities have historically affected governance in Africa (Herbst, 2014), variations in population density continue to hold explanatory power in accounting for differing governance outcomes in the contemporary African context.

6.2 Suggestions for Future Studies

This report closes with some suggestions for future studies which seek to better understand how population density patterns influence governance outcomes at the subnational level in Sub-Saharan Africa. At a basic level, future scholarship should focus on identifying more objective measures of governance quality at the subnational level as this would help researchers understand the disjunct in expected directions of influence between subjective and objective measures which this study identified. These potential measures could make use of – subject to data availability – localised crime data, and proxies for the state of education and health governance such as school enrolment and completion rates, and life expectancies and infant mortality rates at the local or regional level.

Future research on this topic can also add to this study's work by looking at whether the subnational population density-governance relationship has changed over time. This study used one round of the Afrobarometer survey, and one set of population density raster data to study how the relationship operated in 2015. However, future research can enhance our understanding of how this relationship has operated over time. As of October 2020, there had been seven complete rounds of the Afrobarometer survey, while Round 8 is ongoing (Afrobarometer, 2020). A future study can use the same population density data source as this study – the WorldPop Project – and link this data to the geocoded versions of Afrobarometer data to measure how this relationship has changed over multiple years. Following the completion of the Round 8 survey, it will be possible for researchers to study the subnational density-governance relationship in Sub-Saharan Africa over a 20 year period from 2000 to 2020. Such research would be an invaluable contribution to understanding the more recent history of the population density-governance relationship, particularly given the rapid demographic shifts which have occurred on the continent during the early decades of the 21st century. The use of a longitudinal dataset would also enable researchers to better discern the precise nature of the reverse causality which the existing scholarship suggests may exist between population density patterns and infrastructure quality since such a study can also incorporate data about rural to urban migration patterns over time.

For future research which utilises a cross-sectional approach similar to the one used in this study, it would be useful to see whether the dynamic which was observed in this study regarding the density-governance relationship holds in other parts of the world. Specifically, it would be useful to understand whether the finding of population density in Sub-Saharan Africa being positively linked with infrastructure quality, but negatively linked with institutional trust, the rule of law and satisfaction with government is also the case in other parts of the world. There are some similar surveys to Afrobarometer which are carried out in other parts of the world such as Europe (Eurobarometer), and Latin America (Latinobarómetro). Research which links population density data to similar survey-based indicators calculated using data collected by these other surveys would help determine whether this study's findings on the population density-governance relationship are unique to Africa, or are also applicable to other parts of the world.

Appendix A

Sample Countries Information

TABLE A.1: Country and Country Code List

Country	Afrobarometer Country Code	ISO 3 Country Code	Number of PSUs	Number of Afrobarometer Respondents
Benin	BEN	BEN	150	1,200
Botswana	BOT	BWA	150	1,200
Burkina Faso	BFO	BFA	150	1,200
Burundi	BDI	BDI	148	1,192
Cameroon	CAM	CMR	150	1,182
Cote d'Ivoire	CDI	CIV	150	1,199
Eswatini	SWZ	SWZ	150	1,200
Gabon	GAB	GAB	149	1,190
Ghana	GHA	GHA	188	2,391
Guinea	GUI	GIN	150	1,200
Kenya	KEN	KEN	255	2,397
Lesotho	LES	LSO	118	1,200
Liberia	LIB	LBR	150	1,199
Malawi	MLW	MWI	214	2,400
Mali	MLI	MLI	150	1,200
Mozambique	MOZ	MOZ	300	2,400
Namibia	NAM	NAM	117	1,200
Niger	NGR	NER	150	1,200
Nigeria	NIG	NGA	300	2,400
Senegal	SEN	SEN	150	1,200
Sierra Leone	SRL	SLE	147	1,191
South Africa	SAF	ZAF	431	2,390
Tanzania	TAN	TZA	211	2,386
Togo	TOG	TGO	141	1,200
Uganda	UGA	UGA	176	2,392
Zambia	ZAM	ZMB	93	1,199
Zimbabwe	ZIM	ZWE	292	2,400

TABLE A.2: Population Density (km^2) Summary Statistics by Country (2015)

Country	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Benin	2,170.40	4,462.13	18.24	103.31	370.75	1,831.38	20,677.92
Botswana	364.17	569.39	1.00	3.63	83.83	618.29	2,742.78
Burkina Faso	1,159.02	2,212.56	24.67	76.56	186.27	845.41	7,287.21
Burundi	1,128.71	2,518.89	110.21	306.38	420.58	726.58	19,536.01
Cameroon	3,158.25	5,535.62	8.84	163.27	447.35	2,067.42	19,545.31
Cote d'Ivoire	4,995.76	8,161.90	11.43	295.10	2,140.59	7,023.42	45,115.78
Eswatini	528.93	662.82	10.54	82.94	165.59	666.30	1,974.21
Gabon	1,337.87	1,555.40	1.35	16.47	111.78	3,350.59	3,818.47
Ghana	3,210.02	4,931.67	3.87	321.99	884.05	3,511.68	19,899.36
Guinea	1,178.96	2,965.76	6.66	34.07	93.98	569.01	19,637.01
Kenya	3,460.07	7,939.61	1.00	279.93	918.41	3,089.00	81,184.61
Lesotho	492.13	926.00	12.60	34.85	103.72	292.56	4,123.88
Liberia	2,380.68	3,795.25	4.85	37.78	86.22	5,647.31	10,344.42
Malawi	183.65	268.31	1.00	52.04	202.45	252.42	3,163.10
Mali	3,633.32	7,360.62	1.05	73.24	217.93	1,137.93	30,377.46
Mozambique	1,700.26	3,114.11	2.43	63.96	267.65	1,448.08	25,847.70
Namibia	1,446.33	3,515.12	1.00	16.98	118.47	1398.57	16,229.29
Niger	552.07	1,171.59	3.20	66.54	141.03	360.43	6,839.74
Nigeria	3,090.64	6,037.46	9.75	195.20	800.45	3,028.94	41,641.67
Senegal	5,062.97	8,163.08	5.43	118.75	415.02	6,189.75	32,078.69
Sierra Leone	4,005.08	7,207.87	8.82	54.56	114.34	2,618.84	22,071.89
South Africa	3,301.34	4,742.87	1.00	73.37	1,414.16	4,672.56	36,837.85
Tanzania	1,792.05	4,446.61	1.69	82.24	256.13	1,207.53	40,546.57
Togo	2,046.63	2,953.08	10.50	122.31	555.00	1,684.07	10,328.06
Uganda	1,183.58	2,424.24	15.12	183.72	323.45	583.36	13,756.87
Zambia	981.91	2,055.35	3.52	16.87	75.17	817.19	13,673.76
Zimbabwe	1,087.26	1,797.22	4.28	27.63	65.51	2,243.70	8,042.49

Appendix B

Population Density Rasters Additional Information

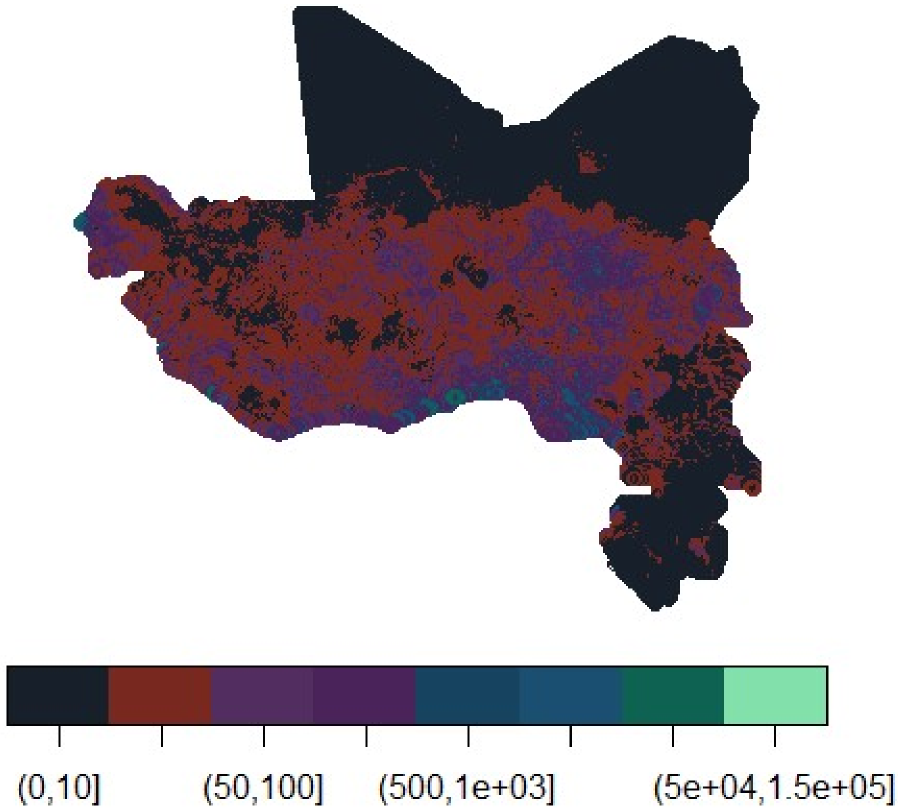
TABLE B.1: Summary Statistics of 2015 Population Density in km^2 Raster Data for 27 African country sample

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Population Density (km^2)	14,743,996	53.30	385.82	0.00	0.41	6.11	31.25	143,471.20

TABLE B.2: Number of Data Points per 2015 Population Density Category within Raster Data for 27 African country sample

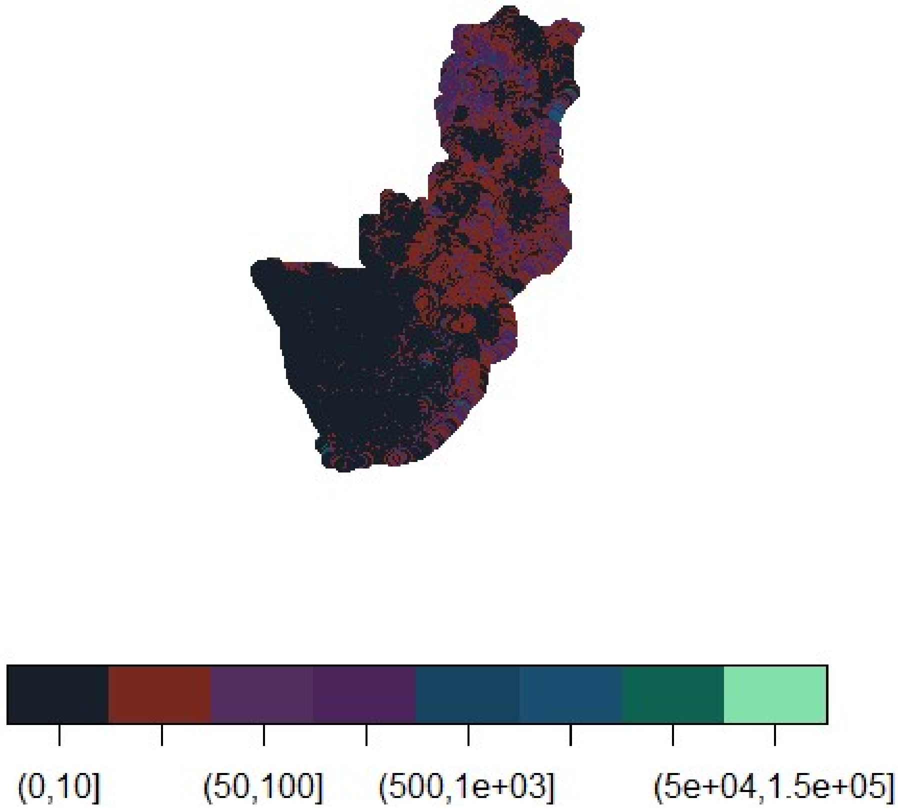
Population Density Category (persons per km^2)	**Number of Points per Category**
0–10	8,258,494
10–50	3,783,898
50–100	1,211,776
100–500	1,187,464
500–1,000	125,322
1,000–5,000	73,789
5,000–50,000	11,495
50,000–150,000	40

FIGURE B.1: Population Density in km^2 Map of West African Countries in Sample (2015)



The West African countries included in Figure B.1 are: Benin, Burkina Faso, Cameroon, Cote d'Ivoire, Gabon, Ghana, Guinea, Liberia, Mali, Niger, Nigeria, Senegal, Sierra Leone and Togo.

FIGURE B.2: Population Density in km^2 Map of South and East African Countries in Sample (2015)



The South and East African countries included in Figure B.2 are: Botswana, Burundi, Eswatini, Kenya, Lesotho, Malawi, Mozambique, Namibia, South Africa, Tanzania, Uganda, Zambia and Zimbabwe.

Appendix C

Spatial Representations of Governance Scores

The dataset used to map the scores reflected in these maps was the one used for the robustness check models.

FIGURE C.1: Map depicting IQI scores at 4,930 Round 6 Afrobarometer PSUs

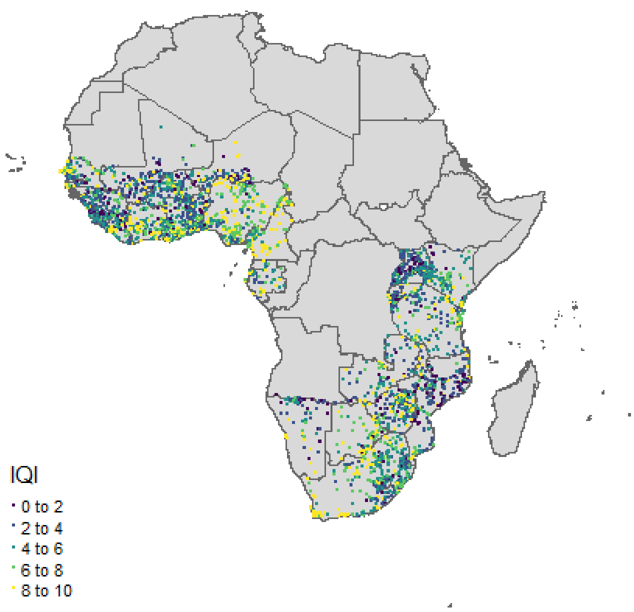


FIGURE C.2: Map depicting ITI scores at 4,930 Round 6 Afrobarometer PSUs

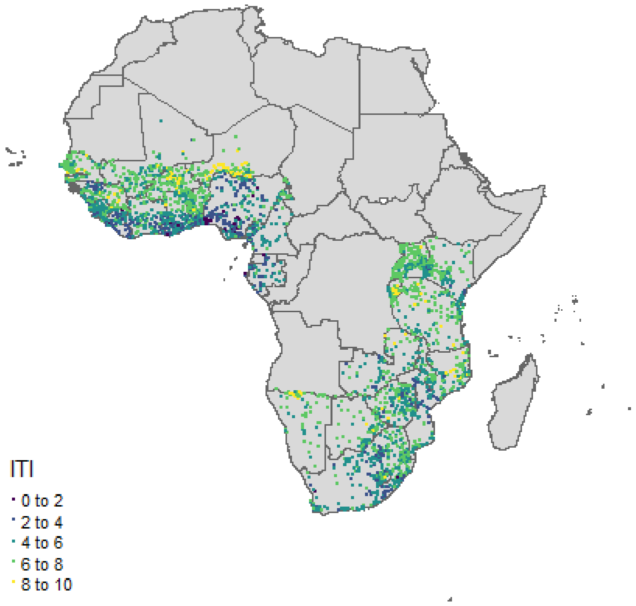


FIGURE C.3: Map depicting RLI scores at 4,930 Round 6 Afrobarometer PSUs

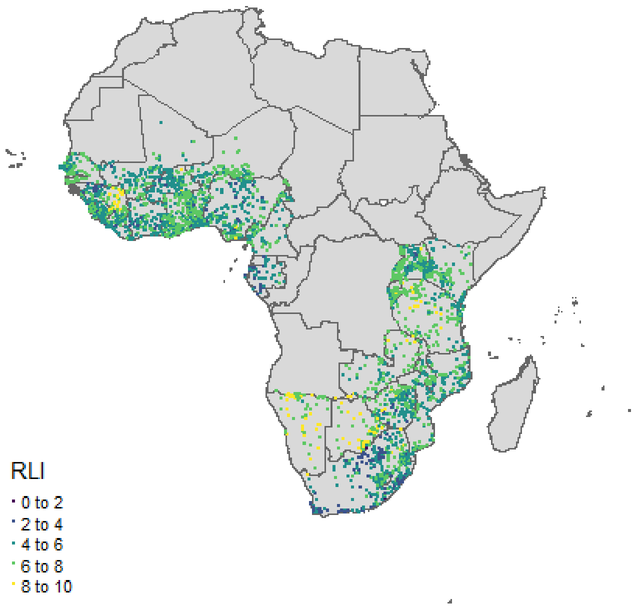
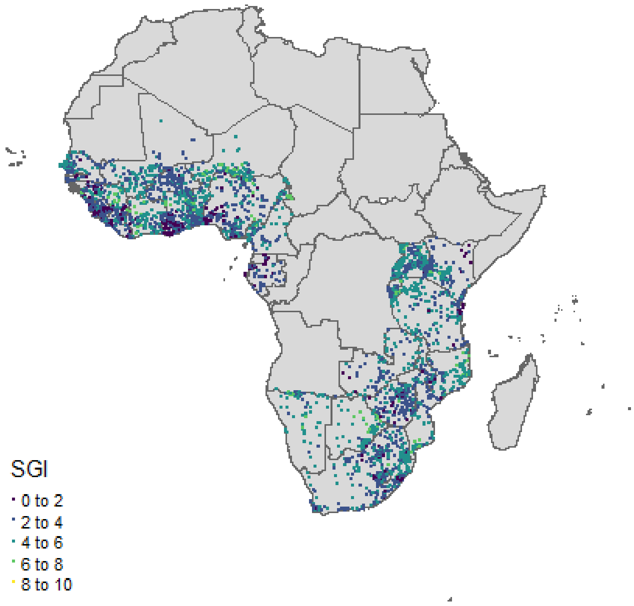


FIGURE C.4: Map depicting SGI scores at 4,930 Round 6 Afrobarometer PSUs



Appendix D

Additional Regression Models

D.1 Urban Inclusive Models

The models included in this section include a control variable denoting whether a respondent stayed in an urban and a rural area. Overall, introducing this variable slightly weakens the predictive power of population density across models and governance indicators and slightly strengthens the overall predictive power of the models.

TABLE D.1: Multivariate Linear Regression Models including Urban control variable

	Dependent variable:			
	IQI	ITI	RLI	SGI
	(1)	(2)	(3)	(4)
log(Population Density)	0.103***	−0.067***	−0.059***	−0.017***
	(0.006)	(0.006)	(0.006)	(0.006)
HS Graduate	0.708***	−0.588***	−0.203***	−0.002
	(0.024)	(0.028)	(0.026)	(0.025)
Some Employment	0.142***	−0.064**	0.094***	−0.090***
	(0.022)	(0.025)	(0.023)	(0.023)
Male	−0.053**	0.013	−0.047**	−0.019
	(0.021)	(0.024)	(0.022)	(0.021)
Urban	2.578***	−0.753***	−0.366***	−0.192***
	(0.025)	(0.029)	(0.027)	(0.026)
$Age^{-0.33}$	−0.117	1.200***	0.120	−0.387***
	(0.091)	(0.105)	(0.097)	(0.094)
Constant	3.409***	3.953***	6.152***	4.806***
	(0.193)	(0.222)	(0.205)	(0.199)
Observations	42,837	42,837	42,837	42,837
R^2	0.339	0.063	0.018	0.004
Adjusted R^2	0.338	0.063	0.018	0.004
Residual Std. Error	2.125	2.446	2.256	2.195
F Statistic	3,653.939***	479.437***	130.090***	27.626***

Note: *p<0.1; **p<0.05; ***p<0.01

TABLE D.2: Mixed-Effects Regression Models including Urban control variable

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
log(Population Density)	0.082***	−0.075***	−0.054***	−0.028***
	(0.006)	(0.009)	(0.008)	(0.008)
HS Graduate	0.084***	−0.331***	−0.103***	−0.007
	(0.008)	(0.028)	(0.026)	(0.025)
Some Employment	0.019***	−0.017	−0.043*	−0.064***
	(0.007)	(0.025)	(0.024)	(0.023)
Male	−0.006	−0.008	−0.047**	−0.015
	(0.006)	(0.021)	(0.020)	(0.019)
Urban	2.955***	−0.458***	−0.152***	−0.001
	(0.028)	(0.039)	(0.034)	(0.038)
$Age^{-0.33}$	−0.014	1.044***	0.287***	−0.149*
	(0.027)	(0.099)	(0.093)	(0.088)
Constant	3.413***	4.147***	5.714***	4.315***
	(0.197)	(0.269)	(0.242)	(0.229)
Observations	42,837	42,837	42,837	42,837
Log Likelihood	−47,955.830	−95,966.550	−93,200.730	−91,323.660
Akaike Inf. Crit.	95,931.660	191,953.100	186,421.500	182,667.300
Bayesian Inf. Crit.	96,018.310	192,039.800	186,508.100	182,754.000
Note:				*p<0.1; **p<0.05; ***p<0.01

TABLE D.3: Mixed-Effects Regression Models with polynomial terms and Urban control variable

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
poly(log(Population Density), 1)	39.710***	−35.838***	−24.514***	−13.430***
	(2.799)	(4.003)	(3.510)	(3.840)
poly(log(Population Density), 2)	7.217***	−9.864***	−1.522	−3.252
	(2.406)	(3.498)	(3.065)	(3.359)
HS Graduate	0.084***	−0.330***	−0.102***	−0.006
	(0.008)	(0.028)	(0.026)	(0.025)
Some Employment	0.019***	−0.018	−0.043*	−0.064***
	(0.007)	(0.025)	(0.024)	(0.023)
Male	−0.006	−0.008	−0.047**	−0.015
	(0.006)	(0.021)	(0.020)	(0.019)
Urban	2.946***	−0.425***	−0.147***	0.009
	(0.028)	(0.041)	(0.036)	(0.039)
$Age^{-0.33}$	−0.014	1.043***	0.287***	−0.149*
	(0.027)	(0.099)	(0.093)	(0.088)
Constant	3.890***	3.704***	5.404***	4.149***
	(0.194)	(0.265)	(0.239)	(0.225)
Observations	42,837	42,837	42,837	42,837
Log Likelihood	−47,943.430	−95,954.290	−93,192.450	−91,314.940
Akaike Inf. Crit.	95,908.850	191,930.600	186,406.900	182,651.900
Bayesian Inf. Crit.	96,004.170	192,025.900	186,502.200	182,747.200

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

D.2 Varying Slope Models

These are mixed-effects models which have both varying intercepts (by country/PSU) and varying slopes by population densities within countries. The results of the varying slope models are quite similar to the varying intercept models found in Table 4.5 and Table 4.6.

TABLE D.4: Mixed-Effects Models with Varying Slopes

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
log(Population Density)	0.407***	−0.149***	−0.073***	−0.028
	(0.061)	(0.028)	(0.014)	(0.022)
HS Graduate	0.116***	−0.366***	−0.117***	−0.002
	(0.008)	(0.027)	(0.026)	(0.025)
Some Employment	0.018**	−0.020	−0.044*	−0.063***
	(0.008)	(0.025)	(0.024)	(0.022)
Male	−0.009	−0.006	−0.046**	−0.015
	(0.006)	(0.021)	(0.020)	(0.019)
$Age^{-0.33}$	−0.024	1.070***	0.291***	−0.150*
	(0.030)	(0.099)	(0.093)	(0.088)
Constant	2.643***	4.368***	5.761***	4.324***
	(0.462)	(0.325)	(0.247)	(0.251)
Observations	42,837	42,837	42,837	42,837
Log Likelihood	−52,248.420	−95,954.410	−93,190.930	−91,274.050
Akaike Inf. Crit.	104,518.800	191,930.800	186,403.900	182,570.100
Bayesian Inf. Crit.	104,614.200	192,026.100	186,499.200	182,665.400

Note: *p<0.1; **p<0.05; ***p<0.01

TABLE D.5: Mixed-Effects Models with Varying Slopes and polynomial terms

	Dependent variable:			
	IQI	ITI	RLI	SGI
Mean Effects	(1)	(2)	(3)	(4)
poly(log(Population Density), 1)	186.342***	−64.828***	−33.022***	−12.152
	(32.225)	(11.438)	(6.280)	(9.425)
poly(log(Population Density), 2)	46.960	−27.444***	−11.480**	−7.311
	(36.345)	(6.464)	(4.569)	(6.456)
HS Graduate	0.113***	−0.359***	−0.114***	0.001
	(0.008)	(0.027)	(0.026)	(0.025)
Some Employment	0.019**	−0.018	−0.043*	−0.064***
	(0.008)	(0.025)	(0.024)	(0.022)
Male	−0.009	−0.006	−0.046**	−0.015
	(0.006)	(0.021)	(0.020)	(0.019)
$Age^{-0.33}$	−0.013	1.059***	0.284***	−0.153*
	(0.029)	(0.099)	(0.093)	(0.088)
Constant	5.022***	3.496***	5.346***	4.159***
	(0.233)	(0.265)	(0.240)	(0.223)
Observations	42,837	42,837	42,837	42,837
Log Likelihood	−51,700.600	−95,926.980	−93,175.720	−91,260.600
Akaike Inf. Crit.	103,431.200	191,884.000	186,381.400	182,551.200
Bayesian Inf. Crit.	103,561.200	192,013.900	186,511.400	182,681.200

Note: *p<0.1; **p<0.05; ***p<0.01

Bibliography

Acemoglu, D., Johnson, S., & Robinson, J. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, *117*(4), 1231–1294.

Achen, C. (2005). Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong. *Conflict Management and Peace Science*, *22*(4), 327–339.

Afrobarometer. (2020). Afrobarometer Data. Retrieved September 4, 2020, from https://www.afrobarometer.org/

Akyeampong, E., Bates, R., Nunn, N., & Robinson, J. (2014). *Africa's Development in Historical Perspective*. Cambridge University Press.

Alence, R. (2017). Political Science and the Study of Africa. *Oxford Bibliographies Online*. Retrieved August 20, 2020, from https://doi.org/10.1093%2Fobo%2F9780199846733-0082

Alesina, A., & Spolaore, E. (2005). *The Size of Nations*. MIT Press.

Angeles, L., & Neanidis, K. (2015). The persistent effect of colonialism on corruption. *Economica*, *82*(326), 319–349.

Austin, G. (2008a). Resources, techniques, and strategies south of the Sahara: Revising the factor endowments perspective on African economic development, 1500–2000. *The Economic History Review*, *61*(3), 587–624.

Austin, G. (2008b). The ‘Reversal of Fortune’ thesis and the compression of history: Perspectives from African and comparative economic history. *Journal of International Development*, *20*(8), 996–1027.

Badenhorst, L. (1951). Population Distribution and Growth in Africa. *Population Studies*, *5*(1), 23–34.

Bandyopadhyay, S., & Green, E. (2012). The Reversal of Fortune thesis reconsidered. *Journal of Development Studies*, *48*(7), 817–831.

Bates, R. (1987). *Essays on the Political Economy of Rural africa*. Univ of California Press.

Bolker, B. (2020). lme4 package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/lme4/index.html

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

Breheny, P. (2020). visreg package. Retrieved September 10, 2020, from https://www.rdocumentation.org/packages/visreg/versions/2.7.0/topics/visreg

Cooper, F. (1996). *Decolonization and African Society the Labor Question in French and British Africa*. Cambridge University Press.

Craig, J. (1984). Averaging Population Density. *Demography*, *21*(3), 405–412.

Day, A., & Day, L. (1973). Cross-National Comparison of Population Density. *Science*, *181*(4104), 1016–1023.

Dijkstra, L., & Poelman, H. (2014). A harmonised definition of cities and rural areas: The new degree of urbanisation. *European Commission Regional and Urban Policy Working Papers*, *1*.

Drakos, K., Kallandranis, C., & Karidis, S. (2016). Determinants of Trust in Institutions: Survey-based evidence from the European Union. *Regent's Working Papers in Business Management, Working Paper 1607*.

Englebert, P., Tarango, S., & Carter, M. (2002). Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries. *Comparative Political Studies*, *35*(10), 1093–1118.

Fassio, O., & Rollero, N., Cand De Piccoli. (2013). Health, quality of life and population density: A preliminary study on "contextualized" quality of life. *Social indicators research*, *110*(2), 479–488.

Food and Agriculture Organization. (2018). Guidelines on defining rural areas and compiling indicators for development policy. *United Nations Press*.

Fosu, A., Bates, R., & Hoeffler, A. (2006). Institutions, Governance and Economic Development in africa: An Overview. *Journal of African Economies*, *15*(suppl_1), 1–9.

Fox, J., & Weisberg, S. (2018). *An r Companion to Applied Regression*. Sage Publications.

Fox, J., & Weisberg, S. (2020a). car package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/car/index.html

Fox, J., & Weisberg, S. (2020b). effects package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/effects/index.html

Frankema, E. (2011). Colonial taxation and government spending in British Africa, 1880-1940: Maximizing revenue or minimizing effort. *Explorations in Economic History*, *48*, 136–149.

Frankema, E., & Jerven, M. (2014). Writing history backwards or sideways: Towards a consensus on African population, 1850–2010. *Economic History Review*, *67*, 907–931.

Gennaioli, N., & Rainer, I. (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth*, *12*(3), 185–234.

Gosh, T., Anderson, S., Elvidge, C., & Sutton, P. (2013). Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-Being. *Sustainability*, *5*, 4988–5019.

Green, E. (2012a). On the Size and Shape of African States. *International Studies Quarterly*, *56*(2), 229–244.

Green, E. (2012b). The Political Demography of Conflict in Modern Africa. *Civil Wars*, *14*, 477–498.

Greyling, T., & Rossouw, S. (2017). Non-economic quality of life and population density in South Africa. *Social Indicators Research*, *134*(3), 1051–1075.

Griffiths, I. (1995). *The African Inheritance*. Psychology Press.

Herbst, J. (2014). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press.

Hijmans, R. (2020). raster package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/raster/index.html

Hlavac, M. (2020). stargazer package. Retrieved September 10, 2020, from https://www.rdocumentation.org/packages/stargazer/versions/5.2.2

Hopkins, A. (2009). The New Economic History of Africa. *The Journal of African History*, *50*(2), 155–177.

ICPAC Geoportal. (2020). Africa - Admin Level 0 Shapefile. Retrieved September 1, 2020, from http://geoportal.icpac.net/layers/geonode%3Aafr_g2014_2013_0#more

Isbell, T. (2017). *Data Codebook for Round 6 Afrobarometer Survey*. Afrobarometer.

Kanbur, R., & Venables, A. (2005). Rising spatial disparities and development.

Kasara, K. (2007). Tax Me if You Can: Ethnic Geography, Democracy, and the Taxation of (a)griculture in Africa. *American Political Science Review*, *101*(1), 159–172.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and Analytical Issues. *Hague Journal on the Rule of Law*, *3*(2), 220–246.

Lloyd, C., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F., . . . Tatem, A. (2019). Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, *3*(2), 108–139.

Lloyd, C., Sorichetta, A., & Tatem, A. (2017). High resolution global gridded data for use in population studies. *Scientific Data*, *4*(1), 1–17.

Lüdecke, D. (2020). sjPlot package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/sjPlot/index.html

Mamdani, M. (1996). *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press.

Manning, P. (2015). Locating Africans on the World Stage: A Problem in World History. *Journal of World History*, 605–637.

McEvedy, C., & Jones, R. (1978). *Atlas of World Population History.* Penguin Books Ltd.

Min, B., Gaba, K., Sarr, O., & Agalassou, A. (2013). Detection of rural electrification in Africa using DMSP-OLS night lights imagery. *International Journal of Remote Sensing*, *34*(22), 8118–8141.

Mveyange, A. (2015). Night Lights and Regional Income Inequality in Africa. *UNU-WIDER Working Papers*.

Nieuwenhuis, R., Pelzer, B., & te Grotenhuis, M. (2020). influence.ME package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/influence.ME/index.html

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

Pebesma, E. (2020). Simple Features for R package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/sf/vignettes/sf1.html

Raleigh, C., & Hegre, H. (2009). Population size, concentration, and civil war. a geographically disaggregated analysis. *Political Geography*, *28*, 224–238.

Robinson, J. (2002). States and Power in Africa by Jeffrey I. Herbst: A Review Essay. *Journal of Economic Literature*, *40*(2), 510–519.

Sanny, J., & Asiamah, G. (2020). Trusted and influential: Religious and traditional leaders can be assets in COVID-19 fight. *Afrobarometer Dispatch Papers*, *Afrobarometer Dispatch No. 355*.

Socioeconomic Data and Applications Center. (2020). Gridded Population of the World, v4. Retrieved September 1, 2020, from https://sedac.ciesin.columbia.edu/data/set/gpw-v4-admin-unit-center-points-population-estimates-rev11/data-download

Stephan, G., & Tedrow, L. (1974). Tribal Territories in Africa: A Cross-Cultural Test of the Size-Density Hypothesis. *Pacific Sociological Review*, *17*(3), 365–369.

Tennekes, M. (2020). tmap package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/tmap/index.html

Tuszynski, J., & Dietze, M. (2020). caTools package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/caTools/index.html

UN Habitat. (2016). World Cities Report 2016. *UN Habitat: Nairobi, Kenya*.

United Nations. (1950). Demographic Yearbook, 1949–50.

United Nations Department of Economic and Social Affairs Population Division. (2020). World Population Prospects. Retrieved September 4, 2020, from https://population.un.org/wpp/

University of Southampton WorldPop datasets. (2020). Population Density/Individual countries 2000-2015 UN adjusted (1km resolution). Retrieved September 1, 2020, from https://www.worldpop.org/geodata/listing?id=77

Van de Walle, N. (2009). The Institutional Origins of Inequality in Sub-Saharan Africa. *Annual Review of Political Science*, *12*, 307–327.

Vengroff, R. (1976). Population Density and State Formation in Africa. *African Studies Review*, *19*(1), 67–74.

Walton, D., Murray, S., & Thomas, J. (2008). Relationships between population density and the perceived quality of neighbourhood. *Social Indicators Research*, *89*(3), 405–420.

Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/dplyr/index.html

Wickham, H., & Pedersen, T. (2020). ggplot2 package. Retrieved September 10, 2020, from https://cran.r-project.org/web/packages/ggplot2/index.html

World Bank. (2020a). Global Roads Inventory Project. Retrieved September 1, 2020, from https://datacatalog.worldbank.org/dataset/grip-global-roads-inventory-project-2018

World Bank. (2020b). World Night Light Annual Composite. Retrieved September 1, 2020, from https://datacatalog.worldbank.org/dataset/world-night-light-annual-composite-2015

World Bank. (2020c). Worldwide Governance Indicators. Retrieved September 5, 2020, from http://info.worldbank.org/governance/wgi