UNIVERSITY
of
GLASGOW

Essays on Cultural and Institutional Dynamics in Economic Development Using Spatial Analysis

by

Timothy Birabi

Submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy

Adam Smith Business School
College of Social Sciences
University of Glasgow
April 2016

Abstract

This thesis seeks to research patterns of economic growth and development from a number of perspectives often resonated in the growth literature. By addressing themes about history, geography, institutions and culture the thesis is able to bring to bear a wide range of inter-related literatures and methodologies within a single content. Additionally, by targeting different administrative levels in its research design and approach, this thesis is also able to provide a comprehensive treatment of the economic growth dilemma from both cross-national and sub-national perspectives.

The three chapters herein discuss economic development from two broad dimensions. The first of these chapters takes on the economic growth inquiry by attempting to incorporate cultural geography within a cross-country formal spatial econometric growth framework. By introducing the global cultural dynamics of languages and ethnic groups as spatial network mechanisms, this chapter is able to distinguish economic growth effects accruing from own-country productive efforts from those accruing from interconnections within a global productive network chain. From this, discussions and deductions about the implications for both developed and developing countries are made as regards potentials for gains and losses from such types and levels of productive integration.

The second and third chapters take a different spin to the economic development inquiry. They both focus on economic activity in Africa, tackling the relevant issues from a geo-intersected dimension involving historic regional tribal homelands and modern national and subnational administrative territories. The second chapter specifically focuses on attempting to adopt historical channels to investigate the connection between national institutional quality and economic development in demarcated tribal homelands at the fringes of national African borders. The third chapter on the other hand focuses on looking closer at the effects of demarcations on economic activity. It particularly probes how different kinds of demarcation warranted by two different but very relevant classes of politico-economic players have affected economic activity quite distinguishably within the resulting subnational regions in Africa.

Table of Contents

List of Tables

List of Figures

Acknowledgement

I would like to express my biggest gratitude to God almighty for keeping my mind, faith and hope going even when it appeared like there was nothing at the end of the tunnel. There were definitely moments alone in my office or with my computer somewhere (anywhere) where I questioned my ability to follow through with this PhD and often received renewed energy and ideas to carry-on. I am very grateful to God for coming through for me in those numerous lonely occasions.

I would also like to thank my parents, Dr Bennett Birabi and Dr Bridget Birabi, for always believing in my academic abilities and for all their encouragements over the phone and in person whenever they had the chance to remind me that I could do it. They have believed in my academic potential even when I didn't think there was anything in my brain and mind worth believing in. I am exceedingly grateful that they have invested so much, financially and emotionally, in me over the years. I can only say thank you very much, as there's no possible way I could repay them. I would also like to appreciate the rest of my family as well in this regard. They were very supportive every step of the way: giving me money whenever I was broke or listening to me whenever I had something to whine about.

I am mostly indebted to my supervisors, Dr Sai Ding and Dr Luis Angeles, for their unwavering support all through my authorship of this thesis. They are certainly the most instrumental people that have ensured my success at this stage through all their timely support reading, commenting, and mentoring me through the entire PhD process. I am certainly indebted to them both and I would like to say a massive Thank you.

I am also very grateful to the Adam Smith Business School for giving me the opportunity to study at such a remarkable institution. For the full three-year scholarship I received to enable me afford my studentship and living as a foreign scholar. I probably wouldn't have been able to be at this point right now in my career if not for such an amazing fortuitous opportunity. To this I am eternally grateful and I hope I can make the world a better place so that such an investment in me ripples unto the rest of humanity.

Finally, I am equally thankful to all the friends and colleagues I have made during my stay in Glasgow. If I could name you all here I would but I believe this isn't possible. I would particularly like to appreciate my life-partner Camille Ying Ying Guo for all her support through the last year of my PhD programme. I would

also like to appreciate Aldo Elizalde one of my oldest and closest friends in Glasgow for all our discussions and interactions before, during and post PhD thesis periods. He and his family have been a very welcomed presence in my Glasgow living experience. I would also like to say a big thank you to Zivile Zekaite, Filipa Fernandes, Amira Elasra, Udichibarna Bose, and every one of my colleagues that were very helpful and listened to me when I needed to discuss and share my challenges during my studies here. You are all in my heart and I would eternally appreciate the experience of studying with you here at the University of Glasgow.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

I also understand that my thesis may be made electronically available to the public. However, the copyright of the content herein belongs to the author. Any materials used or derived from this thesis should be acknowledged appropriately.

Signature:

Printed name: Timothy S. Birabi

x

x

Introduction

Economic development is a very broad multidimensional topic. The literature that attempts to understand the dynamics, patterns and sources of economic growth and development across the world is an almost saturated one with its inception dating as far back to works such as Adam Smith's "The Wealth of Nations" in 1776. Following this extensively researched literature, we find that the themes of history, geography, culture and institutions continue to permeate and persist in various aspects of the discourse. Understanding the sources of economic growth has remained central to unravelling the global patterns of economic development. While we agree on what the proximate growth sources are (human and physical capital accumulation, convergence and technological progress), we are still unclear about what fundamental growth sources are most relevant.

Rather than attempting to identify the most relevant deeper sources of economic development, it might be as fulfilling to understand economic growth dynamics by adopting frameworks that recognize specific fundamentals inherently operational to the localized patterns of development characteristic of one's study. While, variations in patterns of economic development manifest globally to produce comparable dissimilarities in the growth experiences of countries in most of Sub-Saharan Africa apart from most countries in Western Europe, for instance. The dynamics driving and protracting these differences are oftentimes undoubtedly local. Experiential locality is a very interesting phenomenon that can contextualise our study and understanding of economic development across nations.

Four prominent angles in the economic growth literature standout in this regard: institutions, culture, geography and history. These areas often provide economic growth analysts with a very practical and convenient mechanism for individualising and localising diverging growth experiences across nations while allowing us to still maintain our unwavering beliefs about the role of the proximate growth factors in the economic development of countries throughout the world. The most central and observable elements in these four areas are institutions and culture; and depending on how you perceive them they could easily be different representations of the same concept. The applied areas of history and geography on the other hand, provide economic growth analysts with very useful tools for

localising and contextualising divergent growth experiences across a very rich and wide global economic landscape.

What Institutions and Culture are Within the Context of Economic Development

It is not hard to see how the theoretical introduction and discussion of institutional and cultural dynamics in the economic growth of countries and regions cannot be easily separated. In the institutional change or evolution literature some authors such as Gérard Roland in his 2004 paper – 'Understanding Institutional Change: Fast-Moving and Slow-Moving Institutions', identify culture as in fact a slow-changing type of institutional manifestation. In many instances, the study and analysis of institutions and institutional change in economic growth are often consistent with Douglass C. North's definition of institutions as 'the rules of the game… or… humanly devised constraints…' which govern human interaction (North, 1990: 3; also see Hodgson, 2006 for more in this regard). The game as referred here involves both formal (de jure) and informal (de facto) forms of interaction between people(s) of society. Herein, our intention is not an attempt to redefine or reorganise the semantics about the definition and use of these easily ambiguous terminologies. It is not unconvincing to see how language can be considered as both an institution and a cultural element all at the same time. Or how the cultural identity of a group can at the same time represent and embody the political institutional administrative identity characterising that set of people. Such ambiguities can lead to nuances about definitional debates, which although very relevant, are not the objective of the exercise herein.

Our interest is in trying to position the reader within the school of thought opining that this inherent ambiguity concerning how cultural and institutional dynamics are interchangeably treated reflects the deep connections between the manifestations of their presence in human interactions; especially those about economic growth and development. One thing is clear in this literature: institutions are distinguishable from organisations (North, 1990; Hodgson, 2006). However, what institutions are remain more open to context and application than to definition. In many respects, the growth fundamentals of history and geography as often adopted in the literature provide a very useful avenue for contextualising institutions as current, modern or more formal 'de-jure' constraints to human

behaviour while cultures are often perceived as more informal 'de-facto' constraints lingering from the past for any specific set of individuals.

Regardless of the definition or perspective adopted, the fact is that institutional and cultural dynamics affect economic activity through a number of theorised conduits. A lot more of these theories have been analysed discursively compared to the volume of empirical work confirming them. This is very much expected that the theoretical framework most often precede the empirical feats within such a field of enquiry. As already highlighted in the preceding page, Douglas North – one of the pioneers of the New Institutional Economics (NIE) – gave one of the widely adopted working definitions of institutions. Continuing with his explanation, North (1990) opines that institutions as a result, are most effective through their ability to influence 'incentives (political, social, and economic) in human exchange'.

Irrespective of the perceived lack of consensus on a formal working definition for institutions, there is however a conceivable agreement on certain key components of the institutions concept. One of which is that institutions are man-made rules that act as either constraints or provide incentives to human behaviour and interaction. Although Hodgson (2006) gives some critique to North's perspective on institutions being ambiguous, he opines that this perspective predominantly assumes a conceptual similarity between the institutional concepts of constraints and rules. He adds that it is also more biased towards the frequent use of formal rules to the relative neglect of informal rules (Hodgson, 2006: 11). Authors such as Glaeser et al (2004) also fall into this critique of analytically over-emphasizing the constraints element of institutions. This does not necessarily mean that they do not imply that institutions are in fact also able to encourage human behaviour; such discussions may just well be an analytical focus, reflective of the kind of institutions being discussed and not necessarily representative of all institutional types. As already introduced, there is also an apparent consensus (at least by measure of content proliferation) that what constitutes rules can be decomposed into or viewed in *de-jure* and *de-facto* perspectives or as more regularly represented - *formal* and *informal* rules[1]. Consequentially, as an

[1] Boland (1992) refers to these in the form of concrete versus consensus institutions, where the former refer to more formal rules such as constitution and the other more informal rules such as unwritten laws and agreements (p.120).

analytical extension, the relevance of this rule decomposition especially features in the literature with applied institutional themes such as institutional change, enforcements and how different types of institutions inter-relate to effect desired developmental outcomes.

On the *institutions versus organizations* distinction, authors' perspectives appear to reflect the relationship between the social construct or structure (institutions) and the agents (individuals or social units) that make them up. North (1990) says that although both concepts provide an encasement or 'structure to human interaction' they are uniquely different as institutions set the rules of the game while organization set out the strategies through coordination by which interacting individuals achieve set out goals or objectives (pp. 4-5). In effect, a North' view is that *institutions are to the underlying rules for stable interaction* as *organizations are to the players of the game and their strategies for achieving a set purpose*. The distinction thus sets political parties, households, trade unions, and schools, the state, etc. (which are organizations) apart from the actual sets of mechanism ensuring stability, compliance, acceptance and predictability in the coordination of the individuals involved such as constitutional law, contracts, property rights, etc. (which form the institutional bit of the framework). Khalil (1999) provides a more comprehensive review of this analytical dichotomy between institutions and organizations. He also proposes that they are in fact different: that 'organizations adopt institutions ranging from simple routines to *beliefs and cultural ideals*[2] to assist its members achieve stated and sometimes implicit goals' (p.64). He goes on to simplify this distinction by saying that while organizations are defined by their goals or ends focus, institutions are the means employed (ibid). These two views are consistent and often shared by authors in the literature that agree on a conceptual distinction between both social structures even though there are slight differences on how these distinctions are made[3]. Some authors such as Aoki (2007) are not so clear on the distinction – referring to organizations as in fact a type of economic institution (p.1). While

[2] This phrase has been *italicized* to draw the reader's attention to Khalil's (1999) portrayal of institutions as incorporating elements of culture, which displays just how interrelated these themes are and why attempting a comprehensive view of the discussions can be quite herculean.

[3]Khalil (1999) reviews some of these alternative views on the distinction between institutions and organizations a bit further in the same section following his definition of both concepts (see from pages 61-62).

others especially those influenced by writings for everyday usage (such as contemporary dictionary definitions) in fact use both concepts interchangeably[4].

Similarly, another point of definitional divergence in the institutions literature represents one embodying an issue of classification. It is clearly treated in the text by Acemoglu (2009), in his chapter on institutions, where he attempts to distinguish between four fundamentals of growth – *'luck and multiple equilibria, geography, institutions, and culture'*. He proposes an explanation for each growth fundamental as defining the path for the proximate growth factors crucial for economic performance and the causes of the divergence between the growth experiences of countries across the globe. He points out that 'culture as viewed by some social scientists is a key determinant of the values, preferences, and beliefs of individuals and societies… they can be thought of as influencing equilibrium outcomes for a given set of institutions' (p. 122). Acemoglu later points out that although it can be confusingly interpreted that culture is indeed the most crucial, its role in impacting economic performance ought to be relegated to a less potent one compared to institutions for two reasons.

Firstly, because it is difficult to measure what constitutes good culture or social-capital. Secondly, because of the case with the Asian miracle evidencing an instance of diverging economic performances in the midst of a similar Asian cultural or social-capital system (ibid: 122-123). Authors such as Roland (2004) view institutions as a much grander socio-structural layout encompassing cultural elements. In fact he views culture as an embodiment of 'social norms and underlying values', a *'slow-changing'* kind of institution which is distinguishable from *'faster-changing'* institutional types such as political and legal institutions (p.117). This classification of institutional forms into fast and slow changing is also very much compliant with that of formal and informal rules, respectively. This idea mostly features in the discussions on institutional change (see Nye, 2008 for some discussions on the new institutional economics, NIE, in this regard). Greif (2006) proposes a definition broad enough to nest both debates as highlighted above. Hence making institutions appear as a 'super structure' under which such

[4] In this case, you often find the regular usage of institutions including universities when these are in the sense discussed herein referring to organization, while education itself is the institution or like the church and the actual religion conformed to. A similar example is applicable to the commonly assigned international financial institutions (the World Bank, the IMF, etc.), which are more or less big multinational organization (players as North, 1990 terms them) comprising of player countries as participants attempting to affect local institutional frameworks of finance, politics, etc.

social aggregates like organizations, culture, and other forms could be wholly classified and analysed. Greif's proposed definition is that 'institutions are 'a system of rules, beliefs, norms, and organizations that together generate a regularity of (social) behaviour' (as cited in Voigt, 2012: 5).

Introducing Some Ideological Perspectives

A careful look at the central issues as discussed above, one can subtly sense that the source of the debate in the literature is the lingering influence of the earlier views on institutions, now referred to as old institutional economics (OIE), on the current views of institutional influences in the current writings of the new institutional economics (NIE). The earlier OIE view had a broader concept of institutions and as a result considered the relevance of a broader influence of factors from social, political and economic inter-relatedness. While the later view as contained in the NIE approach are narrower and often discuss issues regarding rationality and optimization behaviours of agents and as a result often emphasize a more economic-centric perspective. Resultantly, issues on economic institutions such as property rights, contracts, and other market-based policies pervade these NIE discourses.

As an analytical consequence, we are not surprised that approaches oriented in the 1st category fall into the perspective of what Clark (1998) terms a *'Structures-based approach'*, which gives analytical primacy to the relevance of structures over agents. While those in the second category, which he refers to as *'Agency-centred'* analytically assume primacy of the agents over the structure. As Khalil (1999) diagnostically puts this divide, they are not necessarily two different approaches to viewing institutions - they are ways of approaching two different types of institutions. Such categorization may not also reflect a case of differing substance per se but may be one of shifting paradigms in areas of focus characteristic of the era of the development discourse itself. Earlier proponents of institutional economists such as Veblen and Commons viewed institutions as an inter-relationship between social structures and agents (Hodgson, 2006: 2). They emphasized the aspect of the relationship relating the social structure's influence on the agent: an approach that distinguishes the NIE approach from the OIE approach. This divide in perspective (OIE being collective oriented and NIE being individual oriented) comes through more emphatically in the analytical models

adopted and the policy implications sought after following findings from these respective researches. Hence, its theoretical relevance is most spelt out in the discussions about the potency of institutional based policy, institutional evolution, political institutional forms (democracy and autocracy), economic institutional forms (market-led and state-led systems) and how these affect economic performance than in the substance of the definitions of institutions itself.

As Davis (2009) points out: 'there is a broad consensus about what institutions are not' even though its qualification remains elusive (pp. 2-3). Paldam and Gundlach (2007) refer to two perspectives of the institutions-growth literature: the *'Grand Transition (GT)'* and the *Primacy of Institutions (PoI)* perspectives. Although the discussions under these differing perspectives are more on the relationship between economic performance and institutions, we bring it in here to buttress the point that what distinguish them is not what institutions in fact are but how they change, and how they affect economic growth. The former, GT, being more representative of the OIE perspective and looks at institutions from a broader narrative - more inclined to the notion of social structures, hence viewed as a consequence of economic improvements and not its inducer. While the latter, PoI, is more representative of the more recent NIE perspective which conceives institutions from a more specific oriented narrative - more inclined to the individuals and hence sees economic performance as a consequence of good institutional quality and inducements.

Concluding Remarks

What we have attempted to do so far is, provide some illustration of the lack of centrality in the literature on what institutions (or culture) are and how this translates to varying perspectives on theoretical orientations and analytical frameworks for studying its impact on economic development. Differing perceptions about institutions have reflected more the types and aspects of institutions being analysed rather than differences in the substance of the institutional concept itself. Our position in this thesis herein can be understood in a quote from Acemoglu & Robinson (2008: 2): '… ultimately, the aim of the research on institutions is to pinpoint specific institutional characteristics that are responsible for economic outcomes in specific situations (for example, the effect of legal institutions on the types of business contracts)'. We do not intend to

provide a re-fabrication of the theoretical framework for studying institutional dynamics on economic activity.

That said we hope the reader is able to acquiesce with our perspective that culture and institutions are very much interconnected by the specific historical and geographical contexts under which they are adopted to analyse economic outcomes and development. Take the notion of the geographical territories and boundaries defining a nation-state or any recognized political administration, for instance. Such boundary outlines are defined not only by the physical markers used to identify and distinguish them, but also by the formal and informal rules used to establish and project their local and global recognition. Similarly, norms, values, and other unique elements of culture have been observed and theorised to embody certain collective behavioural codes, which allow consistent clustered patterns usually harnessed in the analysis of economic developmental outcomes - typical of studies such as Michalopoulos (2012). And not surprisingly many institutional and cultural appropriations or manifestations are by their nature contextualised historically and geographically. Resultantly, this makes it near-inept to analyse the development impacts of either of these fundamental growth sources without an intersecting framework involving them all. We could fairly say that: *institutions and cultures are human-designed constructs that enable (or disable) any designated group of people in their interaction with one another as part of a broader society or in their interaction with both the environment (geography) and time (history) in order to ensure regularity in the behavioural outcomes characterising each individual in any of such interactions*.

It is for such reasons that this thesis attempts to firstly, analyse economic development from a global culturally interconnected spatial approach; and secondly, attempts to analyse subnational economic development by drawing on local territorial characteristics of historical ethnic or tribal homelands within the nation-states now spanning the African region.

Thesis Contribution

These two broad focuses of analysis (cross-national and cross-subnational) are treated in three empirical papers or chapters herein; within which the themes of culture, institutions, history and geography permeate. All three chapters (essays or papers) are each comprised of four core sections: an Introduction section, a

Literature Review section, a Methodology and Data section and finally, a Results Interpretation, Conclusion and Weaknesses section.

In chapter one we attempt to analyse economic growth across countries by adopting an augmented neoclassical framework, which incorporates both culture-defined and geography-defined spatial networks. The global dispersion of languages and ethnic groups across the world are used to build these cultural spatial networks. We illustrate these in parallel with a conventional border-neighbour contiguity spatial network model while controlling for arbitrary distance-based spatial clustering in all our models' residuals. The rationale for setting up this analytical framework is to achieve a global integrative design that could allow us analyse economic growth variations across countries using a realistic channel of rooted connections between such countries. Given that group cultural dynamics underpin a considerable amount of interactive behaviour that could be linked to patterns of global productive resource diffusions we attempt to statistically explore the extent of such spatial connections within economic growth across geographic-space and time.

Interestingly, we find after making these culture-defined spatial adjustments to the model that about a fifth of total economic growth accrues from productive-factor effects necessitated by such spatial integrative networks and systems. By incorporating this adjustment to our model we are able to provide a disentangled narrative of economic growth accruals from direct, indirect and total marginal proximate growth factors. Something we could not have been able to neatly infer without a model framework as this. Implications about how such integrative dynamics affect factor accumulation decisions for both poor and rich countries within interacting networks are additionally discussed. It is not impossible to see how poorer countries could easily be out-accumulated for human capital in a globally integrative system for instance, when returns to investments in such productive inputs would unquestionably be more secure in a more developed and stable economy to where they would readily and preferably flow towards.

In chapters two and three, our thesis focuses on analysing economic development within the African continent. Chapter two adopts a historical analytical context within which the effects of national institutions on the economic development of subnational tribal homelands demarcated at the periphery of nation-state boundaries are analysed. By narrowing our focus to the African economies of similar levels and experiences of economic activity, this

chapter minimises the empirical issues that characterise cross-continental and cross-country studies such as carried out in the chapter preceding it. The shared European historical experience, similar cultural dynamics and geographic attributes characterising these now amalgamated countries within this continent are exploited to provide a tangible identified impact of national institutional quality on economic development. This is especially achieved using novel alternatives for measuring economic activity at the subnational level that heavily rely on geographic information systems and the geographic spatial location of identified historic ethnic homelands and modern national territorial delimitations.

The results from the exercise of this second empirical chapter demonstrate that depending on the measure of economic activity adopted, national institutions are consistently positive and statistically relevant for economic activity. Particularly, we find that when the literature's previously used nighttime lights density proxy is adopted, the positive association is observed but the resulting coefficients are not statistically significant. While when our newly introduced average access roads length proxy is adopted, there remains the positive association and the coefficients are consistently and statistically significant. The former result is consistent with the literature adopting this data structure and design at the subnational level, while the latter result is consistent with the literature normally characteristic of the cross-sectional studies. These results do not necessarily contradict each other but demonstrate that both measures capture different aspects of economic development and of the socio-economic and political institutional functioning of these economies.

In chapter three, we assess how different types of tribal demarcations actually affect economic development at the tribal homeland level quite differently. We further attempt to illustrate that these experiential differences are as much aggravated at a pro-European level, as they are reflective of inherent local dynamics characterising either of these demarcation types. Tribal demarcations and forced amalgamations have been a very richly debated issue in African economic research; especially research on the sources of the protracted unfavourable distribution of global economic development towards African countries compared to the rest of the world. Knowing that boundary demarcations and delimitations were not just an act of the initial European colonial history but was also an action of national governments to enable them govern their respective internal territories, this empirical chapter exploits such a design to investigate

the impact of these various demarcation types on subnational economic development. Still exploiting the geographic and cultural dynamics characteristic of the African continent, we adopt the same economic development proxies used in the preceding chapter. However, the economic units of analysis differ from the preceding chapter in that instead of being country-ethnic homeland splits this chapter adopts a provincial-ethnic split. By going down to the state or provincial level of analysis and where this interacts with the location dynamics of historic tribal homelands, this chapter is able to bring economic development analysis closer to where it really matters for the individual experiences of local Africans.

In this final third empirical chapter, we find that tribal demarcations caused by national political actors (intra-national level) are consistently more negatively associated with underdevelopment compared to those affected solely by Euro-colonial inter-national level tribal demarcations. This is contrary to the Euro-colonialist sceptics' narrative that often proposes that the European scramble for African territories and resources left in its wake many issues that would later affect and continue to affect economic developmental efforts in these resulting modern African countries. The findings in this chapter demonstrate that this is not entirely the case. Of course the European scramble and its related activities left a considerable amount of challenges to the development potential of economic activity. However, we illustrate that some of these challenges can also be traced back to some local ethnic related dynamics – such as those related to local trust and tribal favouritism dynamics embedded within the politico-economic structure of subnational and national exchanges.

References

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton University Press.

Acemoglu, D., & Robinson, J. (2008). *The Role of Institutions in Growth and Development* (Commission on Growth and Development No. 10). *Commission on Growth and Development*.

Aoki, M. (2007). Endogenizing Institutions and Institutional Changes. *Journal of Institutional Economics*, *3*(01), 1. doi:10.1017/S1744137406000531

Boland, L. A. (1992). Chapter 8: Knowledge and Institutions in Economic Theory. In *Principles of Economics* (p. 233). London: Routledge.

Clark, W. R. (1998). Agents and Structures: Two Views of Preferences, Two Views of Institutions. *International Studies Quarterly*, *42*(2), 245–270. doi:10.1111/1468-2478.00081

Davis, K. E. (2009). *Institutions and Economic Performance : An Introduction to the Literature* (No. 09-51 - Law and Economics Research Paper Series).

Glaeser, E. L., La Porta, R., Lopez-De-Silanes, F., & Shleifer, A. (2004). Do Institutions Cause Growth? Edward L. Glaeser, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer * June 2004. *Journal of Economic Growth*, *9*(June), 271–303. doi:10.1023/B:JOEG.0000038933.16398.ed

Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge University Press.

Hodgson, G. M. (2006). What Are Institutions? *Journal of Economic Issues*, *XL*(1), 1–25.

Khalil, E. L. (1999). Institutions, Naturalism and Evolution. *Review of Political Economy*, *11*(1), 37–41.

Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review*, *18*(9), 1508–1539. doi:10.1016/j.micinf.2011.07.011.Innate

North, D. C. (1990). *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press.

Nye, J. (2008). Institutions and the Institutional Environment. In É. Brosseau & J.-M. Glachant (Eds.), *New Institutional Economics: A Guidebook* (pp. 67–80). Cambridge University Press.

Paldam, M., & Gundlach, E. (2008). Two Views on Institutions and Development: The Grand Transition vs The Primacy of Institutions. *Kyklos*, *61*(1), 65–100. doi:10.1111/j.1467-6435.2008.00393.x

Roland, G. (2004). Understanding Institutional Change: Fast-Moving and Slow-Moving Institutions. *Studies in Comparative International Development*, *38*(4), 109–131. doi:10.1007/BF02686330

Voigt, S. (2012). How (Not) to Measure Institutions. *Journal of Institutional Economics*, *9*(01), 1–26. doi:10.1017/S1744137412000148

Chapter One:

Economic Growth in a Culturally Interconnected Global Economy

Chapter Abstract

Understanding how countries grow and why the gaps between rich and poor countries persist is an age-old area of enquiry in the growth literature. Even in the current era of increased economic, political and social global integration these gaps still persevere and the spatial patterns of this divide follow similar unchanging routes. In this chapter, we seek to reassess the relationship between economic growth and its proximate factors (convergence, physical and human capital per capita) while adjusting for a form of global integration through culture, which we uniquely design. We do this by adopting a spatial econometric methodology to analyse cultural patterns of global integration embedded in the process of economic growth across countries. To achieve this, two broad specifications of culturally defined spatial weights are designed and used as an alternative to the often-used geographically defined metrics. This method extends the applied spatial econometric growth literature beyond regional level studies (that is, continental and provincial levels) to include a more open and inclusive global model.

Using a spatial autocorrelation panel growth model on a 122 countries spanning from 1990 – 2010, we find that after controlling for global integration in this manner and any arbitrary spatial patterns potentially remaining in the models' residuals: not only are the cross-sectional spatial coefficients (*rho*) statistically significant, there are also positive spatial correlations equally translate into statistically significant indirect spatial marginal growth effects. These indirect marginal effects range in magnitudes between 20% and 30% of the total marginal growth contributions depending on the cultural spatial weighting scheme adopted. As a result, these spatial effects equally affect the conditional convergence rate proportionate to the level of global integration. A number of important implications are deduced and discussed briefly from these findings. Among which is that these statistically substantial spatial effects might provide some illustrative initial evidence explaining why relatively poorer countries are on the average 'out-accumulated' in terms of access to relevant proximate growth factors, such as human capital. And why global economic growth catch-up remains a challenge to developing countries whose rate of convergence are not fast enough to outstrip the resource crowding-out effect which global integration brings in its wake.

Section 1.1 – Introduction

1.1.1 – Objectives

This chapter aims to conduct an exploratory enquiry of economic growth across countries by adopting a framework that controls for the presence of spatial interaction among countries on a global front. Two streams of theory particularly motivate this integrative approach. The first, relates to the locational/geographical economics literature while the second relates with the cultural and institutional economics literature. Although institutions are not of immediate focus herein, their theoretical relevance in functionally linking country geography, spatial distribution of cultural features, and the behavioural expectations of socio-economic (and political) agents with economic outcomes remain remotely relevant (see North & Weingast, 1989; North, 1990 & 2005; Starr, 1991; Beck et al, 2006). Also, depending on the definition and perspective considered, the approach and orientation herein can be of immediate relevance to the institutional framework. Within the institutions literature, economic growth is often linked to fundamental growth factors - geography, culture and institutions - via even deeper channels such as path dependence. However, path dependence remains an equilibrium concept that is mostly unknown and often relegated to the unobservable component of growth regression models (regression models' residuals or the error component). Dismantling path dependence itself is also not the aim of this work. This work seeks to augment the neoclassical economic growth with a relatable global network interactions framework. The spatial global network model adopted in this thesis chapter provides an inclusive framework for revisiting the conditional convergence and resource accumulation arguments; albeit from an integrative economic perspective beyond regional (either continentally or provincially) economic horizons. Resultantly providing an analytical context for individually assessing and comparing traditional proximate growth sources (physical and human capital inputs) from both direct (own-country) and indirect (cumulative cross-country) marginal effects perspectives.

The argument sustained throughout this work is that countries attain a position in the global long-run growth spectrum by the choices they make relative to those made by countries in their respective productive networks. These productive

networks are an encoded embodiment of country-level values, principles, and arrangements, which are difficult to observe directly through any particular index or through geographical proximity alone. By considering countries' slow-changing cultural characteristics while controlling for their fixed geographical attributes, it is possible to demonstrate this claim and further illustrate that country level interactions are increasingly transcending mere border proximity. This is empirically relevant because it is fathomable to imagine that the flows and patterns of productive resources between countries are indeed correlated with both observed forms of shared geographic boundaries and those consistent with cultural or ideological associations. Although it is equally necessary to understand why countries within same productive networks might exhibit similar or integrative global growth patterns, this paper does not yet examine this line of enquiry. This latter line of enquiry: controlling for, exploring, and dismantling the broader institutional and structural mechanisms is characteristic of the recent institutional growth literature. Our current study rather adopts a weak identification strategy using novel network designs adopted herein to provide an interesting means of assigning countries into integrative groups assuming integrative growth processes underpinned by geopolitical and cultural proximities, similarities and affiliations.

Although we have not approached the 'growth' question by attempting to identify specific channels of causation, the spatial specification as analysed provides a methodological improvement to the conventional non-spatially adjusted linear models previously adopted in the empirical growth literature (as demonstrated in LeSage & Pace, 2009; Elhorst, 2010). Our intention is to illustratively explore economic growth while assimilating geographic and cultural patterns of global interrelatedness or interconnections. These networks may certainly underpin or reveal other interesting growth processes, if and when unbundled deeper. As a result, further applications or interpretations of the results obtained herein are not ruled out even though they are not in themselves explored. Nonetheless, a number of variant network regimes are surveyed side-by-side so that the reader can get a sense of a multifaceted richness, which non-network adjusted models might not be able to reveal. The immediate objective herein is to assess economic growth across countries over reasonably short time intervals while controlling for similarities or structural proximities based on key cross-sectional growth fundamentals. This paper integrates time-invariant

elements of connectedness between countries on the basis of shared geographic contiguities, geopolitical and cultural resemblances or affiliations (using language and ethnic mapping schemes) while still maintaining the fixed-effects framework characteristic to cross country-time models. The notion of globalization becomes imminently relevant within this framework, as explored. Although looked at from differing perspectives before (see Dreher, 2006), globalisation remains a dimly explored subject in the spatial growth literature: especially not as treated in this thesis chapter. We propose herein that global cultural attributes are a crucial but weakly explored avenue of inter-national spatial relations in comparison to geographical proximity. Cultural characteristics had previously been adopted either locally to define spatial networks in the sociology literature or globally as additional index controls in the aspatial economic growth literature. Cultural elements are relatively time-invariant compared to other forms of global connections that may be observed[5]. As a result it can be adopted for specifying spatial networks in a similar manner as applied using geographical attributes or metrics. This is not to say that time-varying global integration specifications cannot be modelled; they are certainly feasible and are now being implemented using dynamic spatial model frameworks (see the literature reviewed by Elhorst, 2014: Chapter 4).

At this phase, this research seeks to answer the following research questions:

i. To statistically summarize and estimate economic growth from a global integrative perspective. By exploiting the cross-national distribution of cultural attributes using a spatial country-unit network assignment mechanism via a database of languages and ethnic groups from around the world.

ii. To specifically estimate a spatial panel growth regression model specified to accommodate a framework consistent with the concept of global integration. The intention is to disentangle own-country (direct) growth effects from those effects accruing from a 'factor share' of global inter-country integration.

iii. Finally to ultimately extend the application of spatial growth models beyond regional (both at the provincial and continental levels) growth specifications unto an equally relevant global perspective. This is especially important in the face of the increasing relevance of globalization on economic, socio-cultural, and political fronts. Where

[5] Other forms of global connections may include avenues driven by political, economic factors or other varying aspects of countries' social identities.

productive economic resources are increasingly permeating through country and continental regional boundaries and the notion of geographical proximities are becoming less significant.

1.1.2 – Scope and Contribution

By broadly focusing on disentangling direct and indirect global network effects in the economic growth framework, we attempt to apply the methodological contributions of Pace and LeSage (2009). So rather than extend the Solow-Swan economic growth framework to include the fundamental growth sources like Barro (1991) and his cohorts, we maintain the textbook Mankiw, Romer & Weil, (MRW, 1992) neoclassical structural framework. However, it is extended to include a globally integrative framework. We maintain this structural model because the convergence parameter is central to our discussions as later observed in section 4 of this work. A number of studies have attempted to assimilate cultural and geographic fundamentals into their specifications. However these have been applied using aspatial representations - either by controlling for these attributes as right-hand side indices or by using spatial techniques but limited to the adoption of geography-type network weighting configurations.

The practice of how spatial weights are defined continues to constitute an area of further contribution in the applied spatial econometrics literature (Abreu, Groot & Florak (AGF), 2005). While this remains true, methodologies for estimating and selecting appropriate spatial weights most suited to empirical data remains unrealized in the spatial econometrics literature (Elhorst, 2014). As a result, the convention is to estimate the chosen model using a number of specified spatial weights and see if the estimated coefficients are robustly unchanged, and then choosing the model that most closely predicts the selected data. Hence, adopting a multidimensional approach to spatial weight specifications certainly provides a useful extension to not only the aspatial economic growth models that adopt relatively weaker mono-dimensional spatial controls; but also to the applied spatial econometrics literature, by providing alternative and complementary ways of handling spatial correlations. The empirical and theoretical advantage of spatial model extensions are that they provide a framework to accommodate for increasing interaction or integration opportunities, where countries can be modelled in simultaneous (multiple) relations within globally specified networks.

The premise for applied spatial growth econometrics can remotely be justified from a literature dating back to 'Geographical economics'. Its underlying tenets have certainly evolved remarkably over the decades. Starting with the thoughts and writings on the matter by profound authors such as Montesquieu and Adam Smith, to those now embedded in more recent notions such as the 'equatorial paradox', for instance. A cursory look at the cross-country growth literature brings to the fore methodologies that endeavour to explain global growth divergence by specifying current growth experiences through an exogenous history of each country's previous economic and social experiences. Such experiences are fairly time-invariant and oftentimes geographically confined. These types of discussions can be found in the growth model augmentation[6] literature involving institutions, culture, and geography[7]. The methodologies adopted in these studies are however limited to non-systemic (non-integrative or aspatial) approaches. They also essentially assume countries grow in isolation of each other and are not part of an amalgamating whole or network process, where the progress or regress in a country can have implications for the other countries in interaction with them.

The central theme motivating this work is that: growth is not merely a localized concept, as previously simplified by country-based and regional-based studies (Fingleton & Lopez-Bazo, 2006). Although there are productive elements that are indeed of local relevance such as, the very institutional and structural compositions that uniquely affect each country in the world. There remain significant aspects of growth strategizing and implementation that go beyond country borders and internal operational patterns (see Quah, 1996b[8]).

[6] The term augmentation has been used in this chapter to represent add-ons to the economic growth model specifications. Specifically, such model extensions claim to improve on previous specifications by adding on additional variable(s) of interest to the right-hand of the standard growth equation.

[7] Some of the more influential papers within this literature are Bromley (1995) who suggests that studies of economic development theories tend to be biased resulting in empirical model misspecifications due to the 'teleological' differences of the Sub-Saharan African region from the rest of the transitioning economies of the world; Bloom, Sachs, Collier, & Udry (1998) as per geography; Acemoglu, Johnson, & Robinson (AJR, 2001) as regards geography, cultural, and institutional patterns but does this using an instrumental variable econometric procedure; Easterly & Levine (2003) as regards the direct impacts of geographical features on economic conditions and also indirect effects via institutions are equally explored; Olsson & Hibbs Jr. (2005) as regards geography and bio-geography – they claim this effect is strong directly and not necessarily via cross-sectional variations in institutions; among a host others in this huge literature.

[8] Although his discussions are centred in the European region, the rational can as well be extended unto a global horizon.

In this light, a model that focuses on either extreme of the cross-sectional or temporal spectra would insufficiently explain growth divergence intended for a globally interactive perspective. Particularly given that most countries' political strategies and economic models mostly stem from fairly clustered or global (external) homogenous sources. There is a spatial growth learning process, which cannot be separated from the logical progression of time or history; in fact, they are certainly all intertwined. Therefore working on either dimension to the neglect of the other would likely result in a case of empirical oversimplification or deliberate analytical omission (Garretsen & Martin, 2010). In global politics, the policies recommended to developing countries are oftentimes propagated by development organizations with cross-national economic and political influences. The paradigms of prevailing growth theories, models, and policy actions persist (or retard) globally, based on their successes (or failures) from the experiences garnered from leading economies of the world. As a consequence it is a weak position to limit cross-country growth studies to methods that assume the unobservable growth experiences (latent factors) among countries are disconnected across time. Or to limit such interconnectedness (whenever considered) to smaller network frameworks, such as relegating them to only countries within specific regions like the EU (European Union) or to provinces within specific countries. It is equally meaningful to view growth as part of a more dynamic and permeating global process in which the decisions to accumulate capital (whether physical or human) are motivated by both individual local marginal factor returns and global cumulative marginal returns.

This paper proposes an avenue through which interconnections or networks on a global front could be considered. The degree of country interconnectedness within our proposed framework can provide insight into the extent of global integration effects on economic development possibilities. By looking once more to the tenets of the location hypothesis and the spatial patterns of specifically defined global cultural elements, this paper seeks to empirically decompose the proximate growth factors into local and inter-national components. The link between institutional evolution, culture and geographic patterns provide the theoretical background for the choice of such a design or framework. In the past, fundamental growth factors were used as variables for localizing causality (as instrumental variables). We herein propose that these fundamentals could equally be used to illustrate how countries are spatially connected - as in an integrative

network. Thus providing growth analysts with additional empirical information contained in latent or unobservable growth sources embedded in these deeper growth factors. Just as highlighted earlier, the immediate focus of this exercise is not the identification or disentangling of the links between these fundamental growth factors and economic performance. This empirical study intends to directly observe the global nature and patterns of economic performances across countries based on a number of spatial networks specified using similar conditions pre-set by geographic and cultural allocations.

There are two broad categories of network designs used in defining the spatial connection between the countries sampled in this paper. The first follows the conventional geography-based distance and contiguity based spatial weight specifications. This is done mostly for the purpose of comparison or for reference. The second design, a key contribution of this study, seeks to pair countries based on similarities with respect to key globally distributed attributes of culture such as languages and ethnic groups. While the notion of neighbourhood is often viewed in a singular dimension - as geographical proximity - in practice, it is much more multidimensional. There is no reason to believe that spatial interaction at the country level is in practice solely restricted to geographical proximity. The notion of inter-country neighbourhood as defined by both locational nearness and attribute(s) similarity, can be found in studies as early as "Social Contagion and Innovation: Cohesion versus Structural Equivalence", by Burt (1987). Although a micro-level study from the medical sociology literature, this study nonetheless identified the distinction between the notions of cohesion and structural equivalence. The notion of cohesion defines the interaction between spatial units (in this case, individuals that are medical practitioners) in close geographical proximity. While, structural equivalence on the other hand, describes the kind of spatial interaction between paired units based on similarity in circumstances or qualifying attributes that may not necessarily be associated with geographic proximity. Combining[9] or complementing both network specifications provides us with a useful and relatively more informative avenue for handling relationships that may exist between countries either in or not in close geographical proximity of each other. Sub-Saharan African countries for instance, do not have the same

[9] In this case, interaction has been used in a numerical sense to imply case-to-case (matching) multiplication of network (spatial connection) specifications.

level of economic, political interaction or degree of integration with each other, as do countries within the European Union, where there is a lot more economic integration and there is free movement of productive resources. By inclining towards this analogy we demonstrate that geography is not the sole route via which inter-country integration can be measured and viewed from. Especially given that language and ethnic heritage could provide crucial tools for promoting behavioural and cooperative convergences and collaboration bridging from individual to global spheres of social, political and economic activities. Also with the advancement of technology and development, geographical space and distance is increasingly shrinking with reducing international travel times and improvements in real-time communication channels.

The key avenues of contribution for this thesis chapter can hence be summarised as follows:

i. We introduce and illustrate an empirical avenue through which the global geographic dispersion (positioning) of cultural attributes can affect the economic growth process. This avenue empirically puts to test the hypothesis that the global economic community is connected beyond direct economic channels such as trade openness or through political channels such as bilateral, multilateral, or other dyadic arrangements, which dominate the political economy literatures. We suggest that global cultural mappings can provide us with an interesting mechanism for understanding how economic performance is internationally distributed. Contrary to the belief that cacophonous splits between cultural (ethnic or language) and country nationalism could provide the catalyst for civil wars based on 'irredentist demands' (see Horowitz, 1985 and Michalopoulos & Papaioannou, 2011), it is possible that such international dispersion of identities could act as avenues for promoting economic integration and further global economic development in the long term. Exploring such narratives can be very informative for understanding how developed and developing countries would fair in such a globalised (or globalising) economic landscape.

ii. Specifically, we introduce two intriguing alternative spatial weighting schemes to the spatial growth literature. These help our research analyse how key factors in the neoclassical growth process behaves when countries are allowed to align or integrate globally with other countries based on shared ethnic groups or spoken languages. These cultural based networks are compared with conventional network designs based on geographic contiguity while controlling for arbitrary residual cross-

sectional dependence using a geographic inverse distance metric. This provides an analytical framework for testing the statistical relevance of globalization from a number of closely related perspectives consistent with the openness and political economy discourses. Globalization has remained mostly an economic (or politico-economic) phenomenon and by attempting to define it from differing perspectives, it gives the reader a multifaceted fresh experience of an already overly theorized phenomenon.

iii. We also attempt to empirically apply the interpretative contribution of LeSage & Pace (2009). Following their impact, we introduce the interpretation of spatial regression outputs by separating the main, direct, indirect and total marginal effects of the proximate growth factors. This approach is more informative as it does not look solely at the significance of the spatially interacted coefficients[10], which could be misleading in some cases. By following this approach our research is able to discuss the growth process involving topics such as convergence from a spatial perspective. Thus providing further empirical contribution to methodological discourses in the literature such as Fingleton & López-Bazo (2006).

1.1.3 – Overview of Chapter's Remaining Sections

In this chapter, three sections immediately follow this Introduction section: the first of them, section 2, briefly looks at a review of the literature on how space has featured in the economic growth discourse so far. The second, section 3, lays out the model specification and the variables involved in the central analysis. While finally, section 4 provides an initial summary of the data and reports of the results discussing the key findings in line with closely related aspects of the literature.

[10] This is oftentimes the estimated coefficient obtained from the spatial lag of the dependent variable in the spatial autocorrelation (SAC) model, for instance.

Section 1.2 – Literature Review

1.2.1 – Overview of Some Current Methodological Developments in the Economic Growth Literature

The panel economic growth literature has become increasingly popular over its cross-sectional counterpart since the dawn of the 21st century. This is particularly because it possesses an advantage of being able to control for unobservable country specific attributes. Its relevance and application has grown with the increasing availability of temporal dataset repositories from the 1960s such as the Penn World Table (Feenstra, Inklaar & Timmer, 2013); the Maddison project on historical world GDP figures; the World Bank databank; human capital and education data from the UN (United Nations) databases and the educational attainment dataset by Barro & Lee (2010)[11]. Such data availability has deepened empirical research at both country-specific (time-series) and panel (cross-country-time) horizons. One recurring theoretical premise that brings together these two perspectives of enquiry – long run cross-country studies and time-series or panel studies – is how the notion of convergence is treated. Durlauf, Johnson, & Temple (2005) in chapter 8 of the Handbook of Economic Growth (Volume 1A: p.599) point out that the time-series approach to the implied convergence concept, as embedded in the neoclassical growth framework, is predominantly 'statistical in nature'. They suggest that the methodology therein offers the empiricist the advantage of defining and testing more precisely a wider range of convergence hypotheses. Such 'statistical flexibility' provides the necessary framework for exploiting and exploring a methodological fusion of both cross-country and time-series approaches to the age-old theoretical discussions on global growth divergences or convergences across or between countries. Carlino & Mills (1993) – in which they distinguish between cross-country (ß-convergence) and time-series stochastic convergences – provide some useful insight into some expositions on the convoluted treatment of convergence in the growth literature. There are fundamentally two (broadly nested) core types of convergence typically discussed: ß-convergence and σ-convergence. Where the former relates to a

[11] Although the Barro-Lee educational attainment dataset is populated for 5-year intervals, it is often used in some annual time series studies where the years between each interval are filled using statistical imputation methods.

country's reversion to a hypothesized steady-state level of income, the latter relates to a narrowing of the income spread across countries through time (Durlauf, Johnson, & Temple, 2005). It is often the notion of ß-convergence that predominantly nests the closely related notions of stochastic-, deterministic-, club-, conditional-, and absolute- convergences[12].

In many respects, the methodological tenets of the panel growth modelling procedure are still quite nascent, nevertheless quickly growing. It is also a methodological procedure that still tails behind its relatively older time-series counterpart[13]. Although there have been some significant advancements in the panel data modelling literature - considering unit-root tests that control for cross-sectional dependence[14], for instance. There still remain a number of methodological complexities compared to time-series models that have the algebraic convenience of the construction of a single cross-sectional unit. In their discussion on the advancements made in panel unit-root and cointegration treatments, for instance, Breitung & Pesaran (2008) point out that panel cointegration is still at its infancy and is still at a 'residual-based' focus (p.280). These sort of methodological bottlenecks continue to pose challenges to the cross-sectional temporal discourse especially in applied empirical areas such as in the economic growth literature and in spatial models where the cross-sectional dimension is the focus and not its temporal equivalent. This also explains why temporal averaging (three-, five-, and ten- year) has become a rather popular convenience in the empirical panel literature and atemporal spatial weights are a welcomed alternative or empirical simplification.

More recent methodological developments in the empirical growth literature are increasingly providing tools for tackling the issue of cross-sectional dependence (Anselin, 2001; Pesaran, 2004). In this light, common factor models

[12] For readers particularly curious about the nuances of definitions and methodological interpretations on convergence, please see Islam (2003) – for a relatively more holistically informed review on the subject and Bernard & Durlauf (1995) – for a time series application to the concepts of stochastic and deterministic convergences.

[13] Take for instance, the manner in which error correction models are handled in the time-series framework versus how they are handled in the panel framework. The panel models are still at the stage reliant on a single equation Engle-Granger approach consistent with the Granger Representation Theorem, while that of time-series has advanced to the point of assuming a more general framework as proposed by the Johansen's Vector Error Correction Model (VECM) (Verbeek, 2004; Breitung & Peseran, 2008).

[14] This advancement has come to be generically referred to as 'second generation panel unit-root tests' by popular authors within the literature (see Breitung & Pesaran, 2008).

(following Bai, 2003), spatial models (following Anselin, 1988) and generally panel data models have increasingly been applied to the economic growth enquiry. Also to mention that, panel data models with adjustments for persistent dependent variables raise even more concerns when cross-sectional dependence is present or in suspicion (see Phillips & Sul, 2003). Economic growth researchers are now more alert to such empirical concerns and challenges. Hence adjusting growth models for both temporal convergence and cross-sectional interactions (or integration) is becoming commonplace in the empirical literature and in most teaching texts. Particularly when economic performance embodies both local and global perspectives, as is increasingly likely with the overwhelming manifestations in technological progress and inter-country connections over time. In this paper we believe that spatial models provide a useful framework not only suited to addressing the empirical issue of cross-sectional dependence, but also for surveying varying forms of such interactions or integration.

The roles of geography, history, culture and institutions in the growth discourse are also undoubtedly important and profoundly interwoven. A careful consideration of the approach and manner in which these individual areas of enquiry fit into the literature can very well inform and re-inform future contributions into the broader discourse. With the growing influence of spatial econometric methodologies and its increasing application in the re-investigation of the age-old growth hypotheses, there is increasing hope that these four very interrelated angles of enquiry can be more finely fused thus providing clearer understandings of growth patterns, its sources, including its implications for policy. While the roles of these growth fundamentals - geography, history, culture, and institutions - are pervasively unquestionable, their implication for policy and for accurately informing strategic economic growth planning remains a puzzle for developing countries. As a result, the literature on disentangling growth exploded quite remarkably in the 1990s and early 2000s. Initially, there was relatively more focus on the omitted variable dilemma with the Barro-type informal growth models (Barro, 1991) and later the attention shifted back to the formal growth models but with adjustments for elements such as human capital, social capital, regional spillover channels and etc. following extensions to the neoclassical framework as popularized by AJR (2001) and Glaeser et al (2004).

Researching the channels through which these deeper factors affect growth is a challenging venture and this has received its own share of significant enquiries.

For instance, geographical factors have been argued to be of diminished relevance in explaining economic wellbeing when institutions or policy efficiencies are accounted for (see the findings of Hall & Jones, 1999; Easterly & Levine, 2003, as examples). While in other instances, these same geographical attributes have been found as useful channels: for explaining and validating the geography-trade narrative (Redding & Venables, 2004); for explaining the geography- or climate-endowments-growth hypotheses through angles such as the disease exposure argument (Gallup & Sachs, 2001), its impact on agricultural productivity or on the quality of human capital (Bloom et al, 1998). All of these branches of enquiry share the position that fundamental growth sources are certainly quite important in understanding the path dependences that have defined and continue to define the converging and diverging growth experiences of countries in the global economic landscape.

Space is certainly a crucial element in these discussions and can provide a clearer understanding of the distribution of economic activities around the world. The notion of how space is defined, conceived and accounted for can certainly help shed light on how countries integrate and get confined into particular growth trajectories. It is in this light, that briefly exploring the manner in which spatial econometrics has evolved and has become integrated into the economic growth literature becomes unequivocally relevant.

1.2.2 – Economic Growth and Spatial Econometrics

Empirical variations with geography and space as earlier adopted in the economic growth literature relied mostly on models including such features as exogenous right-hand side controls (Frankel & Romer, 1999; McArthur & Sachs, 2001). These geography-type specifications controlled for spatial heterogeneity from a fainter perspective (one-dimensional conceptualization of space). It is in this light that they significantly deviate from the mainstream spatial econometric models, which are more multidimensional with their approach of coping with the spatial heterogeneity huddle. Typical spatial econometric models represent individual spatial units as a network or web of inter-relationships between sampled units. The resulting estimated spatial coefficients attempt to statistically ascertain the extent of these interactions, on the average.

The adoption of spatial econometrics in the growth literature has mostly been limited to explaining the regional diffusion of technological progress or innovation. These models attempt to assess spillover or indirect (interaction) effects from associated spatial units through the procedural advantage of geography-specific networks. Most likely, its application to the field of growth econometrics lends immediate relevance to the theoretical tenets of the endogenous growth (Romer, 1986; Lucas, 1988; Grossman & Helpman, 1991) and the Schumpeterian lines of reasoning[15]. Following earlier efforts to accommodate geographical dynamics into economic geography by influential authors such as Paul Krugman (1991), Fujita, Krugman, & Venables (1999) and the seminal contributions to the methodological application and interpretation of spatial models by Anselin (1988) and LeSage & Pace (2009), there has been a recent expansion in the literature attempting to augment existing theoretical models with geographically motivated spatial processes.

It is important to note at this juncture, that the term region (regional) as referred to herein represent spatial units defined either within country or between countries, but restricted to those in some form of economic union such as the European Union (EU) or within the same continent (such as Africa or South America). Initial papers in the applied space-growth discourse mostly focused on re-evaluating the endogenous growth models by endogenizing innovation for manufacturing output using a spatially specified network of producing units or countries. Anselin Varga, & Acs (AVA, 1997), for instance, re-evaluated and found evidence of both 'direct' and 'indirect' impacts (through private research and development) between university research and technology innovation for a geography-based knowledge production function for small businesses in the United States Market in 1982. While, Fingleton (2001a & 2001b) re-assessed regional productivity growth, finding that there is evidence of increasing returns which also 'varies indirectly via the rate of technical progress and labour efficiency variations', while his other work focuses on an increasing returns simulation model in the same region (EU countries) with some specific theoretical applications to the 'Verdoon's hypothesis'. Even with further insights into spatial methodologies

[15] These essentially involve discourses on the notion of spillover economics (knowledge spillover). See Quah (1996b) for discussions about some substantive relevance for considering geography based spillover effects when observing regional income dynamics. See also Martin & Sunley (1998) for a review and some key empirical issues as relevant to applied regional endogenous growth theory.

and the considerable improvements made in computational capabilities post Anselin (1988) and LeSage & Pace (2009), there have surprisingly been meagre adaptations to the neoclassical growth framework, of a comparable scale. This is probably due to a perceived lack of substantive interpretational relevance or theoretical justification for spatially adjusted models from a global cross-country neoclassical context. That is, a context beyond applications to regional and provincial (intra-country) growth exploratory frameworks, which are by their design easy to justify spatially given that the sampled units characteristically share common administrative boundaries or policy environments.

Abreu et al (2005) provide a comprehensive review of the literature on 'Space and Growth'. They illustrate how the growth literature has evolved from a spatial methodological concept of 'absolute location' (right-hand side adjusted and instrumental variable type models) based growth model approaches to one of 'relative location' (spatial econometrics) in the last few decades. However, they emphasize that both perspectives, though complementary, can learn from each other - 'absolute location' models can learn to be more methodologically rigorous when considering spatial dependence, while 'relative location' models could learn to be more theoretically motivated in how they specify their aspatial econometric models (ibid: 35; see also Corrado & Fingleton, 2012). Additionally, most of the spatial econometric adaptations analysing growth-led factors mostly adopt the spatial error models (SEM)[16] and spatial autoregressive model (SAR) specifications. In the broader applied spatial econometrics literature, a more general-to-specific approach is gaining appeal as Spatial Durbin models (SDM) and Manski type model specifications are becoming increasingly popular (see LeSage & Pace, 2009; Elhorst, 2010 & 2014). These spatial specifications are more general and as result analytically nest the parametrically smaller SEM, SAR and SAC (spatial autocorrelation) models. They provide the researcher a relatively more flexible opportunity for testing downwards to more parsimonious spatial frameworks (Elhorst, 2010: 13) and reducing the econometric likelihood of specification bias or variable omission that plague typically adopted aspatial models (LeSage & Pace, 2009; Autant-Benard & LeSage, 2011). It is not surprising these general specifications are also most suited to the empirical scope of regional based

[16] The SEMs are an analytical extension of the 'Moran's I' procedure that is motivated through the regression's residual component.

studies, where spatial units share or belong to one form of common international political administration or the other. Intriguingly, the SAR model specification remains most appropriate to neoclassical economic growth specifications because of the underlying assumptions of the model (Fingleton & López-Bazo, 2006).

Despite the remarkable insight from contributions made in the spatial growth literature, there are still areas with considerable scope for further contributions. When it comes to determining and hypothesising realistic spatial networks for cross-sectional interconnections, a lot more can still be done to provide analytical and discursive clarity. Although Lesage & Pace (2010) suggest that it is a 'myth' to suspect there is some bias culminating from model sensitivity to the specification of spatial weight matrices, it is foolhardy to think proper thought should not be given to how realistic these spatial weights are and how they affect result interpretations. The truth is, specifying the nature, type and patterns of spatial connectivity (contained in the spatial weight matrices) are as relevant as estimating, testing and interpreting such model outcomes. As pointed out by Elhorst (2003): that it is important to distinguish between spatial dependence within and between groups. He opines that within group spatial dependence seek to capture spatial associations based on spatial units as geographic neighbours, while between groups spatial dependence may be most suited to capturing spatial interconnections based on comparable characteristics.

As already highlighted above, spatial weight specification and construction remains an avenue for making useful contributions to the literature (Abreu et al, 2005). This is even more so in a global economic landscape, where the relevance of geographical distance between countries is continually shrinking in significance. Apart from the conventional geography based (contiguity and distance) spatial weighting schemes there are fewer applications of practical alternative spatial network specifications in the spatial growth literature. And in terms of application, spatially adjusted models remain mostly either residual-based technology models focusing on knowledge spillover applications (such as, Lopez-Bazo et al, 2004); or they are international trade transmission models in which distance decaying (gravity equations) applications matter (such as Moreno & Trehan, 1997; Fujita et al, 1999). The uses of alternative definitions of proximity such as those with cultural dimensions are often relegated to studies at the micro-type individual and firm level network models. Even as Haynes & Fotheringham (1984:12) point out that the effect of distance may be rightly minimized by the

[unifying] presence of a shared culture, language or information; little has been done to include such significant dimensions to cross-country growth studies. A number of works have identified with this lack such as cited in Moreno & Trehan (1997: 401), where Goodfriend & McDermott (1994) consider technological spillover effects in their model that compositely considers geographical proximity with proximities based on commonalities in language, culture and the presence of commercial relations between country pairs. There are also a lot more applications of this kind in the political economy discourse that not only attempt to use more complex spatial weighting schemes to capture multidimensional interdependencies (such as Beck et al, 2006 following Lacombe, 2004), but also consider non-geographic cultural spatial weight designs to accommodate for alternative forms of interconnections (see Starr, 1991[17]; Simmons & Elkins, 2004).

Section 1.3 – Methodology

1.3.1 – Model Specification

Starting from a standard 3 inputs Cobb-Douglas income production function with the assumptions of constant returns to scale and human capital augmented labour[18] - a vectorized form can be specified as:

$$\boldsymbol{y_{it}} = \boldsymbol{X_{it}\beta} + \in_{\boldsymbol{it}} \quad \text{… eqn.(1)}$$

where,

$y_{it} = the\ dependent\ variable$ usually $measuring\ national\ income$

$X_{it} = a\ vector\ holding\ the\ constant\ and\ the\ physical\ and\ human\ capital\ variables$

$\beta = a\ vector\ holding\ the\ coefficients\ to\ be\ estimated$

$\in_{it} = the\ error\ of\ the\ same\ form\ and\ transformation\ as\ above$

[17] It is worth noting that although Starr (1991) controls for spatial interdependence(s), he does not particularly use a spatial econometric model.

[18] The assumption of constant returns to scale is adopted for its algebraic convenience and interpretational advantage – providing an algebraically feasible model with variables in per-labour terms. Such a model also accords the advantage of yielding fewer estimates to be computed especially relevant as we project this unto a space-time dimension. Although, this is a recognized weakness and has been criticized in the literature (see Klenow & Rodriguez-Clare (1997) and McQuinn & Whelan (2007)), relaxing this assumption does not necessarily map unto significant analytical improvements. Constant returns to scale is also most feasible in a closed economy archetype compared to an open economy where increasing returns are more likely as a result of the integrative mechanism consistent with positive productive network effects (see Fingleton & López-Bazo, 2006).

The model is augmented for spatial heterogeneity and the parameters estimated are in terms of per-labour units. The conventional approach (see Mankiw, Romer & Weil (MRW), 1992 and Islam, 1995 for instance) to solve for the reduced form steady-state parameters of the Investment-to-income ratio via the capital accumulation equation is avoided in this paper. One advantage of modelling the capital variables as stocks is that it obviates the need to solve the model for the steady state levels with respect to their corresponding investment ratios in order to estimate their marginal products. However, the practice of including a convergence coefficient consistent with the neoclassical growth literature is maintained. This is conventionally obtained by a Taylor series approximation around a hypothesized steady state income level consistent with the equilibrium specification (see MRW, 1992: 422-3).

Incorporating this convergence approximation allows one to accommodate the assumption that current observed income is a continuous function of a 'steady state' income level: equation (1) can hence be written as:

$$\boldsymbol{y_{i,t}} = \left(\mathbf{1} - \boldsymbol{e^{-\lambda\tau}}\right)(\boldsymbol{X_{i,t}\beta} + \in_{\boldsymbol{i,t}}) + \boldsymbol{e^{-\lambda\tau}y_{i,t-\tau}} \qquad \text{... eqn.(2)}$$

where,

$\lambda = the\ speed\ of\ convergence\ and\ can\ be\ obtained\ via\ the\ expression\colon -\left(\frac{\boldsymbol{ln(1+\beta)}}{\boldsymbol{\iota\tau}}\right);$

$\tau = 3\colon the\ length\ of\ the\ time\ interval\ used\ in\ the\ temporal\ averaging$

$\iota = 6\colon the\ total\ number\ of\ temporal\ intervals$

Since income (the dependent variable) is usually modelled in per-labour growth units, equation (1) can be transformed into a growth expression by algebraically subtracting the level of Income-per-Labour, lagged one period, from both sides of equation (1).

$$\boldsymbol{y_{i,t} - y_{i,t-\tau} = gy_{i,t}} = \left(\mathbf{1} - \boldsymbol{e^{-\lambda\tau\iota}}\right)\left(\boldsymbol{X_{i,t}\beta} + \in_{i,t}\right) - \left(\mathbf{1} - \boldsymbol{e^{-\lambda\tau\iota}}\right)\boldsymbol{y_{i,t-\tau}} = \boldsymbol{X^{I}{}_{it}\beta^{I}} + \boldsymbol{E_{it}} \qquad \text{... eqn.(3)}$$

Here, $\boldsymbol{X^{I}_{i,t}}$, represents a vector containing the independent variables in $\boldsymbol{X_{i,t}}$ and the temporal lag of the dependent variable, $\boldsymbol{y_{i,t-\tau}}$; and $\boldsymbol{\beta^{I}}$ is simply the matrix, $\left(\mathbf{1} - \boldsymbol{e^{-\lambda\tau\iota}}\right)\boldsymbol{\beta}$. All the variables remain in their log-linear forms; so any first-difference corresponds to the variables in their growth rates.

Following concerns from the literature (Phillips and Moon, 2000) that time series datasets (especially the Penn World Table data) are likely to be non-stationary due to seasonality a 3-year averaging procedure is adopted. This allows for a better exploration of both time and cross-sectional dimensions in

the data. The model is then projected unto a spatial econometric framework by interacting the variable (income growth) and the residual with relevantly specified spatial weight matrices - specifying the nature of the connections between the countries in our sample. Specifically, the residual is repeatedly (for each estimated model) spatially adjusted adopting an inverse-distance matrix design. This allows our framework to consistently explore the spatial patterns (mostly cultural) in the global economic development process that are robust to any arbitrary spatial patterns in the model's residuals.

Given that our spatial assumption is that inter-country associations directly affect income growth, which then translates to systemic decaying influences on the way in which neighbouring productive inputs then affect growth, a spatially augmented model using a spatial autoregressive (SAR) specification is hence most appropriate. However, since we want our estimates to be robust to any arbitrary spatial patterns in the Solow-residual, we adopt a spatial autocorrelation (SAC) model instead. The spatially augmented version[19] of the model in equation (3) can be estimated using the reduced form transformation as shown in equation (5) below:

$$\boldsymbol{gy_{i,t}} = \boldsymbol{\rho W gy_{i,t}} + \boldsymbol{X^{I}_{i,t}\beta^{I}} + \ \ni_{i,t} \qquad \text{… eqn.(4)}$$

The spatially adjusted residual can be seen in equation (4b) immediately below. $\boldsymbol{\varrho}$, is the estimated spatial coefficient using the inverse squared distance specification as contained in the spatial weight matrix, $\boldsymbol{W^{inv}}$.

$$\ni_{i,t} = \ \boldsymbol{\varrho W^{inv}\xi_{i,t}} + \boldsymbol{\mathcal{E}_{i,t}} \qquad \text{… eqn.(4b)}$$

$$\boldsymbol{g_{y_t}} = (\boldsymbol{I_N} - \boldsymbol{\rho W})^{-1}(\boldsymbol{X^{I}_{it}\beta^{I}} + \ni_{it}) \qquad \text{… eqn.(5)}$$

The assumptions here are that the capital factors (physical and human) of production are accumulated internally but are however affected by a global interactive component, which fades away (decays) according to the extent of a country's contact, distance, or connectivity with the other countries in its shared productive network. Our specification allows us to capture this much better as it provides the component, $(\boldsymbol{I_N} - \boldsymbol{\rho W})^{-1}$, which serves as the trickle (fading or ripple) impact mechanism through which growth of interconnected

[19] It is possible to structurally derive and incorporate a spatial dimension into the Solow model directly. However, this paper chooses to adopt a narrative based justification for why cross-sectional dependence is likely present in the empirical distribution of income globally. For examples of papers that provide formal spatial augmentations to the textbook Solow model please refer to Ramirez & Loboguerrero (2002), Fingleton & López-Bazo (2006) and Huggins & Thompson (2014).

countries affect any particular reference economy in the framework. The SAC model, very similar to its SAR counterpart, also provides a more suitable spatially augmented framework for handling an equilibrium income design (Solow-MRW specification) for which the data is measured in units at the country-level (see Elhorst 2014: 7). The model in equation (5) attempts to explain average economic performance (growth in income per capita) based on covariates directly observed in each country and indirectly from closely connected countries accruing from a collective productive process. If country units were under a common geopolitical administration as often characterizing regional (either provincial or continental) spatial growth studies, then an SDM model (which additionally includes the direct impacts of the covariates from neighbouring countries) would have been a more applicable functional form.

Equation (5), above, represents the reduced form version of equation (4). The spatial 'multiplier' or inverse transformation matrix is given as:
$(\boldsymbol{I_N} - \boldsymbol{\rho W})^{-1}$, which helps us identify the total average direct, indirect and feedback effects resulting from the hypothesized spatial process as proposed by this chapter as laid out in the definition of $\boldsymbol{W = spatial\ weight\ matrix}$.

This $\boldsymbol{W}$ matrix is quite informative when viewed in this infinite series expansion expression below:

$$(\boldsymbol{I_N} - \boldsymbol{\rho W})^{-1} = \boldsymbol{I_N} + \boldsymbol{\rho W} + \boldsymbol{\rho^2 W^2} + \boldsymbol{\rho^3 W^3} + \cdots$$

where,

$\boldsymbol{I_N}$: an N x N (number of spatial units) identity matrix that helps to identify from the $\boldsymbol{\beta^I}$ matrix, the main marginal effects accruing to countries from their own productive activities, not induced by the system or network;

$\boldsymbol{\rho W}$: defines the cumulative indirect or spillover effects resulting from first-order spatial neighbours;

$\boldsymbol{\rho^2 W^2} + \boldsymbol{\rho^3 W^3} + \cdots$: defines the total induced (feedback) effects arising from higher-order direct and indirect neighbouring network relationships (effects) for all the respective interacted variables in the reduced form regression in equation (5). The total feedback effect ($\boldsymbol{\rho^2 W^2} + \boldsymbol{\rho^3 W^3} + \cdots$) represents the total impacts on country $\boldsymbol{i}$ resulting from only those effects 'passing through' all neighbouring countries and back to itself, where it originated. It is more like a knock-on or shadow effect incited passively from the more dominant main and indirect effects. This is why the overall direct effect is in fact greater than the actual main effect, $\boldsymbol{\beta^I}$, obtainable from the typical FE-OLS (Fixed-

Effects Ordinary Least Squares) form of the same estimate that normally ignores systemic effects underpinned by processes of spatial integration not accounted for. A summary table of these direct and ancillary effects have been provided in Table A - 1.3.1 below

Effects Type	Description	Equivalent expression (matrix) from Equation (5)
Main Effect	Overall Own-country effects excluding any spatially induced (feedback or battery) effects	$\boldsymbol{I_N} * \boldsymbol{\beta^I} = \boldsymbol{\beta^I}$
Direct Effect	Main Effect + Overall Higher-order direct and indirect (feedback or battery) Effects	$(\boldsymbol{I_N} + \boldsymbol{\rho^2 W^2} + \boldsymbol{\rho^3 W^3} + \cdots) * \boldsymbol{\beta^I}$
Indirect Effect	Overall First-order indirect effects	$\boldsymbol{\rho W} * \boldsymbol{\beta^I}$
Total Effect	Direct effect + Indirect Effect	$(\boldsymbol{I_N} + \boldsymbol{\rho W} + \boldsymbol{\rho^2 W^2} + \boldsymbol{\rho^3 W^3} + \cdots) * \boldsymbol{\beta^I}$

Table A – 1.3.1: A tabular summary of the spatial and non-spatial coefficients of equation (5) above

1.3.2 – Spatial Weight Matrix Specifications

The core hypothesis in the methodology of this thesis chapter is that the underlying spatial process of the world's economic performance is multidimensional. This process is not limited to the traditional approaches of geographical definitions. The world economies have evolved dramatically affecting productive integration in patterns equally consistent with alternative channels of interconnections. These patterns are more reflective of similarities in fundamentals such as integration on the basis of global cultures, political outlook, their origins, and the likes. While geographic space would have mattered in the past, it is not as relevant in present times, especially not when trying to understand the patterns of integrative productive capacities at the macro-level.

Two broad dimensions (angles) of spatial weight specifications are used in this research exercise. The first, for comparative purposes, is based on the classical geographic approach while the other is based on cultural similarities – those particularly focused on the global spatial distribution of ethnic and

languages groups. This provides an avenue for comparing the integrative components of economic activity at the core of the neoclassical framework when exposed to the mechanism of network interactions at an inter-country level. This comparison is made by looking at economic growth in a world in the presence of geographical integration with one in the presence of cultural integration. It is a comparative exercise illustrating the importance of a potentially ignored angle of country-level interconnections and the resulting implications for developed and developing countries, alike.

1.3.2.1 – Specifications for the Geographically Defined Spatial Weights

In general, our $\boldsymbol{W-matrix}$ is an n x n symmetric matrix in the following form:

$$\boldsymbol{W_i} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix} \quad ;$$

where, $\boldsymbol{W}$ is specified as a global array of countries represented both contiguously and in a squared inverse-distance form. For the actual model estimation, the matrices are further spectral-normalized, such that each element in $\boldsymbol{W}$ are divided by a common uniquely acquired scalar (acquired as the largest eigenvalue of the corresponding matrix). This normalization procedure is useful because it allows the $(\boldsymbol{I-\rho W})$ to be nonsingular and invertible for all $\boldsymbol{\rho}$ within its parameter space (see Kelejian & Prucha, 2010 for more technical discussions on the subject).

1.3.2.1a – The Contiguous Weight Matrix Formulation:

$$\boldsymbol{W_1} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where,

$$\boldsymbol{w_{ij}} = \begin{cases} \mathbf{1}, & \textit{if a pair of countries are geographically contiguous neighbours} \\ \mathbf{0}, & \textit{if a pair of countries are not geographically contiguous neighbours} \end{cases}$$

1.3.2.1b – The squared Inverse-distance Weight Matrix formulation:

We adopt an unbounded inverse-distance specification with the attempt to capture any arbitrary spatial effect that could inadvertently still be present in the models' residuals, where a far-reaching global horizon in the error-term is adopted. This is especially useful as it ensures our estimated spatial

autoregressive process in the dependent variable is robust to the presence of any potential unaccounted spatial process in our residuals[20]. The approach would have been distinctly different if the intended study was more localized to a region of countries. In such an instance, using a cut-off radius or adopting a stricter distance regime would have been more ideal. Our specification is relatively stricter than the average inverse-distance specification as our friction parameter is set to 2 (or simply squaring the inverse-distance matrix, element by element). Our approach to adopt an unbounded spatial weight regime here is consistent with Cliff & Ord (1973, 1981).

$$\boldsymbol{W_2 = W^{inv}} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where,

$\boldsymbol{w_{ij} = d_{ij}^{-b}}$;

$\boldsymbol{d_{ij}} = the\ distance\ between\ country,'i',and\ country,'j'$. It is calculated using the haversine distance measure which takes into consideration the spherical nature of the earth's surface and as a result incorporates this 'great-circle' property into the distance measure, such that:

$\boldsymbol{d = r(2\ arcsin(min(1, \sqrt{a}))}$

$\boldsymbol{r = 6,371.009km\ or\ 3,958.761miles}$: This is the mean radius of the earth used in the measure

$\boldsymbol{a = sin^2\phi + cos(\phi_1)\ cos(\phi_2)\ sin^2 v}$

where,

$\phi = \frac{1}{2}(\phi_2 - \phi_1) = \frac{1}{2}(x_2[j] - x_2[i])$

$v = \frac{1}{2}(v_2 - v_1) = \frac{1}{2}(x_1[j] - x_1[i])$

$x_1[j]\ and\ x_1[i] = longitudes\ of\ point\ i\ and\ j, respectively$

$x_2[j]\ and\ x_2[i] = latitudes\ of\ point\ i\ and\ j, respectively$

$b = the\ friction\ parameter$: $which\ takes\ the\ values\ of\ 2,$ in our specification in W_2

[20] See Chen, 2012, for a number of widely used spatial weights specifications based on distance. These weighting types, for instance, show how varying the power of the distance measure (the friction threshold) affects the nature of interactions between the intended spatial units in a model.

1.3.2.2 – Specifications for the Culturally Defined Spatial Weights

Here, our objective is to attempt to capture spatial interaction using spatial weight matrices specified to link countries through global observed patterns of shared ethnic or language groups. We achieve this by pairing countries based on whether there are shared ethnic groups (ethnic neighbours) between them, or whether they have common spoken languages (language neighbours) between them. While the relative population concentration of these cultural elements (languages and ethnic groups) may be relevant in studies interested in distinguishing between varying degrees of global cultural penetration within local populations, our focus herein is on the global distribution of these cultural elements. As a result, we focus on the worldwide spatial dispersion of world languages and ethnic groups rather than the individual proportion of specific ethnic identities or languages within each comprising population of each country. This focus is relevant in two ways: firstly for conducting a research relevant to the empirical reality of increasing (widening) global links between countries (as with globalization, for instance); and secondly, it is technically relevant as our spatial weights are designed to be consistent with the assumption (or condition) that attributes used for specifying them ought to be relatively time invariant such as geographical attributes often are. It is easy to see that the percentage of a country's population speaking a particular language or belonging to a specific ethnic group will more likely change much faster over time compared to its mere occurrence or presence in the said country. Hence, the global spread or dispersion of these cultural elements and not their locational concentration within each country is the dimension from which the cultural spatial weights (links) herein are designed.

1.3.2.2a – Ethnic Neighbours:

$$\boldsymbol{W_{E1}} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where in this specific case,

$$\boldsymbol{w_{ij}} = \begin{cases} \mathbf{1}, & if\ there\ is\ at\ least\ one\ ethnic\ group\ common\ to\ corresponding\ country\ pairs \\ \mathbf{0}, & if\ there\ are\ no\ ethinic\ groups\ common\ to\ corresponding\ country\ pairs \end{cases}$$

We also provide a composite weight specification involving an element-wise multiplication of this weight matrix with the contiguous spatial weight, $\boldsymbol{W_1}$ (specified above), and obtain: $\boldsymbol{W_{E1w1}} = \boldsymbol{W_{E1}} \boxed{.*} \boldsymbol{W_1}$. This latter composite

variant-specification indicates that country pairs are spatial neighbours when they not only have at least one ethnic group in common but also are contiguous neighbours. This allows us to test interaction effects in our spatial weight models in order to confirm whether both patterns are mutually exclusive or if these cultural spatial effects are reinforced by geographic proximity.

$$\boldsymbol{W_{E2}} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where in this specific case,

$$\boldsymbol{w_{ij}} = \begin{cases} \mathbf{1}, & if\ there\ are\ at\ least\ 3\ ethnic\ group\ common\ to\ corresponding\ country\ pairs \\ \mathbf{0}, & if\ there\ are\ less\ than\ 3\ ethnic\ group\ common\ to\ corresponding\ country\ pairs \end{cases}$$

This is a relatively stricter specification than $\boldsymbol{W_{E1}}$. It allows us allay doubts about the generosity and feebleness of the association effects permissible with the $\boldsymbol{W_{E1}}$ specification. The relevant composite weight specification-variant is an element-wise multiplication of this weight matrix with the contiguous spatial weight, $\boldsymbol{W_1}$, obtaining: $\boldsymbol{W_{E2w1}} = \boldsymbol{W_{E2}} \boxed{.*} \boldsymbol{W_1}$.

$$\boldsymbol{W_{E3}} = \begin{bmatrix} \boldsymbol{w_1} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where in this specific case,

$\boldsymbol{w_{ij}} = \sum \boldsymbol{e_s}$;

$\sum \boldsymbol{e_s}$ is the sum of all shared ethnic groups between spatial pairs. This matrix is a more generic one but remains relatively quite as strict a specification as the preceding specification, $\boldsymbol{W_{E2}}$ above. This is because it provides a mechanism for measuring varying relative intensities of associations between country pairs while still maintaining the symmetric quality similar to the preceding specifications. The composite variant specification of this weight matrix is the element-wise multiplication of this weight matrix with the contiguous spatial weight, $\boldsymbol{W_1}$, obtaining: $\boldsymbol{W_{E3w1}} = \boldsymbol{W_{E3}} \boxed{.*} \boldsymbol{W_1}$.

1.3.2.2b – Language Neighbours:

$$\boldsymbol{W_L} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where in this specific case,

$$\boldsymbol{w_{ij}} = \begin{cases} \mathbf{1}, & if\ there\ is\ at\ least\ one\ spoken\ language\ common\ to\ corresponding\ country\ pairs \\ \mathbf{0}, & if\ there\ are\ no\ spoken\ languages\ common\ to\ corresponding\ country\ pairs \end{cases}$$

We equally provide a composite weight specification involving an element-wise multiplication of this weight matrix with the contiguous spatial weight,

$\boldsymbol{W_1}$, as done with the ethnic groups metric variants above obtaining: $\boldsymbol{W_{L1w1} = W_{L1} \boxed{.*} W_1}$.

$$\boldsymbol{W_{L2}} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where in this specific case,

$$\boldsymbol{w_{ij}} = \begin{cases} \mathbf{1}, & \textit{if there are at least 3 spoken languages common to corresponding country pairs} \\ \mathbf{0}, & \textit{if there are less than 3 spoken languages common to corresponding country pairs} \end{cases}$$

The composite weight variant of this specification involving the element-wise multiplication of this weight matrix with the contiguous spatial weight, $\boldsymbol{W_1}$, is: $\boldsymbol{W_{L2w1} = W_{L2} \boxed{.*} W_1}$.

$$\boldsymbol{W_{L3}} = \begin{bmatrix} \boldsymbol{w_{11}} & \cdots & \boldsymbol{w_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{w_{n1}} & \cdots & \boldsymbol{w_{nn}} \end{bmatrix};$$

where in this specific case,

$\boldsymbol{w_{ij} = \sum l_s;}$

$\boldsymbol{\sum l_s}$ is the sum of all spoken languages common to corresponding spatial pairs and defined in the same fashion as $\boldsymbol{W_{E3}}$, above. The concomitant composite variant-specification is the element-wise multiplication of this weight matrix with the contiguous spatial weight, $\boldsymbol{W_1}$, obtaining: $\boldsymbol{W_{L3w1} = W_{L3} \boxed{.*} W_1}$.

It is important to note that the decision to use either 1 or 3 ethnic groups or languages as a criterion for designing the spatial networks herein is very similar in nature to the decision to work with a distance friction parameter of 2 as in our inverse distance specification above. The 1 or 3 criteria acts as a kind of friction parameter as well. Where the higher the benchmark, the tighter the network or the fewer the avenues for spatial interconnections. Hence, a choice of at least 3 languages or ethnic groups between country-pairs would result in a sparser network design compared to that inherent in the choice of at least 1 cultural (ethnic or language) group in common. We choose to start from the slackest option and proceeding to one that is marginally stricter. The choice of 3 is arbitrary and any other friction criteria could have also been likely adopted. The only additional condition is that we allow for a global far-reaching network design; where the use of alternative specifications are for comparison or robustness purposes rather than for the purpose of model selection. It is primarily for this reason, that the unrestricted version of these cultural spatial networks as seen in the $\boldsymbol{W_{L3}}$ and $\boldsymbol{W_{E3}}$ specifications are additionally adopted. These $\boldsymbol{W_{L3}}$ and $\boldsymbol{W_{E3}}$ specifications provide a framework

that harness the full range of shared ethnic and language groups between each country pair using it as a grid intensity index, rather than as an exclusion (inclusion) mechanism. Adopting a good range of choices mitigates concerns about weight selection biases.

Additionally, the spatial weights adopted herein are all spectral normalized representations of the queen contiguity[21], squared inverse-distance (measured in miles), cultural matrices, and their corresponding interacted (or composite) variants, respectively. The distance matrix, as earlier highlighted, is adopted with no particular distance limit or bound. This is so, given that our models are intended to spatially represent integration at a fully extended inter-country level. The empirical relevance of this preference is discussed further in the results and model interpretations section - in section 4 below.

1.3.2.3 – Spectral- normalization: Specification and Justification

Normalizing spatial weight matrices is common practice in the literature particularly because it helps remove any measure-unit effect that may result from the computation of the spatial coefficients. Most papers in the literature adopt row-normalized matrices probably because estimations from such models allow for an algebraic representation of a weighted average of the collective spatial connections thus enabling convenient model prediction and interpretation (Lesage & Pace, 2009). Kelejian & Prucha (2010: 56) critique this position by suggesting that the adoption of row-stochastic spatial weight matrices should be guided by well-considered theory. This is because models based on row-normalized spatial weights may most often result in model misspecification. Their reason being that the normalization technique involves

[21] The queen contiguity matrix specification is less strict than the rook contiguity as it treats spatial units as neighbours either whether they share a common border or a single point rather than the latter, which requires that both conditions be satisfied jointly. A less strict weight definition would naturally provide a better approximation of global inter-country network-growth compared to a stricter one, which would be more appropriate for regional based or intra-country (within country) studies. These studies are often characterized by denser network interactions between the defined spatial units because there are typically more avenues for spatial interactions at lower administrative levels compared to a more global inter-country level focus. This is also the reason why adopting unbounded inverse distance weights do not necessarily pose an empirical challenge as the number of spatial units at the inter-country level is less densely connected compared to sub-administrative level studies at regional intra-country (such as studies for the United States) or intra-regional studies (such as among EU countries). See Chen, 2012, for more discussions on varying the specification of spatial weight and their respective theoretical and empirical implications.

the use of a vector of row-sums based on the number of connections per network rather one based on a scalar which is more suitable for a model with an autoregressive parameter dependent on the total number of cross-sectional units (such as with the spectral radius and min-max normalizations). The difference being the former (row-normalization) often produces an asymmetric matrix resulting in a non-scalar re-scale factor of the auto-regressive parameter, as the normalized matrix is not structurally equivalent to the originally defined weight matrix. The latter normalizations (spectral radius and min-max) however, adopt a uniform normalizing factor (a scalar). This maintains the initial spatial weight's underlying structure; hence is very useful for modelling equations specified as equilibrium relationships or when estimating models based on the classic decaying weight distance hypothesis. In short, row normalizing may be convenient and popular but the asymmetry it produces in the weighting structure implies that the spatial impact of unit $\boldsymbol{i}$ on unit $\boldsymbol{j}$ may not be equivalent to that of unit $\boldsymbol{j}$ on unit $\boldsymbol{i}$ thus obscuring interpretational relevancies tied to specific theories (see Anselin, 1988: 23-24; Elhorst, 2014: 13-14). As a consequence, this paper adopts a spectral normalization weight in its approach as this is a more general and structurally consistent functional form compared to the row-normalization representation.

In the specification of spectral normalized spatial weights, each element of the relevant matrix is normalized using a scalar:

$\boldsymbol{v}^* =$ is usually the largest moduli from the eigenvalues of the correspondingly defined spatial weight matrix

So that the consequential operational matrix, $\boldsymbol{W}$, is obtained by:

$$\boldsymbol{W}_*^{\,v^*} = \frac{w_{ij}}{v^*}$$

Given that the spatially interacted right-hand side variables are expected to be substantially correlated with their non-spatially transformed versions, the paper adopts a Spatial Autocorrelation (SAC) specification: where both the dependent variable (income per capita growth) and the models' residuals are spatially lagged. The rationale for this can be motivated from a number of positions. Most apparent is the argument that increasing economic activity from income growth may readily spillover to neighbouring countries generating convergences via channels such as transfers in economic productive capacity

and political systems or policy choices (see[22] Starr, 1991; Simmons & Elkins, 2004). Also more economically, trade, technological (and skills) transfer, and through increased factor mobility made possible by the integration of countries within regional and global economic and geopolitical organizations, could help us rationalize just why countries may be spatially connected (also see Moreno & Trehan, 1997 for some discussions on similar matters). It is equally not farfetched, for instance, to conceive the rationale that richer countries with relatively more vibrant economies are generally more likely to attract better (in quantity and quality) capital (human and physical). This could be achieved either by actively exporting educational services or by enabling favourable immigration policy choices targeted at accruing and retaining high quality human capital, for instance. Factor mobility, especially as embedded in migratory patterns, provide some observable evidence of how labour and their accompanying human capital have been found to flank to more developed countries from relatively lesser developed ones in search of better socio-economic opportunities (such as higher returns to their labour services). This could certainly generate spatially persistent global growth patterns where relatively more developed economies, possessing higher productive capacities, are able to attract and retain more capital resources at any one time, compared to their less vibrant developing country counterparts. As a result, space – both geographical and otherwise defined could provide a useful mechanism for feasibly exploring the network distribution of such productive capacities. Ideally geographically closer countries would be the initial beneficiaries of any available productive capacity and input spillovers compared to more distant countries. Also, countries that share similar or common historical experiences and heritages, cultural identities or characteristics, and institutional compositions would equally fit this sort of spatial dynamic. A number of papers have attempted to explicitly model capital accumulation as a spatial process within an endogenous growth framework. Huggins & Thompson (2014), for instance, refer to their conceptualization of this as a composite form of capital. This includes a

[22] See these papers for an example of some of these channels discussions that link the political economy literature with the spatial econometrics literature.

constituent network capital, which is quite consistent with our narrated view of paired-sourcing from a global capital pool[23], as highlighted above.

The models herein are estimated using a maximum likelihood estimator. A user-written command, 'xsmle' available via STATA[24], is used to estimate the models with a fixed-effects (FE) specification. The constant term is restricted to capture just individual atemporal effects, which is particularly justifiable as our chosen temporal frequency is at the 3-year averaged dimension. Some of the justifications for choosing spatially augmented models in order to reinvestigate the dynamics of growth divergence across the world have already been discussed in preceding sections. For a methodological justification however, the literature on panel studies identifies overwhelmingly the crucial implications of not accommodating for cross-sectional interdependence across countries in regional (either provincially or continentally) and global empirical estimations (Breitung & Pesaran, 2008). Also there is an additional advantage that spatially augmented models are more robust to omitted variable biases and model misspecifications compared to their aspatial counterparts (Franzese & Hays, 2007; LeSage & Pace, 2009).

1.3.3 – Data and Sample

The conventional economic growth model variables (GDP, labour, physical capital stock and human capital index) have all been sourced from the latest range of the Penn World Table datasets (version 8.0 - Feenstra, Inklaar & Timmer, 2013). The physical capital stock and income variables are measured in real terms, constant at 2005 national prices and in 2005 US dollars. The labour variable is measured as the 'number of engaged persons'. While the human capital variable has been outputted as an index in this dataset series where the average years of secondary schooling attained has been adjusted for returns to education following Psacharopoulos (1994).

The required cross-sectional data for locating the countries (the longitudes and latitudes) is obtained from the Diva-GIS website[25]. The data resource is

[23] This should be finite for any given time and global technological level.

[24] See Elhorst, 2010, for some technical discussions on specifications and alternative estimations of panel models following spatial augmentation.

[25] http://www.diva-gis.org/Data

developed and maintained by a number of geospatial researchers - Robert Hijmans, Edwin Rojas, Mariana Cruz, Rachel O'Brien, and Israel Barrantes, whose objective is to provide free and accessible geospatial data.

Also at the cross-sectional dimension, there are 2 main culturally defined weight matrices. The first is based on the spatial distribution and presence of ethnic groups between and across countries while the second is very similar but uses the presence of language within countries as its reference for designing the spatial network matrices. The variable used for designing the first class of culturally defined spatial weight matrix is obtained from the GREG project (Weidmann, Ketil & Cederman, 2010). It is a dataset that geo-references the ethnic groups drawn from the classical Soviet Atlas Narodov Mira, by employing geographic information systems for each of the ethnic group territories using geographical spatial polygons. From this dataset, we are able to identify major ethnic groups across the world and cluster countries into neighbourhood networks based on the spatial distribution pattern of these ethnic groups. If the countries have a specified number of ethnic group(s) in common, then they are considered ethnic neighbours and hence connected to the same cultural network, if not they are tagged as non-neighbours and unconnected culturally. There are thirteen countries[26] not represented by the GREG dataset but which are presented in the 122-country sample used in this paper. In this case, the GREG ethnic groups' database is cross-referenced with the databases on the websites of the Encyclopædia Britannica, Inc. (2014) and CIA's World Factbook (2014). It is updated where the dataset does not include the countries used in our sample. A similar approach is alternatively adopted for exploring a second dimension of cultural networks based on the spatial distribution pattern of globally spoken languages. From this dimension, this paper relies on the Ethnologue dataset (Ethnologue 17th edition, Lewis et al, 2014). The Ethnologue dataset uniquely identifies major languages spoken in countries across the world. The study herein clusters countries together, if they share at least a specified number of language(s) in common.

[26] These countries are: Armenia, Estonia, Czech Republic, Croatia, Kyrgyzstan, Kazakhstan, Latvia, Lithuania, Slovak Republic, Moldova, Slovenia, Tajikistan and Ukraine.

1.3.3.1 – Time Period

There are a total of 122 countries in sample spanning from 1990 to 2010 (21 years). The time depth has been selected so as to sufficiently maximize the cross-sectional dimension (the number of countries) as much as possible. This is especially useful because the model is focused on exploring the global cross-sectional interconnections between countries and a substantial number of countries in the former Soviet bloc became independent in the early 1990s.

Our three-year averaged sample spans the intervals: 1990-1992, 1993-1995, 1996-1998, 1999-2001, 2002-2004, 2005-2007 and 2008-2010. After losing the first period interval due to lagging the level of income per labour, the total number of observations is 732 (122x6). We do not use an annual sample so as to mitigate the econometric issues surrounding business cycle effects (such as yearly serial autocorrelation) and non-stationarity associated with persistent macro-variables such as income, capital and labour (for instance see, Temple, 1999:131-132). Although the five-year non-overlapping alternative is pervasive in the literature (for instance, Islam, 1995; Caselli et al.,1996; Bond et al., 2001; Ding and Knight, 2008), we opt for the three-year option. This choice is midway between the annual representation and the widely favoured five-year option. Our choice gives us more observations and variations to work with while avoiding the problems associated with using an annual based sample[27].

The regional composition of the dataset can be viewed from the table (Table B - 1.3.2) below and it does show a good global representation of countries across the globe.

Regions	Frequency	Percentage	Cum.
Africa	35	28.69	28.69
Americas	21	17.21	45.9
Asia	32	26.23	72.13
Europe	31	25.41	97.54
Oceania	3	2.46	100
	Total: 122		

Table B – 1.3.2: Tabular Summary of the Countries in our Sample

[27] Also our results are very similar to those obtained using a 5-year average sample as well; even though these have not been reported herein.

Section 1.4 – Data Summary, Results and Interpretation

1.4.1 – Data Summary

Table C - 1.4.1 below provides a simple summary description of the variables used in the estimations of our model over the entire period considered. Two additional indices (the language penetration index and the ethnic penetrations index) have been introduced to allow a quick visual inspection of their more detailed spatial weight matrix specifications. These indices are computed from the cultural weight matrices earlier specified in section 1.3.2.2 above. Both variables range between zero and 1; where zero implies that no local languages or ethnic groups in a country can be found elsewhere in the world, while 1 implies that all the languages or ethnic groups in a reference country can be found in at least one other country in the world. It provides us with a kind of global cultural assimilation index for both languages and ethnic groups; where the most culturally assimilated countries (value close to 1) are the most culturally progressive beyond their respective country boundaries. Giving us with an adequate country-level summary of our respective (languages and ethnic groups) cultural spatial weight matrices which observes the number(s) of shared groups between individual country pairs.

With these indices, we can compare what we observe spatially with other common measures of economic openness such as inward FDI (foreign direct investment) and total trade volumes[28]. We know for instance that measures of economic openness are individually strongly correlated to economic outcomes. As is apparent from Table D - 1.4.2 and Figure A - 1.4.1 below, these cultural indices are quite substantially and consistently positively associated with the FDI-Income ratio and Trade-Income ratio[29]. This is quite interesting as both measures of cultural penetration (assimilation) appear to be consistently positively correlated with both measures of economic openness even though they are not respectively consistently positively correlated with income growth or with our key productive inputs (physical and human capital). This suggests

[28] These two variables are obtained from the World Bank Development Indicators website (version 2015) spanning 1990 to 2010 – same period span as the variables in our sample.

[29] Please refer to the greyed figures in table for specific correlation coefficients, respectively.

that the positive network (spatial) effects, which we later observe from the estimation output below using our cultural spatial weights, might be operating through related mechanisms coherent with economic openness or other similar forms global integration. Also as expected our language penetration index appears to demonstrate twice as much association with the selected measures of economic openness (FDI-Income and Trade-Income ratios) compared to our ethnic penetration index. This could be an indication that language is a more likely mechanism for globally connecting or integrating countries compared to ethnic groups. This is not so surprising, as we would expect following the literature: that ethnic groups are more likely something inherited or attached to their ancestral homelands unlike languages, which could be learned and shared much more easily. This pattern of association with the economic openness measures is also consistent with the 50% (approximately) absolute linear association between both penetration indices.

Variable	Obs	Mean	Std. Dev.	Min	Max
Growth in Income per capita	732	0.0477222	0.1156184	-0.9326577	0.5026374
Lag of logged Income per capita	732	9.442572	1.258493	6.141289	11.65386
Income per capita	732	25326.13	25251.64	505.2304	115040.4
Logged Capital per Labour	732	10.56877	1.355012	7.192822	12.65332
Capital-Labour ratio	732	77721.88	77995.17	1329.851	312800.3
Logged Human Capital Index	732	0.8549338	0.2594019	0.1258426	1.283687
Human Capital index	732	2.426677	0.5780397	1.134104	3.609925
Trade-Income ratio (average from 1990-2010)	121	80.57029	43.04763	21.52435	363.961
FDI-Income ratio (average from 1990-2010)	121	3.536354	3.160522	-0.0218894	19.02137
Language Penetration index	122	0.5381923	0.2521558	0.0102564	1
Ethnic Penetration index	122	0.7435854	0.2925556	0	1

Table C – 1.4.1: Data summary statistics

Where: the income and physical capital variables are real value constant at 2005 prices measured in 2005 USD.

Variables	gypl	laglnypl	lnkpl	lnhc	lang_pentr_index	ethn_pentr_index	fdi_gdp	trade_gdp
Growth in Income per labour	1							
Lagged Income per labour	-0.0641	1						
Capital-Labour ratio	0.1003	0.9325	1					
Human capital index	0.132	0.786	0.7864	1				
Language Penetration index	-0.0131	0.1475	0.1427	0.1696	1			
Ethnic Penetration index	-0.0242	-0.0965	-0.0708	-0.0878	0.5126	1		
FDI to Income ratio	0.0795	0.1322	0.1568	0.227	0.2578	0.1035	1	
Trade to Income ratio	0.016	0.2062	0.2364	0.2608	0.4178	0.2042	0.64	1

Table D – 1.4.2: Correlation matrix of relevant variables

Please note that all other variables are averaged over the period from 1990 to 2010.

From Table C - 1.4.1 above, we also find that although the average global income per labour is between 1990 and 2010 is 24,722.05, the cross-sectional mean-range between the least observation (757.0756) and the maximum observation (102056.5) is very wide. A look at the spatial map of income per labour at selected years (1990 and 2010) and a summary measure of one of our adopted cultural metrics provides an indication of spatial clustering observable through time thus providing some visual case for a spatial panel model. One would expect, following the conditional convergence hypothesis, some cluster-catch-up effect between the poorer and richer regions over time. This should have been even more feasible if the classic pro-globalization narrative of poorer regions catching up to the richer ones were true. If it were true, then the gains accruing to relatively poorer regions from increased global integration, technological accessibility and having knowledge of a sense of what works from natural experiments in the leading economies ought to provide such a catch-up channel. However, this is not the case: poorer regions in the 1990s remain those that are poorest 20 years later in 2010. This pattern happens to be consistent with areas that have lower ratios of global language influence or projection as found in our language penetration index in Figure D - 1.4.4 below. The 'blue' (range) shaded polygons of Figure B - 1.4.2 and Figure C - 1.4.3 correspond to higher income per labour values while those shaded

with the ‘yellow’ to ‘red’ correspond to the lower variable values. The reverse (in terms of colour regimes) is however the case for our language penetration index portrayed in Figure D - 1.4.4 below. The Sub-Saharan African region particularly stands out as a spatio-temporally persistently under-performing region with a relatively lower language penetration index (mostly ‘blues’) and a low-income per labour value region (mostly all ‘yellow’ to ‘red’).

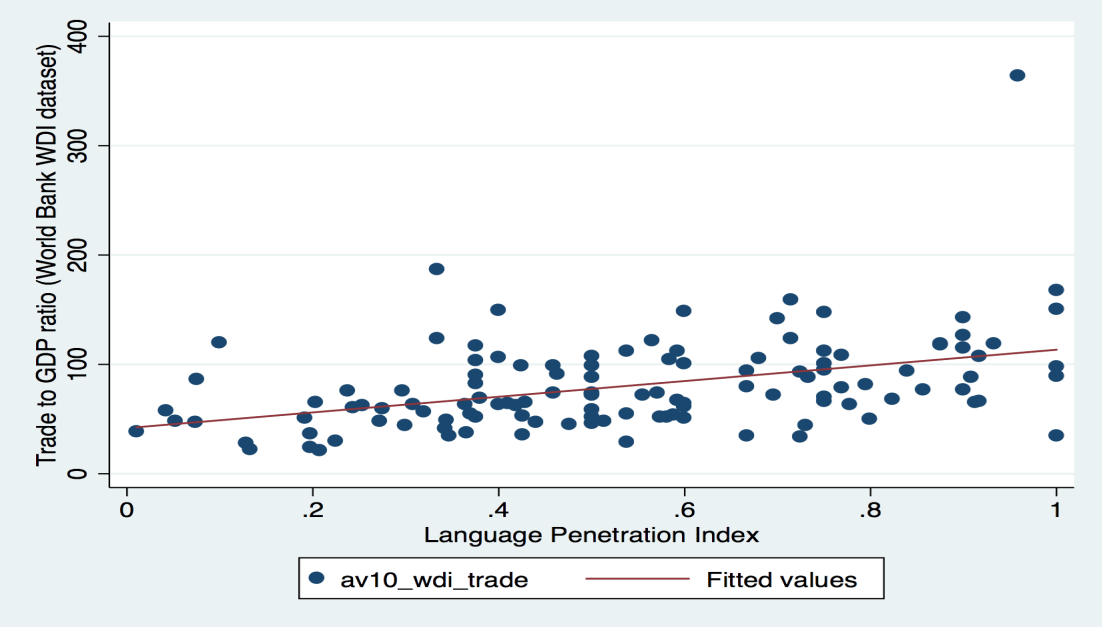

Figure A – 1.4.1: Scatter-plot of our Language penetration index and average Income per labour

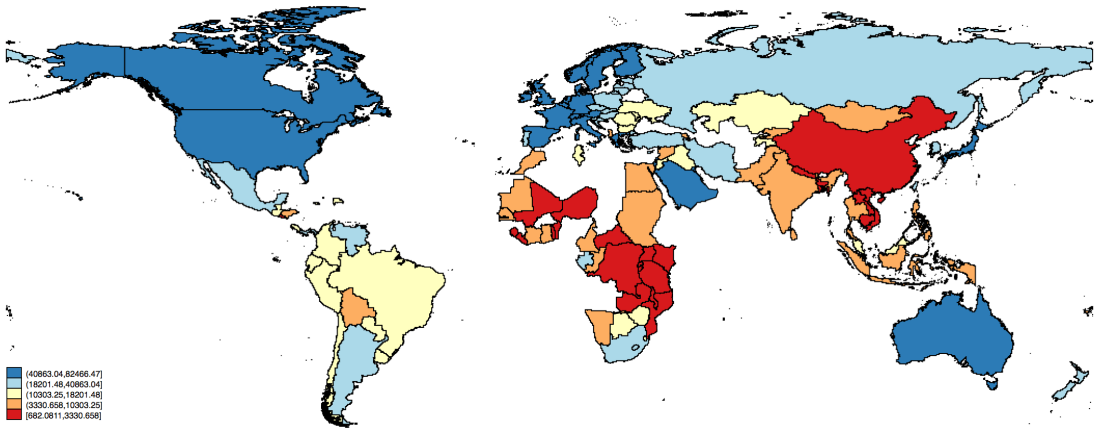

Figure B – 1.4.2: Geographical spatial map of Income-Labour ratio in 1990

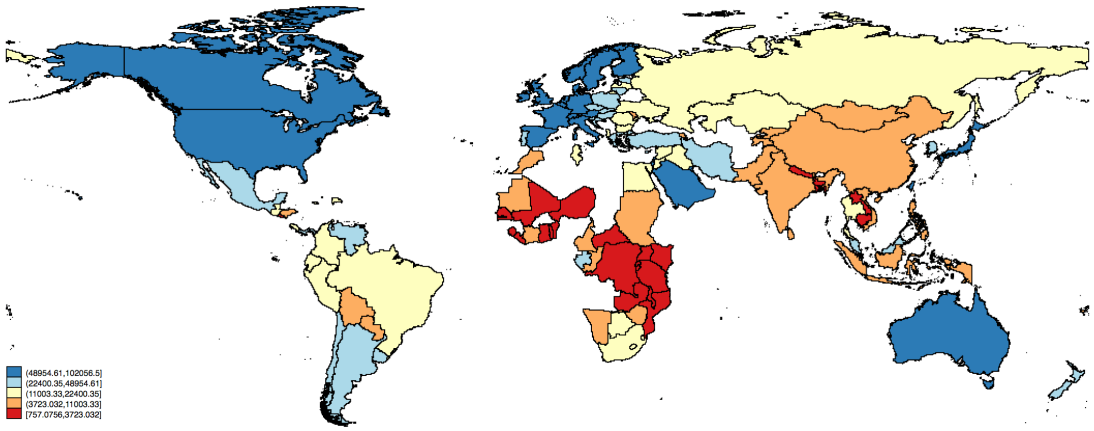

Figure C – 1.4.3: Geographical spatial map of Income-Labour ratio in 2010

Apart from some countries in the Americas regions (South and North America), the spatial distribution of the language penetration index almost certainly approximates the nature of the geographic allocation of income per labour across the world. Although, there is evidence of spatial clustering as portrayed by similar colour-shades found together, this alone does not seem to tell the whole story. As a result, complementing (by adopting composite weight specifications) multiple spatial dimensions of spatial clustering would certainly provide a more informative framework for capturing such multidimensionality that could equally cater for interactions beyond the scope of geographical borders.

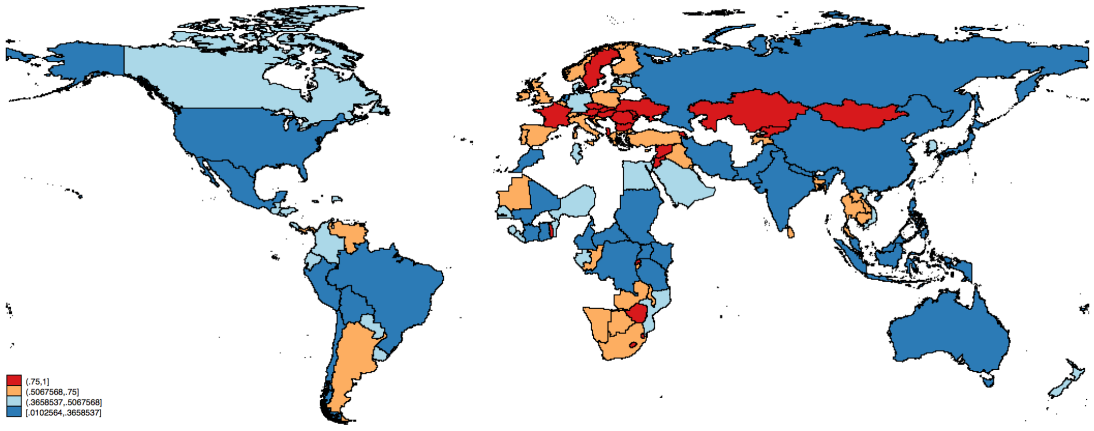

Figure D – 1.4.4: Geographical spatial map of our language penetration index

From Table D - 1.4.2, above, we find that there is in fact a negative relationship between per capita income growth and the lagged level of income per capita, as expected by theory. This suggests some initial evidence for convergence even though this is an absolute correlation coefficient. The size of this coefficient certainly provides some initial factual grounds to suggest

that the possibility of absolute or global convergence may not entirely be as absent as most proponents of the conditional convergence discourse opine. With a meagre absolute correlation coefficient of -0.0641, it indicates an annual average convergence rate of approximately 6.4% over the entire 1990-2010 period. From this Table D - 1.4.2, it can also be seen that all the traditional Solow-variables involved in our model seem to be of the expected signs with regards to their respective correlation coefficients.

1.4.2 – Estimations and their Interpretation

In this section, two classes of cultural spatial weights are adopted to model the spatial interactions (interrelationships) between the countries in our sample. They are compared with 2 base models estimated in A and B in Table E - 1.4.3 below.

1.4.2.1 – Some Discussions on the Base Models

The estimation result in column *A* provides a base estimation of our modelled variables' relationship using a fixed-effects ordinary least squares (FEOLS) specification. While the estimation output in column *B* provides a base spatial autocorrelation specification using the conventional geographic spatial weight matrices[30]. From estimations *A and B* in Table E - 1.4.3 (below) we find that the spatial coefficients are statistically significant for both the spatial-lag of the dependent variable (*rho*) and the error process (*rho_2*). This indicates that not controlling for spatial dependence could most likely result in a misspecification of the actual relationship between the variables. In particular it shows that ignoring an avenue for spatial integration between country units may result in a misspecification bias manifested as an upward bias in the estimated direct (and main) marginal effect coefficients or provide an incomplete story of the global growth process intended for an interconnected socio-economic landscape. It can additionally be observed that the statistical relevance of the spatial-lag coefficients also translates into indirect spatial

[30] In this particular specification, the queen contiguity spatial weight specification is used to spatially augment the dependent variable while the squared inverse-distance spatial weight specification is used to repeatedly spatially augment this model's and the other spatial models' residuals.

marginal effects stemming from the spatial-lag process (refer to Table A - 1.3.1 above for how the process is numerically inferred).

From estimations *A and B* in Table E - 1.4.3, it can be observed that all three proximate growth factors (lagged income level, physical capital and human capital[31]) from the classic MRW (1992) variant of the Solow model are statistically significant and substantially associated with the growth of global income per labour. Quite informatively, we are able to infer the marginal impact of these individual proximate growth factors as shares relative to their own-country and indirect (inter-country) global integration effects. It can be inferred from the total marginal coefficient of the physical capital-labour ratio from estimation *B* that increasing it by two-folds (that is doubling it or a 100% increase) is on the average associated with an approximately 0.33 unit increase in the growth rate of income per labour, when the other covariates are left unchanged. Controlling for the equilibrium adjustment mechanism (or the convergence effect) makes the physical capital per labour coefficient equivalent to the marginal product of capital (MPK) from the base Solow (Cobb-Douglas) production function in equation 1 specified in section 1.3.1 above. The MPK could be obtained by dividing the capital-labour coefficient by the lagged income per labour coefficient (0.33/0.66 = 0.5)[32]. This implied average value of approximately 0.5 for the MPK is robust to the three classes of weight specifications and with respect to both direct and total marginal effects. This value is also consistent with the MPK coefficient implied by the MRW's (1992) base model (0.5). It is nonetheless, slightly bigger than that which they (MRW, 1992) find after controlling for human capital and is much bigger than the empirical prediction (estimate) of the theoretical Solow model of one-third (1/3). By adopting the queen-contiguity spatial weight in equation *B*, we also find that the indirect marginal effect for this capital-labour ratio is only approximately 15%[33] of the total effects, which we have just interpreted above.

[31] All measured in per labour terms.

[32] These have been obtained by rounding-up (in decimal terms) the absolute coefficients of the physical capital and convergence variables in the total effects part of the estimations above. The corresponding formula can be rightly deduced by dividing through the relevant estimated coefficients from the estimating equation with the absolute value of the convergence coefficient: $\frac{\beta\ (\log \text{ of defined Capital variable})}{|1-e^{-\lambda\tau t}|}$.

[33] This is obtained by dividing the indirect marginal effect coefficient for the log (capital-labour ratio) by the corresponding coefficient for the total marginal effects for the same variable.

What this implies is that if a country were mostly in interaction with its border-touching geographic neighbours, there would be a positive network physical capital accumulation effects accruing from interacting within such occurring networks. While the total marginal effect of human capital in model *B* shows that an increase in this index by two-folds is associated with an approximately 0.55 points increase in the income per labour growth rate, holding the other covariates constant. Within this queen-contiguity spatial regime (or specification), it can be inferred that the indirect marginal effect is about 13% of the total marginal effects for the human capital index. While the total marginal convergence effect conditional on the other covariates in this specified model shows that relatively poorer countries catch-up to the richer economies (or retard relative to better previous performances) at the rate of approximately 5.1%[34]. This is approximately two and a half times as much as the 2% hypothesised by the textbook Solow model. Compared to the fixed-effects estimates in estimation *A*, where spatial agglomeration is not considered, we find that the convergence rate is marginally faster by approximately 1.1 percentage unit (with approximately 4.1% as its rate of convergence). The coefficients from the non-spatially adjusted model *A* appear not so different from the main marginal effects in the queen-contiguity spatially adjusted model in estimation *B*. However, it is clear that adjusting for any inherent spatial dimension provides us with a clearer picture of just how important cross-country interdependence is even at a contiguous cross-country neighbourhood level. All the marginal effects for estimations A and *B* appear statistically significant at a minimum statistical level of 5%. This is even after controlling for any arbitrary spatial agglomeration process that may be present in the error process using a squared inverse-distance spatial matrix in estimation *B*.

[34] We obtain this by using the convergence conversion formula, $-\left(\frac{ln(1+\beta)}{\iota\tau}\right)$, introduced in section 1.3.2 above.

		(A) /	(B) Wc/Wi2	(1) WL1/Wi2	(2) WL1_Wc/Wi2	(3) WL2/Wi2	(4) WL2_Wc/Wi2	(5) WL3/Wi2	(6) WL3_Wc/Wi2
Main	**Lag Log of GDP per Labour**	-0.5760***	-0.5653***	-0.5598***	-0.5699***	-0.5638***	-0.5720***	-0.5622***	-0.5713***
	Log of Capital-Labour ratio	0.3050***	0.2806***	0.2823***	0.2837***	0.2840***	0.2845***	0.2847***	0.2841***
	Log of Human Capital Index	0.5767***	0.4786***	0.4424**	0.4800***	0.4701***	0.4837***	0.4419**	0.4981***
Direct	**Lag Log of GDP per Labour**		-0.5699***	-0.5657***	-0.5745***	-0.5691***	-0.5760***	-0.5689***	-0.5741***
	Log of Capital-Labour ratio		0.2867***	0.2891***	0.2897***	0.2904***	0.2902***	0.2919***	0.2894***
	Log of Human Capital Index		0.4817***	0.4462***	0.4831***	0.4738***	0.4865***	0.4463**	0.5000***
Indirect	**Lag Log of GDP per Labour**		-0.0892**	-0.2078***	-0.0822**	-0.1345***	-0.0687*	-0.2555***	-0.0406
	Log of Capital-Labour ratio		0.0457**	0.1086**	0.0420**	0.0692**	0.0349*	0.1339**	0.021
	Log of Human Capital Index		0.0723**	0.1560**	0.0661**	0.1077**	0.0547*	0.1896**	0.033
Total	**Lag Log of GDP per Labour**		-0.6591***	-0.7736***	-0.6567***	-0.7036***	-0.6446***	-0.8244***	-0.6148***
	Log of Capital-Labour ratio		0.3324***	0.3977***	0.3317***	0.3597***	0.3251***	0.4259***	0.3104***
	Log of Human Capital Index		0.5541***	0.6021***	0.5492***	0.5815***	0.5411***	0.6360***	0.5330***
Spatial	**rho** (Coefficient on the spatial lagged dependent)		0.2206***	0.4291***	0.2199**	0.3870***	0.2049**	0.4095***	0.203
	rho_2 (Coefficient on the spatial lagged residual)		0.3302**	0.2815*	0.3727**	0.3163*	0.3970**	0.2931*	0.4311***
Some Statistics	Average Fixed Effects		2.004	1.959	2.014	1.958	2.024	1.955	2.01
	No. of Time Intervals		6	6	6	6	6	6	6
	No. of Countries	122	122	122	122	122	122	122	122
	Total No. of Obs.	732	732	732	732	732	732	732	732
	Model Degree of Freedom	2	125	125	125	125	125	125	125
	No. of Ancillary Parameters		3	3	3	3	3	3	3
	Within R-Square	0.4659943	0.4585582	0.4462433	0.4601094	0.4498521	0.4593622	0.4541153	0.4632114
	AIC	-1722.447	-1732.658	-1741.179	-1732.482	-1737.056	-1729.253	-1737.261	-1723.457
	BIC	-1708.66	-1663.721	-1672.242	-1663.545	-1668.119	-1660.317	-1668.324	-1654.52
	standard deviation of idiosyncratic errors		0.0063***	0.0062***	0.0062***	0.0062***	0.0063***	0.0062***	0.0063***

The dependent variale for all the regressions is the Growth rate of GDP per Labour; ***stands for p<0.001; **stands for p<0.05; *stands for p<0.10; Please take note that the composite weights on the even-numbered estimations above (estimations 2, 4, and 6) have already been defined in Section 3.2.2 of this thesis chapter, while the odd-numbered estimations (1, 3 and 5) represents the uninteracted weights; Also the first of the spatial weights (Wc - queen contiguity; or WL - language networks) represents the one used to lag the dependent variable, while the latter following the backslash (/) represents the squared inverse-distance spatial weight used to lag the error process (Wi2)

Table E – 1.4.3: Economic Growth augmented to show the Estimations with the base-models (A and B) and a Language-type Spatial Agglomeration process

1.4.2.2 – Interpreting the Language based Cultural Network Models

This set of results from estimations *1 through 6* in Table E - 1.4.3 above essentially adopts 3 sub-classes of languages based cultural spatial weights to model the hypothesised spatial dependence inherent in the growth rate of income per labour while still controlling for arbitrary interdependence in the error process. The first subclass of language neighbour matrices (in estimations *1 and 2* with the WL1 and its contiguity interacted variant WL1_Wc, respectively) pair countries as neighbours if they share at least one common language. The second subclass of language neighbour matrix (in estimations *3 and 4* with the WL2 and

its contiguity interacted variant WL2_Wc, respectively) adopts a relatively stricter threshold by pairing countries as neighbours if they share at least three common languages. The third subclass in this subsection (in estimations *5 and 6* with the WL3 and its contiguity interacted variant WL3_Wc, respectively) quite differently, records the total number of languages each pair of countries share in common; thus adopting an intensity-type approach to its specification.

A quick glance at the results from estimations *1 through 6* shows that the non-interacted spatial weight models (*1, 3 and 5*) appear to have stronger and more statistically significant spatial coefficients compared to the interacted spatial weight specifications (*2, 4 and 6*). The direct marginal effects however show a reversal of this observation: estimations *2, 4 and 6* seem to have larger coefficients compared to estimations *1, 3 and 5*. This is expected - mostly because there are fewer channels for interconnections when the basis for cross-sectional interaction is more restrictive - additionally complicated by the condition that each language pair must also be contiguous neighbours. Bearing this in mind, we have chosen to include this composite specification so that we can examine the hypothesis that countries are more spatially connected when they are both culturally similar and within contiguous proximity. As it appears in this set of estimation results, they are not. It appears that countries are interacting on either fronts (culturally or geographically), but not necessarily jointly on both fronts. This provides some statistical affirmation of our earlier position that countries are increasingly interacting beyond their immediate geographical horizons with increasing advancements in technological knowhow and other avenues that have been facilitating both political and socioeconomic integration such as modernisation, cross-country productive-resource mobility and communication.

From estimation *1* we find that the total marginal contribution of physical capital (capital per labour) is approximately 0.3977. This can be translated as: a 10% increase in the level of capital per labour is on the average associated with an increase in the income per capita growth rate by approximately 0.0398 units, all other factors held constant. The ratio share of own-country (direct marginal effects) to inter-country (indirect marginal effects) share in the total marginal effects is approximately 73% to 27%, respectively. The total marginal effect of human capital (as measured by the human capital index) on the other hand, suggests that an increase in the human capital index of a country by two-folds (or 100%), on the average, is accompanied by an increase in the income per capita

growth rate of about 0.602 units. The total marginal effect of the previous period of income per capita remains the biggest individual growth source, which is consistent throughout this work regardless of the spatial network weight adopted. It can be observed that, all things being equal, a two-fold increase in the previous level of income per labour in a country is on the average accompanied by a decrease in the contemporaneous growth rate of income per capita by about 0.774 units. The total income convergence rate implied by this coefficient is approximately 7.1%. It can also be inferred from this that there is significant equilibrium adjustment pressure from external cross-country interaction influences that most likely could be underestimated if global integration is not adequately controlled for. From a cross-sectional perspective, the implication of an indirect convergence coefficient is that richer countries are slowing down on the average relative to their poorer country counterparts, within each specified network. Even though this convergence is of a lesser magnitude compared to that accruing to countries on the average as a result of their own-country (direct) conditional convergence effects. Apart from the indirect marginal effects in estimations *4 and 6*, we find that all the direct, indirect and total marginal effects are statistically significant at a minimum statistical level of 5%.

1.4.2.3 – Interpreting the Ethnic based Cultural Network Models

In similar fashion to the language based spatial networks discussed section 1.4.2.2 above, there are essentially three sub-classes of ethnicity based spatial weights in this sub-section. Designed in exactly the same way: our first subclass of ethnic neighbour matrices (in estimations *7 and 8* with the WE1 and its contiguity interacted variant WE1_Wc, respectively) pair countries as neighbours if they share at least one common ethnic group; the second subclass (in estimations *9 and 10* with the WE2 and its contiguity interacted variant WE2_Wc, respectively) adopts a relatively stricter threshold by pairing countries as neighbours if they share at least three common ethnic groups; finally the third subclass in this subsection (in estimations *11 and 12* with the WE3 and its contiguity interacted variant WE3_Wc, respectively) counts the total number of ethnic groups shared between country pairs, thus advocating an intensity-type approach in its specification as well.

Looking at the estimation results in Table F - 1.4.4 below, we find that the story is very much similar to the discussions already made for the language based cultural matrix on the non-composite geography based network estimations in the previous table (Table E - 1.4.3). As previously emphasized, the more packed the spatial relationship portrayed or hypothesised by the spatial weight specification (network), the more emphatic the cumulative indirect spatial effect as estimated in the corresponding coefficients of the respective proximate growth factors. Also because the patterns and coefficients are very similar across the three subclasses of spatial weights here as was the case with the languages case in Table E - 1.4.3, we focus on interpreting the first of them, estimation *7*. From this subclass of estimations, we also find that the coefficients are generally bigger and that the spatial coefficients (*rho*) on the lag the dependent variable (income per labour growth) is generally larger and consistently statistically significant at the 1% level of significance excluding the *rho* coefficient for estimation *10*, which is significant at 5%. From estimation *7* (for reference purposes) we find that the inferred convergence rate is approximately 5.9% holding the other model's covariates constants. In marginal terms, the indirect marginal effect is about 28.35% ($^{-0.2012}/_{-0.7096}$) of the total marginal effect for the convergence coefficient used in obtaining the rate[35]. The total marginal contribution of the capital-labour ratio to the income per worker growth rate also shows that a two-fold increase in the capital-labour ratio is associated with about 0.35 percentage point increase in the income per worker growth rate. The indirect marginal effects appear to contribute about 28.74% to this total marginal effect. The inferred marginal product of capital from the total marginal effects is about 0.495, which is still approximately the same as that inferred from estimation *A* in Table E - 1.4.3 above. While the total marginal effects coefficient associated with the human capital index shows that an increase in the human capital index by 10% is associated with about 0.0544 percentage point increase in the income per labour growth rate, other covariates held constant. The ratio of the indirect marginal effects to total accrual effects from neighbouring networks is about 27.29%.

Taking a look at the entire set of results, from both Table E - 1.4.3 and Table F - 1.4.4, we find that the ethnic based spatial model specifications appear to

[35] This is not to say that this ratio applies to the convergence rate itself, as both terms are not deduced using the same formulas.

have evidence of stronger global spatial correlations in the income per labour growth rate. This is not surprising, as ethnic groups tend to be a more stable localised medium of identifying groups within country boundaries. As earlier highlighted in section 1.4.1 above, ethnic groups identify themselves mostly ancestrally and by homelands giving them a sense of locality, territoriality and stability. As a result, they would make for deeper but fewer spatial links between spatial units (in this case: countries) compared to languages, which would make for more but shallower connections between country units. This is why even after using an alternative double-criteria approach (cultural and contiguity – estimations *2, 4, 6 and 8, 10, 12*), we find that the latter (the ethnic groups subclass) appears consistently statistically significant with respect to the spatial components compared to the languages subclass. We also find that the ethnic groups subclass of estimations provide the best model-set with the AIC (Akaike Information Criteria) and BIC (Bayesian Information Criteria) generally having lower information criteria coefficients in comparison to both the languages subclass and the base models.

		(7) WE1/Wi2	(8) WE1_Wc/Wi2	(9) WE2/Wi2	(10) WE2_Wc/Wi2	(11) WE3/Wi2	(12) WE3_Wc/Wi2
Main	**Lag Log of GDP per Labour**	-0.5000***	-0.5611***	-0.5185***	-0.5619***	-0.5087***	-0.5549***
	Log of Capital-Labour ratio	0.2431***	0.2802***	0.2610***	0.2835***	0.2554***	0.2769***
	Log of Human Capital Index	0.3913**	0.4663***	0.4234**	0.4691***	0.3763**	0.4599***
Direct	**Lag Log of GDP per Labour**	-0.5083***	-0.5664***	-0.5258***	-0.5665***	-0.5181***	-0.5614***
	Log of Capital-Labour ratio	0.2502***	0.2866***	0.2681***	0.2898***	0.2635***	0.2841***
	Log of Human Capital Index	0.3961**	0.4699***	0.4278***	0.4721***	0.3816**	0.4644***
Indirect	**Lag Log of GDP per Labour**	-0.2012***	-0.0823***	-0.1314***	-0.0639**	-0.2536***	-0.0891***
	Log of Capital-Labour ratio	0.1009**	0.0421**	0.0682***	0.0336*	0.1320**	0.0457**
	Log of Human Capital Index	0.1479**	0.0655**	0.1021***	0.0494*	0.1743**	0.0704**
Total	**Lag Log of GDP per Labour**	-0.7096***	-0.6487***	-0.6573***	-0.6304***	-0.7717***	-0.6505***
	Log of Capital-Labour ratio	0.3511***	0.3287***	0.3362***	0.3233***	0.3954***	0.3297***
	Log of Human Capital Index	0.5440***	0.5354***	0.5300***	0.5216***	0.5559**	0.5348***
Spatial	**rho** (Coefficient on the spatial lagged dependent)	0.5785***	0.2428***	0.4945***	0.2159**	0.5912***	0.3362***
	rho_2 (Coefficient on the spatial lagged residual)	0.0407	0.3105*	0.0863	0.3527**	0.066	0.3051*
Some Statistics	Average Fixed Effects	1.855	1.981	1.816	1.951	1.817	1.96
	No. of Time Intervals	6	6	6	6	6	6
	No. of Countries	122	122	122	122	122	122
	Total No. of Obs.	732	732	732	732	732	732
	Model Degree of Freedom	125	125	125	125	125	125
	No. of Ancillary Parameters	3	3	3	3	3	3
	Within R-Square	0.401866	0.4577431	0.431393	0.462198	0.3948877	0.4558194
	AIC	-1806.558	-1737.474	-1787.063	-1732.612	-1801.512	-1741.87
	BIC	-1737.621	-1668.537	-1718.126	-1663.675	-1732.575	-1672.934
	standard deviation of idiosyncratic errors	0.0056***	0.0062***	0.0058***	0.0062***	0.0057***	0.0062***

The dependent variale for all the regressions is the Growth rate of GDP per Labour; ***stands for p<0.001; **stands for p<0.05; *stands for p<0.10; Please take note that the composite weights on the even-numbered estimations above (estimations 8, 10, and 12) have already been defined in Section 3.2.2 of this thesis chapter, while the odd-numbered estimations (7, 9 and 11) represents the uninteracted weights; Also the first of the spatial weights (Wc - queen contiguity; or WL - language networks) represents the one used to lag the dependent variable, while the latter following the backslash (/) represents the squared inverse-distance spatial weight used to lag the error process (Wi2)

Table F – 1.4.4: Economic Growth augmented to show the Estimations with the Ethnic-type Spatial Agglomeration process

1.4.3 – General Implications of the Findings

This paper arrives at a number of estimations following the presentation tradition of the mainstream applied spatial econometrics literature, which normally compare results from differing (or contending) spatial weighting schemes (see Leenders, 2002). As a result, a general summary of the findings highlighting some relevant implications as deducible with respect to the methodology adopted herein. The relevance of operationalizing spatial agglomeration into present-day growth modelling cannot be overstated. Not only do spatially augmented models provide the statistical benefit of handling methodological bottlenecks such as omitted variable bias (arising from desirable but difficult to measure clustering phenomena), model misspecification, and etc. They also provide growth models with avenues for approximating and observing the nature of spatial connectivity between countries in an ever-increasing globally connected world economy.

Firstly, we generally find that the more spatial units per spatial network, on the average and all things being equal, the more substantial (in magnitude) the indirect marginal effects or spatial spillovers would be. This is especially useful for understanding the nature of the spatial distribution of growth in a world economy with increasing avenues for significant interaction between countries. Specifically, it can be observed that when the extent of spatial agglomeration is minimal as demonstrated using the queen contiguity matrix, the total marginal effect of the capital accumulation variables (physical and human) are relatively smaller compared to when the extent of spatial agglomeration is more, as defined using our cultural network design. The reason why geography and cultural elements may provide statistically significant measures for defining global spatial inter-country connectivity might be because they both underlie many other conventional means via which countries interrelate in the world today. Whether these range from macro-level politico-economic interrelationships and associations to micro-level private ideological mediums of interconnections, we find that they can be traced back to basic connections based on geographical proximities, cultural similarities or a combination (or interaction) of both. This is particularly crucial even as the global connections between countries in a 21st century are non-decreasing in intensity with recent leaps in technological advancements and the continued integration of global policy making (and implementation) processes. Countries today seem ever closer and the concept of

global space is geometrically shrinking even though the geographical metric of distance itself has actually remained unchanged. An alternative way of conceiving these spatial connections especially those based on our culturally defined metrics is to consider the estimated (*rho*) coefficients as varying degrees of fundamental similarities (although, not structurally identified). Considering them this way puts into perspective the argument that these countries, stemming from differing geopolitical institutions, may not necessarily be in direct interaction. However, they may represent an estimation of the measure of institutional (or structural) coordination. This form of interpretation can be duly extended to the concept of geographic nearness at the national level. These cultural and geographic spatial metrics are by their very nature institutionally defined and recognisable – as a result the statistically significant spatial cluster coefficients shown by the respective *rhos* could be capturing global coordination effects at the cross-county level. Interacting these *rhos* with the relevant right-hand side coefficients to obtain the indirect effects helps one understand the extent and direction of the corresponding coordination effects with respect to the type, form, or dimension of spatial interaction being captured. From the results reported above, it can be observed that the spatial patterns in the economic growth process explained by the geographical spatial weights are quite different from the variations in the growth process explained by the culturally defined spatial weights. This is deducible by considering the estimation results from the geography-culture interacted spatial weight models (*2, 4, 6, 8, 10 and 12*). We find that the spatial effects from both the spatial *rho* coefficient and the indirect marginal effects tend to decline in both magnitude and statistical significance when these interacted spatial weight specifications are adopted. If it were the case that they were explaining similar dimensions of variations in the economic growth rate variable, then we would expect reinforcing effects, where the magnitude and statistical significance of the spatial *rho* coefficient and the indirect marginal effects were increasing. Our results indicate that countries appear more similar from a cultural level through the spatial distributive patterns of ethnic groupings and languages than they are via patterns of geographic clusters, as shown by the size of the spatial *rho* coefficients.

Secondly, the relevance of the indirect marginal effects in contribution to the estimated total marginal effects for the proximate growth factors (human capital, physical capital, and growth convergence) tends to range roughly between about

20% to as much as 30%. This suggests that even when the global cross-sectional interconnections between countries are not as strong, there remains about a fifth of the growth process accruing from the summation of global indirect spatial synergic effects. This is why ignoring this potential or possibility for inter-country connections in a panel growth analysis could not only lead to substantial misspecification and bias, but the consequent results might equally be misleading (both to policy and observing analysts) as it would likely under-inform the factors driving the process of economic growth.

Thirdly and closely concerned with the preceding paragraph, a framework ignoring such increasing country level interconnections would fail to fully understand why poor countries get poorer while the richer ones get even richer. It is most likely the case that relatively richer countries in specific spatial networks on the average internalize a lot more returns to their productive resources (such as human capital and physical capital) compared to their relatively poorer counterparts. These relatively richer countries also have the enabling economic environment that attracts global productive resources thus taking advantage of the increased access to resources warranted by an increasingly open and interconnected global economic frontier. This is irrespective of the position that countries might in fact remain separate geographic entities and their corresponding political administrations may have distinctly composed institutional fabrications. Such an open and expansive global economic landscape relatively favours more advanced economies compared to the less developed ones. As a result the convergence phenomenon does not guarantee that poorer economies will catch up to the richer ones as these richer economies are better placed to galvanise competitive productive resources away from these developing countries to theirs.

Fourthly, it is interesting to see how the points raised so far affect the potential of poorer countries to adjust to their respective steady state income levels as embedded in the conditional convergence hypothesis? Although the model estimated herein has not included an average network income level to which each country is adjusting towards. There remains a convergence adjustment pressure at an inter-country level, which may cause countries to adjust much faster than they would have in a hypothetically closed economic setup. Hence, it comes as no surprise that the degree of cross-sectional connectedness is directly proportional to the size and sometimes-evident steady state over-adjustments

that countries face especially in the presence of current increasing globalization trends. This implies that increasing globalization has an un-ignorable part to play in explaining why countries in general take a considerably longer time to recover from global shocks to their economic performances.

Fifthly, the debate on diverging rates of convergence compared to what is theoretically hypothesized by the traditional Solow model remains an empirical quandary[36]. Although the rates obtained herein are on the average faster than those observed in the theoretical and cross-sectional literature (a rate of 2% - see Barro & Sala-I-Martin, 1991; MRW, 1992), they are however comparably slower than those often found in the panel growth literature that record faster convergence rates of up to 20% (see Islam, 1995; de la Fuente, 2002, for instance). On the average, the results herein point to a total convergence rate ranging between approximately 0.04 and 0.09 (4% to 9%). These rates seem within the range of those obtained in Islam (1995) who also finds convergence rates approximately between 4% and 11% for a panel re-estimation of MRW's (1992) study; and Badinger et al (2004) who find a rate of 7% (0.07) for a panel of a 196 EU regions. From the convergence coefficients we also inferred that although conditional convergence is statistically significant to economic growth, its effect could be traced to a combination of own-country and inter-country network convergence components.

1.4.4 – Concluding Remarks

To recap, the results from this study provide indications that the marginal contributions of physical and human capital may be underestimated at the global level or overestimated at the own-country level when spatial dependence or other forms of spatial interactions are not accounted for beyond mere country-specific effects. We find that the global integrative contributions to both physical and human capital are corroboratory to the direct own-country effects of these growth factors. Notwithstanding, there are some concerns that because richer countries are more likely enabled to accumulate capital and other key productive resources, they may quite easily crowd-out these essential productive resources by competitively implementing favourable growth enhancing policies geared at

[36] See Quah, 1996a, for some related technical discussions on the subject.

attracting such resources away from relatively poorer source-countries. The implication of this for a globally relevant resource such as human capital may be undervalued and under-discussed especially when spatial connectivity is not adequately controlled for. Our findings also point to statistically substantial spatially integrative effects with respect to the level of human capital. It is quite a profound growth factor (depending on the extent of inter-country connectivity) and may explain why development recovery remains a challenge for protractedly backward economies such as those in the Sub-Saharan African region.

As a whole, this study has also been able to extend the applied spatial growth literature by considering alternative forms and measures of spatial agglomeration that are consistent with the increasing degree of inter-country integration oftentimes relegated to the globalization and trade discourses. Although focused on a cultural dimension, we find that global spatial diffusion patterns of ethnic groups and languages provide statistically significant explanations for why growth divergences may continue to persist between the developed and developing economies of the world.

1.4.5 – Some Weaknesses

There are a number of weaknesses in this research but two main empirical weaknesses are worth mentioning, as addressing them would certainly improve the quality of a research as ours.

Firstly, the paper does not explicitly address the issue of reverse causality. As a standard empirical economic growth concern, productive capital (physical and human) resources have the potential to affect and be affected by the level or scale of economic activity (see the King & Levine, 1994 'Capital fundamentalism' critique). As a consequence, we cannot conclude that the results obtained and discussed herein are causal. However, as earlier highlighted in the Introduction section of this chapter, this exercise is mostly illustrative (or exploratory). We are mostly seeking to assess (from a systemic approach) the spatial patterns of growth from a perspective of increasing cultural globalization or country level interconnections. In such a globally integrated framework, the spatial econometric methodology is quite relevant at it would reduce the rather likely omitted variable problem inherent in non-spatially adjusted models (see Pace &

LeSage, 2010). Thus it is contextually an improvement to a standard non-spatially adjusted linear model.

One way to address the causality issue in this context, would be to explore possible pathways via which the cultural dynamics exploited in the cultural spatial networks adopted herein could sensibly affect economic activity directly and indirectly through the endogenous proximate growth determiners. A plausible conduit, could be through global policy interactions or global policy sharing, global migration flows and trade at the macro level or through specific FDI channels in a firm-level type study. A major source of challenge in exploring this angle would be trying to obtain such data especially for developing countries (such as in the Sub-Sahara African – SSA – region) where data paucity is a major concern. Equally obvious and difficult, would be the challenge of measuring policy type variables which are not readily available for quantitative-style analysis. Pairing countries based on similarity in policy orientation and implementation would resultantly be subjective. This could pose even further challenges given that policy phenomena are not that temporally (time) varying, as well; thus calling to question its functional relevance in an analytical panel (country-time) framework. An alternative proposal could be to motivate and write an estimation routine for a spatial system GMM model that can provide coefficients for all three spatial effects desired in this chapter herein (direct, indirect and total effects). Such a model would be able to harness the temporal and spatial moment conditions inherent in spatio-temporal frameworks that can efficiently tackle endogeneity while still providing the kind of results sought after in our work herein. It would essentially be a methodological spatial adaptation of the Arellano-Bover/Blundell-Bond estimator (see Kukenova & Neuchatel, 2008, for some technical notes on current advancements on a very similar subject matter).

Secondly, providing a framework for assessing and controlling for the pattern of global spatial clustering does not necessarily provide any particular information on the actual path through which statistically significant spatial patterns translate unto or help identify underlying (fundamental) factors of the global growth process. In essence, it would be more desirable to have a more structurally specified framework that can help to identify paths or mechanisms explaining the actual link between estimated global spatial cluster patterns and the patterns of economic growth itself. However, because the focus of our study rests mostly on attempting to identify global shares of the proximate growth factors while

controlling for spatial global integration. It was not necessarily aimed at attempting to identity these spatial connections with growth via fundamental growth factors such as institutional mechanisms and quality.

References

Abreu, M., De Groot, H. L. F., & Florax, R. J. G. M. (2004). *Space and Growth: A Survey of Empirical evidence and Methods* (No. TI 2004-129/3). *Tinbergen Institute Discussion Paper*.

Abreu, M., De Groot, H. L. F., & Florax, R. J. G. M. (2005). Space and Growth: A Survey of Empirical Evidence and Methods. *Région et Développement*, *21*.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Anselin, L. (2001). Spatial Econometrics. In B. H. Baltagi (Ed.), *A Companion to Theoretical Econometrics* (pp. 310–330). Blackwell Publishing Ltd.

Anselin, L., Varga, A., & Acs, Z. (1997). Local Geographic Spillovers between University Research and High Technology Innovations. *Journal of Urban Economics*, *42*(3), 422–448. doi:10.1006/juec.1997.2032

Autant-Bernard, C., & LeSage, J. P. (2011). Quantifying Knowledge Spillovers using Spatial Econometric Models. *Journal of Regional Science*, *51*(3), 471–496. doi:10.1111/j.1467-9787.2010.00705.x

Badinger, H., Muller, W. G., Tondl, G., & Müller, W. G. (2004). Regional Convergence in the European Union, 1985-1999: A Spatial Dynamic Panel Analysis. *Regional Studies*, *38*(3), 241–253. doi:10.1080/0034340042000211105

Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, *71*(1), 135–171.

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, *106*(May 1991), 407–443.

Barro, R. J., & Lee, J.-W. (2010). *A New Dataset of Educational Attainment in the World, 1950-2010* (No. 15902 - NBER Working Paper Series).

Barro, R. J., & Sala-i-Martin, X. (1991). Convergence across States and Regions. *Brookings Papers on Economic Activity*, *22*(1), 107–182. doi:10.2307/2534639

Beck, N., Gleditsch, K. S., & Beardsley, K. (2006). Space is more than Geography: Using Spatial Econometrics in the Study of Political Economy. *International Studies Quarterly*, *50*, 27–44.

Bernard, A. B., & Durlauf, S. N. (1995). Convergence in International Output. *Journal of Applied Econometrics*, *10*(2), 97–108. doi:10.1002/jae.3950100202

Bloom, D. E., Sachs, J. D., Collier, P., & Udry, C. (1998). Geography, Demography, and Economic Growth in Africa. *Brookings Papers on Economic Activity*, *1888*(2), 207–295. doi:10.2307/2534695

Bond, S. R., Hoeffler, A., & Temple, J. R. W. (2001). *GMM Estimation of Empirical Growth Models* (No. CEPR Discussion Paper No. 3048).

Breitung, J., & Pesaran, M. H. (2008). Unit Roots and Cointegration in Panels. In L. Mátyás László Mátyás & P. Sevestre (Eds.), *The Econometrics of Panel Data* (pp. 279–322). Springer Berlin Heidelberg.

Bromley, D. W. (1995). Development Reconsidered: The African Challenge. *Food Policy*, *20*(5), 425–438. doi:10.1016/0306-9192(95)00036-E

Burt, R. S. (1987). Social Contagion and Innovation: Cohesion versus Structural Equivalence. *American Journal of Sociology*, *92*(6), 1287. doi:10.1086/228667

Carlino, G. A., & Mills, L. O. (1993). Are U.S. Regional Incomes Converging? A Time Series Analysis. *Journal of Monetary Economics*, *32*, 335–346.

Caselli, F., Esquivel, G., & Lefort, F. (1996). Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics. *Journal of Economic Growth*, *1*(3), 363–389. doi:10.1007/BF00141044

Chen, Y. (2012). On the Four Types of Weight Functions for Spatial Contiguity Matrix. *Letters in Spatial and Resource Sciences*, *5*(2), 65–72. doi:10.1007/s12076-011-0076-6

Cliff, A. D., & Ord, K. J. (1973). *Spatial Autocorrelation* (Volume 5.). London: Pion Limited.

Cliff, A. D., & Ord, K. J. (1981). *Spatial Processes: Models and Applications*. London: Pion Limited.

Corrado, L., & Fingleton, B. (2012). Where is the economics in spatial econometrics? *Journal of Regional Science*, *52*(2), 210–239. doi:10.1111/j.1467-9787.2011.00726.x

de la Fuente, A. (2002). On the sources of convergence: A Close Look at the Spanish Regions. *European Economic Review*, *46*(3), 569–599. doi:10.1016/S0014-2921(01)00161-1

Ding, S., & Knight, J. (2008). Can the Augmented Solow Model Explain China's Economic Growth? A Cross-Country Panel Data Analysis. *Journal of Comparative Economics*, *37*(3), 432–452.

Dreher, A. (2006). Does Globalization affect Growth? Evidence from a New Index of Globalization. *Applied Economics*, *38*(10), 1091–1110. doi:10.1080/00036840500392078

Durlauf, S. N., Johnson, P. A., & Temple, J. R. W. (2005). Growth Econometrics. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (Vol. 1A, pp. 555–677). Elsevier. doi:http://dx.doi.org/10.1016/S1574-0684(05)01008-7

Easterly, W., & Levine, R. (2003). Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics*, *50*(1), 3–39. doi:10.1016/S0304-3932(02)00200-3

Elhorst, P. J. (2003). Specification and Estimation of Spatial Panel Data Models. *International Regional Science Review*, *26*(3), 244–268. doi:10.1177/0160017603253791

Elhorst, P. J. (2010). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, *5*(1), 9–28. doi:10.1080/17421770903541772

Elhorst, P. J. (2014). *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*. Springer. Retrieved from http://link.springer.com/content/pdf/10.1007/978-3-642-40340-8.pdf

Feenstra, R. C., Inklaar, R., & Timmer, M. (2013). *The Next Generation of the Penn World Table* (No. 19255). *NBER Working Paper Series*. doi:10.1017/CBO9781107415324.004

Fingleton, B. (2001a). Equilibrium and Economic Growth: Spatial Econometric Models and Simulations. *Journal of Regional Science*, *41*(1), 117–147. doi:10.1111/0022-4146.00210

Fingleton, B. (2001b). Theoretical Economic Geography and Spatial Econometrics: Dynamic Perspectives. *Journal of Economic Geography*, *1*(2), 201–225. doi:10.1093/jeg/1.2.201

Fingleton, B., & López-Bazo, E. (2006). Empirical Growth Models with Spatial Effects. *Papers in Regional Science*, *85*(2), 177–198. doi:10.1111/j.1435-5957.2006.00074.x

Frankel, J. A., & Romer, D. (1999). Does Trade Cause Growth? *The American Economic Review*, *89*(3), 379–399. doi:10.1257/aer.89.3.379

Franzese, R. J., & Hays, J. C. (2007). Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data. *Political Analysis*, *15*(2), 140–164.

Fujita, M., Krugman, P., & Venables, A. J. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge: The MIT Press. doi:10.2307/1061487

Gallup, J., & Sachs, J. (2001). The Economic Burden of Malaria. *The American Journal of Tropical Medicine and Hygiene*, *64*(1, 2), 85–96.

Garretsen, H., & Martin, R. (2010). Rethinking (New) Economic Geography Models: Taking Geography and History more Seriously. *Spatial Economic Analysis*, *5*(2), 37–41. doi:10.1080/17421771003730729

Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do Institutions Cause Growth? *Journal of Economic Growth*, *9*(3), 271–303. doi:10.1023/B:JOEG.0000038933.16398.ed

Goodfriend, M., & McDermott, J. (1995). Early Development. *The American Economic Review*, *85*(March), 116–33. Retrieved from http://www.jstor.org/stable/2117999

Grossman, G. M., & Helpman, E. (1991). Trade, Knowledge Spillovers, and Growth. *European Economic Review*, *35*(2), 517–526.

Hall, R., & Jones, C. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics*, *114*(1), 83–116.

Haynes, K., & Fotheringham, A. (1984). *Gravity and Spatial Interaction Models* (Vol. 2.). Beverly Hills: Sage Publications.

Horowitz, D. L. (1985). *Ethnic Groups in Conflict*. University of California Press.

Huggins, R., & Thompson, P. (2014). A Network-based view of Regional Growth. *Journal of Economic Geography*, *14*(3), 511–545. doi:10.1093/jeg/lbt012

Islam, N. (1995). Growth Empirics: a Panel Data Approach. *The Quarterly Journal of Economics*, (November), 1127–1170.

Kapetanios, G., & Pesaran, H. M. (2005). *Alternative Approaches to Estimation and Inference in Large Multifactor Panels: Small Sample Results with an Application to Modelling of Asset Returns* (No. 1416 - CESifo Working Paper Series).

Kelejian, H. H., & Prucha, I. R. (2010). Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. *Journal of Econometrics*, *157*(1), 53–67. doi:10.1016/j.jeconom.2009.10.025

Klenow, P. J., & Rodriguez-Clare, A. (1997). The Neoclassical Revival in Growth Economics: Has it Gone too Far? *NBER Macroeconomics Annual*, *12*(January), 73–114. doi:Abstract

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, *99*(3), 483–499.

Kukenova, M., & Monteiro, J.-A. (2008). *Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation* (Munich Personal RePEc Archive Working Papers No. 13405 - Munich Personal RePEc Archive (MPRA) Working Paper).

Lacombe, D. J. (2004). Does Econometric Methodology Matter? An Analysis of Public Policy Using Spatial Econometric Techniques. *Geographical Analysis*, *36*(2), 105–118. doi:10.1111/j.1538-4632.2004.tb01128.x

Leenders, R. T. A. J. (2002). Modeling Social Influence through Network Autocorrelation: Constructing the Weight Matrix. *Social Networks*, *24*(1), 21–47. doi:10.1016/S0378-8733(01)00049-1

LeSage, J. P. (2009). *Introduction to Spatial Econometrics*. (N. Balakrishnam & W. R. Schucany, Eds.)*Chapman & Hall/CRC*. Taylor & Francis Group. doi:10.4000/rei.3887

LeSage, J. P., & Pace, R. K. (2010). *The Biggest Myth in Spatial Econometrics* (No. 1725503). *Social Science Research Network*. doi:10.2139/ssrn.1725503

Lewis, P. M., Simons, G. F., & Fennig, C. D. (Eds.). (2014). *Ethnologue: Languages of the World, Seventeenth Edition* (17th ed.). Dallas, Texas: SIL International. Retrieved from http://www.ethnologue.com/17/

Lopez-Bazo, E., Vaya, E., & Artis, M. (2004). Regional Externalities and Growth : Evidence from European Regions. *Journal of Regional Science*, *44*(1), 43–73. doi:10.1111/j.1085-9489.2004.00327.x

Lucas, R. E., & Lucas Jr, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, *22*, 3–42. doi:10.1016/0304-3932(88)90168-7

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, *107*(2), 407–437.

Martin, R., & Sunley, P. (1998). Slow Convergence? The New Endogenous Growth Theory and Regional Development. *Economic Geography*, *74*(3), 201–227. doi:10.1111/j.1944-8287.1998.tb00113.x

McArthur, J. W., & Sachs, J. D. (2001). *Institutions and Geography: Comment on Acemoglu, Johnson and Robinson (2000)* (No. 8114 - NBER Working Paper Series). Retrieved from http://www.nber.org/papers/w8114

McQuinn, K., & Whelan, K. (2007). Solow (1956) as a Model of Cross-Country Growth Dynamics. *Oxford Review of Economic Policy*, *23*(1), 45–62. doi:10.1093/oxrep/grm009

Moreno, R., & Trehan, B. (1997). Location and the Growth of Nations. *Journal of Economic Growth*, *2*(4), 399–418. doi:10.1023/A:1009741426524

North, D. (2005). *Understanding the Process of Economic Change*. Princeton University Press.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

North, D., & Weingast, B. (1989). Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth Century England. *The Journal of Economic History*, *49*(4), 803–832.

Olsson, O., & Hibbs, D. A. (2005). Biogeography and Long-run Economic Development. *European Economic Review*, *49*(4), 909–938. doi:10.1016/j.euroecorev.2003.08.010

Pace, R. K., & LeSage, J. P. (2010). Omitted Variable Biases of OLS and Spatial Lag Models. In P. Antonio, J. Le Gallo, R. N. Buliung, & S. Dall'erba (Eds.), *Progress in Spatial Analysis: Methods and Applications* (pp. 17–28). Springer.

Pesaran, M. (2004). *General Diagnostic Tests for Cross Section Dependence in Panels* (No. 1240 (IZA Discussion Paper)).

Phillips, P. C. B., & Moon, H. R. (2000). Nonstationary Panel Data Analysis: An Overview of Some Recent Developments. *Econometric Reviews*, *19*(3), 263–286.

Phillips, P. C. B., & Sul, D. (2007). Some Empirics on Economic Growth under Heterogeneous Technology. *Journal of Macroeconomics*, *29*(3), 455–469. doi:10.1016/j.jmacro.2007.03.002

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. *World Development*, *22*(9), 1325–1343. doi:10.1016/0305-750X(94)90007-8

Quah, D. T. (1996a). Empirics for Economic Growth and Convergence. *European Economic Review*, *40*(6), 1353–1375. doi:10.1016/0014-2921(95)00051-8

Quah, D. T. (1996b). Regional Convergence Clusters across Europe. *European Economic Review*, *40*, 951–958. doi:http://dx.doi.org/10.1016/0014-2921(95)00105-0

Ramírez, M. T., & Loboguerrero, A. M. (2002). *Spatial Dependence and Economic Growth: Evidence from a Panel of Countries* (No. 206). *Borradores de Economía*.

Redding, S., & Venables, A. J. (2004). Economic Geography and International Inequality. *Journal of International Economics*, *62*(1), 53–82. doi:10.1016/j.jinteco.2003.07.001

Romer, P. M. (1986). Increasing Returns and Long-run Growth. *The Journal of Political Economy*, *94*(5), 1002–1037. Retrieved from http://www.jstor.org/stable/10.2307/1833190

Shehata, E. A. E. (2013). SPGMMXT: Stata Module to Estimate Spatial Panel Autoregressive Generalized Method of Moments Regression. *Statistical Software Components*.

Simmons, B. A., & Elkins, Z. (2004). The Globalization of Liberalization: Policy Diffusion in the International Political Economy. *American Political Science Review*, *98*(1), 171–189. doi:10.1017/S0003055404001078

Starr, H. (1991). Democratic Dominoes: Diffusion Approaches to the Spread of Democracy in the International System. *Journal of Conflict Resolution*, *35*(2), 356–381. doi:10.1177/0022002791035002010

Temple, J. (1999). The New Growth Evidence. *Journal of Economic Literature*, 37(1), 112–156.

Verbeek, M. (2004). *A Guide to Modern Econometrics* (2nd ed.). John Wiley & Sons, Ltd.

Weidmann, N. B., Rød, J. K., & Cederman, L.-E. (2010). Representing Ethnic Groups in Space: A New Dataset. *Journal of Peace Research*, *47*(4), 491–499. doi:10.1177/0022343310368352

Appendix

Appendix 1.1 – Definition of Regression Variables

Variable	Units of Measure	Variable(s) Used from Original Dataset	Source
Lag of logged Income per capita	Log of ratio units	(rgdpna)/(emp)	Penn World Table version 8.0: http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0
Logged Capital per Labour	Log of ratio units	(rkna)/(emp)	Penn World Table version 8.0: http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0
Growth in Income per capita	Growth rate units	(rgdpna)/(emp)	Penn World Table version 8.0: http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0
Logged Human Capital Index	Log of index units bound above zero	hc	Penn World Table version 8.0: http://www.rug.nl/research/ggdc/data/pwt/pwt-8.0

Table G – Appendix-Table 1.1: Definition of Variables Used in the Regression Analyses

Appendix 1.2 – List of Countries Contained in the Analyses' Sample

Sample of Countries Involved in this Chapter's Analyses			
Albania	El Salvador	Lithuania	Slovenia
Argentina	Estonia	Malawi	South Africa
Armenia	Fiji	Malaysia	Spain
Australia	Finland	Mali	Sri Lanka
Austria	France	Mauritania	Sudan
Bahrain	Gabon	Mauritius	Swaziland
Bangladesh	Gambia, The	Mexico	Sweden
Belgium	Germany	Moldova	Switzerland
Benin	Ghana	Mongolia	Syrian Arab Republic
Bolivia	Greece	Morocco	Taiwan
Botswana	Guatemala	Mozambique	Tajikistan
Brazil	Honduras	Namibia	Tanzania
Bulgaria	Hungary	Nepal	Thailand
Burundi	India	Netherlands	Togo
Cambodia	Indonesia	New Zealand	Trinidad and Tobago
Cameroon	Iran, Islamic Rep.	Niger	Tunisia
Canada	Iraq	Norway	Turkey
Central African Republic	Ireland	Pakistan	Uganda
Chile	Israel	Panama	Ukraine
China	Italy	Paraguay	United Kingdom
Colombia	Jamaica	Peru	United States
Congo, Dem. Rep.	Japan	Philippines	Uruguay
Congo, Rep.	Jordan	Poland	Venezuela, RB
Costa Rica	Kazakhstan	Portugal	Vietnam
Cote d'Ivoire	Kenya	Romania	Zambia
Croatia	Korea, Rep.	Russian Federation	Zimbabwe
Cyprus	Kuwait	Rwanda	
Czech Republic	Kyrgyz Republic	Saudi Arabia	
Denmark	Lao PDR	Senegal	
Dominican Republic	Latvia	Sierra Leone	
Ecuador	Lesotho	Singapore	
Egypt, Arab Rep.	Liberia	Slovak Republic	

Table H – Appendix-Table 1.2: List of Countries Contained in the Regression Analyses

Appendix 1.3 – List of Ethnic Groups adopted in the Regression Analyses

List of Ethnic Groups Used in Cultural Spatial Regressions					
Achuale	Bechuanas	Flemings	Kirghis	Muisca, Paese, Cofan, etc.	Swazi
Afghans	Beja	French	Kisi	Muruts and Kelabits	Swedes
Afrikaners	Belarusians	French Canadians	Klemantans	Nanaians	Syria Arabs
Akan	Bemba	Frisians	Kohistanis	Nandi	Tajiks
Alacaluf	Bengalis	Fulbe	Koreans	Nepalese	Talyshes
Albanians	Bhotias	Gagauz	Masai	Ngiri	Tama
Aleuts	Biharis	Georgians	Mashona	Norwegians	Tamangs (Murmis)
Andoa	Bisaya	Gere	Mayorunas	Nubians	Tamils
Angoni	Bo and So	Germans	Mbum	Oases Berbers	Tatars
Arabs of Yemen	Bobangi and Bangala	Gola	Mende	Ojibway	Tats
Arapahos	Boloven	Greeks	Lacandons	Paiconecas	Tay
Araucanians	Bora	Grusi	Lahu	Pamir Tajiks	Tem
Argentinians	Bororo	Guana	Lao	Panare	Terraba
Armenians	Bosnians	Guarañoco	Latvians	Panjabis	Teso
Aromani	Brahui	Guayupe and Marawa	Lenca	Papago	Thai (incl. Thai Nea, Thai Pho)
Ashluslay, Choroti, Mataco, Maca	Estonians	Gurma	Libya Arabs	Sudan Arabs	Toba
Assyrians	Evenks (incl. Hamnigans)	Hani	Limbus	Sui	Togo tribes
Athapaskans	Ewe	Hausa	Lithuanians	Sulu-Samal	Tojolabali
Austrians	Fang	Hindi-speaking peoples of Northern India	Loba	Susu	Tonga
Aymara	Finns	Hondurans	Lobi	Swahili	Tsimshian
Azande	Buryats	Hottentots	Lu	Punans, Beketans	Tsonga
Azerbaijanians	Busa	Hui	Lur	Quechua	Tuaregs
Bagirmi	Bushmen	Hungarians or Magyars	Ma	Rajasthanis	Turkmens
Bakele	Byelorussians	Ibans	Maba (incl. Masalit)	Rhaetoromanians	Turks
Bakomo	Caingua	Iloko	Macedonians	Riau, Palembang	Tuvinians
Bakongo	Catalans	Kui	Macu	Romas	Vankieu
Bakonjo	Cham	Kulango	Maka	Rumanians	Venezuelans
Bakota	Chamba	Kumaonis, Garhwalis	Makonde	Russians	Vietnamese
Baloch	Chane	Kurds	Makua	Saami	Wa
Balozi	Chileans	Kyrgyz	Malavi	Salish	Wachokwe
Baluhya	Chinese (Han)	Jamaicans	Malays of Malaya	Sea Nomads or Sea Gipsies (Orang-Laut)	Wadjagga
Balunda	Chuang	Japanese	Mame	Sedang	Wafipa
Banda	Chuj	Jarai	Mandingo	Semangs	Wai
Baniwe	Circassians	Jews	Maniteneris	Senufo	Wakinga
Bantu-speaking Pygmy tribes	Cocamas	Jhats, Awans	Mano	Serbs	Wanyika
Brazilians	Coreguaje, Sioni, etc.	Joluo	Pase	Sere-Mundu	Wayao
Bribri	Croats	Jordan and Palestine Arabs	Pedi	Siamese	Wayeye

Buginese	Cueretu	Kajars	Persians	Slavs	West Sahara Arabs
Bulgarians	Czechs	Kanuri	Pima	Slovaks	Wolof
Burmese	Danes	Karamojo	Poles	Slovenes	Yamana
Banyaruanda	Diola	Karapapakh	Mexicans	Somba (incl. Berba, Bilapila)	Yao
Banyoro	Ecuadorians	Kayans	Miao	Songai	Yaruro
Barba	English Canadians	Kazakhs	Mnong and Brao	Soninke	Yi
Bari	Englishmen	Kehlao	Moldovans	Southern Lwo	Yoruba
Barundi	Eskimos	Kenyahs	More (Itene)	Stieng	Yukuna
Basques	Indians of India and Pakistan	Khasis	Morocco Arabs	US Americans	Yure
Basubia	Irishmen	Khmers	Moru-Mangbetu	Uighurs	Zulus
Basuto	Iroquois	Khmu	Mossi	Ukrainians	
Bateke	Italians	Kho	Motilones	Ukranians	
Bawenda	Jalaltecs	Kiratis	Mpongwe	Uzbeks	

Table I – Appendix-Table 1.3: List of Ethnic Groups used to Design one of the Cultural Spatial

Source: Weidmann, Ketil & Cederman, 2010, 2010 – http://www.icr.ethz.ch/data/other/greg

Appendix 1.4 – List of Language Groups adopted in the Regression Analyses

List of Languages in Cultural Spatial Regressions					
!Xóõ	Chakma	Herero	Limburgish	Okolod	Talian
A'tong	Cham, Western	Hindi	Limón Creole English	Olusamia	Talinga-Bwisi
Abaza	Chamling	Hmong Daw	Lingala	Omagua	Tamajaq
Abenaki, Western	Chechen	Hmong Njua	Lisu	Ombamba	Tamajaq, Tawallammat
Abron	Chichewa	Huitoto, Murui	Livvi	Oneida	Tamang, Eastern
Achuar-Shiwiar	Chicomuceltec	Hung	Livvi-Karelian	Onondaga	Tamil
Adangbe	Chidigo	Hungarian	Logooli	Orang Seletar	Tanana, Upper
Adele	Chin, Bawm	Hupdë	Lomavren	Ottawa	Tapieté
Adi	Chin, Khumi	Hutterisch	Lombard	Pa Di	Tariana
Adyghe	Chinese, Hakka	Iban	Lozi	Pa-Hng	Tat, Muslim
Afrikaans	Chinese, Mandarin	Idu-Mishmi	Lugbara	Pacoh	Tatar
Afro-Seminole Creole	Chinese, Min Bei	Ifè	Lukpa	Palaung, Pale	Tausug
Aguna	Chinese, Min Dong	Ikposo	Lulogooli	Palaung, Ruching	Tectitec
Aheu	Chinese, Min Nan	Ili Turki	Lumbu	Pana	Teke, Eboo
Aja	Chinese, Pu-Xian	Indian Sign Language	Lun Bawang	Panamanian Creole English	Teke, Ibali
Aka	Chinese, Yue	Indo-Portuguese	Lunda	Pangasinan	Teke-Eboo
Akateko	Chinook Wawa	Ingarikó	Luo	Pangkhua	Teke-Ibali
Akawaio	Chiquitano	Ingush	Luoba, Boga'er	Panjabi, Eastern	Teke-Tege

Akeu	Chiripá	Inuktitut, North Alaskan	Luoba, Yidu	Panjabi, Western	Tektiteko
Akha	Chokwe	Inupiatun, North Alaskan	Luxembourgeois	Pankhu	Telugu
Akposo	Chorote, Iyo'wujwa	Iranun	Lyngngam	Pashto, Southern	Tem
Albanian, Gheg	Chuj	Irish	Lü	Patamona	Teribe
Albanian, Tosk	Chut	Italian	Maasai	Pemon	Teso
Alemannic	Cocama-Cocamilla	Itene	Macedonian	Persian, Iranian	Thai
Aleut	Cocopa	Iu Mien	Machinere	Phu Thai	Thai, Northern
Alsatian	Cofán	Jadgali	Macuna	Phuan	Thangmi
Alur	Corsican	Jakalteko	Macushi	Phukha	Tharu, Chitwania
Amahuaca	Cree, Plains	Jalunga	Madura	Phula	Tharu, Dangaura
Amba	Crimean Tatar	Jamaican Creole English	Magar, Eastern	Phunoi	Tharu, Kathariya
American Sign Language	Croatian	Japanese	Maithili	Phuthi	Tharu, Kochila
Angika	Cubeo	Jarai	Majhi	Piapoco	Tharu, Rana
Angloromani	Cuiba	Javanese	Makhuwa-Meetto	Piaroa	Thulung
Anii	Culina	Jeh	Makonde	Picard	Tibetan
Anufo	Curripaco	Jehai	Makwe	Pidgin Bantu	Tibetan, Central
Anyin	Czech	Jola-Fonyi	Mal	Piratapuyo	Ticuna
Arabic, Eastern Egyptian Bedawi Spoken	Dakota	Judeo-Italian	Malay	Pitcairn-Norfolk	Tidong
Arabic, Gulf Spoken	Dan	Judeo-Tat	Malay, Baba	Plautdietsch	Tinani
Arabic, Hasanya	Danish	Jula	Malay, Cocos Islands	Playero	Tlingit
Arabic, Hassaniyya	Darang Deng	Ju'hoan	Malay, Kedah	Pnar	To
Arabic, Judeo-Iraqi	Daur	Kabardian	Malay, Satun	Pol	Toba
Arabic, Judeo-Moroccan	Deg	Kabiyè	Malay, Standard	Polish	Toba Qom
Arabic, Judeo-Tunisian	Delo	Kachchi	Malayalam	Pomo	Tonga
Arabic, Levantine Bedawi Spoken	Desano	Kaduo	Malecite-Passamaquoddy	Pontic	Totela
Arabic, Libyan Spoken	Dhatki	Kako	Maleng	Portuguese	Tsaangi
Arabic, Mesopotamian Spoken	Dhimal	Kakwa	Mam	Potawatomi	Tshangla
Arabic, Najdi Spoken	Dholuo	Kalanga	Mambwe-Lungu	Prai	Tsimshian
Arabic, North Mesopotamian Spoken	Digaro-Mishmi	Kalmyk-Oirat	Mandaic	Puinave	Tsoa
Arabic, Omani Spoken	Digo	Kamta	Mandinka	Pulaar	Tsonga
Arabic, Shuwa	Diriku	Kang	Mandjak	Pular	Tswa

Arabic, South Levantine Spoken	Ditammari	Kanjobal, Western	Mang	Punu	Tswana
Arabic, Standard	Domari	Kanuri, Central	Maninkakan, Eastern	Puoc	Tucano
Arabic, Western Egyptian Bedawi Spoken	Dotyali	Kaonde	Maninkakan, Western	Purepecha	Tumbuka
Arem	Duano	Karaim	Manjui	Pökoot	Tunebo, Central
Armenian	Dungan	Karapanã	Manyika	Q'eqchi'	Turkish
Aromanian	Duoluo	Kare	Mapudungun	Qabiao	Turkmen
Arpitan	Dutch	Karelian	Mapun	Quechua, Napo Lowland	Turoyo
Arutani	Dzhidi	Karon	Maquiritari	Quechua, South Bolivian	Tuscarora
Ashéninka, Ucayali-Yurúa	Ekegusii	Kashinawa	Maraba	Quichua, Napo Lowland	Tuva
Assiniboine	Emberá, Northern	Kashmiri	Marma	Quichua, Northern Pastaza	Tuyuca
Assyrian Neo-Aramaic	Emberá-Catío	Kazakh	Marwari	Rangpuri	Ukrainian
Asturian	English	Kekchí	Mashi	Riang	Uma' Lasan
Aushi	Epena	Kelabit	Matsés	Romani, Balkan	Urdu
Avadhi	Erromintxela	Kendayan	Mbama	Romani, Baltic	Uyghur
Awa-Cuaiquer	Erzya	Kensiu	Mbandja	Romani, Carpathian	Uzbek, Northern
Awadhi	Ese Ejja	Kenyah, Mainstream	Mbangwe	Romani, Kalo Finnish	Vai
Aymara, Central	Estonian, Standard	Kenyah, Wahau	Mbere	Romani, Sinte	Venda
Ayoreo	Evenki	Khakas	Mbukushu	Romani, Vlax	Venetian
Azerbaijani, North	Extremaduran	Khaling	Mbum	Romanian	Vietnamese
Azerbaijani, South	Fang	Kharia	Meche	Ronga	Vili
Bagheli	Finnish	Khasi	Meitei	Russian	Vlaams
Bagri	Finnish, Tornedalen	Khmer, Central	Mende	Russian Sign Language	Wa, Parauk
Baka	Flemish	Khmu	Michif	Rusyn	Waimaha
Bakati', Rara	Fon	Khowar	Micmac	Rwanda	Wajiara
Balkan Gagauz Turkish	French	Khua	Miju-Mishmi	Sa'ban	Wakhi
Balochi, Eastern	Fulfulde, Adamawa	Khuen	Miranha	Saami, Lule	Walser
Balochi, Southern	Fulfulde, Borgu	Khwe	Miyobe	Saami, North	War-Jaintia
Balochi, Western	Fulfulde, Gorgal	Khün	Mizo	Saami, Pite	Wayuu
Balti	Fulfulde, Maasina	Kickapoo	Mlabri	Saami, Skolt	Welsh
Bamanankan	Fulfulde, Western Niger	Kim Mun	Mnong, Central	Saami, South	Wichí Lhamtés Nocten

Bambara	Fumbira	Kintaq	Mohawk	Saamia	Wolio
Banda, Mid-Southern	Furu	Kisan	Molbog	Sadri	Wolof
Banda, South Central	Gagauz	Kisi, Southern	Moldovan	Saek	Woun Meu
Banda, Togbo-Vara	Galibi	Kiswahili	Monba, Cuona	Salako	Wumbvu
Bandial	Galician	Klao	Mongolian, Halh	Salang	Wyandot
Bangandu	Garifuna	Koch	Mongolian, Peripheral	Salish, Straits	Xaasongaxango
Bangi	Garo	Koda	Monpa, Tawang	Sama, Balangingih	Xasonga
Baniwa	Gascon, Aranese	Kok Borok	Monzombo	Sama, Southern	Xhosa
Banjar	Gbanziri	Kokola	Mpiemo	Samtao	Yaka
Bantawa	Gbaya	Koli, Kachi	Mpyemo	Sangir	Yakan
Barranquian	Gbaya, Northwest	Koli, Wadiyara	Mru	Sango	Yakkha
Baré	Gbe, Maxi	Konkani, Goan	Mugom	Sansi	Yakoma
Barí	Gbe, Waci	Konkomba	Mundari	Sanskrit	Yalunka
Bashkort	Gbe, Western Xwla	Konzo	Mòoré	Santhali	Yaminahua
Basque	Gciriku	Koongo	N'ko	Sanumá	Yanomamö
Bassa	Gelao, Green	Korean	Nago, Northern	Sauria Paharia	Yao
Bateri	Gelao, Red	Kota	Nama	Scots	Yaqui
Bavarian	Gelao, White	Krahn, Eastern	Nambya	Secoya	Yasa
Bedawiyet	Geman Deng	Krahn, Western	Nanai	Selungai Murut	Yeyi
Bekwel	Gen	Krumen, Tepo	Naro	Sembakung Murut	Yiddish, Eastern
Belarusan	Georgian	Kucong	Nawdm	Serahule	Yombe
Bemba	German, Hutterite	Kuhane	Ndali	Seraiki	Yoy
Bembe	German, Pennsylvania	Kulango, Bondoukou	Ndasa	Serbian	Yukpa
Bengali	German, Standard	Kulango, Bouna	Ndau	Serer-Sine	Yulu
Bhojpuri	German, Swiss	Kulina	Ndebele	Shan	Yupik, Central Siberian
Bidayuh, Biatah	Giáy	Kulung	Ndo	Sharanahua	Yámana
Bidayuh, Bukar-Sadong	Glio-Oubi	Kumiai	Ndrulo	Sherpa	Zakhring
Bihari	Gobu	Kumyk	Nengatu	Shina	Zande
Birifor, Southern	Gola	Kuna, Border	Nepali	Shona	Zarma
Birwa	Gourmanchéma	Kunda	Newar	Sila	Zarmaci
Bisa	Grebo, Southern	Kurdish, Central	Ngangam	Silesian, Lower	Zezuru
Bishnupriya	Greek	Kurdish, Northern	Ngbaka	Sindhi	Zhuang, Dai
Bissa	Groma	Kurdish, Southern	Ngbaka Ma'bo	Sinhala	Zhuang, Yang
Bisu	Guahibo	Kuria	Ngom	Sinicized Miao	Zhuang, Yongnan
Bit	Guanano	Kurux	Ngoni	Siona	Zhuang, Zuojiang
Blackfoot	Guaraní, Ava	Kutenai	Ngäbere	Siriano	Zulu

Blang	Guaraní, Eastern Bolivian	Kuy	Nhengatu	Slavic	Éwé
Bobangi	Guaraní, Mbyá	Kyerung	Ninam	Slovak	Ñandeva
Bodo	Guaraní, Western Argentine	Kyirong	Nivaclé	Slovene	‡Kx'au'ein
Bomwali	Guarayo	Kyrgyz	Njebi	Soninke	
Bora	Guarequena	Lachi	Njyem	Sotho, Southern	
Borôro	Gubu	Ladakhi	Nobiin	Spanish	
Bouyei	Gujarati	Ladino	Nonuya	Stieng, Bulo	
Brahui	Gujari	Lahu	Nsenga	Sui	
Brao	Gula	Lahu Shi	Ntcham	Swahili	
Bru, Eastern	Gurani	Lakota	Nubi	Swahili, Congo	
Bugis	Gurung, Western	Lala-Bisa	Nung	Swati	
Bukitan	Gusii	Lama	Nyakyusa-Ngonde	Swedish	
Bulgarian	Gwich'in	Lamba	Nyamwanga	Sylheti	
Buraka	Hagei	Lambya	Nyanja	Sáliba	
Buxinhua	Haida, Northern	Lamet	Nyiha	Sénoufo, Shempire	
Bwile	Hajong	Langbashe	Nyiha, Tanzania	Sô	
Bwisi	Halang	Lao	Nyika	Ta'oih, Upper	
Byangsi	Halang Doan	Lapcha	Nzakara	Taabwa	
Caló	Halkomelem	Lave	Nzema	Tagal Murut	
Carapana	Han	Lendu	O'du	Tai Daeng	
Carib	Hani	Lepcha	Ocaina	Tai Dam	
Cashinahua	Hassaniyya	Lhomi	Occitan	Tai Dón	
Catalán	Hausa	Ligbi	Oirat	Tai Nüa	
Cayuga	Hawrami	Ligurian	Okanagan	Tai Ya	
Ch'orti'	Hazaragi	Limbu	Okiek	Tajiki	

Table J – Appendix-Table 1.4: List of Language Groups used to Design one of the Cultural Spatial Weights

Source: Lewis et al, 2014 – http://www.ethnologue.com/17/

Chapter Two:

Institutions and Economic Development in Africa: A History-Guided Nexus

Chapter Abstract

In this chapter we attempt to evaluate the impact of national institutional quality on economic development in Africa at the subnational level. We assess this age-old research question by adopting a dual-pronged research design motivated by two very influential studies from the literature by Michalopoulos & Papaioannou (M&P, 2014) and Acemoglu, Johnson and Robinson (AJR, 2001). Firstly, by considering the variations in national institutional development explained by key historical exposures to particular aspects of European influences on the African continent, crucial links between past European contact and institutional development are unearthed and or reinforced. Secondly, we proceed to attempt to reveal how these history-predicted measures of institutional development help us understand current cross-sectional differences in economic development on the African continent at the ethnic homeland echelons.

We adopt an alternative measure of subnational economic development to the recently popularised nighttime lights density measure using the spatial distribution of the access roads networks available in each subnational region. Our findings suggest that the nature of European contact experienced by most African countries and their corresponding ethnic homelands provide some insight into the varying patterns of national institutional quality. It is further revealed that these history-treated measures of institutional development obtained herein, in some cases provide statistically significant treatments for economic development in these African territories. Particularly, we find that when the access roads proxy is adopted the institutional quality at the national level is statistically relevant for economic development at the subnational level while when the nighttime lights measure is adopted national institutions appear statistically irrelevant. These findings are not only in agreement with M&P's, 2014 key finding using the nighttime lights density proxy, but also corroborate most of the country level studies that attempt to highlight the importance of national institutional quality for economic development such as AJR, 2001 using the alternative access roads proxy.

Section 2.1 – Introduction

Diagnostically alluding to history and the intuitive use of geographical attributes, have become rather commonplace in attempts at understanding the fundamental sources of the protracted income divides across countries. Although their relevance appears unanimously agreed on (Nunn, 2009), the manner and empirical approach in which both these fundamentals are employed in the growth literature remains the point of departure and a lucrative avenue for insightful research contributions.

The African context provides an interesting regional (continental) case study in the research for understanding persisting underdevelopment for a number of reasons. Firstly, a significant percentage of the continent, particularly in the Sub-Sahara African (SSA) region, is plagued with lingering unresolved political and socioeconomic challenges. The SSA region has one of the gloomiest figures for institutional quality and economic wellbeing in the global economic landscape[37]. Secondly, it is amongst the most ethnically diverse and natural resource rich continents in the world thus continuing to attract varied interests far and wide. Thirdly, the current geopolitical identities and immediate history of the countries within the continent are in fact still within the last two centuries compared to countries such as India and China whose composing geopolitical identities date back more than a millennia. These non-exclusive reasons position the African region as an ideal and relevant research focus for studying the impact of proximate historical, geographical and cultural (ethnic) experiences on current challenges, which affected and continue to affect the institutional and socioeconomic development of most of the countries now spanning the region.

2.1.1 – Objectives

The political demarcation of the continent at the wake of European colonisation provided a break from the past ethnically based geopolitical institutions that characterised the then indigenous societies of Africa (see Asiwaju, 1985; Herbst, 1996). This break also gave birth to the form of modernization and ensuing globalization now shaping the contemporary economic and geopolitical entities

[37] See Bates & Collier, 1995; Easterly, 2001; Collier, 2007; etc. for related discussions on the subject.

presently referred to as nation states or countries within the region (Táíwò, 2010). In light of this institutional and structural evolution, this research seeks to achieve the following:

- Chiefly, our research seeks to link relevant and discernible pre-colonial configurations characterising the bygone existing geopolitical boundaries with current observable differences in levels of national institutional quality.
- The study then goes on to assess how these predicted variations in cross-national institutional quality, warranted by the subsequent modern geopolitical nation states, affect current observable measures of economic wellbeing.
- This duo-pronged design is estimated using a two-step LIML (limited information maximum likelihood) instrumental variable estimator; with the sole purpose of reconsidering the impact of the levels of national institutional quality on subnational economic activity in demarcated pre-colonial ethnic states at the periphery of modern-defined national boundaries.

2.1.2 – Scope and Contribution

The core of the research design adopted in this thesis chapter is not novel but was previously harvested in a recent study by Michalopoulos & Papaioannou (M&P, 2014). In their paper, M&P, 2014 adopt a regression discontinuity approach intended to exploit the exogenous demarcation of the pre-colonial ethnic homelands into different modern national geopolitical institutional boundaries recognised at the dawn of colonization in the 1800s. Their obvious but novel argument (M&P, 2014) being that these unsolicited splits afforded a natural experimental framework for investigating the exogenous connection between the current economic wellbeing of these pre-colonial geopolitical homelands and the modern national institutional quality found in the corresponding halves of their respective present-day countries.

Our research in this chapter uses an ethnic homeland polygonal dataset based on the work of Murdock (1959) and designed in a similar fashion as M&P (2013; 2014): with institutional quality measured at the national level and economic activity at the subnational tribal homelands level. The homeland polygons of demarcated groups are identified by exploiting dynamics available to the geographic intersection of the tribal homeland territories with modern national

boundaries, which then allows us further, examine whether variations in the national institutional quality of these overarching nation-states had any effect on the development of the earlier existing demarcated local groups. Following studies such as Nunn (2008), who adopt a historical narrative investigating the impact of pre-colonial activities (in his case, slave exports) on the modern economic performance of African economies, we explore a narrative that attempts to stitch the overall narrative underpinning these discourses together. These underpinning narratives recognise a common position – that current economic trends have significant connections or associations with history-centric geographic and institutional arrangements (see Acemoglu, Johnson, & Robinson 2001/2012, 2002; Acemoglu & Robinson, 2006; Banerjee & Iyer, 2005, among a plethora of others). Such arrangements remain relevant as they characterised, and most likely continue to do so, the societies that have lived in their historical lands since the advent of significant European (or other foreign) geopolitical or socioeconomic influences.

Our research however, contributes to the literature by introducing an IV approach to the M&P (2013, 2014) core research design. In doing so, we are able to suture together history-informed variations in the experiences of pre-colonial ethnic homelands and their relationships in different forms of European contact with observed modern economic and institutional improvements or their lack of. Thus providing a functional storyline for how such relationship history between the Europeans and the African tribal homelands would have filtered through to contemporary differences in the quality of national institutional levels as observed in the region today. Three main forms of European relationships are observably proposed and harnessed to predict current variations in national institutional quality, within this study. These interactional forms include, but not exclusively:

- Missionary activities: these were targeted at institutionally affecting the cultural fabric of the indigenous ethnic nations' peoples especially on issues of religion and biblical education
- Economic activities: this kind of activity mostly affected the formation of economic based institutions on the pretexts of commerce. In the African case, it was mostly rapacious and extractive in nature; structures were setup to accrue resources from the hinterlands to the coast

- Colonial administrative interactions: which are known to have defined the nature of national boundaries as currently known and set the precedence for national institutional transformations especially in the political sphere

Each of these channels (or instruments) is addressed in more detail in the methodology section, below, with some insightful justifications provided – highlighting why they are relevant in a study such as this. Probing these observed forms of interaction with the pre-independence Europeans provides a good framework for teasing-out relevant aspects of current cross-sectional variations in institutional levels. Particularly as these modern institutional arrangements, were effectively a foreign introduction into the structures and systems of the African indigenous geopolitical lifestyle; they were fundamentally, creations of these 'prodding' Western civilizations embodied by their European representatives. Additionally, significant elements of these pre-colonial ethnic homelands and cultures persist to this date; further providing a valid administrative level for conducting the research herein with respect to our dependent variables of interest. Being coerced into colonially defined nation states, understanding how the resulting national institutions affected current developmental levels within those demarcated pre-colonial groups provides an interesting research inquiry.

2.1.3 – Overview of Chapter's Remaining Sections

The next section (section 2.2) is the Literature review, which provides a brief discussion of the immediate literature relevant to the methodology, motivation and approach adopted herein. The section following section 2.2, section 2.3, provides an overview of the dataset, how it was obtained and setup. Maps and initial summary measures of the essential variables are also provided and briefly discussed to demonstrate some semblance of what is to be expected. Finally in section 2.4, the estimation results are reported and discussed. This is immediately followed by a conclusion and an enumeration of key identified weaknesses and areas for possible extension or improvements to the study as a whole.

Section 2.2 – Background and Literature Review

2.2.1 – Linking Modern Institutions with Economic Development in Africa

There are a number of identifiable conduits via which institutions may have historically affected current economic outcomes and continue to do so for developing countries. For the African economies, one of these influential pathways is through the experience of colonization and other forms of substantial initial European contact. As already hinted, in the preceding section 2.1.2, European colonization of the African continent happens to be one of the most significant modern channels of institutional transformation. Colonisation left in its wake a near shift or replacement in the political administrative structure of pre-existing governance territories closely following the European exploration and slave trading in Africa. There is a huge debate in the growth literature demonstrating that colonization affected economic development in Africa quite variedly. There are positions suggesting that not only did the colonial identity matter for signposting current variations in a country's institutional quality but that the length of the colonial administration also exerted distinguishing influences; further stimulating dissimilar developmental patterns across countries in Africa and the wider global economic landscape (see Glaeser & Schleifer, 2002; La Porta et al, LLS, 2008).

Another closely related causal conduit between institutional and economic development, no doubt, involves geography. This is among one of the strongest mechanisms often adopted in modern researches of institutional influences on economic development. It may certainly well be the reason why it has repeatedly been suggested as more important in explaining variations in economic wellbeing globally than any other conduit ever used in the entire chronicle of the Institutions-Development narrative (see Sachs, 2003). In this pro-geography narrative, there are two distinctly broad routes of inquest. One focuses on the geography of location endowments, while the other emphasizes the geography of natural resource endowments. It is within the theoretical premises of locational endowments (positional advantage) that works such as AJR (2001); Easterly & Levine (2003) and Rodrik et al (RST, 2004), among others attempt to link geography with current developmental levels through their impacts on

institutional quality. The other line of closely related inquisition emphasizes the relevance of geographic channels through natural-resources' contaminating influence interacted with or through institutional quality (see Mehlum et al, MMT, 2006; Acemoglu et al, ARV. 2004; Sarmidi et al, SLJ, 2014) further deepening the Dutch-disease or Resource cause (Sachs & Warner, 1995, 1997) storyline. As a central pivot relating everything together, historical validations have become very relevant in these development discourses. As a result history tends to form the bedrock or analytical adhesive of most identification strategies in the inquest to uncover and unbundle the impacts of fundamental sources of economic development such as institutions.

2.2.2 – Review of the Research Design and Objective

As earlier alluded, this thesis chapter seems to string together a number of already highlighted bits of related literature from the economic growth discourse in an effort to further illuminate on the mechanisms through which national institutions affect subnational economic development. Although so much has already been done in the empirical literature towards presenting some consistency in the analytical endeavours associating pre-colonial societal structures and current economic development in the African economies, a lot more can still be done to identify relevant and informing channels of institutional persistence from the multiple administrative levels found in the region. Nunn & Puga (2012) for instance, argue that difficult geographic terrains provided an exogenous obstacle to over exploitation of human resource through the slave trade era in Africa; AJR (2001) argue that the link between institutional development and modern economic growth can be clarified by observing European residence patterns reliably associated with observable cross-sectional variations in the disease environment or death prevalence in the respective colonial territories across the world; Gennaioli and Rainer (2006) posit that the anthropological characteristics of pre-colonial ethnic groups mirroring their political institutional complexity may still persist today - providing some noteworthy insight into understanding patterns consistent with current institutional quality (especially those pertaining to governance); while Austin (2010) opines, quite similarly but slightly contradictory to the position of AJR (2001), that the settlement patterns of the colonisers were as informative in explaining the variations in economic development as was the

differing colonial administrative policy regimes spanning the African geopolitical territory.

Our paper seeks to re-visit such an already vastly researched economic development inquiry by applying a duo-level identification strategy not previously adopted in a similar fashion, prior to this in the literature. We explore variations in commonly used measures of institutional quality from a different perspective. By initially instrumenting these institutional measures using incisive geographical and historical characteristics of ethnic lands in Africa, a deeper and comprehensive pre-treatment of national institutional quality is attempted. This eventually enables an identified estimation of the impact of national institutional quality on economic activity at the subnational level. Exploiting the intrinsic national institutional variations created by the exogenous colonial demarcation of ethnic groups on the continent (following M&P, 2014), permits the approach herein to nexus cross-sectional variations in the experiences of these demarcated ethnic lands, consequent to the continent's colonial episode, with corresponding current country-level institutional outcomes. There is a forthright story that connects the developmental experiences of the pre-colonial development era with that of the colonial period and which in turn aligns with the modern developmental experiences presently observed.

By adopting an instrumental variable (IV) approach, this research seeks to reconsider this story and fuse together storylines previously analysed individually. The key instruments we adopt depend on variations in the extent of each modern country's and demarcated pre-colonial ethnic group's exposure to particular dealings and aspects of European contact. The makeup of these dealings defined the Europeans' initial and subsequent encounters with the pre-modern ethnic administrative societies existing at the time and those that would later persist. The historical and analytical narrative is that certain types of exposures ought to have positively primed the ensuing post-colonial administrative countries of the African continent towards better or improved modern institutional outcomes whiles others achieved the polar opposite. Very much in line with the narrative proposed by AJR (2001) but critiqued by Glaeser et al (GLSS, 2004), who hypothesise that modern institutional forms (or quality) reflect the kind of not-so-distant historical interactions the Europeans had with their areas of geopolitical influences. AJR (2001), conjecture that modern institutional quality in former

colonies reflects the nature of colonial economic activity that was exerted on the respective colonial territories. That in cases where the organisations of contact involved lots of resource extraction (mostly in Latin America and African colonies), then 'extractive institutions' were setup; while areas where residence arrangements were intended or conducive (such as in Australia and New Zealand), the resultant institutional setup reflected those typical to practices found back in their countries of origin (AJR, 2001: 1374-1376). Glaeser et al (2004) argue however, that better institutions can often be viewed as consequential to economic vibrancy or improved human capital and not necessarily the treatment as quite ordinarily envisioned. Stringing all these positions together from the Institutions-Development nexus makes it apparent why attempts to understand, confirm or affirm the treatment links between institutional and economic developments are pertinent.

2.2.3 – Review of the Background to the Theoretical Framework

European encounter with the African continent over the 19th and 20th centuries affected and influenced the region in a plethora of significant ways. As already mentioned earlier, these influences channelled through structural and institutional dimensions of African life, touching on culture (through education and religion), commerce (through commerce in goods and slaves), and politics (through colonisation). Socio-political and economic changes were spurred and championed to reflect particular dimensions and interests of each channel of influence. A summary of this framework, inspiring our current study, is shown in the layout below:

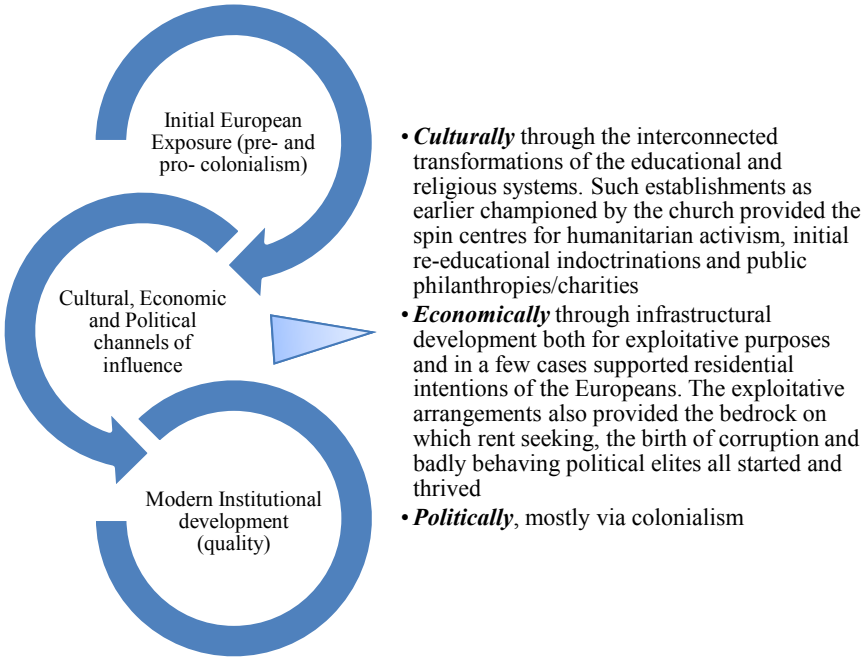

Figure E – 2.2.1: European Contact and Modern Institutional Development

Starting with the channel of cultural influences, Nunn (2010), for instance, points out that the location of missionaries on the African continent not only signposted a number favourable habitable circumstances for the European missionaries such as proximity to clean water sources, areas favourable to imported supplies from their home-countries and fertile soil for sustainable agriculture, etc., but that their decision to locate their missions also reflected areas of necessity - where for instance, slave trading was prevalent and needed action to stir its abolition (pp. 147-148). Nunn (2010) also finds that areas that received greater missionary contact are today more likely to proclaim their religious orientation as Christian. If this history or identification is apt, then we should expect to find distance to where missionary centres were sited to provide some relevant information consistent with current patterns of national institutional quality at the split ethnic homelands. Particularly because, the activities the missionaries advocated were reported to have initialised mechanisms promoting pro-developmental circumstances such as western type re-educational endeavours, humanitarian values of equality, religious values of recognising written laws through pro-attitudes towards religious commandments and the gains from hard 'honest' work, etc. These mechanisms, all things being equal, should eventually have set the precedence for western-type 'good' institutional outcomes as observed and expected in modern developed economies

or societies echoing this developmental era. If however the reverse is the case and we discover that areas farther away from such historical missions are institutionally better off today, this could indicate that the locals' attitudes towards these Western visitors were perceived in similar guise irrespective of their (Europeans) operational motives and approaches.

On the second channel of European-championed economic activities, a number of fairly documented pathways are apparent. But of relevance in this paper are those of slave trading, primary produce and natural resource extraction. Following the theories underpinning studies like AJR (2001) and Nunn (2008), we exploit the position that European influence through economic activities such as slave trading provided a negative shock to current economic activity through institutional development[38]. The position forwarded by this line of inquiry is that the extractive nature of such economic activities garnered by the Europeans in Africa, from the time of the slave trade era to the post-colonial period of natural resource and cash crops trading, created and sustained the path dependence attuned to exploitative relations which would later characterise most of the countries within the region. This path dependence was sustained by a responsive creation and promotion of diminutive institutional foundations which is more likely the source of manifestations of corruption, cross-cultural mistrust and ethnic wars, weak governance, among others, observed in modern African societies (see Whatley & Gillezeau, 2009; Nunn & Wantchekon, 2011 for related research on the subject). Because, the African region had very limited infrastructural development at the advent of European arrival, it is easy to conceive that most European economic activity would have heavily relied on subsequent infrastructural creations such as the explorer (trade) routes. This provides significant premise for attempting to test if the national prevalence of this infrastructure negatively or positively affected subsequent institutional quality. Since this historical routes-network was in relatively scant supply, it is not inappropriate to measure it at the national level. Especially given that its prevalence would have had permeating influences nesting a number of ethnic homelands within its proximity. We attempt to measure exposure to this dimension of European activity by controlling our first-stage model on modern average institutional quality for the number of routes

[38] For some earlier historical discussions on the subject, see Rodney (1973) and Manning (1983), for instance.

within the historic network that fall within the associated modern country territory. If this initial exploration infrastructure provides consistent explanations of the statistical variations in the national institutional quality of the country-split ethnic groups being analysed, then there is some evidence to suggest this initial exposure to ancillary exploitative economic activities correlates with weak modern institutional outcomes.

As highlighted in an earlier paragraph, colonisation in Africa was a relatively more modern exogenous shock to the geopolitical landscape of the administrative units on the continent. As equally portrayed in our chart above, this political channel existed in harmony with either the cultural and economic channels or forms of interaction that post-dated the slave trade era. So it is a bit over-the-top or ambitious to attempt to simultaneously, within our current framework, measure the impact of colonisation on institutional quality. It is even more tenuous to do this when you consider the fact that there were at least four major European colonial administrative identities in the scramble for territory in the African region. Also, all these European countries had their own distinct institutional calibrations. In this light, our model attempts to control for the identities of these colonial administrations as unobservable fixed-effects. In order to account for the unique contributions each colonial identity would have had on their respective colonies in their current post-colonial institutional establishments[39]. Similarly in this light, we also control for the legal origin following Glaeser & Shleifer (2002). We believe that both set of controls are sufficient to control for any colonial dynamics that may affect the other two forms of initial European contact within the African region.

[39] For an easy-read review on the subject matter of colonization and its impacts, see Ziltener & Künzler (2013).

Section 2.3 – Model Framework and Data Sources

2.3.1 – The Model

$$\widehat{\mathrm{Inst}}_{ic} = \hat{\gamma}_1 X_{ic} + \hat{\gamma}_2 Z_{ic} + \mu_e \quad (1)$$

$$y_{ic} = \theta \widehat{\mathrm{Inst}}_{ic} + \beta X_{ic} + \alpha_e + \varepsilon_{ic} \quad (2)$$

where,

- Z_{ic}, is a vector holding the set of excluded instruments to help identify the endogenous variable, average institutional quality, $\widehat{\mathrm{Inst}}_{ic}$. The list include – distance to the nearest historical missionary station, an indicator to capture whether a split ethnic homeland's associated country had a historical European explorer/trade route(s) passing within its current national borders or not, an index measuring a country's state history, and indicator controls to capture the identity of the associated country's major European colonial power and its legal origin;
- X_{ic}, is a vector of control variables characteristic of the literature capturing exogenous geographic features and endowments including population density and the split tribal polygon's homeland area;
- 'μ_e' and 'α_e', are ethnic-specific characteristics and are used to control for ethnic fixed-effects when the fixed-effects IV estimation is adopted;
- ε_{ic}, is the error process hypothesized to follow a white noise IID process. However, because there is a likelihood that the constant variance assumption is violated, the standard errors are thus adjusted for any arbitrary potential heteroscedasticity;
- $\boldsymbol{y_{ic}}$, is the response variable of interest and it attempts to capture the level of current economic activity observed for each split ethnic homeland. There are two main proxies – the average night light density variable and the average roads length variable;
- subscripts: '$\boldsymbol{i}$', identifies the split ethnic polygon or historical homeland, while '$\boldsymbol{c}$', identifies the country, and '$\boldsymbol{e}$', identifies the full tribal state in its un-split form.

We choose to run our estimations at this level of analysis for a number of analytically and methodologically convenient reasons. The most apparent of these being that it allows us control for and hence factor out any inherent pre-colonial tribal characteristics that are latent or unobservable but remain very relevant for explaining

modern growth dynamics at the tribe-to-country levels[40]. This is because this dataset is designed so that only partitioned tribal homelands are considered and as a result partitioned groups are clustered into the same group dynamics when dealing with any single tribal homeland.

In a 'lone' instrumental regression framework, our research design uses the variables in the excluded instruments vector ($\boldsymbol{Z_{ic}}$) in equation (1) to identify exogenous split ethnic and national level variations in institutional quality. By convention, the first-stage equation for the potentially endogenous treatment variable (institutions) also includes the included instruments ($\boldsymbol{X_{ic}}$), which equally act as additional exogenous control-variables in the second stage regression specification on economic development. The instrumental variable framework then allows us provide interpretations of causal treatment effects of institutional quality on our proxies of economic development at the country-split ethnic homeland level. One might argue that simply adopting the single equation framework at the demarcated ethnic level as done in the corresponding base specification model in M&P (2014) might suffice to provide an identified impact of lagged-period average national institutional quality on current tribal homeland economic development (see pp.165). However, we argue that such an approach is insufficient to guarantee an exogenous treatment of national institutions on subnational development. Particularly because the economic development proxy adopted for the subnational split ethnic homelands is the average night light density variable, which is more or less a measure of infrastructural development than it is a reflection of actual per capita income flows. Although, conventional measures of per capita income have been proven to correlate significantly with the night-lights density measure of economic activity (Henderson, Storeygard & Weil, 2012), it remains a relatively slower changing capture of development and not as flow-based as its GDP-type counterpart. Also being relevant to subnational development, it could also include elements, unaccounted for, innate to the cross-sectional level that may equally be correlated with the unexplained components (residual) of the single equation model. These unexplained correlated variations might not be as manifest for an analysis at the aggregated national-level especially

[40] See Gennaioli & Rainer (2007) for an illustration of a very informative empirical work relating specific pre-colonial characteristics with modern economic development dynamics.

given that single measures of both institutions and development are usually adopted to represent an entire country and not representative at a subnational administrative level.

Quite similarly, one could argue that institutional quality by its very nature would dissipate at the national borders relative to the national averages. This is partly because the cost of maintaining national institutional integrity are much more costly and infeasible at the peripheries compared to areas and districts closer to the centre or national capitals. Since our study is focused on exploiting variations made feasible by these demarcated tribal homelands at the fringes of their respective nation state borders, we should also expect an incomplete or fuzzy story from attempting to directly match national institutional quality values to economic development at these peripheral subnational regions. Attitude based institutions are also in most cases slower-moving (Roland, 2004) just like infrastructural development measures; as a result they would not provide sufficient deviations from the past to deserve a valid predetermined identification strategy. Additionally, our level of analysis at the split ethnic homelands indicates by definition that these split groups share common borders. It would be likely that at such borders, national variations in economic wellbeing are less distinct to provide sufficient dissimilarities. Split ethnic homelands still share a common not-so-distant ethnic level geopolitical and socio-cultural history. Hence they are bound to have very similar circumstances or verily absorb cross border advantages via ease of relocation. This is not an elusive idea as border controls in most African regions are quite limited with most of such neighbouring countries belonging to mutual regional economic trade unions such as the EAC, ECCAS, SADC and the ECOWAS[41], etc. Subnational variations in economic activity would resultantly be driven equally by individual exposures warranting the need for an identification strategy that brings into perspective the unique experiences consistent at these split ethnic homelands level.

Our method of institutional filtration or identification, as revealed in the preceding equations above, harnesses key dimensions from the literature in constituting its instruments set. To the best of our knowledge, these 'excluded'

[41] The EAC is an acronym for East African Community; the ECCAS: Economic Community of Central African States; the SADC: Southern African Development Community; and the ECOWAS: Economic Community of West African States.

instruments have not been engaged in a similar fashion in the literature prior to ours. We have done our selection in this manner to allow fundamental multilevel variations to be netted. Such variations as consistent with the unique halves of each split ethnic group and the corresponding national features specifically and ubiquitously characterising them.

2.3.2 – About the Sort of Endogeneity Underpinning Our Research Design

It is imperative to provide some discussion to shed further light on the kind of bias necessitating the identification strategy we have adopted herein. Inherent in the IV methodology are often discussed three sources of endogeneity – simultaneity, omitted variable bias and measurement error (see Wooldridge, 2008:506-536). A look at the structure of the data characterising this study, provides an initial indication that could inform us about some inherent possible sources of such endogeneity. Being a multilevel design framework, it is not farfetched to conjecture that the patterns of bias could be consistent with a combination of simultaneity biases stemming mostly from the country-level and of measurement error accruing mostly from the subnational level. Essentially because, institutions observed at the national level, is being regressed on an income proxy measured at the subnational tribal level. These two biases usually operate in opposing directions causing the conventional OLS-based models to inadequately account for the resulting endogeneity. As a result this yields inconsistent and unsurprisingly misleading estimates.

Taking a more focused view of these two biases in turns, we can conjecture that:

- The first potential bias stems from the position that levels of economic activity and institutional quality simultaneously elicit each other. If on the average ethnic homelands within a country are richer, they certainly can afford better institutions through channels such as increased ability to enforce and monitor collective behaviour in line with decreasing corruption, better rule of law, and improving governance effectiveness. This bias is consistent with the country-level cross sectional studies such as AJR (2001), Rodrik et al (2004), and more recent cohorts.
- The second potential source of endogeneity bias acknowledges that national averages for institutional quality may not exactly or seamlessly map to subnational levels of

economic activity - thus creating an avenue for a possible bias based on measurement error. This is very much related to the notion of the ecological fallacy in applied statistics. Its precise application here is more of a mapping matter: where our right-hand-side (RHS) variable institutional quality observed at the national level might not exactly map unto a previously non-existent national historical experience through the individual experiences of the component ethnic homelands spanning that current country. This further provides a likely indication why our results not using exogenous sources of institutional variations (IV variables) consistent with the historic experiences of subnational ethnic homeland level have generally relatively smaller coefficient estimates and are sometimes statistically insignificant. The argument here is that given we are dealing with national institutional quality outcomes applied to the bordering demarcated tribal homelands, it is more useful to use historical data varying sufficiently at these split levels to predict or map these national institutional values to these split subnational areas. And this link in theory should be able to yield predicted measures of institutional variations distinct from those observable in other fringes of the associated country.

Our discussion in favour of these two sources of endogeneity – simultaneity and measurement error – does not in any way exclude the possibility of the third channel – omitted variable bias. It is possible, as we later acknowledge in the weaknesses subsection of this chapter, that not being able to control for inherent unobservable variations at the lower political administrative levels (such as cultural and other non-institutional structural patterns) could pose further empirical challenges.

2.3.3 – Exclusion Restriction Condition

To effectively attempt to justify our selected instruments, we provide some arguments in this section to dispel major doubts about how our selected instruments measuring aspects of historical European contact could have affected modern economic development indirectly mostly through modern institutional quality within the African subnational regions. Essentially, we attempt to layout a narrative to validate the key untestable assumption of the IV-method that there should not be any systematic relationship (covariance) between the unobservable(s) (error process from equation 2 in the model subsection above)

and our selected instruments (diagrammatic portrayal is provided in Figure F - 2.3.1 below). While there are no analytically watertight instrument(s) in most empirical econometric exercises, we believe that the temporal break in the sort of economic activities carried out in each of the contact phases between the African region and with the key European powers covered in this study provides a good platform to strengthen our narrative. The contact phases referred to include the initial European contact through *the activities of the explorers* – the commercial activities in slaves and essential natural resources that would immediately characterise this period, *the missionary activities* – that mostly initially involved the propagation of the gospel thus promoting mostly cultural changes, and *the colonial periods* – which mostly involved attempts to implant western type institutional and infrastructural changes so as to setup and secure imperialist empires to initially empower or enrich the governing imperialists' home countries (also refer to chart in Figure E – 2.2.1 above). These phases particularly help us demonstrate that the angle and type of economic activity that was the product of these different historical periods were unanimously differentiated from the kind of economic activity characterising modern states within the African region now. To defend this claim, we only need to look at the pervasive theme behind the nationalist drive for independence which manifested across the region from the late 1950s. This unanimous drive reflected an outright rejection of the imperialist extractive status quo for a more inclusive and retentive regime of political governance.

These instruments elaborated below are by no means perfect and with some thought it is quite easy to find weakening empirical and econometric counter-positions. More so when the instruments involved are historical measures that can, in some circumstances, be found to link past variations in levels of economic wellbeing and institutional dynamics with their respective current levels, even when trivially conceived. As an example, Banerjee et al., 2012 highlight that access to transportation networks are endogenous to economic outcomes through their connection with historical cities, for instance. This can pose an empirical issue as Jeffery Wooldridge rightly points out that "even a small amount of correlation between [instrument(s)] and [unobservable(s)] can cause serious problems for the IV estimator" (Wooldridge, 2008: 514). We argue that although most historical instruments are not perfect, our chosen approach or method is

made most relevant as it ties nicely with the central theme of this entire thesis – one of promoting a plausible narrative that attempts to knit together history, geography, culture and institutions.

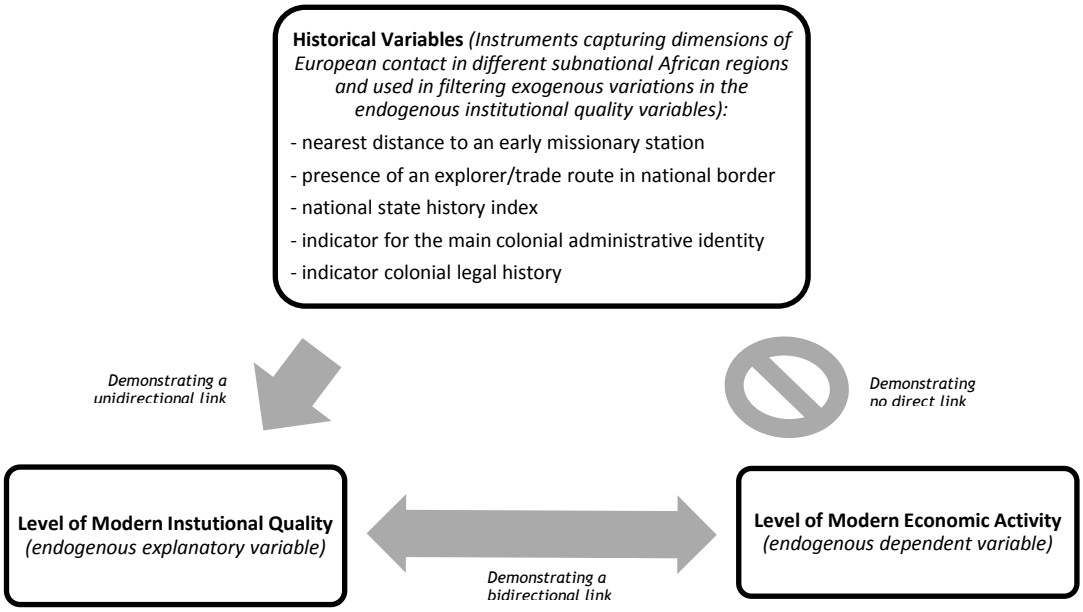

Figure F – 2.3.1: Diagram to Demonstrate the Indirect Link between our Selected Historical Euro-Contact Variables and Current Economic Activity via Modern Institutional Quality

To proceed, we attempt to individually consider each of the excluded instrumental variables used to proxy for European contact and that used to proxy for the national state history. The principal position for the instrumental variable approach proposed herein is that, the variations in the exposures of both subnational split ethnic homelands and their respective aggregated postcolonial national boundaries to differing contexts, types and forms of European contact, can provide a unique dais for linking known (observable) historical experiences with the current institutional evolution and eventual quality at the national level. The key premise put forward by our identification schema hence, is that the multivariate experiences in national institutional quality filtered by these key proposed forms of European contact at the different administrative levels can provide an exogenous causal channel to economic activity at the subnational split-ethnic level. This position is not new to the literature and can be found, for instance, in arguments proposed by Herbst (2000) and further engaged by M&P (2014). They postulate that the reason national institutional impacts may be

negligibly linked to subnational peripheral regions is because a number of factors limited the extent of pre-colonial and colonial relations with significant peoples and lands across the African region. However, this core theme can equally be harnessed to further explain why national institutions may, in fact, be relevant for subnational economic outcomes. Particularly given that we are able to observe relevant patterns and variations in key proximate historic forms of interaction between the institutionally impelling Europeans and the consequently impelled African territories.

Starting with the distance to the early missionary stations and a national indicator of the presence of any historic explorer and trading routes, we are able to capture proximity (or exposure) of an ethnic homeland to influences of missionary and extractive economic activities prior to independence. Because we are aware of the type of activities missionaries engaged in within their areas of location and operation from historical accounts, we can posit that proximity to the cultural, re-educational, and humanitarian activities injected by the missionaries existing at the time provided the bedrock for the implantation and cultivation of early institutional forms that resembled those from the missionaries' host countries. These neo-institutional installations, under normal circumstances, should have provided outlines compatible with modern notions of "good institution" as embodied by the developed countries of the post-colonial powers.

One could argue that the location of these missionary stations were themselves consistent with the location of other pre-colonial footprints such as early infrastructural development for the movement of goods and service to promote the economic activity dominating this era[42]. As a result providing a revealed falsification as a valid instrument given that such pre-colonial developmental injections would likely have equally set the precedence for modern levels of economic activity having persisted to this day - thus justifying what we now observe as higher levels of economic activity in these subnational locations. The problem with this latter counter-position is that the nature and type of economic activity at the pre-colonial economic era was not designed to enrich the local African economies. Instead economic activity then as we have come to understand, was depletory and extractive in nature - often orchestrated to enrich

[42] See Figure G - 2.3.2 below, for a visual sense of this argument.

the respective home economies of the associated European governments. That being said there is no apparent persistence or link between this kind of economic activity from the pre-colonial and early-colonial era with that of the modern era. What we rather had was a pre-colonial era characterized by exploitative economic activities designed to, at best, empower a tiny faction of individuals or micro-units within the exploited local communities to aid the promotion of slave trading and the other resource quarrying ventures. With these micro-unit empowerment schemes, it could be argued that the forms of economic activity during the pre-colonial and early-colonial periods provided the antecedence for the persistence of bad institutional manifestations such as corruption via intangible but real channels such as 'local elitism'. There was a clear break from these earlier periods from the kind of economic activity observed now. In more recent times subnational regions now have higher capacities for canvassing or lobbying for better circumstances and higher levels of economic outcomes from their respective national and subnational governments and administrative representatives.

Though arguably related as highlighted above, the second instrumental variable used herein attempts to capture the proximity of split tribal homelands to the trading routes. Since these early infrastructural feats served a considerably wide geographical area, it suffices to capture their impact at a national level rather than at the demarcated tribal level as we do with the missionary stations instrumental variable. The trading and exploration routes were a European championed infrastructural system built to transport resources and slaves from the hinterlands to the coasts for exports to lands beyond the oceans. Missionaries also located their stations strategically close to areas where deleterious economic activities were highest so as to maximize the impact of their humanitarian presence and benefit from proximity to the transportation of resources from their home countries. However, this doesn't affect the rationale justifying the use of these two instruments. It rather strengthens their concurrent adoption, as it demonstrates that pre-colonial European activities fringed about particular locations irrespective of the ethos characterizing or dominating each activity's basis. It is also pertinent that both instruments are adopted as they provide historically motivated subnational variations in national institutional quality of very distinct but related aspects of European activity in these African regions.

Both measures would equally allow us reveal whether these distinct set of European activities correlated differently or similarly with the national institutional dynamics of the associated countries.

Another very interesting fact related to the data and the choice of the proximity to missionary stations as a good instrument for capturing the formation of good institutional outcomes at the subnational level, is that most of these stations have remained to this day (White, 2010:12; Nunn, 2010:148) and were the locations from which the churches orchestrated their cultural and re-educational exploits in their respective African territories of operation. Being mostly Catholic and early protestants such as the Methodists, Baptists and Anglicans, they provided infrastructural facelifts in the form of church buildings and surrounding compounds from where early affordable re-educational and capacity building activities spurred. Being funded by their respective parent churches and mostly motivated by non-pecuniary objectives, they were also inclined to setup and expand to the non-favourable interior hinterlands to promote pro-European culturally transforming campaigns. This is in contrast with the infrastructure the Europeans erected to serve the economic purposes of that era such as the historical explorer routes (or early railways), which although still visible to this day are mostly non-functional. We can further demonstrate this by overlaying a current map of modern day railway routes over the map for the trading or explorer routes as shown in Figure G - 2.3.2 below. The historical missionary stations are also included to provide a visual idea of how these three experiences are visually correlated. We can rightly see that the two transport routes do not intersect in most cases, thus indicating the break in the objective behind the infrastructural developmental designs between the nature of movement at this historical exploitative period and that of today. It is also important to note that the initial infrastructural infusions by the missionaries promoted the cultural and instructive transformations that ought to have laid the foundations favourable to contemporary institutional conceptions and their eventual development. This however would not have directly fed into increased modern-day economic activity unless through recent favourable institutional development consistent with the current global political atmosphere. This argument is apparent because it is easy to see that while the motive of economic activities have considerably changed between the pre-colonial, colonial and post-colonial eras, the values, principles

and motives of most of the institutional fabrications that empowered the churches (missionaries) prevail to this day from their inception. Following this proposed narrative we can show that the impact of contemporary national institutional quality on modern economic activity could be estimated through the exposure of demarcated subnational ethnic homelands to patterns of the associated layout of historical European missionary stations and the early constructed trading or explorer routes. This therefore demonstrates that these two instruments are at least conceptually exogenous to unobserved factors of current subnational economic activity, all things being equal.

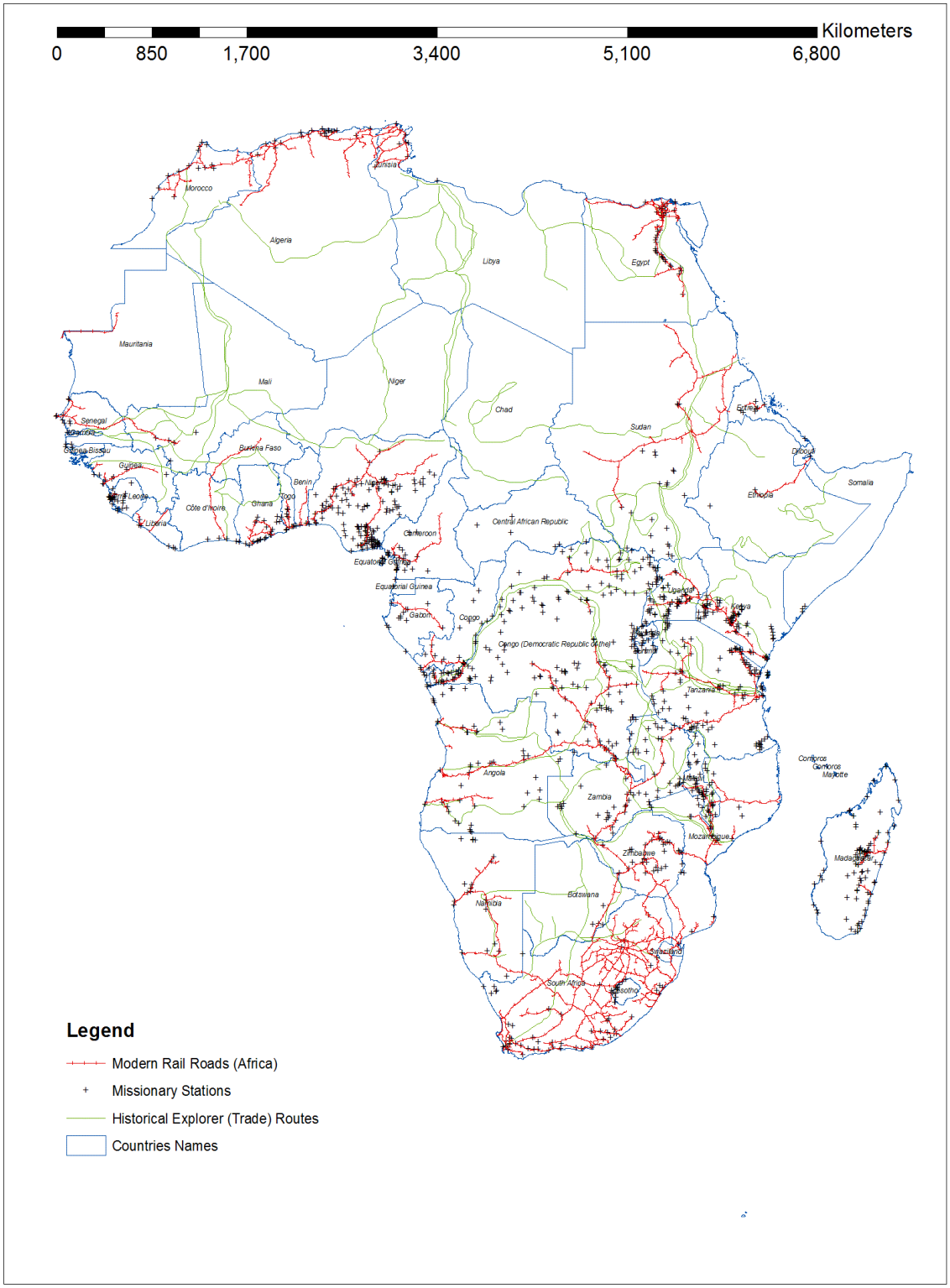

Figure G – 2.3.2: A Geographic Intersection of the Historical Explorer (Trading) Routes, the Modern Rail Routes and the Historical Missionary Stations in Africa

The third instrumental variable attempts to measure the history of a country's national identity associated with a demarcated ethnic homeland[43]. It varies at the same level as the national institutional quality variable and provides an indication of the extent of national administrative or coordinative amiability to the institutional shock of European influence at the time of their en-mass entry and involvement with the African territories from the pre-colonial times onwards. It is expected that modern national territories that had a longer history of multi-tribal coordination very similar to the subsequent post-colonial state would have been more agreeable to the exogenous aggregation of ethnic administrations into new single states with foreign common institutional frameworks. As a result, countries having higher state history indices are expected to on average have higher values of modern national institutional quality, other things (covariates and instruments) being equal (held constant or controlled for). Being more resilient to global shocks, older states are more likely to experiment and be more successful with outward looking ideas such as global political integration; thereby, deepening an efficacious adoption of institutional forms similar to those found in the wealthier Western economies.

It is difficult to find how this variable could map so lucidly unto economic activity at the subnational level via another channel as clearly as it would through its impact via national institutional quality. Unless we observed that there was a pervasive case of a few ethnic groups per country - so that most of the countries were fairly ethnically homogenous - we do not expect ethnic specific pre-colonial geopolitical characteristics to have dominated similar national level characteristics. But African countries were and remain mostly ethnically heterogeneous; as a result subnational geopolitical variations are quite distinct from those at the national level. Making it less likely that unobserved drivers contained in the error-process of the system equation varying at the subnational tribal level would map neatly or directly with this state history index (varying at the national-level); thus buttressing the exclusion restriction validation for such an instrument.

The last set of instruments control for both the colonial identity and the legal origin of the country associated with the split ethnic homeland. These are a

[43] We provide some more information on the source of this variable in the data sources section 2.3.4.3 below.

warranted set of instruments controlling for the colonial and early post-colonial advent of socio-economic and geopolitical systems in tandem with the foreign policies of the global political climate promoted by the leading European economies in the late 17th century onwards. Not only did each European colonial state set-up institutional and structural forms that were fundamentally different from each other, but also these institutions were a ubiquitous break from the pre-colonial geopolitical status quo of the previously disaggregated institutional and structural practices of the already existing tribal states. With such European driven institutional implantations and the concomitant exogenous aggregation of pre-existing tribal lands distinct legal, economic, and political institutional forms and evolutionary paths were stirred.

It is evident how this provides a viable channel for exploring a dimension through which current institutional variations at the national level appear to be pulsating current disaggregated differences in the economic activity of split ethnic homelands as observed in our study. Also, because most of the types of economic activity embarked on by the colonial states in most of the African territories were extractive or locally unpopular, it is difficult to see why unobserved sources of current levels of economic activity would co-vary with this set of instruments rendering them invalid. The evidence of the unpopular nature of the colonialist administrations is buttressed by the unanimous efforts made by the African nation states to seek, demand and fight for self-governance independent of their previous respective colonial governments. However, because the different foundations of institutional arrangements took their roots and persisted in most cases, we find a channel through which this exogenous shock affected current institutional manifestations further impacting modern levels of economic activity.

2.3.4 – Variables and Data Sources

2.3.4.1 – The Dependent Variables Measuring Economic Activity

To advance our analysis herein, we adopt two main proxies of economic development (economic activity) observed at the subnational split ethnic homeland level. The first of these is based on the DMSP-OLS nightlight data from the National Oceanic and Atmospheric Administration (NOAA) data repository[44]. This dataset provides a measure of nighttime light density at a 30 arc-second resolution excluding other noisy light sources from the sun, the moon, glares, cloudy observations, lighting features such as auroras, and etc. The values in the variable are nominal and range between 0 and 63. We use a more recent account of the average over the 2010-2011 period. The potential problems with using the stable night light data, as a time series is that such a measure would definitely be very dependent on the condition of the satellite detecting it and on prevailing atmospheric or cloud circumstances. While it is easy to account for atmospheric and cloud conditions in such a model it is a lot more difficult to guarantee and control for conditions in the health of the detection satellite and that is why this dataset is most useful for cross-sectional purposes for any given point in time.

On the second matter of light stability at night in most African countries, it is common knowledge that electrification is a challenge and in most cases power generation for domestic artificial illumination is usually shared between the national power grid and private power generation sources. With this in mind, it is not startling to expect that rationing is prevalent and is the norm. For this reason, providing an alternative measure of economic activity at the same geopolitical administrative level as the nighttime lights data is pertinent. This is because regions may most likely have suffered protracted blackouts over the period the satellite did its recordings causing the polygonal average of an area to be small even though it might have had as much or even more light at other times throughout the day compared to the other areas measured with relatively higher light intensities. For this variable, we obtain averages for each demarcated ethnic homeland polygon and compute the natural log of the resulting polygonal average pixel values.

[44] http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

The second proxy for economic development (our dependent variable) is a relatively more stable and permanent cumulative measure of infrastructural development. It measures the average lengths of all access road traversing each demarcated ethnic groups' homeland polygon. These access routes include any official, recognised and named access route on land and over water compiled for each split polygon in our sample. This paper has applied this dataset as an alternative (robust) measure to confirm whether the treatment effects we observe permeates key facets fundamental to everyday economic activity. Road infrastructure is a relatively steady capital-intensive developmental feat and it can show to a large extent the level of or potential for economic activity in any average city or dwelling area (or in this case a split ethnic region). We obtain this data as a 'geodatabase' of polylines from the Socioeconomic Data and Applications Center - SEDAC website[45]. It has been compiled globally using data entry sources from each country and as a result the road network compilation ranges from the 1980s to 2010 depending on the countries and what dates they confirm with their respective entries into the repository. The data version we have employed in this research is the Global Roads Open Access Data set (gROADS) v1. The dataset uses as common data model approach on the UN Spatial Data Infrastructure Transport (UNSDI-T) version 2. It is collected from multiple sources and hence the road network representation is available over the span of 1980 to 2010, depending on the country.

In a majority of cases, road network representations from source countries have no confirmed dates. Although this is a weakness, the infrastructural coverage is not systematically correlated with other fundamental factors inherent to individual countries as a number of measures were taken by the dataset's compilers to ensure data validity using Google Earth high-resolution imagery and GPS track data to confirm individual data source entries. Additionally, an update to the dataset was compiled for 27 countries and 6 smaller geographic units by Columbia University's Center for International Earth Science Information Network (CIESIN). They focused mainly on developing countries with relative data paucity. With such a concerted approach, this dataset provides comparability of average road lengths across areas within its scope of coverage. In this study, we compute

[45] http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1

and adopt the variable in a similar fashion to the nighttime lights density data above by obtaining averages for each split ethnic homeland polygon and then take the natural log.

Although the nighttime lights data has been discussed and referred to quite extensively in the literature as a relatively good proxy measure for economic activity, the roads access proxy has not received equal use especially in the economic growth literature. Since we are going to be using this additional proxy in this chapter and the next (chapter 3), we attempt to provide some demonstration of its suitability as a fairly decent proxy for economic activity, as well. Roads density has most popularly been previously used as right-hand side control measures for infrastructural development and urbanization in the economic development literature (Duranton et al., 2014; Faber, 2014; Storegard, 2016)[46]. Roads density has been referred to in the GIS (geographic information systems) discourse as a companion proxy for economic development and urbanization at varied levels of political administrations or spatial resolutions (Mellander et al., 2015). It is also most suitable for developing countries where data paucity is an additional concern (ibid). Additionally economic activity or urbanization is a multifaceted phenomenon, so resorting to a single metric would not tell a well-rounded story. The variations in economic activity that both proxies (nighttime lights and roads access) bring to the study are not entirely overlapping. Both proxies capture diverging components of economic activity; thus using them both allows the study to bring these multi-dimensions into focus. This might be the reason why both proxies do not necessarily correlate as highly with each other as they individually do with our real national income measure – GDP (refer to the plots immediately below). For one it is easily arguable that while electrification could be pervasively relevant to an economy, road access could be biased towards key industries and sectors (see Chander & Thompson, 2000 for such findings on the selective economic impact of highway construction). This does not to dispute the role roads access can play in improving economic activity in general. It is in

[46] These studies however focus on different aspects of the access roads measure such as paved road, highways, distance between cities along road networks, etc. By focuses on average roads access length across demarcated polygonal boundaries, our study provides a type of quality measure as well rather than a quantity capture – which is what a density measure (mere count) would provide. Due to data availability for road types and other related characteristics, obtaining such classifications of roads in the dataset across all the countries in the SSA region for our sample would not be possible. So this provides a convenient adaptation to the dataset's weakness in this regard.

fact quite conceivable to see how economic activity and development can be improved in this regard. Beuran et al. (2015:674) elaborate on a few of these channels in their study, which include – to facilitate key public goods (health and education, for instance) provisioning, facilitate markets access, and employment.

Nighttime lights could pose additional issues as a sole proxy for economic activity in countries such as in Sub-Sahara Africa (SSA) countries where electrification rationing by the national grids of most countries is a pervasive reality (see Bhattacharyya, 2012; Cook et al., 2015). It might be a national or regional strategy to spatially allocate electricity to industrial areas during the daytime and then reallocate power (electricity) to residential areas at night. That said, it would result in such industrial areas being dark at night when the satellite sensing is occurring causing regions with many of these included industrial locations to be persistently darker relative to other locations when they might otherwise be booming with economic activity all year round. Therefore bringing the roads access proxy to the literature provides an additional robustness measure given that roads are a more stable infrastructure based indicator that is not biased in such a manner. One downside for using this measure is that, the roads access measure is a stock over the period for which it has been captured. This is a weakness given that the income measure it is meant to proxy is usually a flow. This weakness is however mitigated because the models we adopt using these proxies are setup in a cross-sectional fashion, which conveniently bypasses the concerns of variations over time.

Below, we provide some plots to show the correlation between summed nighttime lights and GDP (gross domestic product), summed length of access roads and GDP, and summed-up length of access roads and nighttime lights – all aggregated to the national level. The second of the plots does the same thing but looks at the variables in their weighted-average forms so as to facilitate comparison across countries of different population levels and land masses. We decide to look at these variables at the national aggregate levels as there are no reliable measures of economic activity available and comparable across the countries in our adopted sample.

In the first set of these plots (Figure H - 2.3.3, below) we see the linear association (at the levels) between the popular nighttime proxy (in sums) and real national income (GDP) is as strong – at about 93% - as similarly maintained in the

literature (see Ghosh et al., 2013). The linear association at the levels for the summed length of access roads with real GDP is about 61%. While the direct correlation between nighttime lights and the alternative roads access proxy is about 54%. This is a lesser degree of linear association between both proxies compared to their individual linear associations with GDP; which is equally maintained after they have both been weighted by the size of each country's polygon (about 14% in this case). As shown in Figure I - 2.3.4 below, when we consider the access roads proxy in its polygonal average form it is as correlated with real GDP per capita as is the average nighttime lights (about 32% and about 36% for average roads length-to-GDP per capita and average nighttime lights-to-GDP per capita, respectively). As earlier highlighted, while these show that they are both correlated with national income (at both levels and per capita measures) to similar extents, their relatively lower correlations with each other indicate that they might be capturing disparate dimensions of economic activity that may be worth considering.

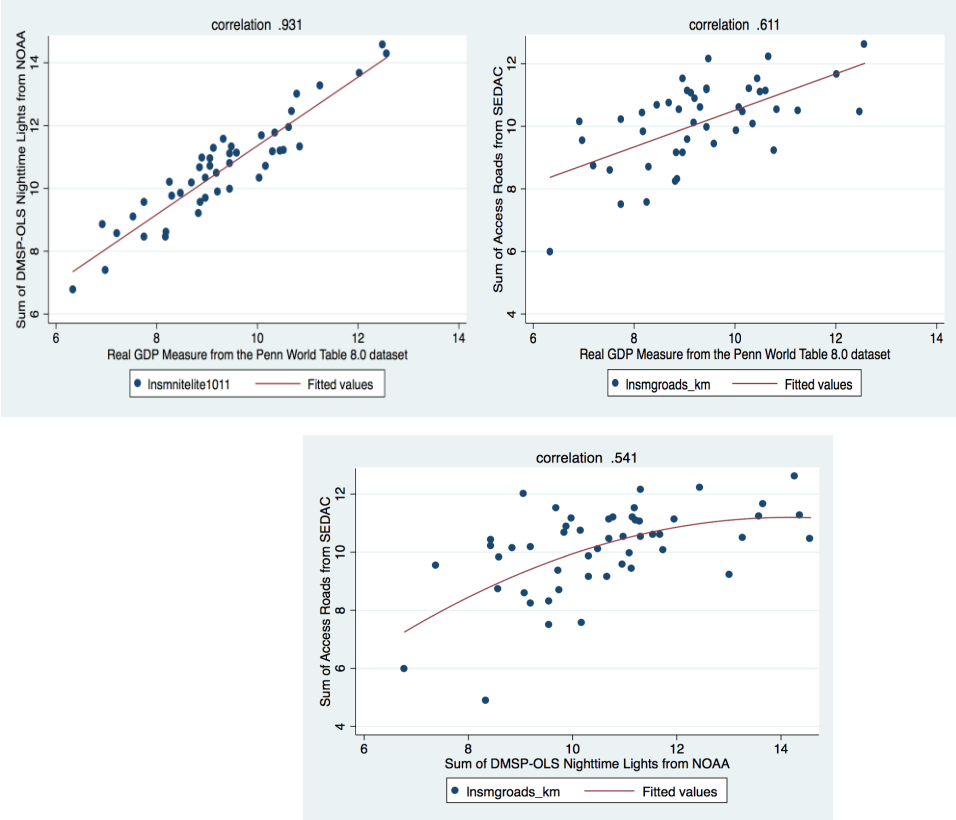

Figure H – 2.3.3: Plots Showing the Linear Association between the 2 Proxies of Economic Activity (polygonal sums) and Economic Activity with Real GDP

Where: the income variable is measured in real national aggregate terms collected from the Penn World Table 8.0, while the nighttime lights and access roads are also in national aggregate levels (they have not been weighted in this figure); plots using 51 African countries and all values have been logged for rescaling reasons

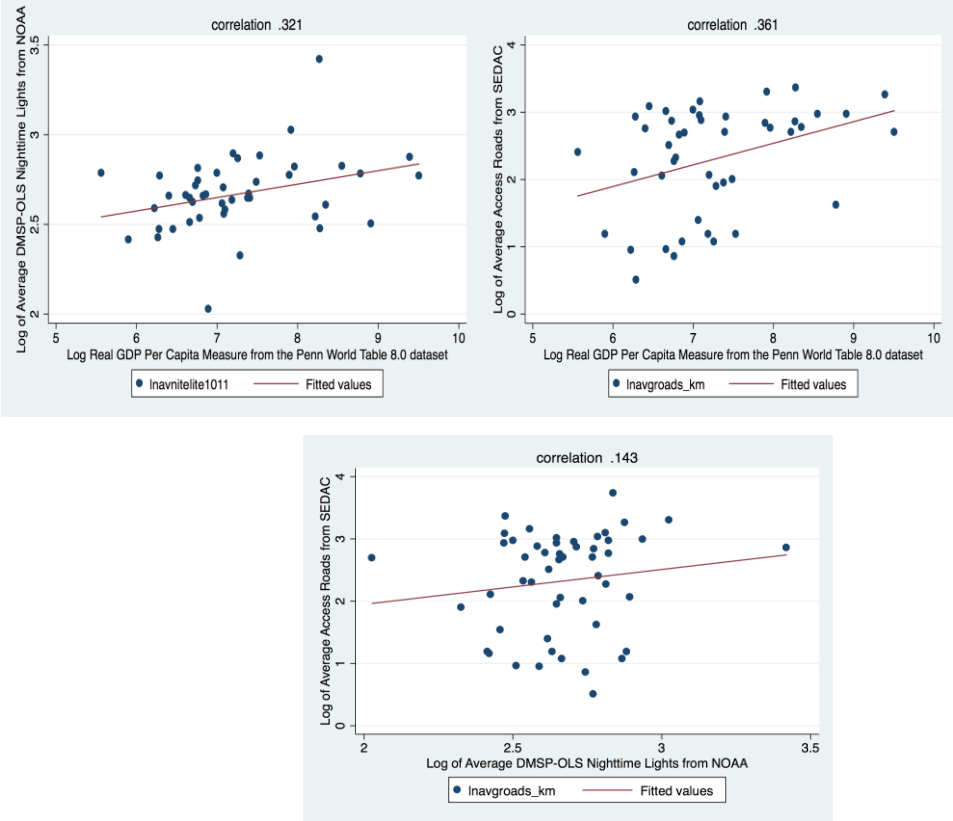

Figure I – 2.3.4: Plots Showing the Linear Association between the 2 Proxies of Economic Activity (polygonal weighted-averages) and Economic Activity with Real GDP Per Capita

Where: the income variable is measured in real per capita national aggregate terms averaged from 1990-2010, while the average nighttime lights (2010-2011) and length of roads access (1980-2010) are measured as aggregate national polygonal-averages; plots using 51 African countries

2.3.4.2 – The Endogenous Treatment Variables Measuring Institutional Quality

There are four institutional quality measures adopted in this paper. Three are from the World Bank Governance Indicators (WBGI) by Kaufmann, Kraay, & Mastruzzi (KKM, 2009, 2011) - the corruption control index, the rule of law index, and the governance effectiveness index. All three indices have been averaged from time of initial availability, 1996 to 2010. The values on these indices range from -2.5 to +2.5, where higher values represent higher institutional quality and vice versa. The fourth measure of institutional quality is an indicator of the quality of governance measure based on the mean value of the ICRG (International Country Risk Guide) data variables: "Corruption", "Law and Order" and "Bureaucracy Quality", scaled between 0 and 1. Higher values in the scale

represent higher quality of government values. This composite measure is obtained from the Quality of Governance Basic Dataset, a project by the University of Gothenberg. We have adopted this 'institutions' metric as an alternative measure of institutional quality in our study. This variable has been averaged over the period 1990 to 2010 at the country level accompanying each split ethnic homeland.

The choice of these four indices to capture institutional quality are by no means comprehensive. There are alternative measures of institutional quality available out there in the literature. We have particularly chosen the first of these two (the corruption control index and the rule of law index) so that our discussions and results are comparable to the immediate literature as per Michalopoulos & Papaioannou (M&P, 2014). While the latter two (WBGI's governance effectiveness index and the ICRG's quality of governance measure) are arbitrarily added for robustness (comparison) purposes. The latter two indices also provide a relatively broad measure (in topic scope) to the subject of institutional quality; thus making it most applicable to our study as concerns governance in a more holistic fashion for the African countries under study. The other measures in the World Bank Governance Indicators (WBGI) such as the voice and accountability, political stability, absence of violence and regulatory quality measures capture relatively more specific dimensions of the institutional quality subject matter and tend to be popular in the literature for studies relating to property rights, political instability and freedom of speech and association topics. Although these topics are equally important we do not include them herein; particularly for the sake of analytical brevity and contextual relevance. As a result we decide to restrict our analysis to these four institutional quality measures or indices.

2.3.4.3 – The Instrumental and independent Exogenous Variables

A bulk of the regressors in this paper's models is from Geographic Information Systems (GIS) sources. This is particularly useful as this study's unit of analysis is at the subnational level and most development-based measures available are usually compiled at the country-nation level. Additionally, GIS data and modelling allows for a more relevant opportunity to quite easily control for geographic influences. Geographic influences have been theorized in the literature to exert quite an effect on key development indicators in the empirical literature (see

Bloom et al, 1998; Sachs, 2001; Easterly & Levine, 2003; Nunn & Puga, 2012; etc.). This chapter's analytical reference point is the split ethnic homeland polygon based on the anthropological work on pre-colonial ethnic societal characteristics by Murdock (1959). Our dataset is very similar in structure to the replication datasets used in the Nunn & Wantchekon (2011) and M&P (2013) studies.

Following similar practice as characteristic in the literature using these dataset types, we match each ethnic or tribal group's identified homeland from the Murdock map to the countries currently recognised over these territories. This is done using a spatial joining tool with respect to the spatial reference information available from both the Murdock map and those of the countries currently overlaying these areas. There are in total 835 tribal homeland polygons identified in the Murdoch (1959) spatially referenced ethnic map for the African region excluding the 8 uninhabited areas located on the map. From this larger sample of tribes, we are able to identify 227 valid-split tribes whose historical homelands fall in at least 2 modern-day independent countries spanning a sample of 47 countries on the African continent. This gives us a full sample of 526 observations in total. We have tagged these demarcated homelands as valid-splits with the additional conditions that each country-split ethnic polygon is greater than a 100 square-kilometres and belong to a tribe with at least 10% of its historical homeland found in more than one modern country's geopolitical territory. This latter criterion is consistent with that adopted by M&P (2014). We have used this threshold to accommodate for trivia splits, which may be the result of any erroneous outcomes of the geocoding process itself. Figure J - 2.3.5 below exhibits the full sample and valid-split sample of the Murdock (1959) tribal homelands map with the countries currently wielding their geopolitical institutional fervour over these African territories, respectively.

Most of the distance-based (distance to border, distance to sea coast, and distance to missionary location) metrics have been calculated using the nearest distance algorithm in ArcGIS 10.1. They represent the distance from the split ethnic homeland's centroid point to the nearest polygon or point of the relevant phenomena of interest. The distance to seacoast is computed using the 'Water Polygons' data available on the Open Street Map Data website[47]. The distance to

[47] http://openstreetmapdata.com/data/water-polygons

capital variable on the other hand was obtained using a centroid-to-centroid geodetic distance algorithm ("geodist") user-written for STATA users. These distance variables are also all measured in kilometres.

The polyline shapefile used for the exploration route analysis is obtained from Nunn & Wantchekon (2011) based on work originally from Roland A. Oliver (2000) and digitized by Matthew T. Ciolek (2001)[48]. The data on the location of missionary stations is obtained from Nunn (2010) based on a map published by William R. M. Roome (1924)[49].

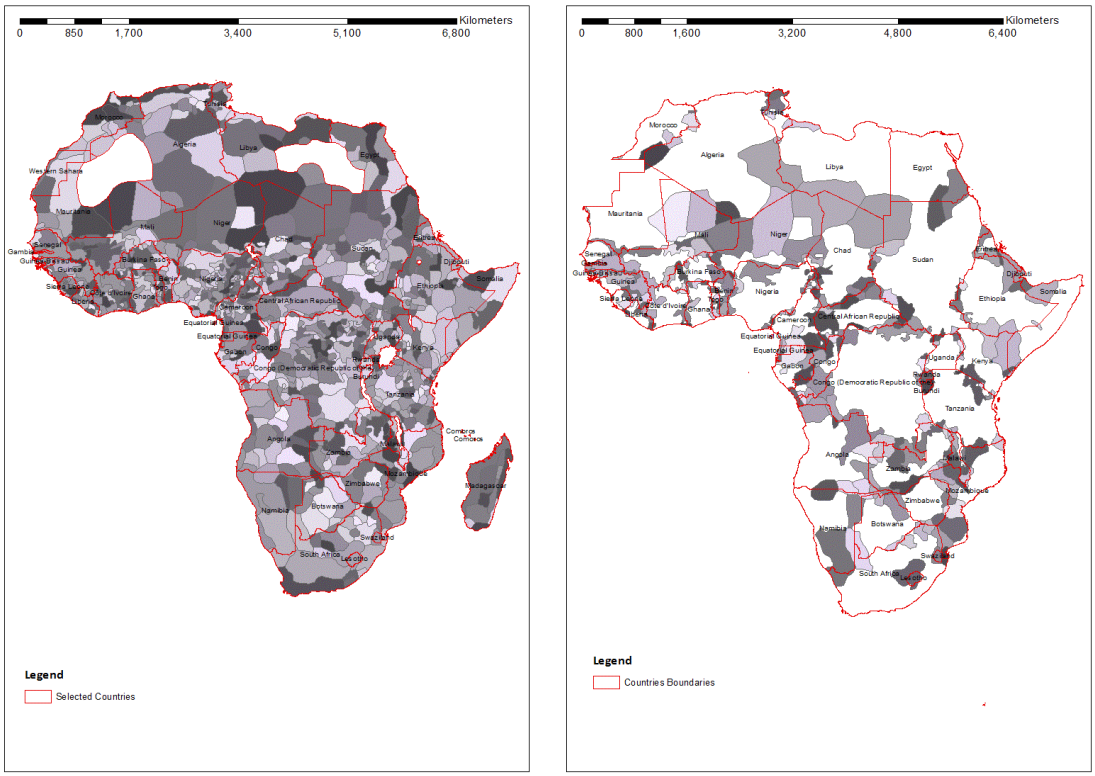

Figure J – 2.3.5: The Original 835 Ethnic Homelands (Murdock, 1959) and the accompanying 233 valid-splits

The state history index, one of the excluded instruments, is the 'statehist' variable version 3.1 obtained from the work of Bockstette, Chanda & Putterman (2002)[50]. This variable provides a proxy measure between the values of 0 and 1, where 0 indicates a country with the least state history and 1 indicates one with

[48] This "data is on the historical location of towns and routes from the Sahara trade…" It has been referenced in Nunn & Wantchekon (2011: 3240). See Figure G - 2.3.2 above, for a visual presentation of the spatial location of these routes.

[49] See Figure G - 2.3.2 above, for a visual of where the stations were located.

[50] http://www.brown.edu/Departments/Economics/Faculty/Louis_Putterman/antiquity%20index.htm

the most state history. The index overall, attempts to capture the extent to which a modern national state was the location of a united nation-state, empire, or any other form of unified political administrative entity every half-century since the first century.

The population density data is obtained from the Andy Nelson (2004) African Population Distribution Database, UNEP GRID Sioux Falls Dataset of population numbers per square-kilometres in 2000. This data is available at the United Nations Environment Programme (UNEP) Environment for Development website[51]. The variable has been natural-logged in the estimations carried out herein.

A number of additional control variables are used to account for the natural endowments of each demarcated ethnic homeland. There are two indicators to control for the natural resource endowments. These are dummy variables to identify the presence of diamond deposits and onshore deposits of oil and gas petroleum products, respectively. The diamond deposits variable is obtained from a shapefile from the Diamond Resources dataset on the USGS website[52]. While the petroleum deposit dataset is available at the Peace Research Institute Oslo (PRIO) website[53]. An additional variable controls for a split group's (split ethnic homeland) proprietorship of inland water; this is measured as the total area covered by a river, stream or lake. This variable is available for download as a raster file, titled Global Lake and Wetlands Database Level 3 (GLWD-3), on the World Wildlife Fund (WWF) website[54].

In convention with the literature, key aspects of a split group's natural ecology are equally controlled for. Four widely treated dimensions are considered: general ecological ambience using measures of mean rain and mean elevation, agricultural ecology measured using a soil productivity index, and a disease ecological measure - using the malaria ecology index. The mean elevation variable measures the average surface elevation for each split ethnic polygon; based on a joint work by the National Oceanic and Atmospheric Administration (NOAA) and the U.S.

[51] http://na.unep.net/siouxfalls/datasets/datalist.php

[52] https://www.prio.org/Data/Geographical-and-Resource-Datasets/Diamond-Resources/

[53] http://www.prio.no/CSCW/Datasets/Geographical-and-Resource/Petroleum-Dataset/Petroleum-Dataset-v11/

[54] http://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3

National Geophysical Data Center, TerrainBase, release 1.0. The soil productivity index measures the average suitability of the soil of each polygon in our sample to the crop best adapted to that soil's type, condition and composition. It is a dataset published by the FAO in 2002 with index values ranging from a minimum of 1 to a maximum of 97; where values increase in tandem with the soils productivity and suitability to the chiefly adapted crop. Finally, the malaria ecology index is based on the work of Kiszewski, Mellinger, Speilman, Malaney, Sachs, and Sachs (2004). It measures the prevalence and exposure of each split ethnic polygon to the malaria endemic.

2.3.5 – Initial Summary Statistics for Variables of Interest

Variable	Obs	Mean	Std. Dev.	Min	Max
Log(0.001 + Average Nighttime lights density index 2010-2011)	526	0.5516535	2.797105	-4.60517	3.270541
Log(Average Roads Length over the 1980-2010 period)	519	2.840186	0.8937715	-0.7694919	5.620769
Country's WBGI Corruption Control Index (average 1990s - 2010)	526	-0.7727009	0.4919724	-1.738412	0.8818495
Country's WBGI Rule of Law Index (average 1990s - 2010)	526	-0.889778	0.5735923	-2.348055	0.6000053
Country's WBGI Governance Effectiveness Index (average 1990s - 2010)	526	-0.8317847	0.5430416	-2.174895	0.5986422
Country's Quality of Governance Index (ICRG QOG - average 1990s - 2010)	451	0.3879278	0.1195069	0.1258818	0.6521715
Country's WBGI Corruption Control Index (average 1990s - 2005)	526	-0.7745004	0.5183785	-1.676455	0.8334835
Country's WBGI Rule of Law Index (average 1990s - 2005)	526	-0.9141688	0.5914214	-2.197272	0.5751488
Country's WBGI Governance Effectiveness Index (average 1990s - 2005)	526	-0.8114545	0.5529397	-2.070641	0.6860611
Country's Quality of Governance Index (ICRG QOG - average 1990s - 2005)	451	0.3960776	0.1248166	0.1280382	0.6793258
Nearest Split Ethnic Centroid Distance to Confirmed Missionary Locations (1,000s KM)	526	0.2070467	0.2319605	0.0015858	1.313903
Indicator for the presence of a historical explorer route passing through country's border	526	0.8840304	0.3204932	0	1
State History Index (Bockstette, Chanda & Putterman, 2002 - discounted at 10%)	487	0.4105477	0.2152004	0.0465868	0.9473389
Never Colonised Indicator	526	0.039924	0.1959669	0	1
British Colony Indicator	526	0.3707224	0.483458	0	1
French Colony Indicator	526	0.4239544	0.4946537	0	1
French Legal Origin Indicator	495	1.591919	0.4919754	1	2

Log (1 + Population Density at 2000)	526	3.006171	1.749962	-4.60517	6.242243
Distance to Sea (Measured in 1000 KMs)	526	0.6301444	0.4522818	0	1.796686
Distance to Country Border (Measured in 1000 KMs)	526	0.0406222	0.0439317	0.0000553	0.320909
Distance to National Capital (Measured in 1000 Kms)	526	0.5504251	0.3947329	0.0050121	1.929665
Mean Elevation (Measured in 1000 KMs)	526	6.016954	1.083722	1.062245	7.795999
Soil Productivity Index	526	9.460107	14.57147	1.307692	88.54546
Malaria Ecology Index	526	13.8157	8.608235	0	33
Log Split Land Area	526	2.00502	1.551917	-2.149397	6.125164
Diamond Mine Present Indicator	526	0.0969582	0.2961825	0	1
Petroleum Well Present Indicator	526	0.0608365	0.2392577	0	1
Indicator for the presence of Inland Water Body (KM)	526	0.7604563	0.4272113	0	1

Table K – 2.3.1: Tabular Summary of the Variables used in the Models

As highlighted earlier, there are 526 observations from 227 tribal groups represented in 47 countries on the African continent within our designed sample. This is the full sample that constitutes the valid-split ethnic homelands. However, the observations in each specification are driven by data availability across the variables utilised in the different estimated models because, not all the variables have observations across the full sample of 526 for the 227 valid-split tribal homelands. The access roads length, institutional quality, legal origins and state history index variables have 'missing' observations in a number of instances. This certainly poses a limitation to the data points available to the estimations carried out, but not so significantly as to alter the findings and results embedded within the discussions carried out. In a similar light, our logged nighttime lights variable would have untenable regions with pre-transformation values of zero. So in conformity with the literature, we adjust this variable by first adding 0.01 to it before taking its natural-log. A summary table of all the variables involved in our regressions are provided in the Table K - 2.3.1 above.

The spatial intersection procedure utilised in this work also helps to mitigate observational loss even further. It unifies each demarcated tribal homeland within each country even when these tribal homelands are further split into disjointed homeland-islands by the associated country's national frontier. We have tagged such disjointed but amalgamated homeland polygons as non-linear splits, while those with no apparent homeland boundary interruptions have been termed as

linear splits[55]. A graphical illustration has been provided below to demonstrate the distinction between these two coined terminologies - linear splits and non-linear splits. As illustrated in Figure K - 2.3.6 below for instance, the tribal land of the *'Dialonke'* is found in three countries (Senegal, Mali and Guinea). But because of the non-smoothness or irregularity of the national borders in most parts of Africa, tribes such as the *'Dialonke'* and *'Foutadjalon'* have their homelands split within and across national border demarcations. The visible lack of tribal border continuity for the *'Foutadjalon' tribe in Guinea* in Figure K - 2.3.6 demonstrates this appropriately, for instance. As a result, it is useful that the spatial intersection tool we adopt incorporates an aggregation or fusion function that identifies and matches tribal homelands at an amassed tribe-to-country level regardless of whether this is from a uniform continuous tribal homeland polygon or not.

An illustration of the linear type split on the other hand is much more predominant and is what one would naturally expect to find when conducting this kind of research. Its type is illustrated in Figure L - 2.3.7 below between the *'Boran'* in Ethiopia and the *'Boran'* in Kenya. Given the prevalence of the linear type splits, it brings to question the widely held claim in the literature, that tribal homelands and their resulting wellbeing were significantly affected by the outcomes of such non-linear intra-group demarcations. It indicates that such claims might be over-propagated far and beyond their actual relevance or importance in explaining local economic developmental dynamics. It appears more likely that other fundamentals at the tribal (ethnic) or micro levels may be at play not necessarily associated with these border related issues observed at the national frontier.

[55] This is done for illustrative purposes and is not characteristic of any associated terminologies used in the accompanying literature.

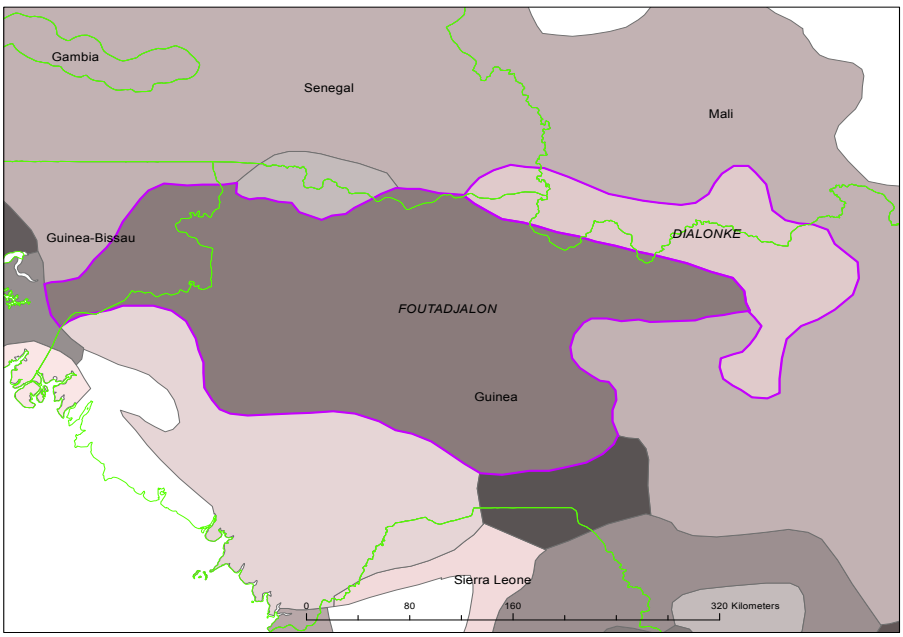

Figure K – 2.3.6: An Illustration of a Non-Linear Split

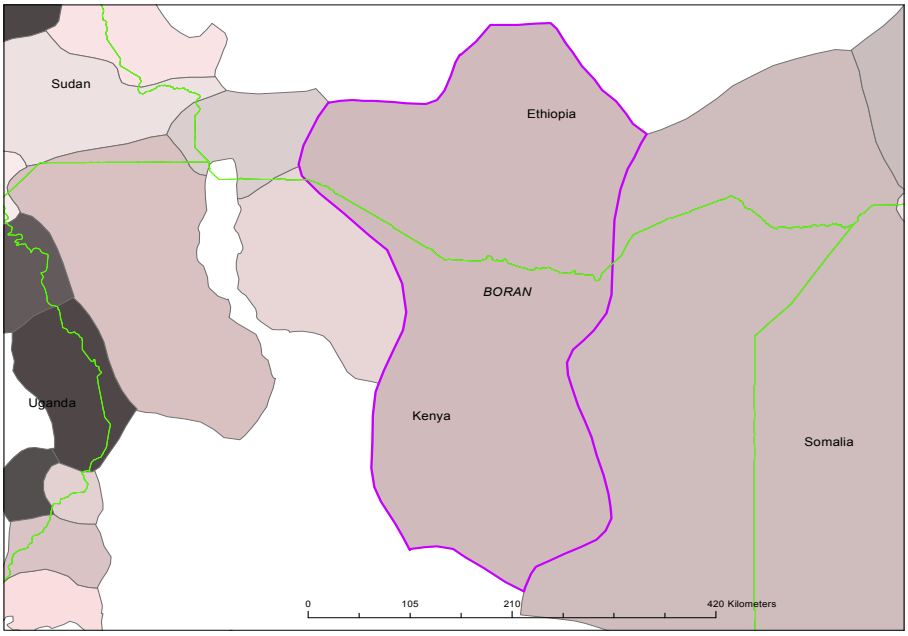

Figure L – 2.3.7: An Illustration of a Linear Split

When it comes to visualizing the measures of economic activity and of institutional quality used in our model estimations, we find some consistency in the claim that there are observable breaks or cross-sectional variations at the split ethnic homelands level. These breaks can be observed at both administrative levels characterising this research; that is both at the national country-level and the subnational split-tribal level. It can also be observed, from Figure M - 2.3.8 below, that there is also a considerable degree of within country variation across

ethnic groups within each country for the dependent variables, which have been advantageously observed at the demarcated ethnic homeland administrative level.

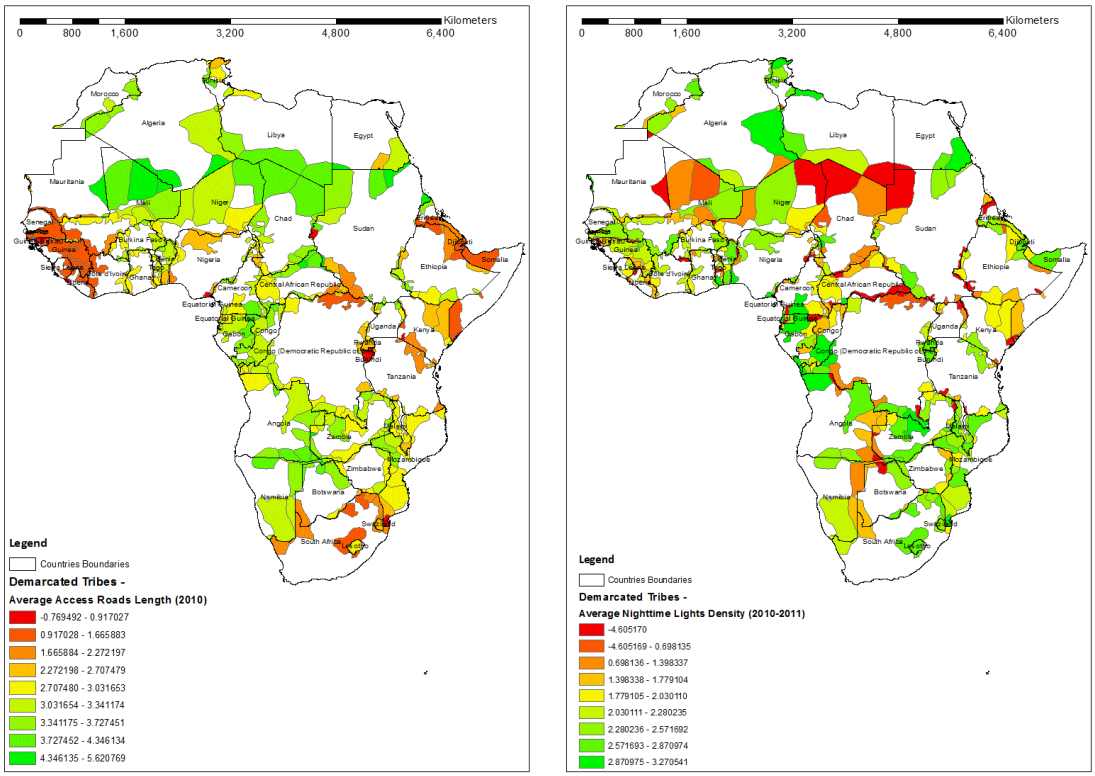

Figure M – 2.3.8: Average Nighttime lights density (Log (1 + Average 2010-2011)) and Average Roads access length (Log (average cumulated up to 2010))

Looking at the institutional quality variables, we are able to also observe the break in institutional quality between splits of the same tribal origins found in different countries much clearer. M&P (2014) use the discontinuity charts and methodology to portray this much clearer; it is nonetheless as clear in the graphical illustration displayed in Figure N - 2.3.9 below. We can see how the colours vary across split ethnic homelands at the country border subnational locations with the World Bank's Corruption control and Rule of law estimates based on Kaufmann et al (KKM, 2009) – averaged over the 1996 to 2010 period. Although the values are changing at the national frontier as expected, the respective colours represent cluster-ranging values and should not change so readily if national institutional outcomes were regionally clustered, for instance.

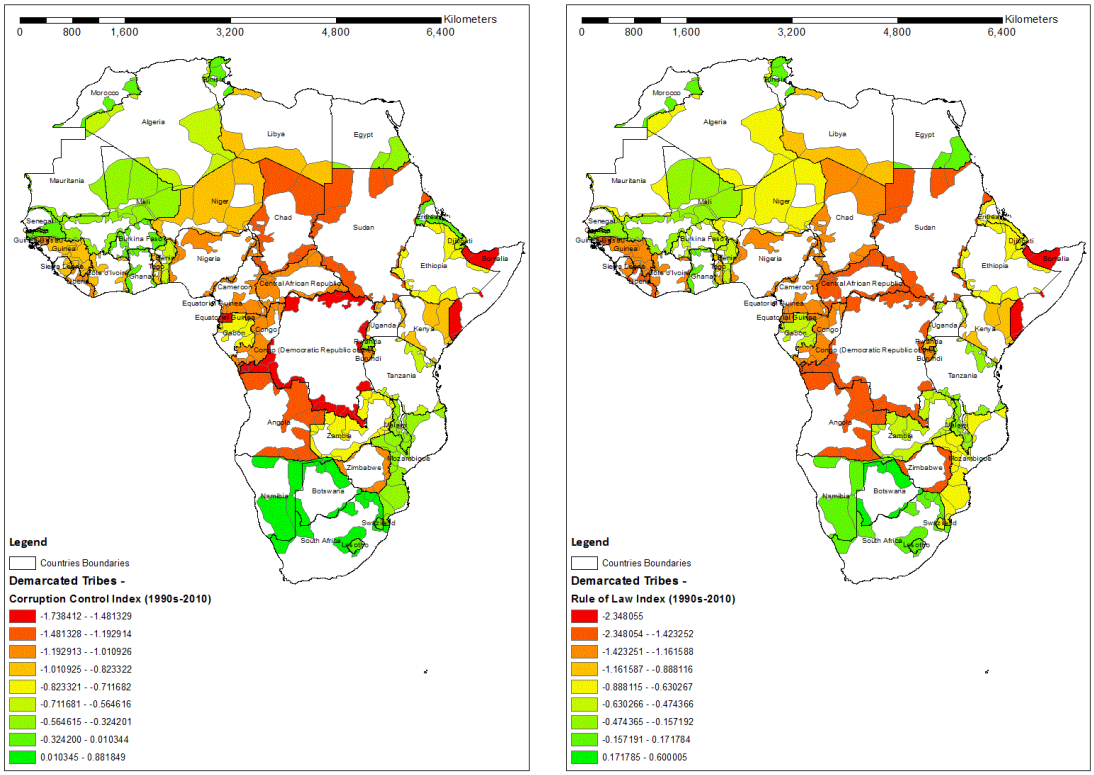

Figure N – 2.3.9: National Average Corruption Control and the Rule of Law Estimate from the World Bank Governance Indicators (1996-2010)

Section 2.4 – Results and Interpretation

2.4.1 – The Base Model

From the results of the base model in Table L - 2.4.1 below, it is evident that national institutional influences are on average positive in their impact on economic activity at the subnational demarcated ethnic homeland level. This observation is consistent across the average nighttime lights density data used in the literature and with the roads access data, which we adopt in this paper. The claim by M&P (2014) that the statistical significance is contestable and sensitive to other relevant controls present in the model especially ethnic fixed-effects remains very much held and consistent with our findings. When ethnic fixed-effects are controlled for, it becomes evident that national institutional quality starts to become mostly statistically insignificant and the magnitudes of these coefficients also reduce in general. This finding is not only consistent with the literature but also with our additional (robust) measure of economic activity,

which we propose in this work. Apart from instances with the second and fourth institutional measures in our average access roads length estimations below, the rest are statistically insignificant once tribal specific characteristics are controlled for.

The trend and behaviour of the coefficients of the institutional variables reducing in both magnitude and significance when tribal fixed effects are controlled for, to some extent, may indicate the relevance of un-observed tribal factors over national institutional impacts at the bordering subnational regions or territories. It may also indicate reasons associated with the possible notion that national institutional dictates are not so welcome or as effective in bordering regions of each country, where a sense of national identities may be less important compared to a sense of tribal identities. This is not such a farfetched inkling for a continent such as Africa, where border controls are relatively weak and a sense of fused belonging and identities characterise border subnational regions. As is quite apparent, their tribal identity happens to be the predominant reality at the national boundaries, which have affected pervasive national homogeneity by introducing diverging national loyalties. Such a conflict in loyalties could certainly affect the effectiveness or reach of national institutional values at such peripheral areas of the associated nation states.

There are certainly other elements or even components of a country's very national institutional makeup driving economic activity for these subnational regions at the peripheries but they might require a clearer identification approach to capture them. What we observe in these initial results could also be a signal explaining an increased tendency for bordered subnational regions to rely on non-national institutional frameworks at the peripheries of national identities. Further reflecting not only the weaknesses of the extent of country-level institutional influences but also the notion that a sense of national institutional identities take a relatively longer time to take root; especially in such subnational regions where tribal histories are much older and hence more crystallised. A deeper assessment of such intricacies is certainly very plausible and can be further carried out adopting methods that systemically harvest the common experiences of these subnational regions present in more than one national institutional identity simultaneously. It is in this light we go ahead to carry out deeper assessments of

a similar framework in the next section, but by adopting history-motivated instrumental variable (IV) estimations.

Dependent Variable: Logged Split Ethnic Land's Average Stable Nighttime Lights Density at 2010-2011 (NOAA data repository)								
VARIABLES	(A) y1	(B) y1	(C) y1	(D) y1	(E) y1	(F) y1	(G) y1	(H) y1
WBGI Corruption Control Index (average 1990s - 2007)	0.455** (0.190)				0.264 (0.268)			
WBGI Rule of Law Index (average 1990s - 2007)		0.529*** (0.172)				0.263 (0.219)		
WBGI Governance Effectiveness Index (average 1990s - 2007)			0.528*** (0.177)				0.294 (0.232)	
Quality of Governance Index (ICRG QOG - average 1990s - 2007)				2.233** (0.998)				1.427 (1.159)
Population Density (at 2000 in 1000s of Peoples)	0.678*** (0.0759)	0.676*** (0.0750)	0.670*** (0.0758)	0.612*** (0.0758)	0.350*** (0.127)	0.347*** (0.126)	0.352*** (0.127)	0.380*** (0.115)
Adjusted R_Square	0.353	0.358	0.357	0.374	0.353	0.355	0.355	0.395
R_Square/Within R_Square	0.368	0.373	0.371	0.391	0.368	0.369	0.369	0.411
F-Statistic	30.21	31.14	31.13	30.01	16.19	16.14	16.28	16.46
No. of Observations	526	526	526	451	526	526	526	451
No. of Clusters	.	.	.	.	227	227	227	223
Dependent Variable: Logged Split Ethnic Land's Average Access Roads Length 1980 - 2010 (NASA SEDAC data repository)								
VARIABLES	(H) y2	(I) y2	(J) y2	(K) y2	(L) y2	(M) y2	(N) y2	(O) y2
WBGI Corruption Control Index (average 1990s - 2007)	0.0702 (0.0657)				0.0627 (0.0733)			
WBGI Rule of Law Index (average 1990s - 2007)		0.205*** (0.0618)				0.129** (0.0571)		
WBGI Governance Effectiveness Index (average 1990s - 2007)			0.156** (0.0687)				0.0985 (0.0633)	
Quality of Governance Index (ICRG QOG - average 1990s - 2007)				0.559 (0.348)				0.552* (0.286)
Population Density (at 2000 in 1000s of Peoples)	-0.227*** (0.0219)	-0.225*** (0.0214)	-0.228*** (0.0216)	-0.217*** (0.0229)	-0.160*** (0.0311)	-0.160*** (0.0315)	-0.159*** (0.0311)	-0.151*** (0.0325)
Adjusted R_Square	0.355	0.370	0.362	0.377	0.231	0.244	0.236	0.276
R_Square/Within R_Square	0.370	0.385	0.377	0.394	0.249	0.262	0.254	0.296
F-Statistic	20.00	21.44	20.75	19.09	7.279	7.399	7.365	6.901
No. of Observations	519	519	519	445	519	519	519	445
No. of Clusters	.	.	.	.	227	227	227	223
OLS Specification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tribal Fixed Effects Specification	No	No	No	No	Yes	Yes	Yes	Yes

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

Controls: Population density, distance to sea, distance to border, distance to associated national capital, mean elevation, soil productivity, malaria succeptibility index, land area, diamond mines indicator, petroleum sites indicator and water bodies indicator

Table L – 2.4.1: Results from the Base Model with an OLS Specification

2.4.2 – National Institutional Impact on Subnational Economic Activity (using an IV specification)

In Table M - 2.4.2 below, we display results adopting an instrumental variable (IV) approach harnessing the divergent historical experiences of each halves of demarcated ethnic homelands, which form the analytical unit of this research design. These distinct historical experiences are more proximate when historically compared to the tribal values and institutions of the characteristic halves. They are also composed of both national and subnational experiential components (mechanisms). The instruments as chosen reflect these equally appropriate prongs, which are believed to have affected the institutional evolution of each demarcated halves of formerly unique ethnic identities or homelands.

As can be observed from the results below, the introduction of this IV-design certainly provides an interesting set of results to the debate in this literature. We seem to find results consistent with both the cross-country and subnational level literatures depending on what measure of economic development adopted. In these results we find that after controlling for the exogenous historical experiences of each ethnic split, there appears to be a pattern consistent with national institutional relevance. This is however specific to the measure of the level of economic activity adopted. The exogenous historical experiences of each ethnic split are sourced from both national and subnational premises relevant to the evolution and sustenance of modern institutional manifestations at the national administrative levels.

These results in Table M - 2.4.2 show that national institutional quality measures are not statistically significant after controlling for ethnic fixed-effects across country at these subnational-bordering regions. This still corroborates the findings from the base OLS estimations above and also the findings from M&P (2014). Interestingly, our initial estimations in estimations 1-4 show that even before introducing the tribal fixed-effects statistical significances are much less than those observed in their A-D and H-K OLS counterparts above. The results differ very much from those in table 4.3 below however. We find that when the dependent variable is altered to the average access roads length variable, the national institutional quality measures stay statistically significant even after controlling for these ethnic level fixed-effects. These are in contrast with M&P's

(2014) findings but are very much in resonance with the cross-country studies such as AJR (2001) and their cohorts.

Our results also show that not only are the magnitudes of the coefficients (depicting the impact of national institutional quality on economic activity at the subnational split ethnic level) increasing when the IV method is adopted, but generally become more statistically significant compared with the OLS estimations above. This finding is maintained even in the presence of unobservable tribal fundamentals. This suggests that an unidentified or unfiltered institutional measure at the national level may appear to be irrelevant to the variations in the level of subnational economic activity at these peripheral regions. It may also suggest a much deeper methodological concern about the sort of endogeneity inherent in the relationship between national institutional dynamics and subnational economic development as earlier discussed in section 2.3.2 above. It could be an empirical confirmation that the issue of measurement error embedded in the endogeneity problem may be driving the bias characterising the relationship between both these variables of interest compared to that of reverse causality.

When we change the measure of economic development to the average access roads length variable however, our findings flip in statistical significance and start to resemble those found in the country-level empirical studies. These findings are supported by the history of the African region and of the literature on the impact of national institutions on peripheral subnational regions. In this literature, it is often hypothesized that national institutions are much more felt in the national capitals and those subnational regions fringing these capitals (see Herbst, 2000; Boone, 2003; etc). With these results using our history channelled IV procedure, we are able to rediscover a place for the relevance of colonial and countrywide institutional outcomes on economic activities even at the geopolitically distant hinterlands. This stands in contradiction with the literature as propagated in the discussions of Herbst (2000), Bates (2008), and revalidated by the results of M&P (2014), for instance. In our findings as shown in Table N - 2.4.3 below, national institutions appear relevant after you control for the endogeneity embedded in the relationship it has with subnational economic activity. This control we show ought to sensibly reflect the historical background of the subnational economic units spanning the subnational ethnic homelands as engaged in this study. Interestingly, the key findings contained in Table N - 2.4.3 are robust to all the

widely used geographic, ecological and demographic controls often adopted in similar research designs in the literature.

Holistically, the results buttress the position that areas that had closer proximity to missionary activities as measured using the distance to missionary stations were not necessarily associated with higher national institutional values as expected. As observed in our first-stage results below each of the results from Table M - 2.4.2 and Table N - 2.4.3 respectively, we find that distance to the missionary stations is positive and mostly statistically significantly related to the respective measures of national institutional quality. This indicates that areas farther away from these stations were on the average associated with national institutional values. This is also very much in line with the findings involving the measure of exploitative European contact using the presence of a European exploratory route for any associated country. The coefficient associated with this instrument shows that national level exposures to the exploitative activities of the Europeans are on the average negatively and statistically associated with national institutional quality. We can find the contribution of these channels by looking at the selected output from the First-stage regressions in the bottom section of each table in Table M - 2.4.2 and Table N - 2.4.3 below.

On the average, we find that national institutional quality positively impacts subnational economic activity at the borders. Specifically, a unit increase in national institutional quality results in between 26% to about 200% increase in the nighttime light density levels compared to approximately 6% to 130% increase in the average roads access length, depending on the measure of national institutional quality and the estimation technique adopted. These results are on the average robust to the inclusion of unobservable ethnic level specific characteristics, which may certainly be of considerable relevance for bordering subnational regions. These regions may not readily be accessible by the national machinery of most countries' institutional enforcement and monitoring programmes given the predominant sparseness of border identity for these peripheral regions in Africa. We observe on the other hand that demarcated areas associated with a nation-state with a higher sense of nationalistic identity are statistically and positively associated with higher modern national institutional quality values. This is not surprising as such areas would have begun existing as multi-ethnic communities before the advent of the Europeans harnessing and

adopting a sense of identities beyond the individual ethnic nations that composed such super-states. We are able to deduce this using the state history index developed by Bockstette, Chanda & Putterman (2002).

We equally find that our instruments after controlling for tribal fixed effects are robust to key statistical tests on the exclusion restriction condition as shown by the '*Jansens*' test of over-identifying restrictions. This provides further statistical validation for the exclusion restriction storyline provided in section 2.3.3 above.

Dependent Variable: Logged Split Ethnic Land's Average Stable Nighttime Lights Density at 2010-2011 (NOAA data repository)								
VARIABLES	(1) y1	(2) y1	(3) y1	(4) y1	(5) y1	(6) y1	(7) y1	(8) y1
WBGI Corruption Control Index (average 1990s - 2010)	1.583**				0.881			
	(0.740)				(0.714)			
WBGI Rule of Law Index (average 1990s - 2010)		0.902*				0.639		
		(0.476)				(0.482)		
WBGI Governance Effectiveness Index (average 1990s - 2010)			0.933**				0.595	
			(0.429)				(0.454)	
Quality of Governance Index (ICRG QOG - average 1990s - 2010)				3.104*				1.987
				(1.633)				(1.698)
Population Density (at 2000 in 1000s of Peoples)	0.675***	0.653***	0.644***	0.622***	0.264*	0.251*	0.261*	0.310***
	(0.0797)	(0.0760)	(0.0764)	(0.0759)	(0.136)	(0.131)	(0.136)	(0.110)
IV Specification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tribal Fixed Effects Specification	No	No	No	No	Yes	Yes	Yes	Yes
Centred R_Square/Within R_Square	0.362	0.387	0.383	0.406	0.406	0.408	0.412	0.450
Adjusted R_Square	0.345	0.371	0.366	0.388	-0.126	-0.122	-0.114	-0.0719
F-Statistic	24.98	26.73	26.71	27.71	13.08	13.40	13.77	14.65
No. of Observations	456	456	456	409	415	415	415	345
No. of Clusters	.	.	.	.	.	.	.	.
Degree of Freedom	12	12	12	12	196	196	196	168
Kleibergen-Paaps Weak Identifcation Test Wald Statistic	26.25	49.90	53.10	145.9	17.94	28.07	20.60	41.39
Jansens Test P-value For Overidentifying Restrictions	0.0548	0.0435	0.0465	0.0200	0.281	0.309	0.315	0.164
Kleibergen-Paaps Underidentification Test LM Statistic	69.38	75.34	91.81	71.82	61.09	72.30	70.77	50.32
Kleibergen-Paaps Underidentification Test LM P-value	0	0	0	0	8.24e-10	0	0	9.37e-08
The C-Statistic (for - Distance to the nearest Missionary station)	8.739	3.459	5.128	0.00731	11.42	5.778	14.46	9.581
P-value of C-Statistic	0.00311	0.0629	0.0235	0.932	0.000726	0.0162	0.000143	0.00197
Anderson-Rubin F-test of significance of endogenous regressors	0.00240	0.00240	0.00240	0.00147	0.115	0.115	0.115	0.0569
P-value of Anderson-Rubin F-test of endogenous regressors	0.00442	0.00442	0.00442	0.00308	0.171	0.171	0.171	0.107
No. of Tribal Clusters					184	184	184	156
First-stage Results for the IV Regressions Above								
Nearest Split Ethnic Centroid Distance to Confirmed Missionary Locations (1,000s KM)	0.3859***	0.2732*	0.2894**	0.0022	1.1993***	0.9688**	1.4879***	0.3421***
	'(0.1386)	'(0.1531)	'(0.1341)	'(0.0270)	'(0.3595)	'(0.4096)	'(0.4146)	'(0.1054)
Indicator for the presence of a European Explorer route in associated country	-0.1146**	-0.2321***	-0.1733**	0.0953***	-0.2977***	-0.5034***	-0.4317***	0.0662*
	'(0.0583)	'(0.0699)	'(0.0828)	'(0.0182)	'(0.0749)	'(0.0922)	'(0.1081)	'(0.0353)
State History Index (Putterman, 2012 - discounted at 10%)	0.2473**	0.2229	0.4022***	-0.0507*	0.8351***	0.8647***	1.0263***	0.0289
	'(0.1190)	'(0.1387)	'(0.1283)	'(0.0273)	'(0.1729)	'(0.2059)	'(0.2038)	'(0.0449)
Adjusted R_Square	0.25	0.333	0.394	0.643	-0.345	-0.152	-0.126	0.296
Centred R_Square/Within R_Square	0.282625	0.361883	0.420171	0.660614	0.316113	0.414282	0.427504	0.655187
F-Statistic	23.61657	43.48632	71.68912	131.7966	11.87181	19.85018	12.2856	37.16442
F-Statistic showing joint strength of Instrumental variables	26.25456	49.89797	53.10294	145.9253	17.94483	28.07402	20.5999	41.39437
No. of Observations	456	456	456	409	415	415	415	345
Tribal Fixed Effects Specification	No	No	No	No	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Full-set of Controls: Population density, distance to sea, distance to border, distance to associated national capital, mean elevation, soil productivity, malaria succeptibility index, land area, diamond mines indicator, petroleum sites indicator and water bodies indicator

Table M – 2.4.2: Results from the Nighttime Lights Density Measure Using Our IV-identification Strategy

Dependent Variable: Logged Split Ethnic Land's Average Access Roads Length 1980 - 2010 (NASA SEDAC data repository)								
VARIABLES	(9) y2	(10) y2	(11) y2	(12) y2	(13) y2	(14) y2	(15) y2	(16) y2
WBGI Corruption Control Index (average 1990s - 2010)	1.225***				0.609***			
	(0.230)				(0.206)			
WBGI Rule of Law Index (average 1990s - 2010)		0.779***				0.387***		
		(0.134)				(0.132)		
WBGI Governance Effectiveness Index (average 1990s - 2010)			0.756***				0.398***	
			(0.132)				(0.116)	
Quality of Governance Index (ICRG QOG - average 1990s - 2010)				2.412***				1.340***
				(0.604)				(0.421)
Population Density (at 2000 in 1000s of Peoples)	-0.186***	-0.199***	-0.208***	-0.199***	-0.145***	-0.155***	-0.147***	-0.164***
	(0.0265)	(0.0233)	(0.0230)	(0.0239)	(0.0336)	(0.0338)	(0.0318)	(0.0324)
IV Specification	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tribal Fixed Effects Specification	No	No	No	No	Yes	Yes	Yes	Yes
Centred R_Square/Within R_Square	0.0192	0.271	0.247	0.334	0.115	0.207	0.171	0.298
Adjusted R_Square	-0.00784	0.251	0.226	0.314	-0.688	-0.511	-0.580	-0.379
F-Statistic	16.59	20.76	19.83	18.45	5.045	5.153	5.882	6.071
No. of Observations	449	449	449	403	406	406	406	338
No. of Clusters	.	.	.	.	.	.	.	.
Degree of Freedom	12	12	12	12	193	193	193	166
Kleibergen-Paaps Weak Identifcation Test Wald Statistic	26.22	49.78	53.01	145.7	16.89	26.97	20.32	40.20
Jansens Test P-value For Overidentifying Restrictions	0.0520	0.00942	0.0117	6.71e-05	0.583	0.400	0.680	0.204
Kleibergen-Paaps Underidentification Test LM Statistic	68.22	74.23	91.87	71.64	57.31	69.10	70.45	49.33
Kleibergen-Paaps Underidentification Test LM P-value	0	0	0	0	4.41e-09	0	0	1.44e-07
The C-Statistic (for - Distance to the nearest Missionary station)	10.02	3.800	5.706	0.0266	10.98	5.747	14.59	10.09
P-value of C-Statistic	0.00155	0.0512	0.0169	0.871	0.000919	0.0165	0.000134	0.00149
Anderson-Rubin F-test of significance of endogenous regressors	0	0	0	0	0.00131	0.00131	0.00131	0.000347
P-value of Anderson-Rubin F-test of endogenous regressors	0	0	0	0	0.00470	0.00470	0.00470	0.00224
No. of Tribal Clusters					181	181	181	154
First-stage Results for the IV Regressions Above								
Nearest Split Ethnic Centroid Distance to Confirmed Missionary Locations (1,000s KM)	0.4044***	0.2841*	0.3020**	0.0043	1.1682***	0.9639**	1.4930***	0.3552***
	'(0.1362)	'(0.1521)	'(0.1331)	'(0.0269)	'(0.3540)	'(0.4086)	'(0.4122)	'(0.1040)
Indicator for the presence of a European Explorer route in associated country	-0.1116*	-0.2310***	-0.1795**	0.0936***	-0.3086***	-0.5254***	-0.4535***	0.0637*
	'(0.0588)	'(0.0701)	'(0.0824)	'(0.0185)	'(0.0761)	'(0.0929)	'(0.1071)	'(0.0362)
State History Index (Putterman, 2012 - discounted at 10%)	0.2353*	0.2193	0.4204***	-0.0486*	0.8373***	0.8905***	1.0717***	0.0338
	'(0.1201)	'(0.1406)	'(0.1299)	'(0.0279)	'(0.1743)	'(0.2077)	'(0.1992)	'(0.0462)
Adjusted R_Square	0.253	0.335	0.397	0.645	-0.342	-0.143	-0.108	0.295
Centred R_Square/Within R_Square	0.28664	0.364341	0.423724	0.662269	0.322321	0.42271	0.440602	0.6577
F-Statistic	23.79585	44.27269	73.10094	132.4316	12.08857	20.14894	11.67797	35.40559
F-Statistic showing joint strength of Instrumental variables	26.21707	49.78453	53.01406	145.735	16.88728	26.97096	20.3214	40.1971
No. of Observations	449	449	449	403	406	406	406	338
Tribal Fixed Effects Specification	No	No	No	No	Yes	Yes	Yes	Yes

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

Full-set of Controls: Population density, distance to sea, distance to border, distance to associated national capital, mean elevation, soil productivity, malaria succeptibility index, land area, diamond mines indicator, petroleum sites indicator and water bodies indicator

Table N – 2.4.3: Results from the Average Access Roads Length Measure Using Our IV-identification Strategy

2.4.3 – Brief summary of findings

In summary, we find that contrary to an un-instrumented model there is considerable evidence to indicate that the impacts of European-filtered national institutional quality is statistically significant for explaining the variations of economic activity in the subnational ethnic homelands of Africa. This is especially so when we provide an alternative measure of economic development using the average access roads length. This finding appears robust to alternative measures of national institutional quality. Although we do not find the same outcome using the average nighttime lights density variable as adopted in M&P (2014), our

findings using this measure corroborates theirs, which uses an alternative identification strategy and empirical methodology.

We find in general from the first-stage regressions that, pre-colonial European contact, on the average, appears to have had a debilitating effect on institutional quality regardless of whether the type of activity encouraged exploitation through commerce or whether it was subtler by adopting the machinery of religion. The latter of these ideas was not an expected outcome but it is plausible. The locals at the time could easily have perceived either form of European infiltrations as invasive and negligent of their inherent pre-existing traditional values. As a result they would have nurtured a sense of disdain very early on and responded negatively towards any form of European influence including the sense of a national identity and the institutional dynamics that would have followed after. We see that a national historical feature, such as the national administrative identity of the country is equally very important and provides a positive explanation of current average institutional quality. Indicating that a nation state's history may be very relevant, as peripheral states would definitely cooperate in the nation state's interest the longer they have been part of such a national historical identity.

By introducing an alternative measure of economic development to the empirical discourse we are able to probe deeper the effect national institutional outcomes have on urban development from multiple dimensions and perspectives. This average access roads length variable is a measure that captures economic activity quite differently from the nighttime lights density variable. Particularly because it does not change that much from year to year but certainly does reflect the potential and level of development and potential for urbanisation for a particular region. Since both measures capture some form of infrastructural development, it is quite interesting to compare the results of national institutional development on their cross-sectional variations across our African subnational demarcated ethnic homelands sample. This finding could also be indicative of the analogy that roads is more of a public infrastructure compared to local or domestic illumination efforts which are more of a private endeavour with the exception of street lighting. As a result, our nighttime lights index is less likely to be correlated with national institutional quality compared to the average roads length index.

2.4.4 - Key Weaknesses of Study

This study has a number of weaknesses; our enumeration as listed below is by no means exclusive or conclusive of itself.

Firstly, the research would have benefited from a much more subnational mono-levelled analysis. Given that the institutional measure is observed and measured at the national level, it does not allow the model to control for other non-measurable country level phenomena such as internal coordination, current international influences, etc., which would have easily been captured using country level fixed-effects. Although a lot has been captured or soaked up using ethnic fixed-effects, including country fixed-effects would have definitely helped to further eliminate any suspicions of an omitted variable issue.

Secondly, the multidimensionality in the data structure created by our key variables being observed at both the country-level and the split ethnic homelands subnational-level inclines our model framework to additional residual noise. Although we have adopted an IV methodology, which could mitigate this issue, a less noisier model framework would have been possible if there was a clear signal mechanism or process in the modern regional geopolitical system that more strongly linked institutional quality at the national level to economic activity at the subnational level apart from the historical signal mechanism adopted herein. For instance, it would have been more useful if there were a national institutional variable measured at the subnational state or intra-regional administrative (provincial) level. This would have provided additional justification or robustness to the results obtained in this study as the link between subnational administrative institutional apparatuses and economic activities at the ethnic homeland is more obvious than claims of direct channels with their respective central governments.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401.

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

Acemoglu, D., Johnson, S., & Robinson, J. A. (2012). The Colonial Origins of Comparative Development: An Empirical Investigation: Reply. *American Economic Review*, *102*(6), 3077–3110. doi:10.1257/aer.102.6.3077

Acemoglu, D., & Robinson, J. A. (2004). Kleptocracy and Divide-and-Rule: A Model of Personal Rule. *Journal of the European Economic Association*, *2*(2-3), 162–192.

Acemoglu, D., & Robinson, J. A. (2006). *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Asiwaju, A. I. (1985). *Partitioned Africans: Ethnic Relations across Africa's international Boundaries, 1884-1984*. C. Hurst & Co. Publishers.

Austin, G. (2010). *African Economic Developemnt and Colonial Legacies*. *International Development Policy - Revue internationale de politique de développement* (Vol. 1).

Banerjee, A., Duflo, E., & Qian, N. (2012). *On the Road: Access to Transportation Infrastructure and Economic Growth in China* (No. 17897). *NBER Working Paper Series*. doi:10.2139/ssrn.2018637

Banerjee, A., & Iyer, L. (2005). History, Institutions and Economic Performance: The Legacy of Land Tenure Systems in India. *The American Economic Review*, *95*(4), 1190–1213. doi:10.2307/4132711

Bates, R. H. (2008). *When Things Fell Apart: State Failure in Late-Century Africa*. Cambridge University Press. doi:10.1017/CBO9781107415324.004

Bates, R. H., & Collier, P. (1995). The Politics and Economics of Policy Reform in Zambia. *Journal of African Economies*, *4*(1), 115–143.

Beuran, M., Gachassin, M., & Raballand, G. (2015). Are There Myths on Road Impact and Transport in Sub-Saharan Africa? *Development Policy Review*, *33*(5), 673–700.

Bhattacharyya, S. C. (2012). Electrification Experiences from Sub-Saharan Africa. In *Rural Electrification through Decentralised Off-grid Systems in Developing Countries* (pp. 131–156). Springer London.

Bloom, D. E., Sachs, J. D., Collier, P., & Udry, C. (1998). Geography, Demography, and Economic Growth in Africa. *Brookings Papers on Economic Activity*, *1888*(2), 207–295. doi:10.2307/2534695

Bockstette, V., Chanda, A., & Putterman, L. (2002). States and Markets: The Advantages of an Early Start. *Journal of Economic Growth*, *7*, 347–369.

Boone, C. (2003). *Political Topographies of the African State: Territorial Authorities and Institutional Choice* (Vol. 1). Cambridge University Press. doi:10.1017/CBO9781107415324.004

Chandra, A., & Thompson, E. (2000). Does public infrastructure affect economic activity? *Regional Science and Urban Economics*, *30*(4), 457–490. http://doi.org/10.1016/S0166-0462(00)00040-5

Collier, P. (2007). *The Bottom Billion: Why the Poorest Countires are Failing and What Can Be Done About It*. Oxford University Press.

Cook, N., Campbell, R. J., Brown, P., & Ratner, M. (2015). *Powering Africa: Challenges of and U.S. Aid for Electrification in Africa. Congressional Research Service (CRS) Report*.

Duranton, G., Morrow, P. M., & Turner, M. A. (2013). Roads and trade: Evidence from the US. *Review of Economic Studies*, *81*(2), 681–724. doi:10.1093/restud/rdt039

Easterly, W. (2001). *The Elusive Quest for Growth: Economists' Adventure and Misadventure in the Tropics*. MIT Press.

Easterly, W., & Levine, R. (2003). Tropics, Germs, and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics*, *50*(1), 3–39. doi:10.1016/S0304-3932(02)00200-3

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies*, *81*(3), 1046–1070. doi:10.1093/restud/rdu010

Gennaioli, N., & Rainer, I. (2006). Precolonial Centralization and Institutional Quality in Africa. In M. Gradstein & K. A. Konrad (Eds.), *Institutions and Norms in Economic Development* (pp. 21–46). MIT Press.

Gennaioli, N., & Rainer, I. (2007). The Modern Impact of Precolonial Centralization in Africa. *Journal of Economic Growth*, *12*(June), 185–234. doi:10.1007/s10887-007-9017-z

Glaeser, E. L., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do Institutions Cause Growth? *Journal of Economic Growth*, *9*(3), 271–303. doi:10.1023/B:JOEG.0000038933.16398.ed

Glaeser, E. L., & Shleifer, A. (2002). Legal Origins. *The Quarterly Journal of Economics*, *117*(4), 1193–1229. doi:10.2307/4132477

Ghosh, T., Anderson, S. J., Elvidge, C. D., & Sutton, P. C. (2013). Using Nighttime Satellite Imagery as a Proxy Measure of Human Well-being. *Sustainability (Switzerland)*, *5*(12), 4988–5019.

Henderson, V. J., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, *102*(2), 994–1028. doi:10.1257/aer.102.2.994

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

Herbst, J. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control* (1st Editio.). New Jersey: Princeton University Press.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2009). *Governance Matters VIII Aggregate and Individual Governance Indicators* (No. 4978). *Policy Research Working Paper* (Vol. 21). doi:10.1080/713701075

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and Analytical Issues. *Hague Journal on the Rule of Law*, *3*(02), 220–246. doi:10.1017/S1876404511200046

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., & Sachs, J. (2004). A Global Index Representing the Stability of Malaria Transmission. *American Journal of Tropical Medicine and Hygiene*, *70*(5), 486–498. doi:70/5/486 [pii]

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2008). The Economic Consequences of Legal Origins. *Journal of Economic Literature*, *46*(2), 285–332. Retrieved from http://www.nber.org/papers/w13608

Manning, P. (1983). Contours of Slavery and Social Change in Africa. *The American Historical Review*, *88*(4), 835–857. doi:10.2307/1874022

Mehlum, H., Moene, K., & Torvik, R. (2006). Institutions and the Resource Curse. *The Economic Journal*, *116*(508), 1–20. doi:10.1111/j.1468-0297.2006.01045.x

Mellander, C., Lobo, J., Stolarick, K., & Matheson, Z. (2015). Night-time light data: A good proxy measure for economic activity? *PLoS ONE*, *10*(10), 1–18.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, *81*(1), 113–152. doi:10.3982/ECTA9613

Michalopoulos, S., & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, *129*(1), 151–213. doi:10.1093/qje/qjt029.Advance

Nunn, N. (2008). The Long-term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics*, *123*(1), 139–176.

Nunn, N. (2009). The Importance of History for Economic Development. *Annual Review of Economics*, *1*(1), 65–92. doi:10.1146/annurev.economics.050708.143336

Nunn, N. (2010). Religious Conversion in Colonial Africa. *American Economic Review*, *100*(2), 147–152. doi:10.1257/aer.100.2.147

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

Nunn, N., & Wantchekon, L. (2011). The Slave trade and the Origins of Mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252. doi:10.1257/aer.101.7.3221

Rodney, W. (1973). *How Europe Underdeveloped Africa*. Dar-Es-Salaam: Bogle-L'Ouverture Publications, London and Tanzania Publishing House. doi:10.2307/217137

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth*, *9*(2), 131–165. doi:10.1023/B:JOEG.0000031425.72248.85

Roland, G. (2004). Fast-Moving and Slow-Moving Institutions. CESifo DICE Report: 2/2004.

Sachs, J. D. (2003). *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income* (No. 9490 - NBER Working Papers Series).

Sachs, J. D., & Warner, A. M. (1995). *Natural Resource Abundance and Economic Growth* (No. 5398 - NBER Working Paper Series).

Sachs, J. D., & Warner, A. M. (1997). Sources of Slow Growth in African Economies. *Journal of African Economies*, *6*(3), 335–376.

Sarmidi, T., Law, S. H., & Jafari, Y. (2014). Resource Curse : New Evidence on the Role of Institutions. *International Economic Review*, *28*(1), 191–206. doi:10.1080/10168737.2013.787110

Storeygard, A. (2016). Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *Review of Economic Studies*, *83*, 1263–1295. doi:10.1093/restud/rdw020

Táíwò, O. (2010). *How Colonialism Preempted Modernity in Africa*. Indiana University Press. doi:10.1080/03086534.2011.598039

Whatley, W., & Gillezeau, R. (2009). *The Impact of the Slave Trade on African Economies*. *World Economic History Congress*. Utrecht. doi:10.1300/J265v03n01_02

White, B. W. (1996). Talk about School: Education and the Colonial Project in French and British Africa, (1860-1960). *Comparative Education*, *32*(1), 9–25. doi:10.1080/03050069628902

Wooldridge, J. M. (2008). Introductory Econometrics: A Mondern Approach. Journal of contaminant hydrology (4th ed.). South Western College.

Ziltener, P., & Künzler, D. (2013). Impacts of Colonialism – A Research Survey. *American Sociological Association*, *19*(2), 290–311.

Appendix

Appendix 2.1 – Definition of Regression Variables

Variable	Units of Measure	Variable(s) Used from Original Dataset	Source
Log(0.001 + Average Nighttime lights density index 2010-2011)	Polygonal-average of Index between 0 and 63	GRID values	Version 4 DMSP-OLS Nighttime Lights Time Series: http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
Log(Average Roads Length over the 1980-2010 period)	Polygonal-average in 1000 Kilometres	LENGTH_KM	Global Roads Open Access Data Set (gROADS), v1 (1980–2010): http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1
Country's WBGI Corruption Control	Country-level Index between -2.5 to +2.5	N/A	World Bank Governance Indicators (WBGI) by Kaufmann, Kraay, & Mastruzzi (KKM, 2009, 2011): http://info.worldbank.org/governance/wgi/index.aspx#home
Country's WBGI Rule of Law Index	Country-level Index between -2.5 to +2.5	N/A	World Bank Governance Indicators (WBGI) by Kaufmann, Kraay, & Mastruzzi (KKM, 2009, 2011): http://info.worldbank.org/governance/wgi/index.aspx#home
Country's WBGI Governance Effectiveness Index	Country-level Index between -2.5 to +2.5	N/A	World Bank Governance Indicators (WBGI) by Kaufmann, Kraay, & Mastruzzi (KKM, 2009, 2011): http://info.worldbank.org/governance/wgi/index.aspx#home
Country's Quality of Governance Index	Country-level Index between 0 and 1	icrg_qog	Quality of Governance (QoG) Basic Jan 2015: http://qog.pol.gu.se/data/datadownloads/data-archive
Nearest Split Ethnic Centroid Distance to Confirmed Missionary Locations (1,000s KM)	Polygonal-average in 1000 Kilometres	NAME - representing the individual missionary stations	Nunn & Wantchekon (2011) based on work originally from Roland A. Oliver (2000) - available in his replication-data section: http://scholar.harvard.edu/files/nunn/files/roome_map_1924.zip?m=1364314848
Indicator for the presence of a historical explorer route passing through country's border	Country-level Dummy variable	Explorer_N - representing names of the individual routes	Nunn (2010) based on a map published by William R. M. Roome (1924) - available in his replication-data section: https://dl.dropboxusercontent.com/u/3610301/Colonial_railways.zip
State History Index (Bockstette, Chanda & Putterman, 2002 - discounted at 10%)	Country-level Index between 0 and 1	statehistn10v3	Version 3.1 based on work by Bockstette, Chanda & Putterman, 2002 : available at http://www.brown.edu/Departments/Economics/Faculty/Louis_Putterman/antiquity%20index.htm
Never Colonised Indicator	Country-level Dummy variable	ht_colonial	Quality of Governance (QoG) Basic Jan 2015: http://qog.pol.gu.se/data/datadownloads/data-archive
British Colony Indicator	Country-level Dummy variable	ht_colonial	Quality of Governance (QoG) Basic Jan 2015: http://qog.pol.gu.se/data/datadownloads/data-archive
French Colony Indicator	Country-level Dummy variable	ht_colonial	Quality of Governance (QoG) Basic Jan 2015: http://qog.pol.gu.se/data/datadownloads/data-archive
French Legal Origin Indicator	Country-level Dummy variable	lp_legor	Quality of Governance (QoG) Basic Jan 2015: http://qog.pol.gu.se/data/datadownloads/data-archive
Log (1 + Population Density at 2000)	Polygonal-average: Discrete Count Variable	grid_code	UNEP/GRID - Sioux Falls Dataset (Africa Population Distribution Database): http://na.unep.net/siouxfalls/datasets/datalist.php
Distance to Sea	Nearest distance for each Polygonal-unit in 1000 Kilometres	From a Geodatabase of Polylines	Water Polygons: Available at the Open Street Map http://openstreetmapdata.com/data/water-polygons

Distance to Country Border	Nearest distance for each Polygonal-unit in 1000 Kilometres	From a Shapefile with National Boundary Outlines	National Administrative Boundaries, v3 (2000), Gridded Population of the World (GPW), v3: http://sedac.ciesin.columbia.edu/data/set/gpw-v3-national-admin-boundaries/data-download
Distance to National Capital	Centroid distance for each Polygonal-unit in 1000 Kilometres	From a Shapefile of National Capital centroid points	ArcGIS Data Gallery - World Cities: http://www.arcgis.com/home/item.html?id=dfab3b294ab24961899b2a98e9e8cd3d
Mean Elevation	Polygonal-averages but used as the log(kilometres)	From Data provided in ArcView ESRI Grid Format	National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colo.: http://nelson.wisc.edu/sage/data-and-models/atlas/maps.php?datasetid=28&includerelatedlinks=1&dataset=28
Soil Productivity Index	Polygonal-average of Index between 1 and 97	GRID values representing Soil Quality	TERRASTAT I, FAO Soil Productivity Index
Malaria Ecology Index	Polygonal-average of Index between a lowest value of 0 and highest value of 38.081	GRID values representing Incidence rates	Malaria Ecology Map based on work by Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004): available at http://www.earth.columbia.edu/people/gmccord/sitefiles/file/malaria_ecology.zip
Log Split Land Area	Polygonal-average in 1000 Kilometres	N/A	Obtained from the Geographic intersection of the Country-level administrative boundaries - obtained from the Diva GIS website http://biogeo.ucdavis.edu/data/world/countries_shp.zip; and the Tribal Homeland Boundaries based on the work of Peter Murdock (1959) - available at http://scholar.harvard.edu/files/nunn/files/murdock_shapefile.zip?m=1364315353
Diamond Mine Present Indicator	Polygonal Dummy variable	NAME - representing the individual sites	Diamond Resources from the Peace Research Institute Oslo (PRIO) website: https://www.prio.org/Data/Geographical-and-Resource-Datasets/Diamond-Resources/
Petroleum Well Present Indicator (On-shore)	Polygonal Dummy variable	NAME - representing the individual sites	Petroleum Dataset from the Peace Research Institute Oslo (PRIO) website: https://www.prio.org/Data/Geographical-and-Resource-Datasets/Petroleum-Dataset/
Indicator for the presence of Inland Water Body (KM)	Polygonal Dummy variable	N/A	Global Lakes and Wetlands Database: Lakes and Wetlands Grid (Level 3): http://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3

Table O – Appendix-Table 2.1: Variable Definition of Regression Analyses Variables

Appendix 2.2 – List of Countries Sampled in the Regressions

Sample of Countries Involved in this Chapter's Analyses	
Algeria	Libya
Angola	Malawi
Benin	Mali
Botswana	Mauritania
Burkina Faso	Morocco
Burundi	Mozambique
Cameroon	Namibia
Central African Republic	Niger
Chad	Nigeria
Cote d'Ivoire	Republic of Congo

Democratic Republic of the Congo	Rwanda
Djibouti	Senegal
Egypt	Sierra Leone
Equatorial Guinea	Somalia
Eritrea	South Africa
Ethiopia	Sudan
Gabon	Swaziland
Gambia	Tanzania
Ghana	Togo
Guinea	Tunisia
Guinea-Bissau	Uganda
Kenya	Zambia
Lesotho	Zimbabwe
Liberia	

Table P – Appendix-Table 2.2: List of Countries Used in the Regression Analyses

Appendix 2.3 – List of the 227 Tribal Homelands Sampled in the Regressions

List of the 227 Tribal Groups Used in the Regression Analyses					
ABABDA	BUEM	HIECHWARE	LIGBI, DEGHA (SE)	NSUNGLI	TIENGA
ADELE	BULOM	HLENGWE	LOBI	NUKWE	TIGON
AFAR	BUSA	HOLO	LUGBARA	NUSAN	TIGRINYA
ALUR	BWAKA	IBIBIO	LUNGU	NYAKYUSA	TLOKWA
AMBA	CHAGA	IFORA	LUVALE	NYANGIYA	TOMA
AMBO	CHAKOSSI	IMRAGEN	MADI	NYANJA	TONGA
AMER	CHEWA	ISHAAK	MAKONDE	NYASA	TRIBU
ANA	CHIGA	IWA	MALINKE	NZANKARA	TRIPOLITANIANS
ANUAK	CHOKWE	JERID	MAMBILA	PANDE	TUBURI
ANYI	DAGARI	JIE	MANDARA	POPO	TUKULOR
ASBEN	DARI	KABRE	MANGA	PUKU	TUMBUKA
ASSINI	DAZA	KANEMBU	MANYIKA	RESHIAT	TUNISIANS
ATTA	DENDI	KAONDE	MASAI	RONGA	UDALAN
ATYUTI	DIALONKE	KAPSIKI	MASALIT	RUANDA	VAI
AULLIMINDEN	DIDINGA	KARA	MASHI	RUNDI	VENDA
AUSHI	DIGO	KARAMOJONG	MASINA	RUNGA	VILI
AVATIME	DIOLA	KARE	MATAKAM	SABEI	WAKURA
AZANDE	DUMA	KGATLA	MBERE	SAHO	WANGA
AZJER	DZEM	KISSI	MBUKUSHU	SAMO	WUM
BABUKUR	EGBA	KOBA	MBUNDA	SANGA	YAKA
BAJUN	EKOI	KOMA	MENDE	SEKE	YAKOMA
BALANTE	ESA	KOMONO	MINIANKA	SHAMBALA	YALUNKA
BANYUN	EWE	KONGO	MOMBERA	SHEBELLE	YAO
BANZIRI	FANG	KONJO	MPEZENI	SHUWA	YOMBE

BARABRA	FON	KONKOMBA	MUNDANG	SONGHAI	ZAGHAWA
BARARETTA	FOUTADJALON	KONO	MUNDU	SONINKE	ZEKARA
BARGU	FUNGON	KONYANKE	MUSGU	SOTHO	ZIMBA
BASHI	GADAMES	KORANKO	NAFANA	SUBIA	
BATA	GIL	KOTA	NALU	SUNDI	
BAYA	GOMANI	KOTOKO	NAMA	SURI	
BERABISH	GREBO	KPELLE	NAUDEBA	SWAZI	
BERTA	GRUNSHI	KRAN	NDAU	TABWA	
BIDEYAT	GUDE	KREISH	NDEMBU	TAJAKANT	
BIRIFON	GULA	KUNDA	NDOGO	TAMA	
BOBO	GUN	KUNG	NDUKA	TAWARA	
BOKI	GURENSI	KUNTA	NGAMA	TEDA	
BONDJO	GURMA	KWANGARE	NGERE	TEKE	
BONI	GUSII	LAKA (ADAMAWA	NGUMBA	TEM	
BORAN	HAMAMA	LAMBA	NGWAKETSE	TENDA	
BRONG	HAUSA	LAMBYA	NSENGA	THONGA	

Table Q – Appendix-Table 2.3: List of Tribe Groups Used in the Regression Analyses

Chapter Three:

The Geopolitics of Ethnic Boundary Demarcations and African Economic Development

Abstract

The study of under-development and tribal boundary demarcations in Africa has previously been limited to the Euro-centric position that the colonial scramble for resource and territory left the inherited post-colonial nation-states fundamentally ill-positioned to develop at a similar pace with the rest of the world. In this chapter, we argue that this narrative is as much intra-national (local, inward or internal) as it is inter-national (cross-national, outward or external). We find using an original dataset design that economic under-development can be observed as badly in subnational African tribal homelands that were solely affected by intra-national or provincial border demarcations as in tribal homelands that were jointly affected by both intra-national and inter-national border demarcations. Interestingly, subnational tribal homelands in Africa that were only demarcated inter-nationally were not significantly negatively associated with economic development as the relevant literature would have us believe; this persists even after controlling for two key politico-economic pre-colonial tribal characteristics.

The local dynamics associated with these patterns are further explored in line with the local trust and ethnic discrimination literature. We find that the patterns of under-development observed using the nighttime lights index and the average roads length measures are consistent with the attendant local trust and ethnic discrimination channels. While local trust measures are consistently negatively and significantly associated with subnational locally demarcated tribal homelands, ethnic discrimination by the government is significantly targeted at or positively associated with tribal homelands that were demarcated intra-nationally (splits at the sub-national frontier only) and similarly at those that were jointly demarcated both intra-nationally and inter-nationally (splits at both sub-national and international frontiers). Our results go to show that local dynamics are as debilitating as externally composed ones and ignoring these in our study of African development could certainly ill-inform any policy attempts interested in achieving economic or politico-economic progress.

Section 3.1 – Introduction

Economic development is a multifaceted phenomenon. The idea that the drivers of economic growth can be classified broadly as proximate and fundamental factors has dominated the empirical literature for quite a while now (Spolaore & Wacziarg, 2013). However, the approaches employed to distinctly assess these classes of growth-stimuli with regards to economic outcomes across economic units remain an avenue for significant contributions to the literature.

Economic development in Africa had over decades been pervasively investigated from a country-unit perspective (Grier & Tullock, 1989; Barro, 1996; Easterly & Levine, 1997; Alesina & Perroti, 1996; Alesina et al, 1996; Englebert, 2000; Acemoglu, Johnson & Robinson (AJR), 2001; Glaeser et al, 2004; etc.). As a result, key fundamental issues affecting the geo-institutional fabric of the growth environment, such as the relevance of ethnic structures for instance, were relegated to aggregated-dimensions consistent with those country-unit constructs. The connection between geography, politics and socioeconomic outcomes are very intricately interwoven - more so within the African context (Bloom et al, 1998; AJR, 2001; Naude, 2004; MacKinnon & Cumbers, 2009; Crescenzi & Percoco (Eds.), 2013; Auer, 2013; etc.). The significance of the African ethno-cultural angle seems particularly dominant, however: mundanely appearing to be the glue holding and orchestrating most levels of societal organisation, even in modern times (Nunn, Robinson & Weigel, 2015). The very institutional fabric underpinning the African development model is not easily investigable without a framework that rightly recognizes the inherent importance of disaggregation and individuality made relevant by the enduring manifestation of ethnic identification in African geopolitical states. This itself contradicts the notion of statehood imposed by the colonial and subsequent postcolonial political construct.

While the literature has capitalized on very worthwhile issues of ethnic heterogeneity and dissension in its assessment of this inherent ethno-institutional interplay, few studies have gone further to actually assess the challenges broader issues of ethnic partitioning bring to bear on economic development. It is common knowledge that the African region is among the most ethnically dense continents in the world. It is also acknowledged that the amalgamations of different such ethnicities into single nation-states following the Berlin conference of 1885 may have contrived the protracted geopolitical environment of institutional infancy

resulting in sustained weak economic outcomes compared to other parts of the world. However, the subject of ethnic demarcations (cleavages) has only recently begun being empirically analysed from a disaggregated subnational administrative-level by the likes of Michalopoulos & Papaioannou (2011) - who address our question from an ethno-country level of administrative interaction (or geo-intersection). By operating at a more disaggregated level, key aspects of this research area can be explored much deeper. This is particularly because the African socio-political and economic landscape is still very much rife with sub-regional themes of ethnic and tribal individualism and dissentions.

3.1.1 – Contribution

In this paper we attempt to assess the relationship between ethnic homeland demarcations and economic activity from a multilevel analytical perspective. It considers this African-relevant empirical question by looking at 3 different kinds of ethno-geographic splits arising from the geopolitical interaction (or geo-intersection) of historic ethnic homelands with modern subnational African states. This approach is multilevel in that it accommodates a framework where ethnic homelands are nested within the administrative jurisdictions of subnational state or provincial boundaries, which are further nested within the central administration and territory of their corresponding countries. As a result of this design, our study of split-ethnic homelands is divided into two administrative dimensions or frontiers. This allows us to obtain four split-ethnic categories, which could be viewed either as a split-intensity index measuring the outward degree of tribal demarcations or as an ordinal classification of related types of tribal demarcations[56]. These split categories correspond to different levels of ethnic homeland demarcations: that is tribal homelands with no splits at both frontiers; those split at the sub-national frontier only; those split at the international frontier only; and those jointly split at both the sub-national and inter-national frontiers, respectively. Consequently, this makes our design subtly yet significantly unique from those that have previously looked at fundamental African development issues from different ethno-country perspectives such as Fearon,

[56] It is important for the reader to note at this point that the terms tribes and ethnic groups are used interchangeably. Also the term state is used to stand-in for the next subnational administrative or provincial level to a country. The country is sometimes referred to as a nation-state, which is distinct from the subnational state or state, as already highlighted.

2003; Nunn (2008); Nunn & Puga (2012); Michalopoulos (2012); Michalopoulos & Papaioannou (2013, 2014, 2011).

We attempt to achieve our primary objective of assessing the impact of ethnic homeland demarcations on the cross-sectional variation of subnational economic activity in Africa by specifically:

i. Looking at the average impact of different levels of ethnic or tribal homeland demarcations on economic development using a novel-fabricated measure of geo-ethno-split intensity index. This index provides a single measure of the intensity of the geographic (geopolitical) demarcations of ethnic homelands within and across countries. It also provides the literature with an alternative measure of ethnic fragmentation based on the administrative level and the point of effect for each respective ethnic homeland. Such a measure not only allows this analysis to capture the average impact of moving laterally across each category but also allows the study to demonstrate the debilitating impact of ethnic cleavages in general on the identification and development of these geopolitical units within the African region

ii. The application of our multilevel approach to the empirical question of ethnic partitions in Africa also affords us the opportunity to evaluate its mixed impacts on economic development. In this design, differing sources or actors of ethnic homeland partitioning within Africa are considered. Thus providing us with an avenue for juxtaposing their individual associations with economic activity across the geopolitical units of interest

iii. Also by comparatively analysing the effects of tribal demarcations from a source or actor perspective, we are able to statistically revisit claims: that the exogenous uninformed scramble for African territories conducted by the European powers at the time of the Berlin conference would have created better current economic outcomes had these ethnic nations been demarcated and individually identified along tribally conscious lines. Such an assessment is pertinent and becomes feasible within a framework as ours. Particularly because, we exploit a pseudo-experimental design in which each country should have had, ex-poste, an opportunity to fabricate geopolitically administrative identities consistent with the homelands and customary practices of its constituent tribes using subnational or internal boundaries following the independence drive of the late 1950s. Although, most countries resulting from the colonial scramble for Africa inherited and maintained their respective outer national boundaries and identities as initially prescribed and later formally agreed on at the 1st Ordinary Session of the Assembly of Heads of States and Government of

> the OAU in Cairo of July 1964, the internal partitioning and governing of each country remained an independent responsibility of each realized multi-ethnic nation-state. The resulting multi-ethnic nation-states certainly had more knowledge of the extent of their respective existing cultural diversity or homogeneity and would have been compelled to create institutions and geopolitical delimitations consistent with their respective internal ethno-polity structures to maximize any possibility for national singularity or nationhood. Taking this narrative to the inquiry affords us the opportunity to unearth possible root causes of underdevelopment beyond the over-dramatized colonialism and imperialism arguments still dominating the African story after over three-decades of widespread regional political independence

The core premise of our study is that: regardless of the type or source, the geopolitical demarcations and polarization of tribal states in Africa have direly affected the development of economic activity in subnational areas across the African region. While this angle of research enquiry has previously been explored at both the cross-national and ethnic-country levels, we are yet to see how differing political agents of such demarcations may have affected economic activities distinctly. There are two main classes of political agents at play in this discourse.

The first of these classes, were the pre-colonial European political agents that divided and amalgamated a significant number of previously autonomous ethnic homelands into whole new governable territories (or polities). These polities were noticeably governed by key European colonial powers that administrated each only almost autonomously; these colonies would later become recognized countries in the post-colonial independence era from the late 1950s.

The second class of agents is the principal administrative governments of the respective post-colonial independent nation-states themselves. There is no reason to believe that these second set of political agents did not know the underlying geopolitical structures of the ethnic homelands spanning their respective national territories. Although the timing, design and motives of their ethnic-demarcation actions may have varied for each country across the region, they nonetheless had a significant impact on the eventual turn of events both socio-politically and socio-economically for the experiences of each resulting tribal homeland.

A lot has been said about just how dismal the ill-informed scramble, conquest and imperialist colonisation of Africa was on the modern development of the countries spanning the region. However, little has been done to actually compare

and demonstrate how different types of tribal demarcations or how the extent of such demarcations collectively affected developmental differences across different African subnational states. With the framework harnessed herein, we can assess this scenario and confirm whether it is the European scramble that is mostly to blame or whether there are other much related underlying local factors of equal attention to be reflected on.

3.1.2 – Analytical Background

The multilevel framework adopted herein is not so farfetched from the socio-political reality of most African economic and political administrative arrangements. Most of these arrangements are multi-tier in nature permitting most country-governments to replicate their machineries at subsequently lower levels of administration (or subnational states such as provinces, protectorates, regions, divisions, local government areas, etc.). These frontline governance constituents or administrative arms usually act as institutional appendages with which respective central governments are able to gain administrative proximity to grass-root levels (or ethnic homelands) for the pooling, provisioning and consolidation of national resources. They are effectively a systemic adaptation of national institutions to the unique and nuanced experiences of internally existing subnational ethnic dynamics. The relevance and justification for intra-national regional territorial delimitations and demarcations can be substantially found in the decentralisation literature. It is often argued therein that the delegation of fiscal, political, and administrative responsibilities to subnational levels of government not only necessitates bringing economic and political systems closer to the people but also ensures the integration of the national political system itself (Litvack, Ahmad & Bird 1998).

For countries in one of the most culturally diverse continents of the world, it is no surprise that national demarcations consistent with subnational ethnic differences and similarities are an inevitable inclination. By analysing the cross-sectional development of economic levels at this tribal-subnational (split) level, we can provide initial insight into the actual outcomes of this sort of political decentralization regardless of the intrinsic motives. When we speak of motives here, we are recognising the position that some of the existing subnational demarcations may have been politically engineered to diffuse the political

influence of competing ethnic groups and not necessarily as engines for promoting power-sharing, governance quality or administrative accountability as we would expect[57]. This is not surprising especially as the unofficial goal of most political parties in ethnically-dense countries are not necessarily directed at ideological supremacy but at ethno-regional identification and empowerment[58]. In fact, ethnic (or tribal) identification, and individual behaviour, political and socioeconomic circumstances are not easily separable – particularly so for multi-ethnic countries in Sub-Saharan Africa (Easterly & Levine, 1997; Eifert et al, 2010).

By geographically inferring the nature, type or source of ethnic homeland demarcations from a subnational perspective this study is able to not only unearth the impact of ethnic boundary fragmentation on economic activity but also reveal how the intensity of such splits correlates with economic development as well. Although some channels are explored in the general discussions around the results of this chapter, causal pathways are not themselves its central focus. With such a functionally relevant level of administrative disaggregation, this work focuses on comparing the cross-sectional average levels of economic activity (development) across differing actors or degrees of geopolitical ethnic homeland splits. Such a non-causal analysis could, at the least, be relevant to absolving or reinforcing the role of the European scramble for African colonies versus those propagated by local non-European dynamics on its regional development. Also if nothing else, this provides the empirical literature with an alternative measure of ethnic homeland fragmentation very different from those previously adopted. We are able to identify major classifications, types or sources of ethnic homeland cleavages; and how these provide differing perspectives on the sources and distribution of economic development across the African continent at a provincial level.

A number of interesting paths of inquest can arise from a research such as this. Its findings could firstly be relevant for its plain purpose of demonstrating the association between differing ethnic homeland demarcations and economic development (the main objective). Or secondly, it could provide a much deeper inference about the key findings especially with respect to uncovering some

[57] We expand on these opposing positions or arguments a few paragraphs down.

[58] This is a very similar notion to that discussed in Wimmer and Cederman, 2009: 317, where they opine that unlike empires, the politics of nation states will hover around the relative distribution of ethnic groups' access to political power.

dynamics inherent to the geopolitical and socioeconomic landscape of the subnational regions involved.

If for instance, it turns out that areas of intra-national (or local) ethnic demarcations seem to be as underdeveloped as those areas affected by European-led demarcations. Then it could indicate that the schism engineered at this level was as debilitating on economic activity as was that done in ignorance by the Europeans. This could be a signpost supporting the narrative that the intra-national demarcations of ethnic homelands in Africa might itself be part of a deeper local political and economic power game orchestrated to weaken potentially competing ethno-political minorities against their dominant (incumbent) counterparts. That dominant ruling tribes (governing incumbents) attempt to localise potential ethnic tensions stirring from ethno-political competition to the local subnational levels thus redirecting the interests of these minority groups away from national politics and such like affairs unto local ones.

A country that embodies such dynamics happens to be one of the most populous and ethnically diverse countries in Africa, Nigeria. This is a country where even the formula about how national resources (political-power and economic wealth) are internally distributed is often rife with ethnic and tribal subjects. A typical illustration of this issue is on the matter of state creation (a major mechanism for national resource allocation), which is mostly done down the lines of ethnic identity and to ameliorate claims of disenfranchisement from national affairs (Vande, 2012 and Adeyemi, 2013). In such cases subnational-state creation (power decentralization) is a buyout effort to decrease the probability for cessation or irredentist demands by affected interest groups.

Also state creation could minimise increased demands for political competition and participation at the national levels. This channel is definitely consistent with the segregation → trust → institutional-quality → economic-development argument put forward by Alesina & Zhuravstaya (2011). This is mostly because we would expect that disenfranchised or politically isolated groups would naturally have minimal confidence (trust) in the politically dominant groups to fairly allocate public (national) goods towards their homelands compared to the homelands of the politically dominant group.

Similarly, it is likely that the extent of ethnic demarcations also maps unto relative underdevelopment through the ethnic-inequality → political-inequality (political discrimination) → unfair allocation of public-goods → regional economic

under-development channel as argued by Alesina, Michalopoulos & Papaioannou (2012)[59]. Both routes linking ethnic discrimination with economic development are very similar or more or less entangled, as they would inevitably result in a loss of trust amongst ethnic groups in cooperative endeavours. This is because public goods provisioning and the equitable or fair allocation of shared national wealth is itself a symptom of the quality of governance or a component of the broader national institutional framework.

On the other hand, it could be the case that while the European-forced demarcations on ethnic homelands appear to have had debilitating consequences for development, the locally fabricated split tribes could have been positively associated with economic outcomes. This is the kind of outcome that proponents of African decentralization expect. It signals that internal partitions were designed with knowledge of the inherent local ethnic homeland structures in order to maximize cooperative functioning both at the local and national echelons. As a result the problems of ethnic domination and fears of local resource exploitation by dominant groups are staved off and those of devolution, regionalization, and ethnic cooperation are thus promoted.

In this paper, we argue that our results seem to support the former much more than the latter. It appears that demarcations at the local echelons happen to be consistently larger and more significantly associated with lower developmental levels. This is in contrast to those areas affected by the uninformed European-only splits or those unaffected by any sort of demarcations at all. This suggests that the findings in the literature providing support for European-induced underdevelopment through the scramble for Africa, as contained in the ethnic fractionalization literatures and their affiliate narratives, may in fact be echoes of deeper subnational-level dynamics independent of that type of European influence as propagated by the anti-imperialist studies.

Further implying that although ethnic cooperation matters in a multi-ethnic structure as in African nation-states, it matters through more local channels than through aggregated national ones, contrary to what the earlier anti-colonialist researches would have us believe. The link between human interaction (institutionally and individually) and land or territoriality is quite a strong one; it often underpins many other issues around strides at state creation and nationalism

[59] Also see Burgess et al, 2013 for a similar study where they show that access roads development projects were under-provisioned for ethnically under-favoured ethnic groups.

(Saltman (Eds.), 2002). After all early evidences of regional ethnic tensions, dissentions and competition for land and resource-wealth in Africa predate European influences of colonialism or imperialism (see Besley & Reynal-Querol, 2014).

3.1.3 – Survey of Chapter's Remaining Sections

The section following this, the second section, is the literature review that provides an overview of selected papers from the literature relevant to ethnic demarcations and economic development. The third section is the model specification and data setup section, which provides information on the models to be estimated, a brief description and summary of the variables used, their sources and also information on how the overall dataset is setup. Section four is the results and interpretation section, which displays the results and attempts to provide some intuitive interpretations in line with narratives introduced and maintained throughout the entire chapter. This final section also includes a general conclusion which recaps on the major findings, their implications and provides some avenues for further research based on the weaknesses of those from this research.

Section 3.2 – Literature Review

Understanding development patterns in Africa has been at the heart of economic development analysis for over 4 decades now. This is because although the region is rife with the abundance of peoples and natural resources, it has been plagued by widespread fundamental challenges to its ability to utilize and harness developmental recipes in the same manner as other countries from other regions of the world having similar endowment constellations. In this section, we attempt to briefly review the literature related to how ethnic demarcations have affected developmental strides in Africa and elsewhere relevant in the world.

3.2.1 – Ethnic Demarcation and Economic Performance

While the empirical literature on ethnicity and African development has previously predominantly focused on channels through conflict (Easterly & Levine,

1997; Wimmer & Cederman, 2009; Montalvo & Reynal-Querol, 2005a & 2005b)[60] via demography-motivated notions of ethnic divisions or clustering (fractionalization, diversity, polarization, etc.), the literature on ethnic homeland boundary alterations (geography-motivated) have remained relatively meagre. The interlacing of geographic endowments (homeland ownership including the accompanying natural resources) and political influence can to a considerable degree mask the underlying relationships between demography-based treatments of the 'ethnicity-development' discourse. Discussion on ethnic demarcations and economic development in Africa has so far accentuated the modern incidence of the European scramble for land, resources and influence in the region. Themes of border artificiality (Asiwaju, 1985; Griffith, 1986; Alesina, Easterly & Matuszeski, 2006[61]; Fall, 2010; Green, 2012), forced-integration ("suffocation") and forced-separation ("dismemberment") (Englebert & Carter, 2002) appear to dominate the primary conduits via which the scramble linked historical homeland demarcations with current strides in economic development. Other sub-channels have been explored and it is not improbable that demarcations at different administrative levels could in some cases generate positive and in others negative associations with the development of economic activity[62].

The debate on local (intra-national) demarcations can themselves be found in the decentralization discourse, which has mostly been non-African. In this literature, decentralization at the subnational administrative levels can both be good and bad for development depending on specific circumstances and on the case study in-view. According to the consolidation discourse, amalgamation of subnational groups results in bigger more efficient mechanisms for the allocation of public goods at their respective grass root interests benefiting from positive economies of scale effects (see for instance Dollery et al, 2006; Slack & Bird, 2013; etc.). While opponents to this amalgamation narrative argue that more administratively fragmented systems representing clear common singular interests

[60] For readers interested in this debate, they could also see the counter-literature arguing that conflict is not necessary significantly associated with ethnic fractionalization as one would expect (Bowen, 1996; Montalvo & Reynal-Querol, 2005a; Collier 2007: 25-26; etc.).

[61] This was later published as Alesina, Easterly & Matuszeski (2011) in the Journal of the European Economic Association.

[62] See Alesina & La Ferrara (2002) for a review of the positive and negative effects of ethnic diversity on economic activity and the underpinning micro-theory.

tend to encourage local participation, administrative responsiveness and accountability to the local interests they serve (see Faguet, 2004, for instance).

It is not hard to see how the ethnic diversification literature can fit into these themes of intra-country demarcations, particularly when such demarcations relate with the amalgamation of smaller ethnic homelands into relatively bigger subnational administrative units or the splitting of bigger ethnic homelands into multiple subnational administrative units. Miguel & Gugerty (2005) shows that ethnic diversity in rural western Kenya was associated with weaker social sanctions resulting in the shrinking of needed public goods provisioning (such as "primary school funding and water-wells maintenance"). Alesina, Baqir & Easterly (1999) do a similar work with datasets on U.S. cities, metropolitan areas and urban counties and find that ethnic fragmentation was inversely associated with key productive public goods provisioning such as education, roads, and etc. This line of narrative relating ethnic diversity (heterogeneity) in subnational areas with economic development through public goods provisioning is very much consistent with that which attempts to connect ethnic diversity with governance quality (see La Porta et al, 1999 and Alesina & La Ferrara, 2005).

At this point one can see how these two literatures (decentralization and ethnic-heterogeneity) knit together: amalgamation appears to be more detrimental to the provisioning of public services (or goods) in more ethnically heterogeneous circumstances compared to decentralization, while the reverse is more likely true for more ethnically homogenous areas where commonness in heritage fuels administrative cooperation. Although, Dippel (2014) in his study on Native American reservations shows that even for ethnically and linguistically homogenous groups, forced integration into a system of shared governance was associated with a 30% decline in present day levels of economic development. It is important to note here however, that these were for groups distinguished at sub-tribal bands implying that while they may be of common ancestry at the tribal level they are quite distinct groups at a finer more detailed level of classification.

The African growth literature is very well researched for how cross-national demarcations of ethnic homelands associate with the economic development of affected areas. Cross-national ethnic demarcations are very much the penetration method in this literature. It is argued that such demarcations affected the homeland and cultural identities of tribes at the periphery of colonially inherited national borders. This further resulted in the polarization of both homelands and

peoples; where the affected areas have been found to not only be associated with weaker national institutions on the average (Michalopoulos & Papaioannou, 2014) but also have been found to have higher average conflict incidences (Michalopoulos & Papaioannou, 2012). Although some authors such as Gennaioli & Rainer (2007), Michalopoulos & Papaioannou (2013) and Fenske (2014) argue that the politico-economic and unique fundamentals of key pre-colonial tribal states made them more affable to modern recipes of political and economic development.

Taking a summative glance at these intersecting discourses on ethnicity and economic development in Africa, we find that the partitioning of ethnic homelands created a regional environment marred by both ethnic divisions[63] (Easterly & Levine, 1997; Alesina et al, 2003; Fearon, 2003) and ethnic inequality or political inequality (Alesina & Zhuravstaya, 2011; Burgess et al, 2013; Hodler & Raschky, 2014). Each of these narratives demonstrate (even though differently) that: the loss of historic homeland arrangements for a significant percentage of African tribes (Asiwaju, 1985) resulted in a loss of politico-economic bargaining influences and created institutional (cultural and territorial) instabilities, which eventually translated into varying levels of economic development at various administrative levels.

[63] This could equally be referred to using any of the compound classifications adopted in the literature: cleavages, fragmentation, fractionalization, polarization and etc.

Section 3.3 – Model Framework and Data Layout

3.3.1 – The Core Model

$$dev_{ejc} = \theta_1 \text{split\_index}_e + \beta_1 X_{ejc} + \beta_2 X_e + \alpha_{c;j} + \ni_{e,j,c} \quad (1)$$

and:

$$\ni_{e,j,c} = U_j + V_c + \varepsilon_{ejc} \quad (2)$$

where:

- $\boldsymbol{split\_index_e}$, is our categorical variable classifying each ethnic homeland into 0 – for those with "no-splits", 1 – for those "split at the sub-national frontier only", 2 – for those "split at the inter-national frontier only" and 3 – for those "jointly split at both the sub-national and inter-national frontiers simultaneously"
- $\boldsymbol{X_{ejc}}$, is a vector of control variables characteristic of the literature capturing exogenous geographic, ecologic and demographic features including a general intercept. These control variables vary at the ethnic-state (subnational) split level; while $\boldsymbol{X_e}$ are historic pre-colonial covariates (political decentralization and settlement pattern indices) that vary at the ethnic or tribal level
- $\boldsymbol{\alpha_{c\,;\,j}}$, controls for either country specific characteristics (c) or state (or subnational) level specific characteristics (j). Such as institutional and administrative dynamics, which may vary from one subnational jurisdiction to another
- $\ni_{\boldsymbol{e,j,c}}$, is the error process conjectured to be composed of three components - $\boldsymbol{\varepsilon_{ejc}}$ follow a white noise IID process, while $\boldsymbol{U_j}$ **and** $\boldsymbol{V_c}$ represent the variance adjustments of the error-process for both the subnational state and country levels, respectively, which we use in estimating our multilevel model variants. $\boldsymbol{U_j}$ **and** $\boldsymbol{V_c}$ allows the model to control for these two sources of variation in the error-term which otherwise would not be captured by $\boldsymbol{\varepsilon_{ejc}}$ resulting in omitted variable issues and such like biases
- $\boldsymbol{dev_{ejc}}$, is the response variable of interest and it attempts to measure the level of current economic activity observed at the demarcated ethnic homeland level. There are two main proxies used in this study – the average nighttime lights density variable and the average access roads length variable. Both have been normalized to represent units per square kilometres, respectively. To minimise measurement problems arising from both the data transformation approach and the actual measure of variable having a significant number of zeroes in it, we also provide estimates using the

variable in its original ordinal discrete form using a fixed-effects negative binomial estimator

- subscripts: $\boldsymbol{e, j \text{ and } c}$, identifies the split ethnic polygon, the corresponding subnational state and the country, respectively

We have chosen to specify the model above to be consistent with a multi-level dataset structure. There is at least one ethnic group or tribal homeland (e) for each subnational state (j) within each country (c). As a result, there are an unbalanced number of observations per group in our designed sample. We choose to estimate the model above using an overarching multilevel (mixed-level) model with a maximum likelihood estimator, "mixed", found in STATA[64]. This provides us with an advantage over conventional longitudinal models given this unbalancedness in our data structure (see Snijders & Bosker, 2012). Adopting such an approach allows us to estimate our desired model coefficients relatively more efficiently while controlling for multi-way nesting and cross- (or in-) variation patterns underlying the model's errors process. As a result we are able to obtain estimates for the differing variances in the error-structure at each level of our analysis (in this case 3). We nonetheless compare our results to the Fixed-effect (FE) estimations conventional to the growth discourse where we control for unmeasured time-invariant heterogeneity consistent at state- (subnational-) or country- specific characteristics such as institutional dynamics, for instance. In such instances, we adjust for any potential multi-way cluster robust standard errors consistent with a three-way structure (ethnic-state-country) following Cameron, Gelbach, & Miller (2011).

[64] This is a computer application like Eviews, Matlab, or R normally used for statistical data analysis and for preparing statistical outputs (graphs and tables).

3.3.2 – The Model for Exploring Some Deeper Local Channels (following Nunn & Wantchekon, 2011)[65]

$$y_{i,e,d,c} = \alpha_c + \theta_2 \text{split\_index}_e + \beta_1 X_{i,e,d,c} + \beta_2 X_{d,c} + \beta_3 X_e + \ni_{i,e,d,c} \quad (3)$$

because the error-term can equally be affected by the data structure, which is multi-nested at the ethnic and district levels, we also go ahead and provide estimates incorporating a mixed-effects structure:

$$\ni_{i,e,d,c} = U_d + V_e + \varepsilon_{i,e,d,c} \quad (4)$$

where:

- $\boldsymbol{X_{i,e,d,c}}$, $\boldsymbol{X_{d,c}}$, $\boldsymbol{X_e}$ are vectors of control variables varying at the individual, district and tribal levels, respectively, mostly following the model originally adopted in by Nunn & Wantchekon (2011). These include the individual's: education levels, occupation, religion, living conditions, age, gender, urban indicator for living location, perception of tribe's political discrimination, perception of tribe's economic conditions; the affiliated district's: ethnic fractionalisation index and fraction of ethnicity within district; and the tribe's: city status in year 1400, pre-colonial explorer contact status, total pre-colonial missions in area; malaria ecology, pre-colonial railway contact, pre-colonial political centralization index, and pre-colonial settlement status indicator. Two respondent level variables are introduced as additional controls (individual perception of political discrimination against own tribe and own tribe's economic conditions), because we conjecture that these are correlated with both trust and patterns of internal tribal homeland demarcations[66]
- $\boldsymbol{\alpha_c}$ is the country-level specific characteristics capturing shared unobserved heterogeneous effects such as institutional and other fundamental dynamics specific to each respondent's country. These may, themselves systemically affect the trust disposition (or inclinations) of each respondent and the patterns of ethnic demarcations characteristic to each respondent's country
- ∋, is the error process hypothesized to have three components for the base model - while $\boldsymbol{\varepsilon_{,i,e,d,c}}$ follows a white noise IID process, $\boldsymbol{U_d}$ **and** $\boldsymbol{V_e}$ represent the variance adjustments of the error-process for both the subnational state level and country

[65] For more details on how the original model is set up and for further information on the variables used in this trust model, please refer to the original paper. We also adopt this latter model to test some of the channels introduced and discussed in section 3.1.2, above.

[66] There are some theoretical and empirical allusions to this position in the literature survey and analytical background sections above.

levels, respectively. $\boldsymbol{U_d}$ ***and*** $\boldsymbol{V_e}$, allows the model to control for two sources of variation in the error-term which otherwise would not be captured by $\boldsymbol{\varepsilon_{i,e,d,c}}$ resulting in omitted variable bias

- $\boldsymbol{y_{i,e,d,c}}$, is alternatingly one of 5 trust measures used in the paper by Nunn & Wantchekon (2011) including an additional variable measuring the extent of government nepotism (or tribal favouritism) following a channel similar to that explored in Alesina, Michalopoulos & Papaioannou (2012). The trust variables include trust for local councils, trust for neighbours, trust for relatives, trust for other tribes' members, trust of own tribe' members, respectively
- subscripts: $\boldsymbol{i, e, d}$ ***and*** $\boldsymbol{c}$, identifies the Afrobarometer[67] respondent, the respondent's ethnic or tribal group as defined by Murdock (1959), the respondent's subnational district, and the respondent's country, respectively

To estimate the model laid out in equation 3 and 4 above, we adopt a maximum likelihood ordinal-logit estimator[68] with a three-level mixed-effects framework (expatiated in equation 4) to capture the survey's multi-cluster and multi-nested dynamics inherent in the data structure. This adjustment in our overarching assumption controls for intra-cluster correlations because the respondents involved in our estimations are resident in different districts and belong to different tribal homelands (or groups).

We use this latter model specification (equations 3 and 4) to test two of the likely channels via which split categories may be associated with specific patterns of underdevelopment as discussed in our analytical background in section 3.1.2 above. Since the exploration of causal channels is not the focus of our current study herein, we focus on the micro-trust and tribal fairness conduits. It has been discussed extensively in the literature how local trust intricacies affect economic development via mechanisms of coordination and institutional quality (see for instance: Alesina & Zhuravstaya, 2011). Additionally, because our dependent variables from our core model of interest (equation 1 in section 3.3.1 above) are

[67] This is the survey-based Afrobarometer Merged Round 3 Dataset available for 18 countries for the 2005 survey stream. The codebook can be found on the http://afrobarometer.org/sites/default/files/data/round-3/merged_r3_codebook2_0.pdf website, for those interested in how the survey questions are designed or framed.

[68] Although Nunn & Wantchekon (2011) use an OLS based model, we choose to use an ordinal-logit model, as we believe this provides a better treatment of the dependent variables – *trust* and *fairness*. These are ordinal variables with four response categories and not continuous variables.

public-good in nature (nightlights and roads), the resulting findings could readily be applicable to the empirical literature relating public goods (services) provisioning with ethnic fractionalisation (see for instance: Miguel & Gugerty, 2005; Alesina, Baqir & Easterly, 1999) and those associated with tribal level politico-economic dynamics that associate access to power with economic empowerment or uneven distribution of economic development (Alesina, Michalopoulos & Papaioannou, 2012).

In both models to be estimated (equations *3.1* and *3.3*), we are unable to control for tribal level fixed effects because our key right-hand side variable of interest, split index, is measured at the ethnic homeland level. As a result we control for two crucial and commonly used pre-colonial politico-economic controls in the literature (level of political complexity and settlement patterns) to account for the varying levels of institutional complexity for the tribal arrangements involved in our study. The only downside to this is that these two variables are obtained from the Murdock (1959) anthropological dataset on the activities of a subsample of 441 tribal homelands from the 834 homelands originally mapped. As a result only about half our full-data sample is considered when we control for these pre-colonial attributes.

3.3.3 – Variables and Data Sources

The dependent variables used in this paper are the average nighttime lights density measure by the DMSP-OLS nightlight data from the *National Oceanic and Atmospheric Administration* (NOAA) data repository and the average roads length data obtained from the Global Roads Open Access Data set (gROADS) v1, by the *Socioeconomic Data and Applications Center* (SEDAC)[69]. Both measures are polygonal averages constructed at the split ethnic-state (subnational) level. The interesting thing is that although both these variables can gauge economic activity at disaggregated split-levels, they could equally be useful for capturing cross-sectional variations in development consistent with the extent of government service or goods provisioning at such levels of governance administration. Nighttime light density would definitely capture significant amounts of street

[69] Please refer to section 2.3.4.1 for a more in-depth discussion demonstrating the suitability of the Roads access variable as a decent proxy for economic activity. Since these two proxies (nighttime lights and roads access) have been used in both chapters 2 and 3, they were discussed in more detail in the preceding chapter to avoid content repetition and any unnecessary verboseness.

lighting within communities, while roads infrastructure is a purely government-provisioned service (or good) at these echelons of society. With such measures our study makes its findings relevant to both the urban economics literature on ethnic heterogeneity and public goods provisioning and the economic development literature on ethnic heterogeneity and development.

The control variables are mostly composed of those often used in the literature for studies with similar analytical data structures. Three broad categories are adopted to classify the control variables used in our study: distance controls (distances to coast, national border, and national capital), geographical and ecological controls (mean elevation, soil productivity index, malaria index, absolute latitude, relative size of the split tribal homeland to the subnational state's geo-size, natural resource endowment controls (indicators for diamond and oil deposits) and demographic control (natural log of the population density).

All the controls are obtained using GIS sources. The mean elevation variable is measured as the average surface elevation for each split or non-split ethnic polygon. It is based on a joint work between the National Oceanic and Atmospheric Administration (NOAA) and the U.S. National Geophysical Data Center, TerrainBase, (version release 1.0). The soil productivity index measures the average suitability of the soil of each polygon in our sample to the crop best adapted to that soil's type, condition and composition. It is a dataset published by the FAO in 2002 with index values ranging from a minimum of 1 to a maximum of 97; where values increase in tandem with the soils productivity and suitability to the chiefly adapted crop. The malaria ecology index is based on the work of Kiszewski, Mellinger, Speilman, Malaney, Sachs, and Sachs (2004). It measures the prevalence and exposure of each split ethnic polygon to the malaria endemic. The onshore petroleum sites and the diamond sites endowments indicators are respectively obtained from the Peace Research Institute Oslo (PRIO) websites[70] [71]. An additional natural endowments variable controls for a represented group's (split or un-split) proprietorship of inland water. It is measured as the average area covered by a river, stream or lake for each represented tribal polygon. This variable is obtained as a raster file, titled Global Lake and Wetlands Database

[70] http://www.prio.no/CSCW/Datasets/Geographical-and-Resource/Petroleum-Dataset/Petroleum-Dataset-v11/

[71] https://www.prio.org/Data/Geographical-and-Resource-Datasets/Diamond-Resources/

Level 3 (GLWD-3), from the World Wildlife Fund (WWF) website[72]. An initial summary description of the variables involved in the estimation of equation 1 is provided in Table R - 3.3.4.1 below.

We arrive at our core data structure by spatially joining (intersection method) a recent map of the subnational administrative territories[73] for each country with the Peter Murdock (1959) map of African tribal homeland predating colonial political activities in the region. There are initially *834* habitable ethnic tribal polygons and *1285* subnational polygonal territories[74] (this includes island territories as well) in the peripheries of the African region. After selecting only those geographical intersections with a minimum of *100* square-kilometres, there are *728* subnational territories nesting the *834* tribal homelands spanning *50* African countries. This results in a total of *2985* observations in our working sample, in its full breadth. The subnational territories map adopted herein is spatially clipped for the African region from the 'RWDB2 Subnational-Ad2 Polygonal Boundaries' (current at: 26-09-2005) dataset and available from the FAO website[75]. For a visual inspection of the maps involved and the relevant territories included in this analysis, a quick map-intersection view is provided in Figure O - 3.3.3.1, below. From this figure, we can see that going to the subnational level provides us with an incredible opportunity for understanding grass-root dynamics that might be affecting the development of African homelands. It clearly provides us with more data points and more variation opportunities for carrying out further in-depth analysis of some of the development theories already introduced and discussed earlier.

The variables based on the Afrobarometer Round 3 data, as earlier highlighted, are obtained mostly from the replication file from the work of Nunn & Wantchekon (2011). We however, merge our tribal level split index with this principal dataset and include 2 additional variables (q80a and q80b, respectively) also found in the Afrobarometer that control for respondent's perception of their tribe's economic

[72] http://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3

[73] This is at administrative level two, where administrative level one is the actual country level while level three is a higher level of administrative territories local to each one from level two.

[74] It is noteworthy that this figure includes the Spanish and Portuguese island territories in the seas North-West of Africa.

[75] http://ref.data.fao.org/map?entryId=936f7280-b06d-11db-8922-000d939bc5d8&tab=metadata

condition and political influence relative to those of other tribal groups. We believe these two variables are correlated with both our split index and the dependents variables (trust and fairness) adopted in our channels-exploration exercise. A summary table of all the variables involved in the estimation of equation *3* (find above) can be found in Table T - 3.3.4.3 below, while a more detailed explanation of the model selection and rational can be found in the paper by Nunn & Wantchekon (2011). This latter dataset used in exploring these two sub-channels comprises of *21480* respondents living in *1287* districts within *184* Murdock-defined tribal homelands spanning *17* countries.

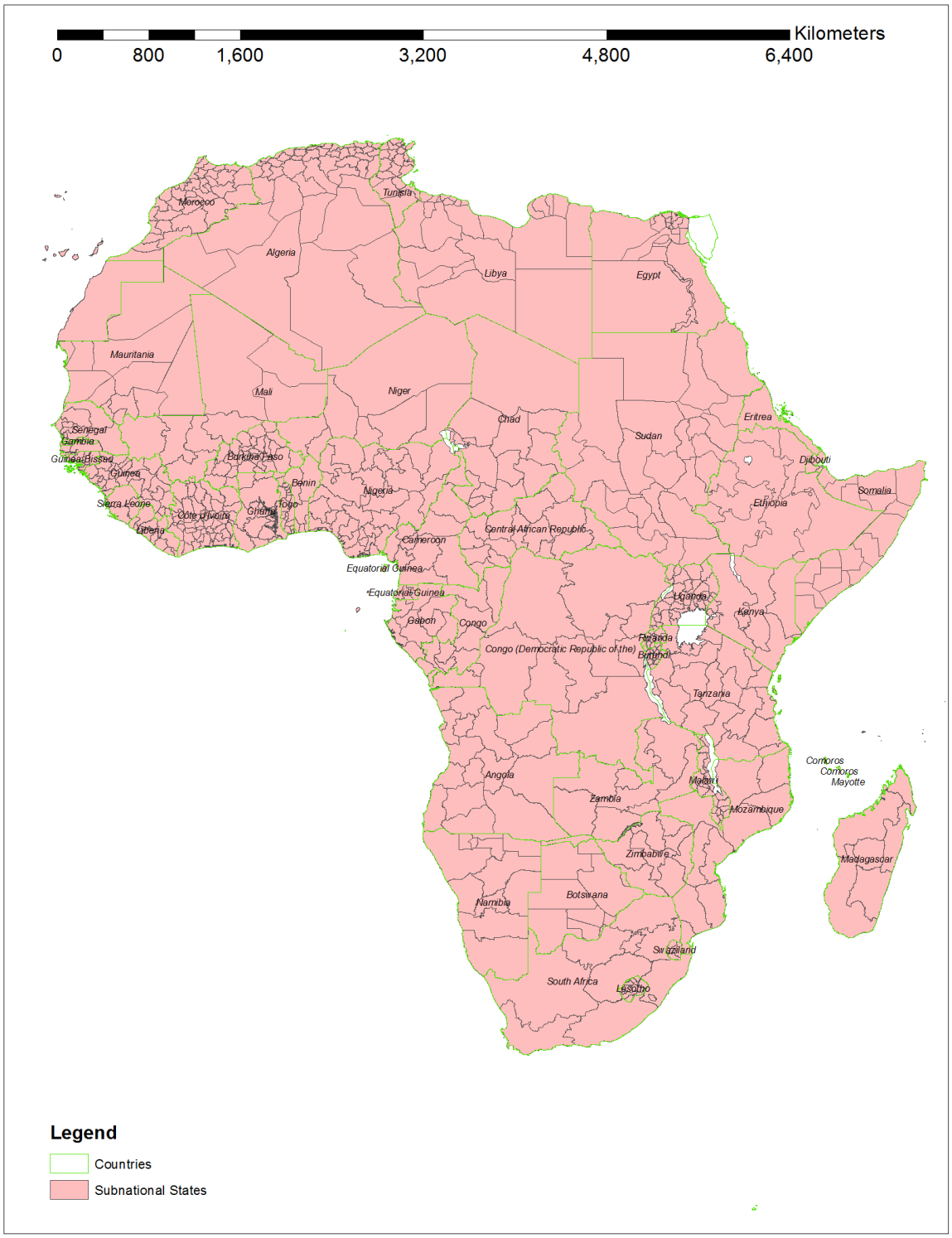

Figure O – 3.3.3.1: Map of National and Subnational Boundaries

Full Variable Name	Obs	Mean	Std. Dev.	Min	Max
Log (Stable Night light-density to split tribal homeland area ratio)	2985	-8.254246	2.894378	-17.00652	-1.589045
Stable Night light-density (ordinal)	2985	6.282412	7.104989	0	60
Log (Average Roads length to split tribal homeland ratio)	2877	-5.24841	1.792807	-10.58359	-0.0465962
Index of the extent of ethnic homeland splitting (ordinal)	2985	2.022111	1.049151	0	3
Log (Population density)	2985	2.98546	2.191487	-4.60517	9.067394
Split Tribal homeland's Distance to sea (in 1000 KMs)	2985	0.5631758	0.4399529	0	1.805956
Split Tribal homeland's Distance to the National border (in 1000 KMs)	2985	0.0270168	0.0318217	7.08E-06	0.3067594
Split Tribal homeland's Distance to the National capital (in 1000 KMs)	2985	0.4384652	0.3420012	0.0026696	1.941609
Mean Elevation (in 1000 KMs)	2985	0.6036329	0.4767962	-0.2323125	2.960833
Soil Productivity index	2985	8.007492	13.21188	1	97
Malaria Ecology index	2985	0.3382968	0.2541887	0	0.9407895
Absolute Latitude	2985	12.69702	9.439829	0.0041917	37.07984
Log (Split Tribal homeland land area to Unsplit Tribal land area ratio)	2985	-2.180032	1.600903	-7.744343	0
Indicator for the presence of Diamond mines	2985	0.0515913	0.2212375	0	1
Indicator for the presence of Petroleum wells	2985	0.0847571	0.2785666	0	1
Log ((1 + Average Inland water area))	2985	2.329027	17.5423	7.22E-06	548.5612

Table R – 3.3.4.1: Tabular Summary of Variables used in Equation 3.3.1

The table above provides a brief tabular summary of the variables involved in the estimations of equation *(1)* in section 3.3.1 above. While Figure P - 3.3.4.1 below, provides a spatial map of our core right-hand side variable of interest (the split index). As is apparent, there is a considerably good amount of variation in this variable although there appears to be some country-based clustering patterns. This sort of intra-group variations arising from the two upper-administrative levels of our dataset (state and country levels) coupled with varying number of observations within each nested level, merits our preferred choice of a mixed-effects model. Also the split index shows that there are more observations in categories 1 and 3 compared to categories 0 and 2. Actually Table S - 3.3.4.2 below summarises this more aptly. Categories 2 and 3 capture influences of split impacts across national borders and jointly shows that about 56% of the ethnic homelands in our sample were divided across countries boundaries. Our study notwithstanding tries to compare this group with those who were only split locally or intra-nationally and those that were not demarcated at all by any of our two

key political actors (either Euro-colonialist or own-country governments); who as a result remained within the geopolitical confines of a specific or common country.

Our Split Index	Freq.	Percent	Cum.
0	147	4.92	4.92
1	1,169	39.16	44.09
2	140	4.69	48.78
3	1,529	51.22	100
Total	2,985	100	

Table S – 3.3.4.2: Table Summary of our Split Index

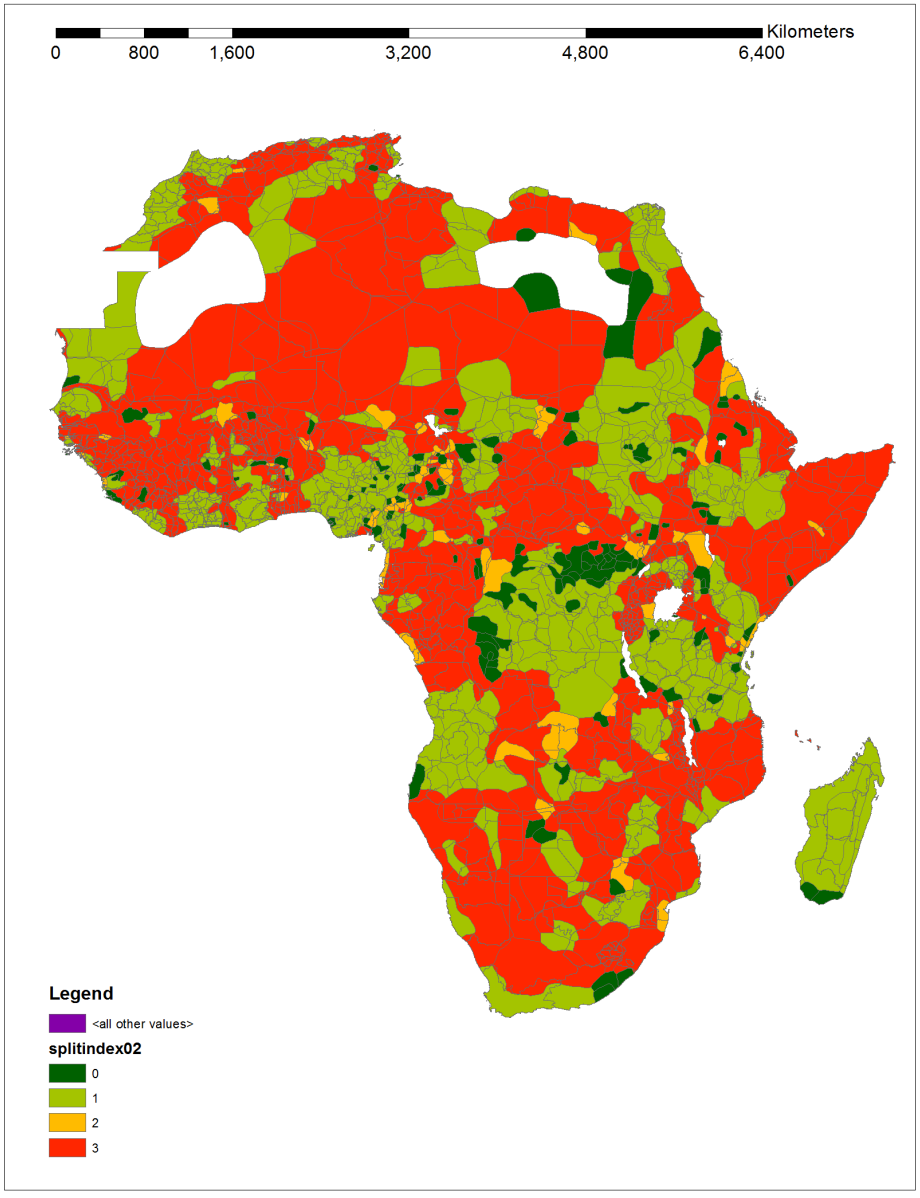

Figure P – 3.3.4.1: Spatial Map of the Split Index Variable

Variable names	Obs	Mean	Std. Dev.	Min	Max
How much do you trust your local council? (q55d)	20008	1.664484	1.101049	0	3
How much do you trust people from other ethnic groups? (q84d)	20081	1.359245	0.9960181	0	3
How much do you trust people your own ethnic group? (q84c)	20280	1.671795	1.002917	0	3
How much do you trust your relatives? (q84a)	20396	2.184252	0.9592221	0	3
How much do you trust your neighbours? (q84b)	20358	1.731506	1.008838	0	3
How often is your ethnic group treated unfairly by the government? (q81)	18966	0.7928925	0.9795828	0	3
Split Index by Tribal group	21480	2.031378	1.052663	0	3
ln(1+Atlantic+Indian Slave Exports) by Tribal group	21480	1.967817	2.319538	0	6.752222
Respondent's Perception of Ethnic Group's Relative Economic Conditions (q80a)	19667	3.157472	0.9856416	1	5
Respondent's Perception of Ethnic Group's Relative Political Influence (q80a)	19194	3.058664	0.9901812	1	5
District-level ethnic fractionalization	21480	0.4038963	0.2968748	0	0.9057617
Proportion of ethnic group in district	21480	0.6020231	0.3477237	0.0027174	1
Tribe's Average Malaria Ecology	21480	11.54582	9.765481	0	34.63975
Total Catholic + Protestant mission per land area	21480	0.000223	0.0003402	0	0.0027638
Indicator for historic contact with European explorer	21480	0.4342644	0.4956715	0	1
Indicator variable for historic integration into the colonial railway network	21480	0.4381285	0.4961687	0	1
Indicator for existence of city among ethnic group in 1400	21480	0.1216946	0.3269404	0	1
Population Density Average for Tribal Group	21480	4.261382	1.000939	0.6981347	7.266834
Log population density during the colonial period - from Murdock	18144	2.561385	1.309979	-4.273627	5.869599
Pre-colonial settlement patterns of ethnicity: from Ethnographic Atlas v30	20295	6.105395	1.250319	1	8
Pre-colonial juris. hierarchy beyond the local community: Ethnographic Atlas v33	20317	2.905744	0.913639	1	4
What is the highest level of education you have completed? (q90)	21402	3.077937	2.008928	0	9
What is your main occupation? (If unemployed, retired or disabled, what was your last main occupation	21392	15.75421	75.88902	0	995
What is your religion, if any? (q91)	21388	28.72817	106.3576	0	995
In general, how you describe your own present living condition?	21406	2.557647	1.203483	1	5

Table T – 3.3.4.3: Tabular Summary of Variables used in Equation 3.3.2

As can be seen in the summary statistics in Table T - 3.3.4.3 above there is a considerable range of variables (individual, district and tribal level) involved in the survey-based model (equation 3, specified above). These have been adopted to explore the connection between our in-designed split index and the core micro-level phenomena that have a theorized effect on the development of economic activity. In the section following this (section 4), we present and discuss the results from our estimations accordingly.

Section 3.4 – Estimation Results and Interpretations

3.4.1 – Tribal Demarcations and Economic Development using Nightlight density

From Table U - 3.4.1.1 and Table V - 3.4.1.2 below we find that our split index appears consistently negatively associated with the nightlight density measures, indicating that the extent of ethnic demarcations are linked with patterns of underdevelopment in the African subnational regions, as represented. Particularly, we find that areas of split-categories 1 and 3 (tribal homelands only demarcated at the intra-national frontier and those jointly demarcated at both intra-national and inter-national frontiers, respectively) appear to be the most affected categories from our split index measure. Both categories appear to have similar magnitudes and sign in their associations with underdevelopment depending on the estimation considered. These results are consistent even after controlling for key selected pre-colonial politico-economic characteristics, such as the tribe's level of political complexity and settlement patterns as provided by Murdock's (1959) anthropological study. While tribal areas of split index category 2 (tribal homelands affected by demarcations at the inter-national frontier only) also remain negative in their associations with development at this nighttime light density measure, but become statistically insignificant after controlling for pre-colonial politico-economic features. Estimations 1 thru 4 differ from 5 thru 8 because the latter explores the dependent variable in its original untransformed ordinal count form (0 - 63) and in the transformed continuous log of the ratio form but excluding the nighttime lights data points observed at the zero level. This is because we suspect that the zero-inflatedness of the nightlight measure increases

the imprecision or the associated standard errors post-transformation[76]. This can be confirmed by this zero-adjusted transformed dependent variable's relatively wider range values of -17.00 (lower range) and -1.58 (upper range) in Table R - 3.3.4.1 above, and the approximately twice as large standard-errors accompanying the associated estimates in estimations 1 thru 4 compared to the estimates of estimations 5 thru 8. With that in mind, we focus on interpreting the results from estimations 6 and 8 showing the split index categories without and with the two key additional pre-colonial tribal politico-economic controls. These two commonly used pre-colonial politico-economic features are the settlement patterns and political complexity of the corresponding tribal groups. Apart from the magnitude of the standard errors, the estimates also do not appear to differ significantly across the estimations from 1 thru 8 thus showing that the omission of the zero-region of the data does not seem to affect the result systematically. It also provides some indication confirming that the transformation of the nightlight ratio using the 0.01 sum-adjustment had quite a measurement effect on the nighttime lights density index measure.

From our estimation results below, we find that not only are different frontiers of tribal demarcation associated with differing local homeland developmental levels but that being demarcated at all appears to have a positive association with underdeveloped in such split homelands on the average. This is still very much in line with the findings in the literature relating underdevelopment in African sub-regions with ethnic homeland fragmentation (Michalopoulos & Papaioannou, 2011). As already highlighted, we show that categories 1 and 3 seem to be the most relatively negatively affected areas development-wise. While category 2, although remaining consistently negative, nonetheless turns statistically insignificant when pre-colonial politico-economic attributes are controlled for. This is a very interesting and a unique finding, as it unearths some deeper information about the driving circumstances of (under-) development in African sub-regions.

We find contrary to the literature that areas split inter-nationally seem to be the least underdeveloped compared to those split locally and those jointly split both locally and inter-nationally. This contradicts the literature on the scramble for Africa and its Euro-centric narrative, which proposes that pre-colonial,

[76] The transformation adopted is: $\frac{nightlights+0.01}{split\ area\ in\ KM}$

colonial, and post-colonial European politico-economic activities in Africa triggered a lot of the underdevelopment currently experienced on the continent (see Asiwaju, 1985; Alesina et al., 2011; Green, 2012). Our findings seem to support a more conservative counter-narrative: that it is as likely the case other locally brewed circumstances are adjacently at play igniting and fuelling these patterns of underdevelopment; and that these may actually be entirely independent of the European scramble and other related influences (see Fall, 2010 for a similar call for such analytical conservatism especially when administrative borders are central to the evidence). Although, it is empirically implausible to explore a substantial counterfactual of comparable [African] territories that were not colonised or entirely unaffected by European political influences. We attempt to demonstrate that the mainstream narrative of the non-positive economic development effects of the European facilitated artificial African national frontiers is not the entire story. This artificial borders argument tends to put the actions of the European countries involved in the Berlin conference of 1885 in the driver's seat of modern regional African economic underdevelopment (see Alesina et al., 2011)[77]. Fall (2010) attempts to analytically demonstrate in his article that this nature versus politics rhetoric has its unique empirical concerns as the naturalness (straight – artificial – *versus* squiggly – natural) of national borders are not entirely independent of the political or economic reality of the national territories involved. By adopting our approach, we sort of go beyond the nature of the border designs to focusing on the political actors involved. This in a way, inherently brings into the spotlight the narrative purporting that the European actors were not aware of the internal historic ethnic dynamics versus the post-independent political administrators who are more likely to understand these local ethnic dynamics. These historical ethnic dynamics have also not really changed that much between both periods under study, thus providing the link between both periods often exploited in this cross-boundary (ethnic frontier and national frontier intersections) type of studies.

As in estimation 8 below, we find that, holding all other covariates constant, areas affected solely by intra-national demarcations tend to be about 0.669 log-

[77] It is important to note here that the artificial borders storyline is not necessarily restricted to the African case; and it in fact researches within the narrative often pool examples from regions and countries across the globe. However, when it is applied to the African case, the border delimitations that result from the agreements of the 1885 Berlin conference and the proceeding OAU's Assembly of Heads of States and Government conference in Cairo of July 1964, we find ourselves apportioning substantial modern underdevelopment blames to the Europeans even following so many years of independent rule.

units or approximately 48.8%[78] darker in nighttime lights density per square kilometre compared to those not split at all; while those areas affected by both intra-national and inter-national demarcations tend to be 1.001 log-units or approximately 63.2% darker than the non-split tribal homelands in terms of nighttime lights density per square kilometre. These values are of similar dynamics to those found in estimation 4 from Table U - 3.4.1.1 below with 0.571 (or about 43.5% darker) and 0.728 (or about 51.7% darker) associated marginal drop in the log-count of the nighttime lights density index from the same group of demarcated tribal homelands compared to the non-split group, respectively. These results are robust to all the relevant controls (covariates) employed herein and are statistically significant at the 1% level of significance. Both these coefficients in most cases (categories 1 and 3) appear to be approximately twice as much as the coefficients in category 2; this can be observed consistently throughout the estimations reported in Table U - 3.4.1.1 and Table V - 3.4.1.2 below.

[78] Taking the exponent of the coefficient gives you these effects in odds ratios when you are dealing with the count-version of the dependent variable as in estimation 4. However, it also gives you the average marginal effect when you are dealing with the rest of the estimations that treat the dependent variable as a continuous variable (ratio form) and follows the log-linear functional form. The actual percentage effects as discussed herein are arrived at using the expression: $1 - e^{(coefficient)}$.

Dependent Variabes = Log (Stable Nightlight density to split tribal homeland area ratio or Stable Nightlight density - ordinal)					
		(1)	(2)	(3)	(4)
VARIABLES		lasnlite001	lasnlite001	rsnlite	rsnlite
Split Index Categories:	Category_1	-0.964***	-0.172	-1.040***	-0.571***
		'(0.245)	'(0.598)	'(0.143)	'(0.187)
	Category_2	-0.508	-0.00809	-0.732***	-0.328
		'(0.354)	'(0.748)	'(0.167)	'(0.218)
	Category_3	-1.016***	-0.136	-1.298***	-0.728***
		'(0.234)	'(0.575)	'(0.146)	'(0.189)
Mean Elevation (in 1000 KMs)		0.135	-0.221	-0.0772	-0.156
		'(0.360)	'(0.324)	'(0.202)	'(0.186)
Soil Productivity index		0.00281	0.00357	0.00138	0.00122
		'(0.00640)	'(0.00596)	'(0.00438)	'(0.00304)
Malaria Ecology index		-1.178	-1.612**	-0.947**	-0.745*
		'(0.769)	'(0.710)	'(0.454)	'(0.397)
Absolute Latitude		0.0389	0.00143	-0.00972	-0.00887
		'(0.0310)	'(0.0439)	'(0.0128)	'(0.0107)
Log (Split Tribal homeland land area to Unsplit Tribal land area ratio)		0.0684	0.179***	-0.268***	-0.252***
		'(0.0560)	'(0.0693)	'(0.0357)	'(0.0492)
Indicator for the presence of Diamond mines		0.393**	0.213	-0.051	-0.119
		'(0.170)	'(0.219)	'(0.104)	'(0.131)
Indicator for the presence of Petroleum wells		0.634***	0.523*	0.154	0.19
		'(0.240)	'(0.278)	'(0.135)	'(0.127)
Log ((1 + Inland water area) to split land area)		0.00488	0.00546	0.00488***	0.00492***
		'(0.00339)	'(0.00339)	'(0.00108)	'(0.000944)
Log (Population density)		0.548***	0.551***	0.371***	0.383***
		'(0.0426)	'(0.0554)	'(0.0550)	'(0.0488)
Split Tribal homeland's Distance to sea (in 1000 KMs)		-0.26	-0.298	-0.335*	-0.341*
		'(0.432)	'(0.408)	'(0.183)	'(0.188)
Split Tribal homeland's Distance to the National border (in 1000 KMs)		-12.13***	-13.67***	-13.86***	-14.50***
		'(1.992)	'(3.291)	'(1.522)	'(1.615)
Split Tribal homeland's Distance to the National capital (in 1000 KMs)		-0.35	-0.384	-0.045	-0.0235
		'(0.444)	'(0.624)	'(0.285)	'(0.328)
Pre-colonial Political Centralization Indicator (based on v33)			0.424**		0.195
			'(0.203)		'(0.129)
Pre-colonial Settlment Pattern Indicator (based v33)			0.0355		0.0931
			'(0.165)		'(0.0694)
	Main Intercept	-7.607***	-7.132***	-6.256***	-7.009***
		'(1.301)	'(1.844)	'(0.512)	'(0.517)
	ln(alpha)			-0.212*	-0.454***
				'(0.111)	'(0.158)
	Error-Variance adjustment at the Country-level			0.303***	0.172*
				'(0.106)	'(0.0907)
	Error-Variance adjustment at the State-level			0.177***	0.176***
				'(0.0256)	'(0.0396)
Estimation Type:					
OLS 3-way Robust Cluster (with Country-level FEs)		*YES*	*YES*	*NO*	*NO*
Mixed Effects Model Negative Binomial (with Ordinal Count DVs)		*NO*	*NO*	*YES*	*YES*
Observations		2,985	1,423	2,985	1,423
R-squared		0.399	0.44	.	.
ll		.	.	-7866	-3848
chi2		.	.	1234	720.8
p		.	.	0	0
Number of groups				50	48
Robust standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

Table U – 3.4.1.1: Estimation Output using the Nightlight density

Dependent Variabes = Log (Stable Night light-density to split tribal homeland area ratio) - ***zeros values excluded***					
VARIABLES		(5) lasnlite	(6) lasnlite	(7) lasnlite	(8) lasnlite
Split Index Categories:	Category_1	-1.006***	-0.644***	-0.987***	-0.669***
		'(0.125)	'(0.183)	'(0.126)	'(0.177)
	Category_2	-0.621***	-0.373	-0.624***	-0.318
		'(0.153)	'(0.262)	'(0.144)	'(0.210)
	Category_3	-1.409***	-1.007***	-1.376***	-1.001***
		'(0.123)	'(0.198)	'(0.120)	'(0.185)
Mean Elevation (in 1000 KMs)		0.0226	-0.0573	-0.0333	-0.0863
		'(0.126)	'(0.127)	'(0.119)	'(0.111)
Soil Productivity index		-0.000702	-0.000865	-0.00109	-0.0012
		'(0.00223)	'(0.00201)	'(0.00186)	'(0.00154)
Malaria Ecology index		-0.655**	-0.642	-0.599**	-0.464*
		'(0.320)	'(0.412)	'(0.290)	'(0.247)
Absolute Latitude		-0.0293**	-0.0243*	-0.0196**	-0.0208***
		'(0.0136)	'(0.0147)	'(0.00805)	'(0.00730)
Log (Split Tribal homeland land area to Unsplit Tribal land area ratio)		-0.454***	-0.484***	-0.452***	-0.477***
		'(0.0283)	'(0.0446)	'(0.0244)	'(0.0363)
Indicator for the presence of Diamond mines		-0.272***	-0.250*	-0.271***	-0.244*
		'(0.0852)	'(0.137)	'(0.0824)	'(0.127)
Indicator for the presence of Petroleum wells		-0.0723	0.0518	-0.0645	0.0667
		'(0.0814)	'(0.0885)	'(0.0721)	'(0.0831)
Log ((1 + Inland water area) to split land area)		0.00638***	0.00508***	0.00611***	0.00509***
		'(0.00185)	'(0.00123)	'(0.00144)	'(0.00105)
Log (Population density)		0.237***	0.250***	0.234***	0.251***
		'(0.0330)	'(0.0427)	'(0.0307)	'(0.0369)
Split Tribal homeland's Distance to sea (in 1000 KMs)		-0.184	-0.244	-0.221*	-0.218
		'(0.152)	'(0.180)	'(0.128)	'(0.134)
Split Tribal homeland's Distance to the National border (in 1000 KMs)		-15.29***	-13.96***	-15.33***	-14.77***
		'(1.376)	'(1.840)	'(1.394)	'(1.776)
Split Tribal homeland's Distance to the National capital (in 1000 KMs)		-0.107	0.104	-0.119	0.00444
		'(0.216)	'(0.277)	'(0.195)	'(0.215)
Pre-colonial Political Centralization Indicator (based on v33)			0.0776		0.0799
			'(0.112)		'(0.101)
Pre-colonial Settlment Pattern Indicator (based v33)			0.13		0.154**
			'(0.0888)		'(0.0729)
	Main Intercept	-5.129***	-6.175***	-5.957***	-6.683***
		'(0.535)	'(0.713)	'(0.291)	'(0.419)
	Error-Variance adjustment at the Country-level			-0.781***	-1.066***
				'(0.153)	'(0.187)
	Error-Variance adjustment at the State-level			-1.659***	-2.86
				'(0.211)	'(3.361)
Estimation Type:					
OLS 3-way Robust Cluster (with Country-level FEs)		*YES*	*YES*	*NO*	*NO*
Mixed Effects Linear Model (with Continuous DVs)		*NO*	*NO*	*YES*	*YES*
Observations		1,980	988	1980	988
R-squared		0.733	0.768	.	.
ll		.	.	-2407	-1152
chi2		.	.	5883	1895
p		.	.	0	0
Number of groups		.	.	50	48

Robust standard errors in parentheses

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

Table V – 3.4.1.2 – Estimation Output using Nightlight - omitting the zero values

3.4.2 – Tribal Demarcations and Economic Development using Average Roads Length

As conducted in Table U - 3.4.1.1 and Table V - 3.4.1.2 above, Table W - 3.4.2.1 below equally provides estimates using a country-level FE-OLS[79] specification with three-way cluster-adjusted standard errors (for estimations 9 and 10). It also includes a mixed-effects specification with error-variance adjustments at the district and tribal levels (for estimations 11 and 12). Given that the dependent variable in this subsection is the log of average access roads length to the represented split or non-split ethnic homeland land's area ratio and that there are no additional adjustments to the variable that could compromise its underlying distributions or significant variations in the estimated coefficients and standard errors, we go ahead and report the estimation results for estimations 10 and 12 respectively.

As observed in the previous subsection using nighttime light density as our dependent variable, we similarly notice in estimation 12 that compared to the non-split tribal homelands, those demarcated internally (intra-nationally) tend to have 0.888 log units (58.9%) shorter road lengths per square kilometre on the average after controlling for other equally relevant covariates. While compared to non-demarcated tribal homelands, those demarcated tribal homelands jointly affected by both inter-national and intra-national frontiers tend to have 1.351 log units (74.1%) shorter average roads to land ratio, this is with the other key covariates left unchanged.

And as in the previous subsection, tribal homelands affected by only inter-national boundary frontiers appear to comparatively be the least affected. Tribal homelands in this category are found to be associated with 0.648 log units (47.7%) marginally shorter average roads length per square kilometres compared to the non-demarcated tribal homelands, all things being equal. These values are not very far from the corresponding marginal effects shown in estimation 10[80] which are -0.953 log units (about 61.4% shorter roads) for category 1, -1.417 log units (about 75.8% shorter roads) for category 3 and -0.675 log units (about 49.1% less) for category 2 in terms of average roads length per square kilometres when two

[79] This means Fixed Effects Ordinary Least Squares specification.

[80] This is an alternative estimation using FE-OLS but with a 3-way (tribal, state, and country levels) cluster-robust standard errors.

key pre-colonial tribal politico-economic characteristics are additionally controlled for. While the estimated coefficients for categories 1 and 3 are found to be statistically significant at 1% level of significant, that of category 2 remains statistically significant at the 5% level of significance.

We also observe similar patterns of general reductions across and between the categories of the split index coefficients with the introduction of the two key pre-colonial politico-economic tribal attributes, earlier highlighted. This provides some indication that both measures of subnational development (average nighttime lights and average roads length) and the split index adopted herein are both linearly associated with levels of pre-colonial politico-economic complexity and that not controlling for them could provide misleading upwardly biased results. Our split index coefficients however remain both economically and statistically relevant even in the presence of all our adopted controls and covariates.

Looking at all the results from the three tables discussed so far (the 2 above and the 1 below) we find that contrary to the Euro-centric literature cross-national boundary delimitations demarcating historic ethnic homelands does not appear to provide the most significant isolated association with underdevelopment in Africa. Our results provide some evidence to suggest that historic tribal homelands also affected by subnational boundary delimitations are associated with specific local dynamics that cause these areas to be as underdeveloped as those most badly affected simultaneously by demarcations at multiple levels or frontiers (that is at both inter-country and subnational frontiers). Local or intra-national sources of such tribal homeland demarcations consistently appear to be associated with negative developmental outcomes. To investigate the underpinnings of this statistical association inherent in our findings a bit further, we look to the urban economics and local-trust literature. We proceed in section 3.4.3 below, by attempting to explore some conditional associations between our split categories and local trust or tribal discrimination by slightly augmenting Nunn & Wantchekon's (2011) base trust-specification.

Dependent Variabes = Log (Average Roads length to split tribal homeland ratio)					
VARIABLES		(9) lg1roads	(10) lg1roads	(11) lg1roads	(12) lg1roads
Split Index Categories:	Category_1	-1.216***	-0.953***	-1.100***	-0.888***
		'(0.140)	'(0.183)	'(0.119)	'(0.179)
	Category_2	-0.885***	-0.675**	-0.854***	-0.648**
		'(0.165)	'(0.308)	'(0.147)	'(0.264)
	Category_3	-1.660***	-1.417***	-1.556***	-1.351***
		'(0.119)	'(0.185)	'(0.102)	'(0.176)
Mean Elevation (in 1000 KMs)		-0.057	-0.119	-0.119	-0.126
		'(0.110)	'(0.166)	'(0.105)	'(0.149)
Soil Productivity index		-0.0039	-0.00514*	-0.00461	-0.00505**
		'(0.00358)	'(0.00278)	'(0.00288)	'(0.00210)
Malaria Ecology index		-0.362	-0.619	-0.361	-0.495
		'(0.264)	'(0.429)	'(0.236)	'(0.378)
Absolute Latitude		-0.0321**	-0.0262	-0.0237*	-0.023
		'(0.0148)	'(0.0266)	'(0.0130)	'(0.0196)
Log (Split Tribal homeland land area to Unsplit Tribal land area ratio)		-0.640***	-0.693***	-0.646***	-0.692***
		'(0.0327)	'(0.0432)	'(0.0246)	'(0.0315)
Indicator for the presence of Diamond mines		-0.445***	-0.437***	-0.431***	-0.424***
		'(0.118)	'(0.148)	'(0.101)	'(0.128)
Indicator for the presence of Petroleum wells		-0.511***	-0.476**	-0.435***	-0.390*
		'(0.177)	'(0.235)	'(0.150)	'(0.208)
Log ((1 + Inland water area) to split land area)		0.00168	0.000844	0.00213	0.00121
		'(0.00184)	'(0.00143)	'(0.00165)	'(0.00127)
Log (Population density)		-0.117***	-0.115***	-0.115***	-0.112***
		'(0.0190)	'(0.0284)	'(0.0165)	'(0.0265)
Split Tribal homeland's Distance to sea (in 1000 KMs)		-0.258	-0.309	-0.247	-0.219
		'(0.293)	'(0.382)	'(0.253)	'(0.329)
Split Tribal homeland's Distance to the National border (in 1000 KMs)		-17.46***	-15.91***	-17.47***	-16.16***
		'(2.256)	'(2.516)	'(2.183)	'(2.280)
Split Tribal homeland's Distance to the National capital (in 1000 KMs)		0.203	0.204	0.256	0.191
		'(0.204)	'(0.291)	'(0.179)	'(0.240)
Pre-colonial Political Centralization Indicator (based on v33)			-0.239**		-0.207**
			'(0.100)		'(0.0971)
Pre-colonial Settlment Pattern Indicator (based v33)			-0.00547		0.00996
			'(0.116)		'(0.0963)
	Main Intercept	-3.277***	-3.867***	-4.123***	-4.588***
		'(0.575)	'(1.071)	'(0.341)	'(0.580)
	Error-Variance adjustment at the Country-level			-0.0681	-0.0759
				'(0.115)	'(0.106)
	Error-Variance adjustment at the State-level			-1.047***	-1.198***
				'(0.134)	'(0.249)
Estimation Type:					
OLS 3-way Robust Cluster (with Country-level FEs)		*YES*	*YES*	*NO*	*NO*
Mixed Effects Linear Model (with Continuous DVs)		*NO*	*NO*	*YES*	*YES*
Observations		2,877	1,379	2,877	1,379
R-squared		0.716	0.755	.	.
ll		.	.	-4021	-1898
chi2		.	.	10129	7831
p		.	.	0	0
Number of groups				50	48
Robust standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

Table W – 3.4.2.1: Estimation Output using the Average Roads Length

3.4.3 – Tribal Demarcations and Local Trust or Ethnic Discrimination

In order to attempt to unravel some of the local mechanisms that could likely be behind why intra-nationally split tribes lean towards similar levels of underdevelopment as those jointly split at the intra-national and inter-national frontiers, we look to individual-level data provided by the Afrobarometer survey. We specifically consider the local trust and tribal discrimination channels as these have both been considerably examined in the literature - both theoretically (see the literature survey by Alesina & La Ferrara, 2005) and empirically (see for instance, Nunn & Wantchekon, 2011; Burgess et al, 2013; Hodler & Raschky, 2014). We choose to reassess our split index as the same categories adopted in the previous subsection above so as to maintain some methodological connection. With this approach, we intend to try to pinpoint some correlated avenues via which underlying local dynamics (such as trust and tribal discrimination) filter into underdevelopment at these subnational regions.

Excluding category 3, the estimated coefficients for our split categories of interest, as regards the trust regressions, appear to be very similar in magnitudes and standard errors. Looking at estimations 14 thru 18 in Table X - 3.4.3.1 below, we find that split category 1 (tribal homelands demarcated at the intra-national frontier) appears to be consistently negatively associated with our selected trust measures (trust for: local councils, adjacently interacting tribal groups, own-tribal groups, relatives and neighbours) and the coefficients appear to be within close proximity of each other ranging from -0.226 log odds (estimation 14) to -0.292 log odds (estimation 16). Category 1 appears to be the dominantly consistent category driving the outputs we observe in our tribal homeland level estimation results in Table U - 3.4.1.1, Table V - 3.4.1.2 and Table W - 3.4.2.1 above, because of the weakening trust base and the relatively higher prevalence of perceived ethnic discrimination accompanying intra-nationally split tribal homelands. Focusing on estimations 14 thru 18, we find that respondents from tribal homelands in split category 1 are marginally on the average approximately 20% to 25% less likely to have higher trust levels for their local councils compared to respondents from non-demarcated tribal homelands. This is statistically significant at the 5% significance level and remains so when you consider other measures of trust at the same demarcation echelons.

Dependent Variabes = Selected Ordinal Categorical Variables measuring a Respondent's attitude towards Tribe related Trust and Favouritism							
		(13)	(14)	(15)	(16)	(17)	(18)
		robu_61	robu_11	robu_21	robu_31	robu_41	robu_51
VARIABLES		tribal_favouritsm	trust_local_council	inter_group_trust	intra_group_trust	trust_relatives	trust_neighbors
Split Index Categories:	Category_1	0.570***	-0.226**	-0.284**	-0.292**	-0.268**	-0.251**
		(0.167)	(0.0971)	(0.114)	(0.129)	(0.128)	(0.124)
	Category_2	0.492**	0.201	0.102	0.0842	-0.194	0.136
		(0.224)	(0.192)	(0.195)	(0.191)	(0.192)	(0.202)
	Category_3	0.545***	-0.106	-0.136	-0.160	-0.308**	-0.196
		(0.168)	(0.101)	(0.116)	(0.132)	(0.123)	(0.122)
ln(1+Atlantic+Indian Slave Exports) by Tribal group		0.0177	-0.0461**	-0.0137	-0.0467**	-0.105***	-0.0518***
		(0.0234)	(0.0195)	(0.0199)	(0.0202)	(0.0216)	(0.0191)
Respondet's Perception of Ethnic Group's Economic Condition (q80a)		0.359***	-0.0855***	-0.0729***	-0.0473**	-0.0551***	-0.0553***
		(0.0256)	(0.0224)	(0.0206)	(0.0223)	(0.0210)	(0.0208)
Respondet's Perception of Ethnic Group's Poltical Influence (q80a)		0.187***	-0.120***	-0.0139	-0.0527**	-0.00699	-0.0255
		(0.0331)	(0.0241)	(0.0222)	(0.0224)	(0.0217)	(0.0213)
District-level ethnic fractionalization		0.0977	0.269**	0.361**	0.267*	0.165	0.178
		(0.165)	(0.131)	(0.143)	(0.141)	(0.137)	(0.142)
Proportion of ethnic group in district		0.192*	0.204**	0.0151	0.0578	0.165*	0.246**
		(0.115)	(0.103)	(0.101)	(0.106)	(0.0996)	(0.108)
Tribe's Average Malaria Ecology		0.0229**	-0.00729	-0.0153**	-0.0195***	-0.00700	-0.0212***
		(0.00914)	(0.00689)	(0.00722)	(0.00684)	(0.00787)	(0.00660)
Total Catholic + Protestant mission per land area		-10.06	33.33	-160.5**	-266.8***	15.22	-3.297
		(94.85)	(77.51)	(78.17)	(85.08)	(85.43)	(85.79)
Indicator for historic contact with European explorer		0.0222	0.278***	0.151**	0.116*	0.0695	0.106
		(0.0819)	(0.0709)	(0.0685)	(0.0675)	(0.0680)	(0.0651)
Indicator variable for historic integration into the colonial railway network		0.0162	-0.145**	-0.101	-0.0572	-0.125	-0.108
		(0.0900)	(0.0699)	(0.0759)	(0.0775)	(0.0760)	(0.0680)
Indicator for existence of city among ethnic group in 1400		-0.187	-0.0857	0.172	0.224**	0.617***	0.441***
		(0.128)	(0.104)	(0.111)	(0.0995)	(0.109)	(0.0957)
Population Density Average for Tribal Group		-0.00285	-0.0909*	-0.135**	-0.229***	-0.170***	-0.271***
		(0.0723)	(0.0521)	(0.0523)	(0.0545)	(0.0598)	(0.0561)
Log population density during the colonial period - from Murdock		0.101**	-0.0503	0.0287	0.0429	-0.0184	-0.0222
		(0.0471)	(0.0336)	(0.0337)	(0.0369)	(0.0352)	(0.0374)
Pre-colonial juris. hierarchy beyond the local community (v33)		-0.0275	-0.0388	-0.0577**	-0.0743**	-0.0853***	-0.0794***
		(0.0388)	(0.0279)	(0.0288)	(0.0296)	(0.0311)	(0.0292)
Constant							
Error-Variance adjustment at the District-level		0.438***	0.141***	0.270***	0.225***	0.152***	0.222***
		(0.0604)	(0.0219)	(0.0332)	(0.0299)	(0.0302)	(0.0295)
Error-Variance adjustment at the Tribe-level		0.0536	0*	0**	0	0	0
		(0.0438)	(0)	(0)	(0)	(0)	(0)
Observations		14,583	14,550	15,195	15,308	15,367	15,341
Number of groups		1,172	1,172	1,174	1,177	1,178	1,178
ll		-11839	-14844	-15589	-15375	-13556	-15361
chi2		2092	3281	2806	2883	3140	8306
p		0	0	0	0	0	0

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

Table X – 3.4.3.1: Estimation Output testing the local Trust and Tribal Discrimination channels

Interestingly, our results show that for local trust measures, the tribal homelands demarcated only at the inter-national frontier (split category 2) appear to be mostly positively associated with trust (excluding the "trust for relatives" index) and are consistently statistically insignificant. We also find that while the homelands jointly split at both intra-national and inter-national frontiers (split category 3) have consistently negative coefficients, they are equally predominantly statistically insignificant, excluding the coefficient associated with the "trust for relatives" measure in estimation 17 (-.308 log odds or approximately 26.5% less likely to perceive higher trust categories), which is not too far off in magnitude to category 1 (-0.194 log odds or approximately 17.8% less likely to have higher trust levels for their relatives). This indicates that the local trust mechanism or channel can provide a convincing rationalisation for why lower levels of economic development are at least associated with tribal homelands demarcated at this intra-national frontier as observed in the earlier estimation tables discussed in the previous subsections above.

Fascinatingly, estimation 13 offers our discussions with an additional conduit providing some local level mechanism via which our split categories may be associated with underdevelopment at the sub-regional levels. We find that respondents from subnational homelands that are internally demarcated (split category 1) and those that were jointly demarcated locally and inter-nationally (split category 3) are respectively about 76.8% and 72.5% more likely to be associated with higher discrimination levels against them by their governments compared to respondents from un-split tribal homelands. This links quite smoothly with the literature suggesting that ethnically segregated groups are linearly associated with relatively weaker institutions and less public good allocations compared to unsegregated tribal groups (Burgess et al, 2013 and Hodler & Raschky, 2014 discuss similar findings). The coefficients 0.57 and 0.545 (for categories 1 and 3, respectively) are both statistically significant at the 1% level. While, as already predominantly observed, category 2 (tribal homelands demarcated at the inter-national frontier only) has the expected positive sign in this instance, remains relatively smaller in magnitude and drops a notch in statistical significance as well.

What our results seem to demonstrate at this stage is that the local trust mechanism is a valid conduit for understanding the underlying dynamics explaining why locally demarcated tribal homelands have consistently lower levels of

development (as measure by nighttime lights and average roads length) compared to non-demarcated groups. While the ethnic favouritism measure seems to provide a valid reinforcement explaining why locally demarcated groups (category 1) and both intra-nationally and inter-nationally demarcated groups (category 3) appear to have lower development levels as measured in our study herein. What is equally consistent is that local dynamics appear to have equally or sometimes more damning developmental outcomes compared to the often-mentioned pressures exerted by the European scramble for territory and resources in African. Also, that focusing on such Euro-centric narratives may not provide an apt or comprehensive treatment (that is treatment in scope and depth) of studies on African development.

3.4.4 – Concluding Remarks

In summary our findings demonstrate that the extent of tribal homeland demarcations in African is detrimentally associated with corresponding developmental levels measured in nighttime lights and average roads length. As earlier posited herein, we show that the European scramble for African territories and resources was not the singular source of underdevelopment in the sub-regions of the continent. Essentially we see that intra-nationally demarcated tribal homelands appear to negatively outperform those tribal homelands solely affected by inter-national demarcations. While those jointly demarcated on both intra-national and inter-national frontiers (our most affected tribal homelands) appear to underperform as bad as the intra-national only demarcated groups. This indicates, as we further demonstrated, that local dynamics were as instrumental in realising such developmental underperformances as those inter-national dynamics, which are often researched in the African scramble literature such as conflict and etc. (see Michalopoulos & Papaioannou, 2011, for instance).

Our results also show that our split index categories are related with two key local dynamics underpinning low developmental outcomes in Africa. We find that the local trust dynamics help explain why intra-nationally demarcated tribal homelands are as underdeveloped as the most split category in our dataset (that is, tribal homelands demarcated both intra-nationally and inter-nationally). While, the local dynamics of ethnic discrimination provides an avenue to further

buttress why both categories 1 and 3 appear similar in magnitude as regards their economic developmental outcomes.

3.4.5 – Avenues for Further Research

This study herein is mostly demonstrative and even while it tries to show how our own-designed split index connects with developmental levels (measured using nighttime lights and average roads length), it does not provide any direct causal statistical framework. Further research that attempts to make such direct causal statistical relationships between our split index (or split categories) and developmental outcomes in these African sub-regions would be most informative. Such a study would allow one to obtain coefficient estimates demonstrating the direct extent or magnitude of these connections.

One way in which causality can be assessed in such a framework as utilised herein would be to attempt to extend the research design beyond the African region to include regions and sub-regions across the world that have been affected by European type colonisation of similar scale. The objective thence would be to group the territories in similar spatial units (or political administrative units), experiences and other relevant characteristics. After which a propensity score matching methodology could be applied in an attempt to obtain average treatment effects on economic activity resulting from our designed classifications of political ethnic territory demarcations. One apparent challenge is that valid counterfactuals embedded in the propensity score matching procedure might not be readily available as there are very different experiences and inherent characteristics typifying the sample units from such a design. For instance, the ethnic dynamics of African tribal homelands are historically more recent than anywhere else in the world. Also, almost all of Africa [apart from Ethiopia and Liberia] was affected by European colonialism so obtaining appropriate counterfactuals within the African continent itself is near impossible[81].

Additionally but not so notably, the data structure and type for each empirical segment of this study is differentiated. This resultantly provides evidence applicable to both homeland and individual (respondent) unit levels. It is pervasively known in the literature that results from varying data aggregation

[81] An appreciation to one of my examiners, whose name I have chosen to withhold at this stage for anonymity reasons, for raising this point at my viva defence.

levels tend to range quite widely and are not easily clearly comparable. Adopting a dataset, in the future, with a uniform aggregation structure studying ethnic demarcations in a similar fashion and integrating causal channels similar to the trust and ethnic discrimination adopted herein would be way more informative than one conducted using datasets of varying aggregation structure or composition. In fact with such a dataset, it would be easier to explore more direct causal pathways for assessing the impact of ethnic demarcations on economic activity as both right hand side controls, instruments (mechanisms) and economic activity would be varying at the same level of variation.

References

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401.

Adeyemi, O. O. (2013). The Politics of States and Local Governments Creation in Nigeria: An Appraisal. *European Journal of Sustainable Development*, *2*(3), 155–174. doi:10.14207/ejsd.2013.v2n3p155

Alesina, A., Baqir, R., & Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, *114*(4), 1243–1284. doi:10.1162/003355399556269

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, *8*, 155–194.

Alesina, A., Easterly, W., & Matuszeski, J. (2011). Artificial States. *Journal of the European Economic Association*, *9*(2), 246–277. doi:10.1111/j.1542-4774.2010.01009.x

Alesina, A. F., Michalopoulos, S., & Papaioannou, E. (2012). *Ethnic Inequality* (No. 18512 - NBER Working Paper Series).

Alesina, A., & La Ferrara, E. (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature*, *43*(3), 762–800. doi:doi:10.1257/002205105774431243

Alesina, A., Ozler, S., Roubini, N., & Swagel, P. (1996). Political Instability and Economic Growth. *Journal of Economic Growth*, *1*, 189–211.

Alesina, A., & Perotti, R. (1996). Income Distribution, Political Instability, and Investment. *European Economic Review*, *40*(6), 1203–1228. doi:10.1016/0014-2921(95)00030-5

Alesina, A., & Zhuravskaya, E. (2011). Segregation and the Quality of Government in a Cross-Section of Countries. *American Economic Review*, *101*(5), 1872–1911. doi:10.1257/aer.101.5.1872

Asiwaju, A. I. (1985). *Partitioned Africans: Ethnic Relations across Africa's International Boundaries, 1884-1984*. Palgrave Macmillan.

Auer, R. A. (2013). Geography, Institutions, and the Making of Comparative Development. *Journal of Economic Growth*, *18*(2), 179–215. doi:10.1007/s10887-013-9087-z

Barro, R. J. (1996). *Determinants of Economic Growth: A Cross-Country Empirical Study* (No. 5698). *Foreign Affairs* (Vol. 76). doi:10.2307/1061363

Besley, T., & Reynal-Querol, M. (2014). The Legacy of Historical Conflict: Evidence from Africa. *American Political Science Review*, *108*(2), 319–336. doi:10.1017/S0003055414000161

Bloom, D. E., Sachs, J. D., Collier, P., & Udry, C. (1998). Geography, Demography, and Economic Growth in Africa. *Brookings Papers on Economic Activity*, *1888*(2), 207–295. doi:10.2307/2534695

Bowen, J. R. (1996). The Myth of Global Ethnic Conflict. *Journal of Democracy*, *7*(4), 3–14. doi:10.1353/jod.1996.0057

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., & Miquel, G. P. (2013). *The Value of Democracy: Evidence from Road Building in Kenya* (No. 19398 - NBER Working Paper Series).

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust Inference with Multiway Clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249. doi:10.1198/jbes.2010.07136

Collier, P. (2007). *The Bottom Billion: Why the Poorest Countries are Failing and What Can Be Done About It*. Oxford University Press.

Crescenzi, R., & Percoco, M. (Eds.). (2013). *Advances in Spatial Science: Geography, Institutions and Regional Economic Performance*. *Springer*. Springer. doi:10.1007/978-3-642-17940-2

Dippel, C. (2014). Forced Coexistence and Economic Development: Evidence from Native American Reservations. *Econometrica*, *82*(6), 2131–2165. doi:10.3982/ECTA11423

Dollery, B., Crase, L., & Johnson, A. (2006). *Australian Local Government Economics*. UNSW Press.

Easterly, W., & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics*, *112*(4), 1203–1250. doi:Doi 10.1162/003355300555466

Eifert, B., Miguel, E., & Posner, D. N. (2010). Political Competition and Ethnic Identification in Africa. *American Journal of Political Science*, *54*(2), 494–510. doi:10.1111/j.1540-5907.2010.00443.x

Englebert, P. (2000). Pre-Colonial Institutions, States, and in Economic Development Tropical Africa. *Political Research Quarterly*, *53*(1), 7–36. Retrieved from http://www.jstor.org/stable/449244

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. doi:10.1177/001041402237944

Faguet, J.-P. (2004). Does Decentralization Increase Government Responsiveness to Local Needs? *Journal of Public Economics*, *88*(3-4), 867–893. doi:10.1016/S0047-2727(02)00185-8

Fall, J. J. (2010). Artificial states? On the Enduring Geographical Myth of Natural Borders. *Political Geography*, *29*(3), 140–147. doi:10.1016/j.polgeo.2010.02.007

Fearon, J. D. (2003). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth*, *8*, 195–222. doi:10.1023/A:1024419522867

Fenske, J. (2014). Ecology, Trade, and States in Pre-Colonial Africa. *Journal of the European Economic Association*, *12*(3), 612–640. doi:10.1111/jeea.12042

Gennaioli, N., & Rainer, I. (2007). The Modern Impact of Precolonial Centralization in Africa. *Journal of Economic Growth*, *12*(June), 185–234. doi:10.1007/s10887-007-9017-z

Glaeser, E. L., La-Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2004). Do Institutions Cause Growth? *Journal of Economic Growth*, *9*(June), 271–303. doi:10.1023/B:JOEG.0000038933.16398.ed

Green, E. (2012). On the Size and Shape of African States. *International Studies Quarterly*, *56*(2), 229–244. doi:10.1111/j.1468-2478.2012.00723.x

Grier, K., & Tullock, G. (1989). An Empirical Analysis of Cross-National Economic Growth, 1951–1980. *Journal of Monetary Economics*, *24*(2), 259–276.

Griffiths, I. (1986). The Scramble for Africa : Inherited Political Boundaries. *The Geographical Journal*, *152*(2), 204–216. doi:10.2307/634762

Hodler, Roland; Raschky, P. A. (2014). Regional Favouritsm. *The Quarterly Journal of Economics*, *129*(2), 995–1033.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. W. (1999). The Quality of Government. *Journal of Law, Economics, and Organization*, *15*(1), 222–279. doi:10.1093/jleo/15.1.222

Litvack, J., Ahmad, J., & Bird, R. (1998). *Fiscal Decentralization in Developing Countries*. Washington, DC: The World Bank. doi:10.1017/CBO9780511559815

Lowes, S., Nunn, N., Robinson, J. a, & Weigel, J. (2015). Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT). *American Economic Review: Papers & Proceedings*, *105*(5), 340–345. doi:10.1257/aer.p20151075

MacKinnon, D., Cumbers, A., & Pike, A. (2009). Evolution in Economic Geography: Institutions, Political Economy, and Adaptation. *Economic Geography*, *85*(2), 129–150.

Michalopoulos, S. (2012). The Origins of Ethnolinguistic Diversity. *American Economic Review*, *18*(9), 1508–1539. doi:10.1016/j.micinf.2011.07.011.Innate

Michalopoulos, S., & Papaioannou, E. (2011). *The Long-Run Effects of the Scramble for Africa* (No. 17620). *NBER Working Paper Series*. doi:10.2139/ssrn.1696195

Michalopoulos, S., & Papaioannou, E. (2012). *National Institutions and African Development: Evidence from Partitioned Ethnicities* (No. 18275 - NBER Working Paper Series).

Michalopoulos, S., & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, *81*(1), 113–152. doi:10.3982/ECTA9613

Michalopoulos, S., & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, *129*(1), 151–213. doi:10.1093/qje/qjt029.Advance

Miguel, E., & Gugerty, M. K. (2005). Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. *Journal of Public Economics*, *89*(11-12), 2325–2368. doi:10.1016/j.jpubeco.2004.09.004

Montalvo, J., & Reynal-Querol, M. (2005). Ethnic diversity and Economic Development. *Journal of Development Economics*, *76*(2), 293–323. doi:10.1016/j.jdeveco.2004.01.002

Montalvo, J., & Reynal-Querol, M. (2005). Ethnic polarization, Potential Conflict, and Civil Wars. *American Economic Review*, *95*(3), 796–816.

Murdock, G. P. (1959). *Africa: Its Peoples and Their Cultural History*. McGraw-Hill Inc.

Naudé, W. A. (2004). The effects of policy, institutions and geography on economic growth in Africa: An econometric study based on cross-section and panel data. *Journal of International Development*, *16*(6), 821–849. doi:10.1002/jid.1129

Nunn, N. (2008). The Long-term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics*, *123*(1), 139–176.

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

Nunn, N., & Wantchekon, L. (2011). The Slave trade and the Origins of Mistrust in Africa. *American Economic Review*, *101*(7), 3221–3252. doi:10.1257/aer.101.7.3221

Saltman, M. (Ed.). (2002). *Land and Territoriality*. *Berg Ethnicity and Identity Series*. Berg Publishers. doi:10.1016/S0962-6298(03)00090-8

Slack, E., & Bird, R. (2013). *Merging Municipalities: Is Bigger Better?* (No. 14 - IMFG Papers on Municipal Finance and Governance).

Snijders, T., & Bosker, R. (2012). *Multilevel Analysis: An Introduction to Basic and Applied Multilevel Analysis* (2nd ed.). Sage Publications.

Spolaore, E., & Wacziarg, R. (2013). How Deep Are the Roots of Economic Development? *Journal of Economic Literature*, *51*(2), 325–369. doi:10.3386/w18130

Vande, P. T. (2012). Ethnicity and the Politics of State Creation in Nigeria. *European Scientific Journal*, *8*(16), 33–51.

Wimmer, A., & Cederman, L.-E. (2009). Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set. *American Sociological Review*, *74*(2), 316–337. doi:10.1177/000312240907400208

Appendix

Appendix 3.1 – List of Variables Used in the Regression Analyses

Full Variable Name	Units of Measure	Variable(s) Used from Original Dataset	Source
Log (Stable Night light-density to split homeland ratio)	Key variable is a Polygonal-average of Index between 0 and 63	GRID values	Version 4 DMSP-OLS Nighttime Lights Time Series: http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
Log (Average Roads length to split tribal homeland ratio)	Key variable is a Polygonal-average in 1000 Kilometres	LENGTH_KM	Global Roads Open Access Data Set (gROADS), v1 (1980–2010): http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1
Index of the extent of ethnic homeland splitting (ordinal)	This Index ranges between 0 and 3. It is uniquely designed by this study. Higher values indicate that the source of boundary demarcation for the tribal homeland is approaching multiple administrative frontiers, where at least one of such frontiers is inter-national	N/A	Obtained from the Geographic intersection of the Provincial-level administrative boundaries - 'RWDB2 Subnational-Ad2 Polygonal Boundaries Map' available at the United Nations' FAO website: http://ref.data.fao.org/map?entryId=936f7280-b06d-11db-8922-000d939bc5d8&tab=metadata ; and the Tribal Homeland Boundaries based on the work of Peter Murdock (1959) - available at http://scholar.harvard.edu/files/nunn/files/murdock_shapefile.zip?m=1364315353
Log (Population density)	Polygonal-average: Discrete Count Variable	grid_code	UNEP/GRID - Sioux Falls Dataset (Africa Population Distribution Database): http://na.unep.net/siouxfalls/datasets/datalist.php
Split Tribal homeland's Distance to sea (in 1000 KMs)	Nearest distance for each Polygonal-unit in 1000 Kilometres	From a Geodatabase of Polylines	Water Polygons: Available at the Open Street Map http://openstreetmapdata.com/data/water-polygons
Split Tribal homeland's Distance to the National border (in 1000 KMs)	Nearest distance for each Polygonal-unit in 1000 Kilometres	From a Shapefile with National Boundary Outlines	National Administrative Boundaries, v3 (2000), Gridded Population of the World (GPW), v3: http://sedac.ciesin.columbia.edu/data/set/gpw-v3-national-admin-boundaries/data-download
Split Tribal homeland's Distance to the National capital (in 1000 KMs)	Centroid distance for each Polygonal-unit in 1000 Kilometres	From a Shapefile of National Capital centroid points	ArcGIS Data Gallery - World Cities: http://www.arcgis.com/home/item.html?id=dfab3b294ab24961899b2a98e9e8cd3d
Mean Elevation (in 1000 KMs)	Polygonal-averages Measured in 1000 Kilometres	From Data provided in ArcView ESRI Grid Format	National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colo.: http://nelson.wisc.edu/sage/data-and-models/atlas/maps.php?datasetid=28&includerelatedlinks=1&dataset=28
Soil Productivity index	Polygonal-average of Index between 1 and 97	GRID values representing Soil Quality	TERRASTAT I, FAO Soil Productivity Index

Malaria Ecology index	Polygonal-average of Index between a lowest value of 0 and highest value of 38.081. This is also used in the Survey regressions and in that case is the average for the respondent's associated Tribe	GRID values representing Incidence rates	Malaria Ecology Map based on work by Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004): available at http://www.earth.columbia.edu/people/gmccord/sitefiles/file/malaria_ecology.zip
Absolute Latitude	Non-negative values of the Latitude of the Ethnic Polygon's associated centroid	Latitude	Obtained from the 'RWDB2 Subnational-Ad2 Polygonal Boundaries Map' by the United Nations' FAO
Log (Split Tribal homeland land area to Unsplit Tribal land area ratio)	Log of the polygonal-average	N/A	Obtained from the Geographic intersection of the Provincial-level administrative boundaries - 'RWDB2 Subnational-Ad2 Polygonal Boundaries Map' obtained by the United Nations' FAO; and the Tribal Homeland Boundaries based on the work of Peter Murdock (1959) - available at http://scholar.harvard.edu/files/nunn/files/murdock_shapefile.zip?m=1364315353
Log Split Land Area	Polygonal-average in 1000 Kilometres	N/A	
Indicator for the presence of Diamond mines	Polygonal Dummy variable	NAME - representing the individual sites	Petroleum Dataset from the Peace Research Institute Oslo (PRIO) website: https://www.prio.org/Data/Geographical-and-Resource-Datasets/Petroleum-Dataset/
Indicator for the presence of Petroleum wells	Polygonal Dummy variable	NAME - representing the individual sites	Global Lakes and Wetlands Database: Lakes and Wetlands Grid (Level 3): http://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3
Log ((1 + Average Inland water area))	Polygonal-average measured in Kilometres	N/A	Global Lakes and Wetlands Database: Lakes and Wetlands Grid (Level 3): http://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3
Historic Ethnic Group's Political Centralization Index	Indicators for each of the 3 Categories	v32	Obtained from the Tribal Homeland Boundaries based on the work of Peter Murdock (1959) - available at http://eclectic.ss.uci.edu/~drwhite/worldcul/world.htm
Historical Ethnic Group's Settlement Pattern	Indicator variables are used for each of the 8 Categories	v30	
How much do you trust your local council? (q55d)	Ordinal Categories ranging between 0 and 3 (higher numbers represent higher outcomes	q55d	These variables are available from the Afrobarometer Merged Round 3 Data for 18 Countries in 2005 at the Afrobarometer website: http://www.afrobarometer.org/data/merged-round-3-data-18-countries-2005
How much do you trust people from other ethnic groups? (q84d)	Ordinal Categories ranging between 0 and 3 (higher numbers represent higher outcomes	q84d	

How much do you trust people your own ethnic group? (q84c)	Ordinal Categories ranging between 0 and 3 (higher numbers represent higher outcomes	q84c	
How much do you trust your relatives? (q84a)	Ordinal Categories ranging between 0 and 3 (higher numbers represent higher outcomes	q84a	
How much do you trust your neighbours? (q84b)	Ordinal Categories ranging between 0 and 3 (higher numbers represent higher outcomes	q84b	
How often is your ethnic group treated unfairly by the government? (q81)	Ordinal Categories ranging between 0 and 3 (higher numbers represent higher outcomes	q81	
Respondent's Perception of Ethnic Group's Relative Economic Conditions (q80a)	Ordinal Categories ranging between 0 and 5 (higher numbers represent higher and worse outcomes	q80a	
Respondent's Perception of Ethnic Group's Relative Political Influence (q80a)	Ordinal Categories ranging between 0 and 5 (higher numbers represent higher and worse outcomes	q80b	
What is the highest level of education you have completed? (q90)	Indicator variables are used for each of the 9 Categories	q90	
What is your main occupation? (If unemployed, retired or disabled, what was your last main occupation	Indicator variables are used for each of the 27 Categories present	q95	
What is your religion, if any? (q91)	Indicator variables are used for each of the 18 Categories present	q91	
In general, how you describe your own present living condition?	Indicator variables are used for each of the 5 Categories present	q4b	
ln(1+Atlantic+Indian Slave Exports) by Tribal group	N/A	ln_exports	
District-level ethnic fractionalization	Observed at the Tribal level and ranges between 0 and 1	district_ethnic_frac	These variables are available in dataset in a replication archive from Nathan Nunn's (Nunn, AER 2011) Research Datasets website at http://scholar.harvard.edu/files/nunn/files/nunn_wantchekon_aer_2011_replication_files_0.zip?m=1364315322
Proportion of ethnic group in district	Observed at the Tribal level and ranges between 0 and 1	frac_ethnicity_in_district	
Total Catholic + Protestant mission per land area	Discrete values at the Tribal level	total_missions_area	

Indicator for historic contact with European explorer	Dummy Variable at the Tribal level	explorer_contact
Indicator variable for historic integration into the colonial railway network	Dummy Variable at the Tribal level	railway_contact
Indicator for existence of city among ethnic group in 1400	Dummy Variable at the Tribal level	cities_1400_dum
Population Density Average for Tribal Group	N/A	lpopd
Log population density during the colonial period - from Murdock	N/A	ln_init_pop_density

Table Y – Appendix-Table 3.1: Variable Definition of Variables used in Regression Analyses

Appendix 3.2 – List of African Countries Sampled in the Regression Analyses

List of sampled African Countries Used in the Analysis		
Algeria	Ghana	Somalia
Angola	Guinea	South Africa
Benin	Guinea-Bissau	Sudan
Botswana	Kenya	Swaziland
Burkina Faso	Lesotho	Tanzania
Burundi	Liberia	Togo
Cameroon	Libya	Tunisia
Central African Republic	Madagascar	Uganda
Chad	Malawi	Zambia
Comoros	Mali	Zimbabwe
Congo	Mauritania	
Congo, Democratic Republic of the	Mayotte	
Cote d'Ivoire	Morocco	
Djibouti	Mozambique	
Egypt	Namibia	
Equatorial Guinea	Niger	
Eritrea	Nigeria	
Ethiopia	Rwanda	
Gabon	Senegal	
Gambia, The	Sierra Leone	

Table Z – Appendix-Table 3.2: List of Sampled African Countries involved in the Regression Analyses

Appendix 3.3 – List of Sampled Tribes Involved in the Core Regression Analyses (Subsection 3.3.1)

List of the Sampled African Tribes Used in the Analysis					
ABABDA	BUSANSI	HLENGWE	KWENA	NAJI	SILA
ABARAMBO	BUTAWA	HOLO	KWERE	NALU	SINZA
ABE	BUYE	HOLOHOLO	KWESE	NAMA	SIRTICANS
ACHOLI	BWAKA	HUNDE	LABWOR	NAMIB	SIWA
ADAMAWA	CAPE HOTTENTOT	HURUTSHE	LAGUAT	NAMSHI	SOGA
ADANGME	CHAAMBA	IBIBIO	LAKA (ADAMAWA	NANDI	SOKO
ADARAWA	CHAGA	IBO	LAKA (NDEBELE)	NARON	SOKOTO
ADELE	CHAKOSSI	IDIO	LALA	NAUDEBA	SOLIMAN
AFAR	CHAMBA	IDOMA	LAMBA	NDAKA	SOMBA
AFO	CHEWA	IFE	LAMBYA	NDAMBA	SOMRAI
AFUSARE	CHIGA	IFORA	LANDUMA	NDAU	SONGE
AHAGGAREN	CHOKWE	IGALA	LANGO	NDEBELE	SONGHAI
AJUKRU	CHOPI	IGBIRA	LELE	NDEMBU	SONGO
AKPOSO	CHUABO	IJAW	LENDU	NDOB	SONGOLA
AKYEM	COMORIANS	IJEBU	LENGOLA	NDOGO	SONGOMENO
ALAGYA	CYRENAICANS	IKASA	LENJE	NDOKO	SONINKE
ALGERIANS	DAFI	ILA	LESE	NDOMBE	SONJO
ALUR	DAGARI	IMRAGEN	LI	NDORO	SOTHO
AMARAR	DAGOMBA	INGASSANA	LIGBI, DEGHA (SE)	NDUKA	SUAFA
AMBA	DAGU	IRAMBA	LIKA	NEFUSA	SUBIA
AMBO	DAKA	IRAQW	LILSE	NEN	SUK
AMER	DAKHLA	ISHAAK	LIMBA	NGALA	SUKU
AMHARA	DAMA	ISOKO	LIPTAKO	NGAMA	SUKUMA
ANA	DAN	ITSEKIRI	LISI	NGANDU	SUMBWA
ANAG	DARASA	ITTU	LOBI	NGBANDI	SUNDI
ANGAS	DARI	IWA	LOGO	NGBELE	SURI
ANTAISAKA	DAZA	IYALA	LOKELE	NGE	SUSU
ANTANDROY	DEFORO	JALO	LOKO	NGERE	SWAZI
ANTESSAR	DEKAKIRI	JANJERO	LOMWE	NGINDO	TABWA
ANUAK	DELIM	JARAWA	LONGUDA	NGIZIM	TAGALI
ANYANG	DENDI	JEBALA	LOTUKO	NGOMBE	TAJAKANT
ANYI	DERA	JEN	LOVEDU	NGONGO	TALODI
ARAD	DIALONKE	JERAWA, CHAWAI (SW)	LOZI	NGONYELU	TAMA
ARAGO	DIAN	JERBA	LUAPULA	NGUMBA	TANALA
ARI	DIDA	JERID	LUBA	NGUMBE	TANGALE
ARUSI	DIDINGA	JIBU	LUCHAZI	NGURU	TANNEKWE
ASBEN	DIGO	JIE	LUGBARA	NGWAKETSE	TASUMSA

ASHANTI
ASSINI
ATTA
ATTIE
ATYUTI

AULLIMINDEN
AUSHI
AUYOKAWA
AVATIME
AVIKAM
AWIYA
AZANDE
AZJER
BABUKUR
BABWA
BACHAMA
BAGA

BAGIRMI
BAHARIYA
BAJUN
BAKAKARI
BAKO
BAKWE

BALANTE
BAMBARA
BAMILEKE
BANDA
BANGANDU
BANGBA
BANGI
BANYUN
BANZA
BANZIRI
BARA

BARABRA
BARARETTA
BAREA
BARGU
BARI

BASA
BASARI
BASHI
BASILA

DILLING
DINKA
DIOLA
DIULA
DOGON

DOROBO
DOROSIE
DRAWA
DUAISH
DUALA
DUI-MENIA
DUMA
DURU
DURUMA
DZEM
DZING
EBRIE

EDO
EGBA
EGEDE
EGYPTIANS
EKITI
EKOI

EKONDA
ESA
EWE
FAJULU
FALI
FANG
FANTI
FANYAN
FERTIT
FEZARA
FEZZAN

FIA
FIGIG
FILALA
FIPA
FON

FOUTADJALON
FOUTATORO
FUNGON
FUR

JOFRA
JUKUN
JUR
KABABISH
KABRE

KABYLE
KADARA
KAFA
KAGORO
KAKA
KALAI
KAM
KAMBA
KAMBATA
KAMBERI
KAMIR
KAMUKU

KANEMBU
KANURI
KAONDE
KAPSIKI
KARA
KARABORO

KARAMOJONG
KARANGA
KARE
KAREKARE
KASONKE
KATAB
KATLA
KAWAR
KEBU
KELA
KEMANT

KENGA
KENTU
KEPERE
KERARISH
KEREWE

KEYU
KGALAGADI
KGATLA
KHARGA

LUGURU
LUIMBE
LUKOKWE
LULUA
LUMBO

LUNDA
LUNGU
LUO
LUPOLO
LUVALE
LUWA
MAAZA
MABA
MACHA
MADA
MADI
MAGUZAWA

MAGWANGARA
MAHAFALY
MAHAMID
MAKA
MAKERE
MAKONDE

MAKUA
MALINKE
MAMA
MAMBILA
MAMPRUSI
MAMVU
MANALA
MANDARA
MANDJA
MANGA
MANGBETU

MANYIKA
MAO
MARGI
MASA
MASAI

MASALIT
MASHASHA
MASHI
MASINA

NGWATO
NKOLE
NKOYA
NKUNDO
NONO

NSAW
NSENGA
NSUNGLI
NUER
NUKWE
NUNUMA
NUPE
NUSAN
NYAKYUSA
NYAMWEZI
NYANEKA
NYANGIYA

NYANJA
NYARI
NYASA
NYIMA
NYORO
NZANKARA

ODODOP
OGADEN
OHEKWE
OKORO
OMETO
ORRI
PADANG
PANDE
PARE
PARI
PEDI

PEMBA
PENDE
PEPEL
PIMBWE
PODOKWO

POGORO
POKOMO
PONDO
POPO

TATOGA
TAWANA
TAWARA
TAZARAWA
TEDA

TEITA
TEKE
TEKNA
TEM
TEMAIN
TEMBU
TEMNE
TENDA
TERA
TESO
TETELA
THONGA

TIENGA
TIGON
TIGRE
TIGRINYA
TIKAR
TIV

TLHAPING
TLHARU
TLOKWA
TOMA
TONGA
TOPOKE
TOPOTHA
TORO
TOTELA
TRARZA
TRIBU

TRIPOLITANIANS
TSIMIHETY
TUAT
TUBURI
TUKULOR

TULAMA
TUMBUKA
TUMTUM
TUNGUR

BASSA
BATA
BATAHIN
BATI
BAUCHI
BAULE
BAYA
BEDE
BEDERIA
BEMBA
BENA
BENDE
BERABER
BERABISH
BERGDAMA
BERIBERI
BERTA
BERTI
BETE
BETSILEO
BETSIMISARAKA
BIAFADA
BIDEYAT
BIJOGO
BINZA
BIRA
BIRIFON
BIRKED
BIROM
BISA
BISHARIN
BOBO
BOFI
BOGO
BOKI
BOLEWA
BOMITABA
BONDEI
BONDJO
BONGO
BONI
BORAN
BORROM
BOSAKA
BOZO
FUT
GA
GAALIIN
GABERI
GADAMES
GAFSA
GAGU
GALIM
GAN
GANDA
GBANDE
GBARI
GERAWA
GERI
GIBE
GIL
GILI
GIMIRA
GIMMA
GISIGA
GISU
GOBU
GOGO
GOLA
GOMANI
GOROA
GOSHA
GREBO
GRUNSHI
GUANG
GUDE
GUIN
GULA
GULE
GUMUZ
GUN
GURAGE
GURENSI
GURMA
GURO
GUSII
GWANDARA
GYRIAMA
HA
HABBANIA
KIKUYU
KILINGA
KIMBU
KIMBUNDU
KIMR
KINDIGA
KINGA
KIPSIGI
KISAMA
KISI
KISSI
KITA
KOALIB
KOBA
KOKO
KOM
KOMA
KOMONO
KONA
KONGO
KONJO
KONKOMBA
KONO
KONONGO
KONSO
KONYANKE
KORANA
KORANKO
KOREKORE
KORO
KOROCA
KOSSI
KOTA
KOTOKO
KOTOPO
KOYAM
KPE
KPELLE
KRACHI
KRAN
KREDA
KREISH
KRU
KUBA
KUFRA
MASONGO
MATAKAM
MATENGO
MATMATA
MATUMBI
MAURI
MAYOGO, BADJO (NE)
MBAE
MBAGANI
MBALA
MBANGALA
MBEMBE
MBERE
MBESA
MBOLE
MBUBULA
MBUGU
MBUGWE
MBUKUSHU
MBUM
MBUNDA
MBUNDU
MBUNGA
MBWELA
MEBAN
MEDJI
MEKYIBO
MENASSER
MENDE
MERARIT
MERINA
MERU
MESSIRIA
MIDOBI
MIJERTEIN
MIMI
MINIANKA
MITTU
MITUKU
MOBA
MOBER
MOMBERA
MONDARI
MONGO
MOROCCANS
POPOI
POTO
PUKU
RANGI
REGA
REGEIBAT
RENDILE
RESHE
RESHIAT
RIF
RIYAH
ROLONG
RONGA
RUANDA
RUARHA
RUFFA
RUMBI
RUNDI
RUNGA
SAADI
SAB
SABEI
SAFWA
SAGARA
SAHEL
SAHO
SAKALAVA
SAKATA
SALA
SAMBURU
SAMO
SANDAWE
SANGA
SANGU
SANUSI
SANYE
SAPO
SEGEJU
SEKE
SELE
SELIM
SENA
SENGA
SENUFO
SERER
TUNISIANS
TURKANA
TURU
TUSYAN
UDALAN
UNGA
UREGU
VAGALA
VAI
VENDA
VERE
VILI
WABA
WAKURA
WALLAGA
WALLO
WANGA
WARA
WARAIN
WARJAWA
WIDEKUM
WOBE
WOLOF
WUM
WURKUM
WUTE
XAM
XOSA
YAHI
YAKA
YAKO
YAKOMA
YALUNKA
YANGERE
YANZI
YAO
YEKE
YERGUM
YESKWA
YOMBE
YORUBA
YUNGUR
ZAGHAWA
ZARAMO
ZEKARA

BRONG	HADENDOWA	KUKU	MORU	SHAIKIA	ZENEGA
BUA	HADIMU	KUKURUKU	MOSSI	SHAMBALA	ZERMA
BUBI	HAMAMA	KULANGO	MPEZENI	SHASHI	ZEZURU
BUDJA	HAMAR	KUMU	MPONGWE	SHAWIA	ZIBAN
BUDU	HAMYAN	KUN	MUBI	SHEBELLE	ZIGULA
BUDUMA	HARARI	KUNAMA	MUM	SHERBRO	ZIMBA
BUEM	HASANIA	KUNDA	MUMUYE	SHILA	ZUANDE, BATU (E)
BUILSA	HAUSA	KUNDU	MUNDANG	SHILLUK	ZULU
BULGEDA	HAWIYA	KUNG	MUNDU	SHLUH	ZUMPER
BULOM	HAYA	KUNTA	MURLE	SHOGO	
BUNDA	HEHE	KUNYI	MUSGU	SHUKRIA	
BUNU	HEIKUM	KURAMA, GURE (NE)	MWERA	SHUWA	
BURA	HEMAT	KURFEI	MZAB	SIA	
BURUNGI	HERERO	KUTSHU	NAFANA	SIDI	
BUSA	HIECHWARE	KWANGARE	NAIL	SIHANAKA	

Table AA – Appendix-Table 3.3: List of Sampled 227 Tribal Groups used in the Main Regression Analyses