Colonial legacies: Shaping African cities*

Neeraj G. Baruah J. Vernon Henderson [†] Cong Peng
London School of Economics

October 29, 2018

Abstract

Differential institutions persisting from colonial rule affect the spatial structure of African cities. Using a sample of 318 cities, Francophone cities have more compact development than Anglophone cities, overall, in older colonial sections, and at a clear extensive margin of development long after the colonial era. Why do colonial origins matter? The British operated under indirect and dual mandate rule, allowing colonial and native sections to develop without coordination. In contrast, integrated city planning and land allocation were a feature of French direct rule. Sprawl matters. Households located in less compact areas have poorer connections to water, electricity, and sewers.

Keywords: colonialism, persistence, Africa, sprawl, urban planning, leapfrog

JEL Codes: H7, N97, O1, O43, P48, R5

---

[*] We gratefully acknowledge the support of an Africa Research Programme on Spatial Development of Cities at the London School of Economics and Oxford University, funded by the Multi Donor Trust Fund on Sustainable Urbanization of the World Bank and supported by the UK Department for International Development. We benefited from helpful comments on an earlier version of the paper by David Weil, Matt Turner, Alex Moradi, Diego Puga, Henry Overman, and Gary Libecap, and suggestions by Andre Shleifer, Alain Durand-Laaserve, and Sameh Naguib Wahba. We also received helpful comments from seminar participants at The World Bank, University of California Berkeley, LSE, University of Oxford, and University of Bilbao. We are grateful to Nicolas Cerkez for his excellent research assistance.

[†] **Corresponding author:** J. Vernon Henderson; J.V.Henderson@lse.ac.uk; 44 (0) 2079556180; Houghton Street, LSE, London, WC2A2AE, UK. Other authors: Neeraj Baruah, n.g.baruah@lse.ac.uk; Cong Peng, c.peng7@lse.ac.uk.

1 Introduction

This paper shows that the spatial structures of a large set of cities in Sub-Saharan Africa are strongly influenced by the type of colonial rule experienced. Our main findings are based on a sample of 318 cities in 15 former British and 13 former French colonies in Sub-Saharan Africa (excluding South Africa). The extent and nature of land development are based on satellite measures of built cover which span 1990-2014 at fine spatial resolution. Compared to Anglophone cities, Francophone cities are more spatially "compact". There are several complementary dimensions to compactness which we investigate. First, we look at "sprawl" in cities taken as a whole today. Otherwise similar Anglophone cities cover 28% more area and have 23% more "openness" - a measure of sprawl from Burchfield *et al* (2006). Second we look at regularity and density of lay-out in older sections of the city likely to be physically influenced by how cities were laid out in colonial times. Francophone cities are laid-out in a more Manhattan-like gridiron fashion with lineal road systems, 4-way intersections and rectangular blocks, with higher density development.

Finally we look at disconnectedness of new developments at the extensive margin of cities built well after the colonial era, to see if there is persistence in the differential in this dimension of compactness beyond a margin where colonial influenced physical layout matters. Such persistence would then be due to persistence in land use management practices and norms decades after the end of the colonial era. Here for reasons detailed later, we focus on the degree to which new developments "leapfrog", or are scattered and spatially disconnected from existing developments. We employ a novel measure of leapfrogging. New patches of development can be broken into in-fill and extensions of prior developments versus disconnected, or leapfrog developments. We define leapfrog developments to be new patches of built cover emerging in a city which are at least 300 meters (or other minimum) away from the edge of any existing development. Anglophone cities have 61% more leapfrog patches compared to otherwise similar Francophone cities. Note, while leapfrogging contributes to sprawl, low density development *per se* need not imply scattered development. Nevertheless we find Anglophone cities both sprawl more overall and have more leapfrog new developments than Francophone ones.

Why does urban spatial form matter? On the direct public policy side, planners argue that compactness lowers the cost of providing public services and urban infrastructure. Compact cities require less infrastructure per person in the form of roads and utilities and the opportunity to operate mass transit systems more effectively, with the planning literature offering assessments of the savings from compactness (e.g., Trubka et al. (2010) and Calderón et al. (2014)). Hortas-Rico and Solé-Ollé (2010) provide econometric evidence on cost savings for Spain. The economics literature also argues that how cities are shaped and sprawl affects how we live: whether we attain efficient density for production in the face of communication externalities (Rossi-Hansberg, 2004;

Helsley and Strange, 2007); how much we pollute (Glaeser and Kahn, 2010); how much time we spend commuting (Harari, 2017); and how we interact socially (Putnam, 2000; Burley, 2016), with sprawl argued to lower positive density externalities, increase pollution and commuting times, and enhance social isolation [1].

In looking at development in African cities today, we add a piece of evidence on how urban form affects public service provision. We use Demographic and Health Survey data on utilities in a variety of African cities and find that families have worse connections to electricity, phone landlines, piped water, and city sewer systems, if they live in areas of a city which have more scattered development, presumably because of increased cost of infrastructure provision. For hundreds of millions of Africans as they urbanize, the institutions and history under which this happens will affect how they live their lives. While we offer a positive view on compactness, for those with a more ambivalent view, the main point remains: colonial origins have a strong lasting influence on the way people live in African cities, and mostly likely other parts of the world.

Why are Francophone cities more compact? Our answer is that the different institutions imposed under the two colonial rules affected and continue to affect urban spatial structure, including road layouts, sprawl and leapfrogging. The substantial literature on the history of urban planning in Africa argues that the French compared to the British adopted more centralised and standardised urban institutions within cities. Much of this literature is based on contrasting the "indirect rule" strategy of the British with French "direct (and assimilative) rule" (Silva, 2015; Njoh, 2015; Crowder, 1964). [2]. Following an economically-oriented style of rule, the British operated under a dual mandate system and dual structure of local government (Oto Peralias and Romero-Ávila, 2017). Home (2015) develops the dual mandate theme in detail for Anglophone Africa and some other parts of the British Empire: "Native authorities would continue to govern the native population, while townships, largely based on the cantonment model, accompanied the colonizers ... Land laws distinguished between on the one hand, the plantation estates and townships of the European colonizers, and, on the other hand, indigenous or customary land under the dual mandate approach...." (p.55, 57).

Driven to establish dominance over their colonies and to promote cultural assimilation, local French rule sought to bring all urban land under one control, supplanting all indigenous institutional structures and practices with French varieties, and bringing all public service provision under the local colonial government. Njoh (2015) provides significant detail and examples of the French political-oriented style of rule (Oto Peralias and Romero-Ávila, 2017). As part of maintaining control over the landscape, the French wanted the different neighbourhoods spatially integrated

---

[1] Using data on German neighbourhoods, Burley (2016) corroborates Putnam's hypothesized correlation between socialization and neighbourhood density, but also presents panel data evidence to suggest that sorting may explain most of this: more social people move to denser neighbourhoods which facilitate socialization. Of course that means greater density is of benefit to social people.

[2] In dealing with local chiefs, British rule was advisory while French rule was supervisory, sapping "the traditional powers of chiefs in the interest of uniformity of the administrative system" (Crowder, 1964)

and linked in a lineal pattern so that from one intersection an official could see 2 km in four directions (Njoh, 2015, chap. 1). Silva (2015) also details how the French adopted centralised and standardised grid systems. Durand-Lasserve (2004) writes about the urban dimension to the direct control strategy: "Customary land management is not recognized.....In former French colonies, this situation is clearly linked with....the French Centralist political model. It is characterised by state monopoly on land, and state control over land markets and centralized land management system....". We interpret these writings as indicating that, at the local level, the French imposed more centralised city planning and land use management than the British.

How might these differences affect spatial lay-out? First Francophone cities in the older colonial sections are more likely to have lineal road and rectangular block structures. As such the literature argues that imposing a road grid pattern leads to greater contiguity (Libecap and Lueck, 2011; Ellickson, 2012; O'Grady, 2014), where the aggregation properties of rectangles without gaps or overlaps promotes contiguity of the spatial structure, reducing sprawl. Second leapfrogging seems less likely at least historically in Francophone cities. British dual mandate already allows for disconnected parts of the city; and, then in these disconnected part, the British had a hands-off approach. The specifics of French centralized city planning allow for neither.

Methodologically, in terms of identification of causal effects of colonial rule, Africa gives an experiment in which initial institutions are given by the happenstance of what colonial power a city fell under, with country borders argued to be imposed from above with no regard for local conditions, history, or past governance (Michalopoulos and Papaioannou, 2016). Even if we think the colonial origin of a city is happenstance, the train of events could have left Francophone versus Anglophone cities ending up on average with very different geographies, which we know affects sprawl and city shape (Burchfield et al., 2006; Harari, 2017). We have a large set of geographic controls to account for these effects and results are robust to many experiments. However there could still be unobserved geographic features of cities or pre-colonial histories which differ systematically between Anglophone and Francophone cities. To meet this challenge to identification of causal effects, we conduct a border experiment in West Africa identifying and matching cities within 100 km of borders between different pairs of Anglophone and Francophone countries to ensure more equality in geographic and historical unobservables. We find as strong effects for this border sample.

The context also lends itself to a second experiment, separating out the effects of colonial influences due to persistence of physical infrastructures and historical city lay-outs versus persistence in influence in new areas of the city of planning and land management practices and norms inherited from the colonial era. Historical spatial lay-out of cities affects urban form for many decades. Public capital stocks are long lived; and rights of way for roads which are key to laying out a city once established are usually followed, given the high costs of acquiring new rights of way in an already built-up area. Persistence due to prior infrastructure could go beyond the older sections of cities, since post colonial lay-outs may follow existing types of patterns as accretions of older developments.

However, our context allows for a clear extensive margin where the primary influence cannot be persistence of the physical. For 111 cities of the 318 of our cities which are included in the World Cities Data set[3], population grew by 550% from 1960 to 2000, with approximately the same growth rate of Anglophone and Francophone cities. For 249 of our cities for which we have a 1975 measure of the built cover area, built cover grew by 145% from 1975 to 2014. We will focus on incidence of leapfrogging in areas of cities developed after 1990, in sections of cities built well beyond the colonial physically influenced city. Here differential effects must arise from colonial institutions and norms governing city planning and land management which persist decades after the end of colonial rule. There are various potential avenues for this: maintaining land use regulation rules and procedures either through inertia or continuation of norms through the training of post-colonial planners and policy makers, or the persistence of legal philosophies which filter into planning procedures (La Porta et al., 2008; Silva, 2015). We cannot distinguish which of these channels matters most, but we do argue there must be persistence of planning norms at a clear extensive margin of spatial development.

This discussion leads to an organization of the paper and presentation of results in four sections, after a discussion of the literature, context and data. First we look overall for cities in 2014 at the degree of sprawl as measured by openness and by overall area of the city. Then we turn to the more likely older and colonial influenced parts of the city, where persistence of physical lay-out and infrastructure and its extensions will drive outcomes. There we examine whether these sections of cities have more of a "Manhattan-like" gridiron structure, with lineal roads, 4-way intersections, and rectangular blocks in Francophone compared to Anglophone cities, and also we compare intensity of built cover. Then we turn to our main results, where we examine a clear extensive margin where persistence due to existing layout and road structures should not be important. We look beyond the 1990 built area of the city, to identify new patches of development. We examine whether Anglophone cities have higher counts of leapfrog patches and a higher ratio of leapfrog to total number of patches of new development. Finally we turn to the implications of sprawl and leapfrogging: the impact on provision of public utilities such as an electricity, piped water, or sewer connection.

---

[3] http://www.econ.brown.edu/Faculty/henderson/worldcities.html

2 Related literature

The related literature not already covered has two distinct veins: the more general literature on colonialism and the literature on local governance and urban form. For the latter we divide the review into theory versus empirical papers.

2.1 Colonialism and institutions

The literature on institutions and persistence (e.g., Banerjee and Iyer (2005) and Guiso et al. (2016)) argues that historical institutional accidents can have a strong impact on modern day outcomes.[4] Historical colonial rule and associated institutional choices have been documented to be significant for contemporary institutions, economic development and political stability (Acemoglu et al., 2001; La Porta et al., 2004). La Porta et al. (2008) show that having French civil law as opposed to British common law imposed resulted in differences in regulatory outcomes, banking procedures, property rights enforcement and the like. They argue that French civil law operates to control economic life, while (Mahoney, 2001) argues that given the ideological differences underlying the two legal systems, French civil law is more "comfortable with the centralized activist government".[5] Oto Peralias and Romero-Ávila (2017) point out that in Africa with its limited extractive opportunities and large indigenous populations compared to some other British colonial regions, even the imposition of British common law was limited, while the French tended to impose direct centralist rule in all colonies. For this paper, it is helpful to note the parallels between statements about French civil law and our characterization of Francophone cities as being managed top down by a central city authority with an eye to imposing regularity in design and lay-out.

2.2 Theory literature on local governance, urban structure and sprawl

Theory papers examine the potential impact of an authority with overall control in metropolitan area governance, as opposed to there being either pure *laissez faire* or decentralised governance. An older literature argues that centralised governance by a benevolent city planner is needed for proper internalisation of externalities and provision of localised public goods overall for a city (Davis and Whinston, 1964; Hochman et al., 1995)[6]. This literature does not deal with compactness. However recent work focused on specific externalities argues that cities under *laissez faire* have insufficient density. Rossi-Hansberg (2004) and Helsley and Strange (2007) show that, in the

---

[4] There is the specific work on Francophone and Anglophone colonial legacy within a small area of Cameroon (split into parts which are former British and French colonies) as affecting wealth and water outcomes Lee and Schultz (2012).

[5] In a different vein, Acemoglu et al. (2011) argue that the imposition of French civil law in the 19th century on areas of Germany which had remnants of feudalism and elitist extractive institutions improved subsequent economic growth.

[6] There is also the huge literature on decentralization of governance within countries. See a general summary in Oates (1999) and for within cities see Helsley (2004) and Epple and Nechyba (2004).

face of communication or social interaction externalities which decay with distance, absent appropriate regulation by a single city authority, cities will lack efficient density of activity near the city centre. More informally, Brueckner (2001) and Brueckner (2005) note that uncoordinated developers will take advantage of the fact that congestion is unpriced and public infrastructure may be subsidised which will lead to sprawl, for example, through ribbon developments sited along government built arterial roads. These are strong arguments that uncoordinated and decentralised land development will result in cities that are less compact, potentially as an empirical prediction. That said, these papers do not directly model the political forces driving the decisions of a central authority.

Our main results concern leapfrogging, which is examined in two theory papers. Turner (2007) examines whether neighbourhoods on the urban fringe will have leapfrog commercial developments. Henderson and Mitra (1996) consider a city with spatially decaying communication externalities across firms and strategic competition by developers setting up new developments on the city fringe. Such developments may be contiguous to old ones or leapfrog. Both papers argue that higher intensity of development in the core city is associated with lower likelihood of land developers engaging in leapfrog development at the extensive margin.

2.3 Empirical literature on land use regulation and urban form

A key paper by Libecap and Lueck (2011) uses a border methodology to study the allocation of rural land in Ohio under a 'metes and bounds' system versus a rectangular survey system. The former is a decentralised system with plot alignments and shapes defined by the individuals claiming the land and topographic constraints, while the latter involves centralised and regularised demarcation of surveyed plots. The authors find subsequent strong coordination benefits and reduced transaction costs due to regularity, which they show metes and bounds is less likely to achieve. Their exploration of land demarcation systems in rural areas suggests land institutions in urban areas may be distinguished by their degree of centralisation and standardisation. The parallels to colonial land demarcation systems were extended more directly to cities by O'Grady (2014), focusing on an example comparing a centralised and standardised rectangular grid demarcation with a more *laissez faire* demarcation system. O'Grady (2014) shows that, for New York City, neighborhoods with a greater fraction of rectangular grids imposed centrally and historically then experienced higher future land values and more compactness, or higher density of use.

Other papers examine persistence in spatial outcomes driven by historical infrastructure investments (e.g. Bleakley and Lin (2012) and Brooks and Lutz (2014).[7] The key paper on sprawl

---

[7] Michaels and Rauch (2017) look at the differential influence of the fall of the Roman Empire in France versus England on urban population size and growth centuries later, based on the notion that French Roman settlements persisted after the fall, while British ones due to political upheavals disappeared almost immediately.

by Burchfield et al. (2006) analyses geographic and historical influences on the degree of land use sprawl in US cities. Shertzer et al. (2016) argue that 1923 Chicago zoning ordinances have a bigger effect on the spatial distribution of economic activity today than geography or transport networks in Chicago. Redfearn (2009) looks at how land use patterns in USA cities are driven by historical uses.

3 Context, data, specification and identification

In this section we describe the context and data, with more details given in the parts of Appendix A1 and in Table A1. Then we turn to a base specification and issues of identification.

3.1 Colonial countries

Our classification of African countries by colonial origin is shown in Figure 1a along with the cities in our sample. The division is not always straightforward. World War I changed the colonial map, with former German colonies being split among the French (e.g., most of Cameroon) and British (e.g. Tanzania), with many complex splits vis a vis modern countries (e.g., Togo). If we think governance procedures and urban plans were developed near the end of the 19th century and early 20th before the end of World War I, those procedures could set the tone for decades to come. We would then face the problem of German influences confounding the picture. Omission of these countries in robustness checks has no impact on results. While some approaches to governance and land allocations are in place well before World War I, cities were typically in infancy, so the pre-World War I influence may be limited.

3.2 Data on land use and cities

We utilize three epochs of land cover data - 1990, 2000, and 2014 - which classify pixels of 38m spatial resolution into different uses where our general focus is on built cover (impervious surface) versus non-built cover (water, various vegetation and crop, barren water and so on). These data are constructed from the Global Human Settlement Layer (GHSL) – a new global information baseline describing the spatial evolution of the built environment, a project which is part of the Global Human Settlement Project by the European Commission and Joint Research Centre (Pesaresi et al., 2013). It is the most spatially global detailed dataset on built cover available today. While the data are based on open access Landsat satellite imagery, other information from publicly available and validated coarse-scale global urban data (MODIS Global Urban Extents, MERIS Globcover and Landscan among others) to more fine-scaled and volunteered geographic information (Open Street

Maps and Geonames) are incorporated[8] (Pesaresi et al., 2016). Available since 1972 (Ban et al., 2015), the GHSL estimates the presence of built-up areas in different epochs (1975, 1990, 2000 and 2014)[9]. For built up cover we have two types in any year, the stock of built cover from the prior period (defined to also be covered in the current and subsequent time periods) and new cover built since the last period, which we use to analyse the nature of new development.

In applying these data, we have a base sample of 333 cities, of which 106 are former Francophone cities and 227 former Anglophone cities, with the latter including 122 Nigerian cities. These cities are reported in Table A1 and shown in Figure 1a. These 333 cities are all cities in the relevant colonial origin countries which are over 30,000 in estimated population in 1990,[10] which have built cover data for years of 1975, 1990, 2000, and 2014. We use *Citypopulation.de* to get city population numbers (based on Censuses), supplemented with data from *Africapolis* for Nigeria. We set 30,000, because across countries and time there is a difference in population cut-off points for reporting on city populations; a 30,000 cut-off provides more consistency in reporting. We also wanted cities to have some degree of maturity to urban spatial development and planning (or lack thereof). We then apply criteria on the extent of persistent cloud cover to get cloud free city-year observations for 2000 and 2014. [11] Removing cities with cloud cover and hence only partial coverage for land cover, in 2000 we have 299 city observations and in 2010 we have 307, with a total of 318 out of 333 cities in one year or the other.

From the base sample, we explore various sub-samples. One is West Africa which is distinct as seen in Figure 1a, in that it contains most of the Francophone countries. Another sub-sample excludes Nigeria which is a third of the sample, to make sure it is not driving the results. A third is to look at newer cities whose origins appear to be colonial (from the French-British era) and founded after 1800, based on web scrapping of information. These cities are denoted in Table A1. We get larger effects for cities which are more subject just to colonial influences. Finally, there is a sample for Open Street Map analysis of all Francophone cities over 300,000 in 2012, with the size bound imposed to ensure more reliable OpenStreet map data which are new to Africa. These 20

---

[8] Landsat data is typically available at 30m spatial resolution. GHSL employs an information fusion operating procedure based on a tiling schema to combine the source Landsat imagery with other data, imposing further restrictions on effective data resolution - the GHSL project adopts a nominal spatial resolution at the equator of 38.21m which best approximates the native 30m of the Landsat imagery.

[9] Pre-processed Landsat scenes were collected for the epochs (1975, 1990 and 2000) from the Global Land Survey (GLS) at the University of Maryland (Giri et al., 2005) and were combined with Landsat scenes for the 2014 epoch to create the spatiotemporal composite. The epochs that characterise the builtup GHSL data approximate the temporal dimension of the GLS data. Epochs signify a time-period range around a given year from which the best available Landsat scene is drawn. For instance, the 1990 epoch for a city i may be drawn from 1988, while it may be 1992 for city j. Processing uses supervised and unsupervised classification based on a combination of data-driven and knowledge-based reasoning. Spectral, textural and morphological features are extracted and a supervised classification method relying on machine learning is employed using a global training dataset derived from various sources at different scales.

[10] These are based on population censuses around 1990 and with growth to 1990 generally based on city population growth rates between two relevant population censuses.

[11] We require the city-year to be 95% cloud free in 1990 for initial stock variables and 100% free in 2000 and 2014 for flow variables. We lose 49 city-year observations from imposing the 0 cloud cover restriction and 11 more from requiring no more than 5% cloud cover in 1990. If we imposed a 0 cut-off in 1990 for cloud cover, there be a loss of another 65 cities. We use the 1990 built cover within our cities at times as a control variable, when looking at flows to 2000 or 2014. Since 1990 defines 2000 pre-built area, in the 2000 analysis any 1990 cloud cover areas in a city are dropped from the calculations for that city.

cities are then propensity matched to 20 Anglophone cities over 300,000 (out of 68) which have similar attributes like populations, growth and coastal location, among others. We will use these to analyse differential urban structure and road lay-outs in the colonial portions of larger Francophone versus Anglophone cites. These cities are listed in the Appendix and mapped in Figure 1b.

3.3 Data on geography and the extent of the city

In applying these data, we must define the spatial extent of cities. Since outcome measures involve aspects of the built environment, we do not want to use a measure based on built cover per se to define the extent of the urban area. We will note later how that biases results, by tending to omit extensive margin developments which are more leapfrog in nature as opposed to infill and extension. We rely on night light readings for Africa (Donaldson and Storeygard, 2016; Henderson et al., 2017) and define the city to be the area within the outer envelope of all areas lit for at least two of the last 5 years from 2008-2012. African cities have generally low light levels, so we do not threshold the lights to be above some cut-off. For smaller cities thresholding excludes obvious built areas (looking at Google Earth) and even some entire cities. In very big cities, blooming is an issue and the lights boundary can include large undeveloped areas and cover satellite towns. In various robustness checks, some reported in Baruah et al. (2017), we experimented with imposing light thresholds, setting distance limits over which we look, and trimming the cities with high maximal and low minimum distances from the centre to the farthest edge. Results are robust to these experiments. We also use night lights to define the city centre, as the brightest lights pixel (about 0.8 x 0.8 km square near the equator) in 1992/93. We also defined smoothed built cover boundaries for cities as described in the Appendix for 1975, 1990, 2000, and 2014. The 1990 measure gives an urban core, beyond which, in the extensive margin, we will find over 98.5% of our post-1990 leapfrog patches.

3.4 Specification

Through the end of Section 6, regressions have the following general form:

$$Y_{ijt} = X_{ij}\beta + Z_{ijt}\theta + \delta Anglophone + d_t + \varepsilon_{ijt} \tag{1}$$

where i is city, j is country and t is time. Some regressions are cross-section so there will be no time dummy d_t; but the later leapfrog ones will be flow measures for 1990 to 2000 and 2000 to 2014. X_{ij} are city i factors which are either time invariant or for which we want a base period measure. Z_{ijt} are time varying factors. The coefficient of interest is δ - the Anglophone differential. For the border experiment in Section 6.2 we adjust the error structure to have fixed effects for 14 cross-border clusters of cities in close spatial proximity. The specification for the DHS work will be discussed in Section 7.

A basic identification issue is whether Anglophone cities differ from Francophone because of colonial origins or because of differential underlying geographic conditions of cities which influence urban layout, regardless of colonial origins, noting that Burchfield et al. (2006), Saiz (2010) and Harari (2017) all show that geography influences urban form[12]. Our X_{ij} controls on geography reflect this concern. We use measures found in different literatures, starting with Burchfield et al (2006). First there is terrain where hilly areas spread out developments around inaccessible topography. We have a basic measure of ruggedness as defined by Nunn and Puga (2012) and of the range of elevation within the city. Water is another constraining feature - we have distance to the coast from the city centre; and, if the city is coastal, the length of coastline (in kms) within its boundary, where extensive coast means more inlets and bays again influencing city shape (Harari, 2017) . In the land portion of the city, we know the fraction of pixels that are inland water (lakes, rivers, wetlands). In some specifications, we then draw a 5 km buffer around the outer edge of the city land area. In that buffer we catalogue the percent of the buffer in different uses in the base period: forest, shrubs, crops, water and wetlands, and sparse and bare vegetation (compared to grasslands). This reflects an issue of the opportunity cost of city land (which could vary systematically between Francophone and Anglophone). For that reason we also control for the rainfall average from 1950-2000.

The hardest items to deal with are growth and economic opportunities for the city. We have initial population size (based on circa 1990 population) and lagged country level GDP per capita, a Z_{ijt} variable. For a city specific growth control, we experimented with various measures of local and national growth of night lights and national urbanization rates which yield similar results, but decided to directly condition outcomes on individual city population growth rates. Faster growing cities may be more likely to sprawl per se and Appendix Table A7 hints at modestly faster average growth among Anglophone cities. We control for the annualized city population growth rate from 1990 to 2012. Doing so loses us about 9% of the sample due to lack of circa 2012 population numbers. Finally we have base 1990 land cover in the city which will appear in leapfrog specifications.

4 Overall patterns in the data for cities as a whole

Using the GHSL Landsat based data, first, we correlate an accepted measure of sprawl with Anglophone colonial origins to see motivating patterns in the data. We examine the *openness index* from

---

[12] There are also social conditions and in a developed country context we might worry about differential attitudes towards use of the automobile and the development of sprawl. First we note that even in seven major Sub-Saharan African cities, automobiles presently account for under 15% of trips (Consortium, 2010). Second, that fraction would have been even smaller in the colonial era.

Burchfield et al. (2006) for the overall city. For openness, following Burchfield et al. (2006), for each built-up 38m x 38m pixel in a city in a year we calculate the fraction of unbuilt pixels in the immediate 1 sq km grid square. These fractions are then averaged across all built pixels in the city. The measure reflects the extent of open space around the typical built pixel in a city. Second we look at the lights area of cities, to see if, *ceteris paribus*, Francophone cities occupy larger areas.

What correlations do we see in the raw data? First, we compare distributions of openness for Francophone versus Anglophone cities,[13], based on graphs in Burchfield et al. (2006). Figures 2a and 2b show the pdf for the distribution of built up pixels in 1990 and 2014 by the percent of land not built in the surrounding one square kilometer (i.e., openness). In both years, the dotted line for Francophone relative to the solid line for Anglophone shows the Francophone pdf's shifted left. Francophone cities tend to have a greater fraction of built pixels in areas with very low openness and a smaller fraction of pixels in areas which are very open, suggesting that Francophone development is more compact and Anglophone more sprawled. Visually it may appear that the differential is smaller in 2014, raising the possibility of some convergence. We did explore this issue, but regression work suggests that, conditional on 1990 openness, there is no distinct Anglophone convergence in openness between 1990 and 2014 (Baruah et al., 2017).

Table 1 examines the Burchfield openness index in 2014 in regressions controlling for geography and other city characteristics and focusing on the Anglophone city effect. Column 1 has no controls; column 2 adds in all geographic controls except those which characterize land use at the city fringe; column 3 adds those controls on land use at the city fringe; and column 4 adds to the geographic controls the listed controls on size and growth in Table 1. Results on geographic controls are in Table A2. The Anglophone effect in column 1 is an increase in openness of 23 % and remains almost the same with all controls added in column 4. Adding in controls has little impact on openness, which suggests fundamental differences in the geography or economies of Francophone and Anglophone cities are not driving the colonial correlation, even though some control variables have expected effects based upon the analysis in Burchfield et al. (2006). Bigger cities have less openness and cities with greater elevation differentials have more. Rainfall (opportunity cost of urban land) or related vegetation measures at the city edge tend to reduce openness. We also note that it could be that French centralised land use control may have responded to differential geography of cities differently than the more decentralised British approach. If we interact all covariates in column 4 with the Anglophone indicator and predict how a typical city with mean characteristics in the overall sample differs under the two regimes, under an Anglophone regime that typical city has 24% greater openness.

In columns 5-8 of Table 1 we then turn to the area of cities as measured by the lights boundary. We have the same pattern of controls as in columns 1-4. In column 8 with full controls, Anglophone

---

[13] These are 307 cities where the Landsat images used are 95% cloud free in 1990 and 100% cloud free in 2014.

cities occupy 28% more land than their otherwise similar Francophone counterparts. Coefficients for control variables are in Table A2. If we interact all covariates in column 8 with the Anglophone indicator and predict how a typical city with mean characteristics in the overall sample differs under the two regimes, under an Anglophone regime that typical city has 37% greater area. One might worry about lights being a noisy measure of city area. Accordingly we rerun the regressions trimming the top and bottom 5% of cities by furthest distance from the centre to any edge. As reported in Table A9 in the Appendix, trimming does lower magnitudes but strong effects remain, whether we measure sprawl by openness or just total area. In short, Anglophone cities as a whole have significantly more sprawl. The task now is to look at the two margins, intensive in the colonial influenced sections of the city and extensive in effectively the areas beyond the 1990 built part of the city.

5 The Colonial portions of cities

We define portions of cities as possibly being under direct colonial influence if they are within 5 km of the city centre. These include the old colonial sections and, through road extensions, the potentially physically colonial influenced sections of the city. For larger cities we know the spatial layouts of the colonial sections, today and some historically from maps. For all cities we know their current intensity of land use, or built cover.

5.1 Road layouts: Anglophone versus Francophone cities

To better understand aspects of colonial influence we start with an example, which first compares Bamako to Accra, and then Brazzaville to Harare. Examples are difficult to construct since the key is to have city-pairs for which we could obtain detailed road maps from about the same year near the end of the colonial era. We also want cities with similar initial and final sizes. All four cities emerged as cities in the late 19th century, Bamako and Brazzaville under French rule and Accra and Harare under British. Starting with Bamako and Accra, their populations are similar in the early 20th century: Bamako at 16,000 in 1920 and Accra at 18,574 in 1911.[14] Accra retains that modest population difference with Accra at roughly 2.3 million and Bamako at 1.8 million today. While Accra is a coastal city, Bamako is on a major river with the initial city on just one side (like a coast line). Bamako had its first (apparently implemented) road plan in 1894 (Njoh, 2007, p. 92) replacing spontaneously prior developed roads with a street network on a classic gridiron with streets intersecting perpendicularly (Njoh, 2001, p. 23). Bamako's urban land was under state control by 1907 with the *"Plan d'une cite administrative - un quartier de Bamako"*, with the

---

[14] For Bamako: "France: Africa: French West Africa and the Sahara". Statesman's Year-Book. London: Macmillan and Co. 1921. pp. 895–903 – via Internet Archive. Colony of French Sudan. For Accra "Population Studies: Key Issues and Continuing Trends in Ghana" S.N.A. Codjoe, D.M. Radasa, and S.E. Kwankje, Sub-Saharan Publishers, Accra, 2014, p.115

state supreme in land allocations and assignment of set plots (Bertrand, 2004). Accra proceeded under the usual British dual mandate without a comprehensive plan until The Town and Country Planning Ordinance of 1945 when, according to Grant and Yankson (2003), "zoning and building codes were strictly enforced to maintain an orderly European character and ambience", especially in the European Central Business District (Ahmed and Dinye, 2011; Grant and Yankson, 2003).

Figure 3a shows the road layout in the older sections of these cities, roughly up to 5km out from the city centre. For Bamako we show the 1963 road layout from tracings of road maps and the road layout today from Open Street Maps (OSM). For Accra we show the roads for 1966 [15] as well as today. Visual inspection suggests several takeaways. First there is physical persistence in both cities - roads that were in place 50 years ago generally remain in place today. Second Bamako presents as having large sections of intense dense, grid-like road structures where sections are interconnected by mostly long lineal roads. And 1963 fringe roads that appear to meander to the north east have in some cases been replaced by grid-like structures. New sections of the city generally are on a rectangular grid structure. In contrast, Accra shows much less grid-like structure with fewer lineal connecting roads between developments even in the colonial parts of the city. And new developments on the fringes of the colonial parts of the city appear to have much less rectangularity and lineal connections than Bamako.

Figure 3b shows a comparison between Brazzaville and Harare. Both cities had a population of about 50,000 in 1945 [16]. These populations have grown to about 2 million or more based on available information. Figure 3b shows current OSM roads and the 1958 and 1954 mapped roads in respectively Harare and Brazzaville. The comments we made on rectangularity and lineal connected roads systems for Accra versus Bamako apply equally well in this comparison. These illustrative mappings suggest evidence of more regular layout and centralised planning in Francophone cities compared to their Anglophone counterparts.

To test whether these differences hold more generally, we took all 20 Francophone cities in Sub-Saharan Africa over 300,000 in 2012, to analyse road layouts from OSM. Since OSM data is relatively new for Sub-Saharan Africa, we restricted our sample to larger cities and to mapping within 3-5 km of the city centre to try to ensure better reporting. We then chose 20 corresponding Anglophone cities over 300,000 out of the 68 in that size range, using a one to one Mahalanobis distance based matching approach without replacement. The covariates include initial city population in 1990, city annual estimated population growth from 1990 to 2012, average rainfall from 1950 to 2000, coastal dummy, and absolute elevation. With matching, means of the matching variables show insignificant differences between Francophone and chosen Anglophone cities (See Table A3 in the Appendix). Also in the end there are 11 Nigerian out of 20 Anglophone cities, ef-

---

[15] The source of both old maps is Bodleian Library of Oxford University. It is digitized by Ramani Geosystems, a GIS firm based on Nairobi

[16] From respectively Robert Edmond Ziavoula, ed. (2006), Brazzaville, une ville à reconstruire, Paris: Éditions Karthala; and the Demographic Yearbook 1955 of the UN

fectively matching Francophone ones concentrated in West Africa. Other samples drawn to reduce the Nigerian count show similar if not stronger results. [17]

For this matched sample we ask if the Francophone colonial sections of cities and immediate extensions have different structures than Anglophone ones, with a more regular and connected road system, which would guide the complementary layout of private investments. Here we give quantitative evidence of the more standardised grid system of Francophone cities. Figure 4 illustrates the process followed and derivation of measures. In part A we have the raw OSM road network data for part of a city and B shows the derived road blocks. Road blocks are categorised by their degree of rectangularity using the minimum bounding rectangular method of Žunić et al. (2012) and Rosin (1999). The minimum bounding rectangle is a rectangle which minimally encloses the actual block polygon. Rectangularity of a block is the ratio of the area of the block to the area of its minimum bounding rectangle - a perfectly rectangular road block would be 1, and the ratio tends to fall as it takes on more complex shapes. Part C of Figure 4 ranks all the blocks in the shot - the dark blocks with rectangularity measures equal to or greater than 0.9 are ones we call rectangular blocks. We chose a cut-off of 0.9 to allow for measurement error and topography in approximating perfect rectangles.

Part D of Figure 4 shows how we define *gridiron blocks*, which is the basis for our main measure and captures contiguity in rectangularity of layout of sections of a city. To be a gridiron block, a block must have a rectangularity index greater or equal to 0.9, be devoid of dangles, and be connected to all neighbouring blocks by 4-way intersections. Dangles are roads off the regular road network which lead to no connection (i.e., dead-end), or blocks with a *cul-de-sac*, dead-end, or T-intersection; and they are illustrated in Part E of Figure 4. Part D of Figure 4 shows in yellow the subset of rectangular blocks which qualify as gridiron. For analysis we calculate the share of gridiron blocks to all blocks in the area in question.

We believe OSM data pretty comprehensively maps roads in these 40 African cities up to about 5 km from city centres, covering both the colonial parts of the city which generally lie within 3 or fewer km of the centre and post-colonial immediate extensions. Further out, mapping is expected to be of poorer quality because of the incomplete nature of volunteered OSM information. In Figure 5, for each of these cities we show the fraction of gridiron blocks out to 5 km with Anglophone cities represented by the darker shades. Francophone cities generally have higher shares of gridiron blocks. The Anglophone outlier, Bur-Sudan (Port Sudan), was a new "planned city" from scratch, like for example Canberra. The visual impression is confirmed by a regression coefficient giving the average Anglophone differential. Anglophone cities average 20 percent points fewer gridiron blocks, from a mean of 17. The sample means are almost the same at 3 and 5 km, so there is no

---

[17] For example, for another project, we had a sample of 55 cities generally over 240,000 for which we obtained SPOT data which was weighted against having too many Nigerian cities and towards greater country (Francophone) coverage. We also matched with Anglophone cities without an explicit requirement that they be over 300,000 which again weighs against Nigeria.

overall diminishing of regularity with the 150 percent increase in area covered.

We note two other things. First results on the share of rectangular blocks are similar to those for gridiron, but we prefer the tougher criterion which captures contiguity. Second we also looked at the share of dangles. Anglophone cities have 3.5% higher shares of blocks (for a mean of 10.7) with at least one dangle to all blocks of the area in question, but the coefficient is only significant at the 11% level. Overall the results suggest a strong colonial influence of centralised control and grid planning, as suggested by Njoh (2015) and Durand-Lasserve (2004), which persists until today.[18]

5.2 Intensity of land use in the colonial portions of cities and immediate extensions

Corresponding to grid-like structures of roads is much greater intensity of land use in the colonial portions and their extensions for Francophone cities compared to Anglophone, indicating much greater compactness. In Table 2, for the full sample of cities, we show ring by ring intensity regressions for 1990, the year nearest the colonial era, as we move out from the city centre in 1 km increments. The dependent variable is the log of the total number of built pixels in each ring. Shown are the coefficients for Anglophone and for an additional covariate beyond those in Table 1: the log of number of available pixels (built or not) in each ring by city, which also allows for differential ring counts based on geography (e.g., cities on a coastline or river vs. more circular non-coastal cities).

For rings 0-1, 1-2, 2-3, 3-4 and 4-5 km Anglophone cities have 39-80% fewer built pixels. They are much less intensely developed. After 5km, the sample starts to drop as we lose smaller cities with no area beyond the given radius. At 5-6 km and beyond out to 11-12 km, there are no significant differences for Anglophone countries. We show the results for the 5-6 km ring; the full set of results and many other intensity specifications are in Baruah et al. (2017). We report one here in Table A4 in the Appendix, where all rings are pooled and we estimate the height and slope to the intensity gradient as we move away from the city centre. In Table A4, Anglophone cities have 89% less intensity at the centre, Francophone cities have a steep slope to the intensity gradient of -0.058, while Anglophone cities are almost flat with a slope of -0.009. The bottom line is the same: Francophone cities have much more intense land use in the older colonial physically influenced portions of cities, while Anglophone cities sprawl.

The evidence so far indicates that Francophone cities are more compact overall and in particular in colonial physically influenced sections of cities. However, openness and intensity raise two problematic measurement issues, which led us to focus on leapfrogging in the next section. First in looking at these measures, it is hard to define the comparative extensive versus intensive margins

---

[18] One issue is whether Anglophone cities were regularly laid out but just not on a rectangular grid, using more diagonal roads with roundabout intersections. We checked the count of roundabouts within 5 km of the centre. On average there is no difference between Francophone and Anglophone cities.

of cities in a simple way. For the same city populations, Francophone cities are smaller and extend less from the city centre. At, say, 5-6 km, many middle size Francophone cities will be ending, or at their edges and extensive margins with low land intensity given by their steep intensity gradients. In contrast, corresponding size Anglophone cities will have city edges further from the centre; and 5-6 km will still be the intensive margin, at an Anglophone overall low intensity of use. Second and critically, our data measure built cover or footprint of impermeable surface on the ground. They do not measure building volumes since we have no data on heights. It could be Anglophone cities near the centre are built higher with more open ground cover, but a high intensity of building volume. Of course that would not be consistent with Anglophone cities covering more area overall. However given these two issues, the next section focuses on a measure of disconnectedness, or leapfrogging, which we can accurately measure and which operates at the extensive margin of cities.

6 Compactness in the (vast) post-colonial extensive margins of cities

Leapfrogging is a concept well established in the literature, but we have a novel measure in this paper. Leapfrogging is a flow measure of leapfrog patches occurring under development from 1990 to 2000 and then under development from 2000 to 2014. How do we define leapfrog development? Using the 1990 to 2000 period as an example, in 1990 we have a set of built pixels, which are typically in clusters. We define the boundary or outer envelope of each cluster of contiguously developed pixels, which we call patches (where some patches are isolated singleton developed pixels). In the illustrative Figure 6, the 1990 developed areas are in light shaded orange. The focus is on newly developed pixels. These also appear as patches of contiguous newly built pixels, which also have boundaries. Around each bounded patch (or singleton) of newly built pixels we draw a 300m buffer, effectively including all pixels or parts of which lie within 300m of the nearest border of the new patch. Then we focus on the areas *within (just) these buffers* around new patches to define three types of new development. If that buffer area is generally contained within an existing development it is called infill (red area in the figure). If it only marginally intersects the existing cover (or is within 300m of it), it is called extension (blue in figure). If does not intersect (within 300m) any existing 1990 development it is called leapfrog (green patch). Our buffer choice of 300m is guided by the literature on “walkable neighbourhoods”. Most notably, Barton et al. (2003) claim a theoretical circular catchment of radius 300m (corresponding to walking time of 5 minutes) as a planning goal for urban amenities and interactions. Thus, leapfrogging occurs when a new urban patch development arises beyond the walkable distance of an existing urban patch. Of course, walkable distance is in the eye of the beholder and we experimented with a different size buffer as reported later under robustness checks.

Given these concepts, we have a measure of the connectedness of urban expansion, or the landscape expansion index (*LEI*) (Liu et al., 2010) where

$$LEI = \frac{A_b}{A_b + A_o} * 100 \tag{2}$$

where A_b is area of intersection between the buffer zone of a new patch with existing built cover, to give the area of already built pixels within the buffer zone of a newly built area. A_o is the area of intersection between (just) the buffer zone itself for new patches with open space, to give the area of open space in the buffer zone. Thus the denominator is total area in the buffer zone and the numerator the built area within that. Infill might be defined as an $LEI > 50$, so at least 50% of the buffer is already built space. Our focus will be on leapfrog patches where the LEI is 0, so there is absolutely no already built space in the buffer surrounding new development. However first in Figure 7, we show the pdf of the LEI measure for patches of new development in all Francophone versus Anglophone cities. The Anglophone measure is more concentrated at the low end of LEI's between 0 and 25%, representing a greater concentration of leapfrog or almost leapfrog developments.

6.1 Primary results

We now turn to statistical analysis and look at the absolute and relative count of leapfrog patches in a city and the area they encompass. Most critical to our claim that we are looking at the extensive margin is the fact that over 98% of all leapfrog patches in the sample lie outside the smoothed land cover boundary of the city in 1990. These are developments in areas new to the city since 1990. LP patches average about 12% of all patches but have high variation across cities (the standard deviation on the variable is 11). While the estimating equation is based on eq. 1, we have two periods for each city. The time dummy, d_t, captures the fact that the second time interval (2000-2014) for LF patches is 4 years longer than the first (1990-2000). We also control for initial built cover of the city in 1990. We cluster errors at the city level to deal with serial correlation. Focusing on flows and the extensive margin may help difference out the influence of key unobserved geographic factors. A finding of greater leapfrogging in Anglophone cities would suggest colonial patterns of disconnected and independent developments in Anglophone compared to Francophone cities persist under today's inherited institutions at a margin well beyond the colonial city circa 1965.

Columns 1-4 in Table 3 show basic results for the logarithm of the count of leapfrog patches in a city. We follow the format of Table 1, where column 1 has no controls other than the time dummy and initial 1990 cover; column 2 adds the main geographic controls; column 3 adds the remaining controls on land use beyond the city edge; and column 4 adds to that the listed size and growth controls. Geographic controls matter as do growth controls; and their coefficients are given in Table A5.

In the main column 4 of Table 3, Anglophone cities have 61% more leapfrog patches. In the specification, there is a small count of about 5% of observations which are zeros which we set to the minimum of 1 (so the log is zero). Results in the Appendix Table A6 show OLS results excluding these zeros, Tobit results, and Poisson count results. The Anglophone magnitude is very similar, especially for the Tobit.[19] We also note the issue again that Francophone cities may respond to differential geography of cities differently than the Anglophone ones. If we interact all covariates in column 2 with the Anglophone indicator and predict how a typical city with mean characteristics in the overall sample differs under the two regimes, under an Anglophone regime that typical city would have 71% more leapfrogging. We also note that effects for the two time episodes of leapfrogging are similar.

In columns 5-8, we show results for the log (count LP patches/ count total patches). The coefficient on the ratio in column 8 is 0.32.[20] This implies the marginal effect on all patches for Anglophone (over Francophone) cities is about 0.29 (0.61- 0.32). Anglophone cities develop more by building in greenfield areas, rather than intensifying already built cover (an intensification the GHSL data don't measure) and have more patchy development. But, given that, they are even more prone to these patches being leapfrog ones. In column 9 we show results for log (average area of LP patches), which checks whether Anglophone patches are somehow bigger, so, for example, they might be easier to service with infrastructure. There is no average size difference in leapfrog patches between the two types of cities. In summary Anglophone cities have more patchy development at the extensive margin, especially leapfrog development, where these leapfrog patches are no bigger than their Francophone counterparts.

6.2 Identification

Are the effects in Table 3 causal? In part we are arguing causality through the weight of different pieces of evidence and the use of a large set of controls and flow data, but biases obviously may remain. Although the insertion of many controls has limited impact on the Anglophone "treatment", the characteristics between the Anglophone and Francophone sets of cities are not balanced in all cases (column 1 , Table A7), suggesting there could be unobservables affecting outcomes which are also unbalanced. To deal with this we turn to a border experiment, to try to compare Anglophone versus Francophone cities facing identical circumstances.

Figure 8 shows West Africa where 5 Anglophone countries share borders with a number of Francophone countries. At these borders there are no significant waterways. We show cities within a 100 km buffer of the borders involved. Results are almost the same if use a 125 km or 150 km

---

[19] The effect is about 25% smaller for the Poisson but that seems due to the Poisson functional specification.

[20] For the typical city under a specification where all covariates are interacted with the Anglophone indicator, the differential is 44%.

buffer. We choose the smaller buffer, but dropping below 100 km loses too many cities. To refine the border experiment, we break border segments into 14 finer portions, grouping nearby cities into natural clusters, to try to control for unobserved geographic or other influences. These clusters are given in Figure 8. Clusters around Ghana and Gambia are country-pair clusters, while those for Sierra Leone and Guinea split naturally into an east and west group. For Nigeria with its huge border, ignoring the green border cluster in southwest Cameroon for the moment, we first match all Anglophone cities to the nearest Francophone city (outside of Anglophone-Francophone Cameroon) creating 8 clusters and then we assign any so far unmatched Francophone cities to the nearest cluster. This algorithm is consistent with the other cluster groupings, except the cities in Burkina Faso have no Ghana counterpart, so these cities are neutralized by their cluster FE. An issue for country borders is that part of the 100 km Cameroon (green border) buffer was under British control after World War I through to the mid-1960's. We did the analysis both excluding and including this area, which is 1 cluster with cities that are in both Francophone and Anglophone Cameroon, with no Nigerian counterpart. Clearly the Anglophone Cameroon cities have conflicting effects: British heritage versus French rule for 50 years. We think it is better to exclude the area, but results do not vary significantly (Baruah et al., 2017).

With these groupings, have we attained balance? Table A7 shows our key covariates from Column 4 of Table 1 regressed on a constant and the Anglophone indicator. Column 1 shows that for the full sample there is a lack of balance for several covariates. All of that disappears for the border sample overall (column 2). When we control for the 14 clusters in column 3, one of the ten covariates has a significantly different mean. That is rainfall, which is not significantly different for Anglophone cities in general nor for the border sample. We believe true rainfall within clusters must be almost identical and that the column 3 difference reflects cross border measurement error based on placement of weather stations and interpolation.

In Table 4 for the border sample, we run the same leapfrog regressions as in Table 3 but with the smaller sample, we limit the controls to the 10 in column 2 of Table 1 plus the control for the log of 1990 initial cover and the time dummy. We show the results for a base case without FE's in row 1 and then in row 2, we add the city-cluster FE's. Results compare 35 Anglophone and 23 Francophone cities without the Anglophone-Francophone Cameroon segment. In the three columns we show the outcomes: log (count of LF), and log (ratio of LF to total count) and log (average area of leapfrog patches), with two rows. In Table 4, the Anglophone degree of leapfrogging is significantly higher in both specifications, with somewhat larger point estimates than in Table 3: 0.79 for the city-cluster FE's in Table 4 versus 0.61 in Table 3. For the ratio of leapfrog to all patches, again magnitudes are higher than in Table 3 (0.54 versus 0.32). As is the case in general, areas of patches do not differ by colonial origin. Table 4 gives strong evidence that it is colonialism and not other factors driving our results for our key measure, leapfrogging.[21]

6.3 Robustness

The next issue is robustness of Table 3 leapfrog results to other considerations. For that we turn to Table 5. In Table 5 in column 1 we show the base case. As in Table 3, in each of the three rows we show the outcomes: log (count of LF), and log (ratio of LF to total count), and log (average area of leapfrog patches). In columns 2 and 3, first we experiment with types of leapfrog measures. Column 2 removes from the counts and areas any developments that are just one (isolated) pixel (38 m x 38m), as an attempt to deal with an obvious source of mis-measurement of built cover. Column 3 uses a buffer around newly built areas of 60m rather than 300m in defining LEI and LF developments. In both cases, the impact on point estimates is fairly minimal. The rest of the columns deal with sampling issues.

Column 4 worries about blooming of night lights in bigger cities, which then add non-urban areas within the lights boundary, where Anglophone versus Francophone differentials might exist for other reasons. Column 4 trims the top and bottom 5% of cities in terms of maximum distance from the centre to any part of the lights boundary. That does diminish effects, but the leapfrog coefficient is still 0.42. We note however (but not shown in the table), that if we defined the area of the city as a smoothed 2014 built area cover, that would bias our results. For the leapfrog count outcome, the coefficient is minimally affected, but the ratio of LF to all patches then has a coefficient of zero. By cutting on smoothed cover, we mechanically tend to exclude areas with more leapfrogging relative to other patches.

Columns 5 and 6 turn to different country samples which are potentially more problematic. Column 5 removes countries which were initially German colonies before being assigned to Britain or France after World War I. Dropping those countries (Namibia, Tanzania, Togo, and Cameroon) has some effect on results but the leapfrog coefficient is still 0.49. Dropping Nigeria which is a big portion of the sample, if anything, strengthens results. Columns 7 and 8 look at colonial and non-colonial origin cities. Colonial origin cities are those built after 1800, which seem to be colonial constructs. The point estimate in column 7 for colonial original cities is almost 78% larger, hinting at stronger differences for cities built from scratch under the different colonial regimes. The intuition is appealing. Column 8 shows that for non-colonial original cities we still get the same effect on leapfrog counts as the base case. Column 9 focuses on the sample of 40 cities for which we assembled OSM data. Here the point estimates on absolute and relative LF counts are in line with the rest of the data. That is reassuring for the applicability of the gridiron results. In Appendix Table A9, we show these same experiments on robustness for Table 1 outcomes; magnitudes vary

---

[21] To complete the picture we also reran Table 1 specifications on openness and area as reported in Table A8. However the sample size is very small (54) and results are insignificant.

some in the columns with smaller samples, but the general pattern holds strongly.

One final experiment involves a look at heterogeneity and other historical infleucnes on spatial lay-out. Similar to Michalopoulos and Papaioannou (2013), we measure pre-colonial political institutions using Murdock (1967)'s "Jurisdictional Hierarchy Beyond the Local Community Level" index. We assign African cities the index of 0-4, where 4 is very strong pre-colonial statehood [22]. Some of our cities include different indexed areas and we did an area weighted average, so the measure is somewhat continuous up to 4. We keep the Anglophone indicator and add the statehood measure and its interaction with Anglophone. Results are in Table A10. Column 1 on leapfrogging is the only one where the interaction variable of interest is significant. It says the Francophone-Anglophone gap narrows as the index of pre-colonial statehood rises. One interpretation is that the French had less success in imposing their institutions and practices in areas where there was prior strong statehood. That seems plausible, but significant results only appear for that outcome.

In sum, the extensive list of robustness checks and the border experiment suggest the quantitative leapfrog and most overall sprawl results hold under different measures and samples and experimental contexts.

7 The so-what question: Public policy relevance

The planning literature argues that having less compact and more irregularly laid out cities raises the cost of infrastructure provision. In Africa we would further argue that such higher costs will lower the likelihood of receiving public infrastructure provision. To assess the reduced form impact, we use data from the Demographic and Health Survey (DHS) on whether a family has a piped water connection (with the alternatives being a shallow or deeper well or having no water connection), an electricity connection, a telephone land line connection, or a (flush) toilet connected to a public sewer system. About 63% and 75% of households in our sample are connected to water and electricity respectively, while having a flush toilet connected to a sewer system or having a landline are at 13% and 6% respectively. We tend to focus on the first 3 outcomes.

DHS uses cluster sampling of 20-30 households in a neighbourhood and we restrict attention to clusters defined by DHS to be in urban areas. We cover about 45,000 households in 60 Francophone cities in 7 countries and 133 Anglophone cities in 11 countries. We constrain DHS surveys to be within 2 years each of 2000 and 2014, the base years for which we measure openness and leapfrogging, and we control for which year before and after the base years a survey is.

The specification differs in two ways relative to equation (1), apart from the outcomes involving infrastructure connections and the specification being a LPM. First, the treatment variables are

---

[22] Index of value 0 indicates stateless societies "lacking any form of centralized political organization"; 1 indicates petty chiefdoms; 2 indicates paramount chiefdoms; 3 and 4 indicate part of large states.

measures of leapfrogging and openness in the local area around where a household lives. Second identification is based on city-year fixed effects, or within city variation in servicing. We are not measuring an Anglophone (versus Francophone) effect per se, since such colonial influences may reflect other public policy elements. Note colonial origin would be neither a valid instrument nor reduced form measure. We are trying to represent a cost effect of lack of compactness on public service provision. However we can and do look at whether there is a differential in impact of openness and leapfrogging from being in a Francophone versus an Anglophone city. Note we will examine the effects of both leapfrogging and openness measures on the likelihood of a connection. While the two measures are related, greater openness might indicate greater distance on average between residences, while leapfrogging indicates that some developments are scattered. Both measures will matter and both suggest connecting households will require more piping, wiring or lines per household entailing more costs.

The challenge in implementation concerns location. Within an urban area, cluster locations are randomized within 2 km by randomly picking a directional ray (angle) from the true cluster centre and then choosing a location randomly along that ray within 2 km of the cluster centre. Under this algorithm, while locations near the true location are more likely to be chosen, the randomized location is equally likely to be in any ring out from the true location up to the 2 km. We draw a 2 km circle around the specified location and look at the effect of more leapfrog patches and openness in that circle on provision, conditional on far it is from the city centre, and other controls. One could view this as a measure of, say, how likely a cluster is to be in a leapfrog patch, but we interpret it as a measure of the overall degree of leapfrogging and openness in the surrounding area. To exploit economies of scale in construction, cities roll out public utilities in large spatial zones. The higher the degree of leapfrog and open development in an area, the less likely it is to be serviced, because roll-out is more costly. Regardless, because of the randomization of location, the variables of interest are measured with error. We could not think of an instrument which both met the exclusion restriction and had power.[23] Given the measurement error involved we did not anticipate seeing strong patterns in the data, but were surprised.

Results are in Table 6 covering 42,748-44,561 households for each attribute. In a linear probability formulation, each attribute has two columns. In the table's reduced form specification, both columns have basic supply and demand controls, with the count of leapfrog developments as the cost factor in the first and that and the share of built cover in the surrounding neighborhood (lack of openness) in the second. An increased count of leapfrog patches significantly reduces the likelihood of connections in 3 of the 4 outcomes, having no effect on piped water. Effects are fairly small. A one standard deviation (5.6) increase in count of LF patches reduces the probability of an electricity connection by 0.018 from a mean of 0.74, although for flush toilets connected to a sewer

---

[23] e.g. Using the propensity of surrounding areas to have LF development for whether cluster is recorded in a leapfrog patch does not.

system the decrease is 0.025 from a mean of 0.12. In the second set of columns we add in the extent of built cover, which has significant impacts on all connections except a landline. Again effects are fairly small. A one standard deviation (0.31) in built cover increases the likelihood of an electricity connection by 0.016, water by 0.025 (from a mean of 0.64), and flush toilet connected to a sewer system by 0.036. In Table A11 we examine whether for either leapfrogging or openness, there is a differential Anglophone effect. In only one of the eight cases (built cover for flush toilets to a sewer system) is the Anglophone interaction term significant. Overall, effects are consistent, suggesting that, despite measurement error, we indeed find a negative relationship between increased sprawl and public utility provision in both Anglophone and Francophone cities.

8 Conclusions

The literature on colonialism in Africa suggests that, compared to the British, the French imposed more comprehensive city wide land use planning, including the lay-out of roads. The theoretical literature in economics suggests that an omniscient and benevolent city planner would create a more compact city which encourages fluidity of movement than *laissez faire* development and that leapfrogging is related to intensity of centre city development. The empirical literature suggests that areas which through centralized control are more regularly laid out on a grid system will have higher levels of future development and/or land values. The African context of colonialism provides an experiment to show that choice of institutions which involve more centralized control within each city, as in Francophone compared to Anglophone cities, lead to more compact cities.

Specifically the paper shows that Francophone African cities have more grid-like structures in their core areas. Anglophone cities have a city wide index of openness which is 23% higher. Their intensity of land use is much lower at the centre and, in contrast to Francophone cities, the intensity of land use gradient is almost flat. Anglophone cities are more sprawled. Correspondingly, with new development, Anglophone cities have about 61% more leapfrog patches, a number that is robust to a border experiment and many experiments with definitions and relevant cuts on the data in terms of samples. There is a consequence to having greater leapfrog development. Such areas are less likely to receive connections to public utilities, such as electricity, phone landlines, piped water, and city sewers.

References

Acemoglu, D., D. Cantoni, S. Johnson, and J. A. Robinson (2011, December). The consequences of radical reform: The French revolution. *American Economic Review 101*(7), 3286–3307.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001, December). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Ahmed, A. and R. D. Dinye (2011). Urbanisation and the challenges of development controls in Ghana: A case study of wa township. *Journal of Sustainable Development in Africa 13*(7), 210–235.

Ban, Y., P. Gong, and C. Giri (2015). Global land cover mapping using earth observation satellite data: Recent progresses and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing 103*, 1–6.

Banerjee, A. and L. Iyer (2005, September). History, institutions, and economic performance: The legacy of colonial land tenure systems in India. *American Economic Review 95*(4), 1190–1213.

Barton, H., M. Grant, and R. Guise (2003). *Shaping neighbourhoods: a guide for health, sustainability and vitality*. Taylor & Francis.

Baruah, N. G., J. V. Henderson, and C. Peng (2017). Colonial legacies: shaping African cities. *SERC working paper*.

Bertrand, M. (2004). Land management and urban development projects: A comparison of experiences in French-speaking and English-speaking West Africa. *International Development Planning Review 26*(1), 83–96.

Bleakley, H. and J. Lin (2012). Portage and path dependence. *Quarterly Journal of Economics 127*(2), 587–644.

Brooks, L. and B. Lutz (2014). Vestiges of transport: Urban persistence at a micro scale. *GWU working paper*.

Brueckner, J. K. (2001). Urban sprawl: Lessons from urban economics. *Brookings-Wharton papers on urban affairs 2001*(1), 65–97.

Brueckner, J. K. (2005). Transport subsidies, system choice, and urban sprawl. *Regional Science and Urban Economics 35*(6), 715–733.

Burchfield, M., H. G. Overman, D. Puga, and M. A. Turner (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics 121*(2), 587–633.

Burley, J. (2016). The built environment and social interactions: Evidence from panel data. Working papers, University of Toronto.

Calderón, F., J. Oppenheimer, and N. Stern (2014). Better growth, better climate–the new climate economy report–the synthesis report. *Technical Report*.

Consortium, T.-A. (2010). *Public transport in Sub-Saharan Africa: major trends and case studies.* International Association of Public Transport (UITP), Brussels.

Crowder, M. (1964). Indirect rule—French and British style. *Africa 34*(03), 197–205.

Davis, O. A. and A. B. Whinston (1964). The economics of complex systems: The case of municipal zoning. *Kyklos 17*(3), 419–446.

Donaldson, D. and A. Storeygard (2016, November). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives 30*(4), 171–98.

Durand-Lasserve, A. (2004). Land for housing the poor in African cities: Are neo-customary processes an effective alternative to formal systems?'. *http://siteresources.worldbank.org/INTURBANDEVELOPMENT/Resources/336387-1268963780932/6881414-1268963797099/durand-lasserve.pdf.*

Ellickson, R. C. (2012). The law and economics of street layouts: How a grid pattern benefits a downtown. *Ala. L. Rev. 64*, 463.

Epple, D. and T. Nechyba (2004). Fiscal decentralization. In J. V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (1 ed.), Volume 4, Chapter 55, pp. 2423–2480. Elsevier.

Giri, C., Z. Zhu, and B. Reed (2005). A comparative analysis of the global land cover 2000 and modis land cover data sets. *Remote sensing of environment 94*(1), 123–132.

Glaeser, E. L. and M. E. Kahn (2010, May). The greenness of cities: Carbon dioxide emissions and urban development. *Journal of Urban Economics 67*(3), 404–418.

Grant, R. and P. Yankson (2003). City profile: Accra. *Cities 20*(1), 65–74.

Guiso, L., P. Sapienza, and L. Zingales (2016). Long-term persistence. *Journal of the European Economic Association 14*(6), 1401–1436.

Harari, M. (2017). Cities in bad shape: Urban geometry in India. *Octobor. http://real. wharton. upenn. edu/˜ harari/Harari_Papers/CityShapeHarariMarch2016_updated. pdf*.

Helsley, R. (2004). Urban political economics. In J. V. Henderson and J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (1 ed.), Volume 4, Chapter 54, pp. 2381–2421. Elsevier.

Helsley, R. W. and W. C. Strange (2007). Urban interactions and spatial structure. *Journal of Economic Geography 7*(2), 119–138.

Henderson, V. and A. Mitra (1996). The new urban landscape: Developers and edge cities. *Regional Science and Urban Economics 26*(6), 613–643.

Henderson, J. V., A. Storeygard, and U. Deichmann (2017). Has climate change driven urbanization in Africa? *Journal of Development Economics 124*(C), 60–82.

Hochman, O., D. Pines, and J.-F. Thisse (1995). On the optimal structure of local governments. *The American Economic Review 85*(5), 1224–1240.

Home, R. (2015). Colonial urban planning in Anglophone Africa. In C. N. Silva (Ed.), *Urban Planning in Sub-Saharan Africa: Colonial and Post-Colonial Planning Cultures*, pp. 53–66. Routledge.

Hortas-Rico, M. and A. Solé-Ollé (2010). Does urban sprawl increase the costs of providing local public services? Evidence from Spanish municipalities. *Urban Studies*.

La Porta, R., F. Lopez-de Silanes, C. Pop-Eleches, and A. Shleifer (2004). Judicial checks and balances. *Journal of Political Economy 112*(2), 445–470.

La Porta, R., F. Lopez-de Silanes, and A. Shleifer (2008). The economic consequences of legal origins. *Journal of Economic Literature 46*(2), 285–332.

Lee, A. and K. A. Schultz (2012). Comparing British and French colonial legacies: A discontinuity analysis of Cameroon. *Quarterly Journal of Political Science 7*(4), 365–410.

Libecap, G. D. and D. Lueck (2011). The demarcation of land and the role of coordinating property institutions. *Journal of Political Economy 119*(3), 426–467.

Liu, X., X. Li, Y. Chen, Z. Tan, S. Li, and B. Ai (2010). A new landscape index for quantifying urban expansion using multi-temporal remotely sensed data. *Landscape Ecology 25*(5), 671–682.

Mahoney, P. G. (2001). The common law and economic growth: Hayek might be right. *The Journal of Legal Studies 30*(2), 503–525.

Michaels, G. and F. Rauch (2017). Resetting the urban network: 117-2012. *The Economic Journal 128*(608), 378–412.

Michalopoulos, S. and E. Papaioannou (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica 81*(1), 113–152.

Michalopoulos, S. and E. Papaioannou (2016, July). The long-run effects of the scramble for Africa. *American Economic Review 106*(7), 1802–48.

Murdock, G. P. (1967). *Ethnographic Atlas*. University of Pittsburgh Press.

Njoh, A. J. (2015). *French Urbanism in Foreign Lands*. Springer.

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. *The Review of Economics and Statistics 94*(1), 20–36.

Oates, W. (1999). An essay on fiscal federalism. *Journal of Economic Literature 37*(3), 1120–1149.

Oto Peralias, D. and D. Romero-Ávila (2017). *Colonial theories of institutional development: toward a model of styles of imperialism*. Contributions to economics. Netherlands: Springer.

O'Grady, T. (2014). Spatial institutions in urban economies: How city grids affect density and development. *Harvard University*.

Pesaresi, M., D. Ehrlich, S. Ferri, A. Florczyk, S. Freire, M. Halkia, A. Julea, T. Kemper, P. Soille, and V. Syrris (2016). Operating procedure for the production of the global human settlement layer from landsat data of the epochs 1975, 1990, 2000, and 2014. Technical report, JRC Technical Report.

Pesaresi, M., G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. Kauffmann, T. Kemper, L. Lu, et al. (2013). A global human settlement layer from optical hr/vhr rs data: concept and first results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6*(5), 2102–2131.

Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.

Redfearn, C. L. (2009). Persistence in urban form: The long-run durability of employment centers in metropolitan areas. *Regional Science and Urban Economics 39*(2), 224 – 232.

Rosin, P. L. (1999). Measuring rectangularity. *Machine Vision and Applications 11*(4), 191–196.

Rossi-Hansberg, E. (2004, January). Optimal land use and zoning. *Review of Economic Dynamics 7*, 69–106.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics 125*(3), 1253–1296.

Shertzer, A., T. Twinam, and R. P. Walsh (2016, July). Race, ethnicity, and discriminatory zoning. *American Economic Journal: Applied Economics 8*(3), 217–246.

Silva, C. N. (2015). *Urban planning in Sub-Saharan Africa: colonial and post-colonial planning cultures*. Routledge.

Trubka, R., P. Newman, and D. Bilsborough (2010). The costs of urban sprawl-infrastructure and transportation. *Environmental Design Guide*.

Turner, M. A. (2007). A simple theory of smart growth and sprawl. *Journal of Urban Economics 61*(1), 21–44.

Žunić, D., C. Martinez-Ortiz, and J. Žunić (2012). Shape rectangularity measures. *International Journal of Pattern Recognition and Artificial Intelligence 26*(06), 1254002.

Table 1: Sprawl: Openness and area

	Openness				Area			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Anglophone dummy	0.229***	0.206***	0.213***	0.231***	0.355**	0.173	0.121	0.277***
	(0.045)	(0.045)	(0.045)	(0.047)	(0.151)	(0.127)	(0.124)	(0.089)
Lag t-1 ln country GDP per capita				0.037				0.013
				(0.053)				(0.084)
Ln annual population growth 90 to 12				-0.854				7.349***
				(1.128)				(2.064)
Ln projected city population 1990				-0.176***				0.867***
				(0.021)				(0.037)
Internal geography	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Fringe geography	No	No	Yes	Yes	No	No	Yes	Yes
R^2	0.077	0.137	0.185	0.374	0.014	0.413	0.423	0.770
N	307	307	307	281	307	307	307	281

Note: Dependent variable is ln openness in year 2014 in columns 1-4, and ln area in columns 5-8. Internal geography controls include ln ruggedness, ln rainfall, ln elevation range, coast dummy, interaction of ln coast length with coast dummy, interaction of ln distance to coast with non-coastal dummy. Fringe geography controls include fraction of river area, fraction of lake area, fraction of forest, fraction of shrubs, fraction of crops, fraction of wetlands and water, fraction of sparse vegetables and bare land. Robust standard errors are applied.

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

Table 2: Intensity by rings in 1990

	(1) 1km	(2) 2km	(3) 3km	(4) 4km	(5) 5km	(6) 6km
Anglophone Dummy	-0.389**	-0.658***	-0.804***	-0.781***	-0.420*	-0.075
	(0.165)	(0.164)	(0.194)	(0.229)	(0.241)	(0.263)
Ln ring total pixel	1.070***	0.534	-0.017	0.274*	0.546***	0.421***
	(0.355)	(0.465)	(0.211)	(0.164)	(0.167)	(0.145)
Anglophone mean	1.313	2.654	2.526	2.245	2.138	2.185
Francophone mean	1.684	3.329	2.877	2.522	2.373	2.424
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.200	0.287	0.395	0.477	0.511	0.429
N	312	316	307	283	254	218

Note: Dependent variable is ln built-up area in 1990 in a ring. City characteristics control variables include ln projected city population in 1990, ln country GDP per capita in 1990. Geographic controls include ln ruggedness, ln rainfall, ln elevation range, coast dummy, interaction of ln coast length with coast dummy, interaction of ln distance to coast with non-coastal dummy, fraction of river area, fraction of lake area, fraction of forest, fraction of shrubs, fraction of crops, fraction of wetlands and water, fraction of sparse vegetables and bare land. Anglophone mean and Francophone mean report mean built-up area in both groups in square kilometers. Standard errors are clustered at city level.

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

Table 3: Leapfrogging

	Ln count of LF				Ln LF minus ln total patches				Ln avg. LF area
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Anglophone dummy	0.907***	0.700***	0.642***	0.607***	0.339***	0.282***	0.246**	0.319***	-0.046
	(0.139)	(0.140)	(0.135)	(0.163)	(0.099)	(0.104)	(0.101)	(0.119)	(0.064)
Ln initial cover 1990	0.656***	0.572***	0.553***	0.278***	-0.081**	-0.126***	-0.143***	-0.311***	0.014
	(0.049)	(0.048)	(0.048)	(0.063)	(0.033)	(0.033)	(0.033)	(0.040)	(0.021)
Year dummy 2014	0.517***	0.508***	0.504***	0.491***	0.119**	0.121**	0.118**	0.119**	0.140***
	(0.065)	(0.065)	(0.066)	(0.068)	(0.053)	(0.054)	(0.055)	(0.056)	(0.034)
Lag t-1 ln country GDP per capita				-0.229				-0.146	0.067
				(0.147)				(0.111)	(0.051)
Ln annual population growth 90 to 12				10.760***				5.222**	1.989
				(3.198)				(2.413)	(1.226)
Ln projected city population 1990				0.730***				0.467***	0.049
				(0.099)				(0.067)	(0.034)
Internal geography	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Fringe geography	No	No	Yes	Yes	No	No	Yes	Yes	Yes
R^2	0.446	0.502	0.517	0.586	0.053	0.099	0.126	0.231	0.117
N	606	606	606	551	606	606	606	551	525

Note: Internal geography controls include ln ruggedness, ln rainfall, ln elevation range, coast dummy, interaction of ln coast length with coast dummy, interaction of ln distance to coast with non-coastal dummy. Fringe geography controls include fraction of river area, fraction of lake area, fraction of forest, fraction of shrubs, fraction of crops, fraction of wetlands and water, fraction of sparse vegetables and bare land. Standard errors are clustered at city level.

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

Table 4: Identification based on border sample

	(1) Ln count of LF	(2) Ln LF minus ln total patches	(3) Ln avg. LF area
Basic controls			
Anglophone dummy	0.952*	0.763**	-0.118
	(0.496)	(0.340)	(0.162)
City cluster FE's			
Anglophone dummy	0.789***	0.537***	-0.210
	(0.290)	(0.178)	(0.135)
R^2	0.673	0.431	0.271
N	108	108	103

Note: Controls include ln initial cover 1990, year dummy 2014, lag t-1 ln country GDP per capita, ln annual population growth 90 to 12, ln projected city population in 1990, ln average ruggedness, and coast dummy. Standard errors are clustered at city level.

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

Table 5: Leapfrogging: Robustness

	(1) Base	(2) No single pixel patches	(3) 60 meter buffer	(4) Trim	(5) No German colonies	(6) No Nigeria	(7) Colonial origin	(8) Non colonial origin	(9) 40 cities
Ln count of LF									
Anglophone	0.607***	0.555***	0.526***	0.420***	0.492**	0.670***	1.079*	0.629***	0.646***
dummy	(0.163)	(0.165)	(0.138)	(0.161)	(0.197)	(0.217)	(0.606)	(0.179)	(0.229)
Ln ratio of LF									
Anglophone	0.319***	0.279**	0.237***	0.227*	0.175	0.336**	0.150	0.362***	0.292
dummy	(0.119)	(0.126)	(0.073)	(0.120)	(0.142)	(0.159)	(0.319)	(0.128)	(0.218)
Ln avg. LF area									
Anglophone	-0.046	-0.034	-0.088**	-0.026	-0.168**	-0.042	0.141	-0.029	0.136
dummy	(0.064)	(0.056)	(0.044)	(0.059)	(0.081)	(0.070)	(0.142)	(0.071)	(0.113)
R^2	0.586	0.582	0.679	0.515	0.606	0.556	0.698	0.550	0.596
N	551	551	551	505	489	330	69	502	49

Note: Columns 1-6 and 8 include same controls as column 4,8,9 in Table 3. Columns 7 and 9 include same controls as in Table 4, except that ln annual population growth 90 to 12 in city level is replaced with ln annual urban population growth 90 to 14 in country level in column 7 due to severe missing data problem in city population in the colonial origin sample. Standard errors are clustered at city level in columns 1-6 and 8, and robust standard errors are applied in columns 7 and 9. Adjusted R^2 and N are reported for ln count of LF.

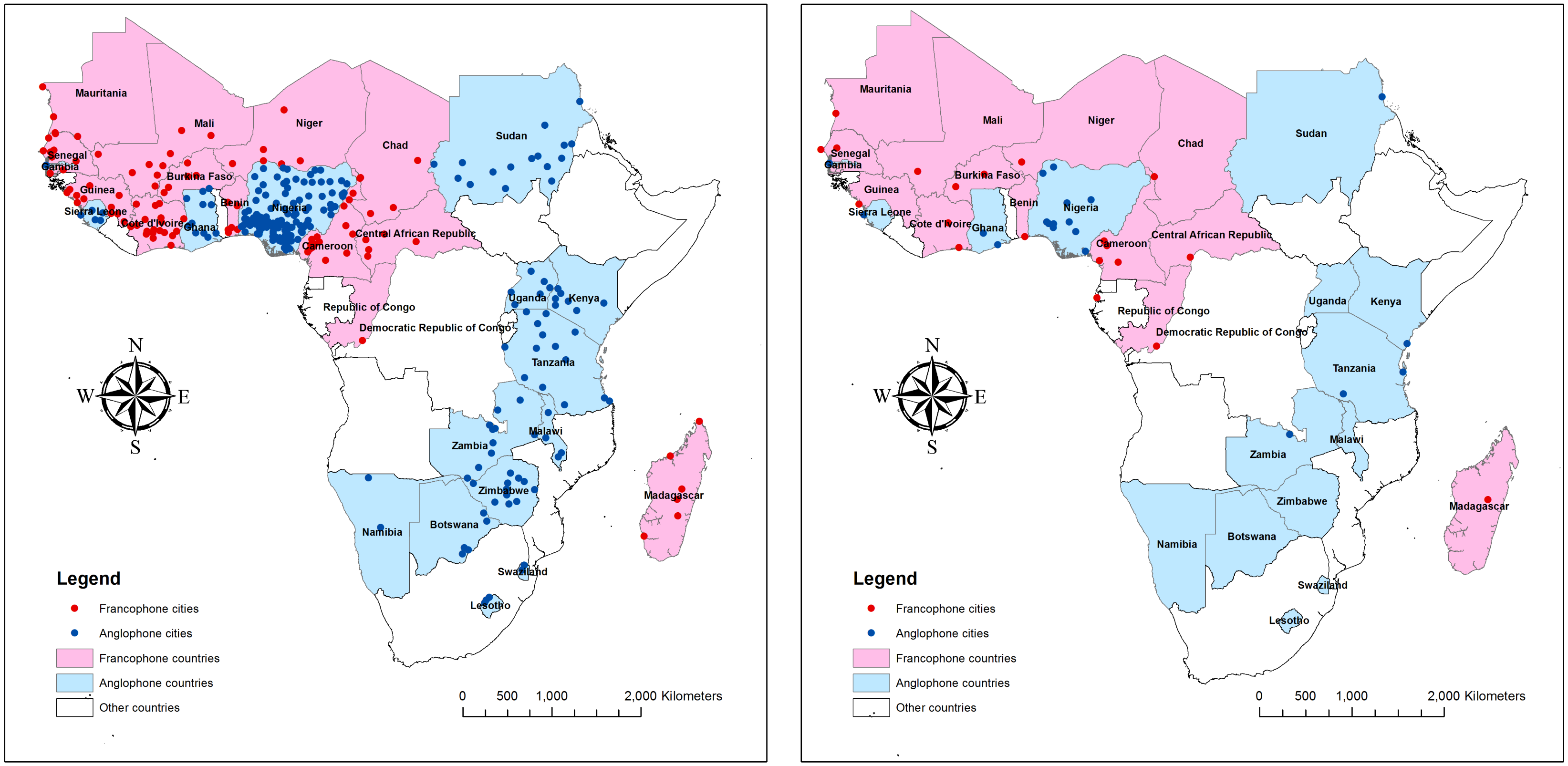
* p <0.1, ** p <0.05, *** p <0.01. Standard errors in parentheses.

Table 6: Public utility connection

	Has electricity		Has piped water		Has flush toilet		Has phone land line	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Count of LF	-0.0033**	-0.0026*	-0.0002	0.0008	-0.0045***	-0.0031***	-0.0019**	-0.0020**
	(0.0014)	(0.0014)	(0.0013)	(0.0014)	(0.0010)	(0.0010)	(0.0008)	(0.0009)
Share of built cover in buffer		0.0526*		0.0797**		0.1168***		-0.0026
		(0.0298)		(0.0364)		(0.0266)		(0.0128)
Ln buffer center distance	-0.0727***	-0.0666***	-0.0700***	-0.0608***	-0.0383***	-0.0248***	-0.0124***	-0.0127***
	(0.0066)	(0.0075)	(0.0088)	(0.0096)	(0.0062)	(0.0073)	(0.0035)	(0.0038)
Ln buffer ruggedness	0.0553***	0.0507***	0.0565***	0.0495**	-0.0054	-0.0157	-0.0012	-0.0010
	(0.0165)	(0.0171)	(0.0204)	(0.0210)	(0.0120)	(0.0124)	(0.0062)	(0.0064)
Buffer has river of lake	0.0209	0.0232	0.0361	0.0395	-0.0057	-0.0007	0.0041	0.0040
	(0.0214)	(0.0214)	(0.0272)	(0.0270)	(0.0174)	(0.0175)	(0.0122)	(0.0121)
Household size	0.0086***	0.0086***	0.0024***	0.0024***	0.0026***	0.0026***	0.0069***	0.0069***
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Sex of household head: Male	-0.0075	-0.0076	-0.0092*	-0.0094**	-0.0159***	-0.0161***	-0.0096***	-0.0096***
	(0.0050)	(0.0050)	(0.0047)	(0.0047)	(0.0041)	(0.0040)	(0.0028)	(0.0028)
Highest educational level of head: Primary	0.0367***	0.0374***	0.0117	0.0128*	-0.0052	-0.0036	0.0126***	0.0126***
	(0.0072)	(0.0072)	(0.0073)	(0.0073)	(0.0045)	(0.0045)	(0.0034)	(0.0034)
Highest educational level of head: Secondary	0.1739***	0.1740***	0.0519***	0.0521***	0.0407***	0.0410***	0.0424***	0.0424***
	(0.0070)	(0.0070)	(0.0071)	(0.0071)	(0.0047)	(0.0047)	(0.0038)	(0.0038)
Highest educational level of head: Higher	0.2877***	0.2877***	0.0578***	0.0577***	0.1652***	0.1652***	0.1124***	0.1124***
	(0.0088)	(0.0088)	(0.0094)	(0.0094)	(0.0089)	(0.0089)	(0.0072)	(0.0072)
Highest educational level of head: Unknown	0.1417***	0.1417***	0.0457**	0.0457**	0.0124	0.0124	0.0027	0.0026
	(0.0247)	(0.0247)	(0.0225)	(0.0225)	(0.0165)	(0.0165)	(0.0150)	(0.0150)
Period dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.320	0.321	0.436	0.437	0.255	0.259	0.126	0.126
N	44517	44517	44561	44561	44500	44500	42748	42748

Note: Period dummies control the difference between the DHS survey years with year 2000 and 2014 when satellite data is available. Period dummies include 1 year before dummy, 1 year after dummy, 2 years before dummy, and 2 years after dummy. Standard errors are clustered at DHS cluster level.

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



(a) Full sample

(b) 40 cities

Table 1: Spatial distribution of sample cities

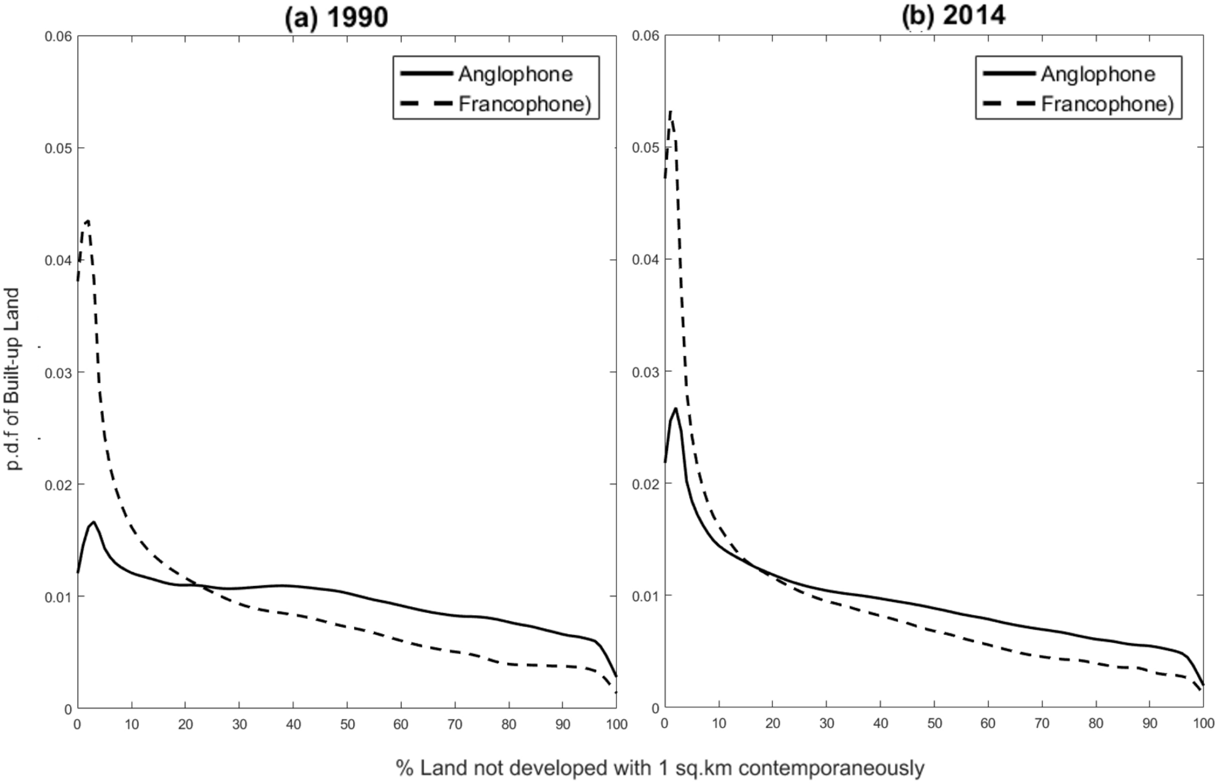
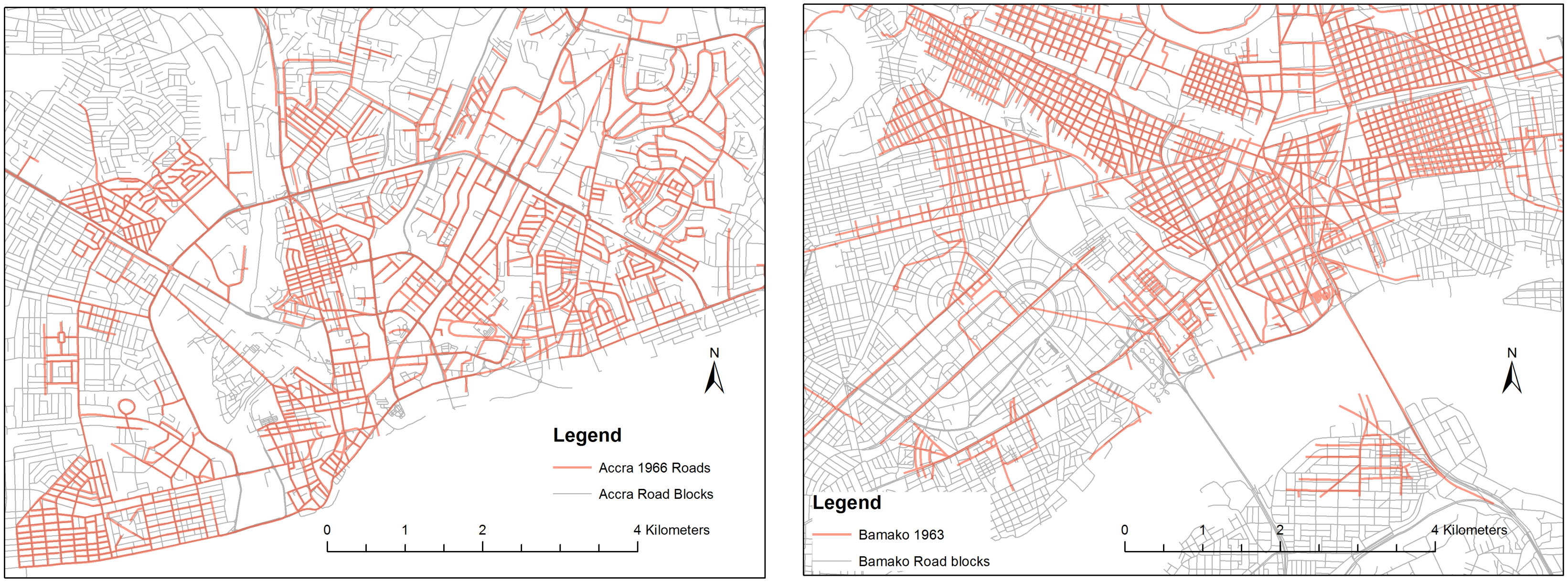


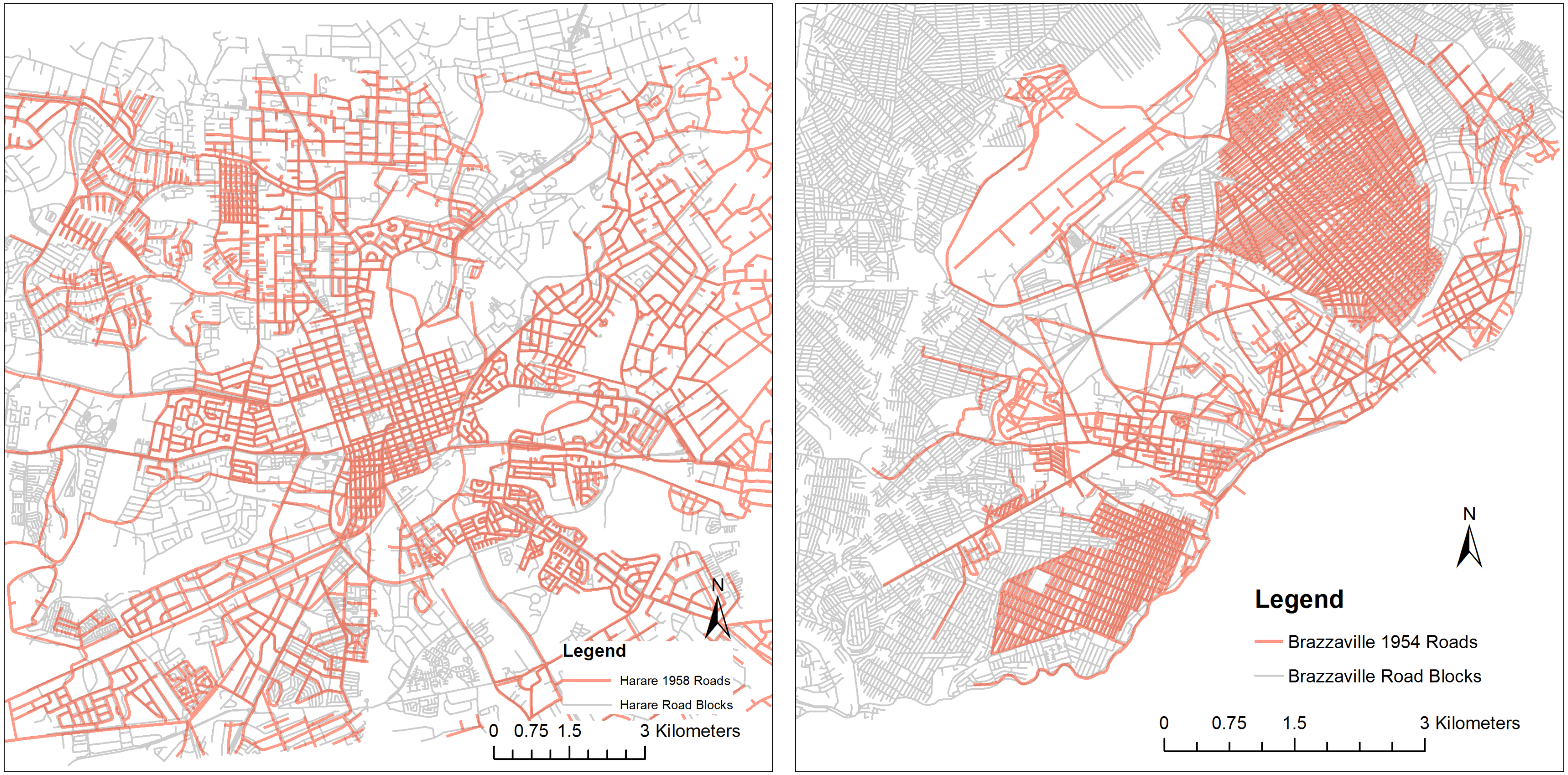
Figure 2: Probability function of Anglophone and Francophone built-up land across areas with different degrees of sprawl for (a) 1990 and (b) 2014



(c) Accra

(d) Bamako

Table 3a: Road blocks in Accra and Bamako



(e) Harare

(f) Brazzaville

Table 3b: Road blocks in Harare and Brazzaville

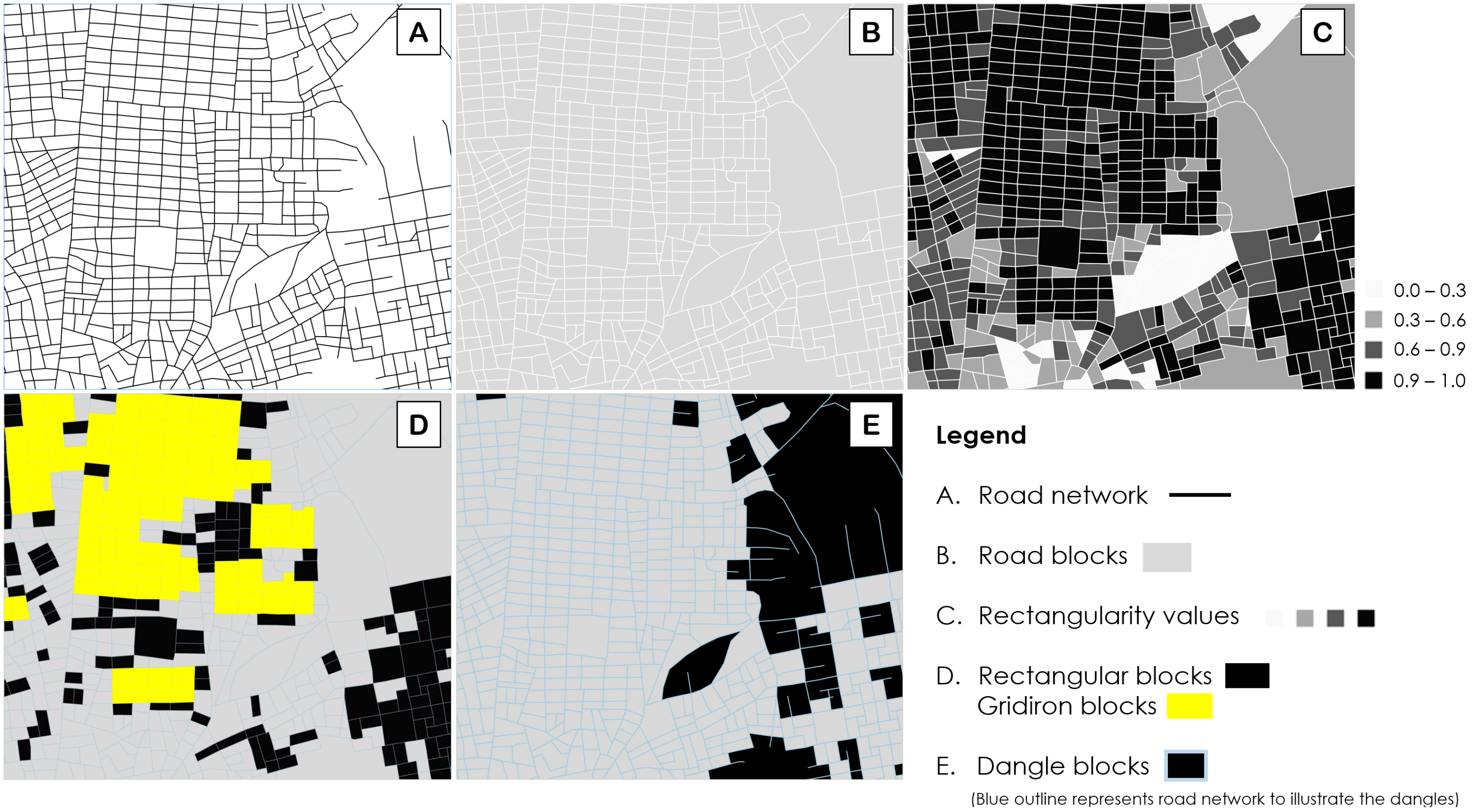


Figure 4: Road blocks and rectangularity

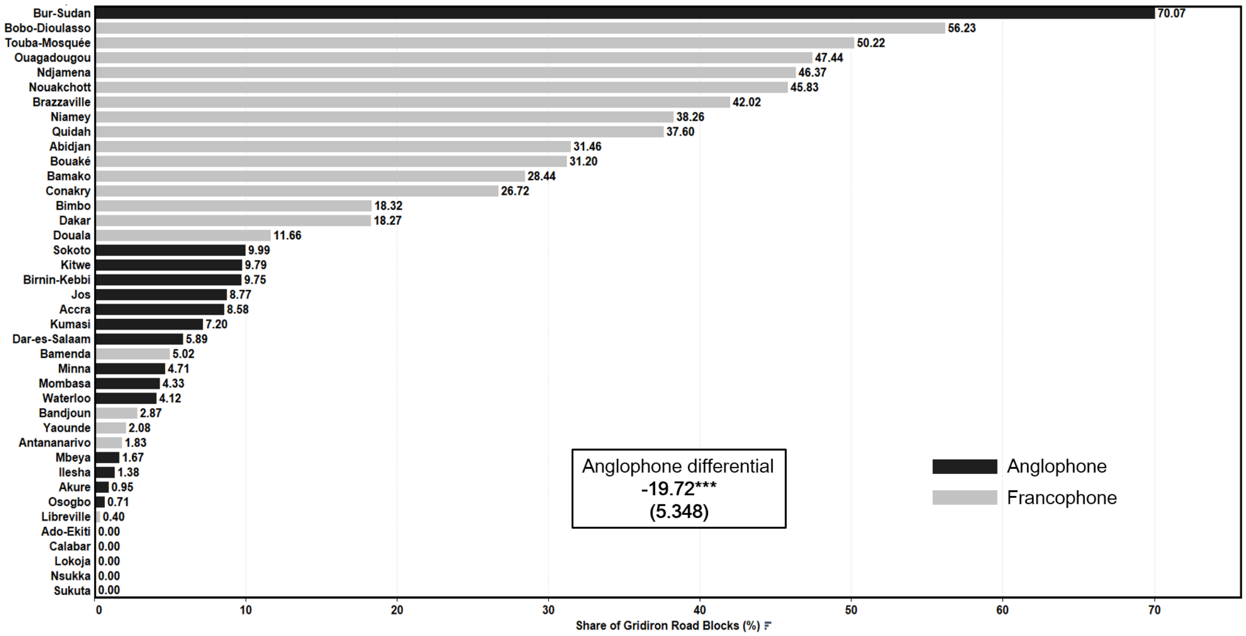


Figure 5: Share of gridiron road blocks within contemporary 5km

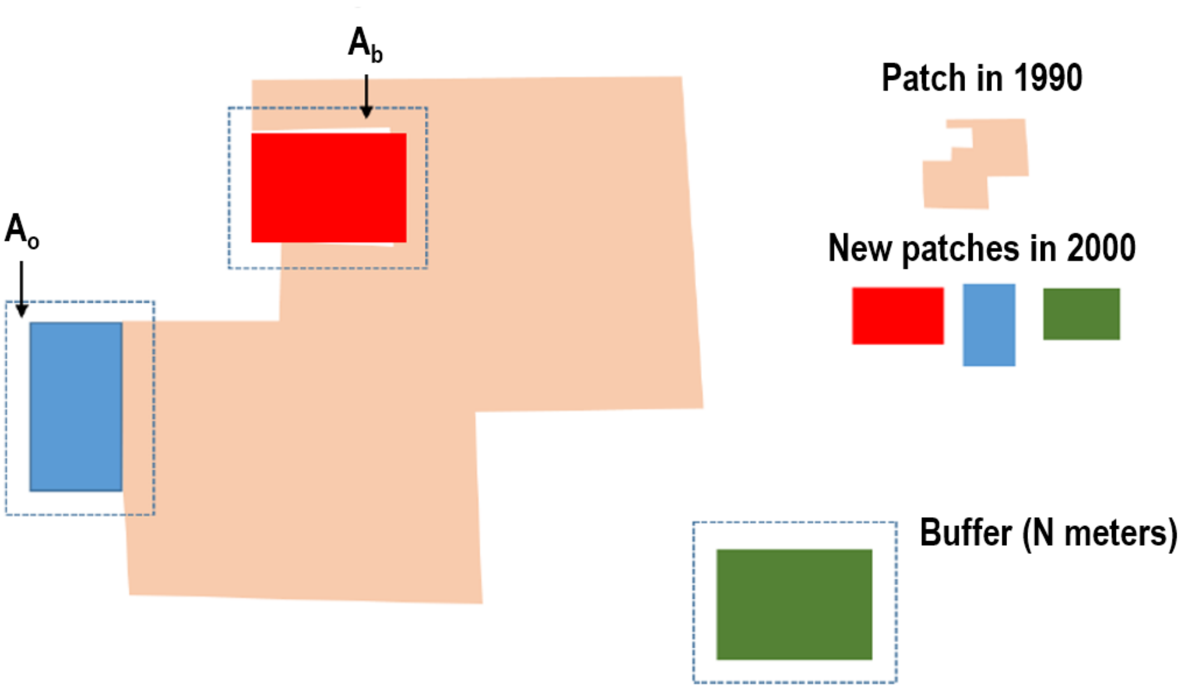


Figure 6: Illustration of using the landscape expansion index (Liu et al., 2010) for defining leapfrog patches

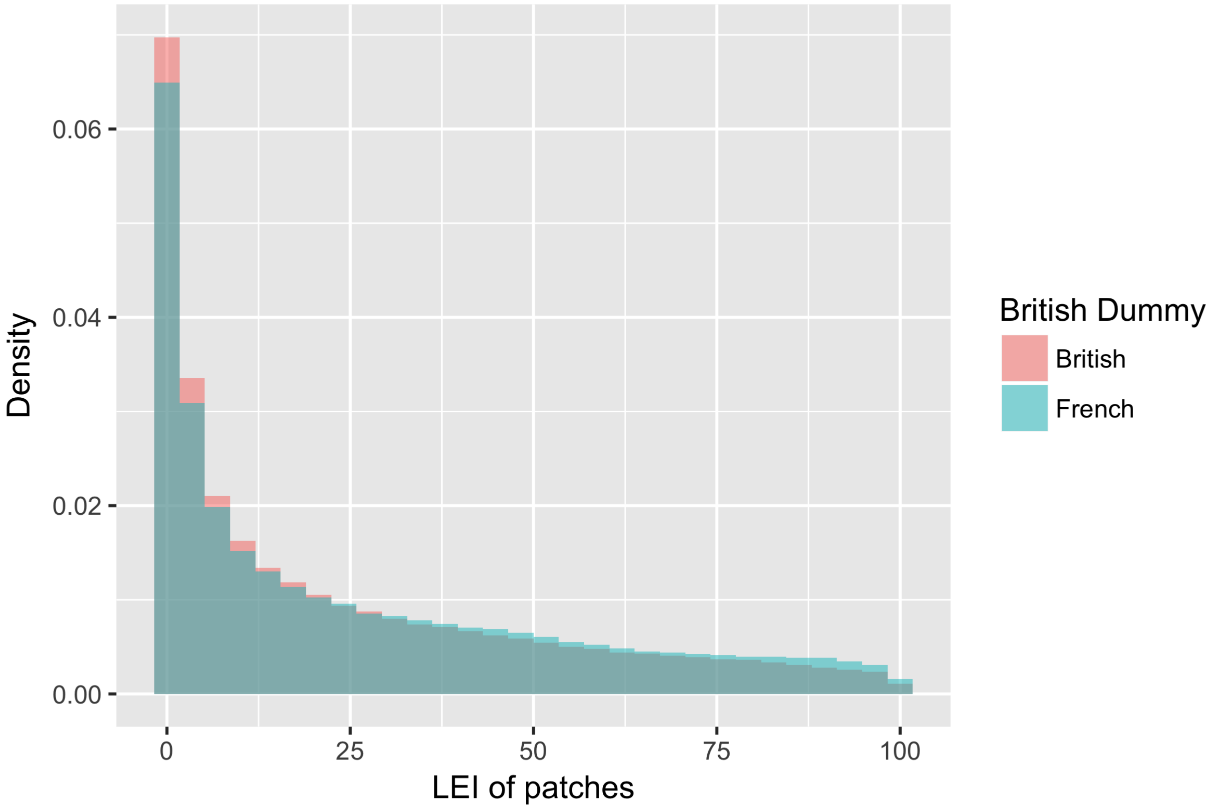


Figure 7: LEI distribution: Anglophone versus Francophone

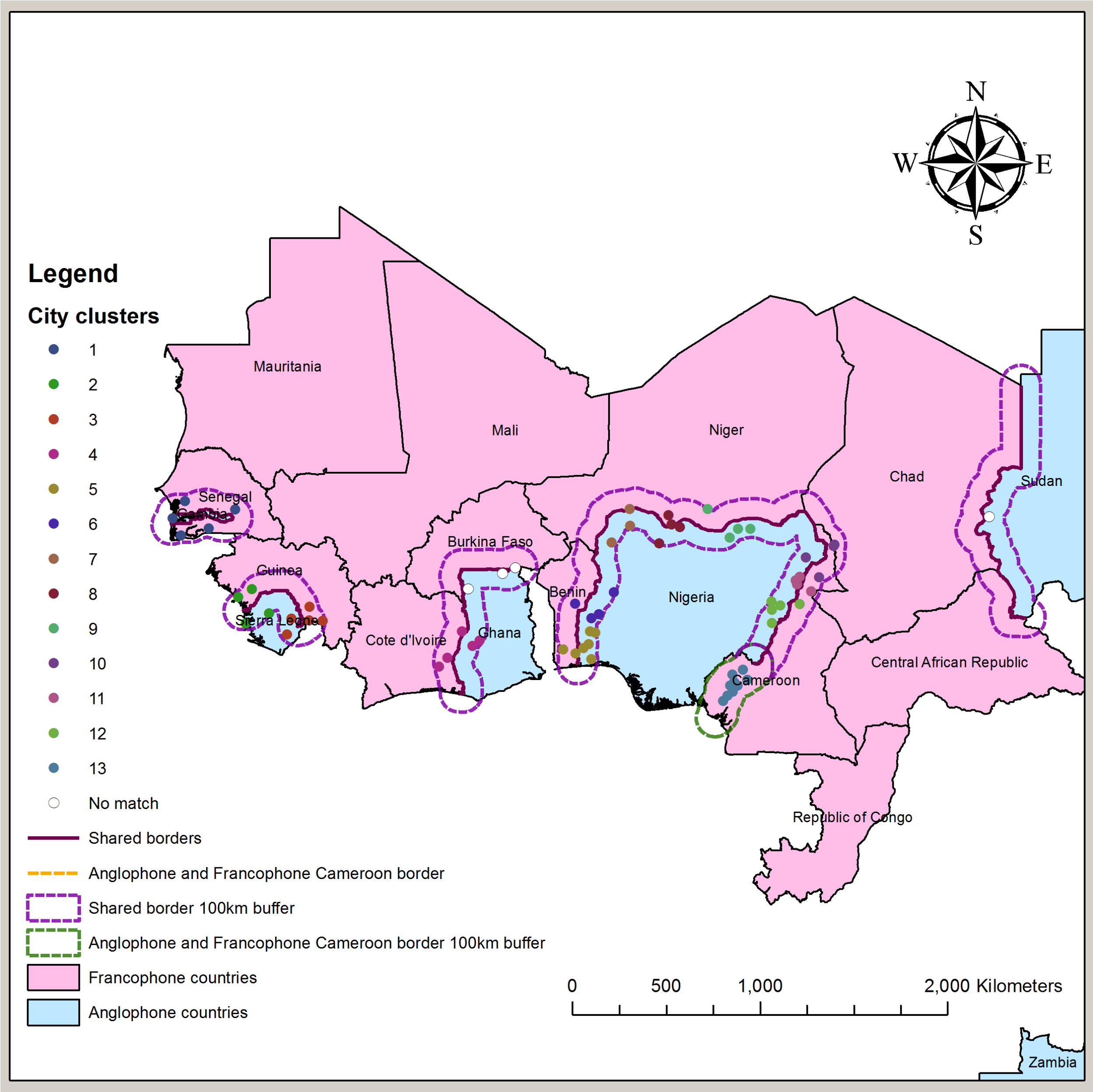


Figure 8: Shared borders

A On-line Appendices

A.1 Data

A.1.1 Data sources

Built-up cover of 38 meter resolution in year 1990, 2000, and 2014 is from Gobal Human Settlement layer (GHSL). Pesaresi, M.; Guo Huadong; Blaes, X.; Ehrlich, D.; Ferri, S.; Gueguen, L.; Halkia, M.; Kauffmann, M.; Kemper, T.; Linlin Lu; Marin-Herrera, M.A.; Ouzounis, G.K.; Scavazzon, M.; Soille, P.; Syrris, V.; Zanchetta, L., "A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results," Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of , vol.6, no.5, pp.2102,2131, Oct. 2013

Historical road blocks data for the 40 Open Street Map cities is derived from digitalising historical maps from Oxford and Cambridge library. Current road blocks data is extracted from Open Street Map (OSM). `https://www.openstreetmap.org`

Population of cities is from two sources: Citypopulation.de (Census data, 39 countries) `http://citypopulation.de` and Africapolis database (for Angola and Nigeria) `http://www.oecd.org/swac/ourwork/africapolis.htm`. The population of an urban area is the sum of the population of all "cities" falling within the lights boundary of an urban area.

National GDP per capita is from Penn World Table 9.0. `https://www.rug.nl/ggdc/productivity/pwt/`

National urban population is from World Bank. `http://wdi.worldbank.org/table/3.12`

River and lake GIS data is from Global Lakes and Wetlands Database. `https://www.worldwildlife.org/pages/global-lakes-and-wetlands-database`

Elevation and ruggedness variables are derived from Aster DEM elevation by NASA. `https://search.earthdata.nasa.gov/search?m=24.1875!3.234375!3!1!0!0%2C2`

Rainfall variables are from "The Climate Data Guide: Global (land) precipitation and temperature: Willmott & Matsuura, University of Delaware." `https://climatedataguide.ucar.edu/climate-data`

Fringe geographic controls are from Land use system of the World. Nachtergaele, F & Petri, Monica. (2008). Mapping Land Use Systems at global and regional scales for Land Degradation Assessment Analysis Version 1.1. `http://www.fao.org/geonetwork/srv/en/graphover.show?id=37048&fname=lus_ssa.png&access=public`

Public utility connections variables including electricity, phone land line, piped water and flush toilet are from Demographic and Health Surveys Program (DHS). `https://dhsprogram.com`

A.1.2 Statistics on variables

	mean	sd	min	max
Ln count of LF	3.17	1.67	0.0	7.6
Ln LF minus ln total patches	-2.46	0.90	-7.0	0.0
Ln average LF area	8.26	0.44	7.3	10.9
Ln openness index 2014	3.75	0.38	2.4	4.6
Ln light area	4.89	1.20	1.6	8.9
Anglophone dummy	0.73	0.44	0.0	1.0
Ln initial cover 1990	15.60	1.60	9.6	20.8
Year dummy 2014	0.51	0.50	0.0	1.0
Lag t-1 ln country GDP per capita	7.64	0.47	6.6	9.2
Ln annual population growth 90 to 12	0.03	0.02	-0.0	0.1
Ln projected city population 1990	11.31	0.92	10.3	15.7
Ln ruggedness	6.89	1.19	3.1	8.7
Ln rainfall	4.48	0.59	1.1	5.6
Ln elevation range	5.22	0.71	3.5	7.3
Coast dummy	0.04	0.20	0.0	1.0
Interaction ln coast length	0.42	2.07	0.0	12.0
Interaction ln distance to coast	12.29	2.65	0.0	14.4
Fraction of river area	0.01	0.03	0.0	0.2
Fraction of lake area	0.01	0.03	0.0	0.3
Fraction of forrest	0.20	0.30	0.0	1.0
Fraction of shrubs	0.17	0.27	0.0	1.0
Fraction of crops	0.40	0.37	0.0	1.0
Fraction of wetlands and water	0.01	0.05	0.0	0.4
Fraction of sparse vege and bare land	0.03	0.15	0.0	1.0
Observations	551			

Note: The sample is the same as column 4 in Table 3

A.1.3 City built cover boundary

We adopted a smoothing algorithm to define the city built cover boundary. First, we measured the area of total built cover for each 500m $\times$ 500m grid. Then the smoothing algorithm gives each grid the average built cover value of its neighbor grids and itself. The neighborhood is all queen and rook neighbors on the grid. Note if there is any grid in a neighborhood that has no built-up cover, the averaged built-up is set to be zero. This condition helps to eliminate scattered built-up and obtain continuous built cover area. Finally, we select the grids with neighbourhoods which average over 10% built cover, and use them to form the final built cover boundary of cities.

A.2 Appendix tables

Table A1: Sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Bohicon	Benin	0	89,553	166,611		
Djougou	Benin	0	47,383	81,341		
Lokossa	Benin	0	30,328	70,048		
Parakou	Benin	0	96,206	216,706		
Pobè	Benin	0	35,163	67,425		
Quidah	Benin	0	921,859	1,922,874		
Toviklin	Benin	0	35,688	66,505		
Francistown	Botswana	1	65,935	109,269	Yes	
Gaborone	Botswana	1	215,068	487,079	Yes	
Kanye	Botswana	1	30,552	47,698		
Molepolole	Botswana	1	35,517	67,791		
Selebi-Phikwe	Botswana	1	45,446	61,570		
Banfora	Burkina Faso	0	41,261	97,859		
Bobo-Dioulasso	Burkina Faso	0	262,478	645,198		
Koudougou	Burkina Faso	0	60,177	99,187		
Ouagadougou	Burkina Faso	0	578,653	2,213,074		
Ouahigouya	Burkina Faso	0	44,462	89,579		
Bafang	Cameroon	0	37,503	33,806		
Bamenda	Cameroon	0	129,657	413,538		
Bandjoun	Cameroon	0	129,500	359,215		
Bertoua	Cameroon	0	48,871	116,686		
Douala	Cameroon	0	935,407	2,691,721		
Dschang	Cameroon	0	39,347	80,013		
Edéa	Cameroon	0	52,976	74,076		
Foumban	Cameroon	0	60,988	96,722		
Garoua	Cameroon	0	154,400	287,668		
Guider	Cameroon	0	35,432	62,750		
Kousséri	Cameroon	0	58,443	108,520		
Kumbo	Cameroon	0	38,606	112,836		
Loum	Cameroon	0	40,726	60,213		
Maroua	Cameroon	0	133,940	243,578		
Mbouda	Cameroon	0	37,434	50,758		
Meiganga	Cameroon	0	32,793	40,857		
Ngaoundéré	Cameroon	0	87,298	198,223		

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Nkongsamba	Cameroon	0	88,275	112,347		
Yaounde	Cameroon	0	771,858	2,744,391		
Bambari	Central African Republic	0	38,985	43,081		
Berbérati	Central African Republic	0	45,426	110,757		
Bimbo	Central African Republic	0	492,970	995,932		
Bossangoa	Central African Republic	0	32,124	39,833		
Bouar	Central African Republic	0	39,766	40,765		
Carnot	Central African Republic	0	32,915	56,765		
Abéché	Chad	0	48,962	109,300		
Moundou	Chad	0	93,710	145,775		
Ndjamena	Chad	0	475,961	1,061,368		
Sarh	Chad	0	71,999	101,946		
Abengourou	Cote d'Ivoire	0	61,400			
Abidjan	Cote d'Ivoire	0	2,312,639	4,395,000		
Akoupé	Cote d'Ivoire	0	38,495			
Bondoukou	Cote d'Ivoire	0	35,283			
Bouaflé	Cote d'Ivoire	0	37,918			
Bouaké	Cote d'Ivoire	0	352,785	536,719		
Daloa	Cote d'Ivoire	0	130,708			
Danané	Cote d'Ivoire	0	34,582			
Dimbokro	Cote d'Ivoire	0	39,581			
Ferkéssédougou	Cote d'Ivoire	0	40,675			
Gagnoa	Cote d'Ivoire	0	112,890			
Issia	Cote d'Ivoire	0	30,922			
Katiola	Cote d'Ivoire	0	34,581			
Korhogo	Cote d'Ivoire	0	115,302			
Man	Cote d'Ivoire	0	94,435			
Odienné	Cote d'Ivoire	0	31,202			
Sinfra	Cote d'Ivoire	0	37,773			
Séguéla	Cote d'Ivoire	0	31,517			

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Yamoussoukro	Cote d'Ivoire	0	139,062			
Libreville	Gabon	0	394,152	694,622		
Sukuta	Gambia	1	357,893	460,450	Yes	Yes
Accra	Ghana	1	2,004,164	3,689,581	Yes	Yes
Bawku	Ghana	1	39,747	63,318		
Bolgatanga	Ghana	1	37,953	69,431		
Dzodze	Ghana	1	52,458			
Ho	Ghana	1	45,396	116,172		
Koforidua	Ghana	1	68,148	129,122	Yes	
Kumasi	Ghana	1	836,568	2,382,131		Yes
Nkawkaw	Ghana	1	35,816	48,870		
Sunyani	Ghana	1	46,279	76,966		
Tamale	Ghana	1	177,409	409,675		
Techiman	Ghana	1	34,094	69,700		
Wa	Ghana	1	45,405	71,967		
Yendi	Ghana	1	34,652	54,365		
Boké	Guinea	0	35,332	58,679		
Conakry	Guinea	0	942,708	1,824,765		
Fria	Guinea	0	41,303	53,703		
Guéckédou	Guinea	0	85,391	64,617		
Kamsar	Guinea	0	55,242	82,002		
Kankan	Guinea	0	80,409	180,127		
Kindia	Guinea	0	85,776	129,993		
Kissidougou	Guinea	0	59,539	86,954		
Labé	Guinea	0	40,570	84,218		
Macenta	Guinea	0	44,266	56,709		
Mamou	Guinea	0	45,178	63,059		
Nzérékoré	Guinea	0	88,082	181,799		
Eldoret	Kenya	1	116,456	285,187		
Garissa	Kenya	1	32,881	161,277		
Kisii	Kenya	1	47,004	74,984		
Kisumu	Kenya	1	194,711	326,009		
Kitale	Kenya	1	56,884	80,007		
Mombasa	Kenya	1	491,834	1,167,440	Yes	Yes
Nairobi	Kenya	1	1,516,055	5,044,352	Yes	
Nakuru	Kenya	1	170,002	336,431	Yes	

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Maputsoe	Lesotho	1	59,779	103,567		
Maseru	Lesotho	1	117,442	178,016	Yes	
Teyateyaneng	Lesotho	1	42,583	61,599		
Antananarivo	Madagascar	0	675,058	1,300,000		
Antsirabe	Madagascar	0	117,026			
Antsiranana	Madagascar	0	54,808			
Fianarantsoa	Madagascar	0	101,428			
Mahajanga	Madagascar	0	99,126			
Toliara	Madagascar	0	75,032			
Blantyre	Malawi	1	372,552	738,274	Yes	
Lilongwe	Malawi	1	268,767	799,762		
Mzuzu	Malawi	1	59,752	159,233		
Zomba	Malawi	1	48,517	99,277		
Bamako	Mali	0	758,125	2,452,195		
Gao	Mali	0	54,413	99,059		
Kayes	Mali	0	55,029	149,909		
Koutiala	Mali	0	55,163	167,010		
Mopti	Mali	0	76,285	134,933		
San	Mali	0	34,466	73,915		
Sikasso	Mali	0	87,024	261,123		
Ségou	Mali	0	92,519	188,365		
Tombouctou	Mali	0	31,338	64,488		
Kaédi	Mauritania	0	31,104	47,803		
Nouadhibou	Mauritania	0	61,209	113,789		
Nouakchott	Mauritania	0	418,294	938,154		
Rosso	Mauritania	0	30,530	50,861		
Oshakati	Namibia	1	34,552	83,432		
Windhoek	Namibia	1	140,410	358,996	Yes	
Arlit	Niger	0	36,261	78,651		
Birni-N'Konni	Niger	0	31,023	63,169		
Maradi	Niger	0	115,144	292,762		
Niamey	Niger	0	427,540	978,029		
Tahoua	Niger	0	52,951	117,826		
Zinder	Niger	0	126,517	235,605		
Aba	Nigeria	1	444,346	1,091,560	Yes	
Abakaliki	Nigeria	1	158,289	439,893		

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Abraka	Nigeria	1	119,940	259,762		
Abuja	Nigeria	1	384,364	3,028,557	Yes	
Ado-Ekiti	Nigeria	1	291,866	647,182		Yes
Afikpo	Nigeria	1	74,524	141,516		
Agbor	Nigeria	1	67,857	129,551		
Aiyetoro	Nigeria	1	43,862	49,195		
Ajaokuta	Nigeria	1	57,702	82,522		
Akure	Nigeria	1	356,210	675,366	Yes	Yes
Akwanga	Nigeria	1	41,705	91,050		
Ankpa	Nigeria	1	39,291	70,006		
Argungu	Nigeria	1	40,367	87,700		
Auchi	Nigeria	1	72,986	147,505		
Azare	Nigeria	1	65,234	124,820		
Bama	Nigeria	1	64,076	107,727		
Bauchi	Nigeria	1	232,939	435,001		
Bida	Nigeria	1	85,084	233,626		
Birnin-Kebbi	Nigeria	1	142,795	347,188		Yes
Biu	Nigeria	1	49,067	105,096		
Calabar	Nigeria	1	159,490	436,394	Yes	Yes
Damaturu	Nigeria	1	36,386	85,027		
Doma	Nigeria	1	42,091	83,383		
Dutse	Nigeria	1	152,198	193,025		
Egbe	Nigeria	1	34,188	89,210		
Egume	Nigeria	1	71,733	133,130		
Ejigbo	Nigeria	1	31,525	92,402		
Ekehen	Nigeria	1	30,566	57,101		
Emure-Ekiti	Nigeria	1	67,364	78,826		
Enugu	Nigeria	1	503,384	912,182	Yes	
Funtua	Nigeria	1	89,954	183,064	Yes	
Ganye	Nigeria	1	58,710	102,167		
Gashua	Nigeria	1	52,963	82,391		
Gboko	Nigeria	1	184,658	362,100		
Gombe	Nigeria	1	191,795	372,804		
Gusau	Nigeria	1	135,788	242,556	Yes	
Hadejia	Nigeria	1	45,276	94,181		
Ibadan	Nigeria	1	1,711,452	2,911,228	Yes	

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Idah	Nigeria	1	82,520	161,370		
Idanre	Nigeria	1	49,885	97,053		
Ife	Nigeria	1	263,879	491,656		
Igbo-Ora	Nigeria	1	31,519	76,914		
Igboho	Nigeria	1	31,854	62,311		
Ihiala	Nigeria	1	96,474		Yes	
Ikare	Nigeria	1	147,132	364,228	Yes	
Ikirun	Nigeria	1	215,476	427,992	Yes	
Ikole	Nigeria	1	56,932	100,183		
Ikom	Nigeria	1	40,718	52,109		
Ikot-Ekpene	Nigeria	1	146,477			
Ikot-Etim	Nigeria	1	87,282	165,044		
Ila	Nigeria	1	43,213	59,975		
Ilesha	Nigeria	1	139,202	332,008	Yes	Yes
Ilorin	Nigeria	1	538,446	833,589	Yes	
Ilutitun	Nigeria	1	45,214	70,917		
Iseyin	Nigeria	1	47,732	174,531	Yes	
Iwo	Nigeria	1	88,314	240,838	Yes	
Jalingo	Nigeria	1	83,219	176,451		
Jega	Nigeria	1	32,799	69,227		
Jibia	Nigeria	1	35,397	56,556		
Jimeta	Nigeria	1	238,746	567,818		
Jos	Nigeria	1	487,013	789,950		Yes
Kaduna	Nigeria	1	849,035	1,139,643	Yes	
Kafanchan	Nigeria	1	41,236	132,111		
Kano	Nigeria	1	1,385,370	3,734,597		
Katsina	Nigeria	1	189,505	425,669	Yes	
Katsina-Ala	Nigeria	1	43,751	74,895		
Kontagora	Nigeria	1	60,584	108,312		
Lafia	Nigeria	1	152,660	312,263		
Lagos	Nigeria	1	6,327,849	14,564,075	Yes	
Langtang	Nigeria	1	65,532	121,295		
Lokoja	Nigeria	1	63,547	375,656		Yes
Maiduguri	Nigeria	1	490,729	694,554	Yes	
Makurdi	Nigeria	1	179,494	301,249		
Malumfashi	Nigeria	1	46,775	58,968	Yes	

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Maya-Belwa	Nigeria	1	30,627	42,151		
Michika	Nigeria	1	48,163	74,898		
Minna	Nigeria	1	98,628	459,441		Yes
Mubi	Nigeria	1	80,666	127,945		
Nasarawa	Nigeria	1	30,873	57,046		
New-Bussa	Nigeria	1	40,675	83,317	Yes	
Nguru	Nigeria	1	44,872	103,062	Yes	
Nkume	Nigeria	1	129,318			
Nsukka	Nigeria	1	638,402	1,918,146		Yes
Numan	Nigeria	1	72,049	77,368		
Obudu	Nigeria	1	59,422	167,241		
Ogbomosho	Nigeria	1	134,065	383,364		
Oguma	Nigeria	1	35,039	72,981		
Ogwashi-Uku	Nigeria	1	42,955	67,482		
Okeho	Nigeria	1	41,304	105,183		
Okenne	Nigeria	1	85,307	376,128		
Okigwi	Nigeria	1	33,699	83,387		
Okitipupa	Nigeria	1	68,819	113,745		
Okpakeke	Nigeria	1	31,662	58,191		
Okpo	Nigeria	1	30,700	59,740		
Omu-Aran	Nigeria	1	47,679	81,069		
Omuo-Ekiti	Nigeria	1	31,118	99,172		
Ondo	Nigeria	1	228,481	426,176	Yes	
Onitsha	Nigeria	1	956,207	8,290,101	Yes	
Ore	Nigeria	1	45,689	102,651		
Oro-Esie-Iludin	Nigeria	1	46,096	75,454		
Osogbo	Nigeria	1	497,049	774,670	Yes	Yes
Otun-Ekiti	Nigeria	1	33,762	41,416		
Oturkpo	Nigeria	1	79,827	147,733		
Owo	Nigeria	1	103,021	186,305	Yes	
Oye-Ekiti	Nigeria	1	60,751	80,981		
Oyo	Nigeria	1	188,026	363,371		
Potiskum	Nigeria	1	46,192	241,243		
Saki	Nigeria	1	74,705	253,572		
Shendam	Nigeria	1	34,042	42,405		
Sokoto	Nigeria	1	310,603	606,753	Yes	Yes

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Takum	Nigeria	1	31,065	53,909		
Uba	Nigeria	1	55,350	70,447		
Ugep	Nigeria	1	34,279	149,847		
Umuahia	Nigeria	1	116,721			
Uromi	Nigeria	1	182,758	365,049		
Uyo	Nigeria	1	197,529	2,513,616		
Vande-Ikya	Nigeria	1	35,671	64,535		
Wukari	Nigeria	1	43,003	83,693		
Yelwa	Nigeria	1	35,055	72,400		
Zaki-Biam	Nigeria	1	54,169	83,361		
Zaria	Nigeria	1	375,845	747,127		
Zuru	Nigeria	1	49,083	110,647		
Brazzaville	Republic of Congo	0	731,625	1,652,847		
Dakar	Senegal	0	1,975,856	3,435,250		
Diourbel	Senegal	0	79,063	104,578		
Kaolack	Senegal	0	153,840	199,066		
Kolda	Senegal	0	36,624	71,134		
Richard-Toll	Senegal	0	36,610	67,954		
Saint-Louis	Senegal	0	118,992	188,160		
Tambacounda	Senegal	0	44,844	90,956		
Touba-Mosquée	Senegal	0	168,853	781,727		
Ziguinchor	Senegal	0	128,061	168,198		
Bo	Sierra Leone	1	76,138	220,890		
Kenema	Sierra Leone	1	66,406	187,158		
Makeni	Sierra Leone	1	48,170	108,671		
Torgbonbu	Sierra Leone	1	95,889	98,014		
Waterloo	Sierra Leone	1	561,004	1,049,768	Yes	Yes
Ad-Damazin	Sudan	1	58,786	255,340		
Ad-Duwaym	Sudan	1	53,580	79,009		
Al-Fashir	Sudan	1	130,226	244,208		
Al-Junaynah	Sudan	1	80,450	229,835		
Al-Manaqil	Sudan	1	60,108	111,669		
An-Nuhud	Sudan	1	52,539	69,668		
Atbara	Sudan	1	121,082	330,905	Yes	
Bur-Sudan	Sudan	1	293,338	421,429		Yes
El-Duein	Sudan	1	64,709	161,998		

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
El-Obeid	Sudan	1	211,433	384,829	Yes	
Gedaref	Sudan	1	178,488	295,201		
Kaduqli	Sudan	1	61,151	68,492		
Kassala	Sudan	1	223,586	318,335	Yes	
New-Halfa	Sudan	1	52,391	66,386		
Nyala	Sudan	1	194,574	606,115		
Sannar	Sudan	1	58,718	266,989		
Ngwenya	Swaziland	1	82,878		Yes	
Tabankulu	Swaziland	1	30,730			
Arusha	Tanzania	1	122,068	416,442		
Bukoba	Tanzania	1	31,826	128,796		
Dar-es-Salaam	Tanzania	1	1,333,413	4,520,658		Yes
Dodoma	Tanzania	1	90,565	213,636		
Kigoma	Tanzania	1	80,568	215,458		
Lindi	Tanzania	1	39,534	78,841		
Mbeya	Tanzania	1	144,556	385,279		Yes
Mtwara	Tanzania	1	68,149	100,626		
Musoma	Tanzania	1	68,356	134,327		
Mwanza	Tanzania	1	193,317	706,453	Yes	
Shinyanga	Tanzania	1	49,960	103,795		
Singida	Tanzania	1	41,807	85,242		
Songea	Tanzania	1	57,908	203,309		
Sumbawanga	Tanzania	1	51,038	124,204		
Tabora	Tanzania	1	96,935	160,608		
Tanga	Tanzania	1	142,799	221,127	Yes	
Zanzibar	Tanzania	1	174,467	501,459	Yes	
Fort-Portal	Uganda	1	32,130	51,795		
Gulu	Uganda	1	34,535	146,233		
Kampala	Uganda	1	803,069	2,269,969	Yes	
Masaka	Uganda	1	47,671	112,864	Yes	
Mbale	Uganda	1	51,446	117,706		
Mbarara	Uganda	1	39,119	164,150	Yes	
Njeru	Uganda	1	96,824	219,039		
Soroti	Uganda	1	40,903	48,069		
Chipata	Zambia	1	52,213	128,045		
Choma	Zambia	1	30,143	54,492		

Continue: Table A1 sample cities and population

City name	Country	Anglphone dummy	Projected population 1990	Projected population 2012	Strict colonial origin sample	40 cities sample
Kabwe	Zambia	1	154,318	207,909	Yes	
Kasama	Zambia	1	47,653	108,492		
Kitwe	Zambia	1	355,793	1,066,992	Yes	Yes
Livingstone	Zambia	1	76,875	143,249	Yes	
Luanshya	Zambia	1	118,143	133,187	Yes	
Lusaka	Zambia	1	813,154	2,000,916	Yes	
Mansa	Zambia	1	37,882	88,890		
Ndola	Zambia	1	329,228	468,324	Yes	
Bulawayo	Zimbabwe	1	611,307	653,337	Yes	
Chinhoyi	Zimbabwe	1	41,969	68,273	Yes	
Gweru	Zimbabwe	1	125,626	154,825	Yes	
Harare	Zimbabwe	1	1,405,753	2,133,801	Yes	
Hwange	Zimbabwe	1	44,297	19,870		
Kadoma	Zimbabwe	1	66,150	91,633	Yes	
Kwekwe	Zimbabwe	1	101,681	136,804	Yes	
Marondera	Zimbabwe	1	37,277	61,998		
Masvingo	Zimbabwe	1	48,780	87,886		
Mutare	Zimbabwe	1	124,697	186,208	Yes	
Zvishavane	Zimbabwe	1	32,571	45,230		

Notes: Two cities are only included in the 40 cities sample, but not included in the 333 cities full sample. They are Bimbo in Central African Republic, Libreville in Gabon.

Table A2: Coefficients of geographic controls for openness index and area

	(1) Openness	(2) Area
Anglophone dummy	0.231***	0.277***
	(0.047)	(0.089)
Lag t-1 ln country GDP per capita	0.037	0.013
	(0.053)	(0.084)
Ln annual population growth 90 to 12	-0.854	7.349***
	(1.128)	(2.064)
Ln projected city population 1990	-0.176***	0.867***
	(0.021)	(0.037)
Ln ruggedness	-0.043	-0.036
	(0.030)	(0.067)
Ln rainfall	-0.149***	-0.983***
	(0.047)	(0.123)
Ln elevation range	0.142***	0.444***
	(0.038)	(0.075)
Coast dummy	1.768	-0.595
	(1.223)	(3.371)
Interaction ln coast length	-0.077	-0.081
	(0.110)	(0.297)
Interaction ln distance to coast	0.086***	-0.062
	(0.033)	(0.077)
Fraction of river area	-0.058	-0.453
	(0.668)	(0.967)
Fraction of lake area	-0.009	0.118
	(0.963)	(1.247)
Fraction of forrest	0.390***	0.124
	(0.092)	(0.179)
Fraction of shrubs	0.337***	0.412***
	(0.081)	(0.141)
Fraction of crops	0.155**	0.230*
	(0.075)	(0.133)
Fraction of wetlands and water	-0.400	2.008**
	(0.399)	(0.938)
Fraction of sparse vege and bare land	0.287*	-0.337
	(0.155)	(0.343)
R^2	0.374	0.770
N	281	281

Note: Robust standard errors are applied.
* p <0.1, ** p <0.05, *** p <0.01. Standard errors in parentheses.

Table A3: Balance test for 40 cities sample

	(1) Project city population 1990 (Thousands)	(2) Ln annual city population growth 90-12	(3) Rainfall	(4) Coast Dummy	(5) Elevation
Anglophone dummy	-214.631	-0.018	-7.533	-0.050	-1.969
	(163.192)	(0.108)	(19.921)	(0.152)	(143.559)
Constant	692.799***	0.882***	120.676***	0.350***	378.372***
	(125.665)	(0.066)	(15.943)	(0.109)	(99.903)
Adjusted R^2	0.018	-0.026	-0.022	-0.023	-0.026
N	40	40	40	40	40

Note: Robust standard errors are applied.

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

Table A4: Built-up cover intensity gradient 2014

	(1) Base	(2) City controls	(3) Full controls
Anglophone Dummy	-0.457**	-0.894***	-0.892***
	(0.195)	(0.226)	(0.237)
Ring distance	-0.028***	-0.052***	-0.058***
	(0.004)	(0.008)	(0.008)
Ring distance × Anglophone	0.013**	0.041***	0.049***
	(0.006)	(0.011)	(0.012)
Ln ring total pixel	0.895***	0.708***	0.710***
	(0.087)	(0.049)	(0.050)
Lag t-1 ln country GDP per capita		-0.051	-0.121
		(0.175)	(0.169)
Ln annual population growth 90 to 12		5.194	5.626
		(4.239)	(4.271)
Ln projected city population 1990		0.488***	0.477***
		(0.088)	(0.090)
Ln ruggedness		0.138	0.120
		(0.137)	(0.149)
Ln rainfall		0.603***	0.304
		(0.225)	(0.202)
Ln elevation range		-0.151	-0.117
		(0.175)	(0.174)
Coast dummy		-0.993	-4.284
		(3.444)	(3.292)
Interaction ln coast length		-0.007	0.215
		(0.301)	(0.281)
Interaction ln distance to coast		-0.159	-0.217
		(0.136)	(0.141)
Fraction of river area		0.359	0.231
		(3.071)	(2.833)
Fraction of lake area		-1.228	-1.957
		(3.081)	(3.233)
Fraction of forrest			-0.065
			(0.413)
Fraction of shrubs			-0.137
			(0.362)
Fraction of crops			0.067
			(0.353)
Fraction of wetlands and water			2.844**
			(1.174)
Fraction of sparse vege and bare land			-1.351*
			(0.696)
R^2	0.193	0.319	0.327
N	4875	4197	4197

Note: Standard errors are clustered at city level.

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

Table A5: Coefficients of geographic controls for leapfrogging

	(1) Ln count of LF	(2) Ln LF minus ln total patches	(3) Ln avg. LF area
Anglophone dummy	0.607***	0.319***	-0.046
	(0.163)	(0.119)	(0.064)
Ln initial cover 1990	0.278***	-0.311***	0.014
	(0.063)	(0.040)	(0.021)
Year dummy 2014	0.491***	0.119**	0.140***
	(0.068)	(0.056)	(0.034)
Lag t-1 ln country GDP per capita	-0.229	-0.146	0.067
	(0.147)	(0.111)	(0.051)
Ln annual population growth 90 to 12	10.760***	5.222**	1.989
	(3.198)	(2.413)	(1.226)
Ln projected city population 1990	0.730***	0.467***	0.049
	(0.099)	(0.067)	(0.034)
Ln ruggedness	-0.007	-0.105	-0.015
	(0.097)	(0.069)	(0.030)
Ln rainfall	-0.396***	-0.180**	-0.069
	(0.141)	(0.082)	(0.051)
Ln elevation range	0.370***	0.152*	0.120***
	(0.111)	(0.083)	(0.037)
Coast dummy	-5.201	-5.667***	-2.102**
	(3.744)	(2.029)	(1.057)
Interaction ln coast length	0.423	0.452***	0.188**
	(0.313)	(0.168)	(0.089)
Interaction ln distance to coast	-0.018	-0.030	0.001
	(0.106)	(0.072)	(0.036)
Fraction of river area	0.528	0.031	-0.901
	(2.034)	(1.455)	(0.806)
Fraction of lake area	-3.297	-3.058	-1.274*
	(2.221)	(2.209)	(0.727)
Fraction of forrest	-0.200	-0.317*	-0.183*
	(0.282)	(0.191)	(0.101)
Fraction of shrubs	0.561**	0.082	-0.150*
	(0.281)	(0.175)	(0.087)
Fraction of crops	-0.055	-0.217	-0.107
	(0.229)	(0.166)	(0.085)
Fraction of wetlands and water	0.881	0.668	0.187
	(1.121)	(0.683)	(0.468)
Fraction of sparse vege and bare land	-1.664***	-1.266***	-0.498***
	(0.470)	(0.372)	(0.165)
R^2	0.586	0.231	0.117
N	551	551	525

Note: Standard errors are clustered at city level.

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

Table A6: Leapfrogging other specifications

	Count of LF	Ln count of LF		Ln LF minus ln total
	(1)	(2)	(3)	(4)
	Poisson	OLS	Tobit	OLS
Anglophone dummy	0.478***	0.497***	0.651***	0.257***
	(0.015)	(0.122)	(0.138)	(0.089)
Ln initial cover 1990	0.349***	0.287***	0.300***	-0.281***
	(0.007)	(0.043)	(0.048)	(0.031)
Year dummy 2014	0.525***	0.501***	0.511***	0.135**
	(0.010)	(0.084)	(0.097)	(0.062)
Lag t-1 ln country GDP per capita	-0.014	-0.184*	-0.241*	-0.072
	(0.011)	(0.110)	(0.125)	(0.081)
Ln annual population growth 90 to 12	6.649***	11.801***	10.761***	6.106***
	(0.256)	(2.583)	(2.951)	(1.889)
Ln projected city population 1990	0.509***	0.641***	0.726***	0.381***
	(0.009)	(0.077)	(0.087)	(0.056)
Ln ruggedness	0.036***	0.051	-0.009	-0.033
	(0.010)	(0.068)	(0.077)	(0.049)
Ln rainfall	-0.325***	-0.366***	-0.413***	-0.128
	(0.015)	(0.111)	(0.124)	(0.081)
Ln elevation range	0.303***	0.363***	0.385***	0.123**
	(0.010)	(0.085)	(0.097)	(0.062)
Coast dummy	-10.786***	-6.868**	-5.453*	-6.568***
	(0.483)	(2.792)	(3.270)	(2.042)
Interaction ln coast length	0.953***	0.469*	0.452	0.463**
	(0.041)	(0.253)	(0.297)	(0.185)
Interaction ln distance to coast	0.023**	-0.117	-0.015	-0.101*
	(0.010)	(0.082)	(0.093)	(0.060)
Fraction of river area	-2.776***	1.734	0.679	1.077
	(0.304)	(1.799)	(2.042)	(1.316)
Fraction of lake area	-1.896***	-3.617**	-3.407*	-3.359***
	(0.228)	(1.652)	(1.929)	(1.208)
Fraction of forrest	-0.380***	-0.194	-0.258	-0.346**
	(0.029)	(0.226)	(0.258)	(0.165)
Fraction of shrubs	0.598***	0.453**	0.578**	-0.043
	(0.027)	(0.220)	(0.251)	(0.161)
Fraction of crops	0.102***	-0.005	-0.057	-0.192
	(0.024)	(0.192)	(0.220)	(0.141)
Fraction of wetlands and water	1.843***	0.610	0.963	0.481
	(0.128)	(1.016)	(1.185)	(0.743)
Fraction of sparse vege and bare land	-0.859***	-1.207***	-1.953***	-0.779***
	(0.059)	(0.410)	(0.439)	(0.300)
Constant	-8.917***	-7.098***	-8.823***	-0.557
	(0.186)	(1.645)	(1.867)	(1.203)
R^2		0.612		0.220
N	551	525	551	525

Note: Columns 2 and 4 show OLS results excluding cities with zero LF patches. Standard errors are clustered at city level.

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

Table A7: Balance test for leapfrogging regressions

	(1) Full sample	(2) Border sample	(3) Border sample
Ln initial cover 1990	-.282	-.5	-.119
	(.187)	(.311)	(.369)
Ln projected city population 1990	.115	-.104	.038
	(.116)	(.253)	(.282)
Ln annual population growth 90 to 12	.003	.004	.001
	(.002)	(.004)	(.004)
Ln ruggedness	.581***	.418	.011
	(.156)	(.28)	(.137)
Ln rainfall	.037	-.011	.129***
	(.081)	(.137)	(.041)
Ln elevation range	.29***	.019	.022
	(.084)	(.19)	(.174)
Coast dummy	-.041	.042	.1
	(.028)	(.065)	(.089)
Fraction of river area	-.007**	-.006	-.004
	(.003)	(.007)	(.004)
Fraction of lake area	.001	.001	.001
	(.003)	(.002)	(.002)
Interaction ln coast length	-.428	.49	1.116
	(.292)	(.711)	(.971)
City cluster FE	No	No	Yes
Observations	318	58	58

Note: Border sample does not include cities in Anglophone and Francophone Cameroon border area. Robust standard errors are applied.

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

Table A8: Openness and area border sample

	(1) Ln openness	(2) Ln area
Basic controls		
Anglophone dummy	-0.109	0.220
	(0.109)	(0.207)
City cluster FE's		
Anglophone dummy	-0.029	0.012
	(0.089)	(0.143)
R^2	0.653	0.806
N	54	54

Note: Border sample does not include cities in Anglophone and Francophone Cameroon border area. Robust standard errors are applied.

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

Table A9: Openness and area: Robustness

	(1) Base	(2) Trim light	(3) No German colonies	(4) No Nigeria	(5) Colonial origin	(6) Non colonial origin	(7) 40 cities
Ln openness index 2014							
Anglophone dummy	0.231***	0.118***	0.211***	0.232***	0.021	0.198***	0.302**
	(0.047)	(0.034)	(0.070)	(0.060)	(0.182)	(0.055)	(0.134)
Ln area							
Anglophone dummy	0.277***	0.168**	0.260**	0.557***	1.191***	0.274***	0.352
	(0.089)	(0.070)	(0.118)	(0.107)	(0.241)	(0.099)	(0.301)
R^2	0.770	0.637	0.782	0.802	0.785	0.736	0.481
N	281	582	248	172	37	255	26

Note: Columns 1-4 and 6 include same controls as column 4 and 8 in Table 1. Columns 5 and 7 include common controls of ln country GDP per capita in 2000, ln projected city population in 1990, ln average ruggedness, and coast dummy. Columns 5 include ln annual urban population growth 90 to 14 in country level due to severe missing data problem in city population in the colonial origin sample. Column 7 include ln annual population growth 90 to 12 in city level. Robust standard errors are applied. R^2 and N are reported for ln area.

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

Table A10: Pre-colonial institution

	(1) Count LF	(2) Ratio LF	(3) Avg LF Area	(4) Openness	(5) Area
Anglophone dummy	0.945***	0.275	-0.021	0.270***	0.406**
	(0.285)	(0.209)	(0.132)	(0.096)	(0.171)
Pre-colonial institutions index	0.421***	0.086	0.041	0.022	0.176**
	(0.133)	(0.084)	(0.059)	(0.038)	(0.080)
Anglophone×Pre-colonial	-0.272*	0.022	-0.034	-0.051	-0.113
	(0.157)	(0.100)	(0.066)	(0.048)	(0.093)
Adjusted R^2	0.634	0.256	0.111	0.349	0.784
N	399	399	376	203	203

Note: Columns 1-3 have same controls as in Table 3. Columns 4 and 5 have same controls as in Table 1. Standard errors are clustered at city level in columns 1-3. Robust standard errors are applied in columns 4-5.

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

Table A11: Public utility connection

	(1) Has electricity	(2) Has piped water	(3) Has flush toilet	(4) Has phone land line
Share of built cover	0.039	0.052	0.005	-0.011
in buffer	(0.043)	(0.047)	(0.028)	(0.017)
Count of LF	0.000	-0.003	-0.001	-0.001*
	(0.004)	(0.005)	(0.001)	(0.001)
Anglophone ×	0.028	0.041	0.197***	0.015
Share of built cover in buffer	(0.055)	(0.070)	(0.045)	(0.023)
Anglophone ×	-0.003	0.004	-0.002	-0.001
Count of LF	(0.004)	(0.005)	(0.002)	(0.001)
Ln buffer center	-0.066***	-0.060***	-0.021***	-0.012***
distance	(0.008)	(0.010)	(0.008)	(0.004)
Ln buffer ruggedness	0.051***	0.052**	-0.008	-0.001
	(0.017)	(0.021)	(0.013)	(0.006)
Buffer has river of	0.023	0.040	0.001	0.004
lake	(0.021)	(0.027)	(0.017)	(0.012)
Household size	0.009***	0.002***	0.003***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Sex of household	-0.008	-0.009**	-0.016***	-0.010***
head: Male	(0.005)	(0.005)	(0.004)	(0.003)
Highest educational	0.037***	0.013*	-0.004	0.013***
level of head: Primary	(0.007)	(0.007)	(0.004)	(0.003)
Highest educational	0.174***	0.052***	0.040***	0.042***
level of head: Secondary	(0.007)	(0.007)	(0.005)	(0.004)
Highest educational	0.287***	0.057***	0.164***	0.112***
level of head: Higher	(0.009)	(0.009)	(0.009)	(0.007)
Highest educational	0.141***	0.046**	0.012	0.003
level of head: Unknown	(0.025)	(0.023)	(0.017)	(0.015)
Periods dummies	Yes	Yes	Yes	Yes
City fixed effect	Yes	Yes	Yes	Yes
Adjusted R^2	0.321	0.437	0.262	0.126
N	44517	44561	44500	42748

Note: Period dummies control the difference between the DHS survey years with year 2000 and 2014 when satellite data is available. Period dummies include 1 year before dummy, 1 year after dummy, 2 years before dummy, and 2 years after dummy. Standard errors are clustered at DHS cluster level.

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