Colonial legacies: Shaping African cities*

Neeraj G. Baruah
Vivid Economics

J. Vernon Henderson[†]
London School of Economics

Cong Peng
Harvard University

September 17, 2020

Abstract

Institutions persisting from colonial rule affect the spatial structure and conditions under which 100's of millions of people live in Sub-Saharan African cities. In a sample of 318 cities, Francophone cities have more compact development than Anglophone, overall, in older colonial sections, and at clear extensive margins long after the colonial era. Compactness covers intensity of land use, gridiron road structures, and leapfrogging of new developments. Why the difference? Under British indirect and dual mandate rule, colonial and native sections developed without coordination. In contrast, integrated city planning and land allocation were featured in French direct rule. These differences in planning traditions persist[1].

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, NBER Summer Institute, OECD, University of California Berkeley, Trinity College Dublin, LSE, University of Oxford, and University of Bilbao. We thank the referees, Rui Du and Frederick Robert-Nicoud for helpful suggestions that improved the final version. We are grateful to Nicolas Cerkez and Eleonore Lorijn for excellent research assistance.

[†] **Corresponding author:** J. Vernon Henderson; J.V.Henderson@lse.ac.uk; Houghton Street, LSE, London, WC2A2AE, UK. Other authors: Neeraj Baruah, neeraj.baruah@vivideconomics.com; Cong Peng, cpeng@hks.harvard.edu

[1] Replication data and code are provided at `https://doi.org/10.5281/zenodo.4033412`

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 which they experienced. Our main findings are for a set 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 four years, 1975, 1990, 2000, and 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 29% more area and have 17% more "openness" - the measure of sprawl from Burchfield et al. (2006). Second, we look at regularity and density of layout at the internal margin: older sections of cities which will be strongly physically influenced by how cities were laid out in colonial times. Francophone cities are laid out in a more Manhattan-like gridiron fashion with much higher density development nearer the city centre.

Third, in the main body of evidence, we look at the disconnectedness of new developments at the extensive margin, in areas of cities built well after the colonial era, to see if there is persistence in compactness. We argue that such persistence must be due to persistence of planning and land use management practices and norms decades after the end of the colonial era. We examine the degree to which new developments in two episodes from 1990-2000 and from 2000-2014 in cities are "leapfrog", or are scattered and spatially disconnected from existing developments, employing a novel measure of leapfrogging. Anglophone cities on average have 54% more leapfrog patches compared to otherwise similar Francophone cities. In summary, we find Anglophone cities compared to Francophone ones sprawl more overall, have less gridiron road structures at the intensive margin, and have more leapfrog developments at the extensive margin. Colonial origin continues to leave a huge footprint on urban form decades after the end of colonial rule.

Do these correlations reflect (reduced form) causal effects of colonial origins? Africa gives an experiment in which happenstance determines 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, 2020), as well as different pre-colonial histories. While we have a large set of geographic and other controls in estimation of effects, there could still be unobserved geographic or historical features of cities, which differ systematically between Anglophone and Francophone cities. To meet this challenge to identification, we conduct a border experiment in West Africa 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.

Having determined that there appear to be strong differences in land use morphology in cities by colonial origin, that raises two major issues we must address at different points in the paper. First, what are the mechanisms promoting these differences? Second, why do spatial patterns of land use matter, in terms of public policy and other outcomes. We start with a discussion of each issue and later in the paper provide some evidence.

Why are Francophone cities more compact? Different institutions imposed under the two colonial rules affected urban spatial structures, including road layouts, sprawl and leapfrogging. The 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, which affected spatial development. The“indirect rule” strategy of the British is contrasted with French “direct (and assimilative) rule” (Silva, 2015; Njoh, 2006; Crowder, 1964).[2]. The British operated under a dual mandate 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). As a result, generally there was no overall integrative land use plan for a city. Any formal planning focused only on the British sections of the city.

In establishing dominance over their colonies and promoting cultural assimilation, the literature argues that French rule in a city 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. Durand-Lasserve (2005) writes “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....”. Njoh (2006) provides details of the French style of rule and city planning for Benin, Cameroon, Cote d’ Ivoire, Guinea, Mali, Madagascar, Mauritania, Niger and Senegal, and Togo, with other details in Oto Peralias and Romero-Ávila (2017). As part of maintaining control over the landscape, the French wanted the different neighbourhoods spatially integrated and linked in a lineal pattern so that from one intersection an official could see 2 km in four directions (Njoh, 2016, chap. 1). Different chapters in Silva (2015) also detail how the French adopted centralised and standardised grid systems. We

---

[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)

interpret these writings as indicating that, at the local level, the French imposed more centralised city planning and land use management than the British. [3]

How might these differences affect spatial structures initially under colonial rule and more particularly at the extensive margin today? Persistence comes from two sources: the long life of many forms of physical infrastructure and persistence in investment "norms" or traditions (Huillery, 2009). On physical infrastructure, roads which drive city-layouts are long-lived. Further, with reconstruction of roads, rights-of-way for roads 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 well beyond the older sections of cities, since postcolonial lay-outs may initially follow existing types of patterns as accretions of older developments. We expect Francophone cities in the older colonial sections and their more immediate extensions are more likely to have lineal roads and rectangular block structures throughout, maintained from the colonial era. A different literature further argues that imposing a gridiron pattern leads to greater contiguity (Libecap and Lueck, 2011; Ellickson, 2012; O'Grady, 2014) and less sprawl. While some sections of Anglophone cities will have grid structures within neighbourhoods (Ross and Bigon, 2018), what distinguishes Francophone cities is the imposition of such a structure on the whole colonial city so that grid-like neighborhoods are linked together in a common gridiron, not as more disconnected gridded pieces. British dual mandate already allowed for disconnected parts of the city; and, then in the disconnected native parts, the British had a hands-off approach. French centralized city planning allowed for neither.

However, our context gives us what we think is a clear extensive margin where the primary influence of colonial rule is highly unlikely to be persistence and accretions of the physical, but rather to be persistence in planning norms inherited from the colonial era. Cities expand outward from a traditional city centre; and, in most African cities, the colonial era part of the city and its more immediate extensions are a small part of the modern city. For 111 cities of the 318 of our cities which are included in the World Cities Data set[4], population grew by 550% from 1960 to 2000, with approximately the same growth rate of Anglophone and Francophone cities. On the spatial extent of cities defined below, for 188 cities for which we could conservatively define built environment boundaries in 1975, the average aerial size of cities increased by about 119% from 1975 to 1990, or cities more than doubled in size by 1990 from a time already typically 10-15 years after independence. Moreover, as cities expand outward from the centre, if the colonial pattern of radial roads were just extended, they must separate more and more, requiring new infill between these roads of many new neighborhoods and their local road networks, governed by contemporaneous planning norms. For the 1990-2000 and 2000-2014 episodes of development, about 83% of

---

[3] Even in contexts where there are strong pre-colonial institutions especially in countries with a heavy Islam influence such as Senegal, researchers such as Ross and Bigon (2018) note the strong French imprint throughout all urban centres in a country.

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

leapfrog developments lie beyond this 1990 boundary. Thus, it seems that differential Anglophone effects must arise from persistence in colonial institutions and city planning and land management practices, not just accretions of the physical. We will show that this persistence in Anglophone effects at the extension margin is not diminishing within the time frame of our data outcomes, 1990-2014.

Where does this persistence really come from? It is hard to distinguish between the influence of planning institutions per se and persistence in social norms developed over generations of colonial rule about how people want to live and be integrated in a city (Giuliano and Nunn, 2017). We are not sure the distinction is necessary or helpful: social norms can drive planning practises and institutional persistence in planning practises might be hard to maintain politically if they contradict social norms. However cities lay out roads which define neighbourhoods, so in a narrow sense persistence starts from planning procedures.

The literature details persistence in planning procedures from the colonial era, using many examples. It is hard to gather comprehensive evidence, because most relevant data on African cities are only available through intensive fieldwork and use of local archives and contacts. That is a reason the literature on the colonial legacy of urban planning focuses on case studies of individual cities as in the chapters in Silva (2015). For our cities, there is no general data source on current or historical cadastre records, shapefiles for municipal boundaries or when city plans were first formulated or amended.[5] Thus we rely on others who have done the local fieldwork.

In examining French West African countries, Njoh (2004) and Silva (2015, Chapter 2) argue that these countries maintained colonial era planning practices well into the 21st century. With planning being much looser in Anglophone countries, Home (2015) and Scholz et al. (2015) argue that Anglophone countries tended to also maintain their colonial era practices into the 21st century. Njoh (2016, Chapter 9) gives examples of how norms were maintained, with the Ministere des Relations Exterieures and the Ministere de l'Urbanisme in France both working to continue influence in Francophone Africa. One avenue was to offer technical assistance and the other to be directly involved in urban planning in former colonies. Specific examples from records include assignment of about 150 French urban planners to Sub-Saharan African locales in the 1980's and France's direct involvement in urban planning in a variety of African contexts in the mid-2000's.

Finally and most critically, we turn to the question of the policy implications of building more compact cities. The planning literature argues 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, with the planning literature offering different assessments of the savings

---

[5] For example on the last item, we conducted a lengthy search of the internet focusing on JSTOR and Google scholar (as well as Wikipedia) by city and country using keywords such as plans, maps, planning and the like. Those searches typically lead to books such as *Africa South of the Sahara 2004* from Europa (2003) and other references in this section. We only found 13 cities which specifically listed a first city plan and date from 1885 to 1960.

from compactness (Trubka et al., 2010; Calderón et al., 2014). Hortas-Rico and Solé-Ollé (2010) provide econometric evidence on public budget cost savings from increased density for Spain, and Carruthers and Ulfarsson (2003) do the same for the USA. Based on a sample of 45,000 households in 193 cities, we make a modest contribution by using Demographic and Health Survey (DHS) data to show that families have worse connections to electricity, phone landlines, piped water, and city sewer systems, if they live in neighborhoods of a city which are more sprawling, presumably because of increased cost of infrastructure provision. Also, *ceteris paribus*, children have more diarrhea, as a health outcome.

The economics literature suggests compactness has many other effects on urban life. Ahlfeldt and Pietrostefani (2019)'s careful meta-study looks at the impacts of higher density, which goes along with greater compactness. They conclude that the literature shows that increased density is significantly associated with reduced energy consumption and vehicle mileage, lowered public sector costs, and increased wages and productivity. On the other hand, they find evidence on crime rates and pollution to be more mixed. This builds on an older literature on how compactness improves communication externalities (Rossi-Hansberg, 2004; Helsley and Strange, 2007), reduces pollution (Glaeser and Kahn, 2010) and commuting times (Harari, 2020) and positively affects how we interact socially (Putnam, 2000; Burley, 2016) [6]. We do not have the data nor the space to look at such micro outcomes. Nevertheless the suggestion is that, for hundreds of millions of Africans, the colonial heritage under which they urbanize will influence the densities and conditions under which they live their lives. It is of course tempting to try to relate compactness to more aggregate outcomes such as GDP per capita or inequality. However as we will see in the next section, the literature on colonialism offers many important ways that these outcomes are influenced by colonial rule, other than just compactness of cities.

This paper presents results in three sections, after a review of other literature and introduction of the context and data. First we look in 2014 at the degrees of sprawl as measured by openness and by overall areas of cities. Second, we turn to the more likely older and colonial physically influenced parts of cities. We discuss evidence on whether larger Francophone cities have more of a "Manhattan-like" gridiron structure compared to similar Anglophone cities; and, for the full sample, we compare the intensity of built cover in older parts of cities. Third, in our main results, 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 patches of new 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. Then we conduct the border experiment and do a number of robustness tests. Finally we examine the impact of

---

[6] 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.

leapfrogging on provision of public infrastructure.

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. Our review is brief.

2.1 Colonialism and Institutions

The literature on institutions and persistence argues that historical institutional accidents can have a strong impact on modern day outcomes (Banerjee and Iyer, 2005; Guiso et al., 2016).[7] 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 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".[8] 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 layout.

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 no control or decentralised governance. Rossi-Hansberg (2004) and Helsley and Strange (2007) show that, in the 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 encourage sprawl, through, for example, ribbon developments sited along government built arterial roads. Overall, theory suggests that uncoordinated and decentralised land development will result in cities that are less compact, offering an empirical prediction in comparing Anglophone and Francophone cities.

---

[7] 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).

[8] 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.

That said, these papers do not directly model the political forces driving the decisions of a central authority, treating that authority as a benevolent dictator. [9]

Our main empirical 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 a lower likelihood of land developers engaging in leapfrog development at the extensive margin. While this could be a mechanism for persistence in newer areas of development, given how far flung and how complex modern cities are now, we think persistence is going to mostly come from continuation of colonial planning processes and norms.

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 under differing 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. Such benefits are explored in a city context by O'Grady (2014), who compares a centralised and standardised rectangular grid demarcation with more ad hoc demarcation systems. 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.

Other papers examine persistence in spatial outcomes driven by historical infrastructure investments or planning (Shertzer et al., 2016; Michaels and Rauch, 2017; Bleakley and Lin, 2012; Brooks and Lutz, 2014). A key paper on sprawl by Burchfield et al. (2006) analyses geographic and historical influences on the degree of land use sprawl in US cities.

3 Context, Data, Specification and Identification

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

---

[9] Davis and Whinston (1964); Hochman et al. (1995) argue the benefits of centralized city governance but do not deal with compactness per se. 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).

3.1 Colonial Countries

Our classification of African countries by colonial origin is shown in Figure 1a along with the cities in our sample. We exclude South Africa since Apartheid itself and the aftermath had such a strong, confounding influence on spatial patterns. The colonial origin 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-à-vis modern countries (e.g., Togo). If one thought governance procedures and urban plans were developed near the end of the 19th century and early 20th before the end of World War I, then we would 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, many cities were in infancy potentially limiting the pre-World War I influence.

3.2 Data on Land Use and Cities

We focus on the three available epochs of land cover data from the Global Human Settlement Layer (GHSL), which start at least 20 years after the end of the colonial era and cover 1990, 2000, and 2014. We classify pixels of 38m spatial resolution into different uses where we distinguish built cover (impervious surface) from non-built cover (water, various vegetation and crop, barren water and so on). The GHSL is a project which is part of the Global Human Settlement Project by the European Commission and Joint Research Centre (Pesaresi et al., 2016)[10].

It is the most spatially global detailed dataset on built cover available. 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) as well as more fine-scaled and volunteered geographic information (Open Street Maps and Geonames) are incorporated (Pesaresi et al., 2013). [11] Available since 1972 (Ban et al., 2015), the GHSL estimates the presence of built-up areas in four different epochs (1975, 1990, 2000 and 2014)[12]. 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 Fran-

---

[10] The 2014 version of the GHSL data which we use can be downloaded from https://ghsl.jrc.ec.europa.eu/download.php

[11] 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.

[12] 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 built-up 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.

cophone cities and 227 former Anglophone cities, with the latter including 122 Nigerian cities. These cities are reported in the first table in Appendix A2. They comprise all cities in the relevant colonial origin countries which are over 30,000 in estimated population in 1990,[13] and which have built cover data for 1990 and beyond. 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 size and maturity. We then apply criteria on the extent of persistent cloud cover to get cloud free city-year observations for 2000 and 2014.[14] Removing cities with cloud cover and only partial coverage of the built surface, 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, in robustness checks, we explore various sub-samples. One is West Africa sample pictured in Figures 1b and 5 for which we do the border experiment to try to answer identification concerns. Another sub-sample excludes Nigeria which is a third of the sample, to make sure it is not driving the results. Another excludes British colonies which were German colonials until just after World War I. There is a sample for Open Street Map analysis of 20 each of Francophone and Anglophone cities over 300,000 in 2012, described later. These cities are listed in the Appendix A2 and mapped in Figure 1b. Finally there is a sub-sample of newer cities founded after either 1800 or 1850, whose origins are more likely to be colonial. These cities are also noted in the first table in Appendix A2.

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. We rely on night light readings (Elvidge et al., 1997) for Africa 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 (Donaldson and Storeygard, 2016; Henderson et al., 2017). African cities generally have low light levels, so we do not threshold the lights to be above some cut-off. For smaller cities, thresholding tends to exclude obvious built areas (looking at Google Earth) and even some entire cities. In very big cities, blooming can be 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) and some later in the paper,

---

[13] We use *Citypopulation.de* to get city population numbers (based on Censuses), supplemented with data from *Africapolis* for Nigeria. Census numbers are for different country-specific census years around 1990, with growth to 2012 based on city population growth rates between two relevant population censuses.

[14] 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 would 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.

we experiment 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. We 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, and 2014. The 1990 measure gives an urban core, beyond which we define as the extensive margin. We will use this 1990 boundary measure in robustness checks and discussion of persistence.

3.4 Specification

Throughout the paper, regressions have the following general form:

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

where i is city, j is country and t is time. The initial regressions are cross section. The later leapfrog ones are flow measures for 1990 to 2000 and 2000 to 2014; there we add a time dummy to the error structure, d_t. 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 where relevant in the leapfrog regressions. The coefficient of interest is δ - the Anglophone differential. For the border experiment in Section 6.2 we will also adjust the error structure to have fixed effects for 14 cross-border clusters of cities in close spatial proximity. In addition, for all relevant tables we do an Oaxaca (1973) decomposition to obtain the differential for Anglophone cities based on the differential in outcomes not explained by differences in characteristics. That is, we estimate

$$\begin{aligned} Y^A_{ijt} &= X^A_{ijt}\beta^A + Z^A_{ijt}\theta^A + \varepsilon^A_{ijt} \\ Y^F_{ijt} &= X^F_{ijt}\beta^F + Z^F_{ijt}\theta^F + \varepsilon^F_{ijt} \end{aligned} \tag{2}$$

and calculate the unexplained part $\bar{X}^F(\beta^A - \beta^F) + \bar{Z}^F(\theta^A - \theta^F)$, as well as the explained part based on differences in mean values of covariates. We are interested in the unexplained part: the colonial impact from differential influences of covariates (including constant terms). Rather than forcing the same coefficients on all covariates except the constant term as in the base, the Oaxaca decomposition allows coefficient differences by colonial origin.

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 (2020) all show that geography influences urban form[15]. Our X_{ij} controls on geography

---

[15] 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.

are the first way we try to deal with this concern; the second will be the border experiment. We use measures found in the literature, starting with Burchfield et al. (2006). First there is terrain where hilly areas create sprawl from inaccessible topography. We have measures of ruggedness as defined by Nunn and Puga (2012) done here at a 30m x 30m resolution and of the range of elevation within the city, also potentially influencing sprawl. Water is another constraining feature. We control for whether a city is coastal, distance to the coast from the city centre, and for coastal cities the length of coastline (in km) within its boundary. More extensive coast means more inlets and bays again influencing city shape (Harari, 2020) and thus the possibility of gridiron structures. For the same reason, within the city, we control for the fraction of pixels that are inland water (lakes, rivers, wetlands). We control for average rainfall and average temperature for 1950-2000 and for an index of soil suitability for cultivation from Ramankutty et al. (2002), all influencing the opportunity cost of land in farming at the city edge. We have base 1990 built cover for the spatial extent of the city which will appear in leapfrog specifications.

We control for local ethnic-linguistic fractionalization (ELF), where increased fractionalization and the potential for conflict may affect sprawl, for example inducing people to spatially separate more within cities for example. The 19th edition of the World Language Mapping System gives polygons showing the extent of any language group circa the early 1990s. Polygons of different languages may overlap and languages are defined in trees by different levels from 1 up to 15. So a level 1 is Indo-European; level 3 is Indo-Aryan; level 7 includes Punjabi and Urdu; and below level 7 are local dialects such as "Eastern Punjabi". For details on the use of these data see Desmet et al. (2018). Within each city we define ELF as

$$ELF_{mi}^{l} = 1 - \sum_{m} (\pi_{mi}^{l})^2 \tag{3}$$

where π_{mi}^{l} is the share of group m in city i at language level l. Ethnolinguistic fractionalization starts at 0 for unilingual cities and rises towards 1. Following Desmet et al. (2018) we generally use level 15 for languages although results do not change with a coarser resolution. We populate the city and construct these shares using the 1975 GHSL data, with resolution at the 1km grid square.

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, despite the potential endogeneity issues. Faster growing cities will be more likely to sprawl per se in the short run and Appendix Table A7 hints at modestly faster average growth among Anglophone cities. We thus decided this was a critical control, and viewed any feedback of sprawl on growth to be second order. Fortunately, results are essentially the same with and without the economic controls. The growth control is specifically the annualized city population growth rate from 1990 to 2012. This loses us about 9% of the sample due to lack of circa 2012 population numbers. We also control for whether the city is a national capital in 1990 or not; and, if not, its distance from the national capital as in Campante and Do (2014). As part of economic opportunities, we control for the malaria index from Anthony et al. (2004).

4 Overall Patterns in the Data for Cities as a Whole

Using the GHSL data, to see motivating patterns in the data, we examine the *openness index* from Burchfield et al. (2006) for the overall city, which is the current accepted measure of city sprawl. 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, which is a measure of pixel openness. 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*, Anglophone cities occupy larger areas.

We first look at the raw data, by comparing distributions of pixel level openness for Francophone versus Anglophone cities,[16] based on graphs in Burchfield et al. (2006). Figures 2a and 2b show the probability distribution [pdf] for all built-up pixels in 1990 and 2014 by the percent of land not built in the surrounding one square kilometer (i.e., pixel openness). In both years, the dotted line for Francophone relative to the solid line for Anglophone shows the Francophone pdf shifted left. The Francophone sample tends 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 aggregate Burchfield city level openness, or sprawl index in 2014 in regressions controlling for geography and other city characteristics. In Table 1 we adopt a presentation format which we will use in Table 3 as well. Column 1 has no controls. Column 2 adds in all geographic and situational controls including distance to the national capital, a malaria index, and a local measure of linguistic fractionalization. Column 3 adds the country lagged GDP per capita control and the listed controls on 1990 population size and growth. Results on these controls are in Table A2. The Anglophone effect in column 1 is an increase in openness of 23%. It drops modestly and insignificantly when geographic and situational controls are added and is little affected

---

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

by economic controls. The column 3 result is that Anglophone countries have 17% more openness. Some control variables in Appendix Table A2 have the expected impacts from Burchfield et al. (2006): bigger cities have less openness and cities with greater elevation differentials have more. We note that that French centralised land use control may have responded to differential geography of cities differently than the more decentralised British approach. However, the Oaxaca decomposition yields unexplained effects which are similar in magnitude to the base regression coefficients in columns 2 and 3.

Columns 4-6 of Table 1 turn to the area of cities as measured by the lights boundary, following the same format as columns 1-3. In column 6 with full controls, Anglophone cities occupy 29% more land than their otherwise similar Francophone counterparts. Coefficients for control variables are in Table A2. The Oaxaca unexplained effect of Anglophone at 27% is close to the base estimate of 29%. One should worry about lights being a noisy measure of larger and smaller city areas. We ran a robustness check in Table A8 in the Appendix trimming the sample by dropping the bottom and top 2.5% each of the sample by distance from the center to the nearest boundary point. Trimming has little effect on magnitudes, whether we measure sprawl by openness or total area. In summary in Table 1, 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 areas beyond the 1990 built part of the city and beyond the physical influence of colonial era city layouts.

5 The Colonial Portions of Cities

In this section we will look at differential densities of land cover between Francophone and Anglophone cities, ring by ring of 1 km widths as we move away from the city centre. We generally think of rings within 5 km of the city centre as potentially being under direct colonial influence. However before doing this we give some motivating examples, both visual and quantitative to illustrate the influence of colonial era Francophone planning.

5.1 Road Layouts: Anglophone Versus Francophone Cities

We start with a simple example, which compares Bamako to Accra, and then Brazzaville to Harare. These cities were chosen a priori with no other comparisons because they were city-pairs for which we could obtain detailed road maps from about the same year near the end of the colonial era and they were paired cities with similar initial and final sizes. These cities first emerged in the late 19th century, Bamako and Brazzaville under French rule and Accra and Harare under British. Starting with Bamako and Accra, their populations were similar in the early 20th century: Bamako at 16,000 in 1920 and Accra at 18,574 in 1911.[17] Accra retains that modest population difference

---

[17] 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

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 coastline).

Bamako had its first (apparently implemented) road plan in 1894 (Njoh, 2006, p. 92), replacing spontaneously prior developed roads with a street network on a classic gridiron with streets intersecting perpendicularly (Njoh, 2001). Bamako's urban land was under state control by 1907 with the *"Plan d'une cite administrative - un quartier de Bamako"*, with the state supreme in land allocations and assignment of set plots (Bertrand, 2004). In contrast, 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" 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 in dark lines from tracings of road maps and we show the road layout today in light grey from OpenStreetMaps. For Accra we show the roads for 1966,[18] as well as today. Visual inspection suggests several takeaways. First there is physical persistence in both cities - roads that were in place 55 years ago generally remain in place today. Second Bamako presents as having large colonial sections of intense dense, gridiron road structures where neighborhoods are interconnected by mostly long lineal roads. New sections of the city generally are also on a rectangular grid structure. In contrast, Accra shows much less gridiron structure with fewer lineal connecting roads between developments even in the colonial parts of the city. Moreover, 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 [19]. 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 gridiron roads systems for Accra versus Bamako apply equally well in this comparison. One stark feature given the difference in scale in the two maps in Figure 3b is the tight packing of rectangular blocks both in the colonial era and today in Brazzaville, which applies everywhere outside of public open spaces.

In the printed Appendix, we examine whether these patterns hold more generally in large cities. We took all 20 Francophone cities in Sub-Saharan Africa over 300,000 in 2012, to analyse road layouts from OpenStreetMaps. Since OSM data are 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 matching Anglophone cities over 300,000 out of the 68 in that

---

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

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

size range, using a one to one Mahalanobis distance based matching approach without replacement. Details are in the Appendix with a balance test in Appendix Table 1. For this matched sample, we see if central portions of Francophone cities have different block structures than Anglophone ones. In the Appendix we detail how we define rectangularity of blocks and then *gridiron blocks*. To be a gridiron block, a block must be rectangular and be connected to all neighbouring blocks by 4-way intersections. Comparing our 20 Francophone to matched Anglophone cities, Anglophone cities average 20 percentage points fewer gridiron blocks, from a mean of 17, implying that Francophone cities average 27% of blocks being strictly gridiron but Anglophone only 7%. The results are almost the same for radii of 3 and 5 km, so there is no overall diminishing of regularity with the 150 percent increase in area covered. More details are in the Appendix.

Overall these illustrations suggest a strong colonial influence of centralised control and planning, as suggested by Njoh (2016) and others, which persists until today. Now we ask if we see these influences in the overall sample, as reflected in intensity of land use nearer city centres.

5.2 Intensity of Land Use in the Colonial Portions of Cities and Immediate Extensions

Corresponding to greater gridiron structures 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 2014, as we move out from the city centre in 1 km increments to 6 km. 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., rings intersecting a coastline or river).

For rings 1-2, 2-3, and 3-4 km Anglophone cities have 46-64% fewer built pixels, with similar Oaxaca unexplained portion effects, implying that Anglophone rings near the city center are much less intensely developed. After 4km, the sample starts to drop as we lose smaller cities with no area beyond the given radius. At 4-5 km and beyond out to 11-12 km, there are no significant differences for Anglophone countries although coefficients are generally negative. A full ring set of results and many other intensity specifications are in Baruah et al. (2017). We report two extensions here.

The first extension is in column 7 of the table, where the ring measures are pooled out to 20km; and we estimate the height and slope to the intensity gradient as we move away from the city centre. Anglophone cities have 73% less intensity at the centre; Francophone cities have a steep slope to the intensity gradient of -0.287; and Anglophone cities have a significantly flatter slope of -0.194. In Table A3, when we estimate the gradients separately for the two samples, we get essentially the same respective slopes (-0.285 and -0.195). Given the height and slope differences, the two gradients cross at about 8 km from the center. Most critically, this raises an issue of how to examine the comparative extensive margin of cities, which we will discuss below.

The second related aspect concerns the one cross-section for 2014 where, besides the count of built pixels, there is a measure of the intensity of building in those built pixels: an estimate of the fraction of the grid square that is covered with built surface from the GHSL. Table A4 shows this for the rings 0-1 to 5-6 km. Anglophone built grid squares are less intensely developed than Francophone ones and significantly so from 1 to 4 km. In summary, Francophone cities have much more intense land use in the older colonial physically influenced portions of cities than Anglophone cities.

We now want to look beyond these sections of the city, to examine persistence through more than persistence of physical colonial layout and its accretions. In doing so, openness and intensity raise problematic issues for defining comparative extensive margins for cities of different colonial origins. For the same city populations, Francophone cities are smaller and extend less from the city centre. At, say, 7-9 km, many middle size Francophone cities will be ending, or at their edges, noting from above that estimates of simple intensity gradients suggest a cross point of the steeper Francophone gradient with the Anglophone gradient at just under 8 km, indicating Francophone cities will typically be ending near that point. In contrast, for corresponding population size Anglophone cities will have city edges much further from the centre. Thus, any lower intensity and greater openness further from the centre for Francophone cities might just indicate that we are near the city edge, not that there is sprawl. [20]

Given this issue in comparing extensive margins across cities with different colonial origins, our main results are based on a measure of disconnectedness, or leapfrogging, which we can accurately measure and which operates almost exclusively beyond the 1990 built cover area of each city. Persistence of Francophone planning traditions and social norms into the modern era should mute the degree of leapfrogging, compared to Anglophone cities.

6 Compactness in the Post-Colonial Margins of Cities

Burchfield et al. (2006) introduced leapfrogging in a rigorous fashion to the economics literature, using their measure of openness to characterise new developments in the USA from 1976 to 1992. Here we introduce a novel and conceptually well based measure of leapfrogging. It is a flow measure of actual new developments, that are spatially separate from existing developments, for developments from 1990 to 2000 and then again from 2000 to 2014. Using the 1990 to 2000 period as an example, in 1990 we have a set of built pixels, which are typically in clusters. We

---

[20] There is also the issue that 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 for the same population.

define the boundary or outer envelope of each cluster of contiguously developed pixels, which we call patches. In the illustrative Figure 4, the 1990 developed areas are the black area. The focus is on newly developed pixels. These also appear as patches of contiguous newly built pixels, which 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 (diagonally lined area in the figure). If it only marginally intersects the existing cover (or is within 300m of it), it is called extension (vertically lined area in the figure). It does not intersect (within 300m) any existing 1990 development it is called leapfrog (cross-hatched area).

Our buffer choice of 300m is guided by the urban 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 as is the whole notion, and we experimented with a different size buffer as reported later under robustness checks.

We estimate equations (1) and (2) for the key outcome: the count of leapfrog patches, where over 83% appear beyond the 1990 smoothed built area boundary, which will tell us whether Anglophone cities, other things being equal, have more leapfrog, or disconnected developments. Note this smoothed boundary will be well beyond the 1975 smoothed boundary, given urban land areas on average more than double between 1975 and 1990 . We also look at the ratio of leapfrog to all new patches of development, to give a sense of whether Anglophone development just has more new patches per se and, hence, less intensification of use within already developed areas (which we cannot measure).

A finding of more scattered development in Anglophone cities would suggest post-colonial patterns of disconnected and independent developments in Anglophone compared to Francophone cities which persist for new developments under today's inherited institutions.

6.1 Primary Results

While the estimating equation for leapfrog patches is based on equation (1), we have two periods of flows 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). Besides the usual controls, we control for initial built cover of the city in 1990, measured as the count of built pixels in 1990 within the 2014 lights boundary. Focusing on flows may help difference out the influence of key unobserved geographic factors. We cluster errors at the city level to deal with serial correlation[21].

We start with the main measure, the count of LF patches. However, a city can have more total patches of greenfield development overall, either LF or contiguous (infill or extension in Figure 4). So a second issue is whether Anglophone cities have higher ratios of LF to total new patches, and hence more scattering of new developments. This also allows us to infer whether Anglophone cities have more new patches in total. Note if $\frac{\partial ln(\text{count LF})}{\partial \text{Anglophone}} = \delta_1$ and $\frac{\partial [ln(\text{count LF}) - ln(\text{count total patches})]}{\partial \text{Anglophone}} = \delta_2$ then

$$\frac{\partial ln(\text{count total patches})}{\partial \text{Anglophone}} = \delta_1 - \delta_2 \tag{4}$$

Note if $\delta_1 = \delta_2$, while Anglophone cities do not have more new patches in total, they still have relatively more leapfrog ones.

Columns 1-3 in Table 3 show basic results for the logarithm of the count of LF 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 geographic and situational controls; and column 3 adds the economic controls. Geographic controls matter as do growth controls in the sense of reducing the magnitude of the Anglophone coefficient from 0.90 in column 1 to 0.54 in column 3. Coefficients on controls are given in Table A5.

In the main column 3 of Table 3, Anglophone cities have 54% 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 across these specifications. We also note the issue again that Francophone cities may respond to differential geography of cities differently than the Anglophone ones. For LF patches, the Oaxaca unexplained portion effect is in fact modestly larger than the base Anglophone coefficient, in Table 3, column 3.

In columns 4-6, we show results for the ln (count LP patches/ count total patches). The coefficient on the ratio in column 6 is 0.33, almost identical to the 0.34 without controls in column 4. With this ratio outcome, controls play little role. This ratio outcome is of interest in of itself, telling us that new patches of development in Anglophone cities are much more likely to be leapfrog ones[22]. Further, if we consider the results in columns 3 and 6 together, we can calculate

---

[21] Clustering standard errors at country-year level barely changes the statistical significance of the results. An alternative is to cluster the errors at country level to account for both spatial and temporal correction, but there are not really a sufficient number of countries to comfortably cluster at this level.

[22] Moreover for this outcome, given Oster (2019), the potential for bias is limited; in the usual rule of thumb the coefficient could be lower by a maximum of just 0.11. The rule of thumb formula in Oster (2019) is

$$\beta^* \approx \tilde{\beta} - \delta[\dot{\beta} - \tilde{\beta}] \frac{R_{max} - \tilde{R}}{\tilde{R} - \dot{R}}$$

the marginal effect of Anglophone over Francophone for total patches, which is about 0.21 (0.539 - 0.325). This implies that, for the same size and growth rate, Anglophone cities develop more by building in greenfield areas (all new patches), rather than intensifying already built cover. Then, given that, they are more prone to these added patches being leapfrog ones. That is, apart from utilizing greenfield development rather than intensification of existing built areas, Anglophone cities have less focus on contiguity of new developments to old.

In column 7 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 new development, especially leapfrog development, where these leapfrog patches are no bigger than their Francophone counterparts.

6.2 Identification

Are the effects in Table 3 causal? Qualitatively, causality might be suggested 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. The insertion of controls in column 3 does affect the magnitude of the Anglophone "treatment"; and the characteristics between the Anglophone and Francophone sets of cities are not balanced in all cases. In column 1 Table A7, Anglophone cities are on average higher up, more rugged, and have less river area. Obviously, if there are differences in observables which matter, there may be colonial origin differences in unobservables which affect outcomes. To deal with this, we turn to a border experiment, to try to compare Anglophone versus Francophone cities facing identical circumstances.[23]

Figure 5 shows West Africa where 5 Anglophone countries share borders with a number of Francophone countries. At these borders there are no significant waterways and the colonial borders drawn historically generally were done so without reference to (or knowledge of) local conditions (Michalopoulos and Papaioannou, 2016). We show cities within a 100 km buffer of the borders involved. Results are almost the same if using a 125 km or 150 km buffer. We choose the smaller buffer to better ensure similar conditions across borders, 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 shown in Figure 5, to try to control for unobserved geographic or other influences.[24] To form clusters, we first match all Anglophone cities to the nearest

---

which applies under strict assumptions about the relation of omitted versus control variables to each other and the treatment variable. In this β^* is the true coefficient. $\tilde{\beta}$ and $\tilde{R}$ are the estimated coefficient and R^2 with controls; while $\dot{\beta}$ and $\dot{R}$ are those with no controls. δ has an upper bound of 1 under equal selection (unobservables and observables equally related to the treatment). R_{max} is the maximum explanatory power obtainable from observed and omitted variables. The rest comes from measurement error or purely idiosyncratic items. We think measurement error could be fairly high for this variable, but even if we set $R_{max} = 1$, the maximum bias by this formula is 0.11.

[23] One can think of other experiments like comparing new political cities in Francophone versus Anglophone countries to see which are more compact. The problem is that the sample is limited and these are very peculiar cases, with layout driven by particular architectural designs, as in Abuja or Bur Sudan.

[24] We note that the areas in these countries of historical Islamic influence or not neatly span the border areas of comparison.

Francophone city and then assign any so far unmatched Francophone cities to the nearest cluster. The outcome is appealing. 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, there are 8 clusters, ignoring the green border cluster in southwest Cameroon for the moment. The cities in Burkina Faso have no Ghana counterpart, so these cities are neutralized by their cluster fixed effect. An issue with 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-1960s. We did the analysis both excluding and including this area, which is one 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? In Table A7, we give our key covariates from column 3 of Table 1 regressed on a constant and the Anglophone indicator. As noted above, 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, only 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 overall 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 within country weather stations and interpolation.

In Table 4 for the border sample, we run the same leapfrog regressions as in Table 3 with the border sample. We limit the controls to eight variables[25]. We show the results for a base case without fixed effects in row 1 and then in row 2, we add the city cluster fixed effects to control for unobservables. Results compare 35 Anglophone and 23 Francophone cities generally having two growth incidents. 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) for each of the two rows.

The results are revealing. In Table 4, the estimate of the differential Anglophone degree of leapfrogging is noticeably higher, not smaller, in both specifications, than in Table 3. Our preferred results are for city-cluster fixed effects where the Table 4 coefficient is 0.80 versus 0.54 in Table 3. For the ratio of leapfrog to all patches, again magnitudes are higher than in Table 3 (0.53 versus 0.33). The net effect on total patches is also higher than in the overall sample: 0.27 (vs 0.21). As is the case in Table 3, areas of patches do not differ by colonial origin. Table 4 gives strong evidence

---

[25] Given the small size of the border sample, we want to consolidate our list of controls to avoid potential low statistical power due to a long list of independent variables. Geographic controls are less important than economic controls as we have minimized the geographic difference in the border and by including city cluster fixed effects, so we keep all economic controls. We kept rainfall as it was unbalanced in column 3 in Table A7. We chose coast dummy over coast length, and chose ruggedness over elevation range. We dropped fraction of rivers and lakes as there are not significant inland rivers and lakes observed in the cities along the shared borders.

that it is colonialism and not other factors driving our results.[26]

6.3 Robustness

How robust are the basic results in Table 3 to other considerations. The key considerations are the following and we will tackle each in turn. Does the Anglophone differential diminish with time, as cities continue to grow even more beyond the colonial era area of the city? Are results sensitive to precise choice of city and country samples and buffering in defining LF patches? Are our results sensitive to different definitions of what area comprises the 2014 city and what are issues in defining that area as they affect results? What happens when we look more closely at ways to think about what is the extensive margin well beyond colonial era physical influence? Finally, do our results differ for colonial era cities, compared to cities subject to histories before the colonial era?

Non-diminution of colonial effects with time. On the first robustness test, we rerun the specifications in columns 3 and 6 of Table 3, allowing for a difference in the colonial origin effect between 90-00 and 00-14 (still controlling for a general time difference between 90-00 and 00-14 with the time dummy). To do this we simply interact Anglophone with the 2014 time dummy. This check is important. If colonial origin effects persist just as strongly for the later 00-14 episode compared to the 90-00, when we are even further out in spatial development of the city and time from the colonial era, this strengthens the extensive margin story. We also note that the spatial area of cities defined by smoothed cover in Appendix A.1.3 grew by 234% from 1990-2000. In results footnoted, this interactive effect for leapfrog patches and the ratio of leapfrog to total patches is always *positive*, indicting that potentially effects actually increase with time when looking at new developments even more distant from the colonial core city. However none of the interactive coefficients are significant.[27] The conclusion is that there is no evidence of diminution of effects with time. Thus, what we see is much more likely to be persistence in planning norms from the colonial era, not simply persistence and accretions in physical development from the colonial era.

Robustness to LF patch definitions. Robustness checks on samples and LF patch definitions are in Table 5. In Table 5 column 1, we show the base case. In each of the three sets of 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, we start by experimenting 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 leapfrog developments. In both cases, the impact on point estimates is fairly minimal and the differential

---

[26] To complete the picture we also reran Table 1 specifications on openness and area. However the cross-section sample size is very small (54) and results are insignificant.

[27] In columns 3 and 6 in Table 3 the interactive coefficients (and s.e.'s) are 0.416 (0.258) and 0.190 (0.156) while the base Anglophone effects are 0.322 (0.192) and 0.226 (0.108) respectively.

in coefficients between rows 1 and 2 which captures the marginal impact of Anglophone on total patches remains about 0.20.

Sample choices. Columns 4-7 deal with sample choice issues. Column 4 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) reduces coefficients, but the leapfrog coefficient is still 0.49 (vs 0.54). In column 5, dropping Nigeria which is a big portion of the sample, if anything, strengthens results. Column 6 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. Finally in column 7, we drop national capitals to show that the major political centres are not driving results.

Robustness to city boundary choices. In Table 5, columns 8 and 9, we start with the issue of defining city boundaries by night light readings. One issue is 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. Another issue concerns outliers: small cities with minimal electricity provision or inordinately large area cities created by long light ribbons (like along transport corridors). In column 8 we define city boundaries not by going to zero lights but by pixel light readings falling below a reading of 5. In column 9 we trim the top and the bottom 2.5% of cities, in terms of maximum distance from the centre to any part of the lights boundary. These experiments reduce coefficients but not significantly so. However, these experiments raise an issue: why are we using light boundaries at all? Because Anglophone cities are so scattered, if we instead defined city extent by a 2014 smoothed built cover boundary as defined in Appendix A.1.3, that would cut Anglophone cities short and bias our results. If we used that boundary, 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.

Defining the extensive margin. This bias just noted when using a built cover measure for city spatial extent presents a problem in defining an exact extensive margin. We tried one experiment which, given the bias, we do just for LF patches. We first defined a doughnut area for all cities between the 1990 smoothed cover boundary and the 2014 lights boundary. Given the inner border of the doughnut is based on smoothed cover and, in economic terms, for scattered Anglophone cities is probably drawn too near the centre, we try to mitigate this by looking just at the results for the 2000-2014 episode. For that episode, 87.5% of LF patches are beyond the 1990 boundary. For this regression based on the Table 3, column 3 specification, the Anglophone coefficient is highly significant at 0.757, noticeably higher than that in Table 3. This is consistent with the results reported above of Anglophone effects potentially increasing over time. In sum, given we are at margins far beyond any accretions of physical extensions of the colonial era area, persistence in differences in planning norms must be driving the Anglophone tendency to contemporaneous scattered development.

Results for colonial era cities. Last, in Table 6, we compare cities we know to be founded in the colonial era versus other cities. Other cities could be cities either founded before the colonial era or cities for which we have no information. We pick two dates to define the start of the colonial era, 1800 and 1850. We set the end of the colonial period as 1965. Founding dates are gleaned through a search of the DBpedia database [28]. Results for both founding dates are shown in Table 6, where we look at the leapfrog measures in Tables 3-5. In columns 1-4 we separate out non-colonial origin cities from colonial origin cities defined on the two dates. The idea is that with an increased ability of the French in new cities to impose their order from scratch, there would be a bigger Anglophone differential. The table confirms this with much bigger effects for both the leapfrog and ratio of leapfrog to total patch outcomes for colonial era cities compared to the rest of the sample, although the ratio coefficient has a more limited degree of statistical significance. Note that for the rest of the sample, our effects are undiminished. This is not surprising based on Ross and Bigon (2018), who argue that French planning was imposed throughout the urban hierarchy, while the British typically had little involvement in cities in which they had little presence (Home, 2015).

7 Public Policy Relevance

The planning literature argues that having less compact and more irregularly laid out cities raises the cost of infrastructure provision. For Africa, we further argue that such higher costs will lower the likelihood of receiving public infrastructure provision at all. 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 land line are at 13% and 6% respectively.

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 two years for which we measure can 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 linear probability model. First, the

---

[28] We queried the property of the established date of cities using their names through the link `http://dbpedia.org/property/establishedDate`

treatment variables are 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. Thus, also, colonial origin would be neither a valid instrument nor reduced form measure. We are trying to represent a cost effect of lack of compactness anywhere on public service provision. 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. Both leapfrogging and openness measures may matter. 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 excavation, piping, poles, and/or wiring per household entailing more costs and lower likelihood of being serviced.

The challenge in implementation concerns location. Within an urban area, for confidentiality reasons, the DHS randomizes cluster locations within 2 km of the true cluster location. They randomly pick a directional ray (angle) from the true cluster centre and then choose 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.

To deal with this, 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 how far it is from the city centre, and other controls. We interpret the measures as reflecting the overall degree of leapfrogging and openness in the surrounding area. 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.[29] Given the measurement error involved we did not anticipate seeing strong patterns in the data, but were surprised.

Results in Table 7 cover 42,748 - 44,561 households for utility attributes. 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. In the first column of a pair, there is the count of leapfrog developments as the cost factor; and, in the second, there are both the LF count and the share of built cover in the surrounding neighborhood. 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 system the decrease is 0.025 from a mean of 0.12. In the second set

---

[29] For example, using the propensity of surrounding areas to have LF development for whether cluster is recorded in a leapfrog patch does not.

of columns, the extent of built cover has significant impacts on all connections except a landline. Again effects are fairly small. A one standard deviation (0.31) increase 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. 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.

Does utility provision matter beyond "convenience"? For sewer and water connections the most critical outcomes are health ones, in particular in the DHS, diarrhea in the last two weeks for children under 5. In columns 9 and 10 we look at this. While a great count of LF patches increases the likelihood of diarrhea the effects are not significant. However in column 10, an increased share of built cover significantly lowers the incidence of diarrhea.

8 Conclusions

The literature on colonialism in Africa suggests that, compared to the British, the French imposed more comprehensive citywide land use planning, including the layout of roads. The literature suggests these planning practices persisted well into the post-colonial era. The African context of colonialism provides an experiment to show that institutions that involve more centralized land use control within each city, as in Francophone compared to Anglophone cities, leads to more compact cities at both the intensive and extensive margins. The differential in older colonial sections of the city may be explained by persistence of physical infrastructure and its layout. However for new developments since 1990, the differential is much more likely explained by persistence in planning traditions inherited from the colonial era.

Specifically the paper shows that Francophone African cities have more gridiron structures in their core areas. Anglophone cities have a citywide index of openness which is 17% higher and cover 29% more land. Anglophone intensity of land use is much lower at the centre and, in contrast to Francophone cities, the gradient for the intensity of land use is flatter. Anglophone cities are more sprawled. Correspondingly, with new development, Anglophone cities have about 54% 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. These Anglophone leapfrog effects do not diminish over time.

The question is whether there is a consequence to having greater leapfrog development and more sprawl. Such areas are more expensive to service and potentially less likely to receive connections to public utilities, such as electricity, phone landlines, piped water, and city sewers in an African context. We find that areas with more sprawl in a city in general have poorer connections, and a worse health outcome, diarrhea, for children.

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.

Ahlfeldt, G. and E. Pietrostefani (2019). The Economic Effects of Density: A Synthesis. *CEPR Discussion Paper 13440*.

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.

Anthony, K., M. Andrew, S. Andrew, M. Pia, S. S. Ehrlich, and S. Jeffrey (2004). A global index representing the stability of malaria transmission. *The American Journal of Tropical Medicine and Hygiene 70*(5), 486–498.

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*.

Campante, F. R. and Q.-A. Do (2014, August). Isolated capital cities, accountability, and corruption: Evidence from US states. *American Economic Review 104*(8), 2456–81.

Carruthers, J. I. and G. F. Ulfarsson (2003). Urban sprawl and the cost of public services. *Environment and Planning B: Planning and Design 30*(4), 503–522.

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.

Desmet, K., J. Gomes, and I. Ortuño-Ortín (2018, June). The Geography of Linguistic Diversity and the

Provision of Public Goods. NBER Working Papers 24694, National Bureau of Economic Research, Inc.

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. (2005). *Land for housing the poor in African cities: Are neo-customary processes an effective alternative to formal systems?*, Chapter 12, pp. 160–174.

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

Elvidge, C. D., K. E. Baugh, E. A. Kihn, H. W. Kroehl, E. R. Davis, and C. W. Davis (1997). Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing 18*(6), 1373–1379.

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.

Europa (2003). Africa South of the Sahara 2004.

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.

Giuliano, P. and N. Nunn (2017, July). Understanding cultural persistence and change. Working Paper 23617, National Bureau of Economic Research.

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. (2020, August). Cities in bad shape: Urban geometry in india. *American Economic Review 110*(8), 2377–2421.

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*. 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*.

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

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.

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

Njoh, A. (2001). *Planning Rules in Post-colonial States: The Political Economy of Urban and Regional Planning in Cameroon*. Nova Science Publishers.

Njoh, A. (2004). The experience and legacy of French colonial urban planning in Sub-Saharan Africa. *Planning Perspectives 19*(4), 435–454.

Njoh, A. (2006). *Planning power: town planning and social control in colonial Africa*. UCL Press.

Njoh, A. (2016). *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.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review 14*, 693–709.

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.

Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: Current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography 11*(5), 377–392.

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

Ross, E. and L. Bigon (2018). The urban grid and entangled planning cultures in Senegal. *Planning Perspectives 0*(0), 1–27.

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.

Scholz, W., P. Robinson, and T. Dayaram (2015). *Colonial Planning Concept and Post Colonial Realities: The Influence of British Planning Culture in Tanzania, South Africa and Ghana*. Routledge.

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)
Anglophone dummy	0.229***	0.172***	0.173***	0.355**	0.282**	0.285***
	(0.045)	(0.044)	(0.055)	(0.151)	(0.116)	(0.093)
Ln annual population			-0.522			8.453***
growth 90 to 12			(1.217)			(2.241)
Ln projected city			-0.174***			0.857***
population 1990			(0.025)			(0.042)
Geographic and Situational Controls	No	Yes	Yes	No	Yes	Yes
Economic Controls	No	No	Yes	No	No	Yes
R^2	0.077	0.212	0.341	0.014	0.492	0.786
N	307	307	281	307	307	281
Oaxaca decomposition						
Explained		0.063*	0.119*		0.052	0.228
		(0.035)	(0.061)		(0.131)	(0.162)
Unexplained		0.166***	0.142*		0.303**	0.267**
		(0.056)	(0.073)		(0.145)	(0.104)

Note: Dependent variable is ln openness in the year 2014 in columns 1-3, and ln area in columns 4-6. Geographic and situational controls include ln ruggedness, ln rainfall, ln elevation range, coastal dummy, interaction of ln coast length with coastal dummy, ln distance to coast if non-coastal, malaria index, land suitability index, ln temperature, ethnic fractionalization index at level 15 in year 1975, distance to national capital in the year 1990 and non-national capital dummy. Economic controls include ln annual population growth from 1990 to 2012, ln projected city population in 1990, and ln country GDP per capita in 1990. Robust standard errors are applied, noting there are too few countries to cluster at the country level.

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

Table 2: Intensity by rings in 2014

	(1) 1km	(2) 2km	(3) 3km	(4) 4km	(5) 5km	(6) 6km	(7) Intensity Gradient
Anglophone Dummy	-0.202	-0.456***	-0.615***	-0.636**	-0.212	-0.136	-0.726***
	(0.136)	(0.125)	(0.180)	(0.269)	(0.321)	(0.325)	(0.204)
Ln ring total pixel	1.321***	1.569***	0.848***	0.750***	0.953***	0.777***	0.628***
	(0.354)	(0.430)	(0.180)	(0.153)	(0.144)	(0.167)	(0.056)
Ring distance							-0.287***
							(0.033)
Ring distance × Anglophone							0.093***
							(0.034)
Anglophone mean	1.734	3.865	4.170	3.916	3.720	3.874	
Francophone mean	1.889	4.279	4.409	3.800	3.403	3.409	
Geographic and Situational Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.181	0.348	0.419	0.447	0.556	0.536	0.481
N	284	286	283	273	251	216	3178
Oaxaca decomposition							
Explained	0.070	0.143	0.205	0.384	0.486	0.432	
	(0.149)	(0.182)	(0.234)	(0.331)	(0.380)	(0.437)	
Unexplained	-0.185	-0.320*	-0.446*	-0.534	-0.177	0.177	
	(0.171)	(0.192)	(0.239)	(0.339)	(0.378)	(0.440)	

Note: Dependent variable is ln built-up area in 2014 in a ring. The first 6 columns stratify the sample by rings. Column 7 shows the intensity gradient for rings up to 20km. Economic control variables include ln projected city population in 1990, ln country GDP per capita in 1990, and city population growth from 1990 to 2014. Geography controls include ln ruggedness, ln rainfall, ln elevation range, coastal dummy, interaction of ln coast length with coastal dummy, ln distance to coast if non-coastal, malaria index, land suitability index, ln temperature, ethnic fractionalization index at level 15 in year 1975, distance to national capital in the year 1990 and non-national capital dummy. Anglophone mean and Francophone mean report mean built-up area in both groups in square kilometers. Standard errors are robust and clustered at city level in column 7 given spatial correlation.

* 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)
Anglophone dummy	0.907***	0.685***	0.539***	0.339***	0.292***	0.325***	-0.013
	(0.139)	(0.138)	(0.167)	(0.099)	(0.101)	(0.118)	(0.061)
Ln initial cover 1990	0.656***	0.533***	0.295***	-0.081**	-0.151***	-0.295***	0.022
	(0.049)	(0.049)	(0.058)	(0.033)	(0.034)	(0.040)	(0.020)
Year dummy 2014	0.517***	0.505***	0.495***	0.119**	0.117**	0.119**	0.137***
	(0.065)	(0.065)	(0.067)	(0.053)	(0.054)	(0.056)	(0.035)
Ln annual population growth 90 to 12			12.114***			5.842**	2.215*
			(3.338)			(2.585)	(1.287)
Ln projected city population 1990			0.712***			0.443***	0.036
			(0.101)			(0.072)	(0.035)
Geographic and Situational Controls	No	Yes	Yes	No	Yes	Yes	Yes
Economic Controls	No	No	Yes	No	No	Yes	Yes
R^2	0.446	0.536	0.602	0.053	0.131	0.230	0.109
N	606	606	551	606	606	551	525
Oaxaca decomposition							
Explained	-0.181	0.079	0.124	0.035	0.105*	0.133	0.063
	(0.113)	(0.159)	(0.201)	(0.023)	(0.060)	(0.102)	(0.043)
Unexplained	0.894***	0.635***	0.613***	0.328***	0.258**	0.314**	-0.022
	(0.139)	(0.150)	(0.183)	(0.099)	(0.104)	(0.129)	(0.061)

Note: Geography controls include ln ruggedness, ln rainfall, ln elevation range, coastal dummy, interaction of ln coast length with coastal dummy, ln distance to coast if non-coastal, malaria index, land suitability index, ln temperature, ethnic fractionalization index at level 15 in year 1975, distance to national capital in the year 1990 and non-national capital dummy. Economic controls include ln annual population growth from 1990 to 2012, ln projected city population in 1990, lag $t-1$ ln country GDP per capita. Finally, there is a control for 1990 initial cover, ln count of built pixels in 1990 within the 2014 boundary. Standard errors are robust and clustered at the city level to capture the serial correlation.

* 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	1.033**	0.820***	-0.131
	(0.432)	(0.302)	(0.169)
City cluster FE's			
Anglophone dummy	0.796**	0.529***	-0.200
	(0.306)	(0.184)	(0.131)
R^2	0.669	0.425	0.264
N	108	108	103

Note: Controls include 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, ln rainfall, a coastal dummy, and ln intial cover 1990. Standard errors are robust and 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) No German colonies	(5) No Nigeria	(6) 40 cities	(7) Non capital	(8) Light 5	(9) Distance trim
Ln count of LF									
Anglophone	0.539***	0.503***	0.414***	0.486**	0.579***	0.650**	0.504***	0.449***	0.390**
dummy	(0.167)	(0.166)	(0.139)	(0.207)	(0.215)	(0.243)	(0.184)	(0.169)	(0.163)
Ln LF minus ln total patches									
Anglophone	0.325***	0.292**	0.201***	0.185	0.300**	0.340	0.370***	0.288**	0.256**
dummy	(0.118)	(0.122)	(0.068)	(0.151)	(0.150)	(0.245)	(0.131)	(0.118)	(0.117)
Ln avg. LF area									
Anglophone	-0.013	-0.031	-0.067	-0.154**	0.007	0.116	-0.006	-0.028	-0.027
dummy	(0.061)	(0.053)	(0.045)	(0.078)	(0.067)	(0.127)	(0.066)	(0.062)	(0.062)
R^2	0.602	0.604	0.700	0.609	0.615	0.585	0.532	0.605	0.573
N	551	551	551	489	330	49	518	551	529
Oaxaca decomposition									
(for count of LF patches)									
Explained	0.124	0.190	0.113	0.078	0.343	-1.015*	0.355*	0.278	0.111
	(0.201)	(0.204)	(0.177)	(0.238)	(0.262)	(0.545)	(0.183)	(0.201)	(0.208)
Unexplained	0.613***	0.522***	0.493***	0.591***	0.587**	1.056**	0.556***	0.450**	0.479**
	(0.183)	(0.178)	(0.145)	(0.222)	(0.248)	(0.512)	(0.203)	(0.185)	(0.187)

Note: Columns 1-5 and 7-9 include same controls as columns 3, 6 and 7 in Table 3. Column 6 includes the same controls as in Table 4. Standard errors are clustered at city level in columns 1-5 and 7-9, and are robust. Adjusted R^2 and N are reported for ln count of LF.

Table 6: Colonial origin and pre-colonial institutions

	1800-1965		1850-1965	
	Colonial origin (1)	Non-colonial origin (2)	Colonial origin (3)	Non-colonial origin (4)
Ln count of LF				
Anglophone dummy	1.653***	0.627***	1.836***	0.619***
	(0.482)	(0.169)	(0.547)	(0.159)
Ln LF minus ln total patches				
Anglophone dummy	0.508*	0.254**	0.374	0.232**
	(0.277)	(0.121)	(0.297)	(0.114)
R^2	0.652	0.481	0.575	0.503
N	63	543	50	556

Note: Colonial origin cities are defined as the cities established between year 1800 and year 1965 in the first two columns and between year 1850 and year 1965 in the last two columns. Controls include ln initial cover 1990, year dummy 2014, lag t-1 ln country GDP per capita, change in ln country level urban population from 1990 to 2014, ln projected city population in 1990, ln average ruggedness, ln rainfall, and a coastal dummy. Adjusted R^2 and N are reported for ln count of LF. Standard errors are robust and clustered at the city level.

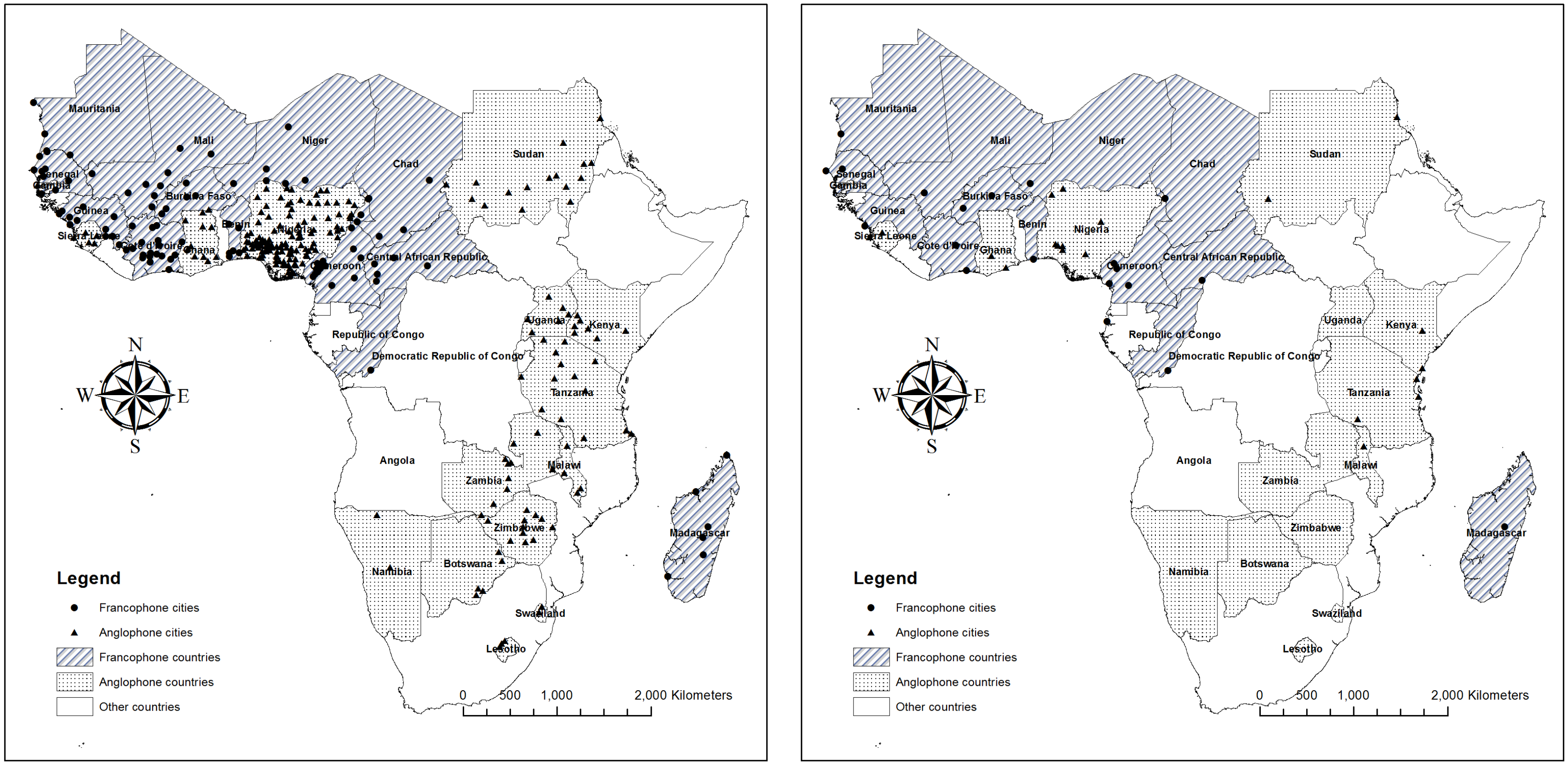
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 7: Public utility connection

	Has electricity		Has piped water		Has flush toilet		Has phone land line		Has diarrhea	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Count of LF	-0.0033**	-0.0026*	-0.0002	0.0008	-0.0045***	-0.0031***	-0.0019**	-0.0020**	0.0008	0.0001
	(0.0014)	(0.0014)	(0.0013)	(0.0014)	(0.0010)	(0.0010)	(0.0008)	(0.0009)	(0.0007)	(0.0008)
Share of built cover in buffer		0.0526*		0.0797**		0.1168***		-0.0026		-0.0420**
		(0.0298)		(0.0364)		(0.0266)		(0.0128)		(0.0178)
Ln buffer center distance	-0.0727***	-0.0666***	-0.0700***	-0.0608***	-0.0383***	-0.0248***	-0.0124***	-0.0127***	-0.0097**	-0.0141***
	(0.0066)	(0.0075)	(0.0088)	(0.0096)	(0.0062)	(0.0073)	(0.0035)	(0.0038)	(0.0041)	(0.0045)
Ln buffer ruggedness	0.0553***	0.0507***	0.0565***	0.0495**	-0.0054	-0.0157	-0.0012	-0.0010	0.0242**	0.0261**
	(0.0165)	(0.0171)	(0.0204)	(0.0210)	(0.0120)	(0.0124)	(0.0062)	(0.0064)	(0.0114)	(0.0112)
Buffer has river of lake	0.0209	0.0232	0.0361	0.0395	-0.0057	-0.0007	0.0041	0.0040	0.0008	-0.0024
	(0.0214)	(0.0214)	(0.0272)	(0.0270)	(0.0174)	(0.0175)	(0.0122)	(0.0121)	(0.0145)	(0.0145)
Household size	0.0086***	0.0086***	0.0024***	0.0024***	0.0026***	0.0026***	0.0069***	0.0069***	0.0036	0.0038
	(0.0006)	(0.0006)	(0.0006)	(0.0006)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0036)	(0.0036)
Sex of household head: Male	-0.0075	-0.0076	-0.0092*	-0.0094**	-0.0159***	-0.0161***	-0.0096***	-0.0096***	0.0290	0.0289
	(0.0050)	(0.0050)	(0.0047)	(0.0047)	(0.0041)	(0.0040)	(0.0028)	(0.0028)	(0.0279)	(0.0275)
Highest educational level of head: Primary	0.0367***	0.0374***	0.0117	0.0128*	-0.0052	-0.0036	0.0126***	0.0126***	0.1097***	0.0972**
	(0.0072)	(0.0072)	(0.0073)	(0.0073)	(0.0045)	(0.0045)	(0.0034)	(0.0034)	(0.0383)	(0.0383)
Highest educational level of head: Secondary	0.1739***	0.1740***	0.0519***	0.0521***	0.0407***	0.0410***	0.0424***	0.0424***	0.0401	0.0399
	(0.0070)	(0.0070)	(0.0071)	(0.0071)	(0.0047)	(0.0047)	(0.0038)	(0.0038)	(0.0288)	(0.0286)
Highest educational level of head: Higher	0.2877***	0.2877***	0.0578***	0.0577***	0.1652***	0.1652***	0.1124***	0.1124***	-0.0558**	-0.0597**
	(0.0088)	(0.0088)	(0.0094)	(0.0094)	(0.0089)	(0.0089)	(0.0072)	(0.0072)	(0.0274)	(0.0271)
Highest educational level of head: Unknown	0.1417***	0.1417***	0.0457**	0.0457**	0.0124	0.0124	0.0027	0.0026	0.0167	0.0054
	(0.0247)	(0.0247)	(0.0225)	(0.0225)	(0.0165)	(0.0165)	(0.0150)	(0.0150)	(0.1826)	(0.1826)
Period dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	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	0.032	0.032
N	44517	44517	44561	44561	44500	44500	42748	42748	17840	17840

Note: Built cover measures for 2000 and 2014 are matched to DHS surveys within 2 years of the respective base dates, 2000 and 2014. Period dummies control for the difference between the DHS survey years and year 2000 or 2014. Period dummies include 1 year before the built cover measure, 1 year after the measure, 2 years before the measure, and 2 years after the measure. 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

Figure 1: Spatial distribution of sample cities

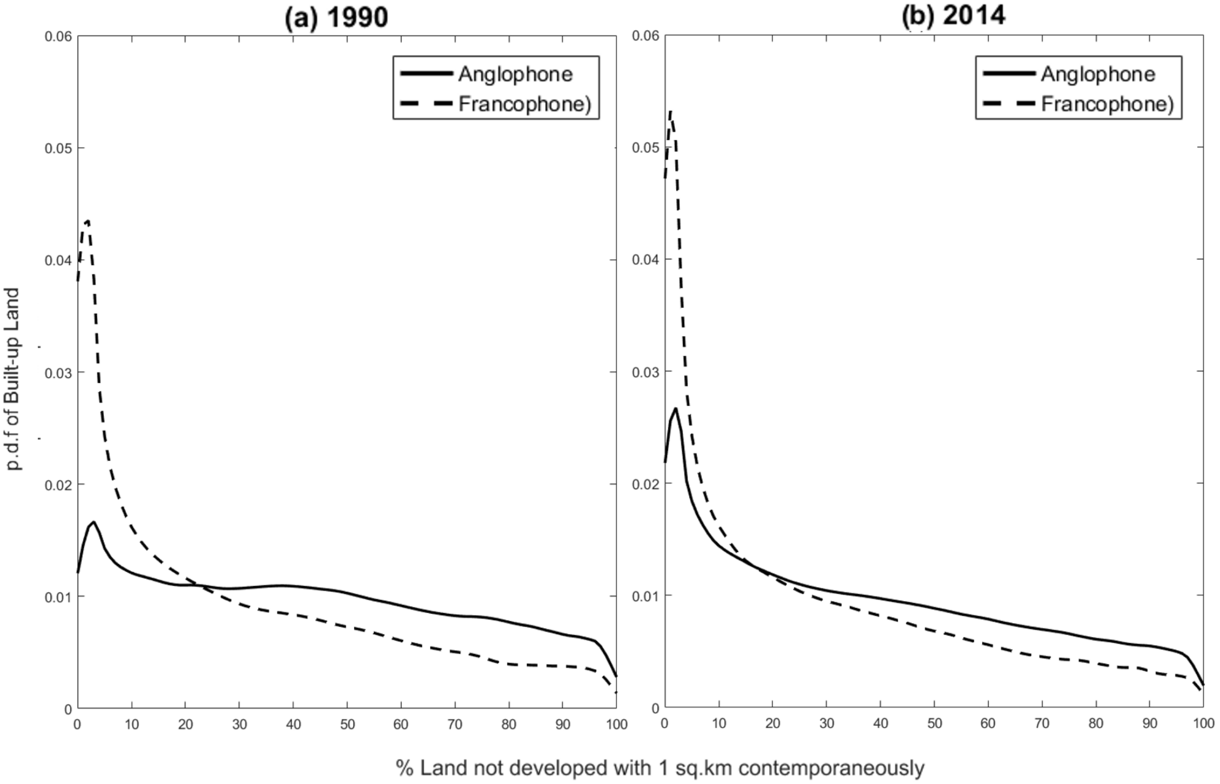
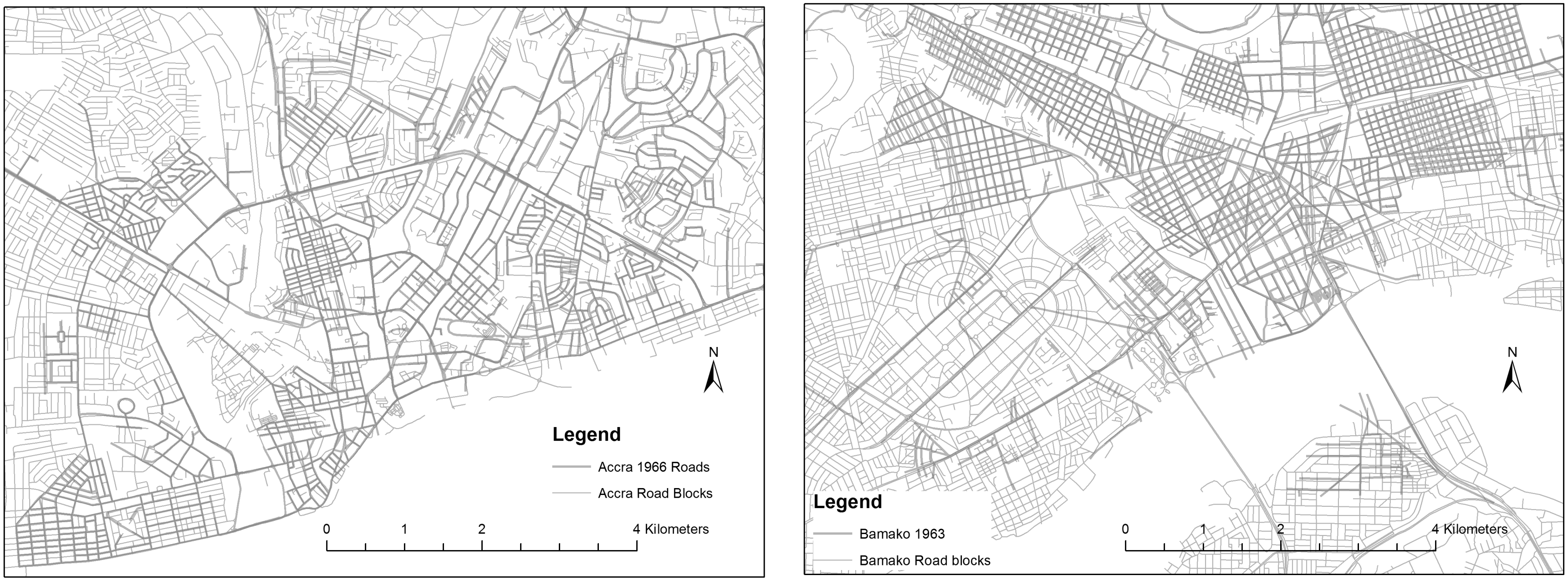


Figure 2: Probability function of Anglophone and Francophone built-up land across areas with different degrees of sprawl for (a) 1990 and (b) 2014



Accra

Bamako

Figure 3a: Road blocks in Accra and Bamako

Harare

Brazzaville

Figure 3b: Road blocks in Harare and Brazzaville

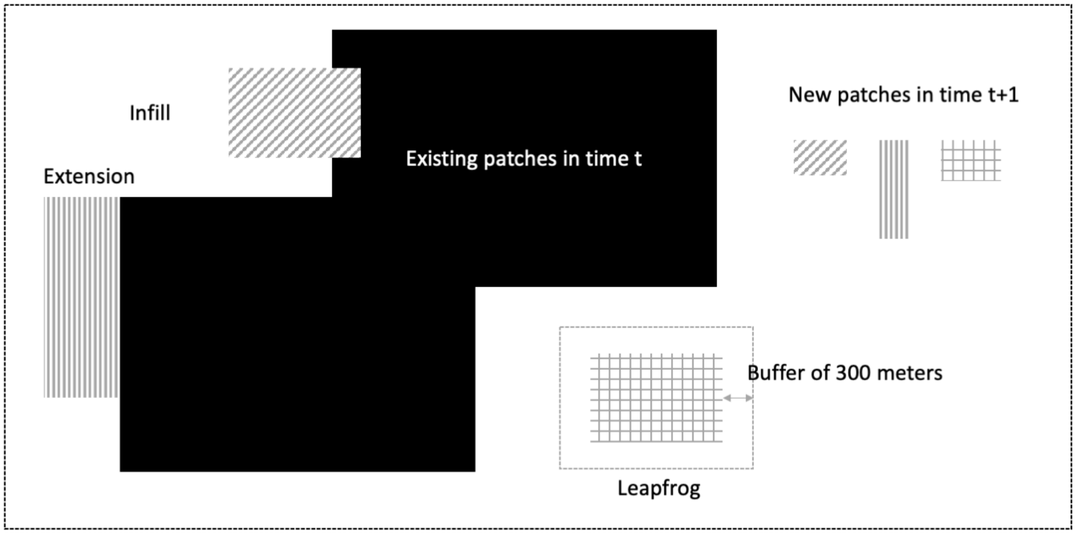


Figure 4: Illustration: Defining leapfrog patches

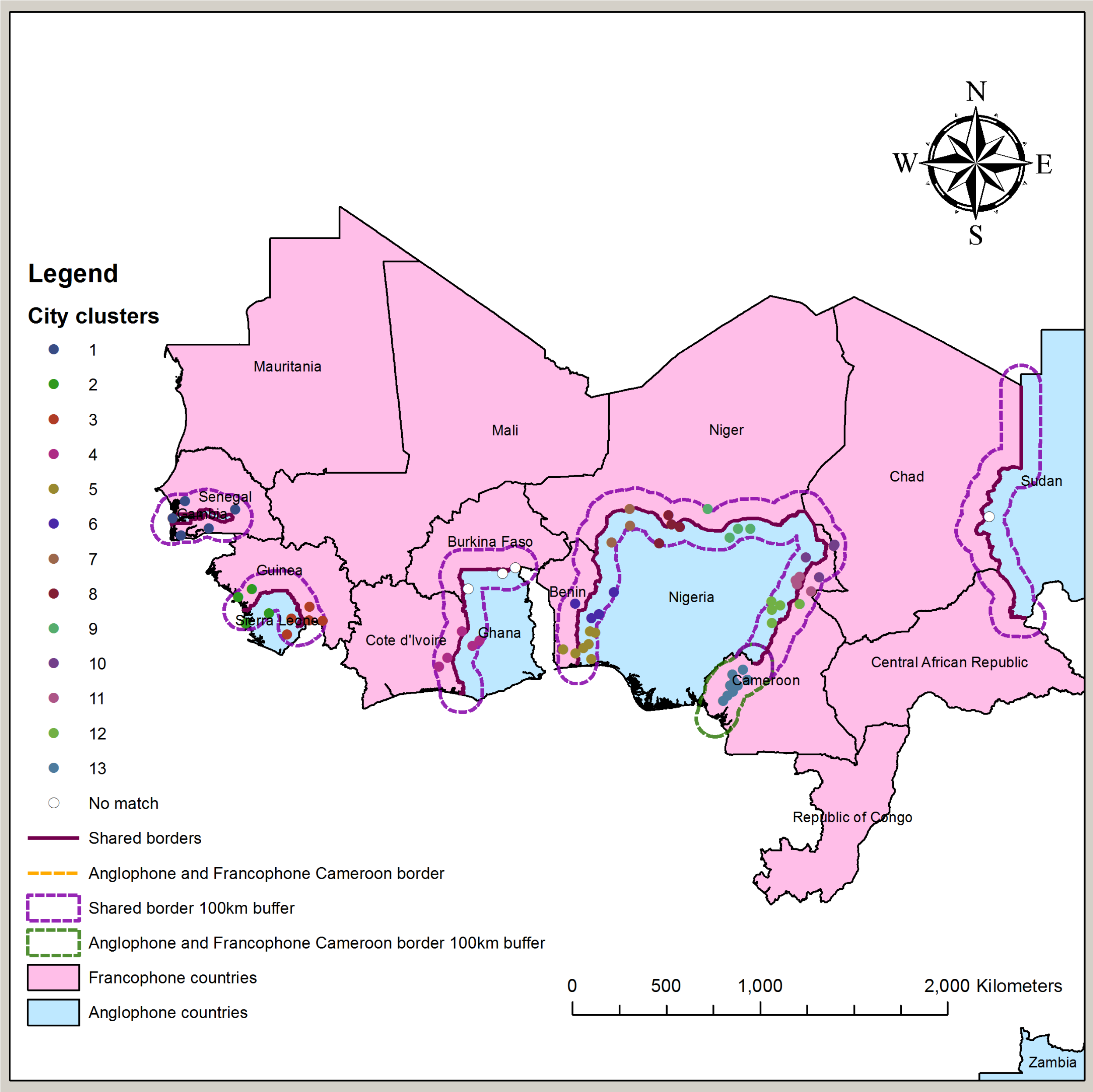


Figure 5: Shared borders

Appendix. Gridiron Structures

To test whether there are differences in colonial era spatial layouts, we took all 20 Francophone cities in Sub-Saharan Africa over 300,000 in 2012, to analyse road layouts from OSM. Since OSM data are 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 elevation. With matching, means of the matching variables show insignificant differences between Francophone and chosen Anglophone cities as shown in Appendix Table 1. In the end there are 11 Nigerian out of 20 Anglophone cities, effectively matching Francophone ones concentrated in West Africa. Other samples drawn to reduce the Nigerian count show similar if not stronger results. [30]

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. Appendix Figure 1 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 part B shows the derived city blocks. City 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 block would be 1, and the ratio tends to fall as it takes on more complex shapes. Part C 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 the figure shows how we define *gridiron blocks*, which is the basis for our main measure and captures contiguity in rectangularity of layout in the city. To be a gridiron block, a block must be rectangular, 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* or T-intersection; and they are illustrated in Part E of the figure. Part D 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, to capture

---

[30] 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.

the degree to which there is an overall, and connected gridiron structure.

We believe OSM data pretty comprehensively map 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 Appendix Figure 2, 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 percentage 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 overall diminishing of regularity with the 150 percent increase in area covered. 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 (2016) and Durand-Lasserve (2005), which persists until today.[31]

---

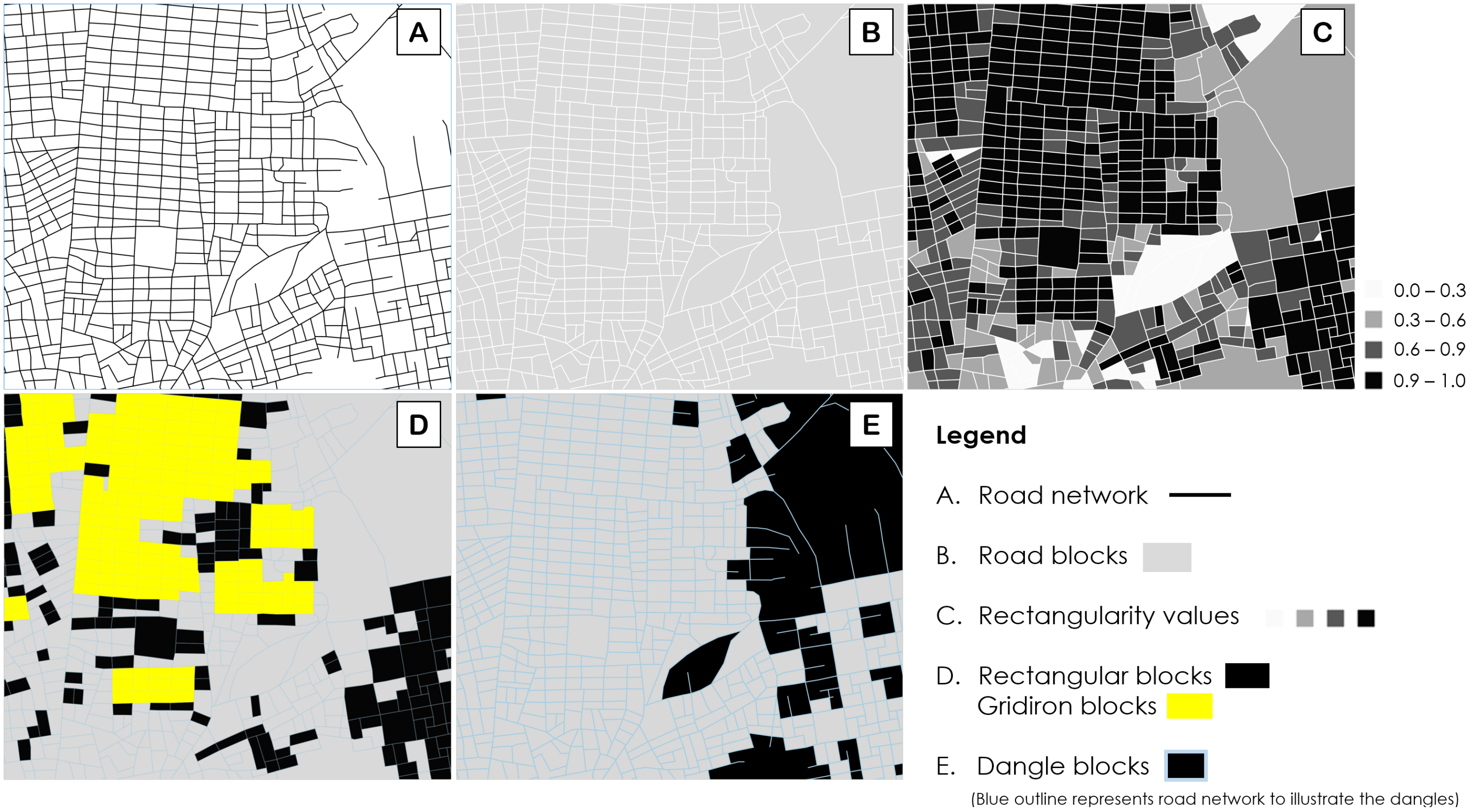
[31] One issue is whether Anglophone cities were regularly laid out but just not on a rectangular grid, using more diagonal roads with roundabout intersections. We checked the count of roundabouts within 5 km of the centre. On average there is no difference between Francophone and Anglophone cities.

Appendix Table 1: Balance test for 40 cities sample

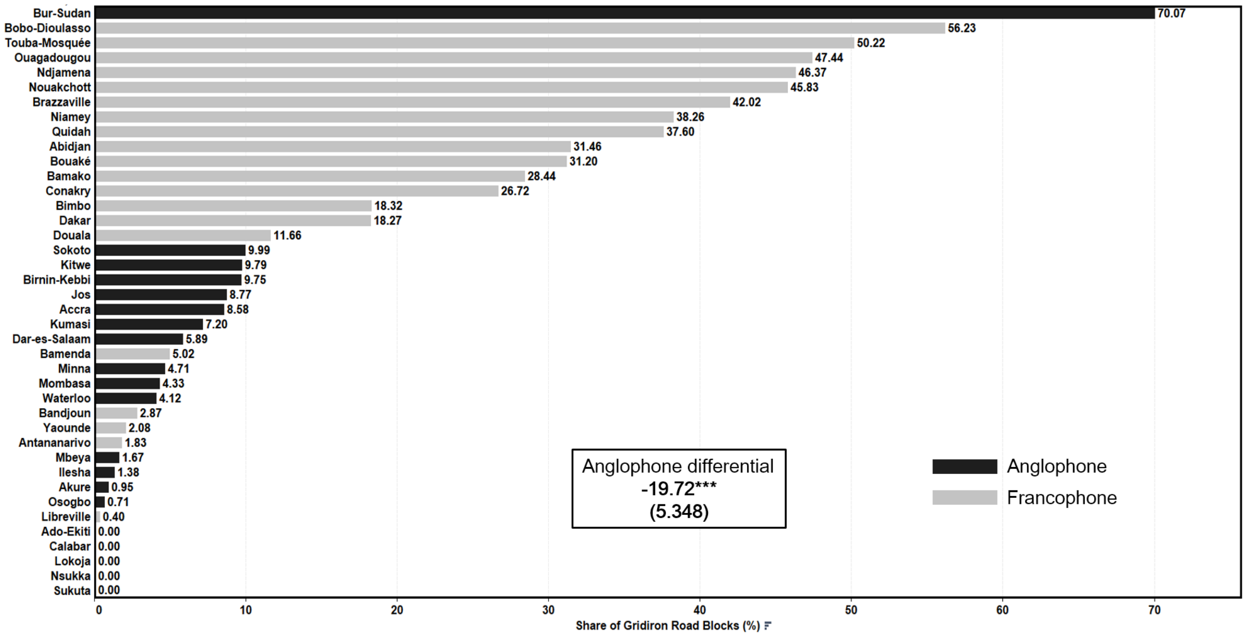
	(1) Project city population 1990 (Thousands)	(2) Ln annual city population growth 90-12	(3) Rainfall	(4) Coast Dummy	(5) Elevation
Anglophone dummy	-214.631	-0.018	-7.533	-0.050	-1.969
	(163.192)	(0.108)	(19.921)	(0.152)	(143.559)
Constant	692.799***	0.882***	120.676***	0.350***	378.372***
	(125.665)	(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.



Appendix Figure 1: Road blocks and rectangularity



Appendix Figure 2: Share of gridiron road blocks within contemporary 5km

Appendix A.

A Online 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 (OSM) cities is derived from digitalising historical maps from Oxford and Cambridge library. Current road blocks data is extracted from 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 convolution smoothing algorithm to define the city built cover boundary. First, scattered built-up settlements are removed using a sieve operation, where pixel aggregations below a certain areal threshold are removed. Second, a low-pass filter was applied to the built cover data at 30m spatial resolution to blur the data using Gaussian smoothing and remove noise. A queen neighbourhood filter of radius 200m was found to be appropriate for this exercise. The Gaussian filter is similar to a mean filter - which replaces a pixel of interest with the mean of its neighbourhood - but it uses a different kernel that represents the shape of a Gaussian ("bell-shaped") hump. Third, we extract the "blurred" built-up blob and convert it to a vector outline using a polygonisation (raster to polygon) process in Geographic Information Systems. Finally, for a given city, the central outline with the biggest area is chosen as the city built cover boundary.

A.2 Appendix Tables

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	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	1850	
Gaborone	Botswana	1	215,068	487,079	1850	
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	1850	
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	40,856		
Ngaoundéré	Cameroon	0	87,298	198,223		
Nkongsamba	Cameroon	0	88,275	112,347	1850	
Yaounde	Cameroon	0	771,858	2,744,390		

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
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	1850	
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	nan		
Abidjan	Cote d'Ivoire	0	2,312,639	4,395,000		
Akoupé	Cote d'Ivoire	0	38,495	nan		
Bondoukou	Cote d'Ivoire	0	35,283	nan		
Bouaflé	Cote d'Ivoire	0	37,918	nan		
Bouaké	Cote d'Ivoire	0	352,785	536,719		
Daloa	Cote d'Ivoire	0	130,708	nan		
Danané	Cote d'Ivoire	0	34,582	nan		
Dimbokro	Cote d'Ivoire	0	39,581	nan		
Ferkéssédougou	Cote d'Ivoire	0	40,675	nan		
Gagnoa	Cote d'Ivoire	0	112,890	nan		
Issia	Cote d'Ivoire	0	30,922	nan		
Katiola	Cote d'Ivoire	0	34,581	nan		
Korhogo	Cote d'Ivoire	0	115,302	nan		
Man	Cote d'Ivoire	0	94,435	nan		
Odienné	Cote d'Ivoire	0	31,202	nan		
Sinfra	Cote d'Ivoire	0	37,773	nan		
Séguéla	Cote d'Ivoire	0	31,517	nan		
Yamoussoukro	Cote d'Ivoire	0	139,062	nan		
Libreville	Gabon	0	394,152	694,622		
Banjul	Gambia	1	357,893	460,450	1800	Yes
Accra	Ghana	1	2,004,164	3,689,581		Yes

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
Bawku	Ghana	1	39,747	63,318		
Bolgatanga	Ghana	1	37,953	69,431		
Dzodze	Ghana	1	52,458	nan		
Ho	Ghana	1	45,396	116,172		
Koforidua	Ghana	1	68,148	129,122		
Kumasi	Ghana	1	836,568	2,382,130		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
Nairobi	Kenya	1	1,516,055	5,044,352	1850	
Nakuru	Kenya	1	170,002	336,431	1850	
Maputsoe	Lesotho	1	59,779	103,567		
Maseru	Lesotho	1	117,442	178,016	1850	
Teyateyaneng	Lesotho	1	42,583	61,599		
Antananarivo	Madagascar	0	675,058	1,300,000		
Antsirabe	Madagascar	0	117,026	nan	1850	
Antsiranana	Madagascar	0	54,808	nan		

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
Fianarantsoa	Madagascar	0	101,428	nan	1800	
Mahajanga	Madagascar	0	99,126	nan		
Toliara	Madagascar	0	75,032	nan		
Blantyre	Malawi	1	372,552	738,274	1850	
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	1850	
Koutiala	Mali	0	55,163	167,010		
Mopti	Mali	0	76,285	134,933	1800	
San	Mali	0	34,466	73,915		
Sikasso	Mali	0	87,024	261,123	1850	
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	1800	
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		
Abakaliki	Nigeria	1	158,289	439,893		
Abraka	Nigeria	1	119,940	259,762		
Abuja	Nigeria	1	384,364	3,028,556	1800	
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

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
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
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	1850	
Funtua	Nigeria	1	89,954	183,064		
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		
Hadejia	Nigeria	1	45,276	94,181		
Ibadan	Nigeria	1	1,711,452	2,911,228	1800	
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	nan		
Ikare	Nigeria	1	147,132	364,228		
Ikirun	Nigeria	1	215,476	427,992		
Ikole	Nigeria	1	56,932	100,183		
Ikom	Nigeria	1	40,718	52,109		

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
Ikot-Ekpene	Nigeria	1	146,477	nan		
Ikot-Etim	Nigeria	1	87,282	165,044		
Ila	Nigeria	1	43,213	59,975		
Ilesha	Nigeria	1	139,202	332,008		Yes
Ilorin	Nigeria	1	538,446	833,589		
Ilutitun	Nigeria	1	45,214	70,917		
Iseyin	Nigeria	1	47,732	174,531		
Iwo	Nigeria	1	88,314	240,838		
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		
Kafanchan	Nigeria	1	41,236	132,111		
Kano	Nigeria	1	1,385,370	3,734,597		
Katsina	Nigeria	1	189,505	425,669		
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		
Langtang	Nigeria	1	65,532	121,295		
Lokoja	Nigeria	1	63,547	375,656		Yes
Maiduguri	Nigeria	1	490,729	694,554		
Makurdi	Nigeria	1	179,494	301,249		
Malumfashi	Nigeria	1	46,775	58,968		
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		
Nguru	Nigeria	1	44,872	103,062		
Nkume	Nigeria	1	129,318	nan		
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		

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
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		
Onitsha	Nigeria	1	956,207	8,290,100		
Ore	Nigeria	1	45,689	102,651		
Oro-Esie-Iludin	Nigeria	1	46,096	75,454		
Osogbo	Nigeria	1	497,049	774,670		Yes
Otun-Ekiti	Nigeria	1	33,762	41,416		
Oturkpo	Nigeria	1	79,827	147,733		
Owo	Nigeria	1	103,021	186,305		
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
Takum	Nigeria	1	31,065	53,909		
Uba	Nigeria	1	55,350	70,447		
Ugep	Nigeria	1	34,279	149,847		
Umuahia	Nigeria	1	116,720	nan		
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		

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
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		
Freetown	Sierra Leone	1	561,004	1,049,768		Yes
Kenema	Sierra Leone	1	66,406	187,158		
Makeni	Sierra Leone	1	48,170	108,671		
Torgbonbu	Sierra Leone	1	95,889	98,014		
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		
Bur-Sudan	Sudan	1	293,338	421,429		Yes
El-Duein	Sudan	1	64,709	161,998		
El-Obeid	Sudan	1	211,433	384,829		
Gedaref	Sudan	1	178,488	295,201		
Kaduqli	Sudan	1	61,151	68,492		
Kassala	Sudan	1	223,586	318,335		
New-Halfa	Sudan	1	52,391	66,386		
Nyala	Sudan	1	194,574	606,114		
Sannar	Sudan	1	58,718	266,989		
Mbabane	Swaziland	1	82,878	nan	1850	
Tabankulu	Swaziland	1	30,730	nan		
Arusha	Tanzania	1	122,068	416,442	1850	
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

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
Mtwara	Tanzania	1	68,149	100,626		
Musoma	Tanzania	1	68,356	134,327		
Mwanza	Tanzania	1	193,317	706,453	1850	
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		
Zanzibar	Tanzania	1	174,467	501,459	1850	
Fort-Portal	Uganda	1	32,130	51,795		
Gulu	Uganda	1	34,535	146,233		
Kampala	Uganda	1	803,069	2,269,969		
Masaka	Uganda	1	47,671	112,864		
Mbale	Uganda	1	51,446	117,706		
Mbarara	Uganda	1	39,119	164,150	1850	
Njeru	Uganda	1	96,824	219,039	1850	
Soroti	Uganda	1	40,903	48,069		
Chipata	Zambia	1	52,213	128,045		
Choma	Zambia	1	30,143	54,492		
Kabwe	Zambia	1	154,318	207,909		
Kasama	Zambia	1	47,653	108,492		
Kitwe	Zambia	1	355,793	1,066,992		Yes
Livingstone	Zambia	1	76,875	143,249		
Luanshya	Zambia	1	118,143	133,187	1850	
Lusaka	Zambia	1	813,154	2,000,916	1850	
Mansa	Zambia	1	37,882	88,890		
Ndola	Zambia	1	329,228	468,324		
Bulawayo	Zimbabwe	1	611,307	653,337	1800	
Chinhoyi	Zimbabwe	1	41,969	68,273	1850	
Gweru	Zimbabwe	1	125,626	154,825	1850	
Harare	Zimbabwe	1	1,405,753	2,133,801	1850	
Hwange	Zimbabwe	1	44,297	19,870		
Kadoma	Zimbabwe	1	66,150	91,633		
Kwekwe	Zimbabwe	1	101,681	136,804	1850	
Marondera	Zimbabwe	1	37,277	61,998	1850	
Masvingo	Zimbabwe	1	48,780	87,886	1850	
Mutare	Zimbabwe	1	124,697	186,208	1850	

Continued on next page

City name	Country	Anglophone dummy	Projected population 1990	Projected population 2012	Colonial origin sample	40 cities sample
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	0.173***	0.285***
	(0.055)	(0.093)
Ln ruggedness	-0.062*	-0.172**
	(0.034)	(0.080)
Ln rainfall	-0.150***	-0.912***
	(0.049)	(0.119)
Ln elevation range	0.137***	0.358***
	(0.038)	(0.077)
Coast dummy	0.776	-0.234
	(1.184)	(3.173)
Interaction ln coast length	0.027	0.007
	(0.107)	(0.273)
Interaction ln distance to coast	0.098***	0.064
	(0.037)	(0.086)
Fraction of river area	-0.142	-0.263
	(0.679)	(1.030)
Fraction of lake area	-0.664	0.831
	(0.970)	(1.174)
Malaria index	-0.002	0.003
	(0.004)	(0.006)
Land suitability	0.235**	0.267
	(0.094)	(0.179)
Log temprature	-0.270	-1.458***
	(0.273)	(0.537)
Non-capital dummy	0.186**	0.038
	(0.090)	(0.164)
Distance to national capital	-0.000**	-0.001***
	(0.000)	(0.000)
ELF level 15 1975	-0.038	0.117
	(0.061)	(0.128)
Lag t-1 ln country GDP per capita	0.082	0.071
	(0.061)	(0.078)
Ln annual population growth 90 to 12	-0.522	8.453***
	(1.217)	(2.241)
Ln projected city population 1990	-0.174***	0.857***
	(0.025)	(0.042)
R^2	0.341	0.786
N	281	281

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

Table A3: Built-up cover intensity gradient 2014

	(1) Anglphone Cities	(2) Francophone Cities
Ring distance	-0.195***	-0.285***
	(0.011)	(0.035)
Ln ring total pixel	0.646***	0.664***
	(0.054)	(0.104)
Lag t-1 ln country GDP per capita	-0.039	-0.039
	(0.174)	(0.308)
Ln projected city population 1990	0.933***	1.194***
	(0.098)	(0.201)
Ln annual population growth 90 to 12	1.011	-4.779
	(4.467)	(6.302)
Ln ruggedness	0.001	0.557***
	(0.180)	(0.180)
Ln rainfall	0.521***	-0.542**
	(0.198)	(0.210)
Ln elevation range	0.038	-0.303
	(0.169)	(0.185)
Coast dummy	-0.331	-1.825
	(2.587)	(3.657)
Interaction ln coast length	0.011	-0.460
	(0.212)	(0.349)
Interaction ln distance to coast	-0.005	-0.655***
	(0.175)	(0.185)
Fraction of river area	-0.454	-3.752*
	(4.063)	(1.980)
Fraction of lake area	3.398	-1.006
	(4.174)	(5.017)
Malaria index	0.030	0.061***
	(0.019)	(0.015)
Land suitability	-0.975**	1.088**
	(0.454)	(0.414)
Log temprature	-1.932**	-5.473**
	(0.979)	(2.370)
Non-capital dummy	-0.740**	-0.558
	(0.312)	(0.516)
Distance to national capital	-0.001**	0.001
	(0.000)	(0.001)
ELF15 1975	0.161	-0.142
	(0.285)	(0.363)
R^2	0.479	0.551
N	2475	703

Note: Sample includes rings up to 20km. Standard errors are robust and clustered at city level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table A4: Intensity of built pixels 2014

	(1) 1km	(2) 2km	(3) 3km	(4) 4km	(5) 5km	(6) 6km
Anglophone Dummy	-0.109*	-0.215***	-0.403***	-0.283**	-0.312**	-0.139
	(0.061)	(0.057)	(0.095)	(0.134)	(0.147)	(0.175)
Ln ring built pixel	1.489***	1.333***	1.361***	1.362***	1.324***	1.310***
	(0.057)	(0.045)	(0.054)	(0.056)	(0.042)	(0.048)
Anglophone mean	1.196	2.514	2.539	2.425	2.229	2.299
Francophone mean	1.463	3.153	3.024	2.569	2.586	2.501
Geographic and Stiational Controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.907	0.930	0.920	0.920	0.929	0.915
N	283	285	279	259	236	209

Note: Standard errors are robust.

* 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	0.539***	0.325***	-0.013
	(0.167)	(0.118)	(0.061)
Ln initial cover 1990	0.295***	-0.295***	0.022
	(0.058)	(0.040)	(0.020)
Year dummy 2014	0.495***	0.119**	0.137***
	(0.067)	(0.056)	(0.035)
Ln ruggedness	-0.190*	-0.204***	-0.034
	(0.107)	(0.076)	(0.032)
Ln rainfall	-0.279**	-0.053	-0.007
	(0.137)	(0.087)	(0.048)
Ln elevation range	0.223**	0.069	0.100***
	(0.110)	(0.082)	(0.036)
Coast dummy	-3.614	-3.707**	-1.156
	(2.854)	(1.572)	(0.980)
Interaction ln coast length	0.539**	0.409***	0.142*
	(0.221)	(0.122)	(0.083)
Interaction ln distance to coast	0.245**	0.113	0.046
	(0.116)	(0.080)	(0.037)
Fraction of river area	-0.365	-0.270	-0.929
	(1.716)	(1.436)	(0.786)
Fraction of lake area	-1.057	-1.455	-0.466
	(1.921)	(2.022)	(0.740)
Malaria index	0.030***	0.013*	-0.001
	(0.010)	(0.008)	(0.004)
Land suitability	0.375	-0.005	-0.112
	(0.291)	(0.193)	(0.106)
Log temprature	-3.506***	-1.736***	-0.040
	(0.795)	(0.542)	(0.282)
Non-capital dummy	0.012	-0.152	-0.044
	(0.203)	(0.148)	(0.091)
Distance to national capital	-0.001***	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)
ELF level 15 1975	-0.280	-0.159	0.004
	(0.184)	(0.131)	(0.068)
Lag t-1 ln country GDP per capita	-0.063	-0.087	0.059
	(0.144)	(0.117)	(0.050)
Ln annual population growth 90 to 12	12.114***	5.842**	2.215*
	(3.338)	(2.585)	(1.287)
Ln projected city population 1990	0.712***	0.443***	0.036
	(0.101)	(0.072)	(0.035)
R^2	0.602	0.230	0.109
N	551	551	525

Note: Standard errors are robust and 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	0.661***	0.472***	0.577***	0.283***
	(0.152)	(0.132)	(0.177)	(0.100)
Ln initial cover 1990	0.299***	0.298***	0.325***	-0.272***
	(0.063)	(0.060)	(0.065)	(0.039)
Year dummy 2014	0.541***	0.502***	0.515***	0.135**
	(0.107)	(0.066)	(0.071)	(0.055)
Ln ruggedness	-0.136	-0.130	-0.203*	-0.129**
	(0.114)	(0.094)	(0.115)	(0.066)
Ln rainfall	-0.271**	-0.233**	-0.283*	-0.026
	(0.119)	(0.118)	(0.144)	(0.075)
Ln elevation range	0.180**	0.246***	0.222*	0.069
	(0.091)	(0.091)	(0.116)	(0.066)
Coast dummy	-7.481***	-5.206*	-3.788	-4.731***
	(2.463)	(2.924)	(3.073)	(1.703)
Interaction ln coast length	0.875***	0.571**	0.582**	0.415***
	(0.207)	(0.234)	(0.242)	(0.140)
Interaction ln distance to coast	0.248**	0.140	0.268**	0.026
	(0.100)	(0.100)	(0.125)	(0.069)
Fraction of river area	-5.020**	0.898	-0.410	1.007
	(2.476)	(1.467)	(1.932)	(1.229)
Fraction of lake area	1.460	-1.520	-1.064	-2.043
	(1.380)	(1.858)	(2.064)	(2.099)
Malaria index	0.040***	0.022**	0.034***	0.006
	(0.009)	(0.009)	(0.011)	(0.006)
Land suitability	0.130	0.115	0.423	-0.165
	(0.250)	(0.256)	(0.304)	(0.172)
Log temprature	-2.680***	-2.877***	-3.793***	-1.189***
	(0.728)	(0.732)	(0.862)	(0.421)
Non-capital dummy	-0.120	0.001	0.044	-0.168
	(0.224)	(0.198)	(0.208)	(0.143)
Distance to national capital	-0.001***	-0.001***	-0.001***	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
ELF level 15 1975	0.026	-0.299*	-0.276	-0.168
	(0.176)	(0.163)	(0.196)	(0.114)
Lag t-1 ln country GDP per capita	0.064	-0.057	-0.066	-0.049
	(0.076)	(0.099)	(0.153)	(0.084)
Ln annual population growth 90 to 12	5.178	12.764***	12.200***	6.481***
	(4.617)	(3.031)	(3.487)	(2.377)
Ln projected city population 1990	0.532***	0.628***	0.705***	0.365***
	(0.094)	(0.094)	(0.106)	(0.065)
R^2		0.633		0.231
N	551	525	551	525

Note: Columns 2 and 4 show OLS results excluding cities with zero LF patches. Standard errors are robust and 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	-0.282	-0.500	-0.119
	(0.187)	(0.311)	(0.369)
Ln projected city population 1990	0.115	-0.104	0.038
	(0.116)	(0.253)	(0.282)
Ln annual population growth 90 to 12	0.003	0.004	0.001
	(0.002)	(0.004)	(0.004)
Ln ruggedness	0.581***	0.418	0.011
	(0.156)	(0.280)	(0.137)
Ln rainfall	0.037	-0.011	0.129***
	(0.081)	(0.137)	(0.041)
Ln elevation range	0.290***	0.019	0.022
	(0.084)	(0.190)	(0.174)
Coast dummy	-0.041	0.042	0.100
	(0.028)	(0.065)	(0.089)
Fraction of river area	-0.007**	-0.006	-0.004
	(0.003)	(0.007)	(0.004)
Fraction of lake area	0.001	0.001	0.001
	(0.003)	(0.002)	(0.002)
Ln coast length	-0.428	0.490	1.116
	(0.292)	(0.711)	(0.971)
City cluster FE	No	No	Yes
Observations	318	58	58

Note: In each column, each row reports coefficients from a regressions with the row header as outcome and Anglophone dummy as independent variable. Border sample does not include cities in Anglophone and Francophone Cameroon border area. Standard errors are robust.

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

Table A8: Openness and area: Robustness

	(1) Base	(2) Distance trim	(3) No German colonies	(4) No Nigeria	(5) Colonial origin 1885	(6) Non capital	(7) 40 cities
Ln openness index 2014							
Anglophone dummy	0.173***	0.159***	0.188**	0.156**	0.332***	0.166***	0.302**
	(0.055)	(0.057)	(0.077)	(0.069)	(0.078)	(0.063)	(0.134)
Ln area							
Anglophone dummy	0.285***	0.278***	0.227*	0.551***	0.104	0.260***	0.352
	(0.093)	(0.098)	(0.120)	(0.120)	(0.318)	(0.098)	(0.301)
R^2	0.786	0.717	0.790	0.821	0.693	0.727	0.481
N	281	266	248	172	54	261	26

Note: Columns 1-4 and 6 include the same controls as in column 4 and 8 of Table 1. Columns 5 and 7 include the controls of ln country GDP per capita in 2000, ln projected city population in 1990, ln average ruggedness, and coast dummy. Columns 5 includes 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 includes ln annual population growth 90 to 12 in city level. Robust standard errors are applied in all columns. Adjusted R^2 and N are reported for ln area.

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