Received: 28 November 2022 | Revised: 27 February 2024 | Accepted: 10 March 2024

DOI: 10.1111/jors.12699

RESEARCH ARTICLE

JOURNAL OF REGIONAL SCIENCE WILEY

Within-city roads and urban growth

P. Brandily[1] | F. Rauch[2]

[1]London School of Economics, London, UK

[2]Heidelberg University, Heidelberg, Germany

Correspondence
F. Rauch, Heidelberg University, Heidelberg, Germany.
Email: ferdinand.rauch@awi.uni-heidelberg.de

Funding information
Department for International Development, UK Government; World Bank Group

Abstract

In this paper we study the role of within-city roads layout in fostering city growth. Within-city roads networks have not been studied extensively in economics although they are essential to facilitate human interactions, which are at the core of agglomeration economies. We build and compute several simple measures of roads network and construct a sample of over 1800 cities and towns from Sub-Saharan Africa. Using a simple econometric model and two instrumental variable strategies based on the history of African cities, we then estimate the causal impact of within-city roads layout on urban growth. We find that over the recent decades, cities with greater road density and road evenness in the centre grew faster.

KEYWORDS

road layout, Sub-Saharan Africa, urban planning, urbanisation

1 | INTRODUCTION

Cities enhance productivity and consumption benefits through various mechanisms described in a large literature in economics. Within this tradition, all mechanisms proposed for agglomeration effects relate to one essential feature of the urban environment: it facilitates short-distance interactions between economic agents. This feature is at the core of the main theories on agglomeration forces, regardless of whether they describe production or consumption. In his seminal typology Marshall (1920) suggested knowledge spillovers, labour market externalities and production linkages as drivers of urban productivity; more recently, Duranton and Puga (2004) highlight the importance of

---

We thank Julia Bird, Steven Brakman, Paul Collier, Victor Couture, Harry Garretsen, Vernon Henderson, Richard Hornbeck, Guy Michaels, Matt Turner, Tony Venables, a number of anonymous referees and various seminar participants for helpful comments. We gratefully acknowledge the generous support of a Global Research Program on Spatial Development of Cities, funded by the Multi Donor Trust Fund on Sustainable Urbanization of the World Bank and supported by the UK Department for International Development.

sharing, matching, and learning. In this paper we investigate the role of roads and streets network within cities. We thus focus on a central component of the urban setting that could directly promote or discourage interactions between agents. More specifically, we study the impact of the density and shape of the road grid within African cities on population growth in recent decades.

To study these specific urban features, we measure various statistics on the road networks within urban centres for a large number of African cities and show the association between some of the layout's features and urban growth. Combining satellite imagery and Open Street Map, an open-source geographic data set mapping roads and streets all over the world, we build simple statistics of the road layouts at a precise geographic scale. We believe our focus on within-city layouts to be a conceptual innovation to the economic literature where the topic has remained largely unstudied so far. We show that measures of the central road grid, in particular its density and evenness, are associated with recent population growth.

We focus on the link between these measures of the within-cities roads layout and urbanisation specifically in the context of Sub-Saharan Africa. We chose this focus for different reasons. First, this is the region of the world that is currently experiencing the fastest rate of urbanisation (Bryan et al., 2020). Understanding the role of this urban infrastructure in the capacity of cities to generate growth better could help designing policies that benefit the billions of people that live or are expected to live in African cities in the next decades. Second, the road layout is more meaningful in an environment that relies heavily on road transportation, and currently does not have an extensive urban train or underground network. Third, it is costly to change a road layout once it is in place. Planning ahead is therefore critical in Africa's urbanisation, which is characterised by low-income levels and where ‘two-thirds of the cities are to be built’ (Bryan et al., 2020). Fourth, various studies report high level of traffic congestion in African cities and the need for better road investments. Hidden external costs of congestion were estimated at up to 5% of cities’ GDPs in Dakar or Abidjan, which is more than twice the estimates for European cities (Cervero, 2013) while ‘in a sample of 30 cities around the world, the 8 African cities rank in bottom 12 spots for road density’ (Lall et al., 2017). Finally, some argue in favour of a ‘grid structure, which (...) enhanced travel efficiency’ (as stated in the last major report from the World Bank on Africa's Cities - idem), although to the best of our knowledge there is little empirical research in economics so far to assess this hypothesis. Our focus on within-city effects complements a wider literature that finds important positive effects of transportation infrastructure connecting cities (recently reviewed e.g. by Cui et al., 2023 and Cao et al., 2023).

To estimate the impact of road layout we rely on two observations: the road layout is both difficult to change and dependent of the context and available transportation technologies at the time of its construction. The persistence of road layout over time is exemplified in cities of the ancient world (like Jerusalem or Paris, among many others, where the Roman's cardo are still important streets today) or the new (as New York where the ‘Commissioners’ Plan’ of 1811 still structures Manhattan today). More recently, this was also one conclusion from a large-scale analysis of cities around the world by Barrington-Leigh and Millard-Ball (2019). In the case of Africa in particular, cities are often much newer. Baruah et al. (2021) provide empirical evidence of layout persistence in African cities. Together, all these observations imply that due to historic reasons some cities may be accidentally stuck with better or worse road grids, which would benefit or hurt their growth potential. If a city by historic accident develops a good road layout it will be rewarded by long lasting population growth, as people move in to benefit from the agglomeration benefits generated in the city centre. If on the other hand a city is stuck with a suboptimal grid, it may reach a peak population level beyond which it finds it difficult to grow any further.

Our main measure of the quality of the road network is a simple measure of road density in the centre of towns and cities, thereby testing whether a lack of road quantity constrains urban growth. We also compute a second and secondary simple measure of the distance to the nearest road for a random point in the city centre, as well as a number of measures of regularity and orientation of the road grid. To increase comparability across different cities we keep the area of city centre for which we compute these measures constant across cities. All measures we use in this paper have the advantage of being intuitive, straightforward to compute and simple to interpret. Our conclusion focuses on the linear effect of our measures of road density and spatial distribution, which are the most robust and quantitatively most important of all the network features we consider. To address measurement error and potential endogeneity, we rely on two instrumental variable strategies based on the age of a city. We accept that the data available to us are limited. In particular, we cannot observe historic road grids for a large panel of cities in a way that is comparable. Thus, we rely on proxy data. Despite the unavoidable limitations in this setting, we think that our proximate measures are still good enough for this analysis to uncover a real effect. Our IV approaches also suffer from weak first stages in some specifications. Yet we think that the consistency of results across OLS and various IV specifications, despite their occasional weaknesses (which we show for completeness), overall points to suggestive evidence concerning the importance of roads in the city centre.

While there is a substantial literature on the effect of roads or other forms of transportation on urban shape (recent examples include Baum-Snow, 2007; Duranton & Turner, 2012; Faber, 2014; Jedwab & Moradi, 2016; Jedwab & Storeygard, 2016; Michaels, 2008 and Storeygard, 2016), this literature typically has concentrated on roads connecting cities with one another. Less is known on roads and road layout within cities (exceptions include Baum-Snow et al., 2017, who consider the effect of urban highways in China, Akbar et al., 2023 who study mobility in Indian cities and Couture et al., 2018 who suggest that restricting traffic would bring welfare gains). Our approach is more focused on within-city variation and includes smaller roads than these existing studies. As we mentioned, the importance of road grid for urban development has long been recognised. Fuller and Romer (2014) write: ‘getting the grid right will likely be more important than enforcing building codes on structures or imposing limits on density’. Yet, to our knowledge, there is a gap in research quantitatively assessing the importance of cities’ road layout in economics. The topic has of course been studied in urban planning or environmental studies. Layout shapes have been shown to be associated with transport decisions, travel time and local CO2 emissions. Notably, in a recent series of articles, Barrington-Leigh and Millard-Ball (2019 and, 2020) also use Open Street Map to provide global-level measures of within-city roads networks and confirm results about transport decisions. Their focus is not on agglomeration economies nor do they specifically study the Sub-Saharan African context. While this region has known a rapid urbanisation in recent decades and will most likely continue to do so, research is still needed on this context. In a literature review on developing-world cities, Bryan et al. (2020) insist on the need for more research on urban mobility in African cities.

Overall, our results suggest that denser road networks are associated with more population growth. Our interpretation of this empirical relationship is that better connected city-centre foster interactions, agglomeration economies and eventually the growth of cities. We also find evidence that, given a specific density of road, the even distribution of roads across space in the centre is also a significant determinant of growth, and more so than the grid-orientation or the number of nodes in the network. The type of features we analyse here is associated with an increased number of destinations that can be reached within a given time frame. We take away from our results that such an increase in connectivity is important for the growth of cities. Although we acknowledge limits to our identification strategies, we hope that the consistency of our results is indeed indicative of a causal relationship between road network and city performance (an identification mostly absent from the literature outside of economics). By these findings, we also provide evidence for a hypothesis advanced by Collier and Venables (2016) that the road network is an important predictor of recent population growth, and that many cities and towns in Sub-Saharan Africa are constrained in their relative growth due to lack of road density.

The paper is organised as follows: Section 2 develops our hypothesis in greater detail, discusses how it connects with the literature and presents our main measures of roads network. Section 3 describes the construction of the datasets used and presents descriptive statistics. Section 4 develops the empirical strategy and discusses the instrumental variable strategies we use. Section 5 presents the main results. We present a series of robustness checks in Section 6. Section 7 concludes.

2 | HYPOTHESIS AND MAIN MEASURES

Our research question, our hypothesis and our choice of instrumental variable are shaped by connecting several findings from the urban economics literature. First, we assume that the returns from agglomeration are largely generated in the city centre. This is one of the standard assumptions used for example in monocentric city models (following Mills, 1981), central place theory (see e.g. Christaller, 1933) as well as other theories on the economics of cities. This suggests that people live around a city centre or central business district, to which they travel to work and enjoy leisure time. The agglomeration returns are overwhelmingly generated in that centre. Consistent with this literature, we pay special attention to the road grid in the centre of cities rather than the whole road grid structure over the entire urban agglomeration. Although we also study the city as a whole for comparison in our robustness section, our preferred measure captures only the centre. Empirical evidence for the monocentric model has been found for a variety of cities, by demonstrating gradients of wages and house prices, as well as from commuter flows (for recent evidence for Africa, see Antos et al., 2016 or Larcom et al., 2017). If anything, monocentricity has been found to be more pronounced in Africa than in the developed world (Cervero, 2013; Grover Goswami & Lall, 2016). Further down in this draft we also provide empirical evidence that indeed an additional square metre of road built in the city centre has a stronger effect on population growth than an additional square metre of road built anywhere in the wider city. To define a city centre we pick the centroid of the brightest pixel of nightlight of a city and draw a circle around it. In our main specification, this circle has a radius of 1 km. This approach has the additional advantage that we do not need to define the edge of a city and how it changes over time, and that having similar sized areas makes the measure more comparable across cities. This circle also ensures that we don't compare centres with the edges of cities. We do however adjust the central area of a city in the case where we find rivers, coasts, or a border (see Section 3.1 for more details).

Second, by defining the city centre as the centroid of the brightest spot of nightlight, we only have one single centre for each town or city. This again is a standard assumption made in monocentric city models. But it is also a restriction we impose to our data despite potential cases of multipolar cities. In the context of our study this restriction is likely to have a limited importance, especially given that we also include many small and medium sized towns, for which monocentricity is even more plausible (given they may not be large enough to support two centres). Antos et al. (2016) provide evidence that is consistent with monocentric centres for a few large African cities.

Third, we take from the literature that the location of the city centre does not change much over time, and that cities typically expand around historic city centres. This is true, for instance, if the foundation of a city followed locational fundamentals that do not change over time. Evidence can be seen in the many cities in the world that are still structured around Medieval or Ancient sites. A recent empirical paper making this assumption is Harari (2020), evidence for the survival of Roman towns in Europe for two centuries can be found in Michaels and Rauch (2018). There is also ample evidence for persistence of Roman roads (e.g. Botasso et al., 2022 or De Benedicts et al., 2023). Following this view, a city with a better working city centre will develop a larger commuter zone around it over time. This observation is important to inform our identification strategy since it allows us to assume that the historic centre and the modern centre coincide.

Fourth, we believe that the road network is of crucial importance to facilitate the agglomeration returns, whatever these returns may be. This could be violated if most of the agglomeration returns happen overwhelmingly within buildings and are less dependent on connection in the centre. Even in this case people would have to get to these buildings, but quantitatively the true agglomeration effects may be poorly approximated by our measures. It could be violated if some strict version of the ‘fundamental law of road congestion’ (Duranton and Turner, 2011) applies, whereby road construction does not relieve congestion and hence does not lower the transport time of human interactions. This fourth point is the main hypothesis of our paper, and we provide empirical evidence that indeed the road measures we use in this paper have predictive power for population growth.

Fifth, we take it as given that there is path dependence of the road layout. Once a layout has been established, it is costly to change it, and the grid from 1 year overlaps strongly with the grid from another year. While there are examples of radical changes in the layout of city maps (Haussmann's renovation of Paris in the 19th century is a famous example), the great majority of cases that come to mind suggest indeed a tendency of lock-in.[1] Using a large data extraction from Open Street Map and covering a large share of world's cities, Barrington-Leigh and Millard-Ball (2019) study within cities roads layout and notably conclude that 'street-network sprawl is a path-dependent process. Because streets are one of the most permanently defining features of cities'. In economics, Baruah et al. (2021) compare colonial and modern street maps for some African cities. They find that 'the spatial structures of cities in Sub-Saharan Africa are strongly influenced by the type of colonial rule experienced' and they point to 'high costs of acquiring new rights of way in an already built-up city'. Their examples that contrast colonial and modern maps show strong persistence of road layout. Comparing 1960s maps of African cities to the same street data we use, they notably conclude that 'roads that were in place 55 years ago generally remain in place today'. Shertzer et al. (2016) argue that Chicago zoning ordinances from 1923 have a larger effect on the spatial distribution of economic activity in the city today than geography or transport networks. If this point was not true, cities would be able to self-correct, our first stages would not be statistically significant, and our measures of the road grid would be poor predictors of future population growth.

Sixth, we consider it likely that the optimal road grid at a given time depends on the available transport technology. The ideal grid in the age of walking and the age of the horse differs from the ideal grid in the age of the car. For example, the latter might require broader roads and might avoid crossings of roads harder. If so, the age of a city influences its initial grid, which in turn influences its grid later on. This observation helps us to develop our instrumental variable, which uses a correlate of the age of a city as an instrument for the initial road grid specification.

Taken together, these six observations imply that initial differences of road layout, even if from a very distant past, can lead to different long run population developments. Due to small differences in geography and local history, some towns may accidentally stumble upon a more successful layout than others. These initial differences depend to some degree on the age of a city, which is observable.

These points do not tell us what a successful layout should look like. We use the Open Street Map data set to measure several features of within-city roads like total length, number of nodes and intersections, orientation of each segment of road (bearing), etc. We then use this measurement either directly or in order to compute relevant characteristics of the layout. In particular, we build a simple measure of density of roads in the city, measured by road kilometres over area considered (which usually is the city centre only). When computing this density variable, we count a road recorded in the open street map data with n lanes n times. Such a simple count of the density of roads is the most direct way to measure the density of the transport network. On average, more roads imply fewer diversions from the shortest distance between two points, and more roads help ease congestion. There could also be too many roads, which we don't believe to be likely in our sample. This strikes us as a simple and straightforward measure of road layout in cities. This measure also has the advantage of not taking any strong stand on the correct relationship between the observable features and the actual travel speed of residents. The latter not being obvious as having roads with multiple lanes may be helpful in bigger cities, but wasteful in smaller ones; crossings may slow traffic down, but enable useful travel combinations, etc. Since part of this question is also empirical, we will study the impact of directly-measured features like the number of intersections on city growth (either controlling for total road length or not).

What we try to capture is how well-connected people and firms in the city centre are. Road density might be misleading if all the roads are concentrated in a small subset of the city. In an extreme example, a city might have a road-heavy parking lot in the centre, but few roads otherwise. Such a city would appear as having a high road density by our density measure, but would not enable fast and straight forward connections. To address the concentration of roads additionally, we compute a variable measuring the concentration of roads. We build the

---

[1] Examples of large and successful cities whose central road layouts still reflect their Ancient or Medieval past abound in Europe. For evidence of path dependence for cities see Michaels and Rauch (2017). For evidence of path dependence of infrastructure in Africa see Kocornik-Mina et al. (2020).

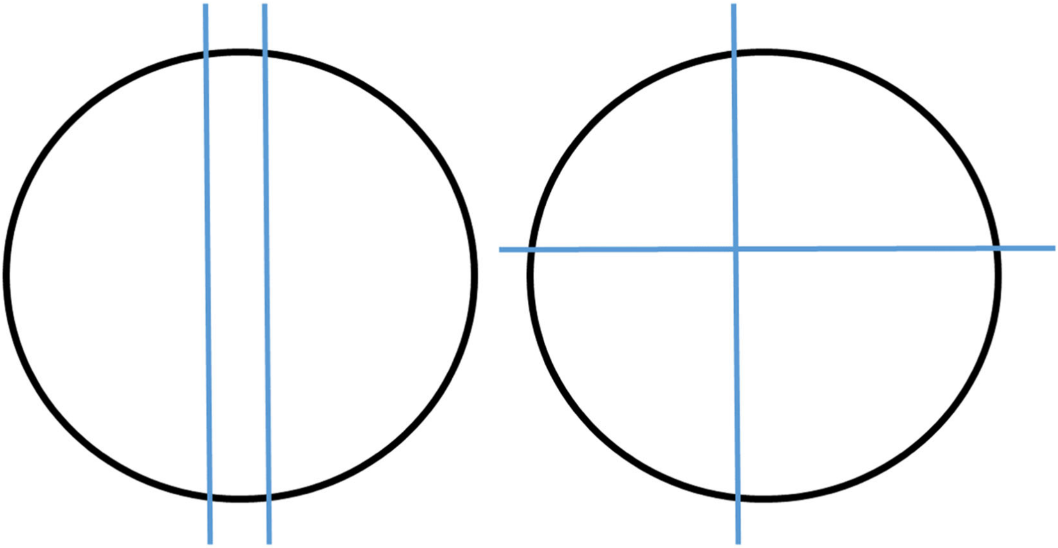

FIGURE 1 This figure presents two cities with stylised road layouts. Both cities have the same road density in the centre by construction. The city on the right has a shorter distance from a random point to the nearest road, a property we call 'evenness'. [Color figure can be viewed at wileyonlinelibrary.com]

evenness approach as a measure of how easy it is to access road network from a random point in the city centre. This variable captures how regularly roads are spread over the area of measurement. Facing a similar problem, Donaldson and Hornbeck (2016) create 200 random points within US counties, calculate the distance from each point to the nearest railroad, and take the average of these nearest distances. Following this algorithm, we create 12 points in each circle, measure the distance to the nearest road for each of these points, and compute the average nearest distance to the road from these 12 points.[2] Because this measure decreases as the roads are less evenly distributed across space, we call it an 'unevenness index'. To illustrate this variable, consider Figure 1. This figure presents two cities with stylised road layouts. Both cities have the same road density in the centre by construction. The city on the right has a shorter distance from a random point to the nearest road.

In addition to these two main variables of density and unevenness, we compute a third variable measuring the orientation of roads. Using the distribution of the (segments of) streets' bearing we compute an Herfindahl index of concentration. This simple and transparent measure takes its highest values when all streets run in either one of two directions, and two only (like a perfect grid does). We call this measure the 'orientation Herfindahl index'. We describe further the construction of all measures in 3.3.

3 | DATA

We build a complete and consistent sample of cities, combining satellite imagery and geospatial population data in addition to other data sources. For each of these cities, we observe centre's road layout and measure a set of descriptive statistics. This section briefly describes datasets and measurement, all sources and processes are detailed further in the data appendices A and B.

---

[2] Using 12 points is a low number in comparison to 200. We do so for computational convenience. As a justification, we also record the mean distance from the centre of the circle to the nearest road, which is effectively the same exercise with that parameter set to one. The correlation between the exercise of 1 and 12 points is around 0.9. Given this large correlation, we believe that all such measures would give fairly similar numeric estimates.

3.1 | Defining cities: boundaries and centres

Our main source for the measurement of cities location and boundaries is satellite imagery of luminosity at night. This source has the double advantage of being complete (i.e. it covers the entire continent) and consistent (i.e. all the cities are defined in the same way). First, we identify lit areas from NOAA's DMSP-OLS satellite record (cf. Supporting information: Appendix A) by keeping pixels that emit light at least twice over the 5 yearly observations from 2008 to 2012. Contiguous lit areas are then aggregated using a GIS software to create unions. Each union represents a city, and we consider its footprint as the city boundaries. Figure 2 illustrates our sample. As the Figure shows, we consider cities in almost all countries of Sub-Saharan Africa, with healthy spatial variation. As a

FIGURE 2 We identify 2,779 cities from 40 countries. [Color figure can be viewed at wileyonlinelibrary.com]

robustness check, we also rely on the 'Urban Centre Database' (UCDB) from the EU Commission JRC to delimit a second sample of African cities. We discuss the definition of this sample in Supporting information: (Appendix D).

Our focus is on city centres, where we assume agglomeration effects to take place. Because, to the best of our knowledge, there is no data set providing precise coordinates of cities' centre we rather follow a systematic data driven definition. For each city, we take as 'centre' the centroid of the brightest area (i.e., where pixels have the highest light value observed in the city). In practice, this brightest area sometimes is a single light pixel, but sometimes is a larger area of contiguous pixels, in which case we compute its centroid. In cases of multiple brightest areas of equal brightness in a city we choose the largest one. This method identifies a unique city centre for each city.

Two reasons at least support this definition of a city centre: first, because nightlights glow, it is often the case that brightest area is close to the geographic centroid of the city. Second, at the same time, nightlights are also known to be linked with economic activity, even at the local level. Figure 3 illustrates the derived city centres for two small towns. The figure also shows the city extent, computed from lights data, and the road network of these towns. Importantly, we chose to use early images of nightlights (1994-1996) to define city centre. This is for both empirical and theoretical reasons. Empirically, African cities emitted less light in the early period which facilitates the identification of the main historical centre. Theoretically we are interested in the lock-in of historical centre and therefore want to identify where this centre was as early as possible. This also reduces the sample to cities that were important enough in the mid-1990 to emit some light. This selection mainly excludes very small towns and is coherent with our instrumental variable strategies (cf. Section 4.2). In Supporting information: (Appendix B.1) we discuss alternative measures of the city centre and show that they yield very similar outcomes.

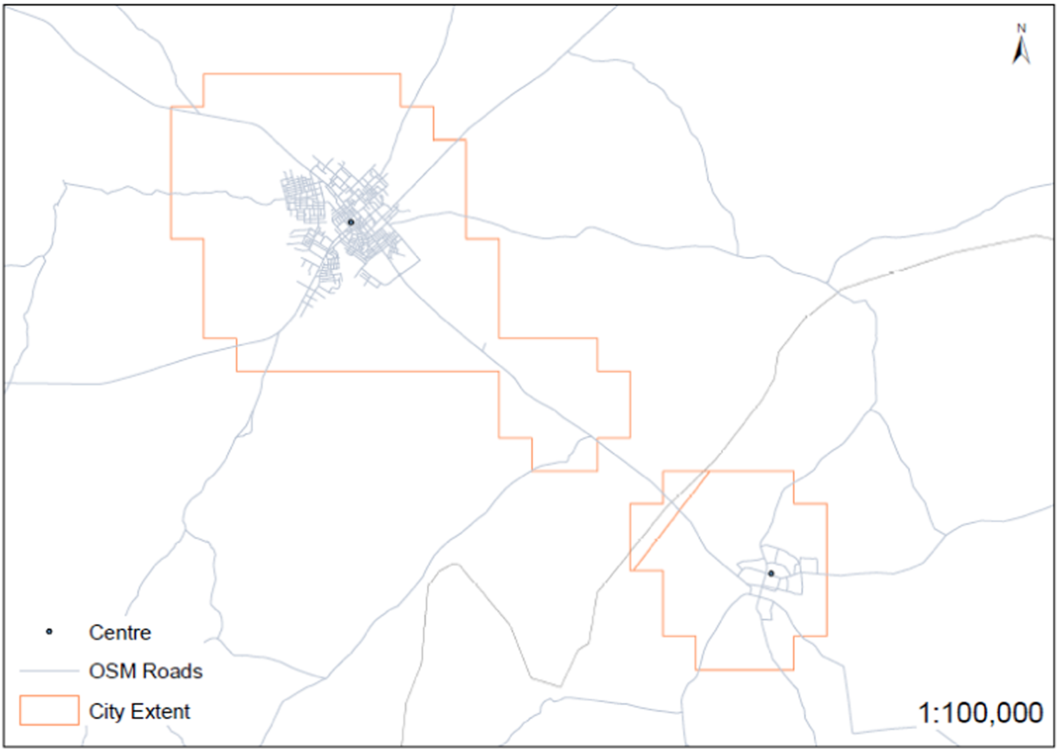

FIGURE 3 Example of city extent definition. City extent and centre are defined using nightlight luminosity (in late—2008–2012—and early—1994–1996—period, respectively). [Color figure can be viewed at wileyonlinelibrary.com]

3.2 | City growth

Our main outcome of interest is the population growth of cities. We measure this outcome relying on the GHS data set provided by the European Commission. The GHS is a spatial grid that provides a population estimate for each cell (of 250 by 250 m resolution) in four target years: 1975, 1990, 2000 and 2015. The population estimates are obtained from the spatial distribution of census data, guided by GHS modelling of built-up presence (derived from Landsat day-time satellite images). For each city in our sample, we sum up all the GHS cells falling within its boundaries and extract the population counts for the four target years. In Supporting information: (Appendix B.3), we discuss this measurement further. Although our focus is on city growth from 2000 to 2015, where we think the data quality is best, we also present extensive robustness checks using alternative outcomes and data sources.

3.3 | Centre and road layout

We construct city-centres' roads layout using the publicly available data from Open Street Map (OSM), a volunteered geographic information project consisting of information from over 2 million users. We downloaded the data in December 2016. As with any user generated data, there is some concern that the data set may be badly measured or biased. In their quantitative evaluation of OSM quality, Barrington-Leigh and Millard-Ball (2017) however conclude that the completeness of OSM was already high as of 2016, including for Sub-Saharan Africa. In Supporting information: (Appendix B.2.1) we discuss this point in more details and present the robustness of our results using only highest OSM-quality African countries. Our main conclusions hold even in this smaller sample. Seidel (2023) also studies OSM bias in Africa by comparing OSM maps with information from other sources, primarily for hospitals. His findings imply that our standard specification with country fixed effects and additionally controls for the initial size of a city (which we always include) would address this bias to some degree. A specification with (within-country) province fixed effects would absorb most of the bias. We provide evidence in this paper that suggests our main coefficients are robust to both these versions.

For each city we gather data on the road network at two different levels. At the city level, we take into account any road that falls within a city boundary; at the centre level, we draw a 1km-radius circle around the centre point and keep only the (segments of) roads that fall within. Figure 4 gives an example of a small city and highlights the derived circle of the defined central part. At both city and centre level, we measure a series of descriptive statistics about the road network. In particular we compute total length of roads, number of nodes[3] and intersections, as well as the density of each of these items (i.e. dividing them by the considered area).[4]

As previously mentioned, we also compute two indices of the network spatial organisation: we call the first index 'unevenness' and the second 'orientation Herfindahl'. For the unevenness index, we first create 12 points symmetrically distributed around the city centre and within each city centre circle and compute the average nearest distance to the road from these 12 points. Figure 5 illustrates the construction of this evenness measure, plotting the 12 points over a city centre (also showing the city and centre road networks). The index is thus expressed in a distance unit. The lower the index (i.e. the distance) the more likely the road network is evenly distributed over space. Second, we follow G. Boeing (2018) in constructing an orientation measure we then summarise in an Herfindahl index. To do so, we first compute the bearing of each edge (i.e. segment of road between two nodes)

---

[3] Intersections are nodes with strictly more than one street emanating, thereby excluding cul-de-sacs, Boeing (2017).

[4] Obviously, when considering only 1 km radius city-centres, road density is just a multiple of total road length. There are however two cases where density is, to our view, a better measure: first, some 1 km centre radius overlap water bodies or foreign countries that we exclude from the denominator; second, we also study city-wide measures and want to take city size into account in that case.

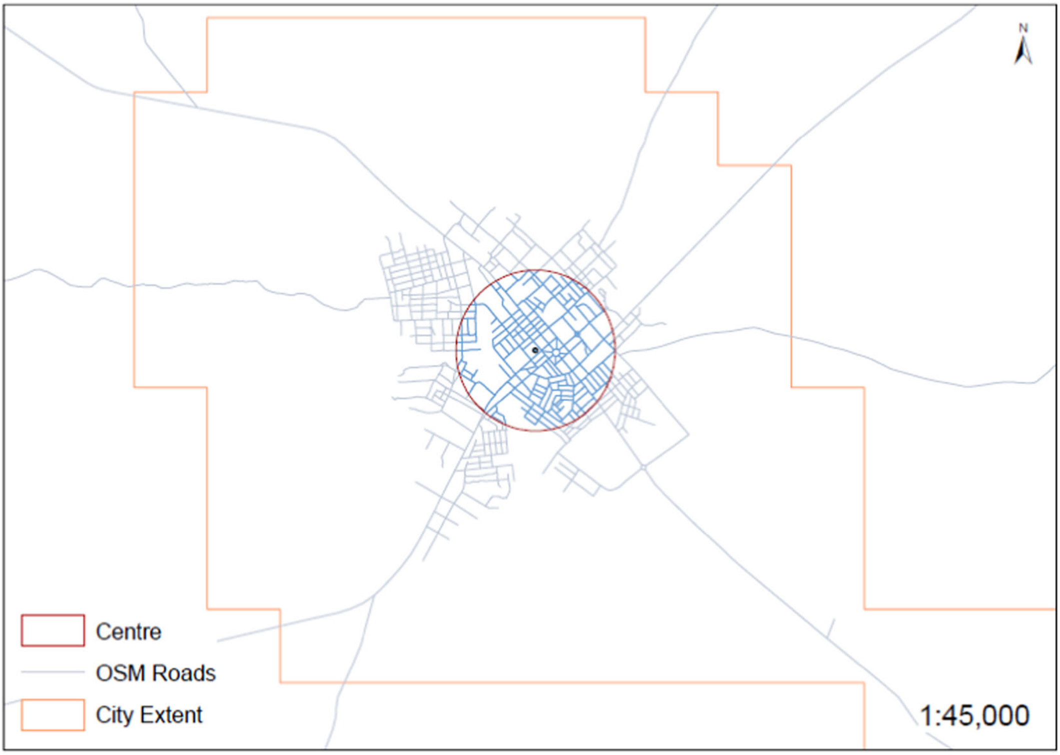

FIGURE 4 Example of road grid selection. Our measure of road layout is based on OSM roads in a 1km-radius circle around the city centre. [Color figure can be viewed at wileyonlinelibrary.com]

within a city centre and then observe the distribution of all edges across equal-sized bins of compass orientation (each bin covers 10 degrees). We then summarise the concentration of roads among potential orientations by taking the Herfindahl index of this edge bearing distribution. Figure 6 illustrates the construction of this orientation Herfindahl. The left two panels show the street grid in two city centres (Umm Ruwaba from Sudan at the top, and Mohale's Hoek from Lesotho at the bottom), defined as one kilometre radius circles around the centre point. The middle panels take the road segments from the first, and sort them by their orientation, weighting segments by their length. The final two panels take this information and display them as standard histograms. It is over these shares that we compute the Herfindahl indices of street orientation. In this context, the Herfindahl index is large if many roads run in parallel, and small if the grid is more chaotic. In practice, the highest values of the orientation Herfindahl correspond to an almost perfect grid (as the one displayed in the top panel of Figure 6). Both the evenness and the orientation Herfindahl are related, and their correlation in the sample is pretty strong (0.42). If one thinks of a perfectly dense and grid-like city-centre, both indices would most likely reach their minimum and maximum (for the unevenness and orientation Herfindahl, respectively). But the two indices still capture different things: one can imagine a city where most roads are concentrated in one single neighbourhood (i.e. not evenly distributed) yet run in parallel (i.e. close to a perfect grid), and conversely. The role of any of these network features on a city growth is an empirical question.

Ideally, we would want to measure the road layout at the beginning of the period over which we measure population growth, and not at the end. For a sample as large as ours this is however data we were unable to obtain. Instead, we use road maps that are closer to the end of the period as a proxy. As we discussed in the introduction, the road grid does not change fast over time, particularly in the city centre. This assertion is supported by

FIGURE 5 Evenness is computed from 12 within-centre points. [Color figure can be viewed at wileyonlinelibrary.com]

substantial evidence, including for African cities. We also note that our instrument corrects potential bias arising from this measurement error.

3.4 | Sample description

By following the steps described in subsections 3.1, 3.2 and 3.3 we obtain a total sample of 1,850 cities from 40 Sub-Saharan countries (cf. Supporting information: Appendix C.1). We exclude two countries from the main analysis. First, we exclude Nigeria because nightlights data suffer from potential issues in the Gulf of Guinea. Second, we exclude Madagascar as its own history is incompatible with one of our historical IVs. In the robustness section we show that results remain robust when these restrictions are not applied, and when further restrictions are made. In Supporting information: (Appendix C) we discuss all the data constraints and decisions that define our sample.

Table 1 provides a break-down of our samples by country. In 2000, the median city size is around 33,000 inhabitants, varying greatly across countries (from 2500 in Djibouti to above 100,000 in Chad and South Sudan, where we identify only 12 and 3 cities, respectively). The average population growth rate over the 2000-2015 period is estimated to be 2.04%. And there again, countries experienced different paths with city population growth spanning between 5.66% a year in Eritrea to negative growth in Swaziland (−1.04%).

As for the road network of these cities, Table 2 provides simple descriptive statistics. Here again, we find a high variation in the total amount of road length within cities (from virtually 0 to 28,621 km) and even within centres

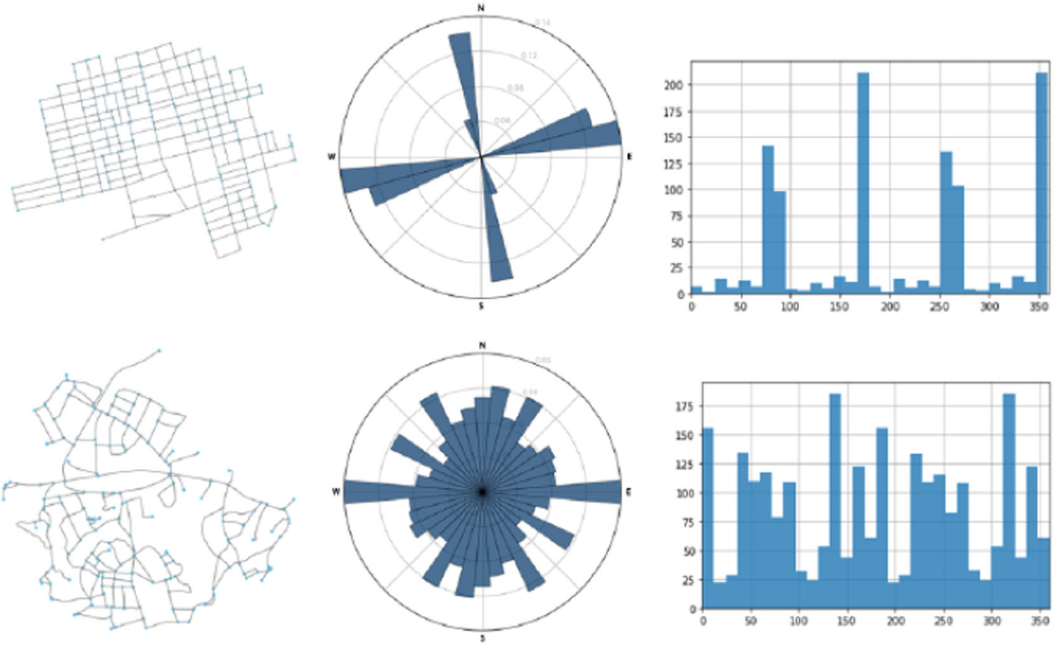

FIGURE 6 The edge bearings histograms (right) summarise the orientation of city centre edges (left). The central panel is a polar representation of the histogram, illustrating the construction of our measure. The first city (top) is Umm Ruwaba (Sudan) and has a typically high Herfindahl index (0.48). Below is Mohale's Hoek city (Lesotho), which is representative of the lowest values in our sample (0.05). [Color figure can be viewed at wileyonlinelibrary.com]

(0–87 km), defined as a circle with radius one. Although the maximum values seem high, they reflect that cities defined by nightlights can reach very large size, magnified by lights glowing. In our sample, the largest area encompasses Lagos as well as peripheral cities like Abeokuta in the north or Ijebu Ode in the east. Table 2 also reports statistics on the distances from cities' centre to the sea, and to the four cities (Dakar, Dijbouti, Cairo and Cape Town) we use for our IV specification.

4 | EMPIRICAL STRATEGY

To uncover the role of street networks in fostering city growth we first present a simple OLS regression model in Section 4.1. Because this model may suffer from standard limitations that limit causal interpretation, we also present and discuss two IV strategies in the following Section 4.2.

4.1 | Baseline model

The estimation equation we use for both sets of results is of the form:

$$\Delta pop = f(\mathit{Street\ network}) + \eta_c + \beta_1 \ln(pop_{t=1}) + \beta_2 \ln(\mathit{dist\ sea}) + \gamma X + \varepsilon. \quad (1)$$

We compute the left-hand side measure of annualised population growth as a log difference. In the main regression the 2 years for which we observe population levels are 2015 and 2000, and so the time difference is 14

TABLE 1 Summary statistics. In total, we consider 1850 cities from 40 countries. In our main specification we usually exclude Nigeria and Madagascar, reducing our sample to 1412 cities. The first two columns detail each country total number of cities and contribution to the sample, respectively. The following three column gives indication of population size in 2000 (median and average) as well as the average population growth rate in cities over the 2000–2015 period.

	Cities	Sample Share (%)	Median Pop.	Av. Pop.	Av. Pop. Growth
Angola	33	1.78	54,130	163,514	4.78
Benin	27	1.46	43,352	58,856	1.47
Botswana	29	1.57	18,382	32,626	0.67
Burkina Faso	27	1.46	44,998	62,813	1.94
Cameroon	69	3.73	34,104	93,522	1.01
Central African Republic	10	0.54	25,628	24,958	0.74
Chad	12	0.65	101,813	109,497	4.10
Congo, Dem Republic of	52	2.81	88,960	177,904	2.30
Congo, Republic of	15	0.81	18,752	65,410	3.75
Côte d'Ivoire	131	7.08	17,728	68,461	1.03
Djibouti	5	0.27	2479	2581	3.86
Equatorial Guinea	2	0.11	47,880	47,880	2.69
Eritrea	5	0.27	4317	21,045	5.66
Ethiopia	113	6.11	50,527	70,516	3.41
Gabon	25	1.35	10,507	18,436	1.93
Gambia	5	0.27	22,254	26,840	1.75
Ghana	88	4.76	41,778	81,420	0.77
Guinea	23	1.24	40,946	52,615	2.54
Kenya	80	4.32	29,856	71,457	2.92
Lesotho	10	0.54	19,638	38,970	0.15
Liberia	2	0.11	37,678	37,678	0.59
Madagascar	33	1.78	15,292	19,509	2.88
Malawi	57	3.08	13,495	30,477	2.79
Mali	30	1.62	30,466	46,283	2.94
Mauritania	15	0.81	17,271	20,955	2.50
Mozambique	62	3.35	33,707	68,073	1.64
Namibia	42	2.27	5246	13,361	1.71
Niger	34	1.84	27,652	69,991	3.21
Nigeria	405	21.89	50,889	150,988	1.60
Rwanda	9	0.49	63,907	57,128	2.40
Senegal	54	2.92	29,195	74,327	2.20

(Continues)

TABLE 1 (Continued)

	Cities	Sample Share (%)	Median Pop.	Av. Pop.	Av. Pop. Growth
Sierra Leone	6	0.32	92,277	80,577	1.93
South Sudan	3	0.16	101,410	142,367	1.64
Sudan	52	2.81	91,025	141,978	3.63
Swaziland	7	0.38	13,838	30,454	−1.04
Tanzania, United Rep of	100	5.41	26,104	569,60	2.59
Togo	22	1.19	51,132	62,932	0.50
Uganda	38	2.05	33,671	54,929	2.97
Zambia	56	3.03	45,171	72,605	2.12
Zimbabwe	62	3.35	22,623	50,864	0.95
Total	1850	100.00	32,972	87,565	2.04

years. f is a function of our main measures of the road network in a city centre. We sometimes use country fixed effects or province (i.e. within-country areas) fixed effects, denoted as η_c above. We define these provinces as the level-1 administrative subdivisions from the GADM data set, which corresponds to areas such as NUTS2 in the European Union.[5] In our main specifications we control for the (log) initial population density of a city $ln(pop_{t=1})$ and (log) distance to the sea $ln(dist.\ sea)$, which is important to condition on for one of our instruments. Finally, we also show results using a richer set of control variables including city-specific ruggedness, elevation, minimum temperature, maximum temperature, average temperature, precipitation, distance to the big lakes and climatic conditions for malaria. We use robust standard errors throughout.

The coefficient associated with $f(Street\ network)$ identifies the causal impact of a feature of the city-centre road network on city's population growth only if that feature is independent of any factor influencing city growth that are not controlled for. This would be violated if cities growing faster (or slower) change their road network. This is another reason why we focus on the city centre, which is less prone to new road construction than city fringes would be. We sometimes include city-specific control variables that could capture geographic endowments influencing both the city-centre road network and population growth. In robustness checks we also include market access measures.

Table 3 displays the change in the coefficient for different versions our main specification, and highlights that the magnitude does not seem to be greatly affected by the inclusion of control variables, country fixed effects or different ways of selecting the sample. The independent variable of interest here is city-centre road density, our main variable. In this table, we are first more interested in the comparison across specifications; we interpret the coefficient results more thoroughly in Section 5. Our conclusion from Table 3 is that the main coefficient of interest is positive, significant and that it remains close to unchanged across alternative specifications. Put differently, the estimate of our favourite specification (column (2)) is not statistically different from any of the alternatives displayed. More specifically, in column (1) we exclude the country fixed effect; in column (3) we use a large set of control variables, including data on ruggedness, elevation, minimum, maximum and average temperature, precipitation distance to the lakes and to the coast, and malaria risk. In column (4) we add (within- country) province fixed effects; in column (5) (and (6)) we also add the cities of Nigeria (and Madagascar) to our sample.

---

[5] The GADM is the 'Global Administrative Areas Data set', cf. details in data appendix.

TABLE 2 Descriptive statistics of the main variables for 1850 cities of the main sample.

	N	Mean	sd	Min	Median	Max
Whole City						
Total Road Length (km)	1850	243	878	0	73	28621
Metres of Road per Sq. Km	1850	2491	2177	0	1850	17,569
Area (Sq. Km.)	1850	94.6	279	2	40	7408
City Centre (1 km radius)						
Total Road Length (km)	1850	19.6	16.1	0	15	87
Metres of Road per Sq. Km	1850	6403	5160	2	5205	27,703
Area (Sq. Km.)	1850	3.10	0.21	1	3	3
Dist. Centre to closest Road (m)	1850	133	167	0	62	975
Road Orientation Herfindahl	1850	0.19	0.19	0	0	1
Average Dist. Points to Road (m)	1850	189	155	10	150	1078
Distances from City Centre						
Distance to Dakar (km)	1850	4066	2071	78	3487	8272
Distance to Djibouti (km)	1850	3583	1451	34	3773	6507
Distance to Cairo (km)	1850	4007	1046	918	3875	6738
Distance to Cape Town (km)	1850	4338	1311	628	4838	7168
Distance to the closest (km)	1850	2065	830	34	2135	3678
Distance to the sea (km)	1850	451	350	0	389	1711

TABLE 3 Baseline model and OLS estimations. *d centre* measures road density in the city centre. The set of controls include variables measuring ruggedness, elevation, minimum, maximum and average temperature, precipitation, distance to the lakes and the coast and malaria risk. Robust standard error in parentheses. Stars denote significance at 10 (*), 5 (**) and 1 (***) percent.

	(1) Δpop 2000–2015	(2) Δpop 2000–2015	(3) Δpop 2000–2015	(4) Δpop 2000–2015	(5) Δpop 2000–2015	(6) Δpop 2000–2015
d centre	0.0014***	0.0012***	0.0012***	0.0012***	0.0011***	0.0011***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
ln pop 2000	−0.0060***	−0.0069***	−0.0086***	−0.0081***	−0.0068***	−0.0068***
	(0.0015)	(0.0016)	(0.0017)	(0.0018)	(0.0014)	(0.0014)
ln dist sea	0.0006	−0.0005	0.0000	−0.0004	−0.0004	−0.0007
	(0.0005)	(0.0007)	(0.0013)	(0.0017)	(0.0015)	(0.0014)
Observations	1,450	1,450	1,450	1,450	1,856	1,890
Country FE		Yes	Yes			
Province FE				Yes	Yes	Yes
Controls			Yes	Yes	Yes	Yes
Nigeria included					Yes	Yes
Madagascar included						Yes

4.2 | Instrumental variable strategies

One concern with the empirical strategy based on Equation 1 is the potential role of reverse causality. If cities that grow faster also build more road or improve the OSM data set faster, then our OLS model would not identify the causality of the road network on population growth. To circumvent this concern, we implement two related instrumental variable (IV) strategies using a city's age as an instrument for its city-centre road network. As discussed in Section 2, the current layout of a city-centre is likely to be a direct product of the initial, historic layout. In the IV regressions we control for initial population density itself in these regressions. We expect to see some road network features changing across cities of different ages. For example, a city founded before the invention and adoption of cars would have a higher density suitable for older technologies. We expect more modern cities, built in the era of increasing adoption of cars, having bigger roads and also smaller road densities.

To measure the foundation age of a city, we rely on the Africapolis (2019) data set, that records historic population for a large number of cities in Sub-Saharan Africa. The oldest year in that data set is 1950, and our instrument consists of a simple dummy variable, indicating those cities that existed in that year according to this data set (i.e. that had 1 or more inhabitants in 1950). One downside of this 'age data set' is that it is only available for a subset of our cities covered by Africapolis and thus reduces our sample (see Supporting information: Appendix B.3 for more details). However, we don't see any reason that this selection is related to our main variables of interest, and hence likely to introduce selection bias.

There could be concerns that the age of a city might influence recent population growth through channels other than the road layout in the city centre. To mitigate this partially, we measure the age crudely as an indicator variable. To address this concern more thoroughly, we also construct a second, related instrument, measuring the distances from the cities of Cape Town, Cairo, Dakar and Djibouti. These four cities represent connection nodes of the two main colonial powers of the continent. If a significant part of colonial expansion was conducted via these ports, distance to these four cities should influence the year of creation of modern cities on the continent. In these specifications we control for the distance to the sea and again for initial population density. The control for the distance to the sea ensures that it is not the distance to ports in general that drives this instrument, but only distance to specific historic ports. As argued for the first instrument: if the age of a city correlates with the initial road layout, then the distance to these four cities will also correlate with the initial road layout. Using the distance from cities to each of these four points is a simplistic vision of history that captures enough information that it might function as an instrument.

Urbanisation in Africa was facilitated to a large degree by the colonial powers, especially the British and the French, the two main parties in the scramble for Africa. As late as 1880, about 80 percent of Africa were ruled by Africa's own kings and queens. Colonial powers were found near the coast, with a British stronghold around the port of Cape Town, and French holdings around Dakar (Boahen, 1985). The British invasion of Cairo of 1882 established a second stronghold on the continent for the British that lasted well into the 20th century (Hourani et al., 2004). Soon it became the vision of the British to connect these two centres, and build the 'Cape to Cairo' Railway line. Historically, these two cities played an important part as connection points of British colonial Africa, and many expeditions, travels and trade originated in one or the other. France's influence on the eastern coast of Africa was initiated by treaties with the rulers of what is today Djibouti from 1883, and the creation of outposts in modern day Dakar. Soon the French ambition became to establish an East-West link between its ports in Dakar and Djibouti. French and British expansions clashed near the town of Fashoda, in the Fashoda incident of 1898 (Bates 1984). In schoolbooks, this episode in history is sometimes illustrated with maps that show the cities of Cairo, Cape Town, Dakar and Djibouti, with four arrows meeting near Fashoda. This simplified schoolbook view of colonial expansion in Africa is the model we follow in the construction of our instruments, taking the distance to these colonial origin ports as instrument for the age of a town, and consequently its initial road grid.

4.2.1 | First-stage regressions

Table 4 shows the correlation between our main right hand side variables, which are the measures of road density and unevenness, and the distance to the four instrument cities. This is effectively a first stage regression. As expected, we observe a significant negative or no relationship between the distance to the ports and road density in the city centre in Columns (1) - (4), which could reflect that newer cities that were designed with modern transportation in mind rely on a less dense road grid. It is negative in the case of three of our four cities and insignificantly different from zero at five percent level of statistical significance in the case of Cape Town.[6] When we jointly include all four cities in one regression and add various control variables as in column (5) we continue to observe negative coefficients, typically at a high level of statistical significance. Coefficients become weaker when we include country fixed effects, as in column (6). This seems plausible to us since we expect the distance to these colonial ports to matter less within countries than for the longer distances across the whole continent. However, it creates a weak instrument problem for results with country fixed effects. Hence, we show the main results later with and without country fixed effects. In our IV strategy we prefer to use the four distances as separate variables to simply using the minimum distance given that British and French colonial legacies could have influenced cities in different ways (Baruah et al., 2021). In columns (7) and (8) we repeat the estimation from columns (5) and (6) but using unevenness on the left-hand side, rather than density in just the city centre. Again, we find a strong relationship without country fixed effects and a much weaker one when we include country fixed effects.

In columns (9) and (10) we provide more direct evidence that distance to these historic ports is indeed correlated with the age of a city. For this analysis we use the data from the ‘Africapolis’ project. Keeping only cities in our data set that are also in the Africapolis reduces the sample to 1,107 observations. We then compute an indicator for cities that, according to Africapolis, did not have any population in 1950. As columns (9) and (10) show, there is a significant relationship between these distances and the age of a city. Once we include our control variables and country fixed effects, the relationship is positive, which confirms that cities that are further away from these ports are more likely to be new. Columns (11) and (12) are then first stage regressions when using the age indicator for cities as an instrument. Finally, in Columns (11) and (12) we show that this simple indicator variable for cities that are older than 1950 correlates strongly significantly with the density of roads in the city centre and unevenness in the expected way.

4.2.2 | Exclusion restriction

These IV strategies could uncover the causal relationship between the road network and population growth only under the exclusion restriction that the age of a city influences its recent population growth only through its impact on the road layout. In our preferred 2SLS estimation we always include initial population size and distance to the sea as well as country (or province) fixed effects. This ensures that only the effect of a city's age through these two channels that is time-invariant but country-specific is taken into account.

If, as we argued, age and location of cities are related, then geographic factors could potentially influence city growth through other channels than just the city-centre road layout. The inclusion of the distance to the coast does not only address the concern that the four ports (Cairo, Cape Town, Dakar and Djibouti) still play a large role as transportation hubs today, but it also picks up other influences the coast may have on the local economy. We also note that there are large ports available along the African coast, apart from these four

---

[6] On average, the coefficient in columns (1)-(4) is −0.87, perhaps close to the distance coefficient of −1 sometimes estimated for distance coefficients (Rauch, 2016).

TABLE 4 First stage regressions. *new* is an indicator for cities founded after 1950. Variables *ln dist* refer to log distance to either Cairo, Cape Town, Dakar, Djibouti, or the sea. *d centre* measures road density in the city centre; *uneven* is the unevenness index. The set of controls consists of variables measuring ruggedness, elevation, minimum temperature, maximum temperature, average temperature, precipitation, distance to the big lakes, distance to the coast and climatic conditions for malaria. Stars denote significance at 10 (*), 5 (**) and 1 (***) percent.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	d centre	d centre	d centre	d centre	d centre	d centre	uneven	uneven	new	new	d centre	uneven
new											−2.284***	54.42***
											(0.356)	(9.682)
ln dist sea	−0.0318	−0.121*	−0.198***	−0.0783	−0.0833	−0.0894	5.369	10.79**	0.0279***	0.0188	0.0999	8.439*
	(0.0734)	(0.0726)	(0.0760)	(0.0715)	(0.133)	(0.143)	(4.029)	(4.677)	(0.00829)	(0.0175)	(0.166)	(4.947)
ln pop 2000	1.901***	1.886***	1.821***	1.892***	1.979***	1.955***	−44.48***	−43.95***	−0.135***	−0.146***	1.534***	−31.36***
	(0.0937)	(0.0925)	(0.0928)	(0.0947)	(0.0975)	(0.107)	(2.862)	(2.856)	(0.00991)	(0.0104)	(0.149)	(3.848)
ln dist Cap				0.365	−5.494***	−3.271*	115.5***	32.13	0.0800	0.563***		
				(0.303)	(0.893)	(1.824)	(33.76)	(59.52)	(0.0986)	(0.212)		
ln dist Dhak	−0.399**				−2.651***	−1.431**	56.9***	17.08	0.113**	0.0226		
	(0.164)				(0.419)	(0.620)	(17.41)	(20.25)	(0.0447)	(0.0701)		
ln dist Dji		−0.738***			−0.545	0.383	−9.919	−43.01	−0.0847**	0.0552		
		(0.225)			(0.545)	(1.637)	(10.85)	(27.23)	(0.0360)	(0.0516)		
ln dist Cai			−2.826***		−8.564***	−7.877***	128.9***	141.6**	0.203*	0.673**		
			(0.548)		(1.164)	(2.658)	(31.51)	(69.81)	(0.114)	(0.264)		
Observations	1450	1450	1450	1450	1450	1450	1450	1450	1107	1107	1107	1107
Controls					Yes	Yes	Yes	Yes		Yes	Yes	Yes
Country FE						Yes		Yes		Yes	Yes	Yes

cities.[7] In addition to seaports, airports now play a substantial road in connecting Africa, which further reduces the importance of these four historic ports in contemporary trade. Also, to address concern relating to the location of cities within the continent we also provide robustness specifications in which we control for market access measures of each city in our data set. Including such measures barely change the magnitude and significance of our coefficients of interest.

Another concern for our instrumental variable strategies is that other colonial legacies could equally depend on age. For example, there may be physical infrastructure other than roads whose age depends on these distances in a similar way, and that have a similar effect on future development. We can't dismiss this concern entirely but note some limits to this concern. First, as we argued previously, we think that the road layout is particularly prone to lock-in, compared to other types of infrastructure. Additionally, road construction and city planning were of central importance to all potential colonial investments. Second, building reconstruction rates in Africa are very high. Michaels et al. (2021) estimate an annual replacement rate of five percent for housing in Tanzania, Henderson et al. (2016) find a replacement rate of 3.6 percent in Kenya. Given these high replacement rates, most houses and most other physical infrastructure has been rebuilt several times since the early 20th century, and have much time to converge from an initial steady state to a new one. Path dependent behaviour of infrastructure quality is far less likely in such a rapidly changing environment. Of course, also the physical characteristics (the 'quality' and size) of the roads were changed and updated multiple times in the great majority of our towns and cities. It is, however, the layout and location of roads, the plan, that we think is harder to change and that persists even if roads are reconstructed many times. Third, a large share of public infrastructures in Sub-Saharan Africa happens to be concentrated in capital cities (Bekker & Therborn, 2012), which are a small share of the sample, and which are excluded from the analysis in one of our robustness checks. If public policy programmes are less common outside these cities, then there are fewer channels through which administrative legacies could manifest. Note also that, in our baseline specifications we always include a country (or province) fixed effect. This also addresses other concerns related to institutional legacies. Any institutional setting that is set (and invariant) at the country (province) level would be absorbed by the fixed effects. Finally, we are reassured by the fact that the key results from the IV specifications are robust to many alternative specifications or sample changes.

5 | RESULTS

5.1 | Results from OLS estimation

To estimate the impact of various features of the city-centre road network on cities' growth and prosperity, we first estimate Equation 1. In this specification, we always include a country fixed effect as well as (log) initial population density and minimum distance to the sea (cf. Section 5.1 for more details). We investigate the relationship between our simple variables of road network and population growth.

Table 5 reports the results. Column (1) shows that a greater road density in the city centre correlates positively and strongly statistically significantly with population growth in the period 2000–2015. The measure of road density in the city centre 'd centre' is measured in km/km^2 and has a mean of around 6. Increasing the road density in the city centre by 1 km per km^2, is associated with more than one-tenth of a percent of higher population growth annually. Accumulated over 15 years, this adds up to a total population that is more than 1.55 percent larger

---

[7] The busiest ports of Africa are (in order descending in size according to total cargo volume): Richard's Bay, Saldanha Bay, Alexandria, Damietta, East Port, Apapa, Casablanca, Skikda, Pointe Noire, Mombasa, Abidjan, Tanger, Bejaia, Jorf Lasfar and Tin Can Island (AAPA - American Association of Port Authorities, 2015).

TABLE 5 Results from the baseline OLS regression model. *d centre* measures road density in the city centre; *n centre* is the number of roads' nodes in the city centre; *uneven* is the unevenness index; *orientation* H is the orientation Herfindahl index. Robust standard error. Stars denote significance at 10 (*), 5 (**) and 1 (***) percent.

	(1) Δpop 2000–2015	(2) Δpop 2000–2015	(3) Δpop 2000–2015	(4) Δpop 2000–2015	(5) Δpop 2000–2015	(6) Δpop 2000–2015	(7) Δpop 2000–2015	(8) Δpop 2000–2015	(9) Δpop 2000–2015	(10) Δpop 2000–2015
d centre	0.0012***		0.0013**		0.0009***		0.0012***	0.0007	0.0012	0.0011*
	(0.0003)		(0.0006)		(0.0003)		(0.0003)	(0.0007)	(0.0008)	(0.0006)
n centre		0.00003***	−0.00000					0.0000	0.0000	0.0000
		(0.0000)	(0.00002)					(0.0000)	(0.0000)	(0.0000)
uneven				−0.00003***	−0.00002**			−0.00002**	−0.00001	−0.00001*
				(0.0000)	(0.0000)			(0.0000)	(0.0000)	(0.0000)
orientation H						−0.0083	−0.0002	0.0037	−0.0004	0.0037
						(0.0052)	(0.0048)	(0.0052)	(0.0051)	(0.0052)
ln pop 2000	−0.0069***	−0.0064***	−0.0069***	−0.0061***	−0.0072***	−0.0050***	−0.00691***	−0.00712***	−0.00691***	−0.00616***
	(0.00163)	(0.00162)	(0.00160)	(0.00152)	(0.00167)	(0.0014)	(0.0017)	(0.0016)	(0.0016)	(0.0015)
ln dist sea	−0.0005	−0.0005	−0.0005	−0.0007	−0.0006	−0.0006	−0.0005	−0.0006	0.0008	0.0006
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0014)	(0.0005)
Observations	1450	1450	1450	1450	1450	1450	1450	1450	1450	1450
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Province FE									Yes	

because of this additional road length. Below we multiply this effect with our preferred estimate. In column (2) we show the estimated impact of the number of nodes (i.e. any intersection with one road or more) within the city-centre also correlates positively with city growth but column (3) shows that this effect disappears when the density of roads is considered. In other words, the number of nodes correlates with city growth insofar as it is a direct product of the total amount of road. In columns (4) to (7) we turn to our two indices of spatial organisation, the unevenness and orientation Herfindahl indices, respectively. We describe these indices in Section 3.3. The main result in column (4) indicates that as the evenness of the road grid (respectively, the index) increases (decreases) so does the population growth over 2000–2015. In column (5) we also add city-centre road density as an independent variable. The concern here is that two cities with identical road density as measured by our *d centre* measure may nevertheless have quite different flows in the centre. Similar to the result in column (4), the negative sign associated to the unevenness index indicates that larger average distances to the nearest road harm population growth, holding road density constant, as expected.[8] In column (6) we estimate the correlation between the orientation Herfindahl index and population growth: a negative correlation appears that is not statistically significant at conventional levels. This coefficient shrinks and remains insignificant in column (7), when road density is controlled for. As explained in Section 3.3, both the unevenness and the orientation Herfindahl indices eventually measure different things, although they correlate quite strongly (0.42). One could for instance think of a city-centre where all roads run north to south (highest Herfindahl index) yet are all concentrated in one specific neighbourhood (highest unevenness); of these two forces, the unevenness seems a dominant factor for city growth.

In column (8) we estimate a model including all the road network features at once (i.e. street density, number of nodes, the unevenness index and the orientation Herfindahl index). We draw two main conclusions from this exercise: first, the only feature whose impacts remain significant is the unevenness index. Such results indicate that the spatial distribution of the network matters even after the total amount of roads and nodes, as well as their overall bearing concentration are considered. Second, although the significance of the road density collapses, the point estimate remains quite similar in most specifications. We take this result as an indication of the robustness of our result and of the importance of road density. Note also that city-centre road density directly correlates with all other independent variables, and it is therefore not surprising that the standard errors increase in column (8) compared to column (1). In column (9) and (10) we substitute either (within-country) province fixed effects or no fixed effect, respectively, to the country fixed effect. Column (9) is a strategy to reduce the potential measurement bias (Seidel, 2023; cf. Section 3.3) while column (10) mirrors subsequent specification in the remainder of this paper. In both cases, the effect of the unevenness index is reduced and remains significant only at the 10% level in column (10). The coefficient of road density remains close to other estimates in terms of magnitude, but is weaker in terms of statistical significance. What Table 5 shows is that of the variables tested, only *d centre* and *uneven* show a significant and somewhat robust correlation with population growth, while the other variables don't.[9] We continue the more detailed investigation with these two central variables.

5.2 | Results from the 2SLS estimation

We next turn to the IV estimates corresponding to the OLS results derived so far. First, we use an indicator variable for new cities as an instrument for the road layout in the city centre in Table 6. This variable indicates cities that were founded after 1950 according to the Africapolis data set (cf. Section 4.2 for more details). Given that not all

---

[8] The unevenness index is expressed in metres so the coefficient in column (4) can be read as: within a given country, cities where the evenness is 1 m smaller also had a population growth that is 3×10^{-5} higher (given initial city population and its distance to the sea coast). Accumulated over 15 years this implies a population that is more than 0.045 percent larger.

[9] Consider the reported coefficients in Column (10) of Table 5. In combination with the standard deviations reported in Table 2 we can see that an increase by one standard deviation has an effect of 0.018 for *d centre*, of 0.0017 for *uneven*, but only 0.00007 for *orientation* on the outcome.

TABLE 6 IV using age as instrument. *d centre* measures road density in the city centre; *uneven* is the unevenness index. The first stage F statistic reports the F test of the excluded instrument, which is an indicator for cities that were founded after 1950. The set of controls consists of variables measuring ruggedness, elevation, minimum temperature, maximum temperature, average temperature, precipitation, distance to the big lakes, distance to the coast and climatic conditions for malaria. Stars denote significance at 10 (*), 5 (**) and 1 (***) percent.

	(1) Δpop 2000–2015	(2) Δpop 2000–2015	(3) Δpop 2000–2015	(4) Δpop 2000–2015	(5) Δpop 2000–2015	(6) Δpop 2000–2015	(7) Δpop 2000–2015	(8) Δpop 2000–2015
d centre	0.00574***		0.00653***		0.00920**		0.00621***	
	(0.00217)		(0.00176)		(0.00391)		(0.00159)	
uneven		−0.0002***		−0.0003***		−0.0002***		−0.0003***
		(0.0000)		(0.0000)		(0.0000)		(0.0000)
ln dist sea	0.00110	0.00030	−0.00017	−0.00075	0.0019	0.0022	0.00016	0.00298
	(0.00071)	(0.00075)	(0.00098)	(0.0011)	(0.0017)	(0.0015)	(0.0017)	(0.00195)
ln pop 2000	−0.0153***	−0.0135***	−0.0169***	−0.0159***	−0.0245***	−0.0160***	−0.0187***	−0.0174***
	(0.00483)	(0.00410)	(0.00433)	(0.00392)	(0.0084)	(0.0040)	(0.00402)	(0.00361)
Observations	1107	1107	1107	1107	1107	1107	1107	1107
Country FE			Yes	Yes			Yes	Yes
Controls					Yes	Yes	Yes	Yes
First stage F	18.7	15.8	45.2	26.8	8.3	14.6	42.1	30.5

our towns and cities feature in the Africapolis data set, this reduces the size of our sample to 1,107. This instrument follows the same logic as the port distances, which is that older cities are more likely to be stuck with suboptimal grids; but it measures the age of a city directly. The advantage of this strategy is that we get a strong first stage, with the exception of the specification in column (5). All coefficients on road density are positive and statistically significant, while magnitudes on unevenness are negative and statistically significant. Magnitudes are larger than in the OLS estimate, which might be explained by a reverse causality problem in the OLS regression, and some scope for measurement error. Qualitatively, this table supports the OLS findings on density and unevenness and shows that both are strongly associated with population growth in recent years. As expected, a city with a denser road grid has a higher population growth, while a more uneven road grid leads to less population growth. This confirms that a city with a regular street grid consisting of roads that are evenly distributed has favourable conditions to grow.

Table 7 presents results corresponding to the OLS results in Table 5, but this time estimated as 2SLS using distances to colonial ports as instruments. Coefficients retain the expected signs, positive for road density and negative for unevenness. This table suffers from weak first stages in columns with country fixed effects. We still report the coefficient for comparison. We note that coefficients are similar in magnitude to the ones estimated using the other instrument. Overall, both IV strategies seem to confirm the qualitative results from our baseline OLS estimation. The impact of city-centre road length density remains positive and significant with the 2SLS estimations. Similarly, the coefficient associated to the unevenness index indicates that better spread roads have a positive impact on city growth over the 2000-2015 period and 2SLS only confirms this conclusion. For both measures and both instruments, the point estimate of the coefficient is greater in 2SLS than in OLS, suggesting that the potential bias in OLS attenuates the true effect of roads. We acknowledge than these two IVs have limitations, but we find it worth reporting that they yield comparable conclusions regarding both the direction of the effect of road network and the direction of the potential bias in OLS. Coefficients again have the expected signs as in previous tables.

In terms of magnitude, a coefficient of 0.005 on road density implies that if a city increases its central road density by 1 km per unit circle area, its total population grows by an additional 0.5 percent per year. We can compare this with the mean of 6.4 km in our sample, and a standard deviation of 5.2. Accumulated over 15 years, our preferred estimate thus leads to a population that is 7 percent larger in a city that has an additional kilometre of road in the centre. On unevenness, a coefficient of −0.0002 suggests that increasing the unevenness measure by 100 reduces population growth by 0.02 percent per year. Given that the mean unevenness is around 200, this is a less quantitatively strong effect than the one we report concerning road density.

6 | ROBUSTNESS CHECKS

In this subsection, we verify that our results remain robust when we vary some of the choices guiding our main specification and sample selection. These results are shown in Table 8. The 6 columns reports our 6 main estimations: the main independent variable is road density in columns (1) to (3) and unevenness in columns (4) to (6); for both independent variables we present the results associated to our basic OLS (columns (1) and (4), respectively), our distance IV (columns (3) and (6)) and our age IV (columns (2) and (5)). Panel A displays the baseline results (i.e., a specification that uses the full set of control variables including log initial population and distance to the sea as well as a country fixed effect). Given the country fixed effects, the distance IV results have low statistical power throughout this table, as in previous results.

Panel B repeats the main exercise but excludes capital cities. It could be that capital cities follow unique development paths within their countries, given their better institutions, public goods provisions, and international connections (Bekker & Therborn, 2012). As Panel B shows, excluding capital cities does not change the coefficients in either magnitude or statistical significance in any meaningful way.

TABLE 7 IV using distances as instrument. *d centre* measures road density in the city centre, *uneven* is the unevenness index. The First stage F test statistic reports the multivariate F test of the excluded instruments. The set of controls consists of variables measuring ruggedness, elevation, minimum temperature, maximum temperature, average temperature, precipitation, distance to the big lakes, distance to the coast and climatic conditions for malaria. Stars denote significance at 10 (*), 5 (**) and 1 (***) percent.

	(1) Δpop 2000–2015	(2) Δpop 2000–2015	(3) Δpop 2000–2015	(4) Δpop 2000–2015	(5) Δpop 2000–2015	(6) Δpop 2000–2015	(7) Δpop 2000–2015	(8) Δpop 2000–2015
d centre	0.00242***		0.00513**		0.00333***		0.00358	
	(0.000828)		(0.00253)		(0.00104)		(0.00237)	
uneven		−0.0000		−0.00028**		−0.0000*		−0.000163*
		(0.0000)		(0.0001)		(0.0000)		(0.0000)
ln dist sea	0.000694	0.000326	−0.000256	−0.00120	0.000962	0.00103	0.000109	0.00154
	(0.000486)	(0.000530)	(0.000671)	(0.000997)	(0.00122)	(0.00127)	(0.00131)	(0.00168)
ln pop 2000	−0.00802***	−0.00630**	−0.0143***	−0.0167***	−0.0116***	−0.00850***	−0.0132**	−0.0134***
	(0.00239)	(0.00259)	(0.00547)	(0.00603)	(0.00288)	(0.00289)	(0.00520)	(0.00463)
Observations	1450	1450	1450	1450	1450	1450	1450	1450
Country FE			Yes	Yes			Yes	Yes
Controls					Yes	Yes	Yes	Yes
First stage F	58.6	12.2	4.8	7.1	32.6	6.4	9.2	7.2

TABLE 8 Robustness specifications. Each Panel in this table repeats the main results in the version with all the control variables used before, with modifications as indicated. We focus on two main independent variables: *d centre* and *uneven*. Columns (1) and (4) report OLS results, columns (2) and (5) results from the distance IV and columns (3) and (6) results from the age IV. First stage F-statistics are reported next to the number of observations. Panel A reproduces the baseline results. Panel B excludes capital cities. Panel C adds controls for market access. Panel Panel D includes Nigeria. Panel E includes Nigeria and Madagascar. Panel F uses variables for the whole city rather than just the centre. Panel G uses level outcomes instead of growth outcomes. These could be interpreted as growth from an initial log level of zero. Panel H uses growth starting in 1975. Panel I uses growth in lights rather than population. Panel J uses population data from GWP4. Panel J uses a centre circle with radius of 500 m instead of 1 km. Panel L re-estimates the main specifications in an entirely different data set from UCDB. The latter specification uses control variables for initial log population and log distance to the sea, as well as country fixed effects.

	(1) **Δpop** **2000–2015**	**(2)** **Δpop** **2000–2015**	**(3)** **Δpop** **2000–2015**		**(4)** **Δpop** **2000–2015**	**(5)** **Δpop** **2000–2015**	**(6)** **Δpop** **2000–2015**
A Baseline							
d centre	0.00120***	0.00358	0.00621**	uneven	−0.00003***	−0.0002*	−0.00026***
	(0.00027)	(0.00237)	(0.00159)		(0.0000)	(0.0000)	(0.0000)
Observations/F-stat	1450	1450/9.2	1107/42.1		1450	1450/7.2	1107/30.5
B Exclude capital cities							
d centre	0.00110***	0.00251	0.00598***	uneven	−0.00003***	−0.00011	−0.00025***
	(0.000263)	(0.00198)	(0.00157)		(0.0000)	(0.00008)	(0.00006)
Observations/F-stat	1412	1412/9.2	1070/42.1		1412	1412/7.2	1070/30.5
C Include market access controls							
d centre	0.00108***	0.00275	0.00611**	uneven	−0.00003***	−0.00011	−0.00025***
	(0.000262)	(0.00212)	(0.00164)		(0.0000)	(0.00008)	(0.00007)
Observations/F-stat	1412	1412/11.2	1070/35.7		1412	1412/7.4	1070/28.5
D Include Nigeria							
d centre	0.00114***	0.00272	0.00471***	uneven	−0.00003***	−0.00013	−0.00025***
	(0.00023)	(0.00220)	(0.00109)		(0.0000)	(0.00009)	(0.00006)
Observations/F-stat	1856	1856/11.3	1466/59.9		1856	1856/6.8	1466/28.0
E Include Nigeria and Madagascar							
d centre	0.00115***	0.00311	0.00471***	uneven	−0.00003***	0.00015	−0.00025***
	(0.000231)	(0.00221)	(0.0011)		(0.0000)	(0.00010)	(0.00006)
Observations/F-stat	1890	1890/11.8	1466/59.9		1,890	1,890/7.0	1,466/28.0
F City wide variables							
d centre	0.00209***	0.0107	0.0259***				
	(0.00045)	(0.0089)	(0.00804)				
Observations/F-stat	1450	1450/5.8	1107/15.1				
G Outcome in levels	lnpop2015	lnpop2015	lnpop2015		lnpop2015	lnpop2015	lnpop2015
d centre	0.0169***	0.0501	0.0869***	uneven	−0.00046***	−0.00229*	−0.00365***

(Continues)

TABLE 8 (Continued)

	(1)	(2)	(3)		(4)	(5)	(6)
	Δpop	**Δpop**	**Δpop**		**Δpop**	**Δpop**	**Δpop**
	2000–2015	**2000–2015**	**2000–2015**		**2000–2015**	**2000–2015**	**2000–2015**
	(0.0038)	(0.0332)	(0.0222)		(0.00011)	(0.000136)	(0.00093)
Observations/F-stat	1450	1450/9.2	1107/42.1		1450	1450/7.2	1107/30.5
H Outcome from 1975	Δpop	Δpop	Δpop		Δpop	Δpop	Δpop
	1975–2015	1975–2015	1975–2015		1975–2015	1975–2015	1975–2015
d centre	0.000175***	0.00003	0.00587***	uneven	−0.00041***	0.0000	−0.00025***
	(0.000170)	(0.00239)	(0.000821)		(0.0000)	(0.0000)	(0.00003)
Observations/F-stat	1377	1377/3.4	1061/56.2		1377	1377/2.8	1061/47.3
I Outcome Δ light 2021-2014	Δlight	Δlight	Δlight		Δlight	Δlight	Δlight
d centre	0.00363***	0.00326	0.00608*	uneven	−0.00006**	0.00068	−0.00029*
	(0.00062)	(0.00793)	(0.0032)		(0.0000)	(0.00069)	(0.00016)
Observations/F-stat	1406	1406/4.9	1073/35.4		1406	1406/2.9	1073/21.7
J Outcome measure from GwP4							
d centre	0.000422***	0.00135	0.00119**	uneven	−0.0000**	−0.0000	−0.00005**
	(0.000114)	(0.00205)	(0.00568)		(0.0000)	(0.0000)	(0.00002)
Observations/F-stat	1450	1450/2.8	1107/41.3		1450	1450/3.7	1107/28.4
K Centre circle 500 m							
d centre	0.000906***	0.00318	0.00508***	uneven	−0.00005***	−0.00019	−0.000594**
	(0.00027)	(0.00210)	(0.00139)		(0.0000)	(0.00018)	(0.000238)
Observations/F-stat	1376	1376/8.4	1061/28.5		1376	1376/4.0	1061/7.6
L UCDB sample							
d centre	0.000884***	0.00265**	0.00719***	uneven	−0.00016	0.00016	−0.00065***
	(0.000137)	(0.00122)	(0.00245)		(0.00011)	(0.00015)	(0.00022)
Observations/F-stat	1214	1214/17.9	983/9.3		1214	1214/7.9	983 12.1
M Only countries with similar number of cities in baseline and UCDB samples							
d centre	0.0013***	0.00022	0.00719***	uneven	−0.00004***	0.00011	−0.00031***
	(0.0003)	(0.00171)	(0.00200)		(0.00001)	(0.0001)	(0.00010)
Observations/F-stat	866	866/12.1	739/27.1		866	866/6.2	739/19.99

Panel C adds measures of market access to the baseline model. We measure market access by distance weighted populations of all other cities in our data set. The weights we use are −1, −2 and −4, as in Maurer and Rauch (2023); and we include all three market access measures in logs to the main model. Controlling for market access should capture the geography-induced trade factors that favour (or impede) city growth. This is a potentially important channel which also threatens the exclusion restriction of the distance instrument. We find that these additional control variables don't affect results meaningfully.

Panels D and E change the sample, by introducing the countries of Nigeria and Madagascar, respectively. We argued against the inclusion of the cities in Nigeria because of potential measurement errors, but also provide results when including one or the other to be comprehensive. These inclusions increase the sample size. Results stay similar, except for the distance IV, which now becomes insignificant in the case of the unevenness measure. We find this weakening of the distance instrument to be plausible, as the reason why the exclusion restriction applies in the case of Sub-Saharan Africa does not apply to Madagascar, which had a colonial history that did not rely on expansion from the four colonial ports we use there. Hence, the inclusion of this country adds noise to these specifications, which in turn weakens the strength of the IV estimation.

In Panel F we replace city-centre measure of road density by city-wide measure of road density. This checks whether potential errors in identification of the historical centre of cities affects results. We do not compute the unevenness variable for this sample, since conceptually it can't be compared easily for cities of different areas. The sign of the coefficient in each robustness check remains similar to the baseline, implying that cities with a greater road density overall grew faster. One of the IV specifications is no longer statistically significant. The magnitudes also show larger estimates for IV specifications.

In Panel G we use population levels as outcome variable instead of population growth rates to show that our results do not depend on the potentially noisy population data for the year 2000. All coefficients have the expected sign, and they remain statistically significant. These results can be interpreted as growth from a starting point close to zero for every city.

In Panel H we replace the main dependent variable (annualised population growth rate over 2000-2015) by the annualised population growth rate over the longer 1975 to 2015 period. The road density results remain positive and significant, apart from one of the IV specifications which turns insignificant. The unevenness results don't have enough power in this specification. So, these results are consistent with our baseline results when significant, but they suffer from reduced statistical power.

In Panel I we use nightlights as main outcome variable rather than population. Nighttime lights capture economic activity in general and thus offers insights complementary to results that rely on population density data alone. We use the nightlights data version VIIRS (see Supporting information: Appendix A) and compute city-level growth of total light intensity (i.e. the sum of pixels' values) from 2014 to 2021. Not all coefficients are statistically significant, but they are consistent with our baseline sample where they are, with a positive relation between roads density and a negative relationship with unevenness.

In Panel J we replicate results using outcome data from gridded world population version 4, an alternative source of population density data.[10] Again we find that coefficients are consistent with the baseline results where they are significant, and overall statistical power is reduced.

In Panel K we change the definition of the city centre by reducing the circle that defines the centre from radius 1 km to radius 500 m. This reduces the area of the city centre to a quarter. The coefficients remain strongly statistically significant and give the same sign as the baseline. The magnitudes of coefficients remain fairly similar to our baseline specification.

In Panel L we replicate results completely in an alternative data set, the UCDB sample. We want to emphasise that this is a replication in a completely separate and newly created alternative data set. While coefficients are less significant than in the equivalent baseline regressions, they have the same sign as the baseline specification when they are statistically significant.

The number of cities reported by country differs between our baseline sample and the UCDB sample. So, in Panel M we restrict the set of countries used to those where the difference in the number of observations is no greater than 50 percent between the two samples. Again, we observe coefficients of similar signs and magnitudes as in our baseline specifications.

---

[10] Both GHS and GPW4 are imputed from census data. We discuss the role of that imputation in Supporting information: (Appendix 2.3.2).

7 | CONCLUSION

Having a road layout in the centre of a city that facilitates interaction between people is an essential urban public good. Providing this good is an important policy concern, particularly in the fast growing and urbanising Africa, where virtually all urban transport uses roads. In this paper we show empirically that indeed the layout of roads in the centre of a city influences population growth of African cities. Cities that have either a road density that is too low in the city centre, or an uneven road network in the centre grow less than other cities. The magnitude of the effect is large, such that a city that has an additional kilometre of road in the centre defined as area of a circle with radius of 1 kilometre will have a population that is 7% larger after 15 years.

Better connectivity leads to higher agglomeration economies, which are likely the mechanism that drives these results. Cities with dysfunctional urban connectivity fall behind. Given the importance of the problem and the large magnitudes we estimate, policy makers in affected cities are most likely aware of the problem. They seem unable to solve the issue by retrofitting better road layouts. Correcting infrastructure ex-post seems prohibitively costly. Rather than fixing existing cities, it might be more feasible to carefully plan and protect road networks in cities while they are still small, before they are locked into suboptimal equilibria.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

Open Access funding enabled and organized by Projekt DEAL.

REFERENCES

AAPA - American Association of Port Authorities, "World Port Rankings 2015", downloaded from http://www.aapa-ports.org on March 22nd, 2017.

Africapolis, Aggomeration data, downloaded from http://africapolis.org/data in March, 2019.

Akbar, P., Couture, V., Duranton, G., & Storeygard, A. (2023). Mobility and congestion in urban India. *American Economic Review*, *113*(4), 1083–1111.

Antos, S. E., Lall, S. V., & Lozano-Gracia, N. (2016). The morphology of African cities. *World Bank Policy Research Working, Paper 7911*.

Barrington-Leigh, C., & Millard-Ball, A. (2017). The world's user generated road map is more than 80% complete. *PLoS One*, *12*(8), e0180698.

Barrington-Leigh, C., & Millard-Ball, A. (2019). A global assessment of street-network sprawl. *PLoS One*, *14*(11), e0223078.

Barrington-Leigh, C., & Millard-Ball, A. (2020). "Global trends toward urban street-network sprawl", Proceedings of the National Academy of Sciences, Jan, 117 (4) 1941–1950.

Baruah, N. G., Henderson, J. V., & Peng, C. (2021). Colonial legacies: shaping African cities. *Journal of Economic Geography*, *21*(1), 29–65.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, *122*(2), 775–805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *The Review of Economics and Statistics*, *99*(3), 435–448.

Bekker, S. B., & Therborn, G. (2012). *Capital cities in Africa: Power and powerlessness*. Hsrc Press.

Boahen, A. (1985). "Africa and the colonial challenge", in Boahen (edt.) 'General History of Africa VII', United Nations Educational, Heinemann Books, London.

Boeing, G. (2017). OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, *65*, 126–139.

Boeing, G. (2018). "Urban spatial order: Street network orientation, configuration, and entropy".

Bryan, G., Glaeser, E., & Tsivanidis, N. (2020). Cities in the developing world. *Annual Review of Economics*, *12*, 273–297.

Cao, S., Ma, G., & Mao, H. (2023). Is distance really dying? Transportation and knowledge spillovers. *Journal of Regional Science*.

Cervero, R. B. (2013). Linking urban transport and land use in developing countries. *Journal of Transport and Land Use*, *6*(1), 7–24.

Christaller, W. (1933). "Die zentralen Orte in Süddeutschland (the central places in southern Germany)." *Jena: Gustav Fischer*.

Collier, P., & Venables, A. J. (2016). Urban infrastructure for development. *Oxford Review of Economic Policy*, *32*(3), 391–409.

Couture, V., Duranton, G., & Turner, M. A. (2018). Speed. *The Review of Economics and Statistics*, *100*(4), 725–739.

Cui, J., Li, T., & Wang, Z. (2023). Research collaboration beyond the boundary: Evidence from university patents in China. *Journal of Regional Science*, *63*(3), 674–702.

Donaldson, D., & Hornbeck, R. (2016). Railroads and American economic growth: A "market access" approach. *The Quarterly Journal of Economics*, *131*(2), 799–858.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, pp. 2063–2117). Elsevier.

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, *101*(6), 2616–2652.

Duranton, G., & Turner, M. A. (2012). Urban growth and transportation. *The Review of Economic Studies*, *79*(4), 1407–1440.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, *81*(3), 1046–1070.

Fuller, B., & Romer, P. (2014). *Urbanization as opportunity*. The World Bank.

Grover Goswami, A. G., & Lall, S. V. (2016). 'Jobs in the city: Explaining urban spatial structure in Kampala', *World Bank Policy Research Working Paper* 7655, World Bank, Washington D.C.

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

Henderson, J. V., Regan, T., & Venables, A. "Building the city: Sunk capital and sequencing" (Discussion paper 11211, Centre for Economic Policy Research, London, 2016).

Hourani, A. H., Philip Shukry, K., & Mary Christina, W. (2004). The modern Middle East: A eeader (2nd ed.). London: I.B. Tauris. ISBN 1-86064-963-7.

Jedwab, R., & Moradi, A. (2016). The permanent effects of transportation revolutions in poor countries: evidence from Africa. *Review of Economics and Statistics*, *98*(2), 268–284.

Jedwab, R., & Storeygard, A. "The heterogeneous effects of transportation investments: Evidence from Sub-Saharan Africa 1960–2010." Unpublished manuscript, Tufts Univ., Medford, (2016).

Kocornik-Mina, A., McDermott, T. K. J., Michaels, G., & Rauch, F. (2020). Flooded cities. *American Economic Journal: Applied Economics*, *12*(2), 35–66.

Lall, S. V., Henderson, J. V., & Venables, A. J. (2017), Africa's cities: Opening doors to the world, World Bank, Washington DC.

Larcom, S., Rauch, F., & Willems, T. (2017). The benefits of forced experimentation: Striking evidence from the London underground network. *The Quarterly Journal of Economics*, *132*(4), 2019–2055.

Marshall, A. (1920). *Principles of economics*. Macmillan.

Maurer, S., & Rauch, F. (2023). Economic geography aspects of the Panama Canal. *Oxford Economic Papers*, 75(1), 142–162.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *Review of Economics and Statistics*, *90*(4), 683–701.

Michaels, G., Nigmatulina, D., Rauch, F., Regan, T., Baruah, N., & Dahlstrand, A. (2021). Planning ahead for better neighborhoods: Long-run evidence from Tanzania. *Journal of Political Economy*, *129*(7), 2112–2156.

Michaels, G., & Rauch, F. (2018). Resetting the urban network: 117–2012. *The Economic Journal*, *128*(608), 378–412.

Mills, D. E. (1981). Growth, speculation and sprawl in a monocentric city. *Journal of Urban Economics*, *10*(2), 201–226.

Rauch, F. (2016). The geometry of the distance coefficient in gravity equations in international trade. *Review of International Economics*, *24*(5), 1167–1177.

Seidel, A. (2023). A global map of amenities: Public goods, ethnic divisions and decentralization. *Journal of Development Economics*, *164*, 103113.

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

Storeygard, A. (2016). Farther on down the road: Transport costs, trade and urban growth in Sub-Saharan Africa. *The Review of Economic Studies*, *83*(3), 1263–1295.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brandily, P., & Rauch, F. (2024). Within-city roads and urban growth. *Journal of Regional Science*, 1–29. https://doi.org/10.1111/jors.12699