Geometry, cities and lights: Evidence from the developing world

Rodrigo González Laiz*

Bachelor Thesis (TFG)

June 2020

Abstract

Using nighttime satellite images, this paper studies the effect of city shape on economic performance in the developing world, particularly in Sub-Saharan Africa and Asia. I find that more compact cities have better economic outcomes in Asia. However, I am not able to draw a definite conclusion in the case of Africa which may be evidence of a distinct urban development pattern. Furthermore, I find nonlinear (decreasing) effects in the case of India and China.

Advisor: Clara Santamaría

Economics Department, UC3M.

*I want to thank Clara Santamaría for her valuable comments and guidance. I am also grateful to Mariaflavia Harari for her helpful suggestions.

Contents

1 Introduction

Urbanization is an unstoppable process and a crucial part of socio-economic development. Nowadays there are more people living in cities than in rural areas and this trend is very likely to continue. But why do cities exist? Why do humans only occupy a small percentage of all land available on earth? These are some of the oldest and most important questions in urban economics and economic geography.

A helpful way of thinking about a city is as the outcome of agglomeration economies against the costs of urban congestion, i.e. pollution. More specifically, agglomeration economies refer to the benefits that arise when firms and people locate together. Three mechanisms are usually identified behind agglomeration economies: *matching* (workers and firms), *sharing* (input-output linkages) and *learning* (localized information and knowledge spillovers) (Duranton & Puga 2004). The existence of agglomeration economies is usually reflected in urban wage premiums although other sources such as local institutions, public capital and geographic amenities (Combes et al. 2010) or selection of more productive workers (Combes et al. 2008) are also relevant to explain higher urban wages.

Although urbanization happens everywhere, there are big spatial differences, often associated with income. In fact, a well-documented fact in urban economics is that the urbanization growth rate peaks at per capita income levels of around \$3.000 - \$5.000. Accordingly, in 2010 Sub-Saharan Africa and Asia had urbanization rates below 50% but approximately 80% of people in Europe, North America and Latin America lived in cities.

Urbanization has accelerated throughout history. Jedwab et al. (2014) estimates that developing countries have experienced the same urbanization growth in half the time, comparing Europe in the period 1800-1910 and developing countries in the period 1950-2010. Moreover, according to the United Nations (UN 2015), urban population in developing countries is expected to increase in 2 billion by 2050 mainly driven by Asia (1 billion) and Africa (800 million). This presents many opportunities as well as challenges for these two continents.

There is an extensive literature that studies the relationship between urban concentration, city size and economic growth (Spence et al. 2008, Henderson 2003). The conclusion of this literature is that there exists a strong positive link between urban concentration and economic growth *when cities are able to exploit the benefits of scale and specialization*, which results in higher productivity and job creation. However, the relationship between urban concentration and economic growth in developing countries is not robust. Particularly, evidence from Africa seems to be pointing in the opposite direction (Collier 2017, Castells-Quintana 2017) although more research is needed.

The goal of this paper is to empirically study a somewhat forgotten property of cities: their *urban shape*. Urban shape, also referred as urban geometry, is defined as the geometric characteristics of the boundary of the city. There are hundreds of metrics to quantify characteristics of shapes but in this paper I will focus on *compactness*[1]. Urban planners have always recognized the importance of urban shape as a determinant of urban productivity and urban livability through two channels: service provision and transit accessibility. Service provision refers to the delivery of public services such as water or electricity whereas transit accessibility is related to the transportation system in a broad sense. Therefore, cities with "better" shapes are usually defined by shorter distances within the city which can potentially affect commuting times and public service provision.

To the best of my knowledge, the first paper that causally estimates the effect of city shape is Harari (forthcoming). In her excellent paper she finds that more compact Indian cities are characterized by larger populations, lower wages and higher housing rents which imply that compactness is valued as a consumption amenity. In other

[1] In the next section I will clearly define what compactness is and how to measure it.

related work Duque, Lozano-Gracia, Patino & Restrepo (2019) study the impact of urban form in Latin American cities and they find that the shape of the city, the inner-city connectedness and fullness have a statistically significant influence on the productivity level of the city. Similarly, Tewari & Godfrey (2016) find there is a clear link between more compact urban growth and stronger economic performance in India.

In this paper I attempt to contribute to this recent literature by estimating the causal effect of urban shape on economic performance (measured by the density of light per km^2) in developing countries. In particular, I provide evidence for Sub-Saharan Africa (hereafter Africa) and most of Asia with a total of 68 countries. I choose to study these regions because, as said before, Africa and Asia are the regions which have more to gain from a successful urbanization.

The results show that more compact cities are associated with better economic performance in Asia, as found in other research, but that the relationship in Africa is not clear. This evidence is in line with the thesis that African urbanization is different to what we have observed before.

One of the main challenges of the paper is the likely endogeneity of city shape. In order to address this issue, I perform an instrumental variable estimation with fixed effects using a synthetic instrument developed in Harari (forthcoming). The instrument, which exploits both cross-sectional and temporal variation, relies on exogenous changes that the shape of the city experiences over time due to topographic obstacles (steep terrain, water, ...etc.) along its expansion path. I also perform an instrumental variable quantile regression to uncover potential distributional effects.

The rest of the paper is organized as follows. Section 2 provides a description of the data and the techniques used to measure both urban shape and economic performance. Section 3 presents descriptive statistics. Section 4 describes the methodology in detail. Section 5 presents the results. Section 6 concludes and points to future research.

2 Data

Urban geometry

Satellite images are widely used to study multiple topics such as urbanization (Henderson 2003, Burchfield et al. 2006), poverty measurement (Engstrom et al. 2017) or population modelling (Schiavina et al. 2019) among many others. In Economics, the work of Henderson et al. (2012) seems to be a turning point when researchers have started to exploit this source of information to its full potential, although previous research using satellite images exists. A review of the applications of satellite images in Economics can be found in Donaldson & Storeygard (2016).

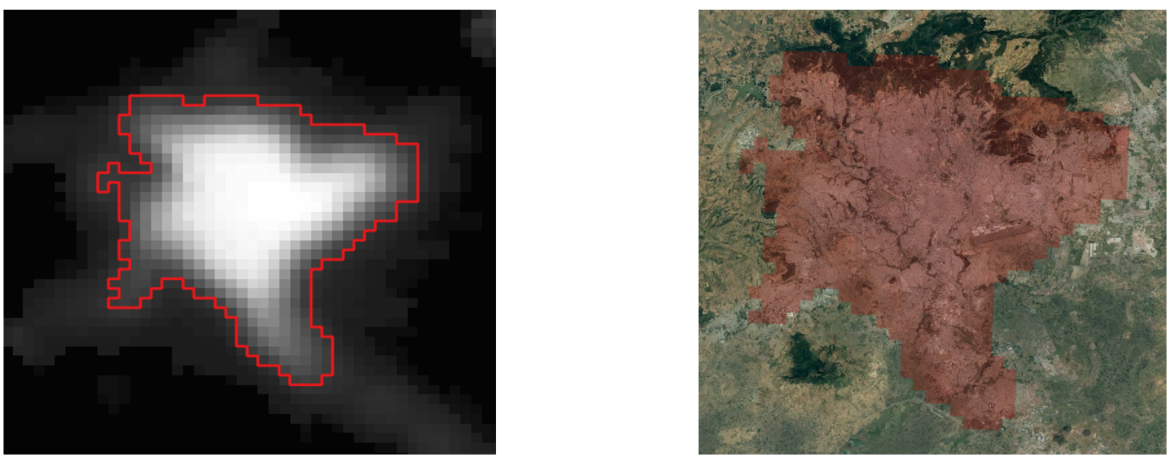
To measure the urban shape of the cities in my sample I use the Defense Meteorological Satellite Program/Operational Linescane System stable night-time light (DMSP-OLS NTL) dataset, which consists on a series of satellite images taken during nighttime that record the intensity of light, have global coverage and span two decades (1992-2012) with a resolution of around 1 km.

There are multiple ways to extract a city's geometry but they often are grouped into two categories: threshold based and classification based. I use a threshold based approach due to the possible blooming effect of the NTL data. This approach consists in determining the status (urban vs non urban) of a pixel based on if the value of the pixel is above or below a specified threshold. With this idea in mind, I consider a city to be the *patch of contiguous lighted pixels that lie above a predetermined threshold*. Figure 1 provides an example for the city of Addis Ababa (Ethiopia). The left image shows the NTL raw image and the extracted geometry on top of it and the right image shows the extracted geometry on top of a daytime satellite image from Google Maps.

One important issue is that I am examining a very large and heterogeneous set of countries and therefore it is not suitable to use the same threshold in every location to define what is a city. Therefore, I compute an "optimal" threshold for each country following the methodology implemented in Roberts et al. (2015) which uses an additional global layer and a histogram analysis to identify the best threshold based on the concept of user's accuracy.

Another issue worth mentioning is that the NTL satellite images are not ready to use immediately. Specifically, there are four problems that need be addressed before they are used to analyze urban change: inter-calibration, temporal patterns, blooming and saturation. There is an extensive literature in the remote sensing field that tries to account for these problems (Li & Zhou 2017). A detailed explanation about how to deal with each problem as well as other details about the NTL satellite data can be found in the appendix.

Figure 1: Urban geometry of Addis Ababa (Ethiopia)



Source: Own elaboration with NTL data.

Shape metrics

I am interested in the shape of cities, sometimes called urban geometry. There are multiple ways to describe the shape of a polygon (city), but in this paper I focus on *compactness*. Compactness is defined as the extent to which a polygon's shape departs from that of a circle, the most compact shape for a given area. In particular, I use two compactness metrics (cohesion and spin) which are borrowed from Angel et al. (2010). The cohesion index is defined as the average (Euclidean) distance between all pairs of interior points whereas the spin index is defined as the average of the squared of the (Euclidean) distance between all interior points and the centroid. Figures 5 and 6 provide a graphical explanation of these two measures.

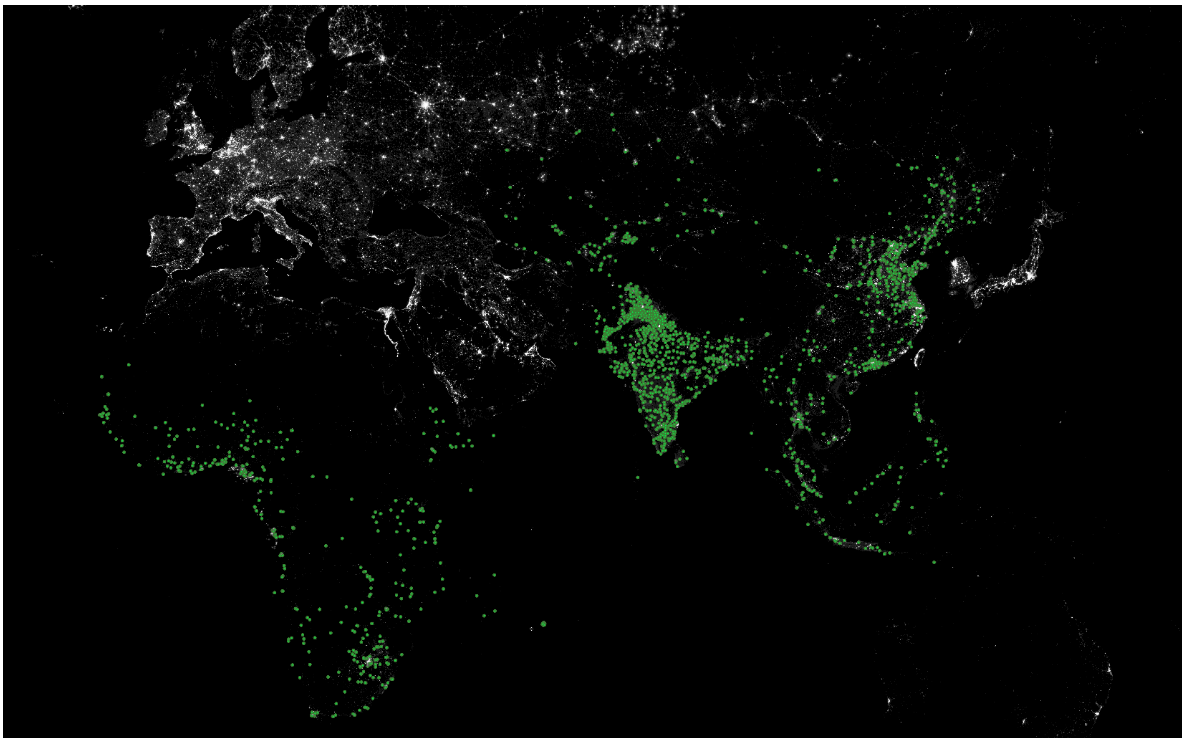
As Harari (forthcoming) points out, any compactness metric based on distances within a polygon is mechanically correlated with the polygon's area. Therefore, if I want to disentangle the effect of the geometry from city size, I need to control for the area that cities occupy. In this paper I employ normalized versions of the two compactness metrics, which makes them invariant to the area of the polygon (size of the city). Moreover, because cohesion and spin have a high degree of correlation I define the variable *compactness* as the average between them ($compactness = \frac{1}{2}cohesion + \frac{1}{2}spin$). The variable *compactness* takes values in the interval (0,1], where 1 corresponds to the most compact shape (perfect circle). I will use this variable throughout my analysis.

Economic performance

I measure economic performance using a variable defined as the night light density divided by area in square kilometers, which is retrieved from NTL satellite images. There has been recent work that concludes that night-time lights provide a good proxy for income because consumption of goods during the night require light, and as per capita income rises, so do light emissions (Henderson et al. 2012). In other work, Michalopoulos & Papaioannou (2013) show that night light density per km^2 has a strong correlation with development and household wealth indicators. Similarly, note that this measure of economic performance is plausibly the best measure of economic performance we have today at the sub-national level in developing countries because official information of this nature is very scant and unreliable.

In summary, I have a panel of around 1630 cities covering most of Africa (sub Saharan Africa) and Asia in the period 1992-2012 with variables describing their urban shape, area, and economic performance. Figure 2 shows graphically where cities are located.

Figure 2: Location of cities in space



Source: Own elaboration with NTL data.

3 Descriptive evidence

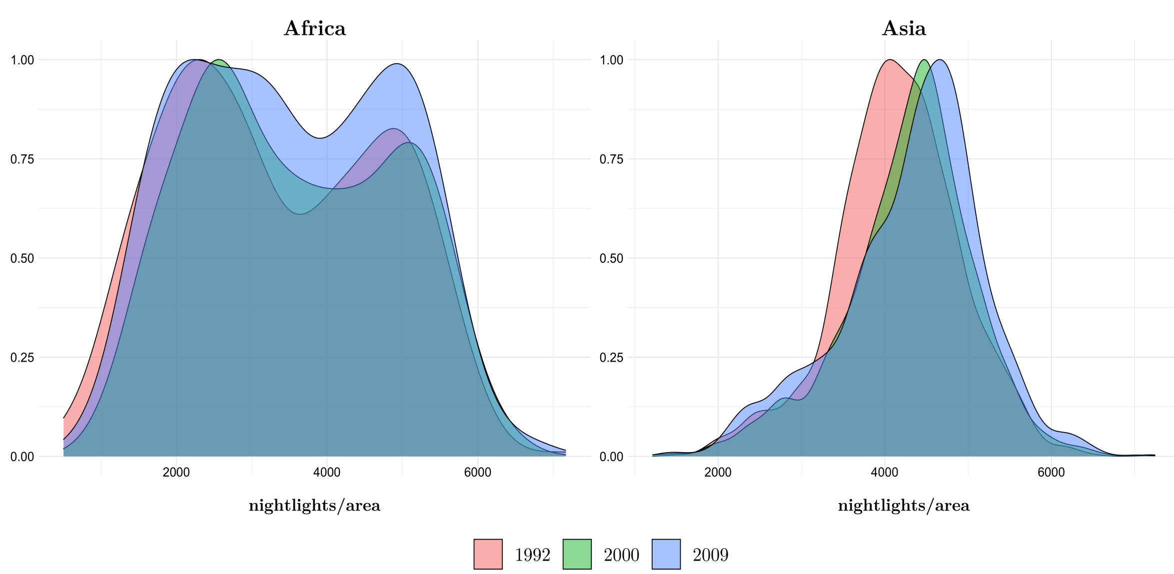
In this section I present stylized facts regarding the economic performance, size and compactness of the cities in my sample. Table 1 describes the main variables in selected years 1992, 2000 and 2012. Table 2 disaggregates the same information by subregion.

The first take away is that *cities have become richer and bigger*. Specifically, table 1 shows an increase of 9.3% in average density of light per km^2 during the period 1992-2012 and a rise in urban area from 64.5 km^2 in 1992 to 214 km^2 in 2012, a 228% increase.

Moreover, table 2 shows that on average Asian cities have approximately 22% more density of light per km^2 than Africa. Furthermore, the regions with the lowest density of light per km^2 are Central Asia, East Africa and Central Africa. These results are reassuring because they are consistent with aggregated official statistics and, therefore, give credibility to the use of nighttime lights as a proxy for economic performance.

Figure 3 displays the scaled kernel density of the density of light per km^2 for Africa and Asia during the period 1992-2012. These density plots indicate that the distributions of nighlights are quite different. On the one hand, Asia has a concentrated distribution around the median, which increased from 4162 to 4654 in the period 1992-2012, an 11% increase. On the other hand, Africa has a more spread distribution with two modes, which indicates higher inequality of income per capita between cities. Also median density of light in Africa went up from 3096.6 in 1992 to 3710 in 2012, a 20% increase. Similarly, figure 7 displays the non scaled kernel density plot which accounts for the difference in the number of cities between continents (386 in Africa and 1250 in Asia).

Figure 3: Density of light per km^2 in Africa and Asia



Looking at table 2, we can observe that African cities are larger than Asian cities with an average city size of 173.8 and 111.8 km^2 respectively. This could be related to the fact that Africa's urban development is dominated by primate cities (Henderson & Kriticos 2018). The subregion with the smallest city size on average is South Asia with 67.9 km^2, which is interesting given that India is part of South Asia. On the other hand, the subregions with the largest urban areas are Southern Africa with 271.5 km^2 and surprisingly, Central Asia with 249.3 km^2

In regards to urban shape, table 1 indicates that cities are more round than elongated and that have become less compact over time although the change is small. Specifically, cities became more compact between 1992 and 2000 but their shape deteriorated in the period 2000-2012 which could indicate a trend of low density growth similar to what other studies have found (Huang et al. 2007, Duque, Lozano-Gracia, Patino, Restrepo & Velasquez 2019). Table 2 shows that Asian cities are on average more compact than African cities (0.903 vs 0.882). Moreover, the subregion with the least compact cities (0.851) is South East Asia, probably due to the presence of a large number of islands.

Table 1: Descriptive statistics

	year	count	mean	std	min	25%	50%	75%	max
density	1992	1636	3964.1	1006.6	501.5	3491.0	4088.5	4625.3	7177.0
	2000	1636	4119.9	990.2	1007.0	3589.1	4288.0	4775.8	7241.9
	2012	1636	4331.9	1027.4	1013.1	3765.1	4572.3	5024.0	7137.3
area	1992	1636	64.5	198.9	0.9	8.9	20.2	46.9	4357.9
	2000	1636	106.8	323.8	0.9	20.2	37.6	77.3	5153.1
	2012	1636	214.1	775.8	0.9	33.9	67.0	159.1	8070.4
compact	1992	1636	0.891	0.088	0.425	0.857	0.921	0.954	0.992
	2000	1636	0.905	0.098	0.341	0.886	0.944	0.966	0.992
	2012	1636	0.888	0.103	0.406	0.855	0.929	0.957	0.994

Notes: density refers to the density of light per km^2; area is measured in km^2; compact refers to the compactness index described in section 2.

4 Methodology

In this section I describe the empirical strategy used in the paper. I show that using a instrumental variable estimation along with a fixed effects model, I am able to causally estimate the effect of urban shape on economic performance. Moreover, I motivate and explain why it is important to estimate an instrumental variable quantile regression model to uncover potential distributional effects.

Urban shape is determined both by exogenous factors such as topographic obstacles as well as endogenous factors such as policy choices. Therefore, one of the main challenges of this paper is to account for the potential endogeneity of urban shape and calculate credible causal estimates.

Urban shape can be endogenous for a variety of reasons and it is not clear what will be the sign of the bias. On the one hand, cities with stronger state capacity and well-functioning institutions, which usually have better economic outcomes, are likely to have better strategies regarding urban planning. This would result in a positive correlation between compact cities and economic performance but it would no be due to the effect of compactness. On the other hand, cities that grow faster may have worse shapes because of the difficulty of managing fast growth development. This would result in a negative correlation between economic performance and compactness but it would be unrelated to the true effect of compactness.

To address the issue of endogeneity, I employ an instrumental variables estimation with fixed effects for which I borrow an instrument developed by Harari (forthcoming). The (synthetic) instrument, which exploits both cross-sectional and temporal variation, relies on exogenous changes that the shape of the city experiences over time due to topographic obstacles along its expansion path. Specifically, the instrument tries to isolate the changes in urban shape that are *only* driven by topography and to exploit cross-sectional variation, the instrument *predicts* urban growth following a mechanical model of city expansion.

This results in a *potential footprint* for each city and year that represents the possible shape that a city could have had if it were only constrained by topography. Finally, the instrument is defined as the compactness index of the potential footprint which I use to instrument the compactness index of the observed city shape.

The estimation strategy is given by

$$log(Y_{it}) = \beta compactness_{it} + \gamma_i + \gamma_t + u_{it}$$

$$compactness_{it} = \delta potential\ compactness_{it} + \alpha_i + \alpha_t + \epsilon_{it}$$

The variable Y is defined as the density of light per km^2 (the proxy I use for economic performance), *compactness* refers to the compactness index (described in section 2) of the observed city shape and *potential compactness* measures the compactness index of the potential footprint (instrument). In both specifications I include city fixed effects (γ_i, α_i) to control for any time invariant unobserved heterogeneity as well as time fixed effects (γ_t, α_t). The parameter we are interested is β which plausibly estimates the causal effect of city compactness on economic performance. Harari (forthcoming) also discusses possible threats to the identification strategy such as the possible direct effect of geography and initial conditions or pre-existing trends but she concludes they do not affect the validity of the instrument.

4.1 Instrument construction

The following steps need to be taken to create the instrument, assuming a *common average expansion rate across all cities*:

- Estimate the following regression:

$$log(area_{it}) = \gamma_i + \gamma_t + v_{it}$$

- Compute the predicted area $ar\hat{e}a_{it}$ of city i and year t and calculate the predicted radius [2] as:

$$\hat{r_{it}} = \sqrt{\frac{ar\hat{e}a_{it}}{\pi}}$$

- Estimate the growth rate of the radius $\hat{r_{it}}$ and create concentric circles according to this rate around the minimum bounding circle of the city in the first year in your sample (in my case 1992).

- For each circle, extract the largest patch of contiguous developable land (land that has a slope below 15% and is not a water body), also referred as the potential footprint, and compute its compactness index as defined in section 2[3].

4.2 Instrumental Variable Quantile Regression (IVQR)

Classical regression focuses on estimating the conditional expectation function $E[Y|X]$. This function can be quite complex but it is always specific to a certain location of the conditional distribution of Y. Quantile regression (Koenker & Bassett Jr 1978) on the other hand estimates conditional quantile functions $Q_y(\tau|X)$ that allow us to study the conditional distribution of Y at different locations. This is extremely useful as it gives us a more complete picture of the relationship between the variables in our analysis.

To estimate classical regression we minimize the usual average quadratic loss given by

$$\hat{\beta}_{OLS} = \underset{\beta}{argmin} \sum_{i=1}^{N} (Y_i - X_i'\beta)^2$$

To estimate quantile regression we minimize an asymmetric loss function, called *check function* given by

$$\hat{\beta}_{\tau} = \underset{\beta}{argmin} \sum_{i=1}^{N} \rho_{\tau}(Y_i - X_i'\beta)$$

[2]The area of a circle has the following formula: A = πr^2

[3]Detailed information about the instrument can be found in Harari (forthcoming)

$$\rho_\tau = u(\tau - I(u < 0))$$

Therefore the minimization problem of the quantile regression can be rewritten as:

$$\hat{\beta}_\tau = \underset{\beta}{argmin} \sum_{y_i > X_i'\beta} \tau |Y_i - X_i'\beta| + \sum_{y_i \leq X_i'\beta} (1 - \tau)|Y_i - X_i'\beta|$$

The idea of quantile regression has many extensions (Chen et al. 2018) but one of the more useful is its combination with instrumental variables methods. There are multiple ways of building a quantile model with endogeneity but in this paper I use the methodology developed in Chernozhukov & Hansen (2005, 2006, 2008).

In this set of papers Chernozhukov & Hansen (2005, 2006, 2008) provide the necessary conditions to nonparametrically identify quantile treatment effects (QTE) using instrumental variables. The most important assumption is the restriction of how structural errors (rank variables) vary across different potential states of the endogenous variables which in its strongest form is called rank invariance.

In this paper I use the method developed by Harding & Lamarche (2009) which combines panel data with endogenous independent variables using the methodology developed in Chernozhukov & Hansen (2008). One important caveat of this method is that in the IVQR model for panel data, the individual effects ($\alpha(\tau)$) are indexed by τ, the τ-th quantile of the conditional distribution of Y. This means that it is not appropriate to think about them as "fixed effects" because the estimated α's will change over the distribution of Y. Therefore the authors propose to think about them as a hybrid between fixed effects and random effects because they allow for the presence of individual factors that are correlated with the independent variables but they are also allowed to change over the conditional distribution of Y.

5 Results

This section reports the results obtained. First I describe the first stage of the instrumental variables estimation to assess whether the instrument is relevant. Then I show the difference between OLS and IV and finally I describe the results of the instrumental variable quantile regression model for India and China.

Table 3 reports the first stage regression results which relates urban geometry (endogenous variable) to the instrument described in section 4. I find that the instrument is relevant for the whole sample (column 1) because the F statistic is above the usual threshold of 10. However, this result hides important information. Specifically, the instrument is relevant in Asia (F statistic above 10) with the exception of South East Asia (columns 2, 4-7), including India and China (columns 11-12) but it *is not* relevant (F statistic below 10) for Africa (columns 3, 8-10, 14-15).

When the instrument(s) are not correlated wit the endogenous variable(s), we normally have a problem of *weak instruments*, which means that the instrument does not produce enough exogenous variability in the endogenous variable. The problem with weak instruments is that the IV coefficients will be biased and usual inference is not valid. There is a growing literature of weak-IV inference that tries to account for this problem (Andrews et al. 2019) but in this case the F statistics are so low in the case of Africa that I do not think much can be done. Therefore, unfortunately, I will not be able to draw any definite conclusion about the effect of city shape in Africa.

5.1 Economic performance: OLS and IV

Table 4 and table 5 presents OLS and IV estimates respectively. The OLS estimates show that there is a positive but non significant correlation between city compactness and economic performance in the whole sample (column 1). Nevertheless, similarly to the first stage regressions, there is a lot heterogeneity. On the one hand, there is a positive and significant correlation in most of Asia (columns 2, 4-7, 11-12) with the exception of South East Asia (column 5) which is negative but non significant. On the other hand, the correlation in Africa is mostly negative although non significant (with the exception of Southern Africa). Nevertheless, some selected countries such as South Africa and Nigeria show a negative and significant correlation which means that in these countries more compact cities on average are associated with worse economic performance.

The magnitude of the coefficients is better understood in terms of standard deviations (they can be found in table 2). As an example, an increase in one standard deviation in the compactness index of Asian cities (table 4, column 2), is associated with an increase in economic performance of (0.05*0.095) 0.0475%.

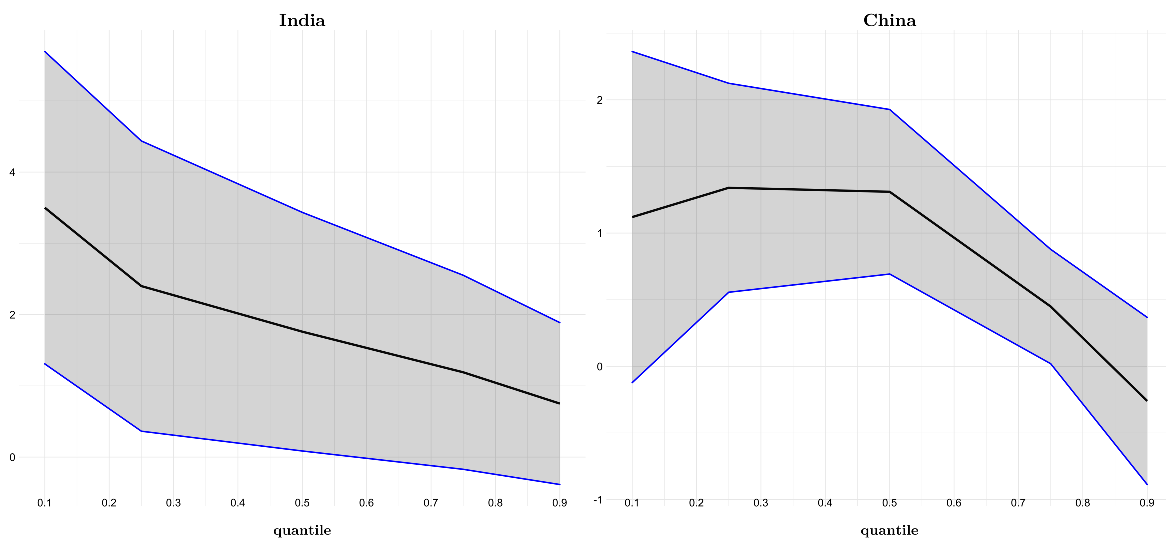
The instrumental variables estimates can be found in table 5. As discussed before, the only estimates that are reliable are the ones with a strong first stage (columns 1,2,4,6-7,12-13) which mainly correspond to Asia. Therefore I can only reliably estimate the effect of city shape in Asia. Given this, I find the effect of city compactness in Asia is indeed positive and bigger than the OLS estimate but with the same sign. In particular, an increase in one standard deviation of the compactness index in Asia (table 5, column 2) increases on average the economic performance in (1.362*0.095) in 12.94%. However, there are big differences between regions. For example, the region with the highest effect is South Asia (SA) with a 36.74% increase whereas Central is the region with the lowest effect with only a 0.39% increase.

5.2 IVQR: The case of India and China

Table 6 presents the results of the instrumental variable quantile regression model (IVQR) for India and China. Figure 4 shows the results graphically. The effect of city shape is not constant and it varies quite a lot depending on the location in the distribution of the density of light per km^2 (dependent variable). In other words, the instrumental quantile regression estimates indicates that there is substantial heterogeneity in the effect of city shape on economic performance.

More specifically, I find that the effect is smaller in higher quantiles both in India and China. Nevertheless, the results are uniformly positive and significant which allow to conclude that more compact cities positively affects their economic performance across the entire distribution in both countries.

Figure 4: IVQR results in India and China



5.3 Discussion of the results

First stage results show a positive and significant correlation between the shape of the actual footprint and the shape of the potential footprint in Asia. Intuitively, it makes sense to observe a positive correlation between these two variables because if there exist topographic obstacles one would think that a city would experience a more leapfrog urban development pattern. Africa's results on the other hand are somewhat surprising because the correlation is sometimes positive (table 5, columns 8, 10, 14, 15), sometimes negative (table 5, columns 9, 11) but *it is never significant.* This means that I find no relationship between how constrained the land is and the shape of the city which it is in itself interesting.

It is important to understand this result. One explanation would be that when constructing the instrument I use an average expansion rate across *all* cities, both in Africa and in Asia. However, there may be heterogeneity in the rate at which cities grow and this could affect the relevance of the instrument. Another reason is that perhaps the NTL satellite images are not precise enough to truly measure the shape of cities in Africa because some parts of African cities may not have adequate lighting, which could bias my results. Finally, there may be another reason that future research could try to explain.

OLS estimates are also interesting because they show a difference between Africa and Asia as well. In Asia, the correlation between city compactness and economic performance is on average positive whereas the correlation in Africa is almost always negative (except Southern Africa) although non significant. Finally, IV estimates show that there is a positive causal effect of city compactness on economic performance on average in Asia, similar to what other work has found (Tewari & Godfrey 2016, Duque, Lozano-Gracia, Patino & Restrepo 2019).

Evidence presented in this paper may support the idea that African urban development is different. Recent literature has found that Africa's urbanization has not been very successful and that many African cities are in a low-development trap due to multiple reasons such as low productivity, decrease in livability, lack of functional land markets or low infrastructure investments (Collier 2017, Lall 2017). In fact, I would hypothesize that city shape will not have a major effect in African cities until the levels of infrastructure reach a certain threshold, but this is out of the scope of the paper.

6 Conclusion and future research

In this paper I study the effect of city shape on economic performance in the developing world, particularly in Africa (Sub-Saharan Africa) and Asia. Using an instrumental variables model with fixed effects I am able to account for the endogeneity of urban shape and I find a positive effect of city shape on economic performance in Asia but I am not able to draw definite conclusion in the case of Africa due to weak instruments. Moreover, using an instrumental variables quantile regression model I find nonlinear (decreasing) effects in the case of India and China.

One important avenue for future research should be the study of the specific case of Africa. The evidence presented in this paper suggests that Africa's urbanization is different, and that maybe city shape is not as important as in other locations, but of course this is not conclusive. I would also like to highlight that satellite images provide invaluable knowledge that no other source of information can and I think economists and social scientist in general should use them more widely, specially to understand dynamic global processes.

A APPENDIX

A.1 DMSP-OLS NTL satellite images

The Defense Meteorological Satellite Program/Operational Linescane System stable night-time light (DMSP-OLS NTL) dataset contains a series of nighttime satellite images that record the intensity of light, have global coverage and they are available in the period 1992-2012. Each pixel is described by a digital number between 0 and 63 where 63 reflects the maximum amount of light possible. In this paper I use the stable or ordinary product (NTL) as opposed to radiance calibrated product (RC) because the ordinary product has a larger time span.

To take advantage of their full potential, some steps need to be taken before they are used. In particular four issues arise: intercalibration of time series, temporal adjustment, blooming effect and saturation of luminosity. A review can be found in Li & Zhou (2017).

One problem of the images is that they are not comparable due to absence of on-board calibration. This means that without applying some intercalibration method it is not possible to track urban change through time. Most of the intercalibration approaches are based on the framework proposed by Elvidge et al. (2009) which are based on (1) selection of the reference region (2) determination of the reference satellite and year for calibration and (3) model development for intercalibration. In this paper I used the method developed by Zhang et al. (2016) which consists [4] on "selecting data points along a ridgeline—the densest part of a density plot generated between the reference image and the target image—and then use those data points to derive calibration models to minimize inconsistencies in the NTL time series".

[4] The intercalibrated satellite images are available on https://urbanization.yale.edu/data. I thank the authors for making them public.

The next step was to create annual composites when two satellites recorded information for the same year by averaging the corresponding images. The idea is to remove any unstable lit pixels that may be wrongly specified Liu et al. (2012). This is sometimes called intra-annual composition. Moreover, I performed a two-way inter-annual series correction to be able to consistently analyze urban change in rapidly evolving regions such as India or China as in Liu & Leung (2015).

A.2 Tables

Table 2: Descriptive statistics

			density		**area (km2)**		**compactness**		**pcompactness**	
	region	count	mean	std	mean	std	mean	std	mean	std
Asia	ALL	26254	4311.6	802.8	111.8	422.6	0.903	0.095	0.968	0.052
	CA	1512	3295.2	595.2	249.3	316.7	0.894	0.098	0.981	0.038
	SA	14830	4164.0	747.0	67.9	150.1	0.919	0.084	0.977	0.041
	EA	7014	4624.3	487.1	167.8	725.2	0.893	0.097	0.957	0.068
	SEA	2898	4840.3	1028.2	129.5	387.1	0.851	0.118	0.946	0.057
Africa	ALL	8106	3546.4	1344.4	173.2	568.0	0.882	0.106	0.954	0.064
	CAF	1260	2834.2	1226.5	122.3	191.7	0.866	0.118	0.937	0.070
	EAF	2100	2809.2	1034.8	157.3	283.9	0.886	0.102	0.946	0.073
	WAF	2373	3754.3	1094.9	116.2	317.6	0.896	0.104	0.967	0.043
	SAF	2373	4369.1	1326.7	271.5	946.6	0.872	0.103	0.957	0.066

Notes: CA: Central Asia; SA: South Asia; EA: East Asia; SEA: South East Asia; CAF: Central Africa; EAF: East Africa; WAF: West Africa; SAF: Southern Africa. density refers to the density of light per km^2; area is measured in km^2; compact refers to the compactness index described in section 2; pcompactness is the compactness index of the potential footprint.

Table 3: First Stage

	Dependent variable: compactness														
	ALL	Asia	Africa	SA	SEA	EA	CA	CAF	WAF	SAF	EAF	India	China	Nigeria	South Africa
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
potential compactness	0.078	0.101*	0.027	0.084	0.026	0.212***	0.112***	0.083	-0.117	0.106	-0.054	0.188***	0.211***	0.414	0.058
	(0.052)	(0.053)	(0.046)	(0.069)	(0.122)	(0.080)	(0.019)	(0.088)	(0.163)	(0.082)	(0.064)	(0.067)	(0.080)	(0.599)	(0.130)
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	34,360	26,254	8,106	14,830	2,898	7,014	1,512	1,260	2,373	2,373	2,100	11,403	6,972	819	1,533
F Statistic	29.208	33.285	1.235	11.751	0.300	39.538	83.103	1.643	2.892	6.085	1.968	53.659	38.981	3.275	0.858

Notes: This table shows the correlation between the instrument (potential compactness) and the endogenous variable (compactness). SA: South Asia; SEA: South East Asia; EA: East Asia; CA: Central A; CAF: Central Africa; WAF: West Africa; SAF: Southern Africa; EAF: East Africa. Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 4: OLS

	Dependent variable: log(density of light per km^2)														
	ALL	Asia	Africa	SA	SEA	EA	CA	CAF	WAF	SAF	EAF	India	China	Nigeria	South Africa
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
compactness	0.015	0.050**	-0.102	0.123***	-0.083	0.067***	0.156	-0.168	-0.137	0.019	-0.178	0.200***	0.066***	-0.480***	-0.103**
	(0.022)	(0.020)	(0.068)	(0.034)	(0.069)	(0.016)	(0.106)	(0.152)	(0.121)	(0.070)	(0.228)	(0.031)	(0.016)	(0.181)	(0.052)
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	34,360	26,254	8,106	14,830	2,898	7,014	1,512	1,260	2,373	2,373	2,100	11,403	6,972	819	1,533

Notes: This table shows the correlation between the compactness index and the density of light per km^2 as a proxy for economic performance. SA: South Asia; SEA: South East Asia; EA: East Asia; CA: Central A; CAF: Central Africa; WAF: West Africa; SAF: Southern Africa; EAF: East Africa. Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 5: IV

	Dependent variable: log(density of light per km²)														
	ALL	Asia	Africa	SA	SEA	EA	CA	CAF	WAF	SAF	EAF	India	China	Nigeria	South Africa
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
compactness	3.630***	1.362***	21.311	4.374***	-1.233	0.885***	0.398***	5.594	1.968	4.514**	-19.767	1.355***	0.881***	-2.795*	2.118
	(0.715)	(0.286)	(19.365)	(1.289)	(2.756)	(0.156)	(0.128)	(4.871)	(1.698)	(1.892)	(14.152)	(0.222)	(0.156)	(1.580)	(2.451)
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	34,360	26,254	8,106	14,830	2,898	7,014	1,512	1,260	2,373	2,373	2,100	11,403	6,972	819	1,533

Notes: This table shows the IV estimates of the effect of city compactness on economic performance. SA: South Asia; SEA: South East Asia; EA: East Asia; CA: Central A; CAF: Central Africa; WAF: West Africa; SAF: Southern Africa; EAF: East Africa. Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01.

Table 6: IVQR

	quantile	coefficient	std. error	t-value
India	0.10	3.500	1.2510	2.797762
	0.25	2.400	1.0800	2.222222
	0.50	1.760	0.7300	2.410959
	0.75	1.190	0.4827	2.465299
	0.90	0.751	0.3360	2.235119
China	0.10	1.120	0.6340	1.766562
	0.25	1.340	0.4001	3.349163
	0.50	1.310	0.3150	4.158730
	0.75	0.449	0.2190	2.050228
	0.90	-0.260	0.3200	-0.812500

Notes: Results of the IVQR model. Robust standard errors.

A.3 Figures

Figure 5: Cohesion index

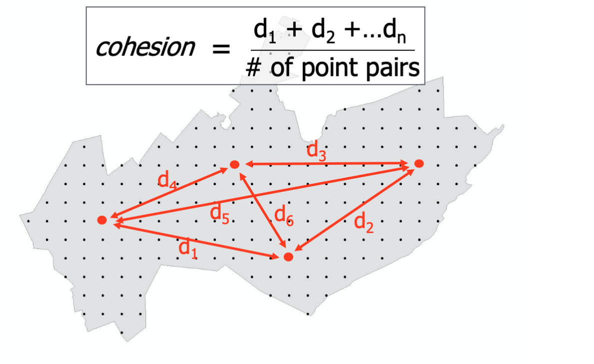
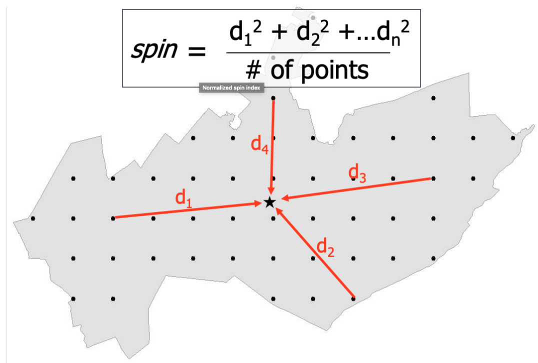
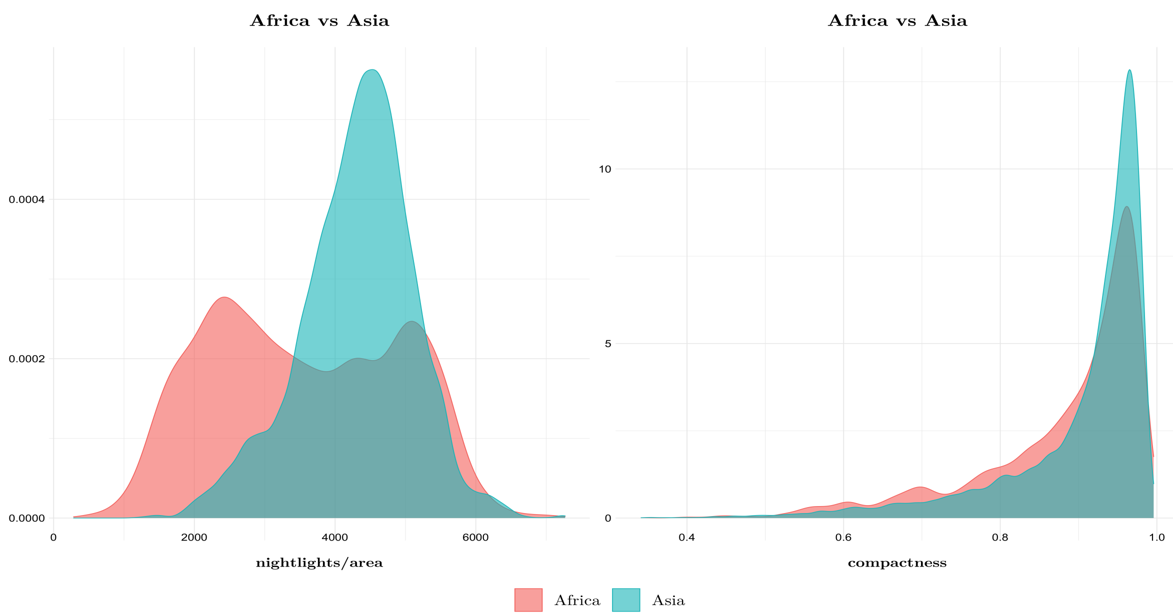


Figure 6: Spin index



Source: Images taken from Angel et al. (2010)

Figure 7: Density of light per km^2 and compactness index



References

Andrews, I., Stock, J. H. & Sun, L. (2019), 'Weak instruments in instrumental variables regression: Theory and practice', *Annual Review of Economics* **11**, 727–753.

Angel, S., Parent, J. & Civco, D. L. (2010), 'Ten compactness properties of circles: measuring shape in geography', *The Canadian Geographer/Le Géographe canadien* **54**(4), 441–461.

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

Castells-Quintana, D. (2017), 'Malthus living in a slum: Urban concentration, infrastructure and economic growth', *Journal of Urban Economics* **98**, 158–173.

Chen, L., Dolado, J. & Gonzalo, J. (2018), 'Quantile factor models'.

Chernozhukov, V. & Hansen, C. (2005), 'An iv model of quantile treatment effects', *Econometrica* **73**(1), 245–261.

Chernozhukov, V. & Hansen, C. (2006), 'Instrumental quantile regression inference for structural and treatment effect models', *Journal of Econometrics* **132**(2), 491–525.

Chernozhukov, V. & Hansen, C. (2008), 'Instrumental variable quantile regression: A robust inference approach', *Journal of Econometrics* **142**(1), 379–398.

Collier, P. (2017), 'African urbanization: an analytic policy guide', *Oxford Review of Economic Policy* **33**(3), 405–437.

Combes, P.-P., Duranton, G. & Gobillon, L. (2008), 'Spatial wage disparities: Sorting matters!', *Journal of urban economics* **63**(2), 723–742.

Combes, P.-P., Duranton, G., Gobillon, L. & Roux, S. (2010), Estimating agglomeration economies with history, geology, and worker effects, *in* 'Agglomeration economics', University of Chicago Press, pp. 15–66.

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

Duque, J. C., Lozano-Gracia, N., Patino, J. E. & Restrepo, P. (2019), 'Urban form and productivity: What is the shape of latin american cities?'.

Duque, J. C., Lozano-Gracia, N., Patino, J. E., Restrepo, P. & Velasquez, W. A. (2019), 'Spatiotemporal dynamics of urban growth in latin american cities: An analysis using nighttime light imagery', *Landscape and Urban Planning* **191**, 103640.

Duranton, G. & Puga, D. (2004), Micro-foundations of urban agglomeration economies, *in* 'Handbook of regional and urban economics', Vol. 4, Elsevier, pp. 2063–2117.

Elvidge, C. D., Ziskin, D., Baugh, K. E., Tuttle, B. T., Ghosh, T., Pack, D. W., Erwin, E. H. & Zhizhin, M. (2009), 'A fifteen year record of global natural gas flaring derived from satellite data', *Energies* **2**(3), 595–622.

Engstrom, R., Hersh, J. & Newhouse, D. (2017), *Poverty from Space: Using High-Resolution Satellite Imagery for Estimating Economic Well-Being*, The World Bank.
URL: *https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-8284*

Harari, M. (forthcoming), 'Cities in bad shape: Urban geometry in india', *American Economic Review* .

Harding, M. & Lamarche, C. (2009), 'A quantile regression approach for estimating panel data models using instrumental variables', *Economics Letters* **104**(3), 133–135.

Henderson, J. V. & Kriticos, S. (2018), 'The development of the african system of cities', *Annual Review of Economics* **10**, 287–314.

Henderson, J. V., Storeygard, A. & Weil, D. N. (2012), 'Measuring economic growth from outer space', *American economic review* **102**(2), 994–1028.

Henderson, V. (2003), 'The urbanization process and economic growth: The so-what question', *Journal of Economic growth* **8**(1), 47–71.

Huang, J., Lu, X. X. & Sellers, J. M. (2007), 'A global comparative analysis of urban form: Applying spatial metrics and remote sensing', *Landscape and urban planning* **82**(4), 184–197.

Jedwab, R., Christiaensen, L., Gindelsky, M. et al. (2014), 'Rural push, urban pull and... urban push? new historical evidence from developing countries', *Institute for International Economic Policy: Washington, DC, USA* .

Koenker, R. & Bassett Jr, G. (1978), 'Regression quantiles', *Econometrica: journal of the Econometric Society* pp. 33–50.

Lall, S. V. (2017), 'Renewing expectations about africa's cities', *Oxford Review of Economic Policy* **33**(3), 521–539.

Li, X. & Zhou, Y. (2017), 'Urban mapping using dmsp/ols stable night-time light: a review', *International Journal of Remote Sensing* **38**(21), 6030–6046.

Liu, L. & Leung, Y. (2015), 'A study of urban expansion of prefectural-level cities in south china using night-time light images', *International Journal of Remote Sensing* **36**(22), 5557–5575.

Liu, Z., He, C., Zhang, Q., Huang, Q. & Yang, Y. (2012), 'Extracting the dynamics of urban expansion in china using dmsp-ols nighttime light data from 1992 to 2008', *Landscape and Urban Planning* **106**(1), 62–72.

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

Roberts, M., Stewart, B., Prakash, M. & McWilliams, K. (2015), 'Global night time lights urban extents and growth patterns product', *Alpha Version World Bank* .

Schiavina, M., Freire, S. & MacManus, K. (2019), 'Ghs population grid multitemporal (1975, 1990, 2000, 2015) r2019a', *Eur. Comm. JRC* .

Spence, M., Annez, P. C. & Buckley, R. M. (2008), *Urbanization and growth*, World Bank Publications.

Tewari, M. & Godfrey, N. (2016), 'Better cities, better growth: India's urban opportunity'.

UN, D. (2015), 'World urbanization prospects: The 2014 revision', *United Nations Department of Economics and Social Affairs, Population Division: New York, NY, USA* **41**.

Zhang, Q., Pandey, B. & Seto, K. C. (2016), 'A robust method to generate a consistent time series from dmsp/ols nighttime light data', *IEEE Transactions on Geoscience and Remote Sensing* **54**(10), 5821–5831.