Urban Form and Productivity

What Is the Shape of Latin American Cities?

Juan C. Duque
Nancy Lozano-Gracia
Jorge E. Patino
Paula Restrepo
Abstract

This paper examines the linkages between urban form and city productivity, using alternative metrics for urban form and applying them to a comprehensive sample of Latin American cities. Although most of the literature has concentrated on the effects of population density (compact versus sprawling urban development), this paper seeks to assess whether different dimensions of a city's urban form, such as shape, structure, and land-use, affect its economic performance. The paper finds that the shape of the urban extent, the inner-city connectedness, and fullness have a statistically significant influence on the productivity level of the city.

This paper is a product of the Social, Urban, Rural and Resilience Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/research. The authors may be contacted at nlozano@worldbank.org and prestrepocadavid@worldbank.org.
Urban Form and Productivity: What Is the Shape of Latin American Cities?

Juan C. Duque (✉)
Research in Spatial Economics (RiSE-group), Department of Economics, Universidad EAFIT, Medellin, Colombia. e-mail: jduquec1@eafit.edu.co

Nancy Lozano-Gracia, Social, Urban, Rural and Resilience. The World Bank, Washington, USA. e-mail: nlozano@worldbank.org

Jorge E. Patino, Research in Spatial Economics (RiSE-group), Department of Economics, Universidad EAFIT, Medellin, Colombia. e-mail: jpatinoq@eafit.edu.co

Paula Restrepo, Social, Urban, Rural and Resilience. The World Bank, Washington, USA. e-mail: prestrepocadavid@worldbank.org

Keywords: productivity · urban form

JEL Classification: R11 · R14
1 Introduction

Early work on urban economics Parr (1979, 1987) recognized the links between urban form and economic performance. The spatial structure of cities or urban form is thought to have important influence on the emergence of agglomeration economies and congestion costs, and, hence, on a city’s level of productivity, but also the sustainability of cities and their quality of life (Squires, 2002). The same channels – matching, learning and sharing – that are used to explain the emergence of agglomeration economies (Duranton and Puga, 2004) are also thought to be at the core of the links between urban form and city productivity. Holding city size constant, cities can use land and space in very different ways. Local governments, through land use regulations and other urban policy instruments, can influence the way economic activities, urban infrastructure, and households are located (World Bank, 2009). These differences in the occupation of urban space have consequences in terms of the transport modes used (i.e. favoring the use of private vehicles versus public transport), commuting times in cities, as well as on the type and intensity of human interaction. Denser cities are thought to improve labor productivity through better matching of firms and workers and enhanced interactions which facilitate the spread of tacit knowledge, both of which are thought to occur more easily the closer are people and firms located together (Ciccone and Hall, 1996; Rosenthal and Strange, 2004; Cervero, 2001). However, despite the links being well established through a theoretical framework, there are few empirical studies that examine the relationship between urban form and economic performance; further, there is still little evidence linking the presence of agglomeration economies with the form that cities take.

With over 80 percent of LAC’s population living in cities, understanding the links between city form and productivity is of paramount importance for policy makers in the region. The challenges of topography may have prevented dense development in some cases, as many Latin American cities are located in rugged topographies with natural barriers. Furthermore, rapid urbanization combined with limited infrastructure investments may have also led to urban forms that pose a barrier to the rise of agglomeration economies, limit firm interaction, and make it hard for workers to reach their jobs. Better understanding the links between urban
form and productivity can shed light on whether urban policy has a role to play in supporting city productivity and through that, economic growth.

In this paper we examine the linkages between urban form and city productivity using different urban form metrics for a comprehensive sample of Latin American cities with more than 50,000 inhabitants in the year 2010. We use a consistent time series of Defense Meteorological Satellite Programs - Operational Linescan System (DMSP-OLS) nighttime lights (NTL) imagery to identify city extents, characterize urban structure and create a proxy measure of a city’s economic performance. In terms of methodological approach, our first base model includes the estimation of urban productivity using time-lagged measures of urban form and other control variables. In a second model we follow Harari (2016) and explore the use of synthetic instruments that use the potential shape of city as a starting point, and calculate the city form indicators based on such potential form. The contribution of this paper is twofold. First, the analysis of linkages between urban form and city productivity has important policy implications in regards to the potential benefits or consequences of land use, transport and other city planning policies that can influence city form. Second, it contributes to the advancement in the measurement of urban form, by using open data available at a global scale that allow for benchmarking of cities across the globe.

The rest of the paper is organized as follows. Section 2 provides a literature review. Section 3 presents the empirical models. Section 4 describes the source data and the construction of the proxy measure of a city’s economic performance, the measures of urban form, and the construction of the synthetic instrument. Section 5 presents the empirical results. Finally, section 6 presents our conclusions.

2 Literature review

Density, size and sprawl have been commonly referenced in economic literature as some of the key aspects of urban productivity. The most recurrent conclusion is that less dense cities face high commuting rates (Wheeler, 2001), have higher marginal costs of transporting intermediate goods (Ciccone and Hall, 1996), and lower knowledge spillovers (Lynch, 1981; Jaffe et al 1993 and Glaeser 1998). All of which can affect a city’s productivity levels.
However some authors argue that achievements such as improved highways, public transit services (Glaeser, 2014; Chatman and Noland, 2014) and advances in communication technologies (Partridge et al, 2009) have helped to reduce the productivity costs of sprawling.

When thinking of urban form, urban economists commonly use population density and city size as proxies. However, the main limitation of this kind of variables is that they assume a uniform distribution of people across space. Such an assumption ignores the fact that the urban form is the result of the interaction of individuals, firms, and government decisions on where to live, where to locate, and where to place infrastructure. Such interactions are often reflected in the layout of the city, e.g. the street layout, and the land use patterns across the urban space. In other words, urban form and land use patterns cannot be properly identified from density (Melo, et al., 2017). Nowadays, the evolution of areas such as Geographic Information Science and Computational Geometry, which has been reflected in friendlier and inexpensive analysis tools, together with the increase of availability of georeferenced data, has allowed economists to look beyond density and size when thinking of the effect of urban form on city productivity.\(^1\) An example of this trend can be found in Harari (2016), who takes a slightly different approach and estimates the impact of urban form on city productivity. In her work, Harari uses DMSP/OLS Night-time Lights to delineate the urban extents of a set with over 450 Indian cities and uses those extent boundaries to calculate geometry-based shape metrics with an open source library written in Python 2.5 for ESRI’s ArcGIS, The Shape Metrics Tool, described in Angel et al. (2010). Contrary to other studies, she finds that more compact cities (in terms of the geometry of the urban extent as opposed to population density) are characterized by larger populations, lower wages, and higher housing rents. This suggests that a city’s residents value compactness as a consumption amenity. By contrast, firms do not appear to be directly affected by city shape in their location choices and there is no evidence of a significant impact on the productivity of firms.\(^2\) Most recently, in their analysis of urban development patterns of Indian cities, Tewari, Alder and Roberts (2016) find that there is a robust and positive relationship between a city’s initial level of

---

\(^1\) A similar phenomenon has occurred with the use of remote sensing data by economists.

\(^2\) Lall, Shalizi and Deichmann (2003) found similar results from an analysis of agglomeration economies and productivity in Indian industries. The authors find that benefits for firms to locate in dense urban areas do not appear to offset the associated costs.
compactness – as measured through different urban form metrics - and its subsequent economic growth, where a city’s rate of economic growth is estimated on the basis of nighttime lights data.

Another recent development on the calculation of shape metrics in a GIS context is the Metropolitan Form Analysis Toolbox developed by the City Form Lab (Amindarbari and Sevtsuk, 2013), an open source toolbox for ArcGIS that includes eight urban form and land-use metrics that can be used to track growth and change in cities: Size and Density, Coverage, Polycentricity, Compactness, Discontinuity, Expandability, and Land-Use Mix. Libraries such as the Metropolitan Form Analysis Toolbox and The Shape Metrics Tool have been used to analyze large samples of cities: Angel, Sheppard et al (2005) analyzed the urban extents of 3,493 cities in 1990 and 2000. Amindarbari and Sevtsuk (2013) calculate urban form and land-use metrics for Los Angeles, Singapore, Jakarta, Guangzhou and Chengdu. Other studies on urban form, that do not necessarily use the libraries mentioned above, are: Bertaud and Malpezzi (2003) calculate a range of measures of urban form for 48 major metropolitan areas around the world, and analyzed the impact of income, population, and the nature of the regulatory regime as potential determinants of urban form; Huang et al (2007) calculate seven spatial metrics of urban form for 77 metropolitan areas in Asia, US, Europe, Latin America and Australia; they found that urban agglomerations in the developing world are more compact and dense than their those in Europe or North America.

Urban planners, in the search of an integrated theory of city planning, have explored multidisciplinary approaches that integrate areas such as economics, network science, and geometry, to understand the impact of size, scale, and shape on city sustainability (Batty, 2008). This process led them to have a more elaborated concept of urban form. As in Batty and Longley (1994), urban form, from an urban planning perspective, “represent the spatial pattern of elements composing the city in terms of its networks, buildings, spaces, defined through its geometry mainly, but not exclusively, in two rather than three dimensions” [pp:42]. Thus, a proper characterization of city form should include not only information on the shape of its border, but also characteristics of its urban texture and land use planning (a similar warning is made by Prosperi et al., 2009). In fact, the characterization of urban form as a combination of external shape and internal structure was suggested for the very first time
by Whyte (1968). Based on these characteristics, the authors differentiate between natural/organic and planned/regular/artificial/geometric cities. Planned cities are characterized by straight streets forming a grid, which can include some radiality with the incorporation of streets that symmetrically converge to a point in the city. They also show a clear segregation of uses and are the result of large-scale planning processes. They also tend to have circular borders. Organic cities grow slower than those that are planned, and their geometry is the result of individual decisions and follows more closely the natural landscape. It has curved roads and open spaces randomly located. The pattern shows no evidence of planned growth and it is radially concentric in structure (Batty, 2008).

Our contribution to the literature consists of using the multidimensional definition of urban form from urban planning to attain an integral understanding of the influence of urban form on city productivity. We also contribute with new empirical evidence of this relationship using a large set of Latin American cities.

3 Models

To examine the effect of urban form on productivity, we estimate the following model:

\[ \ln Y_{it} = \alpha S_{it} + \beta T_{it} + \gamma L_{it} + \xi X_{it} + \theta_i + \epsilon_i, \]  

where \(Y_{it}\) measures productivity of city \(i\) at time \(t\); \(S\) is a vector of urban shape variables; \(T\) is a vector of urban texture variables; \(L\) represents the land use pattern within the city; and \(X\) is a vector of control variables, included the intercept; \(\theta\) are the city fixed effects; and \(\epsilon\) is the error term.

One of the main challenges to be tackled in this paper is the endogeneity of urban form when estimating the relationship between productivity and city shape; city form in fact, can be thought of as being the result of the interaction between decisions taken by firms, households, and government, and hence both urban form and productivity result from the interplay between agglomeration economies and congestion forces. In the simplest example, cities that grow dense (or in a concentric pattern around a central business district) can facilitate agglomeration economies by easing proximity of firms and, hence, fostering productivity.
But highly productive cities are also more likely to have governments willing to invest in city centers, and hence promote denser growth patterns. For identification purposes, this work will explore two alternative strategies:

**Strategy 1:** We implemented instrumental variables techniques for a cross-section of cities. The productivity variable is measured using information from the latest available year, and the explanatory variables use data from an earlier year, to mitigate the direct simultaneity between the dependent and independent variables.

\[
\ln Y_{it} = \alpha S_{i(t-k)} + \beta T_{i(t-k)} + \gamma L_{i(t-k)} + \xi X_{i(t-k)} + \theta_i + \epsilon_i, \tag{2}
\]

**Strategy 2:** We use an instrumental variables approach, exploiting both temporal and cross-sectional variation in city shape. In this step, we use the time variation of the NTL data to build a panel of time-city observations. In defining the instrumental variable, we follow Harari (2016) and construct a synthetic instrument that uses the potential shape of the city as a starting point, and calculates the city form indicators based on such potential form. Following Harari (2016), the identification relies on the fact that exogenous changes in city form over time can result from encountering topographic obstacles along its expansion path. The estimation has the following form:

\[
\ln Y_{it} = \alpha S_{it} + \delta X_{it} + \theta_i + \nu_t + \epsilon_{it}, \tag{3}
\]

\[
S_{it} = \sigma \hat{S}_{it} + \zeta X_{it} + \omega_i + \varphi_t + \pi_{it}, \tag{4}
\]

where $\hat{S}$ is a vector of city shape indicators derived from the potential urban extent, $\omega$ are the city fixed effects; $\varphi$ are the time fixed effects; and $\pi$ is the error term.

For completeness, we present the steps proposed by Harari (2016) to estimate the potential urban extent when assuming a common average expansion rate across all cities:

1. For the first year, the potential urban extent is the largest patch of developable land (i.e., excluding water bodies and steep terrains) within the minimum-bounding circle of that first year.
2. Estimate the predicted area of city $i$ in year $t$, $\text{area}_{it}$, with the following regression:
\[ \log(\text{area}_{i,t}) = \theta_i + \gamma_t + \epsilon_{i,t}, \]  

(5)

where \( \theta_i \) and \( \gamma_t \) are city and year fixed effects. In order to estimate this regression, it is necessary to have the actual areas of the urban extents of all the studied cities for, at least, two different years.

3. The estimated urban extent of city \( i \) in year \( t \) consists of the largest patch of developable land within the circle of ratio \( \hat{r}_{i,t} \) calculated as:

\[ \hat{r}_{i,t} = \sqrt{\frac{\text{area}_{i,t}}{\pi}} \]  

(6)

4 Data

4.1 Study region

Our analysis focuses on Latin American and Caribbean (LAC) cities with populations over 50,000 people in 2010. While LAC is the most urbanized region on the planet, and the built-up areas continue to expand despite deceleration in population growth (UN-Habitat, 2012), regional averages hide large variability in the distribution of urban population across countries. Further, the region is characterized by large diversity of climatic conditions, including from tropical to temperate climates; the terrain also shows high variability in altitudes due to the presence of the Andean Mountain Range. That diversity in climate and topographic conditions translates into a high diversity in the shape that urban areas have taken. On the one hand, there are cities located in the mountains constrained by hilly terrain; on the other hand, there are also some cities located in the lower lands in flat areas, constrained only by the presence of large rivers or lakes, and yet other cities have grown located at the Caribbean and Pacific coasts, with shapes that reflect the coastline.

4.2 Urban extent delineation

4.2.1 The DMSP-OLS images
We use the Defense Meteorological Satellite Programs - Operational Linescan System (DMSP-OLS) nighttime lights (NTL) data to outline urban areas. The NTL data report the recorded intensity of Earth’s surface lights. Nighttime lights products have high correlation to human activities (Hsu, Baugh, Ghosh, Zhizhin, & Elvidge, 2015), and have been previously used for regional and global analysis of urbanization (Cheng et al., 2016; Pandey, Joshi, & Seto, 2013; Sutton, Cova, & Elvidge, 2006; Zhang & Seto, 2011; N. Zhou, Hubacek, & Roberts, 2015; Y. Zhou et al., 2015), population modeling (Anderson, Tuttle, Powell, & Sutton, 2010; Lo, 2001), and economic performance (Cao, Wu, Kuang, Huang, & Wang, 2016; Forbes, 2013).

There are two different nighttime light products from DMSP-OLS that can be used to delineate urban areas: the stable or ordinary product (NTL), and the radiance-calibrated (RC) product. For this work we use the latter, given that such data include a correction for the saturation issue likely to be an important issue for most large cities in the region. Further, previous work has suggested that the RC information provides a better proxy for socioeconomic variables than the stable products (Hsu et al., 2015; Ma, Wu, Li, Peng, & Liu, 2014). Radiance-calibrated yearly composites for 1996, 2000 and 2010 obtained from NOAA National Centers for Environmental Information were used to delineate urban extents. These composites have a spatial resolution of 30 arc-seconds (about 1 km at the Equator).

Another known issue of the DMSP-OLS products is the “overglow” effect: dim lighting detected from light in surrounding areas of cities because the scattering of lights in the atmosphere (Wu, Ma, Li, Peng, & Liu, 2014). A novel deblurring process was applied to address the issue of overglow in the radiance-calibrated products. This process involves the use of two sequential filters, a standard deconvolution and the frequency of illumination maxima, to withdraw the light from the surroundings back and restacking it vertically on its source pixels at city centers (Abrahams, Lozano-Garcia and Oram, 2016).

4.2.2 Calibration and interannual correction of DMSP-OLS RC NTL images

We use the deblurred DMSP-OLS RC yearly composites for the years 1996, 2000 and 2010 to delineate urban extents in Latin America. DMSP-OLS RC NTL from different years
cannot be compared because of the lack of the sensor’s on-board calibration (Cao et al., 2016; Hsu et al., 2015; Pandey et al., 2013; Zhang & Seto, 2011). An empirical intercalibration was performed to enable the comparison of these images. We follow (Hsu et al., 2015) and used the same area as reference: Los Angeles metropolitan area. This area was selected for two reasons: it has for long been a mature metropolis and hence the light change is negligible; second, because of its size and the large variability of light intensity within its extent, it provides samples of high DN values in the city center, as well as low DN values in suburban areas. We use the 2000 image as reference, and estimate linear regression models and second-order regression models. The second-order models showed only negligible improvements compared to the linear model, which is preferred for its simplicity (Table 1).
We set the intercept as 0 in the calibration models to ensure that all the background areas have the same zero value in the three images. While the calibration solves the comparison issue, there are additional inconsistencies in lit pixels among the three dates that need to be addressed. These inconsistencies are seen as lit pixels that change to unlit pixels in the next date image at the same geographic location, and as abnormal fluctuations in pixel values across dates (Cao et al., 2016). We applied the inter-annual series correction proposed by (Cao et al., 2016) to ensure that the lit pixels detected in an image do not disappear in a later date, and that the lit pixel DN values for each date are not smaller than the pixel DN value at the same location of a previous date. The inter-annual correction model after intercalibration is composed of two rules: (1) when a pixel DN value is equal to zero in an image, that pixel DN value for the previous date at the same geographic location should also be equal to zero, and (2) when a pixel DN value is not equal to zero in an image, the pixel DN value must be greater than the DN value of the lit pixel in the same geographic location in the previous date. This correction model is described in equation (7).

\[
DN_{(n,i)} = \begin{cases} 
0, & DN_{(n+1,n)} = 0 \\
DN_{(n-1,i)}, & DN_{(n+1,i)} > 0 \text{ and } DN_{(n-1,i)} > DN_{(n,i)} \\
DN_{(n,i)}, & \text{otherwise}
\end{cases}
\]  

(7)

Where \(DN_{(n-1,i)}, DN_{(n,i)},\) and \(DN_{(n+1,i)}\) are the DN values of the \(i\)-th lit pixel in the deblurred RC NTL image after intercalibration in the \((n-1)\)-th, the \(n\)-th, and \((n+1)\)-th years, respectively. And \(n = 1996, 2000, \text{and} 2010.\)

4.2.3 Urban extent delineation

We use the inter-annual corrected deblurred DMSP-OLS NTL RC images to outline urban extents in Latin America and the Caribbean. We applied a DN threshold to define what is considered an urban area in the NTL imagery. Threshold-based methods have been
extensively used to extract urban areas from NTL data (Ellis and Roberts, 2015; Harari, 2016; Li and Zhou, 2017; Pandey et al., 2013; Sutton, et al., 2010; Zhang and Seto, 2011). The central premise in this method is that urban areas are generally a concentration of people and income such as there is a DN value (a threshold) that can be used to define and delineate urban areas from NTL imagery (Tewari et al., 2016). The threshold was chosen after careful evaluation of the longitudinal profiles for the year 2000 for a sample of LAC cities; as an additional check, the extents were compared to the Built-up Global Human Settlements Layer (GHS) for the same year at 250 meters of spatial resolution (GHS_BUILT_LDS1990_GLOBE_R2016A_54009_250, Freire & Pesaresi, 2015). For this work we found a threshold of DN equal to 110 in the processed NTL images (i.e., after deblurring, and interannual calibration and correction). Because the GHS built-up layer uses Landsat images as input and does not consider at any point layers of NTL, we feel confident using it as an independent data source to set the threshold. Figure 1 shows three examples of the extracted urban extent and the built-up layer used as reference.

**Figure 1.** Examples of urban extents extracted from the 2000 NTL image over the reference 2000 GHS Built-up layer. From left to right: Buenos Aires (Argentina), Mexico D.F. (México), and Rio de Janeiro (Brazil).

We applied the threshold and obtained binary images of urban areas in LAC. The three images were converted to vector format to create the polygons that outline the extent of cities. Further, we applied a buffer of 10 meters to the outlined polygons to encompass all the image cells that belonged to the same urban area. We used the LandScan 2012 data set (Bright et al., 2013) as reference to identify and extract the NTL urban extents with more than 50,000 people in 2010. We obtained 940 urban extents with more than 50,000 people in Latin
America for the year 2010 and extracted the urban extents for the same locations from the processed NTL images of the years 1996 and 2000 (Figure 2). According to (Chuvieco, 2016), an object must be several times larger than the pixel size to be delineated properly from a remote sensing image. Since the pixel size of the DMSP-OLS NTL RC images is 30 arc seconds (almost 1 by 1 Km near the Equator), we feel justified to exclude those urban extents that were smaller than 3 Km² in 1996. That resulting final sample therefore includes 919 urban extents in each year (Figure 2).

4.3 Measuring urban productivity

The use of NTL in socioeconomic studies emerged as a way to address the lack of economic measures at disaggregated scales. The pioneer contribution by Henderson et al (2012) shows that NTL data can be used as a proxy of GDP within and across countries. Since then, other authors have used light density from NTL imagery to measure economic performance and welfare (Bleakley and Lin, 2012; Michalopoulos and Papaioannou, 2013; Lowe, 2014; Storeygard, 2016; Pinkovskiy, 2013).

In this paper we follow Tewari et al. (2016) to calculate a measure of productivity ($\bar{Y}$) from NTL data. We use the density of radiance within the urban extent, $dr_{ntl\_2010}$, as our measure of productivity, computed as the sum of the NTL DN values in 2010 divided by area, in square kilometers, of the urban extent in 2010.
4.4 Measuring urban form

Following the literature on urban planning, we adopt an integral definition of urban form that includes three dimensions: Shape of the urban extent, internal urban structure and land use pattern (Prosperi et al., 2009 and Batty, 2008). In this subsection we present our proposal for measuring urban form within the framework of these three dimensions. For each dimension we present the variables proposed for its measurement, the rationale about the potential mechanisms through which each variable impacts economic performance, and the hypothesis to be tested in the empirical results.

According to Thompson (1952) a perfect circle has geometric properties such as minimum surface area, and maximum accessibility from and to any interior point. Angel et al. (2010) translate these geometric concepts to city shape and argue that the shape of a city affects its
efficiency, equitability and sustainability. It is also proven that the city’s annual costs, per household, of public infrastructure and services are lower for circular/compact cities compared to fragmented/irregular/sprawling cities. Thus, circular cities can reach efficiency at lower cost. Based on those premises, we hypothesize that cities whose shape approaches a perfect circle can reach economic efficiency cheaper and faster. It is important to make clear that measuring the geometric shape of the city using its boundary does not capture what is happening within the city in terms of accessibility or land use. (For example, a geometrically compact city may have an inefficient transport network that prevents taking full advantage of being a circular city.) These internal characteristics of the city are covered by the dimensions urban structure and land use.

From a geometric perspective, a shape metric is usually focused in one of the three following geometric characteristics: degree of roundness, smoothness of its perimeter, and fullness (Angel et al., 2010). None of those geometric characteristics can be captured via the usual density-based measures of urban sprawl, but they affect aspects such as trip lengths and accessibility, and through that, have the potential of affecting productivity levels in a city. We used the Shape Metrics Toolbox4 to calculate the shape metrics of the urban extent polygons. This tool is a Python script that runs out of ArcToolbox in ArcGIS 9.3, and we upgraded the script to properly run in ArcGIS 10.x. All the urban extent metrics calculated with the Shape Metrics Tool take values from 0 to 1, with 1 being a perfect circle: The Exchange_index measures how much the urban extent has deviated from its compact shape towards irregular non-compact, and potentially inefficient, forms. Low values of the Exchange_index tend to capture elongated urban extents. The Perimeter_index measures how smooth is the perimeter of the urban extent. Low values of this index indicate that the urban perimeter is irregular, which can affect the free mobility of residents located near the periphery towards other periphery locations within the urban extent. It is expected that cities located in the rugged topography of Los Andes show growth patterns that follow ridges and valleys, which can create organic shapes that are associated to low values of the perimeter index. Note that these two measures, Exchange_index and Perimeter_index, are

4 This software is intellectual property of the Center for Land Use Education and Research (CLEAR) at the University of Connecticut (http://clear.uconn.edu/tools/Shape_Metrics/index.htm).
complementary since they allow differentiating between an irregularly compact city and a smoothly compact city, the latter being closer to a perfect circle.

*Fullness* measures the presence of built-up areas within the urban extent as a fraction of the urban extent area. We used the GHS built-up 1990 raster layer at 250 meters resolution (GHS_BUILT_LDS1990_GLOBE_R2016A_54009_250, Pesaresi et al., 2015; Pesaresi et al., 2016) to calculate this variable. The GHS built-up layer values are expressed as decimals from 0 to 1, and correspond to the fraction of the pixel that is covered by a building (Pesaresi et al., 2016). Fullness was measured as the mean value of all the pixels of the GHS built-up 1990 layer within the urban extent. This measure can be understood as the complement of the sprawl index proposed by Burchfield et al. (2006) defined as the measure of absence of built-up area. Based on the extensive literature on the relationship between compact cities and productivity, we would expect a “full-city” to be one that is also a more compact city, hence allowing for greater interaction and therefore higher agglomeration economies which ultimately increase productivity. However, a city that is “too-full” might also be suggestive of lack of public space, which can both be a dissamenity and reflect lack of planning. For these reasons we can expect a non-linear relationship between fullness and productivity.

Following the above-mentioned rationale, our hypotheses regarding the effect of these indicators of city shape on productivity are as follows:

- **H1**: An increase of the *Exchange_index* suggests a more circular urban extent, which is expected to be associated with higher productivity levels of the city.
- **H2**: An increase of the *Perimeter_index* implies smooth perimeters that are associated with higher productivity levels of the city.
- **H3**: An increase of the *Fullness_index* implies a less fragmented urban layout that increases productivity levels of the city, but at extremely high levels of fullness can affect productivity.

As mentioned in Section 2, the internal structure of the city may play a considerable role in its productivity levels. Natural/organic cities have different dynamics than planned/regular/artificial/geometric cities; thus, it is important to include variables that capture the different urban structures in order to quantify its impact on productivity. Such structures affect the way in which people and products move within the city. Previous
literature has found empirical evidence between urban structure and economic performance; for example, Mills and Hamilton (1989) and Bogart (1998) found that better accessibility to labor force and efficient transport infrastructure reduces time and costs, which increases productivity levels; and Bertaud (2004) shows that shorter, and therefore cheaper, internal trips increase the levels of urban efficiency.

In order to capture such characteristics of the city structure, we used the OSMnx Python library (Boeing, 2017) to compute three geometry-based network topology variables: Circuity_avg, Intersection_density and Street_density. The Circuity_avg divides the sum of the length of all edges of the street network by the sum of the great-circle distances between nodes incident to each edge (Boeing, 2017). If all the streets in the network are straight, the value of the circuity average (Circuity_avg) is 1. Higher values account for the presence of curvy streets, indicating a more organic urban layout. As shown in Giacomin and Levinson (2015), for the most populated Metropolitan Statistical Areas in the United States, low circuity is associated with more efficient and shorter trips. Using also metropolitan areas in the US, Huang and Levinson (2015) show that transit circuity affects accessibility of transit networks.

The second and third variables are the Intersection_density and the Street_density. These two variables inform about how fine grained and connected the network is, thus informing about the ease of movement across the city (Boeing, 2017). The existence of dense intersection patterns may allow for high rates of walking and non-motorized modes use, taking into account the possibility of having more controlled street crossings and more available access points (Cervero, 1997).

From the discussion above, we formulate the following hypotheses related to the structure of the urban extent:

---

5 The OSMnx Python library automatically downloads and analyzes street networks from OpenStreetMap data. These networks do not include underground metro systems.
**H4:** Regular urban structures (associated to low values of \( \text{Circuity}_{\text{avg}} \)) are usually associated with more efficient and shorter trips, which reduce agglomeration costs and may lead to higher levels of productivity.

**H5:** An increase of either \( \text{Intersection}_{\text{density}} \) or \( \text{Street}_{\text{density}} \) implies higher levels of street coverage that facilitates the mobility of people and products increasing the connectivity levels, and therefore, leading to higher productivity levels of the city.

Finally, as shown in Fallah et al (2011), the consideration of the distributional aspects of the population within the urban extent provides information on the land use pattern of the city. The authors proposed a measure of urban sprawl that allows differentiating between cities with an even distribution of its population, from cities with highly concentrated population. For completeness, we present in equation (8) the measure proposed by Fallah et al (2011), and adopted in this work:

\[
\text{Sprawl} = ((L\% - H\%) + 1) * 0.5 \tag{8}
\]

Dividing the urban extent in small areas, \( L\% \) is the share of the urban population living in a small area with density below the median density calculated for the entire set of analyzed urban extents (in our case the 940 LAC cities). Since our population data are in raster format, we considered each 250 m pixel of the GHS layer as a small area. This measure ranges from 0 to 1, with 1 indicating a greater level of sprawl. Fallah et al (2011) find that higher levels of sprawl are associated with lower levels of productivity.

**H6:** The distribution of population density within a city affects its economic performance. According to Fallah et al (2011), an uneven distribution of population density, associated with high levels of Sprawl, can be linked to deteriorating socio-economic outcomes, inefficient provision of public goods, and lower productivity levels.

As a summary, Table 2 presents the proposed variables, their description, and data source for calculation, and Table 3 presents examples of urban areas with high, medium and low values of these seven variables that describe urban form. Based on our hypotheses, our benchmark-highly efficient city would have the following values for our variables:
• *Exchange_index* = 1 and *Perimeter_index* = 1; suggesting a perfectly circular border of the urban extent.

• *Fullness_index* ≈ 1; suggesting a dense/compact city.

• *Circuity_avg* = 1; suggesting a planned/regular/geometric urban structure.

• *Intersection_density* → ∞ and *Street_density* → ∞; suggesting high levels of connectivity.

• Sprawl = 0; suggesting a homogeneously distributed population across the urban extent.

Table 4 shows the correlation coefficients between population density and urban form variables. Most of the urban form variables are weakly correlated with population density, suggesting that the urban form variables provide complementary information about the urban form, not captured by population density alone. The highest correlation is found for Sprawl, followed by intersection density. The direction of the correlation between population density and sprawl is similar as that found in earlier work for US metropolitan areas (Fallah, Partridge, & Olfert, 2011) with the correlation being higher in absolute terms for Latin American and Caribbean cities (-0.65 for LAC compared to -0.42 for US metro areas).
| Dimension       | Sub-dimension | Variable            | Description                                                                 | Data source for calculation                                                                                                                                                                                                 |
|-----------------|---------------|---------------------|----------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| I. Urban Shape  | Roundness     | nExchange_1996      | Share of the total area of the urban extent that is inside the equal-area   | Urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 image                                                                                                                                                         |
|                 | Smoothness of perimeter | nPerimeter_1996     | Ratio of the perimeter of the equal-area circle and the perimeter of the     | Urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 image                                                                                                                                                         |
|                 | Fullness      | Fullness_1996       | Fraction of the total area of the urban extent that is built-up.            | Urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 image; Built-up raster layer from GHS (GHS_BUILT_LDS1990_GLOBE_R2016 A_54009_250_v1_0 at 250 meters of spatial resolution) |
| I. Urban Structure | Urban structure | Circuity_avg_1996  | From the street network: the average ratio between an edge length and the   | OpenStreetMap street network data within the urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 image                                                                                                                                 |
|                 | Connectivity  | Intersection_density_1996 | From the street network: the node density of the set of nodes with more      | OpenStreetMap street network data within the urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 image                                                                                                                                 |
|                 | Connectivity  | Street_density_1996 | From the street network: the sum of all edges in the undirected representation of network graph divided by the urban extent area | OpenStreetMap street network data within the urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 image.                                                                                                           |
| III. Land use   | Sprawl        | Sprawl_1996         | Normalized difference between the share of areas with population density    | 1990 Population count at pixel level from GHS (GHS_POP_GPW41990_GLOBE_R2015A_54009_250_v1_0 at 250 meters of spatial resolution) within the urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 data. |

**Table 2.** Variables for describing urban form (vectors $S$, $T$ and $L$).
Table 3. Examples of urban areas with high, medium and low values of the indexes that describe urban form.

|                        | High                  | Medium               | Low                   |
|------------------------|-----------------------|----------------------|-----------------------|
| Exchange index         | ![High Exchange Index](image1) | ![Medium Exchange Index](image2) | ![Low Exchange Index](image3) |
| Perimeter index        | ![High Perimeter Index](image4) | ![Medium Perimeter Index](image5) | ![Low Perimeter Index](image6) |
| Fullness index         | ![High Fullness Index](image7) | ![Medium Fullness Index](image8) | ![Low Fullness Index](image9) |
| Circuity avg           | ![High Circuity Avg](image10) | ![Medium Circuity Avg](image11) | ![Low Circuity Avg](image12) |

*Continue next page*
Table 3. Examples of urban areas with high, medium and low values of the indexes that describe urban form (continuation)

| Intersection density | High | Medium | Low |
|----------------------|------|--------|-----|
| Street density       |      |        |     |
| Sprawl               |      |        |     |

Table 4. Correlations between population density (per km2) and urban form variables.

|                          |        |
|--------------------------|--------|
| 1996 norm. exchange      | -0.0524|
| 1996 norm. perimeter     | -0.2143*|
| 1990 fullness            | 0.3422*|
| 1996 intersection density| 0.5002*|
| 1996 street density      | 0.4451*|
| 1996 circuity            | -0.1187*|
| 1990 sprawl              | -0.7123*|

Note: * Denotes significance at 5%

4.5 Potential shape of urban extents

As stated in section 3, we follow Harari (2016) to obtain a panel of city-time potential urban extents and to calculate their shape metrics. Harari (2016) considered “undevelopable” terrain the areas that are occupied by a body of water, or that have a slope above 15%. We
use the digital elevation model from the NASA Shuttle Radar Topographic Mission (SRTM) version 4, with a resolution of 90 meters (Jarvis, Reuter, Nelson, & Guevara, 2008), to calculate slopes; and the Global MODIS Raster Water Mask (Carroll, Townshend, DiMiceli, Noojipady, & Sohlberg, 2009a, 2009b) to account for the presence of bodies of water. We implemented the process to define the potential urban extent in ArcGIS from the extracted urban extents for 1996, considered as the real urban extents. We first determine the “minimum-bounding circle” that contains the real urban extent and locate its centroid. Then we use the $r_{l,t}$ estimated from equations (5) and (6) with the buffer function to obtain the estimated circles for each date. In all cases we use the estimated radius for the first date (1996). This is because we observed contraction of the potential urban extents when using the minimum-bounding circle for the first date (1996) and the estimated ratio for the following two dates (2000 and 2010). Second, we defined the developable terrain. Initial tests with a 15% slope threshold proved to produce fragmented patches of developable land in many cities, resulting in potential urban extents that were extremely small compared to reality. While architectural development guidelines in developed countries typically use the 15% slope threshold as constraint threshold for urban development (Saiz, 2010), given that the Andean Mountain Range crosses many countries in Latin America, in the region, urban development guidelines have been more tolerant and cities have grown in areas with considerably steeper slopes. An analysis of the distribution of slope values in a sample of urban extents extracted for the year 1996 suggested that in fact the median slope in the areas considered is 21%; hence, in the analysis we use a threshold of 20% slope to delimitate the undevelopable land areas. Finally, we manually select the largest patch within the circle as the potential urban extent for each date. Figure 3 shows an example of an actual urban extent and its potential shapes in the three dates.

**Figure 3.** Lima 1996 extracted urban extent outline in dark red and potential urban extent areas for 1996 (light red), 2000 (orange), and 2010 (yellow). Base map: © Google.
4.6 Control variables

In order to isolate the predictive power of the variables describing urban form and to reduce omitted-variables bias, we include in the model a number of control variables including city size, locational variables, natural and urban amenities, as well as city and year fixed effects. A number of different data sources are used to compute control variables. Population data in gridded format for 1990 were obtained from the Global Human Settlement Layer (GHS, at 250 meters of spatial resolution) produced by the Joint Research Centre (JRC) of the European Union (Freire & Pesaresi, 2015; Pesaresi et al., 2016). Natural amenities were calculated using several GIS layers. Average annual temperature and precipitation were obtained from the WorldClim web portal (Hijmans et al., 2005). Topographic variation was computed from the Topographic Ruggedness Index global gridded data\(^6\) (Nunn & Puga, 2012). Water bodies and elevation data were used to code constraints to urban expansion, and to calculate dummy variables for location near the sea (coast) or inland bodies of water such as large rivers and lakes. We used the 250 meters resolution raster MODIS Water Mask (Carroll et al., 2009a, 2009b) and the SRTM 90m Digital Elevation Database v4.1 (Jarvis et al., 2008). Pollution data in gridded format (PM\(_{10}\) total emissions for 2008) were downloaded.

\(^6\) [http://diegopuga.org/data/rugged/](http://diegopuga.org/data/rugged/)
from the Emissions Database for Global Atmospheric Research (EDGAR)\(^7\) of the JRC.

Finally, we used OpenStreetMap data to account for a number of urban amenities like hospitals, universities, among others. (© OpenStreetMap contributors\(^8\)). Table 5 presents a complete list of all control variables.

5 Empirical results

5.1 Strategy 1: Lagged model

Table 6 presents the correlations between the seven variables that we use to describe the three dimensions of urban form, and Table 7 presents a series of 10 alternative specifications for equation (2).\(^9\) The main equation is estimated using OLS and assuming intra-group correlation; i.e., that the residuals are correlated within geographical clusters, but uncorrelated between clusters. We cluster at the country level and have a total of 32 countries in the sample.

Even though the three indexes used to describe the shape of the urban extent seek to measure different aspects of urban form (roundness, smoothness and fullness), the correlations in Table 6 show that there is very high positive correlation between \textit{exchange\_index} and \textit{perimeter\_index} (0.74). We also found high correlation between the variables \textit{intersection\_density} and \textit{street\_density}. Both variables measure the level of connectivity within a city and therefore have a very high correlation of 0.95. For this reason, we test alternative specifications, as shown in Table 7 where columns 1 to 5 use \textit{intersection\_density} and columns 6 to 10 use \textit{street\_density}. Finally, there is a high and negative correlation between \textit{fullness} and \textit{sprawl} (-0.71); for this reason we use them separately.

\(^{7}\) http://edgar.jrc.ec.europa.eu/gallery.php?release=v42&substance=PM10&sector=TOTALS

\(^{8}\) http://www.openstreetmap.org/copyright

\(^{9}\) Table A1 shows the descriptive statistics of the variables.
Table 5. Control variables (Vector $X$)

| Dimension       | Variable            | Description                                                                 | Data source for calculation                                                                                                                                 |
|-----------------|---------------------|-----------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|
| I. Urban population | Pop_1990          | Sum of population counts within the urban extent                            | 1990 Population count at pixel level from GHS (GHS_POP_GPW41990_GLOBE_R2015A_54009_250_v1_0 at 250 meters of spatial resolution) within the urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 data. |
|                 | Pop_1990_2         | The square of population counts within the urban extent                     |                                                                                                                                                             |
| II. Location    | Dist_Border_kms_1996 | Distance to the near international border                                  | Country borders from the World Bank LAC Spatial Database and urban extents from deblurred and corrected DMSP-OLS NTL RC 1996 data.                              |
| III. Natural Amenities | Temp_2010      | Average annual temperature (C)                                              | WorldClim, Bioclimatic variables, BIO1: annual mean temperature within the urban extents from deblurred and corrected DMSP-OLS NTL RC 2010 data.              |
|                 | Precip_2010        | Average annual precipitation (mm)                                           | WorldClim, Bioclimatic variables, BIO12: annual precipitation within the urban extents from deblurred and corrected DMSP-OLS NTL RC 2010 data.               |
|                 | TRI_2010           | Topographic ruggedness index                                                | Topographic ruggedness index global gridded data (Nunn & Puga, 2012) within the urban extents from deblurred and corrected DMSP-OLS NTL RC 2010 data.           |
|                 | Coast_2010         | Dummy for location at the coast                                             | MODIS Water Mask (Carroll et al., 2009a, 2009b) and urban extents from deblurred and corrected DMSP-OLS NTL RC 2010 data.                               |
|                 | Water2_2010        | Dummy for location near a large river or lake                              |                                                                                                                                                             |
|                 | Pollution_ton/km2_2010 | Pollution of PM10 in the urban extent                                       | PM10 total emissions for 2008 (EDGAR) within urban extents from deblurred and corrected DMSP-OLS NTL RC 2010 data.                                              |
| IV. Urban Amenities | N_Arts_Centers_2016 | Number of arts centers                                                     |                                                                                                                                                             |
|                 | N_Cinemas_2016     | Number of cinemas                                                          |                                                                                                                                                             |
|                 | N_Theatres_2016    | Number of theatres                                                          |                                                                                                                                                             |
|                 | N_Museums_2016     | Number of museums                                                           |                                                                                                                                                             |
|                 | N_Pubs_2016        | Number of pubs                                                              |                                                                                                                                                             |
|                 | N_Restaurants_2016 | Number of restaurants                                                       |                                                                                                                                                             |
|                 | N_Cafes_2016       | Number of cafes                                                             | World Bank LAC OpenStreetMap geodatabase                                                                                                                   |
|                 | N_Comm_centers_2016 | Number of community centers                                                |                                                                                                                                                             |
|                 | N_Hospitals_2016   | Number of hospitals                                                         |                                                                                                                                                             |
|                 | N_Libraries_2016   | Number of libraries                                                         |                                                                                                                                                             |
|                 | N_Universities_2016 | Number of universities                                                      |                                                                                                                                                             |
|                 | Metro_Stations_2016 | Number of metro stations                                                    |                                                                                                                                                             |
| Lkm_Cicleways_2016 | Kilometers of cicleways within the urban extent |
The results confirm the convenience of having a more integral definition of urban form that goes beyond the use of population density. Regarding the shape of the perimeter of the urban extent, the models show that, ceteris paribus, a higher level of compactness is associated with higher productivity levels. The coefficients of exchange\_index and perimeter\_index are both positive and significant. These results provide support for hypotheses H1 and H2 regarding the shape of the perimeter of the urban extent.

Changing our focus from the periphery to the inner city, we covered four aspects: the level of urbanization (fullness), the urban structure (circuity), the level of connectivity (intersection density or street density) and the land use pattern (sprawl):

- The models show a nonlinear relation between the level of fullness and productivity: the coefficient of fullness is positive and significant, showing how less interrupted urban layout (i.e., low urban sprawl) increases the productivity levels. Yet, the coefficient of fullness\(^2\) is negative and statistically significant, suggesting that an excessive fullness may reflect a lack of public space and/or a lack of planning, setting the conditions for the appearance of diseconomies of agglomeration. Thus, this finding supports hypothesis H3.

- Regarding the urban structure of the city, the results show no evidence in favor of hypothesis H4. The coefficient of circuity is not significant in all specifications presented in Table 7. This implies that the level of productivity of the city does not depend on whether the city has a reticular or an organic urban structure.

- Connectivity, measured as either the street\_density or the intersection\_density, appears positive and highly significant at the 1 percent level, suggesting that, ceteris paribus, a well-connected city is a more productive city. These results support hypothesis H5.

- We also find empirical evidence in favor of our hypothesis H6: the sprawl variable, as measured by Fallah et al (2010), is negative and statistically significant, at the 10 percent level. Therefore, there is some evidence that the distribution of population density within Latin American cities can affect their level of productivity. It is important to note that the average level of sprawling in our sample is very similar to
the one reported by Fallah et al (2010), 0.61 vs. 0.63; as well as their standard deviation 0.18 vs. 0.12.

Finally, surprisingly and as shown in the results summarized in Table 7, we do not find evidence of the existence of economies of agglomeration or congestion costs, since the variables population and population\(^2\) are not significant. After performing an outlier analysis,\(^{10}\) we found that these specific results were driven by eight extreme outliers: Ciudad de México (México), São Paulo (Brazil), Buenos Aires (Argentina), Rio de Janerio (Brazil), Santiago (Chile), Lima (Peru), Bogotá (Colombia), and Guadalajara (México), which populations are extremely high compared to the rest of the cities. After dropping these outliers from the sample, the coefficient of population becomes positive and statistically significant, showing that higher levels of population are associated to higher levels of productivity; and the coefficient of population\(^2\) becomes negative and significant, indicating the presence of congestion effects on productivity. The conclusions regarding hypotheses H1 to H6 remain unchanged. The results of the estimations without outliers are presented in Table A3.

| 1996 norm. exchange (1) | (1)   | (2)   | (3)   | (4)   | (5)   | (6)   | (7)   |
|-------------------------|-------|-------|-------|-------|-------|-------|-------|
| 1996 norm. perimeter (2) |       | 0.7453* |       |       |       |       |       |
| 1990 fullness (3)       | -0.1573* |       | -0.3100* |       |       |       |       |
| 1996 intersection density (4) | 0.1427* | -0.0309 | 0.3429* |       |       |       |       |
| 1996 street density (5) | 0.1527* | -0.0465 | 0.4361* | 0.9470* |       |       |       |
| 1996 circuity (6)       | -0.1993* | -0.1172* | -0.0586 | -0.3901* | -0.4351* |       |       |
| 1990 sprawl (7)         | 0.0422 | 0.2216* | -0.7122* | -0.5454* | -0.5939* | 0.2032* |       |

Note: * Denotes significance at 5%

5.2 Strategy 2: Instrumental variables

The implementation of instrumental variables techniques was carried out using the two-stage least squares estimator (2SLS), allowing intra-group correlation at the country level. In Table 8 we report the 2SLS estimations, which use the log of normalization of potential urban form variables.

---

\(^{10}\) For the outliers identification, we used the algorithm BACON (Blocked adaptive computationally efficient outlier nominators) devised by Billor et al (2000) and implemented in STATA by Weber (2010).
(exchange_index and perimeter_index) as instruments for actual urban form (the results without outliers are reported in Table A6). We present in Tables A4 and A5 the estimates of the first stage, with and without outliers.\textsuperscript{11}

We began by discussing the instruments diagnostic test reported at the bottom of Table 8. Regarding the relevance of the instruments, the first stage of regressions results show that the instruments for actual urban form have considerable explanatory power. The explanatory power is tested using the F-tests, and we found values above 10 indicating that these instruments are strongly related to actual urban form. In fact, the results from estimating the first stage, reported in Tables A4 and A5, show that the potential urban form is a highly significant and positive predictor of the actual urban form, for the exchange_index as well as the perimeter_index. To further inspect the relevance of the instruments we carried out the Kleibergen-Paap test of under-identification (Kleibergen and Paap, 2006), which tests whether the model is identified, where identification requires that the excluded instruments are correlated with the endogenous regressor. The values of this test for the two models indicate that the test rejects the null hypothesis of under-identification at a 5% level of significance, suggesting that the instruments are relevant. We also perform a weak instrument test to test whether the instruments are correlated with the endogenous regressors, but only weakly. Since we allow intra-group correlation, the relevant statistic in this case is the Kleibergen and Paap (2006) rank Wald F statistic. The results reveal that the statistic values are higher than the Stock and Yogo (2005) critical values, rejecting the null hypothesis of weak instruments.

\textsuperscript{11} Table A2 shows the descriptive statistics of the variables.
Table 7. Estimates of the relationship between urban form and productivity (OLS). With outliers. (Dependent variable: ldr)

|                          | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     | (7)     | (8)     | (9)     | (10)    |
|--------------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 1996 norm. exchange      | 0.424***| 0.394***| 0.392***| 0.353***|         |         |         |         |         |         |
|                          | (0.0601)| (0.0643)| (0.0563)| (0.0611)|         |         |         |         |         |         |
| 1996 norm. perimeter     | 0.432** |         | 0.404** | 0.414** |         |         |         |         |         |         |
|                          | (0.1765)|         | (0.1654)| (0.1801)|         |         |         |         |         |         |
| 1990 fullness            | 0.710***| 0.604** | 0.594** |         | 0.695** | 0.600** | 0.588** |         |         |         |
|                          | (0.2487)| (0.2454)| (0.2560)|         | (0.2680)| (0.2621)| (0.2779)|         |         |         |
| 1990 fullness²           | -0.601***| -0.466**| -0.446**|         | -0.652***| -0.527**| -0.506**|         |         |         |
|                          | (0.1997)| (0.2035)| (0.2058)|         | (0.2107)| (0.2122)| (0.2206)|         |         |         |
| 1996 circuity            | 0.157  | 0.138  | 0.009  | 0.127  | 0.106  | 0.321  | 0.316  | 0.230  | 0.316  | 0.306  |
|                          | (0.8991)| (0.8638)| (0.8912)| (0.8763)| (0.8435)| (0.9297)| (0.8954)| (0.9235)| (0.9057)| (0.8753)|
| 1996 intersection density| 0.461***| 0.474***| 0.507***| 0.485***| 0.495***|         |         |         |         |         |
|                          | (0.0804)| (0.0784)| (0.0860)| (0.0907)| (0.0887)|         |         |         |         |         |
| 1996 street density      |         |         |         |         | 0.040***| 0.041***| 0.044***| 0.042***| 0.042***|         |
|                          |         |         |         |         | (0.0065)| (0.0064)| (0.0070)| (0.0074)| (0.0072)|         |
| 1990 sprawl              | -0.128* | -0.134*|         | -0.051 | -0.056 |         |         |         |         |         |
|                          | (0.0648)| (0.0776)|         | (0.0604)| (0.0733)|         |         |         |         |         |
| 1990 population          | -0.010  | -0.000  | -0.010  | -0.007 | -0.002 | -0.012 | -0.002 | -0.012 | -0.008 | -0.000 |
|                          | (0.0110)| (0.0068)| (0.0086)| (0.0094)| (0.0060)| (0.0106)| (0.0063)| (0.0085)| (0.0091)| (0.0055)|
| 1990 population²         | 0.000  | -0.000  | 0.000  | 0.000  | -0.000 | 0.000  | 0.000  | 0.000  | 0.000  | -0.000 |
|                          | (0.0001)| (0.0001)| (0.0001)| (0.0001)| (0.0001)| (0.0001)| (0.0001)| (0.0001)| (0.0001)| (0.0001)|
| Constant                 | 5.566***| 5.605***| 5.730***| 5.292***| 5.322***| 5.242***| 5.250***| 5.423***| 5.056***| 5.063***|
|                          | (0.8586)| (0.7410)| (0.9005)| (0.8773)| (0.7344)| (0.8952)| (0.7775)| (0.9390)| (0.9136)| (0.7696)|
| Country dummies          | Y       | Y       | Y       | Y       | Y       | Y       | Y       | Y       | Y       | Y       |
| N                        | 919     | 919     | 919     | 919     | 919     | 919     | 919     | 919     | 919     | 919     |
| R-squared                | 0.287   | 0.286   | 0.281   | 0.288   | 0.288   | 0.293   | 0.293   | 0.289   | 0.295   | 0.295   |

Note: Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1; All models include geographical characteristics as measured of natural amenities, namely, distance in km to international border, temperature, precipitation and coast indicator. Intersection_density has been divided by 100; street_density has been divided by 1,000; population has been divided by 100,000; and the control dist_border_kms has been divided by 1,000.
We now turn to consider the estimates of the impact urban form on productivity. For the case of exchange_index, we note that both the OLS and 2SLS estimates reveal that a higher level of urban compactness is associated with higher levels of productivity. The coefficient on exchange_index is positive and highly significant at the 5% level and the 2SLS coefficient is higher than what was found with the OLS estimate. The 2SLS coefficient on exchange_index suggests that a one standard deviation increase in the level of urban compact shape is associated with an 8-percentage point increase in productivity. Regarding perimeter_index estimates, the results show that while the OLS estimate is positive and significant in statistical terms, the 2SLS coefficient is not significant.

| Table 8. Estimates of the relationship between urban form and ldr by OLS and 2SLS |
|------------------|------------------|------------------|------------------|
|                  | OLS (1)          | 2SLS (IV: Norm. potential exchange) (2) | OLS (3)          | 2SLS (IV: Norm. potential perimeter) (4) |
| Norm. exchange   | 0.672***         | 0.958**          |                  |                  |
|                  | (0.1305)         | (0.4298)         |                  |                  |
| Norm. perimeter  |                  | 0.366***         |                  | -0.370           |
|                  |                  | (0.1198)         |                  | (0.3930)         |
| Country dummies  | Y                | Y                | Y                | Y                |
| Year dummies     | Y                | Y                | Y                | Y                |
| N                | 2,757            | 2,757            | 2,757            | 2,757            |
| R-squared        | 0.274            | 0.142            | 0.262            | 0.090            |

**Instrument relevance**  
1. **First-stage statistics**  
   F-stat\(^a\)  
   F-stat P-val  
   62.37  
   0.000  
2. **Under-identification test**  
   Kleibergen-Paap LM stat\(^b\)  
   Chi-sq P-val  
   4.481  
   0.0343  
   5.601  
   0.0179  
3. **Weak identification test**  
   Kleibergen-Paap rk Wald F stat\(^c\)  
   62.37  
   29.22  

\(^a\) F-test of excluded instruments (nexchange or nperimeter) in the first stage model. F-stat above 10 indicates that the instrument has considerable explanatory power.  
\(^b\) The Kleibergen-Paap rank LM test of under-identification tests whether the excluded instruments are correlated with the endogenous regressor.  
\(^c\) The Kleibergen-Paap rank Wald test of weak identification tests the significance of the excluded instruments in the structural equation. The critical values for this test are from Stock and Yogo (2005). The values of this test are higher than the Stock and Yogo (2005) critical values, suggesting that the instruments are not weak.
6. Conclusions

This paper proposes a more integral way to measure urban form and its relationship with economic performance, measured as the density of radiance within the urban extent, \( dr_{ntl\_2010} \), computed as the sum of the NTL DN values in 2010 divided by area, in square kilometers, of the urban extent in 2010. Instead of using population density as a proxy for urban form, we use seven variables that cover the three dimensions of urban form: shape of the urban extent, structure of the urban texture and land use pattern. For the empirical evidence we use two strategies for dealing with the endogeneity between urban form and productivity: First, using lagged explanatory variables as in Fallah et al (2010); and second, using instrumental variables as in Harari (2015).

Using information from 919 urban extents in Latin America, the results show that the shape of the urban extent, the inner-city connectedness, the level of urbanization and the population level have a statistically significant influence over the productivity level of the city. We also found evidence of the presence of a congestion effect on productivity since the results for \( population^2 \) and \( fullness^2 \) are both significant and negatively associated with productivity. According to these results, a compact, dense and well-connected city meets important conditions to be highly productive.

An important consequence of decomposing the concept of urban form into three dimensions is that a non-compact city can reach high levels of productivity by guaranteeing a high level of inner-city connectedness; or, it may happen that a compact city that is poorly connected can show low levels of productivity. Another feasible situation using our way to characterize urban form is that compact cities that are poorly urbanized can have low productivity. In summary, our empirical evidence shows that each city can find its way towards a higher productivity level by analyzing the status of each dimension of its urban form and implementing the corresponding strategies to improve its current conditions.

7. References

Abrahams, A; Lozano-Gracia, N; Oram, C (2016). Deblurring DSMP Nighttime Lights. Working paper. The World Bank, Washington, DC.
Anderson, SJ; Tuttle, BT; Powell, RL; Sutton, PC (2010). Characterizing relationships between population density and nighttime imagery for Denver, Colorado: issues of scale and representation. *International Journal of Remote Sensing*, 31(21): 5733–5746. https://doi.org/10.1080/01431161.2010.496798

Angel, S; Sheppard, SC; Civco, DL (2005). The dynamics of global urban expansion. Transport and Urban Development Department, The World Bank, Washington, DC.

Angel, S; Parent, J; Civco, DL (2010). Ten compactness properties of circles: Measuring shape in geography. *Canadian Geographer*, 54(4): 441–461. https://doi.org/10.1111/j.1541-0064.2009.00304.x

Batty, M (2008). The Size, Scale, and Shape of Cities. *Science* 319, 769.

Batty, M; Longley, P (1994). Fractal Cities: A Geometry of Form and Function Academic Press, London.

Billor, N; Hadi, AS; Velleman, PF (2009). BACON: Blocked adaptive computationally efficient outlier nominators. *Computational Statistics & Data Analysis*, 34: 279–298.

Bleakley, H; Lin, J (2012). Portage and Path Dependence. *Quarterly Journal of Economics*, 127: 587-644.

Boeing, G (2017). OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks. Available at SSRN: https://ssrn.com/abstract=2865501

Bright, EA; Rose, AN; Urban, ML (2013). LandScan 2012. Oak Ridge, TN: Oak Ridge National Laboratory. Retrieved from http://www.ornl.gov/landscan/
Bunnel, T; Drummond, LBW; Ho, KC (2002). Critical Reflections on Cities in Southern Asia. Brill Academic Press, Singapur.

Cao, Z; Wu, Z; Kuang, Y; Huang, N; Wang, M (2016). Coupling an intercalibration of radiance-calibrated nighttime light images and land use/cover data for modeling and analyzing the distribution of GDP in Guangdong, China. Sustainability (Switzerland), 8(2). https://doi.org/10.3390/su8020108

Carroll, M; Townshend, J; DiMiceli, C; Noojipady, P; Sohlberg, R (2009a). Global Raster Water Mask at 250 meter Spatial Resolution, Collection 5: MOD44W MODIS Water Mask. College Park, Maryland: University of Maryland.

Carroll, M; Townshend, J; DiMiceli, C; Noojipady, P; Sohlberg, R (2009b). A New Global Raster Water Mask at 250 Meter Resolution. International Journal of Digital Earth, 2(4).

Center for International Earth Science Information Network - CIESIN - Columbia University, International Food Policy Research Institute - IFPRI, The World Bank, and Centro Internacional de Agricultura Tropical - CIAT. 2011. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Settlement Points. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4M906KR.

Cervero, R (2001). Efficient Urbanisation: Economic Performance and the Shape of the Metropolis. Urban Studies, 38(10):1651-1671.

Cervero, R; Kockelman, K (1997). Travel demand and the 3Ds: Density, Design and Diversity. Transportation Research. Part D Transp. Environ., 2(3):199–219.

Chatman, D; Noland, R (2014). Transit Service, Physical Agglomeration and Productivity in US metropolitan Areas”, Urban Studies, 51(5): 917-937.
Cheng, Y; Zhao, L; Wan, W; Li, L; Yu, T; Gu, X (2016). Extracting urban areas in China using DMSP/OLS nighttime light data integrated with biophysical composition information. *Journal of Geographical Sciences*, 26(3), 325–338. https://doi.org/10.1007/s11442-016-1271-6

Chuvieco, E (2016). Fundamentals of Remote Sensing. An Environmental Approach (2nd ed.). Boca Raton: CRC Press.

Ciccone A; Hall RE (1996). Productivity and the density of economic activity. *American Economic Review*, 86: 54–70.

Couch, C; Ldontidou, L; Petschel-Held, G (2007). Landscapes, land-use change & policy. *Urban sprawl in Europe*. Blackwell Publishing, Oxford.

Duranton, G; Puga D. (2004). “Micro-foundations of urban agglomeration economies”, in J.V. Henderson and J.-F. Thisse, eds., *Handbook of Urban and Regional Economics*, Volume 4, New York: North Holland.

Ellis, Peter, and Mark Roberts. 2015. *Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability*. World Bank Publications.

Fallah, B; Partridge, M; Olfert, M (2011). Urban sprawl and productivity: Evidence from US metropolitan areas, *Paper in Regional Science*, 90(3): 451-473.

Forbes, DJ (2013). Multi-scale analysis of the relationship between economic statistics and DMSP-OLS night light images. *GIScience & Remote Sensing*, 50(5): 83–499. https://doi.org/10.1080/15481603.2013.823732

Freire, S; Pesaresi, M (2015): GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015). European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/jrc-ghsl-ghs_pop_gpw4_globe_r2015a.
García-López, MA; Muñiz, I (2013). Urban spatial structure, agglomeration economies, and economic growth in Barcelona: An intra-metropolitan perspective. *Papers in Regional Science*, 92(3): 515-535.

Gilbert, A (1996). The Mega-City in Latin America. United Nations University Press, USA (1996)

Harari, M (2016). Cities in bad shape: Urban geometry in India. Working paper. The Wharton School, University of Pennsylvania.

Henderson, JV; Storeygard, A; Weil, DN (2012). Measuring Economic Growth from Outer Space. *American Economic Review*, 102 (2): 994–1028. doi:10.1257/aer.102.2.994.

Hijmans, RJ; Cameron, SE; Parra, JL; Jones, PG; Jarvis, A (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25: 1965-1978.

Huang, J; Lu, XX; Sellers, JM (2007). A global comparative analysis of urban form: Applying spatial metrics and remote sensing. *Landscape and Urban Planning*, 82: 184–197.

Hsu, FC; Baugh, KE; Ghosh, T; Zhizhin, M; Elvidge, CD (2015). DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. *Remote Sensing*, 7(2): 1855–1876. https://doi.org/10.3390/rs70201855

Jarvis, A; Reuter, HI; Nelson, A; Guevara, E (2008). Hole-filled SRTM for the globe. Version 4, available from the CGIAR-CSI SRTM 90m Database. Retrieved from http://srtm.csi.cgiar.org

Kleibergen, F; Paap, R (2006). Generalized Reduced Rank Tests using the Singular Value Decomposition. *Journal of Econometrics*, 133(1): 97-126.

Lee B; Gordon P (2007). Urban Spatial Structure and Economic Growth in US Metropolitan Areas. In 46th Annual Meetings of the Western Regional Science Association, at Newport Beach, CA. Accessed March 15, 2017.
Lee, B; Gordon, P (2011). Urban structure: its role in urban growth, net new business formation and industrial churn, *Région et Développement*, 33: 137-159.

Li, X; Zhou, Y (2017). Urban mapping using DMSP/OLS stable night-time light: a review. *International Journal of Remote Sensing*, 0(0): 1–17. https://doi.org/10.1080/01431161.2016.1274451

Lo, CP (2001). Modeling the Population of China Using DMSP Operational Linescan System Nighttime Data. *Photogrammetric Engineering and Remote Sensing*, 67(9): 1037–1047.

Lowe, M (2014). The Privatization of African Rail. Working Paper, Massachusetts Institute of Technology, Boston MA.

Ma, L; Wu, J; Li, W; Peng, J; Liu, H (2014). Evaluating saturation correction methods for DMSP/OLS nighttime light data: A case study from China’s cities. *Remote Sensing*, 6(10): 9853–9872. https://doi.org/10.3390/rs6109853

Meijers, E; Burger, M (2010). Spatial structure and productivity in US metropolitan areas, *Environment and planning A*, 42: 1383-1402.

Melo, P; Graham, D; Levinson, D; Aarabi, S (2017). Agglomeration, accessibility and productivity: Evidence for large metropolitan areas in the US. *Urban Studies*, 54(1): 179-195.

Michalopoulos, S; Papaioannou, E (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, 81(1):113-152.

Nunn, N; Puga, D (2012). Ruggedness: the blessing of bad geography in Africa. *The Review of Economics and Statistics*, 94(1): 20–36.
Pandey, B; Joshi, PK; Seto, KC (2013). Monitoring urbanization dynamics in India using DMSP/OLS night time lights and SPOT-VGT data. *International Journal of Applied Earth Observation and Geoinformation*, 23: 49–61. https://doi.org/10.1016/j.jag.2012.11.005

Partridge, MD; Rickman, DS; Kamar, A; Olfert, MR (2008). The geographic diversity of U.S. nonmetropolitan growth dynamics: A geographically weighted regression approach. *Land Economics*, 84: 241–266

Parr, JB (1979). Regional economic change and regional spatial structure: some interrelationships, *Environment and Planning A*, 11(7): 825 – 837.

Pesaresi, Martino; Ehrilch, Daniele; Florczyk, Aneta J.; Freire, Sergio; Julea, Andreea; Kemper, Thomas; Soille, Pierre; Syrris, Vasileios (2015): GHS built-up grid, derived from Landsat, multitemporal (1975, 1990, 2000, 2014). European Commission, Joint Research Centre (JRC) [Dataset] PID: http://data.europa.eu/89h/jrc-ghsl-ghs_built_ldsmt_globe_r2015b

Pesaresi, M; Ehrlich, D; Ferri, S; Florczyk, A.J; Freire, S; Halkia, M; Syrris, V (2016). Operating procedure for the production of the Global Human Settlement Layer from Landsat data of the epochs 1975, 1990, 2000, and 2014; JRC Technical Report EUR 27741 EN. Ispra (VA), Italy. https://doi.org/10.2788/253582

Pinkovskiy, M (2013). Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night. Working paper, Massachusetts Institute of Technology, Boston MA.

Quigley, J (1998). Urban Diversity and Economic Growth. *Journal of Economic Perspectives*, 12: 127-138.

Rauch, J (1993a), Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics*, 34: 380-400.
Roberts, M; Blankespoor, B; Deuskar, C; Stewart, B (2016). Urbanization and Development: Is Latin America and the Caribbean Different from the Rest of the World?. The World Bank.

Rosenthal, SS; Strange, WC (2004). Evidence on the nature and sources of agglomeration economies, in J.V. Henderson and J.-F. Thisse, eds., Handbook of Urban and Regional Economics, Volume 4, New York: North Holland, 2119-2171.

Rubiera, F; González, V; Pérez, J (2017). Urban Sprawl in Madrid? An Analysis of the Urban Growth of Madrid During the last Quarter of the Twentieth Century. Letters in Spatial and Resource Sciences, forthcoming.

Saiz, A (2010). The Geographic Determinants of Housing Supply. Quarterly Journal of Economics, 125(3): 1253–1296. https://doi.org/10.1162/qjec.2010.125.3.1253

Squires, GD (2002). Sprawl: Causes and Consequences and Policy Responses. The Urban Institute Press, Washington, D.C.

Sutton, PC; Cova, TJ; Elvidge, CD (2006). Mapping “Exurbia” in the conterminous United States using nighttime satellite imagery. Geocarto International, 21(2): 39–45. https://doi.org/10.1080/10106040608542382

Sutton, PC; Goetz, AR; Fildes, S; Forster, C; Ghosh, T (2010). Darkness on the edge of town: Mapping urban and peri-urban Australia using nighttime satellite imagery. The Professional Geographer, 62(1): 119–133. https://doi.org/10.1080/00330120903405006

Stock, J; Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In D. Stock, K. A. A. J. H. (ed) Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg. Cambridge: Cambridge University Press

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. doi: 10.1093/restud/rdw020
Thompson, DW (1952). On Growth and Form (2nd edition, vol. 1). Cambridge: Cambridge University Press.

Tewari, M; Alder, S; Roberts, M (2016). Patterns of India’s urban and spatial development in the post-reform period: an empirical analysis. Working paper.

Uchida, H; Nelson, A (2008). Agglomeration Index: Towards a New Measure of Urban Concentration. World Development Report: Reshaping Economic Geography, 19. Retrieved from http://siteresources.worldbank.org/INTWDR2009/Resources/4231006-1204741572978/Hiro1.pdf

UN-Habitat (2012). State of Latin American and Caribbean Cities 2012. Towards a new urban transition. Naples: UN-Habitat.

Weber, S (2010). Bacon: An effective way to detect outliers in multivariate data using Stata (and Mata). The Stata Journal, 10(3): 331–338.

Whyte, W (1968). The Last Landscape, Garden City, NY: Doubleday.

World Bank (2009). World Development Report 2009: Reshaping Economic Geography. World Bank. © World Bank. https://openknowledge.worldbank.org/handle/10986/5991 License: CC BY 3.0 IGO.

Wu, J; Ma, L; Li, W; Peng, J; Liu, H (2014). Dynamics of urban density in china: Estimations based on DMSP/OLS nighttime light data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(10): 4266–4275. https://doi.org/10.1109/JSTARS.2014.2367131
Zhang, Q; Seto, K.C (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment*, 115(9): 2320–2329. https://doi.org/10.1016/j.rse.2011.04.032

Zhou, Y; Smith, S. J; Zhao, K; Imhoff, M; Thomson, A; Bond-Lamberty, B; Elvidge, CD (2015). A global map of urban extent from nightlights. *Environmental Research Letters*, 10(5): 54011. https://doi.org/10.1088/1748-9326/10/5/054011

Zhou, N; Hubacek, K; Roberts, M (2015). Analysis of spatial patterns of urban growth across South Asia using DMSP-OLS nighttime lights data. *Applied Geography*, 63: 292–303. https://doi.org/10.1016/j.apgeog.2015.06.016
Appendix

**Table A1.** Descriptive statistics of the data used in equation (2).

| Variable | p25 | Median | p75 | Mean | Std. Dev. | Min | Max |
|----------|-----|--------|-----|------|-----------|-----|-----|
| $d_r$   | 360.067 | 514.331 | 733.474 | 587.184 | 318.228 | 170.060 | 3,759.463 |
| 1996 norm. exchange | 0.725 | 0.828 | 0.879 | 0.787 | 0.121 | 0.266 | 0.947 |
| 1996 norm. perimeter | 0.620 | 0.720 | 0.767 | 0.674 | 0.140 | 0.090 | 0.888 |
| 1990 fullness | 0.494 | 0.630 | 0.763 | 0.618 | 0.192 | 0.014 | 0.996 |
| 1990 population | 21,204 | 35,440 | 73,436 | 117,762.1 | 523,030.5 | 259 | 9,597,255 |
| 1996 intersection density | 0.470 | 0.641 | 0.841 | 0.659 | 0.289 | 0.003 | 1.850 |
| 1996 street density | 7.988 | 10.452 | 12.989 | 10.474 | 3.792 | 0.136 | 20.669 |
| 1996 circuity | 1.019 | 1.028 | 1.046 | 1.037 | 0.034 | 1.002 | 1.419 |
| 1990 sprawl | 0.478 | 0.595 | 0.721 | 0.598 | 0.177 | 0.108 | 1.00 |
| Distance in km to international border | 0.152 | 0.414 | 0.765 | 0.556 | 0.542 | 0.000 | 2,461 |
| Temperature | 18.954 | 22.133 | 25.476 | 21.742 | 4.331 | 5.486 | 28.897 |
| Precipitation | 838.185 | 1,244.574 | 1,532.889 | 1,239.906 | 667.526 | 0.519 | 7,250.133 |
| Coast | 0.000 | 0.000 | 0.000 | 0.174 | 0.379 | 0.000 | 1.00 |

**Table A2.** Descriptive statistics of the data used in equation (3) and (4).

| Variable | p25 | Median | p75 | Mean | Std. Dev. | Min | Max |
|----------|-----|--------|-----|------|-----------|-----|-----|
| $d_r$   | 329.377 | 465.960 | 676.922 | 545.439 | 310.412 | 138.902 | 3,759.463 |
| Norm. actual exchange | 0.720 | 0.823 | 0.879 | 0.786 | 0.121 | 0.266 | 0.952 |
| Norm. actual perimeter | 0.596 | 0.712 | 0.760 | 0.660 | 0.143 | 0.090 | 0.888 |
| Norm. potential exchange | 0.888 | 0.977 | 1.000 | 0.927 | 0.104 | 0.245 | 1.000 |
| Norm. potential perimeter | 0.505 | 0.796 | 1.000 | 0.726 | 0.272 | 0.095 | 1.000 |
| Temperature | 18.979 | 22.147 | 25.475 | 21.756 | 4.323 | 5.486 | 28.897 |
| Coast | 0.000 | 0.000 | 0.000 | 0.173 | 0.379 | 0.000 | 1.00 |
### Table A3. Estimates of the relationship between urban form and productivity (OLS). Without outliers.

(Independent variable: ldr)

|                  | (1)       | (2)       | (3)       | (4)       | (5)       | (6)       | (7)       | (8)       | (9)       | (10)      |
|------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 1996 norm. exchange | 0.625***  | 0.594***  | 0.588***  | 0.549***  |           |           |           |           |           |           |
|                  | (0.0912)  | (0.0932)  | (0.0907)  | (0.0927)  |           |           |           |           |           |           |
| 1996 norm. perimeter | 0.906***  |           |           |           |           |           |           |           |           |
|                  | (0.1034)  |           |           |           |           |           |           |           |           |
| 1990 fullness    |           | 0.821***  | 0.692***  | 0.648**   |           |           |           |           |           |
|                  |           | (0.2573)  | (0.2456)  | (0.2366)  |           |           |           |           |           |
| 1990 fullness²   |           | -0.788*** | -0.638*** | -0.586*** |           |           |           |           |           |
|                  |           | (0.2225)  | (0.2193)  | (0.2025)  |           |           |           |           |           |
| 1996 circuity    | 0.003     | -0.041    | -0.015    | -0.057    | 0.150     | 0.109     | 0.067     | 0.149     | 0.102     |
|                  | (0.7643)  | (0.7339)  | (0.771)   | (0.7553)  | (0.7904)  | (0.7628)  | (0.8102)  | (0.7802)  | (0.7488)  |
| 1996 intersection density | 0.411*** | 0.408***  | 0.454***  | 0.404***  | 0.388***  |           |           |           |           |
|                  | (0.0652)  | (0.0567)  | (0.0730)  | (0.0806)  | (0.0738)  |           |           |           |           |
| 1996 street density |           |           |           |           |           | 0.035***  | 0.035***  | 0.040***  | 0.035***  |
|                  |           |           |           |           |           | (0.0052)  | (0.0047)  | (0.0058)  | (0.0062)  |
| 1990 sprawl      | 0.088     | 0.133     | 0.143     | 0.185     |           |           |           |           |           |
|                  | (0.1182)  | (0.1264)  | (0.1181)  | (0.1246)  |           |           |           |           |           |
| 1990 population  | 0.132*    | 0.200***  | 0.098*    | 0.133**   | 0.192***  | 0.125     | 0.191***  | 0.125*    | 0.181***  |
|                  | (0.0748)  | (0.0637)  | (0.0559)  | (0.0643)  | (0.0518)  | (0.0752)  | (0.0638)  | (0.0651)  | (0.0520)  |
| 1990 population² | -0.012**  | -0.015*** | -0.009**  | -0.012**  | -0.014*** | -0.011*   | -0.014*** | -0.009**  | -0.012*** |
|                  | (0.0055)  | (0.0051)  | (0.0042)  | (0.0050)  | (0.0044)  | (0.0056)  | (0.0051)  | (0.0043)  | (0.0044)  |
| Constant         | 5.399***  | 5.233***  | 5.129***  | 4.962***  | 5.159***  | 5.065***  |           |           |           |
|                  | (0.8200)  | (0.7106)  | (0.7334)  | (0.7488)  | (0.6160)  | (0.8473)  | (0.7538)  | (0.7631)  | (0.6459)  |

Country dummies: Y Y Y Y Y Y Y Y Y Y

| N       | 911     | 911     | 911     | 911     | 911     | 911     | 911     | 911     | 911     |
| R-squared | 0.309   | 0.319   | 0.297   | 0.312   | 0.321   | 0.314   | 0.323   | 0.304   | 0.317   | 0.325   |

Note: Robust standard errors clustered at the country level in parentheses.

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

All models include geographical characteristics as measured of natural amenities, namely, distance in km to international border, temperature, precipitation and coast indicator.

Intersection_density has been divided by 100; street_density has been divided by 1,000; population has been divided by 100,000; and the control dist_border_kms has been divided by 1,000.
Table A4. Estimates of the first stage of instrumental variables regressions. With outliers.

|                  | OLS Dependent var: Norm. actual exchange | OLS Dependent var: Norm. actual perimeter |
|------------------|------------------------------------------|-------------------------------------------|
| Norm. potential exchange (1) | 0.263*** (0.0333) | 0.118*** (0.0218) |
| Country dummies   | Y                                       | Y                                         |
| Year dummies      | Y                                       | Y                                         |
| N                | 2,757                                   | 2,757                                     |

Note: Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1
All models include geographical characteristics as measured of natural amenities, namely, temperature and coast indicator.

Table A5. Estimates of the first stage of instrumental variables regressions. Without outliers.

|                  | OLS Dependent var: Norm. actual exchange | OLS Dependent var: Norm. actual perimeter |
|------------------|------------------------------------------|-------------------------------------------|
| Norm. potential exchange (1) | 0.265*** (0.0336) | 0.100*** (0.0212) |
| Country dummies   | Y                                       | Y                                         |
| Year dummies      | Y                                       | Y                                         |
| N                | 2,733                                   | 2,733                                     |

Note: Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1
All models include geographical characteristics as measured of natural amenities, namely, temperature and coast indicator.
Table A6. Estimates of the relationship between urban form and ldr by OLS and 2SLS. Without outliers
(Dependent variable: ldr)

|                | OLS (1) | 2SLS (IV: Norm. potential exchange) (2) | OLS (3) | 2SLS (IV: Norm. potential perimeter) (4) |
|----------------|---------|-----------------------------------------|---------|-----------------------------------------|
| Norm. exchange | 0.672*** | 0.970**                                 |         |                                         |
|                | (0.1311) | (0.4313)                                |         |                                         |
| Norm. perimeter|         |                                         | 0.371*** | -0.487                                  |
|                |         |                                         | (0.1377) | (0.4662)                                |
| Country dummies| Y       | Y                                       | Y       | Y                                       |
| Year dummies   | Y       | Y                                       | Y       | Y                                       |
| N              | 2,733   | 2,733                                   | 2,733   | 2,733                                   |
| R-squared      | 0.275   | 0.143                                   | 0.262   | 0.080                                   |

**Instrument relevance**

1. **First-stage statistics**
   - F-stat\(^a\) 62.02 22.26
   - F-stat P-val 0.000 0.000

2. **Under-identification test**
   - Kleibergen-Paap rk LM stat\(^b\) 4.505 6.223
   - Chi-sq P-val 0.0338 0.0126

3. **Weak identification test**
   - Kleibergen-Paap rk Wald F stat\(^c\) 62.02 22.26

Note: Robust standard errors clustered at the country level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

All models include geographical characteristics as measured of natural amenities, namely, temperature and coast indicator.

\(^a\) F-test of excluded instruments (nexchangep or nperimeterp) in the first stage model. F-stat above 10 indicates that the instrument has considerable explanatory power.

\(^b\) The Kleibergen-Paap rank LM test of under-identification tests whether the excluded instruments are correlated with the endogenous regressor.

\(^c\) The Kleibergen-Paap rank Wald test of weak identification tests the significance of the excluded instruments in the structural equation. The critical values for this test are from Stock and Yogo (2005). The values of this test are higher than the Stock and Yogo (2005) critical values, suggesting that the instruments are not weak.