Non-trivial relationship between scaling behavior and the GDP microstructure in Indonesian cities

Genta Kuno^{1*}, Pradipto²

¹ Center for Southeast Asian Studies, Kyoto University, Yoshidashimoadachi-cho, Sakyo-ku, Kyoto, 606-8304, Japan

² Yukawa Institute for Theoretical Physics, Kyoto University, Kitashirawaoiwake-cho, Sakyo-ku, Kyoto, 606-8502, Japan

These authors contributed equally to this work.

* kuno@asafas.kyoto-u.ac.jp

Abstract

Urban scaling analysis has shown that various aggregated urban quantities obey power-law relationships with the population size. Despite the rapid progress, direct empirical evidence that shows how the power-law exponents β depend on the cities microstructure has been lacking. Moreover, urban scaling studies are hardly reproduced in developing countries due to inadequate official statistics. We tackle these issues by performing urban scaling analysis on Indonesian cities using globally harmonized functional cities delineations and global-scale gridded Gross Domestic Product (GDP) datasets. First, we observe that the GDP and area of Indonesian cities scale linearly with the population size. For GDP in particular, the deviations from the scaling law follow a geographical pattern. Second, we determine the economic hotspots in each city and observe that the area of the hotspots scales mildly sublinear with the population size. Interestingly, the GDP of hotspots scales sublinearly with the population size, indicating a physical growth of hotspots that is not followed by the increasing returns due to scaling. Third, by classifying the cities based on the GDP microstructure in two dimensions (heterogeneity and spatial dispersion) and examining the scaling exponents of each class, we discover a non-trivial relation between scaling behavior and the GDP microstructure. Spatial dispersion strongly affects the scaling behavior in heterogeneous cities, while such effect is weakened for homogeneous cities. Finally, we find that the scaling effect in terms of economies of scale (sublinearity of area) and increasing returns (superlinearity of GDP) is stronger for cities with spatially compact GDP distribution.

Introduction

A comprehensive understanding regarding the fundamental process of urban system is essential to achieve sustainable urban development. The recent paradigm of viewing cities as complex systems [1] has opened the way to elucidate the intricate mechanism beneath the urban system. One of the most celebrated findings within this paradigm is the *urban scaling* framework that explains the systematic variation of urban aggregate indicators with the population size in accordance with the power-law relations

$$Y = Y_0 x^\beta, \quad (1)$$

where Y is the aggregated urban quantities, Y_0 is a constant prefactor, x is the population of the city, and β is the scaling exponent [2–8]. Despite its limitations [9, 10],

three arguments can be made from urban scaling based on the value of β that takes similar values for classes of quantities: (i) β greater than 1 (superlinear scaling) for social interactions driven quantities that experience the increasing return due to scaling, (ii) β less than 1 (sublinear scaling) for quantities that are characterized by the economies of scale, and (iii) $\beta \approx 1$ for quantities associated with individual human needs. Such power-law exponents have been recovered theoretically from micro-level interactions [11, 12], as well as the fractal dimensions of road networks and population distributions [13]. Numerous studies have also been conducted to understand how the scaling exponent β and its deviation depend on another urban indicator or classification. It has been shown that β are affected by cities boundary [10], dragon-king effect [14], spatial hierarchy of regions [15], and commuting network between cities [16]. Others explore how deviation from the scaling law demonstrates a geographical pattern in Indian cities [5, 8]. Then, the dependence of β for economic quantities on its internal classification has been shown in several studies such as Refs. [17–19] for income levels in Australia and United States, Ref. [20] for Gross Domestic Product (GDP) per industry types in China, and Ref. [21] for income of different occupations.

Some studies also tried to clarify the role of the spatial microstructure of cities on scaling behaviors. Surprisingly enough, a clear empirical evidence that directly connects the scaling exponent β for GDP and cities microstructure has yet to be found. One measure of city microstructure is the spatial dispersion of various urban quantities that characterized its activity centers i.e the hotspots. Such a measure can capture the degree of polycentricity of a city (a monocentric city is less dispersed than the polycentric counterpart). It has been implicitly suggested that higher polycentricity weakens the increasing return due to scaling in US cities [22]. Then, a stochastic and out-of-equilibrium model of cities has been proposed to demonstrate the transition from monocentric to polycentric structure [23]. The model predicts that the number of hotspots in a city scales sublinearly with the population size, which has been confirmed empirically in Spanish cities through a non-parametric hotspots detection technique based on the density of mobile phone users [24]. An implicit connection has been made through the model of Ref. [23], where the model demonstrates that the diseconomy due to congestion scales superlinearly with the population and such diseconomy is weaker in cities with higher polycentricity [25]. Ref. [26] analyzed economic hotspots based on night-time light intensity in the United States, European Union, and Chinese cities. They observed an inverted U-shape curve between productivity and compactness i.e as cities become more compact, the GDP per km² increases until critical compactness and then decreases afterward. A more straightforward approach to connect the exponent β for economic output and cities microstructure has been reported in Ref. [27], where the authors elaborate the role of mesoscopic (2 km² grid size) interactions and active population in macroscopic aggregated quantities. They demonstrated a positive correlation between Moran’s I index and β through numerical simulations. Yet, such correlation has not been supported by empirical observations. On top of that, measurement of spatial dispersions like Moran’s I index alone cannot distinguish between polycentric structure and discontinuous development [28, 29]. Therefore, it is necessary to also consider the heterogeneity of the urban quantities upon which the hotspots are defined.

Most of the urban scaling studies have been conducted for cities in developed countries with adequate official statistics. It is concerning that only a few are done in developing countries [5, 6] since the major parts of rapidly urbanizing areas of the globe are in the developing world [30]. Also, cities in low- and middle-income countries have suffered inequality worse than cities in high-income countries [31]. To overcome this challenge, we adopt the recently published dataset of Functional Urban Areas (FUA) [32, 33], which has been used for global-scale urban scaling analysis on GDP [34].

The availability of such harmonized definition of functional cities enables us to set our locus on cities in Indonesia, a typical case of a developing country without adequate official definition of cities, while also lacking derivative socio-economic data, yet approximately 150 million people (56% of the total population) belongs to urban population [35]. In this paper, we perform urban scaling analysis in Indonesian cities for area and GDP based on gridded data with 30 arc-sec ($\approx 1\text{km}^2$) resolution [36]. We then define the economic hotpots from gridded GDP data and discuss their scaling behavior. Finally, we classify cities based on the method of Ref. [29] that captures both the heterogeneity and spatial dispersion for gridded GDP data. With 238 Indonesian cities in the FUA dataset, we gain the merit of having a sufficient number of cities to create a classification, while still grasping the country-specific context. Moreover, the usage of a global-scale publicly available dataset ensures reproducibility of our analysis in other countries as well. More details are elaborated in Methods.

Fig 1. Scaling law in Indonesian cities. Plots of urban indicators of 238 Indonesian functional cities against population size. (a) GDP. (b) Area of the cities. (c) Area of the urban center and the commuting zone. Dashed lines represent the fit from log-normal model with free δ .

Results

Scaling behavior of GDP and area of Indonesian cities

We start with general scaling analysis as we plot the GDP, area, and the corresponding area of the urban center and the commuting zone against the population in Fig. 1. One can see that the for all quantities, the fluctuation tends to be larger in small population regime. Such data where the fluctuation has systematic variation should not be fitted with simple linear least-squares for log-transformed quantities. Instead, the data is fitted using the log-normal distribution with free δ [37] (See Methods for details). First, the GDP in Fig. 1(a) has approximately linear scaling ($\beta = 1.004 \pm 0.072$), violating the generally consistent observation of superlinearity of economic activities in both developed [2, 4] and developing countries [6, 38].

The linear scaling of GDP in Indonesian cities is similar to the one observed in Indian cities, where an association between a linear scaling of GDP and geographic gap in urban productivity is noted [8]. We investigate the geographic difference in GDP by defining the deviation from the scaling law ξ_i as $\xi_i = \log(Y_i/Y_0 N^\beta)$ [5, 8]. Map of ξ_i for GDP of each city can be seen in Fig. 2. Then, we classify the archipelago into regions based on their timezone. Looking at the Central and East Indonesia, a dominant presence of underperforming cities can be observed with some exceptions in the eastern part of Kalimantan, which are characterized by extractive industries similar to cities in eastern Sumatra. Looking at West Indonesia, Sumatra and Java have an even distribution of underperforming ($\xi_i < 0$) and overperforming cities ($\xi_i > 0$), while the

Fig 2. Geographical pattern of GDP scaling law deviations. Map of deviations from the GDP scaling law ξ_i . Red circles represent underperformance w.r.t. scaling law ($\xi_i < 0$) and blue circles represent overperformance ($\xi_i > 0$). The size of the circles represents the absolute magnitude of the deviation. The colors of islands represent different regions per time-zone (grey for GMT+7 or West Indonesia, green for GMT+8 or Central Indonesia, and pink for GMT+9 or East Indonesia). Several main islands in the archipelago are annotated as: West Indonesia (GMT+7) consisting of islands of Sumatra, Java, and the western part of Kalimantan. Central Indonesia (GMT+8) consisting of the eastern part of Kalimantan, islands of Bali, Nusa Tenggara, and Sulawesi. Then East Indonesia (GMT+9) consisting of islands of Maluku and Papua. Right inset: Violin plot of the distribution of scaling deviation grouped by the domestic time-zones.

western part of Kalimantan is dominated by underperforming ones. There are several exceptional overperforming cities in the eastern part of Sumatra, where the economy of these cities is driven by oil mines and refineries. Furthermore, the magnitude of deviations is lower in Java, the island with largest share of cities and population. This implies that the suppressed fluctuation in higher population regime (Fig. 1) is coupled with the geographical constraint. The pattern is summarized in the violin plot of the distribution of ξ_i per domestic time-zone (inset of Fig. 2), as the distribution is shifting to the left (more underperforming) from west to east with a few and almost no positive deviation (overperforming cities) in East Indonesia. This observation portrays the economic and development disparities between Java and the rest of Indonesia [39–41].

A slightly sublinear scaling for area ($\beta = 0.991 \pm 0.05$) with higher β than the typically observed β for area in developed countries [2, 4] is observed (Fig. 1(b)), indicating an inefficient physical growth. Each city consists of urban center and the corresponding commuting zone. To understand the underlying mechanism beneath the overall inefficiency of physical growth, we examine the scaling law of urban center and commuting zone areas respectively in Fig. 1(c). We find a slightly superlinear scaling for urban center area ($\beta = 1.044 \pm 0.039$) and the area of commuting zone scales sublinearly with the population ($\beta = 0.948 \pm 0.066$). While the physical expansion of commuting zone is as efficient as expected by the urban scaling theory, that of the urban center is not. Thus, it is the urban center area that disobeys the economies of scale effect due to scaling and should be clarified further.

Scaling behavior of economic hotspots

City microstructure is characterized by the activity centers i.e the hotspots. Previous attempts define the hotspots from population density [24, 29]. However, population density does not necessarily represent the amount of activity [27]. Thus, our approach to capture the amount of activity is by directly analyzing the GDP density, which

Fig 3. Scaling law of economic hotspots in Indonesian cities. Plots of (a) GDP and (b) Area of hotspots (blue) and non-hotspots (red) against population size, dashed lines represent the fit from log-normal model with free δ .

represents the economic activity. This aligns with Ref. [26], where they define the economic hotspots through night-time light intensity. This way, the hotspots are defined as the local maxima in the surface of GDP density. From the gridded GDP data, we define the hotspots as the cells with GDP value larger than the Louail-Barthelemy (LouBar) threshold GDP^* [24] (See Methods). Conversely, the non-hotspots are the cells with $GDP \leq GDP^*$. In Fig. 3 the scaling behaviors of the area and GDP of both hotspots and non-hotspots are examined. For area (Fig. 3(a)), we found that the hotspots scale sublinearly ($\beta = 0.964 \pm 0.058$) with population size, while the non-hotspots obey a superlinear scaling ($\beta = 1.045 \pm 0.066$). Note that similar to β of urban center area, β of hotspots area in our observation is higher than that was observed in previous research in developed countries [24, 26]. The hotspots tend to cover the core of the urban center, while the non-hotspots tend to cover parts of urban center and the commuting zone. The area of commuting zone scales sublinearly with the population (cf. Fig 1(b)), which implies that the non-hotspots part of the urban centers scales superlinearly with the population. Therefore, the superlinearity of the urban center area contains the contributions from the mildly sublinear area of the hotspots part and the superlinear area of the non-hotspots part.

While the non-hotspots violate the scaling effects in terms of economies of scale, Fig. 3(b) for GDP shows that they experienced the increasing return due to scaling ($\beta = 1.072 \pm 0.055$). Such superlinearity of non-hotspots GDP indicates a transition from semi-urban to urban economics since small cities in Indonesia covers various kind of settlement entities including the peri-urban areas with mixtures of urban and rural features [42]. In contrast, the GDP of the hotspots scale sublinearly with population size ($\beta = 0.925 \pm 0.092$). This indicates that the massive physical expansions of urban hotspots in Indonesian cities are not followed by the economic counterparts.

Non-trivial relations between the GDP microstructure and the scaling exponents β

Let us discuss the connection between GDP microstructure and scaling exponents. Here we are interested in two dimensions of GDP microstructure: spatial dispersion and heterogeneity. The spatial dispersion or the compactness of GDP can be captured by calculating the spreading index η , which is the ratio between the average distance of hotspots and the average distance of all cells in the city [29]. Then, the heterogeneity of the GDP distribution within cities is measured through the normalized Gini index G_{norm} (See Methods for a complete description of G_{norm} and η). In Fig. 4(a), we

Fig 4. Classification of Indonesian cities based on the GDP microstructure.

(a) Spreading index of hotspots η versus normalized Gini coefficient G_{norm} , circle size represents population size, color represents the classifications. Green, brown, purple, and blue for Homo-Poly, Hete-Poly, Hete-Mono, and Homo-Mono classes, respectively. (b) Spatial distributions of GDP (in 30 arc-sec grid size) of the representative cities of each class defined in (a): (top left) City of Jombang, representative of Homo-Poly class, with $G_{\text{norm}} = 0.44$ and $\eta = 1.02$. (top right) City of Pamekasan, representative of the Hete-Poly Class, with $G_{\text{norm}} = 0.67$ and $\eta = 0.89$. (bottom left) City of Padang, representative of the Homo-Mono class, with $G_{\text{norm}} = 0.54$ and $\eta = 0.47$. (bottom right) City of Makassar, representative of the Hete-Mono class, with $G_{\text{norm}} = 0.71$ and $\eta = 0.36$.

represent the cities in the plane (G_{norm}, η) . We find a negative correlation between the two variables, which means that for higher GDP heterogeneity (or higher inequality), the GDP distribution becomes more compact. Consistent with past attempts [29, 43], no relation among G_{norm} and η with the population size is found. We then split the cities into quadrants by medians of G_{norm} and η . This way, we have four classes of cities based on their GDP microstructure: (i) 85 cities with relatively low G_{norm} and high η , which are denoted as *Homo-Poly* class (green symbols in Fig. 4(a), top left in Fig. 4(b)). Homo-Poly cities are associated with a homogeneous and spatially dispersed structure (more polycentric) of GDP distribution. (ii) 34 cities with high G_{norm} and high η , denoted as *Hete-Poly* class (brown symbols in Fig. 4(a), top right in Fig. 4(b)). Hete-Poly cities have a heterogeneous and spatially dispersed GDP distribution, which are often associated with leapfrog development [28]. (iii) 34 cities with intermediate G_{norm} and low η , denoted as *Homo-Mono* class (blue symbols in Fig. 4(a), bottom left in Fig. 4(b)). Cities in Homo-Mono class tend to have a monocentric structure of GDP distribution with intermediary heterogeneity. (iv) 85 cities with high G_{norm} and low η , denoted as *Hete-Mono* class (purple symbols in Fig. 4(a), bottom right in Fig. 4(b)). Cities that belong to the Hete-Mono class have a heterogeneous and compact GDP distribution.

Then, the scaling behavior of cities in each class is examined. From Fig. 5 that shows β of the corresponding classes, we observed a non-trivial relation between GDP microstructure and the scaling exponents since each class has a distinct scaling behavior. One can see that the microstructure strongly affects the scaling exponents of all quantities of interest with a couple of exceptions: (i) Scaling exponents of urban center area and hotspots area are only weakly affected since no huge differences in β , yet transition from sublinear to superlinear scaling takes place. (ii) Scaling exponents of non-hotspots GDP is almost independent of microstructure since neither transition nor

Fig 5. Non-trivial relationship between scaling exponents and GDP microstructure. Scaling exponents β of each city class in Fig. 4(a), with the exponents of all cities for guidance (grey circles). Grey shaded region contains the aggregate urban quantities. Red shaded region contains urban quantities of hotspots and non-hotspots.

huge differences are observed. The microstructure-independent β of non-hotspots GDP means that the semi-urban to urban transition takes place regardless of the microstructure. For aggregated quantities, a pattern exists such that β for areal quantities (total area, urban center area, and commuting area) of heterogeneous cities (Hete-Poly and Hete-Mono classes) are strongly affected by η , while that of the homogeneous cities (Homo-Poly and Homo-Mono classes) are only weakly affected. Then, β of GDP and hotspot GDP for monocentric cities are always larger than that for polycentric cities. Meanwhile, β of commuting area, area, urban center area, and non-hotspots area of monocentric cities tend to be smaller than that for polycentric cities.

For spatially dispersed or polycentric cities (Hete-Poly and Homo-Poly classes), β for areal quantities are always larger than that for GDP, which are the opposite of theoretical predictions [11–13] and previous observations in various countries [2–6, 8]. The aggregated and hotspots GDPs of polycentric cities scale sublinearly with populations, meaning that they experience a decreasing return due to scaling. Then, polycentric cities have weaker economies of scale effect from scaling since β for aggregated areal quantities are always larger than the monocentric counterparts. Moreover, Hete-Poly class is the only class that violates the shared trends in other classes, where β gets gradually larger beginning from commuting zone area, area, then urban center area. On the other hand, the scaling behavior for spatially compact (monocentric) cities (Hete-Mono and Homo-Mono classes) obeys the general trends and theoretical predictions i.e β for areal quantities are smaller than that for GDP. Yet, both classes have distinct hotspots scaling behavior, where the hotspots GDP of Homo-Mono cities scales superlinearly with the population size yet the hotspots area obeys a linear scaling. On the contrary, hotspots GDP of Hete-Mono cities obeys linear scaling while their hotspots area scales sublinearly with the population size. In other words, a trade-off relation exists for hotspots in monocentric Indonesian cities such that the increasing returns due to scaling comes at the expense of lacking the scaling effect in

terms of economies of scale and vice versa. Regardless of their heterogeneity, hotspots in monocentric cities in Indonesia experience the physical expansions of hotspots that are not followed by the economic growth.

Discussions

Our analysis has shown the lack of economic growth despite massive physical expansions in Indonesian cities. This might be caused by two possible mechanisms: (i) As Indonesian cities grow in populations, the amount of social interaction in larger cities exceeds the optimum value so that the cost or dissipation of such interactions overcome the benefits [11]. Such non-monotonic behavior of social interactions might have a relation with the inverted U-shaped effect of hotspots compactness on economic growth [26]. (ii) There are socio-economic barriers that could not be captured through our analysis with approximately 1 km² grid size and LouBar threshold since they neither capture the GDP spread *within* hotspots nor the variation at finer resolutions, which may epitomize the Modifiable Area Unit Problem [44,45]. Within the Indonesian context (or other developing countries), the over-cost accumulation of interaction in large cities can be associated with the so-called *urban involution* state which is often featured by proliferation of informal settlements and economies [46–49]. It has been shown that cities have distinct topological complexities at block-level, which are able to characterize informal settlements [50,51]. One should discuss in the future how the urban scaling exponents are affected by this block-level topological complexity, since it might constrain the social interactions that are necessary to recover the superlinearity of urban economic outputs [11,12].

We have discovered a non-trivial relationship between microstructure (heterogeneity and spatial dispersion) of urban economy and scaling exponents in Indonesian cities. We found that the scaling behavior of aggregated quantities in heterogeneous cities strongly depends on the spatial dispersion of GDP, while homogeneous cities are weakly affected. The non-trivial relationship reveals that the scaling effect in terms of economies of scale (sublinearity of area) and increasing returns (superlinearity of GDP) is stronger for cities with spatially compact (monocentric) GDP distribution, which has been suggested earlier by Ref. [22,27]. It can be hypothetically assumed that all social interactions that lead to superlinearity take place in the hotspots. When the hotspots are spatially adjacent, the contact probability per person is higher than the spatially non-adjacent (spreading) case. Furthermore, a polycentric city can be naively regarded as an adjacent collection of smaller cities, where each of them have their own urban cores. Thus, when polycentric cities grow in size, the surrounding areas of each core sprawls, leading to larger physical expansions. Although monocentric cities align better with the predictions from urban scaling theory, this does not necessarily imply that monocentricity improves the quality of life. For instance, it has been shown that compared to the polycentric cities, monocentric cities suffer more from diseconomy due to congestion [9]. Also, some observed that the increasing return from scaling comes hand in hand with inequalities [18,19]. Our results suggest that future theoretical models should consider and confirm whether a certain degree of heterogeneity and spatial dispersion is inherently required to recover the scaling effects in terms of increasing returns and economies of scale. It is also necessary to investigate whether the relationship between GDP microstructure and scaling exponents β also emerges in other countries as well to disentangle this relationship from country-specific geographical effects. Fortunately, our usage of open-source and globally harmonized definition of cities and gridded GDP datasets ensures the reproducibility of our analysis throughout the globe.

Methods

Data sources

Our city definition is based on the recently released GHSL-FUA dataset [33], consisting of functional urban areas (FUAs) across the globe with a harmonized definition for the year 2015. Throughout this paper, we denote FUAs as cities. Each city consists of urban center and their surrounding commuting zone. The urban center are determined based on the settlement model by GHSL-SMOD [32], and is defined as contiguous grid cells (1 km² resolution) with minimum total population of 5000, where each cell must satisfy: (i) minimum 50% built area or (ii) more than 1500 inhabitants. Then, the commuting zone is approximated based on the global friction matrix [52]. The GHSL-FUA dataset contains the population counts, boundary polygons, total areas, as well as the urban center areas of 238 cities in Indonesia. In addition, the commuting zone area is obtained by subtracting the area of the urban center from the total area of the city. We use the gridded global GDP datasets with 30 arc-sec resolution (1 km² at the equator) for the year 2015 [36]. The GDP dataset is constructed based on the national and subnational GDP data and the gridded population data. Each grid in the dataset contains the GDP using purchasing power parity rates (GDP-PPP) in 2011 international US Dollars. To obtain the GDP data of each city, we overlay the gridded GDP data with city boundary polygons and sum the values at every cell. Territorial boundaries in Fig. 2 are constructed from GADM database of Global Administrative Areas [53]

Log-normal model

We estimate the scaling exponent β using the method in Ref. [37]. Equation (1) can be rewritten as

$$\mathbb{E}(Y|x) = Y_0 x^\beta, \quad (2)$$

with fluctuations around the model $\mathbb{V}(Y|x)$ defined as

$$\mathbb{V}(Y|x) = \mathbb{E}(Y^2|x) - \mathbb{E}(Y|x)^2 \quad (3)$$

$\mathbb{E}(Y|x)$ is the expectation of Y with respect to x . Here, we are interested in the case where $\mathbb{V}(Y|x)$ also scales with population size such as

$$\mathbb{V}(Y|x) = \gamma \mathbb{E}(Y|x)^\delta, \quad (4)$$

with γ is constant prefactor and $\delta \in [1, 2]$. The best fit then can be obtained by minimizing the negative of log-likelihood $\ln \mathcal{L}$ with log-normal distribution

$$\ln \mathcal{L} = \sum_{i=1}^N -\ln(\sigma(x_i)\sqrt{2\pi}) - \ln Y_i - \frac{(\ln Y_i - \mu(x_i))^2}{2\sigma^2(x_i)}. \quad (5)$$

$\mu(x)$ and $\sigma^2(x)$ are the mean and variance of the log-normal distribution and can be calculated from scaling parameters [37]

$$\mu(x) = \ln Y_0 + \beta \ln x - \frac{1}{2}\sigma^2(x), \quad (6)$$

$$\sigma^2(x) = \ln[1 + \gamma(Y_0 x^\beta)^{\delta-2}]. \quad (7)$$

The minimization is done using L-BFGS-B algorithm [54] in SciPy [55] with random initial parameters for 20 realizations. Finally, the errors are computed using bootstrap. All of these procedures are done using the publicly available code [56].

LouBar threshold, Normalized Gini index, and spreading index

The LouBar threshold GDP^* is determined using non-parametric method as follows. For any quantity Q with corresponding Lorenz curve $L(F)$, Q^* is the value of Q that corresponds to F^* , where F^* is the intersection of the x -axis ($L = 0$) with the tangent of the Lorenz curve at $F = 1$. The LouBar threshold gives us a more restrictive criterion of a hotspot than the typical naive average value $\bar{Q}$ as a threshold since it also depends on the dispersion of Q (Q^* increase as the maximum value of Q increases). Thus, a grid is classified as a hotspot if the GDP in the grid satisfies $GDP > GDP^*$. After such hotspots are determined, we sum their GDP to get the total GDP of hotspots. Then, as in previous studies [24, 26], we also obtain the number of hotspot grids in each city, which is treated as the total area of hotspots in a city considering that one grid has the dimension of approximately 1 km².

The heterogeneity of the GDP distribution within cities with total number of cells N_g is measured through the normalized Gini index G_{norm} , defined as $G_{\text{norm}} = G/G_{\text{max}}$. G is the classical Gini index and G_{max} is the maximum Gini index of each city, defined as $G_{\text{max}} = (N_g - 1)/N_g$ i.e the maximum value of G when all of the GDP are concentrated in one cell [29]. The spatial organization of hotspots can be captured by calculating the spreading index η , which is the ratio between the average distance of those hotspots and the average distance of all cells in the city,

$$\eta = \frac{\frac{1}{N_{g,GDP^*}} \sum_{i,j} d(i,j) \Theta(GDP_i - GDP^*) \Theta(GDP_j - GDP^*)}{\frac{1}{N_g} \sum_{i,j} d(i,j)}, \quad (8)$$

where $\Theta(x)$ is the Heaviside step function. For cells i and j , GDP_i and GDP_j represent the GDP value on cells i and j , respectively. N_{g,GDP^*} is the number of cells with GDP larger than GDP^* , and $d(i,j)$ is the distance between cells i and j .

References

1. Batty M. The New Science of Cities. The MIT Press; 2013.
2. Bettencourt LMA, Lobo J, Helbing D, Kuhnert C, West GB. Growth, innovation, scaling, and the pace of life in cities. Proceedings of the National Academy of Sciences. 2007;104(17):7301–7306. doi:10.1073/pnas.0610172104.
3. Bettencourt L, West G. A unified theory of urban living. Nature. 2010;467(7318):912–913. doi:10.1038/467912a.
4. Bettencourt LMA, Lobo J. Urban scaling in Europe. J R Soc Interface. 2016;13(116):20160005. doi:10.1098/rsif.2016.0005.
5. Sahasranaman A, Bettencourt LMA. Urban geography and scaling of contemporary Indian cities. J R Soc Interface. 2019;16(152):20180758. doi:10.1098/rsif.2018.0758.
6. Meirelles J, Neto CR, Ferreira FF, Ribeiro FL, Binder CR. Evolution of urban scaling: Evidence from Brazil. PLoS ONE. 2018;13(10):e0204574. doi:10.1371/journal.pone.0204574.
7. Rybski D, Arcaute E, Batty M. Urban scaling laws. Environment and Planning B: Urban Analytics and City Science. 2019;46(9):1605–1610. doi:10.1177/2399808319886125.

8. Sahasranaman A, Bettencourt LM. Economic geography and the scaling of urban and regional income in India. *Environment and Planning B: Urban Analytics and City Science*. 2019;48(3):540–554. doi:10.1177/2399808319879463. 329–331
9. Louf R, Barthelemy M. Scaling: Lost in the Smog. *Environ Plann B Plann Des*. 2014;41(5):767–769. doi:10.1068/b4105c. 332–333
10. Arcaute E, Hatna E, Ferguson P, Youn H, Johansson A, Batty M. Constructing cities, deconstructing scaling laws. *J R Soc Interface*. 2015;12(102):20140745. doi:10.1098/rsif.2014.0745. 334–336
11. Bettencourt LMA. The Origins of Scaling in Cities. *Science*. 2013;340(6139):1438–1441. doi:10.1126/science.1235823. 337–338
12. Yang VC, Papachristos AV, Abrams DM. Modeling the origin of urban-output scaling laws. *Phys Rev E*. 2019;100(3). doi:10.1103/physreve.100.032306. 339–340
13. Molinero C, Thurner S. How the geometry of cities determines urban scaling laws. *J R Soc Interface*. 2021;18(176). doi:10.1098/rsif.2020.0705. 341–342
14. Cabrera-Arnau C, Bishop SR. The effect of dragon-kings on the estimation of scaling law parameters. *Sci Rep*. 2020;10(1). doi:10.1038/s41598-020-77232-6. 343–344
15. Mori T, Smith TE, Hsu WT. Common power laws for cities and spatial fractal structures. *Proc Natl Acad Sci USA*. 2020;117(12):6469–6475. doi:10.1073/pnas.1913014117. 345–347
16. Alves LGA, Rybski D, Ribeiro HV. Commuting network effect on urban wealth scaling. *Sci Rep*. 2021;11(1). doi:10.1038/s41598-021-02327-7. 348–349
17. Sarkar S, Phibbs P, Simpson R, Wasnik S. The scaling of income distribution in Australia: Possible relationships between urban allometry, city size, and economic inequality. *Environment and Planning B: Urban Analytics and City Science*. 2016;45(4):603–622. doi:10.1177/0265813516676488. 350–353
18. Sarkar S. Urban scaling and the geographic concentration of inequalities by city size. *Environment and Planning B: Urban Analytics and City Science*. 2018;46(9):1627–1644. doi:10.1177/2399808318766070. 354–356
19. Mora EH, Heine C, Jackson JJ, West GB, Yang VC, Kempes CP. Scaling of urban income inequality in the USA. *J R Soc Interface*. 2021;18(181):20210223. doi:10.1098/rsif.2021.0223. 357–359
20. Ramaswami A, Jiang D, Tong K, Zhao J. Impact of the Economic Structure of Cities on Urban Scaling Factors: Implications for Urban Material and Energy Flows in China. *Journal of Industrial Ecology*. 2017;22(2):392–405. doi:10.1111/jiec.12563. 360–363
21. Sarkar S, Arcaute E, Hatna E, Alizadeh T, Searle G, Batty M. Evidence for localization and urbanization economies in urban scaling. *R Soc open sci*. 2020;7(3):191638. doi:10.1098/rsos.191638. 364–366
22. Meijers EJ, Burger MJ. Spatial Structure and Productivity in US Metropolitan Areas. *Environ Plan A*. 2010;42(6):1383–1402. doi:10.1068/a42151. 367–368
23. Louf R, Barthelemy M. Modeling the Polycentric Transition of Cities. *Phys Rev Lett*. 2013;111(19). doi:10.1103/physrevlett.111.198702. 369–370

24. Louail T, Lenormand M, Cantu Ros OG, Picornell M, Herranz R, Frias-Martinez E, et al. From mobile phone data to the spatial structure of cities. *Sci Rep.* 2014;4(1). doi:10.1038/srep05276. 371
372
373
25. Louf R, Barthelemy M. How congestion shapes cities: from mobility patterns to scaling. *Sci Rep.* 2014;4(1). doi:10.1038/srep05561. 374
375
26. Xu W, Chen H, Frias-Martinez E, Cebrian M, Li X. The inverted U-shaped effect of urban hotspots spatial compactness on urban economic growth. *R Soc open sci.* 2019;6(11):181640. doi:10.1098/rsos.181640. 376
377
378
27. Dong L, Huang Z, Zhang J, Liu Y. Understanding the mesoscopic scaling patterns within cities. *Sci Rep.* 2020;10(1). doi:10.1038/s41598-020-78135-2. 379
380
28. Tsai YH. Quantifying Urban Form: Compactness versus 'Sprawl'. *Urban Studies.* 2005;42(1):141–161. doi:10.1080/0042098042000309748. 381
382
29. Volpati V, Barthelemy M. The spatial organization of the population density in cities; 2018. arXiv:1804.00855. 383
384
30. Cohen B. Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society.* 2006;28(1):63–80. doi:https://doi.org/10.1016/j.techsoc.2005.10.005. 385
386
387
31. UNU-WIDER. World Income Inequality Database (WIID); 2018. Available from: <https://www.wider.unu.edu/database/world-income-inequality-database-wiid4>. 388
389
390
32. Florczyk A, Ehrlich D, , Pesaresi M, Maffenini L, Schiavina M, et al. GHSL data package 2019. Publications Office of the European Union; 2019. Available from: <http://dx.doi.org/10.2760/729240>. 391
392
393
33. Moreno-Monroy AI, Schiavina M, Veneri P. Metropolitan areas in the world. Delineation and population trends. *Journal of Urban Economics.* 2021;125:103242. doi:10.1016/j.jue.2020.103242. 394
395
396
34. Ribeiro HV, Oehlers M, Moreno-Monroy AI, Kropp JP, Rybski D. Association between population distribution and urban GDP scaling. *PLoS ONE.* 2021;16(1):e0245771. doi:10.1371/journal.pone.0245771. 397
398
399
35. United Nations. United Nations World Urbanization Prospects: 2018 revision database; 2018. Available from: <https://population.un.org/wup/>. 400
401
36. Kumm M, Taka M, Guillaume JHA. Gridded global datasets for Gross Domestic Product and Human Development Index over 1990–2015. *Sci Data.* 2018;5(1). doi:10.1038/sdata.2018.4. 402
403
404
37. Leitão JC, Miotto JM, Gerlach M, Altmann EG. Is this scaling nonlinear? *R Soc open sci.* 2016;3(7):150649. doi:10.1098/rsos.150649. 405
406
38. Brelsford C, Lobo J, Hand J, Bettencourt LMA. Heterogeneity and scale of sustainable development in cities. *Proc Natl Acad Sci USA.* 2017;114(34):8963–8968. doi:10.1073/pnas.1606033114. 407
408
409
39. Dove MR, Kammen DM. Vernacular Models of Development: An Analysis of Indonesia Under the “New Order”. *World Development.* 2001;29(4):619–639. doi:10.1016/s0305-750x(01)00002-x. 410
411
412

40. Wajdi N, van Wissen LJG, Mulder CH. Interregional Migration Flows in Indonesia. *Sojourn: Journal of Social Issues in Southeast Asia*. 2015;30(2):371–422.
41. Tadjoeddin MZ. Inequality and Exclusion in Indonesia: Political Economic Developments in the Post-Soeharto Era. *Journal of Southeast Asian Economies*. 2019;36(3):284–303.
42. McGee TG. The emergence of desakota regions in Asia: expanding a hypothesis. In: Ginsburge NS, Koppel B, McGee TG, editors. *The Extended Metropolis: Settlement Transition in Asia*. Honolulu, HI: University of Hawaii Press; 1991. p. 3–25.
43. Cottineau C, Finance O, Hatna E, Arcaute E, Batty M. Defining urban clusters to detect agglomeration economies. *Environment and Planning B: Urban Analytics and City Science*. 2018;46(9):1611–1626. doi:10.1177/2399808318755146.
44. Gehlke CE, Biehl K. Certain Effects of Grouping upon the Size of the Correlation Coefficient in Census Tract Material. *Journal of the American Statistical Association*. 1934;29(185A):169–170. doi:10.1080/01621459.1934.10506247.
45. Briant A, Combes PP, Lafourcade M. Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics*. 2010;67(3):287–302. doi:10.1016/j.jue.2009.09.014.
46. Geertz C. *Agricultural Involution: The Processes of Ecological Change in Indonesia*. Association of Asian Studies. Monographs and papers. University of California Press; 1963.
47. Geertz C. *Peddlers and Princes: Social Change and Economic Modernization in Two Indonesian Towns*. Anthropology/Economics. University of Chicago Press; 1963.
48. McGee T. From Urban Involution to Proletarian Transformation. *cgq*. 2005;29(77):295–308. doi:10.7202/021725ar.
49. Davis M. *Planet of Slums*. Verso; 2006.
50. Brelsford C, Martin T, Hand J, Bettencourt LMA. Toward cities without slums: Topology and the spatial evolution of neighborhoods. *Sci Adv*. 2018;4(8). doi:10.1126/sciadv.aar4644.
51. Soman S, Beukes A, Nederhood C, Marchio N, Bettencourt L. Worldwide Detection of Informal Settlements via Topological Analysis of Crowdsourced Digital Maps. *IJGI*. 2020;9(11):685. doi:10.3390/ijgi9110685.
52. Weiss DJ, Nelson A, Gibson HS, Temperley W, Peedell S, Lieber A, et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*. 2018;553(7688):333–336. doi:10.1038/nature25181.
53. GADM. Database of Global Administrative Areas v3.6; 2018. Available from: <https://gadm.org/>.
54. Byrd RH, Lu P, Nocedal J, Zhu C. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J Sci Comput*. 1995;16(5):1190–1208. doi:10.1137/0916069.

55. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. 2020;17:261–272. doi:10.1038/s41592-019-0686-2.

56. Leitão JC, Miotto JM, Gerlach M, Altmann EG. Python code to analyse scaling laws; 2016. Available from: <https://doi.org/10.5281/zenodo.49367>.

Acknowledgements

G. K is supported by Grant-in-Aid for JSPS Fellows Project/Area Number 20J23402.