

Jiang, Yi

Working Paper

Spatial dynamics and driving forces of Asian cities

ADB Economics Working Paper Series, No. 618

Provided in Cooperation with:

Asian Development Bank (ADB), Manila

Suggested Citation: Jiang, Yi (2020) : Spatial dynamics and driving forces of Asian cities, ADB Economics Working Paper Series, No. 618, Asian Development Bank (ADB), Manila, <https://doi.org/10.22617/WPS200220-2>

This Version is available at:

<https://hdl.handle.net/10419/246695>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/3.0/igo/>

SPATIAL DYNAMICS AND DRIVING FORCES OF ASIAN CITIES

Yi Jiang

NO. 618

August 2020

**ADB ECONOMICS
WORKING PAPER SERIES**

ADB Economics Working Paper Series

Spatial Dynamics and Driving Forces of Asian Cities

Yi Jiang

No. 618 | August 2020

Yi Jiang (yijiang@adb.org) is a principal economist in the Economic Research and Regional Cooperation Department of the Asian Development Bank (ADB).

This paper was prepared as background material for the *Asian Development Outlook 2019 Update* theme chapter on “Fostering Growth and Inclusion in Asia’s Cities.” The paper benefited greatly from the guidance and comments of Gilles Duranton, Rana Hasan, and Yasuyuki Sawada. The author thanks Marjorie Remolador for outstanding GIS work and Renz Adrian Calub for excellent research assistance. Seminar participants at ADB also provided very useful suggestions.



Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO)

© 2020 Asian Development Bank
6 ADB Avenue, Mandaluyong City, 1550 Metro Manila, Philippines
Tel +63 2 8632 4444; Fax +63 2 8636 2444
www.adb.org

Some rights reserved. Published in 2020.

ISSN 2313-6537 (print), 2313-6545 (electronic)
Publication Stock No. WPS200220-2
DOI: <http://dx.doi.org/10.22617/WPS200220-2>

The views expressed in this publication are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank (ADB) or its Board of Governors or the governments they represent.

ADB does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use. The mention of specific companies or products of manufacturers does not imply that they are endorsed or recommended by ADB in preference to others of a similar nature that are not mentioned.

By making any designation of or reference to a particular territory or geographic area, or by using the term “country” in this document, ADB does not intend to make any judgments as to the legal or other status of any territory or area.

This work is available under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO) <https://creativecommons.org/licenses/by/3.0/igo/>. By using the content of this publication, you agree to be bound by the terms of this license. For attribution, translations, adaptations, and permissions, please read the provisions and terms of use at <https://www.adb.org/terms-use#openaccess>.

This CC license does not apply to non-ADB copyright materials in this publication. If the material is attributed to another source, please contact the copyright owner or publisher of that source for permission to reproduce it. ADB cannot be held liable for any claims that arise as a result of your use of the material.

Please contact pubsmarketing@adb.org if you have questions or comments with respect to content, or if you wish to obtain copyright permission for your intended use that does not fall within these terms, or for permission to use the ADB logo.

Corrigenda to ADB publications may be found at <http://www.adb.org/publications/corrigenda>.

Notes:

In this publication, “\$” refers to United States dollars.
ADB recognizes “China” as the People’s Republic of China.

The ADB Economics Working Paper Series presents data, information, and/or findings from ongoing research and studies to encourage exchange of ideas and to elicit comment and feedback about development issues in Asia and the Pacific. Since papers in this series are intended for quick and easy dissemination, the content may or may not be fully edited and may later be modified for final publication.

CONTENTS

TABLES AND FIGURES	iv
ABSTRACT	v
I. INTRODUCTION	1
II. DATA	2
A. Delineating the Physical Area of Human Settlements	2
B. Identifying Natural Cities	4
C. Measuring the Populations of Natural Cities	9
D. Other Data Used	10
III. URBANIZATION IN ASIA AND THE PACIFIC	10
A. Urbanization Rates from 1992 to 2016	10
B. Urbanization and Economic Growth	13
IV. THE URBAN SYSTEM	17
A. Distribution of Population across City Sizes	18
B. Primate Cities	24
C. Zipf's Law in Selected Countries	28
V. GROWTH OF CITIES AND EMERGENCE OF CITY CLUSTERS	32
A. Simple Stylized Facts about City Dynamics	32
B. Testing Gibrat's Law	34
C. Factors Driving City Growth	36
D. Emergence of City Clusters	38
VI. CONCLUSION	42
APPENDIXES	45
1 Developing the Natural City Dataset with Nighttime Lights LandScan Data	45
2 Comparison of Estimates of Urbanization Rates	49
REFERENCES	51

TABLES AND FIGURES

TABLES

1	Correlates of Urbanization Rates and Progress	15
2	Counts of Natural Cities by Size	23
3	Primate Cities in Asia and the Pacific, 2016	25
4	Correlates of Primate Cities	27
5	Zipf's Law Regressions	30
6	Number of Cities by Changes in Size, 2000–2016	33
7	Changes in City Size Quintile, 2000–2016	33
8	Testing Gibrat's Law	35
9	Driving Factors for City Growth, 2000–2016	37
10	Natural Cities Merged to become City Clusters, 2016	39
11	Largest City Clusters with Population above 10 Million, 2016	41
A1	Defense Meteorological Satellite Program–Operational Linescan System Satellites and Their Coverage Period	45

FIGURES

1	Raw Nighttime Lights Image of Asia and the Pacific, 1992 and 2016	3
2	Natural Cities Extracted from Nighttime Lights of Metro Manila and Surrounding Areas, Philippines, 1992 and 2016	6
3	Number of Natural Cities by Country	7
4	Spatial Development of Selected Natural Cities	8
5	Shares of Urban Area and Urban Population across 43 Economies Studied	11
6	Shares of Urban Area and Urban Population across 38 Economies Studied	12
7	Correlation of Urbanization Rate and Gross Domestic Product per Capita, 2000 and 2016	14
8	Distribution of City Area, Population, and Density across 43 Economies Studied	19
9	Distribution of City Area, Population, and Density across 38 Economies Studied	20
10	Distribution of Population across Different Sizes of Cities	24
11	Zipf's Law Testing by Country	31
12	City Clusters with Populations of 10 Million or More, 2016	40
A2.1	Estimates of Urbanization Rates based on Natural Cities versus World Urbanization Prospects Estimates, 2000	49
A2.2	Estimates of Urbanization Rates based on Natural Cities versus World Urbanization Prospects Estimates, 2016	50

ABSTRACT

This paper introduces a new city-level panel dataset constructed using satellite nighttime light imagery and grid population data. The dataset contains over 1,500 cities covering 43 economies of Asia and the Pacific from 1992 to 2016. With the dataset, we perform a variety of analyses for Asia and the Pacific as a whole as well as five individual countries in the region. The exercise produces some novel evidence on several interrelated topics including urbanization status and patterns, relations between urbanization and economic growth, evolution of urban systems, primate cities, testing Zipf's law and Gibrat's law, the drivers of city growth, and emergence of city clusters.

Keywords: city cluster, city growth, Gibrat's law, nighttime lights, primate city, Zipf's law

JEL codes: R11, R12, O18

I. INTRODUCTION

Since the 1980s, Asia and the Pacific has grown to become the most economically dynamic region in the world. Currently producing about one-third of global gross domestic product (GDP), the region is expected to account for over half of the world's output by 2050 (Kohli, Sharma, and Sood 2011). It is believed that this rapid economic growth is closely associated with urbanization across the region. Not only have we seen established cities such as Mumbai and Shanghai join Hong Kong, China; Singapore; and Tokyo in playing a key role in global markets, we have also observed relatively younger cities such as Bangalore and Shenzhen grow into the technological epicenters of Asia.

Of course, economic development does not only occur in these major cities. It is the system of cities in a country that facilitates structural transformation, catalyzes productivity improvements and stimulates technological innovations. Cities in different locations, of various sizes, and with different industrial compositions, interact with each other with different specializations or functions. A comprehensive look at all of them could help us better understand urbanization processes and trends across Asia and the Pacific, and thus gain some insights into economic development within individual countries.

To the best of our knowledge, such a comprehensive examination is not yet available in Asia and the Pacific, largely due to a lack of city-level data, which is comparable across countries and time. According to the United Nations (UN) Department of Economic and Social Affairs, Population Division (2018), four types of criteria are typically used in official definitions of urban areas: administrative boundaries, economic parameters, population size and/or density, and urban characteristics. Due to various combinations of these criteria, there are 13 known ways to define urban areas among the 233 economies in the world.¹ The number of actual definitions of cities is much greater than 13 as differing numeric thresholds are applied to these criteria. As a result, when one sees that the urbanization rate is 40% in country A and 30% in country B, it is actually hard to conclude that A has higher a urbanization level than B. We hope to address this gap by constructing and analyzing a large-scale city dataset across Asia and the Pacific.

Satellite imagery that has captured the nighttime lights of human settlements since 1992 was introduced into economic research by Chen and Nordhaus (2011) and Henderson, Storeygard, and Weil (2012). The authors demonstrated that photographic data could be used to study subnational economic activity as well as a proxy for economic growth, especially in low- or middle-income countries, where reliable data is scarce. Thereafter, the data have been creatively applied to study important economic issues (e.g., national institutions and subnational development in Michalopoulos and Papaioannou 2014), although there is some debate as to whether or not nighttime lights are an appropriate indicator of economic activity, especially in large urban areas (Mellander et al. 2015).

Instead of relying on nighttime lights to measure the economic activity of cities in Asia and the Pacific, we use the imagery to delineate the extent of urban agglomerations, which we call “natural cities” in order to distinguish them from officially defined cities. We then measure the population of the natural cities by filling in each area with grid population data from LandScan.² Using this methodology, we created a panel dataset that contains more than 1,500 natural cities in 43 economies in Asia and the

¹ For instance, 59 economies use administrative designations as the sole criterion for defining a city, while another 62 combine the administrative criterion with others to distinguish between urban and rural areas.

² The LandScan data can be accessed at <https://landscan.ornl.gov>.

Pacific between 1992 and 2016.³ The geographic scale and population size are consistently defined across space and time for these cities.

We perform a variety of analyses at different geographic levels using the dataset. First, we report the urbanization rates from 1992 to 2016 for the region as a whole, as well as for each of five countries, which have the largest numbers of natural cities in the dataset, that is, the People's Republic of China (PRC), India, Indonesia, Japan, and Pakistan. We also examine how urbanization rates and progress are correlated with some key country characteristics. This information is presented in section III. In section IV, we focus on urban systems, concerning ourselves with how urban populations are distributed across cities of different sizes, the status and evolution of primate cities, and whether the classic city size-rank rule holds in Asian countries. In section V, our investigations focus on individual natural cities. Here, we are interested in changes in the absolute and relative city sizes in terms of population and area, how the growth of a city is correlated with its initial size, and what characteristics affect city growth. We also present some stylized facts about city clusters, which emerge when two or more natural cities expand to become connected to one another.

II. DATA

The primary units studied in this paper are what we call natural cities. They are urban agglomerations that are not defined by administrative or political boundaries, but are instead identified based on satellite imagery that has captured the nighttime lights of human settlements since 1992. Below we provide a summary of the procedure used to construct the dataset of natural cities. More technical details about the data development are outlined in Appendix 1.

A. Delineating the Physical Area of Human Settlements

Satellites from the United States (US) Air Force Defense Meteorological Satellite Program (DMSP) with Operational Linescan System (OLS) sensors recorded the intensity of Earth-based lights and stored them in a digital archive from 1992 to 2013. Since 2013, the DMSP-OLS data were succeeded by those recorded by the Visible and Infrared Imager/Radiometer Suite (VIIRS) flown on spacecraft launched under the Joint Polar Satellite System, a program that consolidates the polar-orbiting spacecraft of multiple agencies of the US. Scientists at the National Oceanic and Atmospheric Administration (NOAA) process the raw data and distribute a yearly version of nighttime light (NTL) data to the public.⁴ We used this public NTL data from NOAA to delineate the physical extent of human settlements in Asia and the Pacific for 1992, 1995, 2000, 2005, 2010, and 2016.

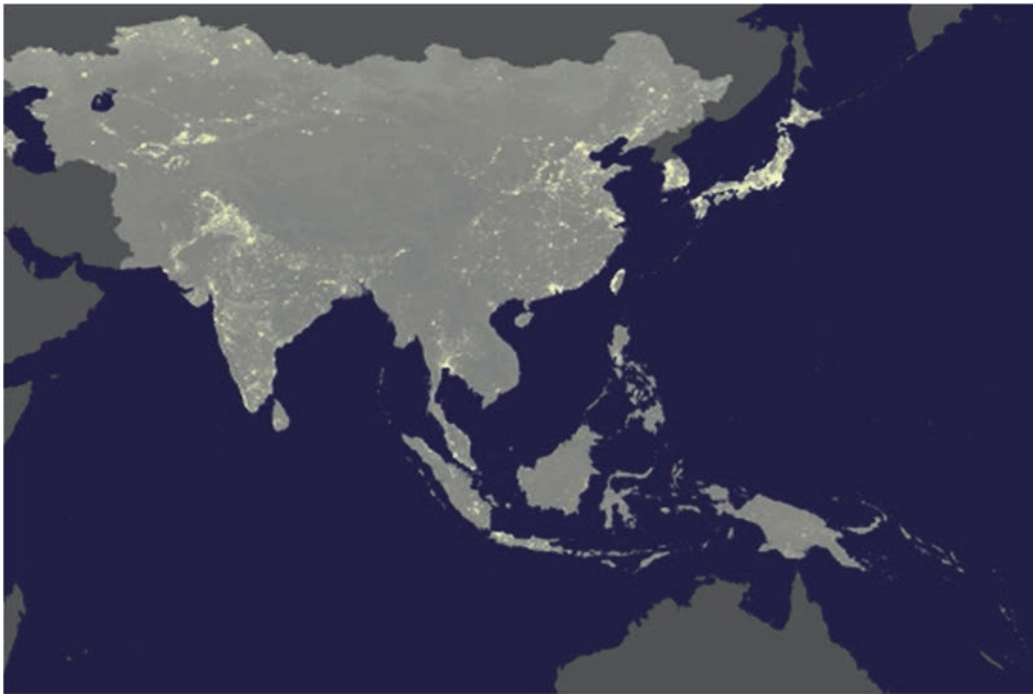
The DMSP-OLS NTL data are available in every latitude-longitude grid of 30 arc-seconds, equivalent to about 0.86 square kilometers (km²) at the equator. The luminosity of each pixel is represented by an integer between 0 and 63, with 0 indicating no light and 63 a censored value for very high luminosity. The VIIRS NTL data are at a higher resolution, with each pixel equal to about 0.22 km² at the equator. The

³ The economies include 42 developing members of the Asian Development Bank (ADB) along with Japan. Kiribati, the Marshall Islands, and Tuvalu have no reliable data identified and are therefore not included in the dataset. In the paper, we refer to these 43 economies as the Asia and Pacific region.

⁴ The data was accessed at the website of the National Geophysical Data Center of NOAA at <https://ngdc.noaa.gov/>.

Figure 1: Raw Nighttime Lights Image of Asia and the Pacific, 1992 and 2016

a. 1992



b. 2016



Source: Author's creation.

public version contains average radiance values, which have the nonlight background set to zero, but are not top-coded like the DMSP-OLS NTL data. Figures 1a and 1b show the NTLs of Asia and the Pacific in 1992 and 2016, respectively. One can see that the images are sharper and the illuminated areas are larger and brighter in 2016.

When we focus in on the region encompassing Metro Manila and surrounding areas in the Philippines (as illustrated in the left panels of Figures 2a and 2b), we see that the boundaries of the illuminated areas are blurry in 1992, while the degree of blurriness is significantly reduced in 2016. This blurriness is known as “blooming” or “overglowing.” It is caused by the relatively coarse spatial resolution of the OLS sensors, the large overlap in the footprints of adjacent OLS pixels, and the accumulation of geolocation errors in the compositing process (Small, Pozzi, and Elvidge 2005). On the other hand, the VIIRS images have much less blooming due to their higher resolution. As a necessary step, we adopted the latest methodology developed in Abrahams, Oram, and Lozano-Garcia (2018) to deblur the DMSP-OLS imagery.

With the deblurred NTL data, we delineated human settlements based on a luminosity threshold equal to zero, which means that pixels with positive luminosity values were all considered human settlements.⁵ We then aggregated those with small gaps between them (1 km for DMSP-OLS data and 0.5 km for VIIRS data) into one polygon to allow for measurement errors as well as unlit areas (such as roads) within an integrated human settlement. The exercise yielded between 88,000 and 187,000 geocoded polygons across the region in various years. The middle panels of Figures 2a and 2b show all the delineated polygons in the area around Metro Manila in 1992 and 2016, respectively. One noticeable difference between the two charts is that many small polygons in 1992 have grown into or merged to form larger polygons by 2016.

B. Identifying Natural Cities

Majority of the polygons obtained above with NTL are very small and discrete, likely representing rural settlements. To identify urban areas from them, we referred to the database of the Global Rural Urban Mapping Project (GRUMP). This database contains geocoded locations, their names, populations, and the upper administrative divisions they belong to of over 70,000 human settlements across the world.⁶ We focused on 1,964 units in Asia and the Pacific that had a population greater than 100,000 in 2000, and identified 1,412 NTL-based polygons in 1992 that either cover these GRUMP units or turn to be the most relevant ones near the units, as revealed by visual checking. These polygons were treated as natural cities and named after their corresponding GRUMP units or the unit with the largest population if a natural city contains multiple GRUMP units.⁷ The number of the natural cities was lower than the number of units from GRUMP because some GRUMP units were located closely to

⁵ It is noted that different thresholds are adopted to draw urban boundaries in the literature. Examples include 5 in Zhang and Seto (2011); 13 in Ellis and Roberts (2015); and in Zhou, Hubacek, and Roberts (2015); 33 in Tewari, Alder, and Roberts (2017); and 35 in Harari (2016). A positive threshold is needed if one uses the pre-deblurring data to delineate urban extent. However, a uniform positive threshold across years may not yield a consistent definition of urban scope given that different sensors (OLS versus VIIRS) were used over the period and the same sensor performed differently over its lifecycle. Moreover, a sensible threshold value to define urban extent, if it exists, should probably vary across countries.

⁶ GRUMP data were generated by the Columbia University Center for International Earth Science Information Network in collaboration with the International Food Policy Research Institute, the World Bank, and Centro Internacional de Agricultura Tropical, through combining census and geospatial datasets. It can be accessed at <http://sedac.ciesin.columbia.edu/data/collection/grump-v1>. It is worth noting that the units in the GRUMP database are not from the same administrative level within, as well as across, countries.

⁷ Thereafter, when we mention a city by name, we refer to the natural city encompassing it instead of the administratively defined city, unless specifically indicated.

each other and thus covered under the same polygons. Such cases arose in the relatively advanced areas of developing economies—such as the Pearl River Delta area centered around Guangzhou City in the PRC—as well as in developed economies such as the metropolitan areas surrounding Tokyo in Japan or the capital city of Taipei, China.

To maximize country coverage and include large cities that were missing in the GRUMP database, we added to the data 115 polygons that were either related to major cities in small countries (mostly in the Pacific) or had an area greater than 100 km² in 2000 despite the associated GRUMP units having a small 2000 population. We reached a final set of 1,527 natural cities across Asia and the Pacific, which should cover most urban agglomerations that were already sizable in population or area by 2000. On the other hand, areas that were small in population and/or area in 2000 may have been left out of our data.⁸

Returning to Figure 2a (right panel), our data identified nine natural cities in the region encompassing Metro Manila and surrounding areas in the Philippines in 1992: Angeles, Batangas, Lipa, Lucena, Metro Manila, Olongapo, San Pablo, San Pedro, and Tarlac. The largest was Metro Manila, which contained 29 GRUMP units such as Makati, Manila, and Quezon City, etc. The second-largest natural city was Angeles, which actually spanned two officially independent cities, Angeles and San Fernando.

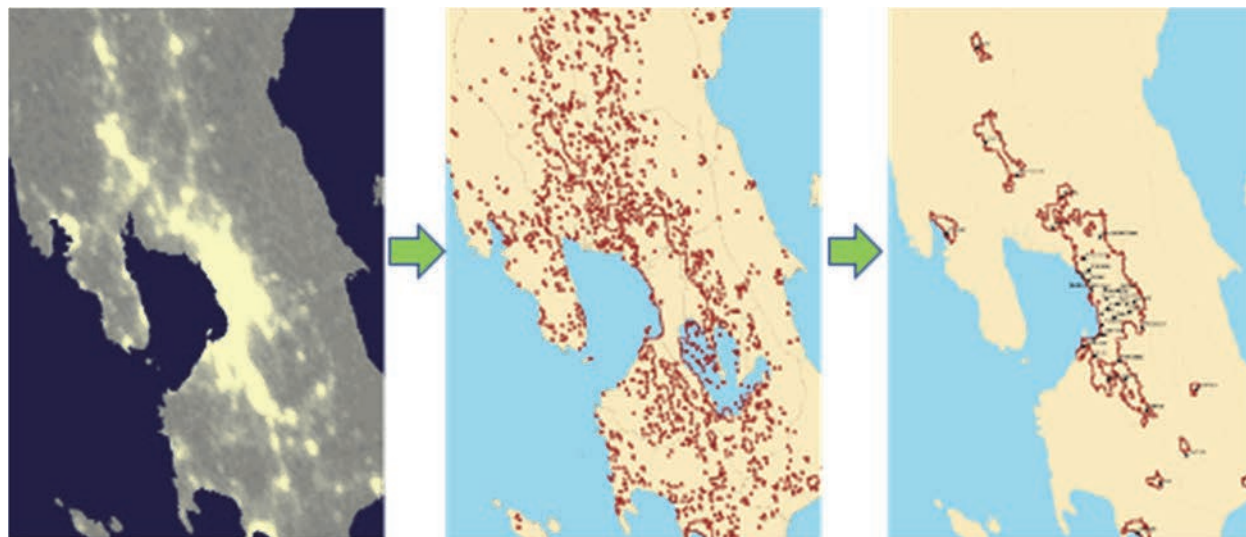
Some natural cities had expanded and became connected with other natural cities over time. This is illustrated in the right panel of Figure 2b by what had happened between Metro Manila, Angeles, Lipa, and Tarlac by 2016. These connected urban areas are considered as city clusters and discussed in more detail in subsection V.D. However, we divided these connected natural cities where the luminosity was the lowest, in order to obtain the footprint of each. By doing so, we managed to retain the same number of natural cities across years, maintaining these natural cities as our primary units for analysis.

Figure 3 shows the distribution of the 1,527 natural cities across the 43 economies covered in this study. Perhaps not surprisingly, countries with large populations were home to more natural cities. The PRC and India had the highest numbers of natural cities, 680 and 320, respectively, followed by Indonesia, Japan, and Pakistan with 92, 68, and 63 natural cities, respectively. These five countries combined made up 80% of the natural cities in Asia and the Pacific (thereafter referred to as top five countries).

⁸ Note the size criteria adopted were for the GRUMP units in 2000. A natural city's population could still be below 100,000 in 2000, or even after 2000, as the GRUMP population, often sourced from official statistics, might be for a very different urban scope.

Figure 2: Natural Cities Extracted from Nighttime Lights of Metro Manila and Surrounding Areas, Philippines, 1992 and 2016

a. 1992



b. 2016



Note: From left to right: Raw nighttime light images, human settlements with positive luminosity value, and natural cities that contain one or more GRUMP units with population greater than 100,000 in 2000.

Source: Author's creation.

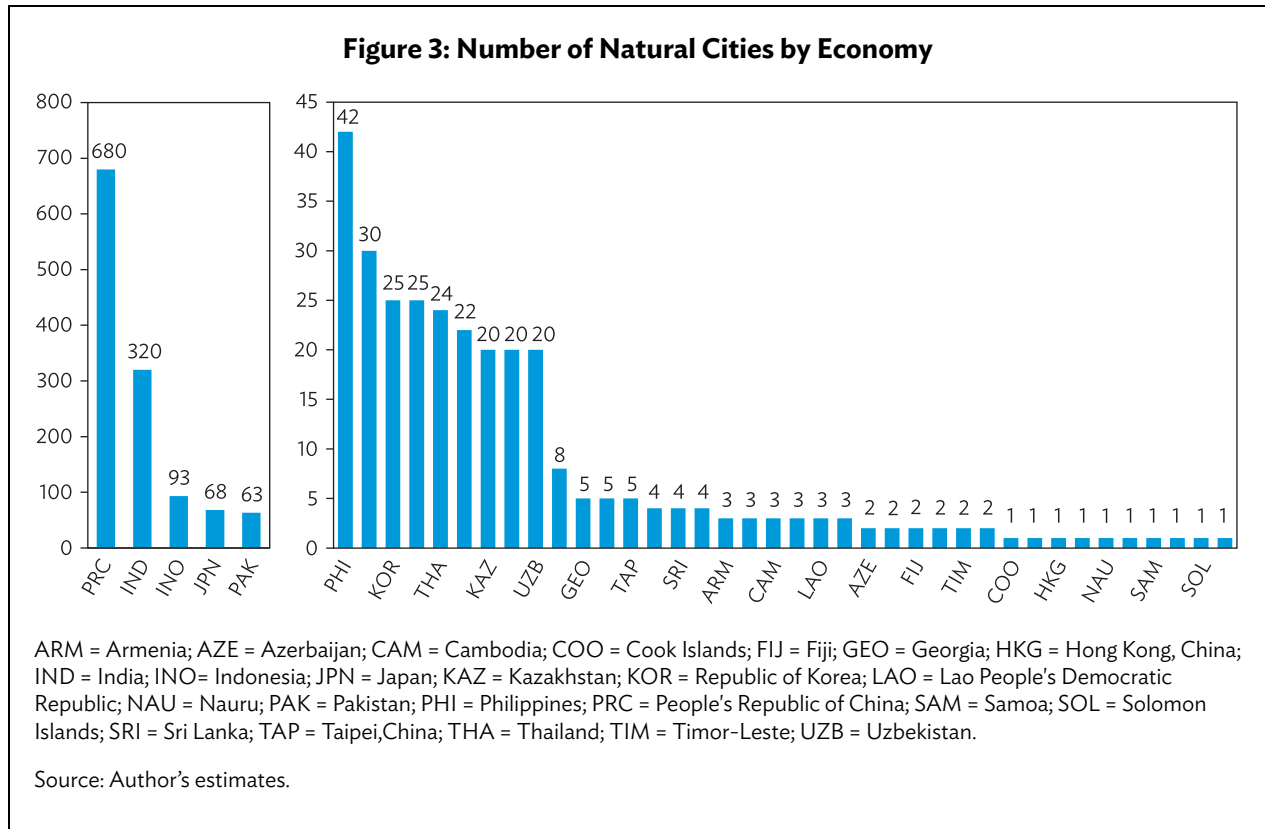
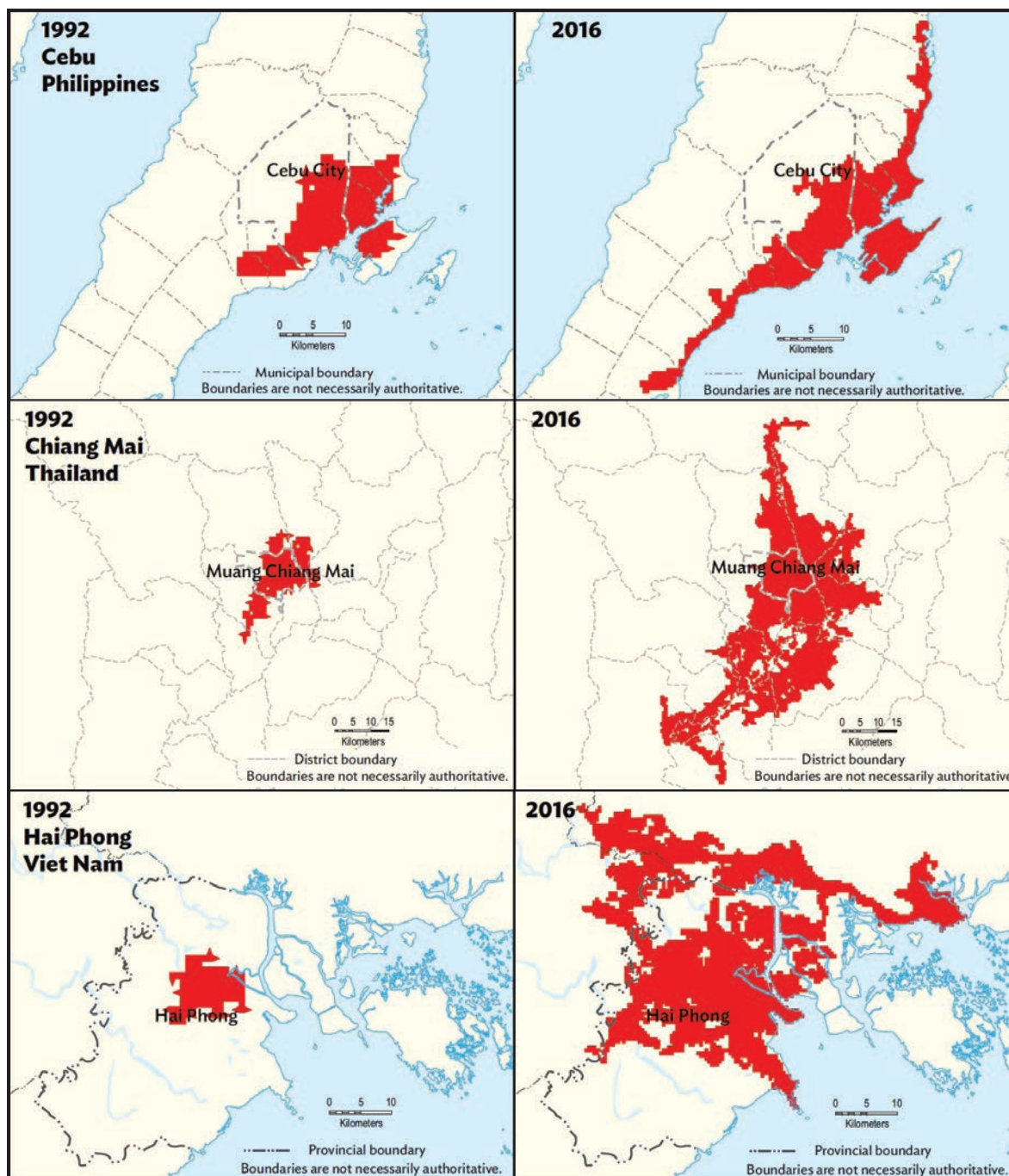
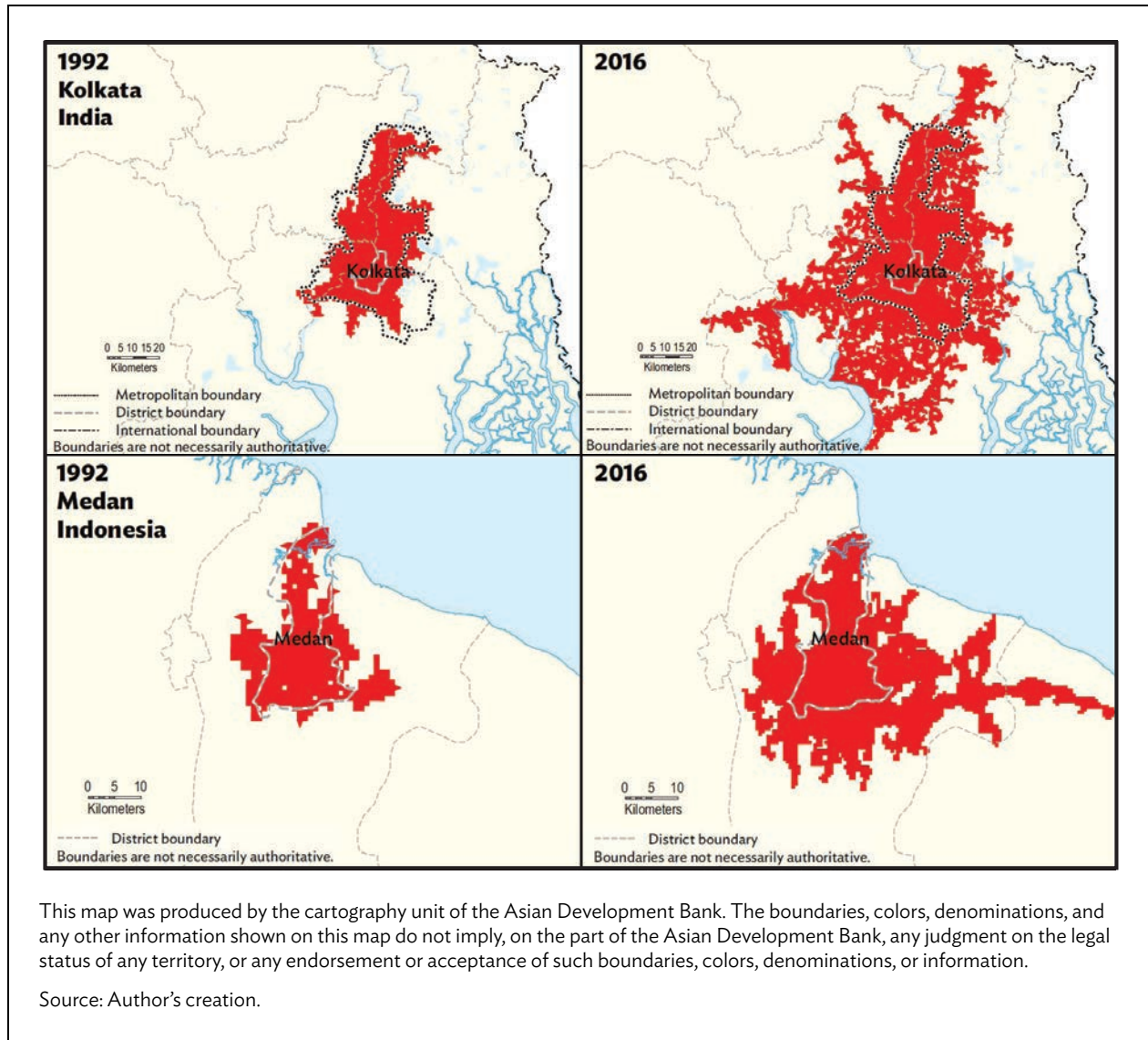


Figure 4 illustrates how the footprint of each of five selected natural cities has evolved since 1992 and how it differs from the boundary of the corresponding administrative division. First, all five natural cities grew considerably in area from 1992 to 2016: at one end of the range, Cebu growing by 58% and, at the other, Hai Phong growing by 930%. Moreover, the spatial expansion of the natural cities was by no means constrained by the administrative boundaries (dashed lines). In 2016, each of the five natural cities had a significant portion of its area lying outside of the corresponding administrative unit. Natural city sprawl reached beyond administrative borders by 73% for Cebu City, 91% for Muang Chiang Mai District, 46% for Hai Phong City, 70% for Medan City, and 64% for the Kolkata Metropolitan Area.

Figure 4: Spatial Development of Selected Natural Cities



continued on next page

Figure 4 *continued*

C. Measuring the Populations of Natural Cities

Since there is no official accounting of population for the natural cities, we filled the delineated areas of the natural cities with grid population data from LandScan. LandScan provides global population counts at a spatial resolution of approximately 1 km², which are generated through spatial modeling and image analysis with inputs from census data, high-resolution imagery, land cover, and other spatial data such as various boundaries, coastlines, elevations, and slopes. Essentially, the census population counts are disaggregated to each cell with a multivariate dasymetric modeling approach. Data precision is improved through manual verification and modification as well as refinements to the input datasets. LandScan data have been widely used in fields such as demographics, urban planning, and remote sensing.

We overlaid the natural city polygons with the grid population data. The population of a natural city is the sum of all cells falling within or intersecting with the city contour. The LandScan data start in 1998, so we have estimated populations for natural cities in all years of analysis, except 1992 and 1995.

D. Other Data Used

At the country level, total population, total surface land, GDP per capita, and sectoral outputs were obtained from the UN, World Bank World Development Indicators, ADB Statistical Database System, or the World Economic Outlook of the International Monetary Fund. At the natural city level, we determined whether or not a city has a seaport based on the World Port Index developed by the National Geospatial-Intelligence Agency in the US. For weather indicators, we referenced the United Kingdom's Climatic Research Unit, which publishes global monthly gridded weather data with a 0.5-degree variance, from 1901 onward. We used this monthly data to obtain the annual average daily precipitation and annual maximum and minimum temperatures for each grid. The averages across the grids surrounding the centroid of the natural city were taken as the weather measures for the natural city. For city clusters, we identified the level-1 administrative division (i.e., the highest subnational level such as state and province) each GRUMP settlement unit belongs to, and then calculated the number of level-1 administrative divisions each city cluster interacts with.

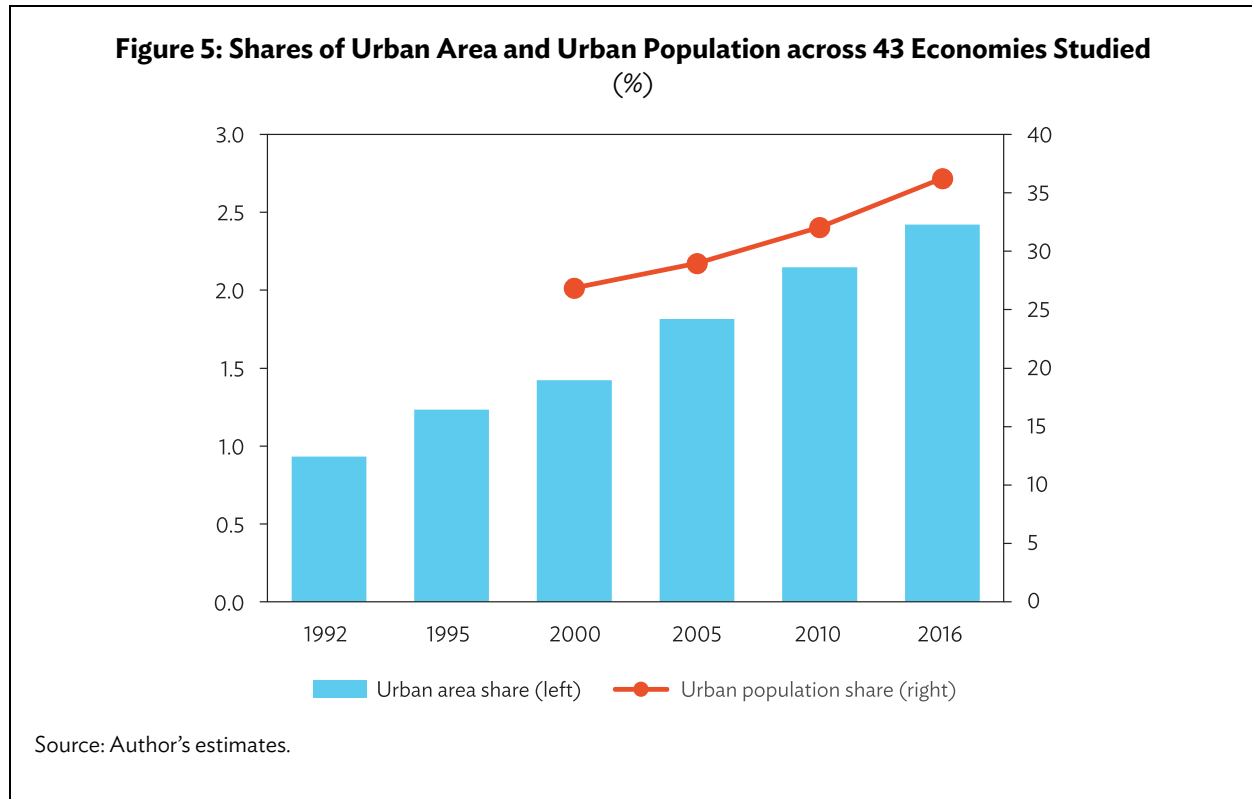
III. URBANIZATION IN ASIA AND THE PACIFIC

A. Urbanization Rates from 1992 to 2016

Our data suggest that the total urbanized area of the 43 economies studied increased from 230,000 km² in 1992 to 610,000 km² in 2016, with an average annual growth rate of 4%. The total land area of the 43 economies studied is around 25.2 million km², so the natural cities together account for 0.9% in 1992 and 2.4% in 2016 of the total land area.

The number of natural city inhabitants increased from 0.93 billion in 2000 to 1.48 billion in 2016, with an average annual growth rate of 3%. Hence, growth of urban habitation has significantly outpaced growth of the overall population, which increased at a rate of 1% per annum, from 3.5 billion in 2000 to 4.1 billion in 2016. The region's urbanization level, measured as the ratio of total natural city population to the total population, has risen steadily from 27% in 2000 to 36% in 2016 (Figure 5).

The data suggests a lower urbanization rate for the region, as compared to 47% calculated with the UN's World Urbanization Prospects (WUP) data. The causes for the gap could be multifaceted. For example, WUP data, which is largely based on official statistics, may count cities based on administrative boundaries. It is common, however, that there are both urban and rural areas within an administrative unit. On the other hand, the natural city data may leave out a number of towns and small cities given the selection criteria we applied to the GRUMP units. Appendix 2 compares the urbanization rates estimated with the two datasets at country level, and further explores possible explanations for the gaps. Essentially, the natural city data is not comparable with official statistics since the two are based on completely distinct definitions of urban areas. We prefer to view the natural city data as a new, complementary source of information on urbanization in Asia and the Pacific.



When we examine the urbanization trends of the top five economies as well as the other 38 in the region, we can see some clear differences (Figure 6).

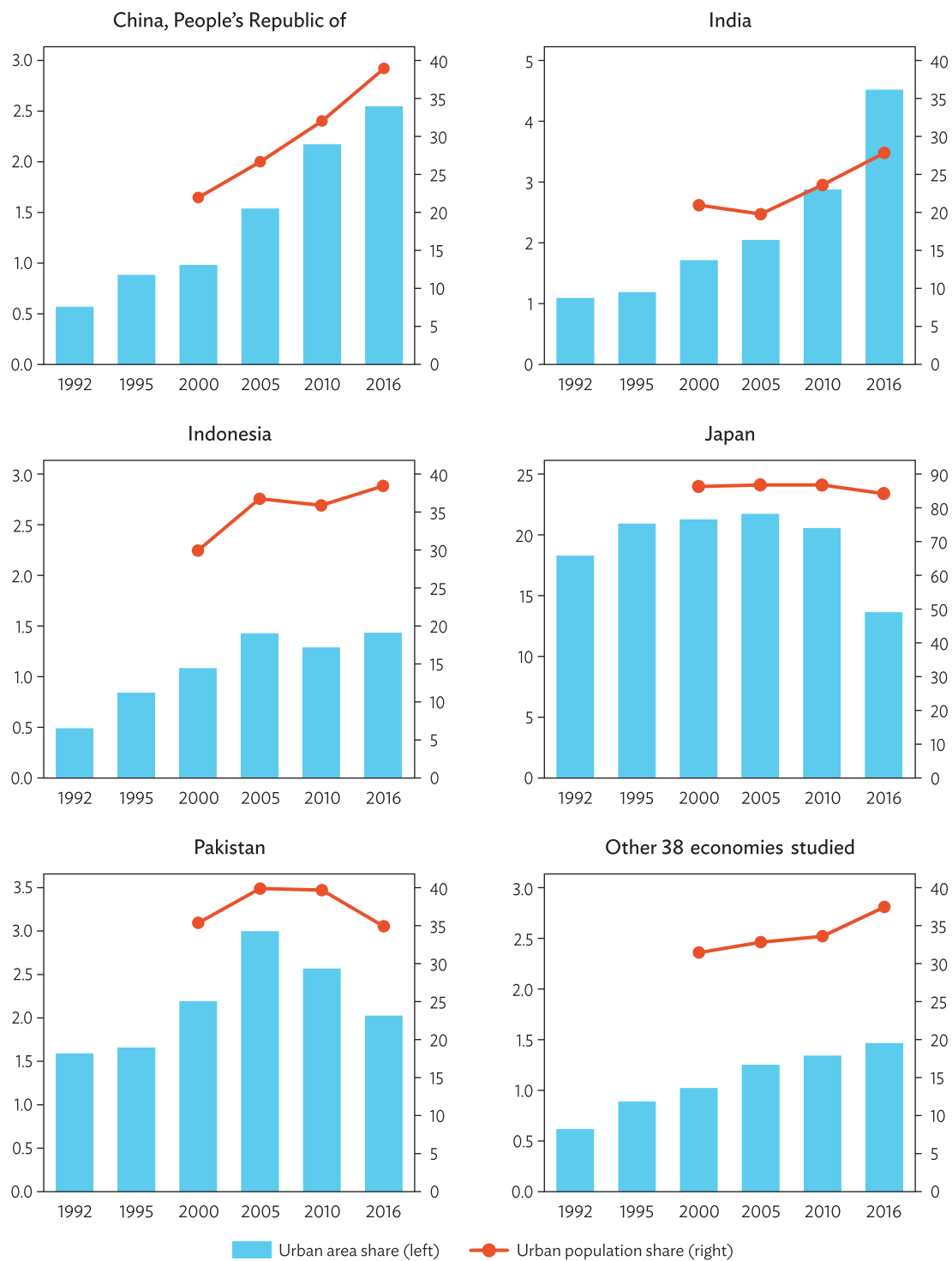
With its significant economic development, the PRC has experienced rapid urbanization since 1992. The country's urbanized area as a proportion of its total land grew from 0.6% in 1992 to 2.5% in 2016, and the share of urban population to total population increased from 22% in 2000 to 39% in 2016, with about 260 million people becoming urban residents since 2000. Both growth rates exceed those of the region as a whole.

India's urbanization has also advanced significantly. The country's proportion of urbanized land increased from 1.1% in 1992 to 4.5% in 2016. There were 368 million people living in cities by 2016, an increase of 148 million urban inhabitants since 2000. However, due to fast growth in the overall population, the proportion of urban inhabitants increased only moderately, from 21% in 2000 to 28% in 2016.

When their figures are combined, the PRC and India account for nearly 75% of the population who became urbanized from 2000 to 2016 in the region. It is therefore clear that region-wide urbanization trends are largely driven by these two countries.

Our data suggest that Indonesia experienced a steady advance in urbanization between 1992 and 2016 except a slight reverse in the late 2000s. The urban share of total land increased from 0.5% in 1992 to 1.4% in 2005, then fell to 1.3% in 2010. This measure had, however, reverted to the 2005 level by 2016. Correspondingly, the proportion of urban population to total population dropped slightly from 36.7% in 2005 to 35.9% in 2010. By 2016, this figure had reached 38.4%, 2.5 percentage points higher than the 2010 level.

Figure 6: Shares of Urban Area and Urban Population across 38 Economies Studied
(%)



Source: Author's estimates.

Japan and Pakistan exhibited de-urbanization trends from 2005 to 2016. Japan had a high urbanization rate in the 1990s, when 21% of the country's land area was urban and 86% of its population lived in cities. However, by 2016, the urban share of total land had decreased to 14% and city dwellers made up 84% of the total population. With the number of urban inhabitants remaining more or less the same, the data suggests that Japanese cities lost quite amount of peri-urban areas and actually became more densely populated. Pakistan presents a different story. In 2005, the country had relatively high urbanization rates compared to other developing Asian countries. By 2016, however, the proportion of urban land to total land had fallen from 3% to 2% and the urban share of total population had fallen 5 percentage points to 35%. Depicting the overall trend, the aggregate figures by no means imply that the cities within Pakistan contracted uniformly. Actually, there are considerable dynamics and variabilities in growth across the country's cities, which we will cover in the following sections.

Finally, the urbanization trends in the other 38 economies in the dataset have been upward in general, albeit at a slower pace than in the PRC or India. Considering these countries together, the share of urban population to total population increased from 34% in 2010 to 38% in 2016, signaling an expedited urbanization process.

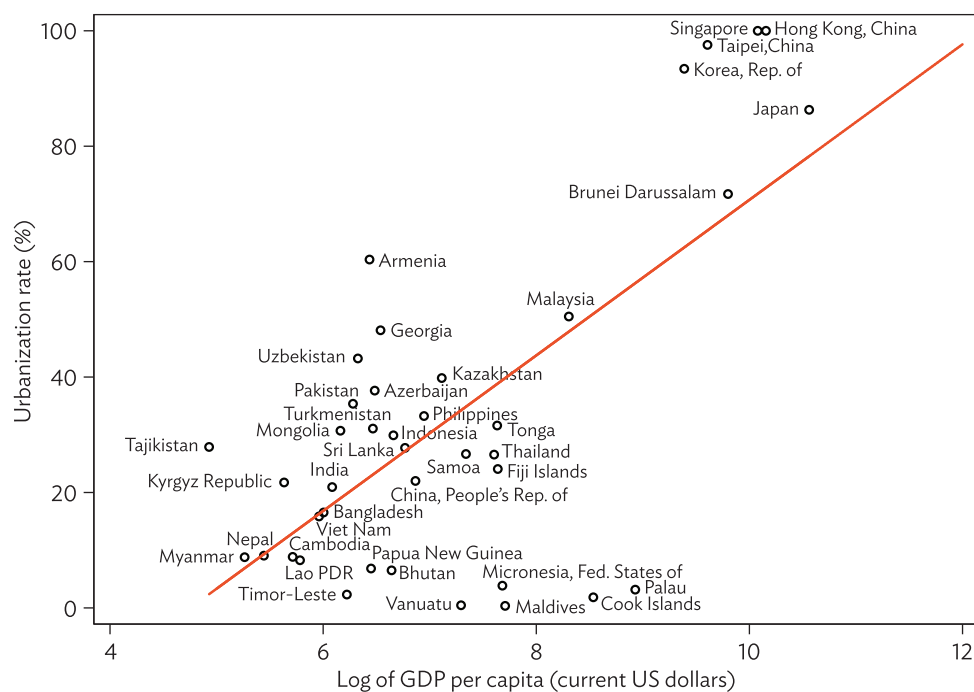
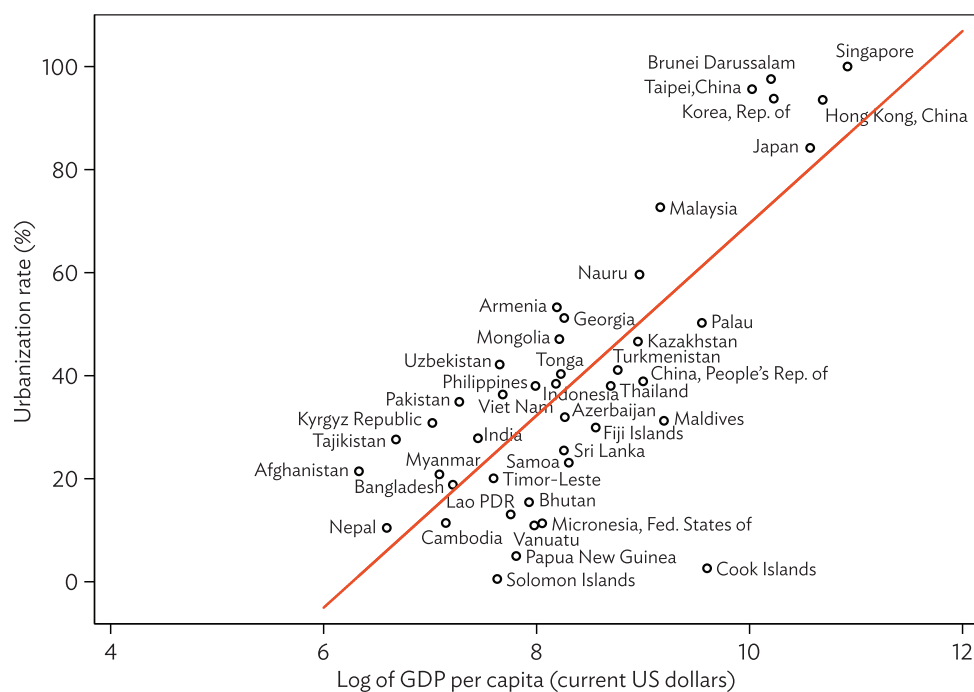
B. Urbanization and Economic Growth

One of the most widely recognized facts in economic development is that urbanization strongly correlates with income. The developed economies of East Asia, Europe, and North America all have high urbanization rates. It is also commonly observed that countries undergoing robust economic growth experience rapid urbanization at the same time. We examine here whether the relationship between urbanization and economic development holds with our NTL data, how other factors such as total population and land are correlated with urbanization, and how the relationship varies between urbanization in terms of urban habitants and urbanized land.

Figures 7a and 7b take urbanization rates as a percentage of total population that live in natural cities and plot them against the log of GDP per capita with linear fitted lines for 2000 and 2016, respectively. As expected, there exists a clear and positive relationship between the two measures. The slope of the fitted line is 0.135 in 2000 and 0.187 in 2016, both statistically significant at 1% level. These estimates imply that, on average, a 10% increase in GDP per capita is associated with an increase in urbanization of 1.4 percentage points in 2000 and 1.9 percentage points in 2016.

Two groups of countries seem to deviate in the opposite directions from the fitted line in 2000. Countries located in Central and West Asia—Armenia, Azerbaijan, Georgia, Kazakhstan, Tajikistan, and Uzbekistan—had urbanization rates higher than the level predicted by our estimates. This may be because they all belonged to the former Soviet Union, which promoted industrialization accompanied by urbanization. Meanwhile, several Pacific developing member countries (DMCs)—the Cook Islands, the Federated States of Micronesia, Maldives, Palau, Papua New Guinea, Timor-Leste, and Vanuatu—had urbanization rates that were substantially below what their GDP per capita would have suggested. These Pacific DMCs are generally dependent on natural resources and/or external remittances, which could reduce the benefits of, and necessity for, agglomeration in urban areas.

By 2016, the relationship between urbanization rates and GDP per capita had become tighter. The R-squared of the regression increased considerably from 0.45 in 2000 to 0.60 in 2016. The latter indicates that 60% of the variation in urbanization rates across the region can be

Figure 7: Correlation of Urbanization Rate and Gross Domestic Product per Capita, 2000 and 2016**a. 2000****b. 2016**

GDP = gross domestic product, US = United States.

Source: Author's estimates.

explained by the variation in GDP per capita. The two groups of countries— former Soviet Union members and Pacific DMCs (except the Cook Islands)—also moved closer to the fitted line in 2016. In short, our NTL city data supports the conventional wisdom that urbanization and economic activity go hand in hand.

Beyond the simple correlation between urbanization rates and income level, it may also be interesting to explore how other country factors such as total population, land area, and economic structure are correlated with urbanization rates, and whether GDP per capita still plays a critical role when these factors are taken into account. To answer these questions, we ran the following regression:

$$UR_c = a + b_1 \log GDP pc_c + b_2 \log pop_c + b_3 \log area_c + b_4 \log agr share_c + b_5 \log ind share_c + \gamma_s + e_c \quad (1)$$

where UR equals percentage of urbanized population, $GDP pc$, pop , and $area$ represent GDP per capita, population, and land area, respectively, $agr share$ and $ind share$ are the shares of GDP produced in agriculture and industrial sectors proxying for economic structure, γ_s represents dummies for the subregions (Central and West Asia, East Asia, Pacific, South Asia, and Southeast Asia), c denotes country, and e_c is the random error term.⁹ Equation (1) is estimated for 2000 and 2016, separately, with results shown in columns (1) and (2) of Table 1.

Table 1: Correlates of Urbanization Rates and Progress

	(1)	(2)	(3)	(4)	(5)	(6)
	Share of Urbanized Population		Log(UrPop ₂₀₁₆ /UrPop ₂₀₀₀)	Share of Urban Land		Log(UrLand ₂₀₁₆ /UrLand ₂₀₀₀)
Variables	2000	2016		2000	2016	
Log of total population	0.0525** (0.0219)	0.0154 (0.0184)		0.0547** (0.0235)	0.0489** (0.0187)	
Log of GDP per capita	0.162*** (0.0412)	0.174*** (0.0491)		0.0816 (0.0488)	0.0523 (0.0497)	
Log of total land area	-0.0424** (0.0205)	-0.0173 (0.0168)	-0.377*** (0.123)	-0.0669** (0.0269)	-0.0653*** (0.0202)	-0.0144 (0.0483)
Log(total population 2016/total population 2000)			5.006*** (1.536)			-1.351 (0.924)
Log(nonurban population 2000)			0.211*** (0.0600)			0.00279 (0.0314)
Log(GDP 2016/GDP 2000)			0.674 (0.434)			0.733** (0.277)
GDP share from agriculture 2000	0.00575 (0.00356)	0.00157 (0.00599)	-0.0302** (0.0145)	0.00418 (0.00387)	0.00105 (0.00469)	0.0116 (0.0110)
GDP share from industry 2000	0.00319 (0.00235)	0.00122 (0.00322)	-0.0386** (0.0155)	-0.000472 (0.00232)	-0.000768 (0.00190)	0.00839 (0.0103)
Central and West Asia	0.111 (0.146)	0.0274 (0.0917)	-0.436 (0.448)	-0.0349 (0.114)	-0.0634 (0.104)	0.0274 (0.284)
South Asia	-0.208 (0.136)	-0.113 (0.0994)	-0.902* (0.513)	-0.163* (0.0941)	-0.140 (0.0954)	0.138 (0.447)

continued on next page

⁹ We did not apply population as weights to the regressions because our attempt here is to understand the relationship at the country level, so each country is treated equally.

Table 1 *continued*

	(1)	(2)	(3)	(4)	(5)	(6)
	Share of Urbanized Population		Log(UrPop ₂₀₁₆ /UrPop ₂₀₀₀)	Share of Urban Land		Log(UrLand ₂₀₁₆ /UrLand ₂₀₀₀)
Variables	2000	2016		2000	2016	
Southeast Asia	-0.138 (0.121)	-0.0472 (0.0843)	-0.371 (0.403)	-0.0694 (0.106)	-0.0483 (0.0963)	0.344 (0.262)
Pacific	-0.267* (0.141)	-0.182* (0.103)	-0.350 (0.600)	-0.186 (0.117)	-0.196* (0.113)	0.534 (0.329)
Constant	-1.277** (0.616)	-1.086 (0.642)	2.506*** (0.778)	-0.591 (0.633)	-0.316 (0.664)	-0.401 (0.382)
Observations	38	36	38	38	36	38
R-squared	0.829	0.807	0.642	0.682	0.699	0.399
Adjusted R-squared	0.774	0.740	0.510	0.580	0.595	0.176

GDP = gross domestic product, Pop = population, Ur = urban

Notes: Robust standard error in parentheses. ***p<0.01, **p<0.05, *p<0.1.

Source: Author's estimates.

First of all, the R-squared of the regression reached 0.83 for 2000 and 0.81 for 2016, both of which are strikingly high given the sample size. Compared to the R-squared with GDP per capita as the only regressor (0.45 for 2000 and 0.60 for 2016), the estimated models suggest that, in addition to development level, the other country factors and subregional dummies can explain a significant portion of the variation in urbanization rates.

For 2000, a country with a greater population, smaller land area, and higher GDP per capita, was likely to have a larger proportion of its residents living in urban areas. The coefficient estimates of these variables are all statistically significant. Holding these variables constant, a country's economic structure did not seem to be correlated with the urbanization rates. Relative to countries in East Asia, those in Central and West Asia had higher urbanization rates, while countries in South Asia, Southeast Asia, and the Pacific had lower urbanization rates, but only the estimate for the Pacific countries is statistically significant.

Moving to 2016, the magnitude of coefficients of population and land diminished and became statistically insignificant, whereas GDP per capita remained highly significant. An increase of 6% in GDP per capita implied a 1 percentage point increase in the proportion of urbanized population. Consistent with Figure 7b, the differences in urbanization rates across subregions were reduced, although countries in the Pacific still lagged behind those in East Asia.

The relationship between these covariates and the share of land for urban use could be distinct if urbanized land does not move parallelly with urbanized population during the course of urbanization. Hence, we also estimate equation (1) with the proportion of land urbanized as the dependent variable. The results in columns (4) and (5) of Table 1 suggest that the model could explain nearly 70% of the variation in the dependent variable. Similar to the population urbanization rates, higher shares of urbanized land is associated with greater total population and smaller land area. However, the estimates of these variables are stable and statistically significant for both 2000 and 2016 in the urban land model. The share of urbanized land also goes up with GDP per capita, but statistically insignificantly. This is probably because, with rising income level, not only do more people live in cities, they also live spatially closer to each other. The estimates for subregion dummies suggest that East Asian countries devote more land as a percentage of their total land area to urban areas as compared to countries in other subregions. However, these estimates generally lack statistic precision.

The above analyses look at urbanization from a static perspective. To understand how the advance of urbanization is correlated with some key factors across the region, we estimated the following model with the log change in the urbanized population or land from 2000 to 2016 as dependent variable:

$$\log \frac{UrPop_{2016}}{UrPop_{2000,c}} = a + b_1 \log \frac{Pop_{2016}}{Pop_{2000,c}} + b_2 \log non - UrPop_{2000,c} + b_3 \log \frac{GDP_{2016}}{GDP_{2000,c}} + b_4 \log agr\ share_{2000,c} + b_5 \log ind\ share_{2000,c} + b_6 \log area_c + \gamma_s + e_c \quad (2)$$

wherein we included population growth, the proportion of the population initially not urbanized, GDP growth, initial economic structure, land area, and subregion dummies as explanatory variables. It is expected that the growth of GDP is positively related to the growth of urban population in light of the view that urbanization and growth are “mutually self-reinforcing processes” (Martin and Ottaviano 1999).

The model shows significantly higher explanatory power for the advance of urbanization in regard to population than for land (R-squared 0.64 versus 0.40). Column (3) of Table 1 indicates that faster growth in total population and GDP, and a greater proportion of population not urbanized in 2000, resulted in faster urban population growth. However, the estimate for GDP growth is not statistically significant. On the other hand, countries with more land area and higher initial shares of outputs from agricultural and industrial sectors tended to have the population migrate to cities more slowly. Across subregions, East Asian countries had the most rapid increase in urban population, while South Asian countries had the least.

The results in column (6) of Table 1 for urban land growth is very different. Most coefficient estimates are smaller in magnitude and statistically insignificant, except the one for GDP growth, which is slightly bigger than its counterpart for urban population but statistically significant. In sum, we do get some evidence that economic growth and urbanization are closely linked and move together.

It is worth noting that the above results should be interpreted as correlation rather than causality. The level or advance of urbanization could also have substantial influence on some of the covariates in the models such as GDP per capita, GDP growth, and economic structure. More data and analyses are needed to obtain the causal effects of these variables on urbanization.

IV. THE URBAN SYSTEM

Cities in a country are connected to one another through flows of goods, services, and people, thereby constituting a system. The urban system consists of large, medium, and small cities; cities that host all types of industries, and others specializing in a few products and services; and cities that incubate new products, and others that focus on the production of mature products (Duranton and Puga 2001). An urban system is more dynamic than people normally perceive. As firms and jobs move across cities frequently, the industrial composition of cities shifts with them. As a consequence, the size of a city relative to other cities changes continuously, though not as rapidly as industries move. In other words, the size rankings of individual cities are far from being fixed, though the overall distribution of city size tends to be stable over time (Duranton 2007).

Developing economies are generally in the middle of the urbanization process. While their urban systems may be as dynamic as those of developed economies, they may not yet have reached a steady state of size distribution. However, studying the evolution of a country's urban system could help us learn about the direction in which its urbanization is heading. We therefore focused on three aspects of urban systems in the region: the distribution of population across city sizes, primate cities, and Zipf's law.

A. Distribution of Population across City Sizes

Figures 8a and 8b show changes in the distribution by area from 1992 to 2016 and by population from 2000 to 2016, respectively, in the 43 economies across Asia and the Pacific. Not surprisingly, both distributions of city size, in terms of physical land area and population, shifted to the right considerably. The average (median) area of a natural city increased from 154 (28) km² in 1992 to 400 (157) km² in 2016, representing a 4% (7.4%) average increase per annum. The average (median) population of a natural city rose from 610,000 (180,000) to 970,000 (320,000), or by 2.9% (3.7%) per annum.

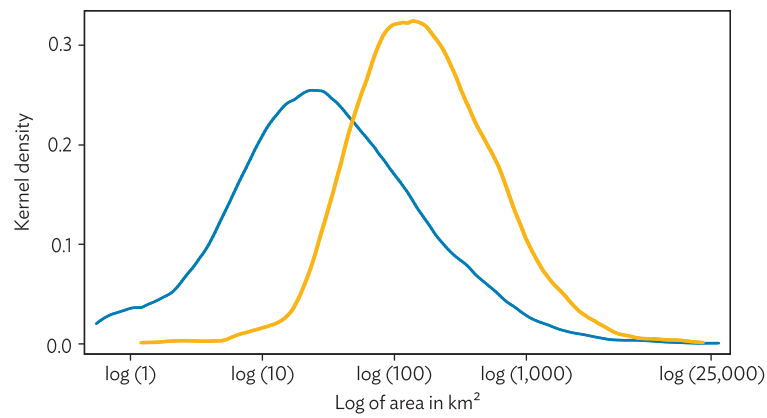
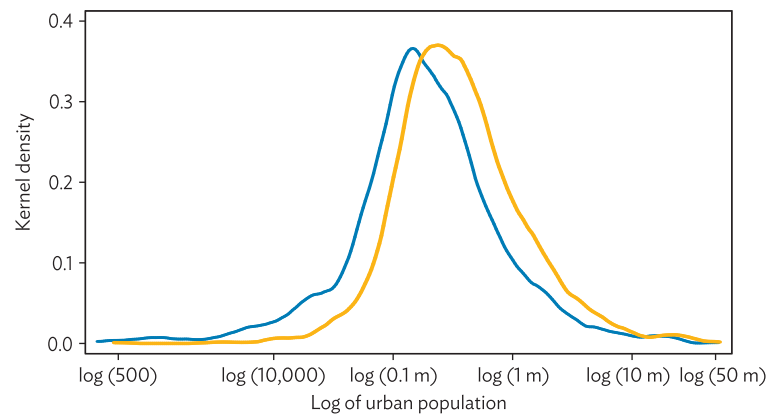
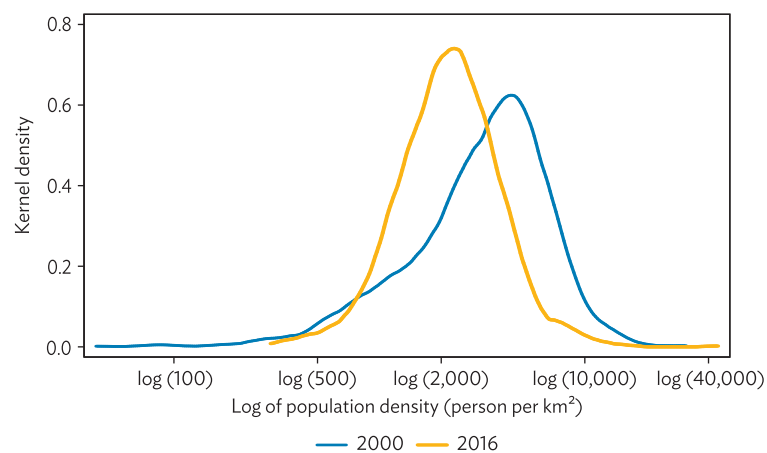
The figures imply that the physical areas of cities have grown faster than the population sizes. As a result, we see the density distribution moved to the left in Figure 8c. The average (median) density of population within the natural cities decreased from 3,910 (3,490) people per square kilometer in 2000 to 2,570 (2,178) people per square kilometer in 2016. Meanwhile, the variance of the density also decreased with fewer cities having extremely low or high densities.

Examining individual countries reveals considerable heterogeneity across the region. Figures 9a and 9b show that the overall sizes of cities in India, Indonesia, the PRC, and the grouping of the other 38 economies in the study, have all become larger in terms of both area and population. In contrast, the size distributions of Japan shifted to the left, though moderately, regardless of city area or population. The case of Pakistan is more complex. The areas of cities in Pakistan have generally grown between 1992 and 2016, but to a much lesser extent than most other Asian countries (except Japan). However, the distribution of urban population seems to have become more polarized. The number of very small cities and very large cities increased, with a decline in the number of mid-sized cities.

The distribution changes in city area and population translate into three types of changes in urban density in the region (Figure 9c).

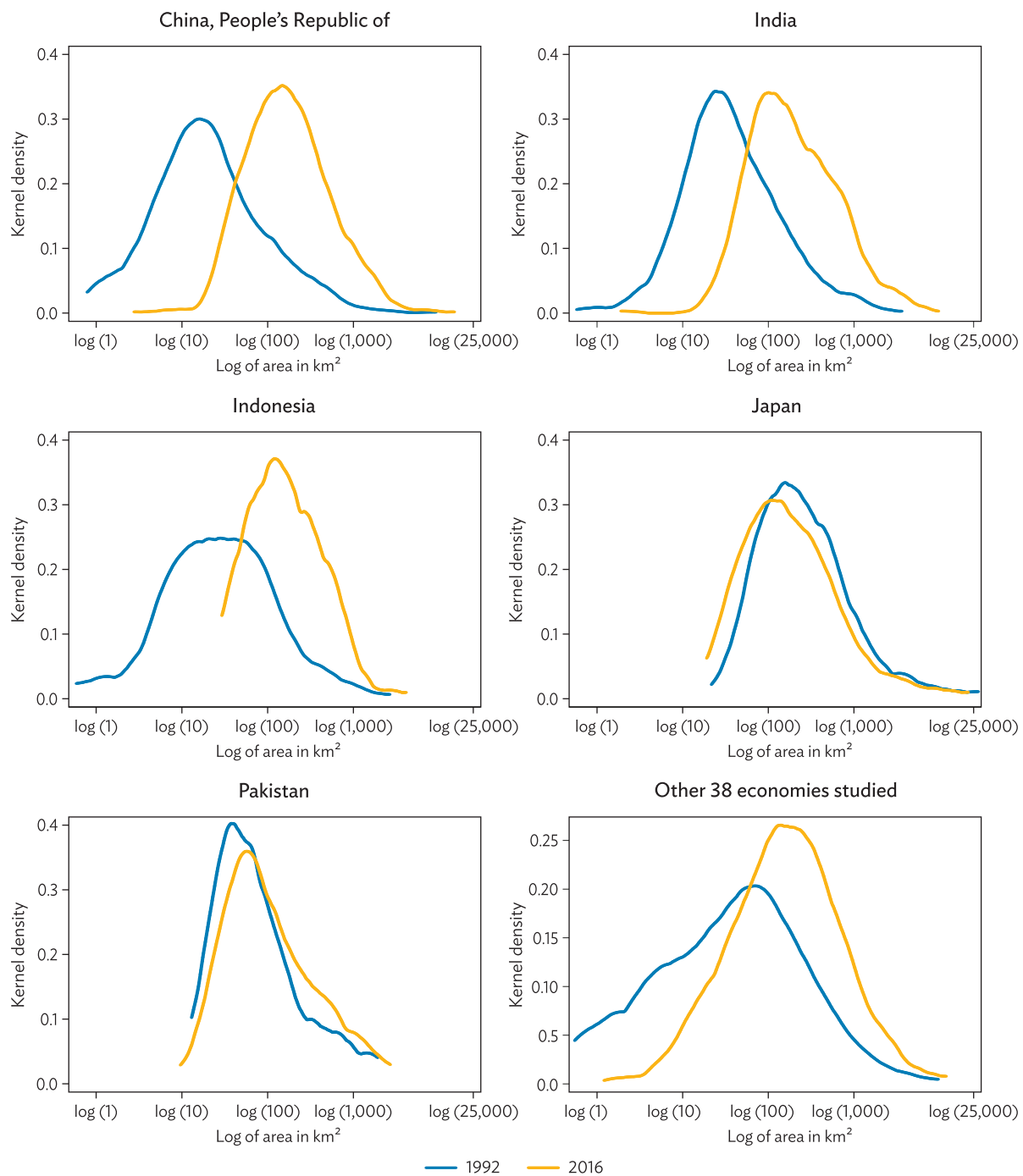
In India and the PRC, there was a significant decrease in density from 2000 to 2016, which also drove the region-wide density decrease discussed above. The phenomenon in the PRC has been documented in earlier literature (e.g., Henderson, Quigley, and Lim 2009) and the trend does not appear to be reversing. It may be partly explained by the fact that leasing land, converted from rural to urban use, has become increasingly important as a fiscal revenue source for local governments in the PRC, even as population inflows lag. In India's case, it is possible that strict land regulations, such as low floor area ratios in city cores, have expanded urban sprawl. This urbanization of land without the urbanization of population causes concerns about the inefficient use of land and foregone benefits of agglomeration.

For Indonesia and the group of other 38 economies in the dataset, the distributions of urban density show striking stability over time. It is interesting to note that a number of cities in Indonesia had their density move closer to the middle of the distribution, with much fewer cities showing extremely low or high density.

Figure 8: Distribution of City Area, Population, and Density across 43 Economies Studied**a. Distribution of city area****b. Distribution of city population****c. Distribution of city density**

km² = square kilometer, m = million.

Source: Author's estimates.

Figure 9: Distribution of City Area, Population, and Density across 38 Economies Studied**a. Distribution of city area**

km² = square kilometer.

Source: Author's estimates.

continued on next page

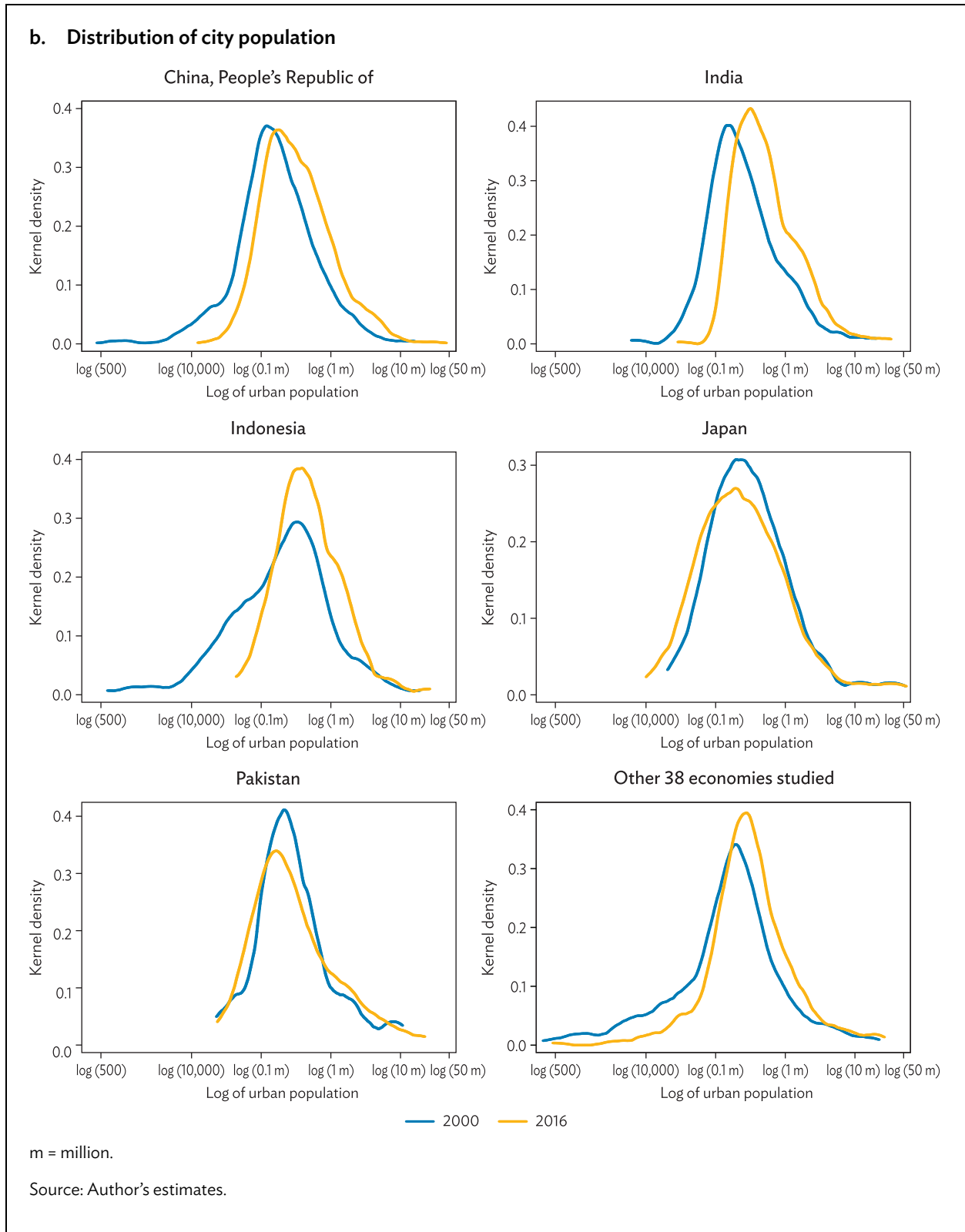
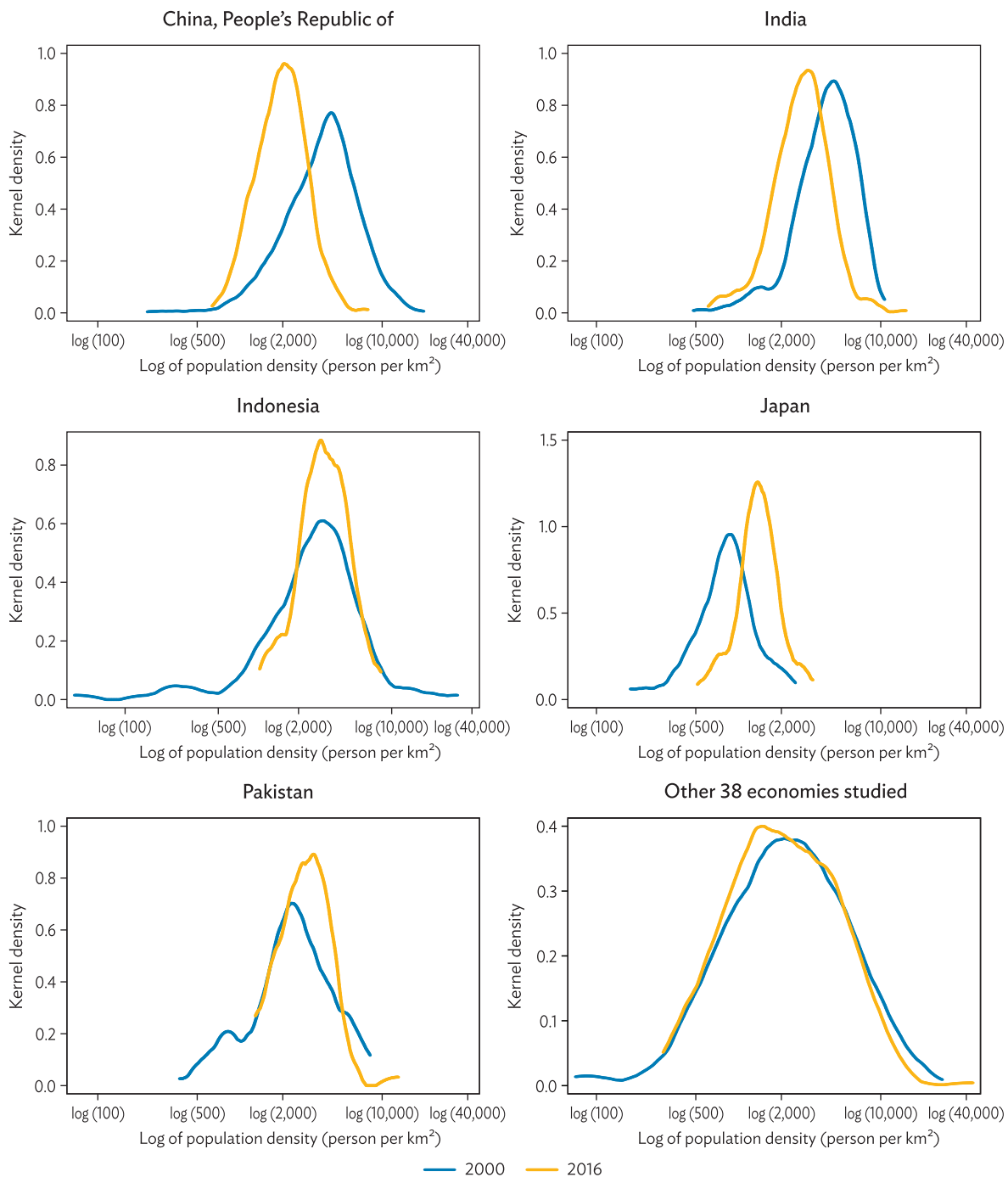
Figure 9 *continued**continued on next page*

Figure 9 continued

c. Distribution of city density

km² = square kilometer, m = million.

Source: Author's estimates.

Japan and Pakistan belong to the third type, whereby cities have generally become more densely populated as the distributions of city density shift to the right in both countries.

Table 2 shows the counts of natural cities in each of the five population categories: (i) below 0.1 million, (ii) 0.1 million to 0.5 million, (iii) 0.5 million to 1 million, (iv) 1 million to 5 million, and (v) above 5 million. The latter two are referred to as big cities and mega cities, respectively. In 2000, there were only 27 mega cities in the dataset for Asia and the Pacific. More than half of all the natural cities were home to 0.1 million to 0.5 million people, while another quarter had populations of fewer than 100,000. By 2016, the category of 0.1 million to 0.5 million was still the largest. However, the number of cities with fewer than 100,000 people shrank by more than 60%, and the three larger categories expanded significantly. The number of big cities and mega cities increased by 82% and 56%, respectively.

Table 2: Counts of Natural Cities by Size

Size	All 43 Economies Studied	China, People's Republic of	India	Indonesia	Japan	Pakistan	Other 38 Economies Studied
Year 2000							
Less than 0.1 m	419	219	58	30	11	9	92
0.1–0.5 m	794	343	185	41	37	39	149
0.5–1 m	155	67	34	11	11	6	26
1–5 m	130	46	36	10	6	6	26
5m and above	27	5	7	1	3	3	8
Total	1,525	680	320	93	68	63	301
Year 2016							
Less than 0.1 m	161	82	1	4	20	10	44
0.1–0.5 m	842	380	177	51	29	35	170
0.5–1 m	245	110	64	16	9	6	40
1–5 m	237	95	69	19	7	10	37
5m and above	42	13	9	3	3	2	12
Total	1,527	680	320	93	68	63	303

m = million.

Source: Author's estimates

Similar patterns and trends are observed in individual scenarios for India, Indonesia, the PRC, and the group of 38 economies. In the PRC, for example, the number of big cities and mega cities more than doubled with the number of cities, with populations below 0.1 million declining by 63% from 2000 to 2016.

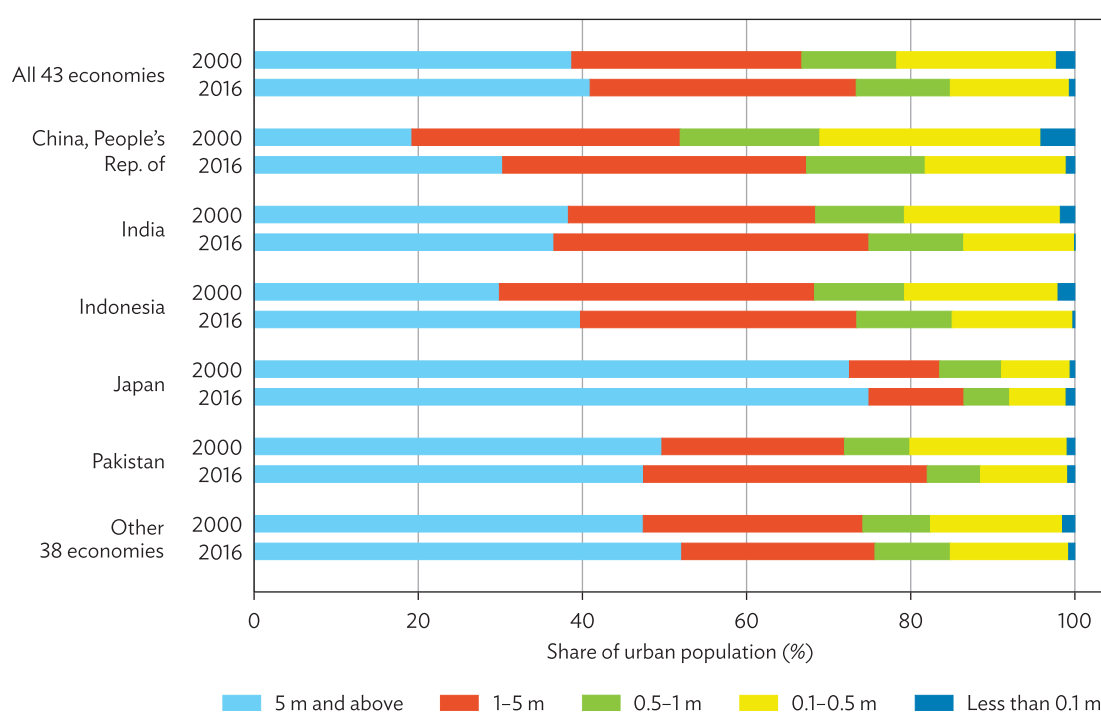
By contrast, in Japan, the category for the smallest cities expanded at the expense of the second and third categories (0.1 million to 0.5 million and 0.5 million to 1 million). This could result from people who once lived in cities of the second and third categories moving to big cities or mega cities—with these medium-sized cities dropping into the first category—as well as a decline in Japan's total urbanized population. In Pakistan, the number of mega cities fell from 3 to 2, although big cities increased from 6 to 10.

Figure 10 outlines how different city sizes account for the total urban population in Asia and the Pacific, in the top five countries individually, and in the remaining 38 economies as one group. In 2016, about 73% of the urban population across the region was living in cities with 1 million or more inhabitants. The proportions were similar in India (75%), Indonesia (74%), and the group of

38 economies (76%). The PRC had a lower proportion at 67%, although this figure rose from just 52% in 2000. The lower proportion in the PRC was partly because the country's mega cities accounted for only 30% of the urbanized population in 2016, while that percentage was 37% in India, 40% in Indonesia, 47% in Pakistan, and a staggering 75% in Japan.

The chart also shows that the proportion of urban population living in big cities and mega cities increased while the proportion residing in cities of fewer than 0.5 million people decreased from 2000 to 2016. The exception is that the mega cities in India and Pakistan host a smaller share of urban population in 2016 as compared to 2000, although the share of urban inhabitants in big cities has gone up remarkably in both countries.

Figure 10: Distribution of Population across Different Sizes of Cities



m = million.

Source: Author's estimates.

B. Primate Cities

Primate city refers to the largest city in any given country. A majority of these cities are the capitals of their respective countries. It is often the primate city that attracts the most attention when interest is expressed in a particular country. Because of the long history the city usually possesses, as well as the concentration of wealth and power within it, the primate city can be seen representative of the country in terms of politics, business, tourism, and culture. Primate cities are not short of attention from urban researchers either. Economists try to explain the driving forces behind these cities, and concern themselves with the consequences of primate city favoritism within the urban system and the overall economy (Ades and Glaeser 1995, Henderson 2005, Duranton 2008).

Table 3 provides some stylized facts about primate cities in Asia and the Pacific. We focus on 33 economies that had at least 2 natural cities present in our data.¹⁰

First, in all 33 countries, the primate cities in 2016 were also the largest back in 2000. The stability of the primate cities to some extent reflects the stability of urban systems in these countries. Of the 33 primate cities, six were not the capital of their respective countries: Almaty in Kazakhstan, Guangzhou in the PRC, Ho Chi Minh in Viet Nam, Karachi in Pakistan, Mary in Turkmenistan, and Yangong in Myanmar. However, Yangong and Almaty, had previously been the capital. The other four all have long histories as urban settlements in each country.

Second, the size of primate city varies greatly across countries, but they all account for a high share of total urban population except in the PRC and India. Several primate cities had population over 10 million with the largest, Tokyo, exceeding 50 million in 2016. On the other end of the spectrum, cities with less than 1 million population were already the largest in their countries. Despite the tremendous size difference, these cities, except Guangzhou and Delhi, account for 20% or more of the total urban population in their countries with quite a few reaching above 80%. The degree of primacy could also be viewed through the size difference between the largest city and the second largest city. In several countries such as Armenia, Cambodia, Nepal, and Sri Lanka, the size gaps between the primate city and the city next to it are staggering.

Table 3. Primate Cities in Asia and the Pacific, 2016

Economy	Annual Urban Population Growth (%)	Primate City	Capital City	Population (million)	Share of Urban Population (%)	Annual Growth (%)	Second Largest City	Share of Urban Population (%)
Afghanistan	10.7	Kabul	Y	4.8	64.4	9.9	Herat	8.8
Armenia	-1.1	Erevan	Y	1.4	91.4	-0.5	Leninakan	6.3
Azerbaijan	0.2	Baku	Y	2.7	87.0	0.3	Ganja	13.0
Bangladesh	2.2	Dhaka	Y	15.1	49.2	2.8	Chittagong	12.2
Bhutan	7.8	Thimphu	Y	0.1	74.8	9.0	Phuntsholing	25.2
Brunei Darussalam	3.5	Bandar Seri Begawan	Y	0.4	87.0	4.9	Kuala Belait	8.4
Cambodia	3.3	Phnom Penh	Y	1.6	90.3	2.7	Battambang	7.4
PRC	4.2	Guangzhou	N	45.5	8.5	6.7	Shanghai	4.5
Fiji	2.0	Suva	Y	0.2	75.4	1.9	Lautoka	24.6
Georgia	-0.7	Tbilisi	Y	1.3	65.7	1.3	Batumi	10.9
India	3.3	Delhi	Y	33.0	8.9	3.1	Kolkata	6.5
Indonesia	2.9	Jakarta	Y	26.4	26.4	2.1	Bandung	7.0
Japan	-0.1	Tokyo	Y	54.5	50.9	0.2	Osaka	17.6
Kazakhstan	2.1	Almaty	N	2.1	25.3	3.1	Chimkent	9.7
Kyrgyz Republic	3.6	Bishkek	Y	1.1	58.7	4.0	Osh	28.6

continued on next page

¹⁰ Economies excluded are Hong Kong, China; Singapore; the Cook Islands; Maldives; the Federated States of Micronesia; Palau; Samoa; Solomon Islands; and Vanuatu.

Table 3 *continued*

Economy	Annual Urban Population Growth (%)	Primate City	Capital City	Population (million)	Share of Urban Population (%)	Annual Growth (%)	Second Largest City	Share of Urban Population (%)
Lao PDR	4.5	Vientiane	Y	0.7	78.9	4.2	Savannakhet	11.3
Malaysia	4.2	Kuala Lumpur	Y	7.9	34.9	4.4	Johor Bahru	8.5
Mongolia	4.2	Ulaanbaatar	Y	1.3	88.7	4.4	Erdenet	5.9
Myanmar	6.4	Yangon	N	5.8	52.4	5.5	Mandalay	17.6
Nepal	2.2	Kathmandu	Y	2.3	74.8	3.2	Birganj	8.7
Pakistan	2.0	Karachi	N	22.4	33.2	4.7	Lahore	14.2
Papua New Guinea	0.4	Port Moresby	Y	0.3	70.2	0.9	Lae	29.8
Philippines	2.6	Metro Manila	Y	24.3	61.7	2.2	Cebu	7.1
Korea, Rep. of	0.6	Seoul	Y	26.5	55.2	1.1	Busan	13.9
Sri Lanka	0.2	Colombo	Y	4.7	87.7	0.6	Kandy	5.9
Taipei, China	0.3	Capital city of Taipei, China	Y	21.6	95.9	0.4	Ilan	1.9
Tajikistan	2.1	Dushanbe	Y	1.7	69.6	3.0	Khujand	17.4
Thailand	2.9	Bangkok	Y	17.0	64.8	3.1	Nakhon Pathom	4.7
Timor-Leste	17.1	Dili	Y	0.2	95.0	17.3	Baucau	5.0
Tonga	2.1	Nuku'alofa	Y	0.0	89.6	3.6	Neiafu	10.4
Turkmenistan	3.2	Mary	N	0.8	35.4	4.0	Ashgabat	33.6
Uzbekistan	1.4	Tashkent	Y	3.3	24.5	-0.2	Andizhan	16.4
Viet Nam	6.4	Ho Chi Minh	N	12.8	37.3	5.5	Ha Noi	23.8

Lao PDR = Lao People's Democratic Republic, PRC = People's Republic of China.

Source: Author's estimates

Table 3 also shows great variability in the average growth rates of primate cities between 2000 and 2016. They range from -0.5% to 17.3% per annum. Comparing the growth rates of the primate city and total urban population reveals the trends of primacy in the economy. For those with primate city growing faster than the total urban population, such as Bangladesh, Kazakhstan, Nepal, and Pakistan, the degree of primacy has been strengthened. There are countries moving the other direction including Afghanistan, Indonesia, Myanmar, Viet Nam, etc.

It is interesting to examine how the size of the primate city correlates with its country characteristics. Intuitively, the primate city should be bigger if the country has a large population. However, if the country is larger in land area, the primate city may be smaller because there may be more cities established across the country, giving people a wider choice of where to live. GDP per capita, which is highly correlated with urbanization level, may be correlated with the size of a primate city in a nonlinear manner. At an early stage of a country's development, the primate city offers disproportionately more economic opportunity. As further economic growth spreads opportunities across the country, the growth of the primate city slows. Holding GDP per capita constant, the economic structure of a country matters as well. A service-oriented economy may require a more concentrated urban population. All else being equal, a primate city that is the national capital might be larger than its peers because its political influence attracts people from other areas of the country. Finally, primate cities that are also port cities might grow bigger due to the presence and expansion of trade-related industries.

To quantify these relationships, we ran a regression on the 33 countries in question (Table 4). The dependent variable is log of population of the primate city. The primary explanatory variables include log of a country's total population, log of total land, log of GDP per capita, indicators of whether the primate city is the national capital and/or has a port, and subregional dummies (column 1).

Table 4: Correlates of Primate Cities

Variables	(1)	(2)	(3)
Log of total population	0.806*** (0.0979)	0.759*** (0.0861)	
Log of total land area	-0.172* (0.0873)	-0.111 (0.0745)	-0.179* (0.0928)
Log of GDP per capita	0.0152 (0.0980)	-0.155 (0.189)	-0.168 (0.311)
Log of nonurban population			0.356** (0.129)
Log of urban population minus largest city			0.307*** (0.0954)
Share of agricultural value added to GDP (%)		-0.0385 (0.0241)	-0.0333 (0.0328)
Share of industrial value added to GDP (%)		-0.0184* (0.0103)	-0.00882 (0.0166)
Capital city dummy	0.432* (0.229)	0.398* (0.217)	0.505* (0.255)
Port city dummy	0.861*** (0.227)	0.871*** (0.289)	0.853** (0.303)
Central and West Asia	-0.211 (0.392)	-0.286 (0.415)	-0.747 (0.460)
South Asia	-0.845* (0.427)	-0.800* (0.417)	-1.006* (0.533)
Southeast Asia	-0.727** (0.345)	-0.516 (0.357)	-0.744* (0.373)
Pacific	-2.288*** (0.479)	-2.107*** (0.418)	-2.668*** (0.471)
Constant	3.354** (1.401)	5.814** (2.427)	9.108** (3.558)
Observations	33	30	30
R-squared	0.966	0.974	0.958
Adjusted R-squared	0.953	0.958	0.928

GDP = gross domestic product.

Notes: Dependent variable is log of population of the primate city. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: Author's estimates.

Despite the small sample size and the gigantic range of the dependent variable (thinking of Tokyo and Dili), the model explains 97% of the variation in the primate city size. A country's total population was strongly and positively correlated with the primate city's population. With a 100% increase in the former, the latter increased by 81%. As expected, the coefficients of land area were negative, but only marginally statistically significant. GDP per capita was positively related to the size of the primate city, but not statistically significant. With all other control variables, the primate cities that were national capitals were, on average, 43% larger than the primate cities that were not capitals, and

the primate cities that have a port were 86% larger than those without one. Compared to countries in East Asia, those in South Asia, Southeast Asia, and the Pacific have statistically smaller primate cities.

Column 2 includes shares of GDP attributed to agricultural and industrial value addition as measures of development stage. Both have negative coefficient estimates, as expected, and the coefficient for the share of the industrial sector is significant at 10% level. The larger the share of economic output derived from industries, the less concentrated the population in the top city. This is perhaps because industrial firms can be more geographically dispersed than firms providing services to urban populations.

Following Ades and Glaeser (1995), we replaced the total population with the nonurbanized population and the urbanized population excluding the primate city in the model (column 3). Both variables are positively associated with size of the primate city and statistically significant. The elasticities of primate city population to the nonurbanized population and urbanized population are 36% and 31%, respectively. The results are comparable with those in Ades and Glaeser (1995), which examined a global sample between 1970 and 1985.

Overall, we found that the primate city was generally larger in a country with greater total population and smaller land area. More industrialized economies tended to have smaller primate cities, although the relationship of income level to the size of the primate city is not clear. The primate city was considerably bigger if it was the country's capital and/or had a port. This suggests that political favoritism and trade may play a role in the formation and expansion of primate cities.

The regression results can be illustrated by a comparison between the Philippines and Viet Nam. Both countries are located in Southeast Asia, are of similar size in terms of land area and population, and belong to the lower-middle-income country grouping. Their largest cities, Manila and Ho Chi Minh, respectively, are both port cities. Manila, however, is a capital city whereas Ho Chi Minh is not. In 2016, the Philippines was more service oriented than Viet Nam, with the share of GDP from the service sector higher by 9 percentage points (60% versus 51%). The urban primacy was also more prominent in the Philippines than in Viet Nam. Growing from 5.4 million in 2000, Ho Chi Minh's population reached 12.8 million in 2016, but it remained about half the population of Metro Manila.

C. Zipf's Law in Selected Countries

A striking empirical regularity about cities, which has been documented in literature (e.g., Rosen and Resnick 1980), is that the city size distribution of a country follows Zipf's law. This law states that the population of the N th largest city in a given country is $1/N$ times the population of the largest city. This is equivalent to characterizing the city size distribution as a power law distribution with a coefficient of minus 1 when the country has a large number of cities.

The underlying theory suggests that Zipf's law is more relevant to describe city size distribution in a society wherein the urbanization reaches a steady state (Gabaix 1999). To what extent the rank-size rule is applicable to rapidly urbanizing countries remains a question. Meanwhile, Zipf's law could shed light on how such countries should consider their public policies to shape their urban systems in the medium to long term. There have been a few studies examining the city size distribution in the countries of developing Asia. They include Chauvin et al. (2017); Schaffar and Dimou (2012); Soo (2005, 2014); and Colmer (2016). These studies, however, use either city data based on administrative boundaries or data of larger areas (e.g., districts in India) to test Zipf's law.

Here we present analyses regarding Zipf's law in India, Indonesia, Japan, Pakistan, and the PRC, based on our natural city data for 2000 and 2016. Following Gabaix and Ibragimov (2011), we examine the relationship between the logarithm (log) of city population and the log of rank *minus one-half*, which corrects bias arising from estimating the relationship between the log of population and the log of the rank. In other words, we estimate a simple regression:

$$\log\left(n_i - \frac{1}{2}\right) = \alpha + \beta \log S_i + \varepsilon_i \quad (1)$$

where n_i is the rank and S_i is the size of population of city i . The coefficient of city size, β , is equal to -1 if Zipf's law holds. If the coefficient is less than -1 (absolute value greater than 1), it implies that the small cities are too big or the large cities are too small as compared to what Zipf's law predicts. If the coefficient is greater than -1 (absolute value less than 1), Zipf's law suggests that the small cities are too small and/or the large cities are too big.

Table 5 shows the estimation results for 2000 and 2016 for the top five countries in our dataset. In the upper panel, we have included all the natural cities in the regressions. First, all the coefficient estimates were greater than -1 and statistically highly significant, with India's estimates closest to -1 and Japan's generally deviating the most. This could suggest some degree of excessive concentration of urban population across these countries. Second, the coefficients have moved closer to -1 by 2016 for India, Indonesia, and the PRC. In particular, the estimates change from -0.51 to -0.86 for Indonesia. This is probably due to faster growth of smaller cities relative to larger cities in these countries, as suggested in Table 2. On the contrary, the urban systems in Japan and Pakistan experienced further, though moderate, deviation from Zipf's law from 2000 to 2016.

Table 5: Zipf's Law Regressions

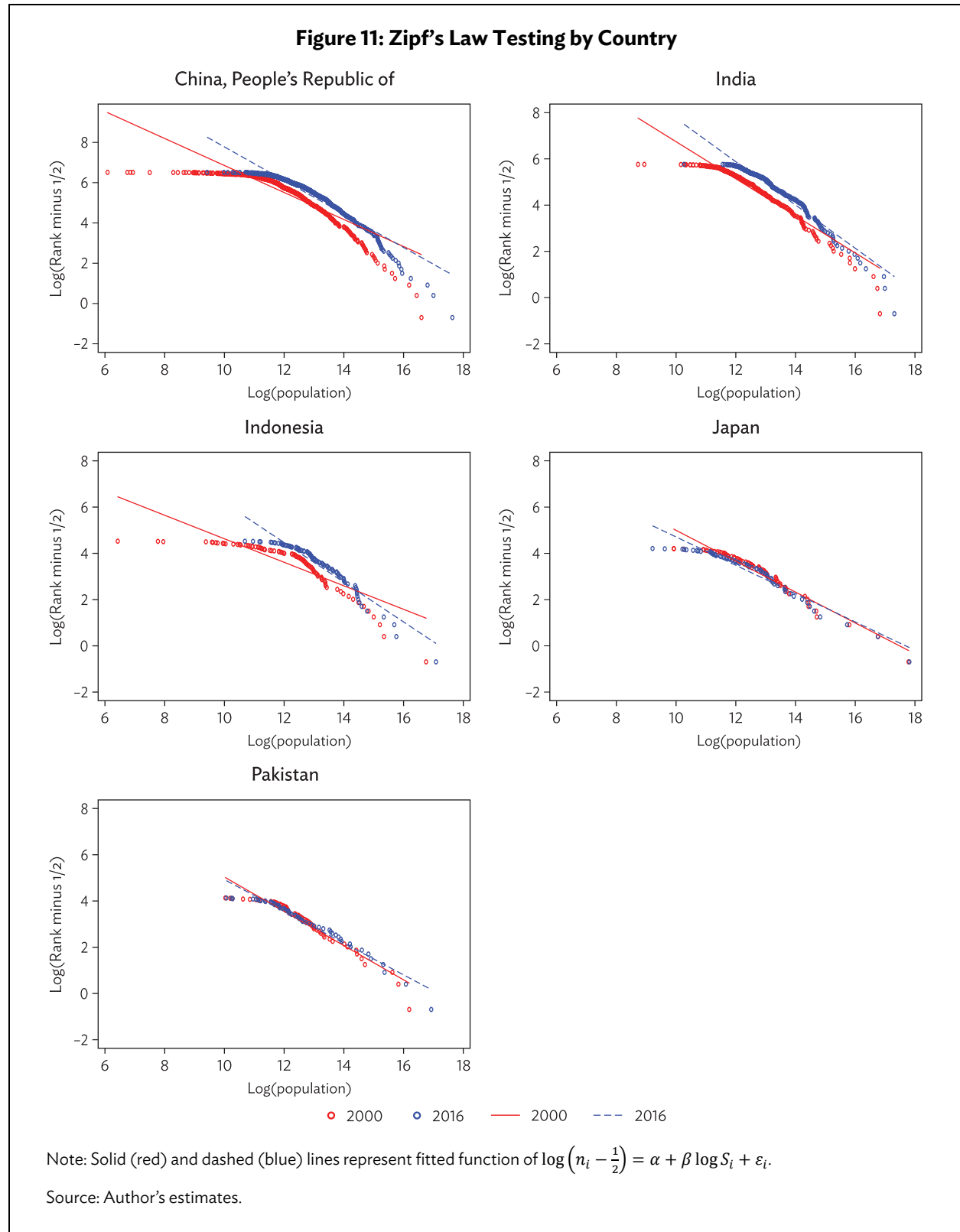
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables	PRC 2000	PRC 2016	India 2000	India 2016	Indonesia 2000	Indonesia 2016	Japan 2000	Japan 2016	Pakistan 2000	Pakistan 2016
Full sample										
Log (population)	-0.671*** (0.0352)	-0.832*** (0.0227)	-0.802*** (0.0340)	-0.935*** (0.0289)	-0.508*** (0.0690)	-0.858*** (0.0529)	-0.668*** (0.0359)	-0.613*** (0.0359)	-0.741*** (0.0576)	-0.687*** (0.0423)
Constant	13.56*** (0.424)	16.10*** (0.286)	14.76*** (0.426)	17.10*** (0.378)	9.714*** (0.841)	14.76*** (0.693)	11.67*** (0.463)	10.84*** (0.452)	12.45*** (0.724)	11.79*** (0.530)
Observations	680	680	320	320	93	93	68	68	63	63
R-squared	0.796	0.919	0.923	0.955	0.742	0.913	0.938	0.933	0.918	0.952
Population above 100,000										
Log (population)	-1.068*** (0.0207)	-0.945*** (0.0196)	-0.938*** (0.0231)	-0.950*** (0.0244)	-0.968*** (0.0336)	-0.936*** (0.0393)	-0.766*** (0.0173)	-0.743*** (0.0169)	-0.884*** (0.0443)	-0.764*** (0.0319)
Constant	18.63*** (0.257)	17.62*** (0.249)	16.55*** (0.289)	17.30*** (0.317)	15.77*** (0.441)	15.82*** (0.513)	13.00*** (0.217)	12.60*** (0.216)	14.35*** (0.553)	12.84*** (0.398)
Observations	461	598	262	319	63	89	57	48	54	53
R-squared	0.978	0.960	0.980	0.965	0.969	0.949	0.982	0.984	0.980	0.983

PRC = People's Republic of China.

Notes: Dependent variable is log of rank of the city subtracting 1/2. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Author's estimates.

Figure 11 plots raw data as well as the fitted line for each country by year. The raw data demonstrate some concave relationship between log of city size and log of adjusted rank: the left portion



of the curves is flatter with smaller cities, while the slopes steepen among larger cities. The curvature is particularly prominent for Indonesia and the PRC in 2000. Across the charts, the turning points mostly occur around log of population equal to 12, which corresponds to a population of 160,000 people.

The charts suggest that cities above a certain scale follow the linear rank-size relationship more closely. Chauvin et al. (2017) examined urban areas with 100,000 or more inhabitants, and confirmed the relevance of Zipf's law in India and the PRC. Therefore, we reestimated equation (1) with samples restricted to cities with more than 100,000 inhabitants in each country-year (lower panel of Table 5). The R-squared indicates significant improvements in model fitting as compared to the results using full samples. Ranging between -0.94 and -1.07 , the estimated β for India, Indonesia, and the PRC get substantially closer to -1 in both years. These results are in line with those in Soo (2005), who obtained estimates close to -1 for India and Indonesia found when using urban agglomeration data. For Japan and Pakistan, the estimates also become closer to -1 , but the gaps remain significant. This suggests that, even when not counting the very small cities (fewer than 100,000 people), population has tended to concentrate in larger cities in these two countries.

In summary, tests of Zipf's law suggest that small cities are too small and/or big cities are too big in all top five countries. This may be because smaller cities follow different growth processes from those of the larger cities in these countries. When we restricted samples to cities with a population above 100,000, Zipf's law holds reasonably well and stably for India, Indonesia, and the PRC. In Japan and Pakistan, however, small cities still seem to be too small and/or large cities too large as compared to what Zipf's law predicts. Given the distinct development stages these two countries are at, the phenomenon may be explained by different reasons, which are worth investigating in future studies.

V. GROWTH OF CITIES AND EMERGENCE OF CITY CLUSTERS

The previous two sections have presented and analyzed the overall patterns and trends of urbanization in the region, and examined cities from an urban system perspective. However, these aggregate results could mask pronounced differences in the evolution of individual cities. It is thus worth looking at the dynamics of individual cities. This section is concerned with how cities in the region have grown and what factors have played a driving role. We also illustrate the emergence of city clusters—urban agglomerations that rise beyond the scale of natural cities.

A. Simple Stylized Facts about City Dynamics

While a majority of the natural cities in our dataset experienced growth in both land area and population—substantial growth for quite a few of them—there were also cities that contracted in land area and/or population. Table 6 provides the counts of cities that have expanded and/or contracted in terms of land area and population. It sets out these counts for the dataset as a whole, for the top five countries, and for the group of the remaining 38 economies. The majority of natural cities (1,241 out of 1,525) expanded in both area and population from 2000 to 2016. This was especially the case in India, Indonesia, the PRC, and the group of 38 economies. However, only two cities in Japan and 17 cities in Pakistan expanded on both measures. In these two countries, the most common scenario was cities contracting in both area and population. This is consistent with our earlier finding about declining urbanization rates in Japan and Pakistan since the mid-2000s.

Table 6: Number of Cities by Changes in Size, 2000–2016

Size Changes	All 43 Economies Studied	China, People's Rep. of	India	Indonesia	Japan	Pakistan	Other 38 Economies Studied
Expansion of population and land	1,241	621	301	76	2	17	224
Population expansion and land contraction	65	5	4	9	14	10	23
Population contraction and land expansion	104	47	5	2	1	11	38
Contraction of population and land	115	7	10	6	51	25	16
Total	1,525	680	320	93	68	63	301

Source: Author's estimates

The physical size of a city and its number of inhabitants generally move in the same direction. The two categories in which land area and population both either grow or contract jointly account for 89% of cities in our sample. However, the other two categories are not negligible. There were 104 cities that experienced growth in land coverage with a reduction in population, of which 47 were in the PRC. This could arise from local governments being incorrectly incentivized to expand their urban footprints despite a net outflow of population. Another 65 cities attracted more people to settle in them while losing some urban land, with 14 such cities identified in Japan and 10 in Pakistan. The resulting rise in density could lead to increased benefits of agglomeration and land-use efficiency, but may incur additional costs due to congestion. Sound city governance is important to mitigate the latter.

Table 7 offers another perspective on city dynamics. In each of the top five countries, cities are assigned to a size quintile by population and area for 2000 and 2016. The table tracks the numbers of cities that have moved to a different quintile from 2000 to 2016. Regardless of population or area

Table 7: Changes in City Size Quintile, 2000–2016

	China, People's Rep. of	India	Indonesia	Japan	Pakistan
Population quintile					
Down	142	82	21	7	10
Stay	409	175	49	53	44
Up	129	63	23	8	9
% of cities changing	39.8	45.3	47.3	22.1	30.2
Area quintile change					
Down	164	80	20	15	15
Stay	356	169	51	37	32
Up	160	71	22	16	16
% of cities changing	47.6	47.2	45.2	45.6	49.2
Total number of cities	680	320	93	68	63

Source: Author's estimates.

ranking, we see a high proportion of cities move across quintiles. The percentage falls between 40% to 50% except for the Japanese and Pakistan cities in terms of population size. This exercise sheds light on the relative size changes of cities within a country. Underlying these figures is the movement of large numbers of people and industries across cities. The next two subsections investigate what forces have been driving these dynamics.

B. Testing Gibrat's Law

Gibrat's law states that the population growth of a city is independent of its initial size. In other words, cities of different sizes follow homogeneous growth processes, that is, common mean and variance. When Gibrat's law holds, the distribution of cities in a certain size range will follow Zipf's law in the steady state (Gabaix 1999, Eeckhout 2004). If we estimate a simple model,

$$\Delta_{t_0}^t \log pop_i = \alpha + \beta \log Pop_{i,t_0} + \epsilon_i \quad (2)$$

where the dependent variable is the log change of population of city i between t_0 and t , and the regressor is log of initial population in t_0 , Gibrat's law implies that the coefficient, β , should not be statistically different from zero.

Academic literature shows a variety of results when testing Gibrat's Law. For instance, Chauvin et al. (2017) found that the growth of India's districts between 1980 and 2010 was negatively correlated with the initial population. On the contrary, using town-level data from multiple censuses, Colmer (2016) found that the coefficient was significantly positive. A similar finding is documented in Hasan, Jiang, and Kundu (2018). Whereas all these studies reject Gibrat's law for India, they point to opposite outcomes. The negative correlation suggests that smaller cities have grown faster than larger cities in India, and vice versa.

The upper panel of Table 8 shows the estimation results of equation (2) for the region as a whole, for each of the top five countries, and for the group of the remaining 38 economies.¹¹ Of these seven cases, Gibrat's law was rejected in six of them.

The average urban growth rates from 2000 to 2016 were negatively correlated with the initial urban population in 2000 for India, Indonesia, the PRC, and the group of 38 economies. This is in line with the evidence that cities with fewer than 100,000 people represented the largest category in 2000 and diminished by 2016, while the categories for cities above 0.5 million all expanded in these countries (Table 2). The estimates also conform to the findings of Chauvin et al. (2017) for India and the PRC, indicating violation of Gibrat's law. A possible explanation is that, with rapid urbanization, these countries have not yet reached a steady state.

In contrast, populations grew disproportionately in the larger cities of Japan. This is consistent with the earlier evidence that people left smaller Japanese cities for larger ones, resulting in greater concentration of urban population. The coefficient estimate for Pakistan is small at 0.01 and statistically insignificant, suggesting Gibrat's law holds in Pakistan.

¹¹ To reduce the influence of extreme values on the estimation, we excluded cities with an average annual growth rate greater than 0.2, despite that our data validation does not identify problems with them. They were primarily very small cities in 2000.

Table 8: Testing Gibrat's Law

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	All 43 Economies Studied	China, People's Republic of	India	Indonesia	Japan	Pakistan	Other 38 Economies Studied
Full sample^a							
Log (Urban population, 2000)	-0.180*** (0.0128)	-0.160*** (0.0186)	-0.233*** (0.0279)	-0.332*** (0.0510)	0.0552** (0.0259)	0.0104 (0.0602)	-0.225*** (0.0277)
Constant	2.814*** (0.160)	2.608*** (0.231)	3.615*** (0.352)	4.869*** (0.637)	-0.926** (0.359)	-0.0914 (0.779)	3.363*** (0.337)
Observations	1,496	671	318	87	68	63	289
R-squared	0.391	0.137	0.239	0.402	0.043	0.001	0.645
Adjusted R-squared	0.375	0.136	0.237	0.395	0.0280	-0.0158	0.596
Cities with population > 100,000							
Log (Urban population, 2000)	-0.0551*** (0.0118)	-0.0165 (0.0177)	-0.130*** (0.0246)	-0.207*** (0.0639)	0.0185 (0.0208)	0.0830 (0.0546)	-0.0444* (0.0228)
Constant	1.190*** (0.154)	0.788*** (0.231)	2.250*** (0.314)	3.219*** (0.843)	-0.370 (0.293)	-1.042 (0.723)	0.963*** (0.301)
Observations	1,083	457	262	62	47	50	205
R-squared	0.239	0.001	0.086	0.139	0.010	0.032	0.429
Adjusted R-squared	0.216	-0.000822	0.0825	0.124	-0.0119	0.0117	0.342

^a The sample excludes cities with annual average growth rates greater than 0.2.

Notes: Dependent variable is log of urban population in 2016 divided by urban population in 2000. All 43 economies and the other 38 economies regressions include country fixed effects with standard errors clustered by country. Country specific regressions estimate robust standard errors. ***p<0.01, **p<0.05, *p<0.1.

Source: Author's estimates.

The Zipf's law estimation differed markedly when we restricted the sample to cities that had a population above 100,000 in 2000. We did this for Gibrat's law as well, and the results are shown in the lower panel of Table 8. First, we see that the magnitude of coefficients dropped substantially across the board, except for Pakistan. This implies that for cities in the upper end of the distribution, the growth rates were more equalized. The faster growth of cities with fewer than 100,000 people largely accounted for the results in the upper panel. Second, the estimates for the PRC and Japan became insignificant, with magnitudes below 0.02, so Gibrat's law seems to be applicable to the relatively large cities in those two countries. Third, the estimate for Pakistan increased by eightfold from 0.01 to 0.08. As the lack of statistical significance may be explained by the small sample size, the result suggests that larger cities grew faster than relatively smaller ones in Pakistan.

It is worth pointing out that the results from testing Gibrat's law do not contradict the population distributions across city sizes as presented in Figure 10. First of all, if Gibrat's law holds, the population grows proportionate to the initial city size. The absolute increase of inhabitants in larger cities could therefore dramatically exceed the increase in smaller cities. As a result, the larger cities will account for a much greater share of the total urban population over time. This is what we see for Japan, Pakistan, and the PRC in Figure 10. Even when the smaller cities grew faster—as was the case in India, Indonesia, and the group of 38 economies—the absolute increase in larger cities could still outnumber the increase in smaller cities, leading to a moderate shift of urban population toward the former.

While Gibrat's law fails to hold in general across Asia and the Pacific, our investigations have shown that cities with populations above 100,000 in 2000 had more homogeneous growth through to 2016 than did the smaller cities. Overall, the trend for urban population mobility could be described as smaller cities growing faster than larger cities, with the exception of Japan and Pakistan, while the distribution of urban population continues to be skewed toward larger cities.

C. Factors Driving City Growth

The interest in city growth goes beyond Gibrat's law. Cities act as engines of growth through improving input-output linkage, better matching workers to jobs, and more rapidly spreading knowledge. These benefits are all reflected in higher incomes for urban inhabitants. The continuous growth of a city means that people from outside still can find opportunities of being more productive to earn higher income in the city, despite a possible increase in living costs. On the contrary, unless there is a negative shock to the amenities, a shrinking city strongly indicates that productivity has stagnated and new opportunities become scarce there. It is thus important to know what types of cities will grow quickly and what factors drive urban growth. Such analysis could help reduce misallocation of public resources as governments need to plan ahead and invest public resources to support city development.

While it is ideal to perform a comprehensive study—as those authored by da Mata et al. (2007) on Brazil; Duranton (2016) on Colombia; and Hasan, Jiang, and Kundu (2018) on India—we had limited access to variables at the natural-city level. We therefore specifically looked at how city growth was affected by initial population and density, maximum and minimum temperature, precipitation, and geography proxied by the distance to the nearest port. The regression model is built on the one for Gibrat's law:

$$\Delta_{t_0}^t \log pop_i = \alpha + \beta \log Pop_{i,t_0} + X_i \gamma + \mu_i + \epsilon_i \quad (3)$$

where X_i is the vector of city-level variables in addition to the initial population. Country fixed effect, μ_i , is included when the sample involves multiple countries.

Table 9 shows the estimation results for the region as a whole, for each of the top five countries, and for the group of the remaining 38 economies. The model is estimated with the full sample (upper panel) and the sample excluding natural cities that had populations below 100,000 in 2000 (lower panel).

First of all, the results for Gibrat's law are generally robust after controlling for the additional city covariates. The smaller cities grew faster in India, Indonesia, the PRC, and the group of 38 economies under both samples. The negative coefficient for the PRC's larger-city sample became statistically significant with the additional controls. This scenario could be driven by government policies that intend to limit the size of large, dense cities. The estimates for Japan and Pakistan are not distinguishable from zero, implying a proportionate growth of cities in these countries. As a result, populations were increasingly concentrated in the larger cities.

Urban density plays an important role in influencing the flow of people into cities in Indonesia, the PRC, and the group of 38 economies. Holding initial size constant, the less densely populated cities grew faster in these countries. Higher density suggests challenges in land access and lack of amenities due to congestion, so cities with easier access to land and/or lesser congestion are more attractive to firms and people. The estimates for India and Japan, and Pakistan excluding small cities, were positive, but not statistically significant.

Table 9: Driving Factors for City Growth, 2000–2016

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	All 43 Economies Studied	China, People's Republic of	India	Indonesia	Japan	Pakistan	Other 38 Economies Studied
Full sample^a							
Log (Urban population, 2000)	−0.156*** (0.0291)	−0.138*** (0.0154)	−0.206*** (0.0235)	−0.301*** (0.0448)	0.0406 (0.0377)	−0.0326 (0.0653)	−0.181*** (0.0471)
Log (Urban density, 2000)	−0.225** (0.0852)	−0.322*** (0.0338)	0.105 (0.0699)	−0.237*** (0.0668)	0.120 (0.160)	−0.0872 (0.123)	−0.374*** (0.0843)
Log (Rainfall, 2000)	0.0992* (0.0505)	0.00470 (0.0484)	0.218*** (0.0783)	−0.162 (0.289)	0.208 (0.331)	−0.0865 (0.150)	0.0103 (0.196)
Max temperature, 2000	0.0171 (0.0119)	0.0218** (0.00983)	0.0261** (0.0115)	−0.0209 (0.154)	−0.0416 (0.0277)	−0.0612 (0.0929)	0.00415 (0.0448)
Minimum temperature, 2000	0.0102* (0.00501)	0.0161*** (0.00263)	−0.0441*** (0.00994)	−0.0479 (0.0559)	0.00959 (0.0194)	−0.00245 (0.0940)	0.00405 (0.00665)
Log (Distance to nearest port)	−0.115** (0.0453)	−0.0734 (0.0486)	−0.0751 (0.0992)	0.412 (0.249)	−0.174 (0.181)	−1.023 (1.222)	0.116 (0.0896)
Log (Distance to nearest port) ²	0.0161* (0.00821)	0.00717 (0.00583)	0.00830 (0.0120)	−0.0848* (0.0435)	0.0130 (0.0344)	0.0967 (0.105)	−0.0241* (0.0138)
Constant	3.777*** (0.681)	4.559*** (0.437)	1.922*** (0.727)	8.073* (4.315)	−0.207 (1.561)	6.169** (2.562)	5.544*** (1.821)
Observations	1,421	671	318	87	67	63	215
R-squared	0.445	0.395	0.374	0.491	0.203	0.115	0.727
Adjusted R-squared	0.431	0.388	0.360	0.446	0.109	0.00209	0.684
Cities with population > 100,000							
Log (Urban population, 2000)	−0.0569** (0.0222)	−0.0481*** (0.0156)	−0.120*** (0.0216)	−0.183** (0.0691)	0.0148 (0.0399)	0.0288 (0.0720)	−0.0304 (0.0290)
Log (Urban density, 2000)	−0.187** (0.0659)	−0.259*** (0.0366)	0.0176 (0.0732)	−0.162* (0.0893)	0.166 (0.116)	0.0153 (0.114)	−0.365** (0.149)
Log (Rainfall, 2000)	0.0716* (0.0378)	−0.0365 (0.0550)	0.0981 (0.0779)	0.0963 (0.337)	−0.0557 (0.195)	−0.142 (0.133)	0.136 (0.121)
Max temperature, 2000	0.0122 (0.0111)	0.0137 (0.0101)	0.0101 (0.0108)	−0.215 (0.132)	−0.0226 (0.0232)	−0.0534 (0.102)	−0.0229 (0.0276)
Minimum temperature, 2000	0.00893* (0.00475)	0.0156*** (0.00292)	−0.0359*** (0.0102)	−0.00312 (0.0601)	0.00404 (0.0149)	−0.000994 (0.0976)	0.0314*** (0.0103)
Log (Distance to nearest port)	−0.111*** (0.0299)	−0.0578 (0.0441)	−0.00788 (0.0964)	0.396 (0.282)	−0.0752 (0.141)	−1.042 (1.234)	0.0357 (0.103)
Log (Distance to nearest port) ²	0.0157** (0.00612)	0.00535 (0.00543)	0.00239 (0.0116)	−0.0772 (0.0479)	0.00312 (0.0265)	0.108 (0.102)	−0.00680 (0.0153)
Constant	2.352*** (0.532)	3.154*** (0.470)	1.953*** (0.681)	10.75*** (3.743)	−0.521 (1.044)	3.828* (2.007)	3.728* (1.909)
Observations	1,025	457	262	62	47	50	147
R-squared	0.298	0.271	0.237	0.288	0.195	0.163	0.565
Adjusted R-squared	0.279	0.260	0.216	0.196	0.0504	0.0232	0.488

^a The sample excludes cities with annual average growth rates greater than 0.2.

Notes: Dependent variable is log of urban population in 2016 divided by urban population in 2000. All 43 economies and the other 38 economies regressions include country fixed effects with standard errors clustered by country. Country specific regressions estimate robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Source: Author's estimates.

Climate characteristics are found to influence the choices people make in terms of which city to move to in India and the PRC, whose cities could offer a variety of climatic options. In particular, cities with more rainfall have grown faster in India when the full sample is studied. This is probably because cities with more rainfall face lesser shortage in water resources, which poses a challenge to the country in general. However, the importance of rainfall drops when the analysis focuses on cities with population greater than 100,000 in 2000. Temperature, especially the minimum temperature, affected city expansion in both countries, but in opposite ways. Regardless of which sample we examine, the results suggest that urban migrants in India favored cities with lower minimum temperature and those in the PRC preferred higher minimum temperatures.

The economic geography theory highlights the importance of geography to agglomeration and spatial distribution of economic activities. For example, a shorter distance to a seaport means being closer to international markets, which could promote agglomeration of producers and consumers, and thus city growth. Our estimations suggest this hypothesis generally holds true for the region as a whole and in India, Japan, Pakistan, and the PRC, although the country-specific estimates are not statistically precise. The positive estimates for the quadratic term of the distance imply that the advantage of this proximity decreases as the distance increases. Indonesia and the group of 38 economies have opposite coefficient estimates for the distance variables. This may be due to that fact that some of these countries like Indonesia, Philippines, and Viet Nam have many seaports (31 out of the country's 93 natural cities have seaports), so the advantage of access to international markets is not captured by the distance variables.

Generally speaking, we found that cities with lower populations, lower density, better climatic conditions, and in close proximity to a seaport grew faster from 2000 to 2016. However, the impacts of these factors varied greatly, and were even the opposite, when comparing different countries. The results should therefore not be interpreted as being universal, although they appear informative and relevant. More research on city growth in the region, particularly that covering additional driving factors such as a city's demographics, human capital, and market access, should be fruitful.

D. Emergence of City Clusters

Over time, many natural cities have not only expanded in physical area and in population, but also become spatially connected with one another. We refer to these urban agglomerations, each encompassing two or more connected natural cities, as city clusters. Not only are the natural cities within a city cluster connected by transport infrastructure, but the former farmland along the transport arteries was also developed for nonagricultural use, which was well captured by the NTL data.¹²

Some city clusters may be closer to the concept of metropolitan cities because the distance between the constituent natural cities is not too far for people to commute on a daily basis. Others may comprise natural cities that are beyond feasible commuting distance of each other, but the NTL patterns suggest the cities are economically intertwined. Since commuting distances can be negated by the advance of technology, and also because there are important commonalities between metropolitan cities and city clusters, we do not further distinguish between the two in this study.

¹² Given the way we identified natural cities, some of them could have in fact been city clusters in 1992. Such examples include Guangzhou, Osaka, Seoul, and Tokyo, which contain multiple administrative cities that could have been independent natural cities if traced further back. However, these were rare cases, especially in the developing world at that time.

City clusters are relevant and important because they represent the most vibrant urban areas in their respective countries and in the region. Hosting vast numbers of people and firms that are hugely diverse, city clusters could generate colossal flows of goods, services, and ideas. The agglomeration benefits of individual cities could be enhanced and amplified through the formation of city clusters. On the other hand, making city clusters competitive and livable poses considerable challenges. To connect people across the cities, more reliable infrastructure networks are needed, and this requires concerted efforts in planning, financing, and managing infrastructure. Massive coordination and cooperation requirements also extend to land-use regulation, provision of public services, managing traffic and environment, and market development. The task becomes formidable when considering that each constituent city is generally governed by an independent administrative authority, with some of them even belonging to different upper-level jurisdictions.

Our aim here is not to offer policy advice regarding city cluster governance, but to instead provide an up-to-date overview of city clusters in Asia and the Pacific, with a focus on those of enormous size. Characterizing such city clusters could provide an understanding of the scale and complexity in managing them.

We identified 1,527 separate natural cities in 1992. By 2016, while 1,038 natural cities continued to stand on their own, the other 489 natural cities had become connected to form 129 city clusters across Asia and the Pacific. Altogether, these city clusters had 977 million residents inhabiting a land area of 400,000 km², accounting for 65.8% of the population and 65.6% of the land area of all natural cities in 2016. Table 10 shows breakdowns by number of natural cities each city cluster comprises. The majority of these clusters (74) consisted of two natural cities, while there were 22 clusters encompassing three natural cities and 10 clusters encompassing four natural cities. The extreme case covering 53 natural cities is the cluster surrounding Shanghai in the PRC, which is often referred to as the Yangtze River Delta Area.

Table 10: Natural Cities Merged to become City Clusters, 2016

Number of Natural Cities in each City Cluster	Number of City Clusters	Total Number of Natural Cities Encompassed
2	74	148
3	22	66
4	10	40
5	4	20
6	5	30
7	1	7
8	1	8
9	3	27
10	3	30
11	2	22
12	2	24
14	1	14
53	1	53
Total	129	489
Number of spatially independent natural cities by 2016		1,038
Total natural cities		1,527

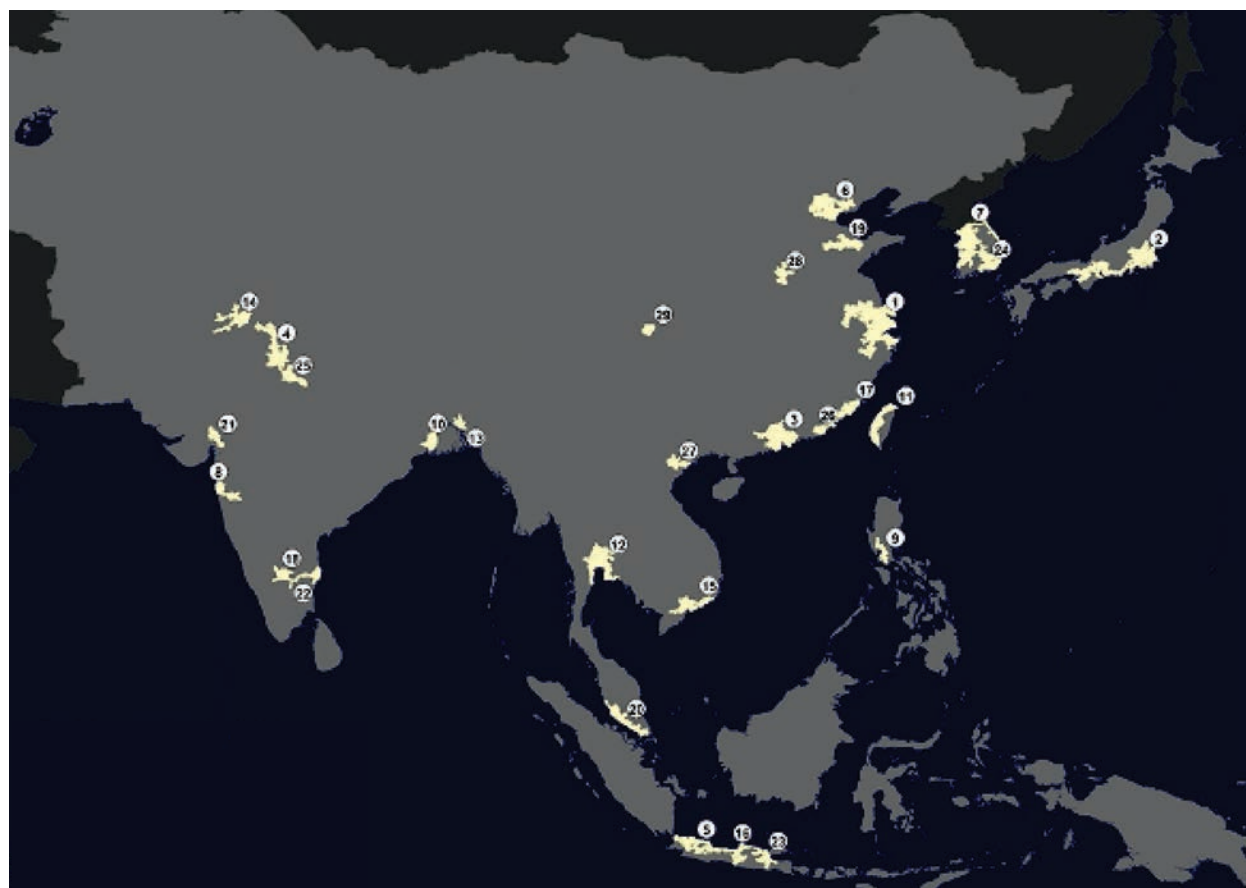
Source: Author's estimates.

Figure 12 shows the largest 29 city clusters across the region, each home to more than 10 million people in 2016 and numbered to indicate its ranking by population. Of the 29 city clusters, eight are in the PRC, seven in India, three in Indonesia, two each in the Republic of Korea and Viet Nam; there is one each in Bangladesh; Japan; Malaysia; the Philippines; Taipei, China; and Thailand. Our data show that there were 19 natural cities reaching a population of 10 million or more in 2016, of which 18 belonged to some city clusters. Karachi of Pakistan was the only natural city with more than 10 million population that did not form a city cluster with other cities.

We named each city cluster by joining the names of the largest two natural cities in the cluster. For instance, the cluster centered around Metro Manila included three other natural cities, namely Angeles, Lipa, and Tarlac. It is named Metro Manila–Angeles because Angeles is larger than Lipa and Tarlac.

Table 11 provides some descriptions of these gigantic urban agglomerations, that is, city clusters with population above 10 million in 2016, ranked by their population size. The largest one in 2016, both in

Figure 12: City Clusters with Populations of 10 Million or More, 2016



The boundaries, colors, denominations, and any other information shown on this map do not imply, on the part of the Asian Development Bank, any judgment on the legal status of any territory, or any endorsement or acceptance of such boundaries, colors, denominations, or information.

Note: The numbers indicate rankings of the city clusters by population in 2016. The corresponding cluster names could be found in Table 11.

Source: Author's estimates.

terms of population and land area, was Shanghai–Nanjing, which covered an area of 45,335 km² and had 91.5 million inhabitants. If treated as a country, this cluster would rank 16th in the world for population, sitting between Viet Nam and Germany. The clusters of Tokyo–Osaka and Guangzhou–Huizhou ranked second and third for both area and population. Dhaka–Sabhar was the smallest cluster when measured by land area, despite the fact that it is three times the size of Singapore.

The sixth column in Table 11 indicates the number of natural cities each cluster contains. Comparing this with Table 10, we note that clusters spanning more than 6 natural cities are all included in the list. A few clusters—including Tokyo–Osaka; capital city of Taipei, China–Puli; and Metro Manila–Angeles—are big, even though they contain only two or three natural cities. This is because they consist of large natural cities, such as Tokyo; Osaka; capital city of Taipei, China; and Metro Manila, which were already clusters of smaller cities in 1992.

Table 11: Largest City Clusters with Population above 10 Million, 2016

Rank	Economy	City Cluster	Total Population (million)	Total Area (km ²)	Number of Natural Cities	Number of Level-1 Admin Div	Type
1	PRC	Shanghai–Nanjing	91.5	45,335	53	4	II
2	Japan	Tokyo–Osaka	73.3	27,798	2	20	III-2
3	PRC	Guangzhou–Huizhou	52.3	19,465	12	4	III-2
4	India	Delhi–Chandigarh	43.4	15,379	10	6	I
5	Indonesia	Jakarta–Bandung	38.6	6,981	11	3	II
6	PRC	Beijing–Tianjin	37.8	17,023	12	4	III-2
7	Korea, Rep. of	Seoul–Taejon	32.2	18,330	9	11	I
8	India	Mumbai–Pune	30.8	5,514	4	2	II
9	Philippines	Metro Manila–Angeles	26.5	3,313	4	4	I
10	India	Kolkata–Habra	24.7	4,769	3	1	I
11	Taipei, China	Capital city of Taipei, China–Puli	21.7	12,135	2	20	I
12	Thailand	Bangkok–Nakhon Pathom	20.7	16,864	10	5	I
13	Bangladesh	Dhaka–Sabhar	18.1	2,128	3	2	I
14	Pakistan	Lahore–Faisalabad	16.5	4,814	9	1	I
15	Viet Nam	Ho Chi Minh–My Tho	16.0	8,768	5	4	I
16	Indonesia	Semarang–Cirebon	15.0	4,121	14	4	III-m
17	PRC	Quanzhou–Xiamen	14.3	5,752	9	1	III-2
18	India	Bangalore–Tumkur	13.7	4,779	4	2	I
19	PRC	Jinan–Zibo	12.5	7,516	11	1	III-m
20	Malaysia	Kuala Lumpur–Johor Bahru	12.4	8,922	7	5	I
21	India	Ahmadabad–Vadodara	12.3	3,954	4	1	I
22	India	Chennai–Vellore	12.2	3,538	3	1	I
23	Indonesia	Surabaya–Malang	11.9	3,867	8	2	I
24	Korea, Rep. of	Pusan–Taegu	11.6	8,925	6	7	I
25	India	Agra–Aligarh	11.4	6,211	6	1	I
26	PRC	Shantou–Chaoyang	11.2	3,310	6	2	I
27	Viet Nam	Ha Noi–Hai Phong	10.3	3,839	3	4	I
28	PRC	Zhengzhou–Xinxiang	10.3	4,758	10	1	I
29	PRC	Chengdu–Xindu	10.0	2,967	6	1	I

Admin Divs = administrative divisions, km² = square kilometer, PRC = People's Republic of China.

Note: In the last column, “III-2” represents type III city clusters with two leading cities, while “III-m” represents city clusters with more than two leading cities.

Source: Author's estimates.

The seventh column in Table 11 shows the number of level-1 administrative divisions involved in each cluster. Level-1 administrative division refers to the subnational level right below the national government. It is state in some countries such as India and province in others such as the PRC. The number of administrative divisions exceeded the number of natural cities for some clusters because their natural cities sprawled across multiple administrative boundaries. The counts show that over two-thirds of the clusters involved two or more level-1 administrative divisions, and more than half involved three or more divisions. This suggests that intergovernmental coordination at the central government level is often necessary in order to improve efficiencies and reduce negative consequences of city clusters.

The last column in Table 11 classifies the clusters by the degree of dominance of the largest natural city in the cluster. Specifically, type I clusters have one dominant city with the population of the second-largest city less than 5 million and half that of the dominant city. In type II clusters, the largest city is still dominant with the second largest below half of its size. However, the second-largest city is big with a population ranging between 5 million and 10 million. Type III clusters feature multiple leading cities, whose populations are all greater than 10 million, or where no one city exceeds the next by 100% of the population.

The majority of the clusters listed in Table 11 belonged to type I. For instance, the Delhi natural city had about 33 million inhabitants, while the next city, Chandigarh, had 4.2 million. Three clusters in the list were classified as type II. In the Shanghai–Nanjing cluster, Shanghai was clearly the leading city with population over 24 million. Nanjing and Hangzhou, the second- and third-largest cities in the cluster each having a population of over 6 million, serve as subcenters of the cluster. Mumbai–Pune and Jakarta–Bandung were the other two cases, wherein Mumbai and Jakarta functioned as the primary leading cities in their respective clusters, while Pune and Bandung were the secondary ones. Six clusters were considered to be type III, of which Beijing–Tianjin, Guangzhou–Huizhou, Quanzhou–Xiamen, and Tokyo–Osaka all had two leading cities, while Jinan–Zibo and Semarang–Cirebon had multiple leading cities.¹³

While our classification of cluster types looks slightly simplistic, it could shed light on the structure of the clusters. This structure is likely to evolve over time and could be correlated with the functional and market division among cities within each cluster. Design of optimal governance across a cluster may take into account such structural characteristics.

VI. CONCLUSION

Through constructing and analyzing a unique dataset of more than 1,500 cities in Asia and the Pacific, we have presented some stylized facts about the status and dynamics of urbanization across the region from 1992 to 2016. Our investigations focused on three levels— region and country, urban systems, and individual cities—to produce the following findings.

¹³ In Guangzhou–Huizhou cluster, the leading cities are Guangzhou and Shenzhen, each with a population greater than 10 million, despite that Shenzhen was considered part of the natural city of Guangzhou according to the NTL in 1992.

First, urbanized land increased from 0.9% of the total land area of the region in 1992 to 2.4% in 2016. The urbanization rates across the region, measured by the proportion of the population living in natural cities, increased from 27% in 2000 to 36% in 2016. The estimated urbanization rates are highly correlated with the UN estimates at the country level, although ours tended to be lower for several countries. While most countries have experienced a steady advance of urbanization, significant de-urbanization, in terms of both population and land area, was observed in Japan and Pakistan from the mid-2000s.

We found a close relationship between urbanization rates and GDP per capita in the region across time. However, the countries in Central and West Asia tended to have higher urbanization rates at their GDP per capita levels, while some Pacific DMCs appeared the opposite in 2000. The deviations were mitigated as the urbanization–GDP relationship became tighter by 2016. Simple regressions confirm that GDP per capita was a strong predictor for urbanization rates, though to a lesser extent for the proportion of urbanized land to total land. In addition, a country's total population was positively correlated, and total land was negatively correlated, with urbanization rates. As far as urban population growth is concerned, total population growth, GDP growth, and initial nonurban population all had positive impacts, while countries with more land and higher share of outputs from the agricultural or industrial sector tended to have slower urbanization rates.

Asian cities have generally become larger in both geographic area and number of inhabitants from 1992 onward. Consistent with the changes in city size distribution across the 43 economies, nearly three-quarters of the urban population was living in cities of more than 1 million inhabitants in 2016. The distribution of city density shifted to the left, mainly driven by India and the PRC, whose cities expanded much faster in terms of land area than in population. In Japan and Pakistan, there was an increase in very small cities and a decrease in medium-sized cities. Cities became much more densely populated in these two countries.

Primate cities play a unique role in a country's urban system. Among the 33 primate cities identified in this study, 27 are national capitals. Further analysis found that the size of primate city was positively correlated with its country's total population, urban population excluding itself and nonurban population. The primate city tended to be larger if it was a national capital or port city.

Zipf's law seems to hold for cities in India, Indonesia, and the PRC with a population greater than 100,000, but it fails when smaller cities in those countries are taken into account. It does not hold at all for Japan and Pakistan, with estimated coefficients suggesting their small and medium cities are too small, and their large cities are too large, compared to what Zipf's law predicts.

When individual cities were examined, our analysis showed that there were quite a few cities that had contracted in terms of population and/or land area, although the majority of cities had expanded on these two measures. The relative size of cities, measured by the size quintile, appeared highly volatile across the top five countries in the region. Gibrat's law was rejected in general for the region as a whole as well as for individual large countries except Pakistan. Despite that smaller cities have grown faster in most countries, except Japan and Pakistan, the distribution of urban population continued to be skewed toward larger cities.

At the city level, we found that cities with lower initial populations, lower density, better climatic conditions, and in close proximity to a seaport, tended to grow faster from 2000 to 2016. It is worth further studying to understand the causal relationship between these and other factors and city growth.

Finally, we documented that nearly one-third of the natural cities in our study had expanded to form 129 city clusters across Asia and the Pacific by 2016. Of these, 29 clusters were home to more

than 10 million people each and differed in terms of internal structure and composition. While representing the most vibrant areas of the region, these city clusters require significant improvement in governance and coordination to be more competitive and livable.

We find our natural city dataset consistent and credible, generating novel and sensible evidence. We hope that stakeholders will find some information presented in this paper relevant to policy discussions in areas such as urban development, infrastructure investment, and public service delivery. More in-depth analyses using this dataset in conjunction with other data sources are desirable to foster evidence-based policy making.

APPENDIX 1: DEVELOPING THE NATURAL CITY DATASET WITH NIGHTTIME LIGHTS LANDSCAN DATA

A. Defense Meteorological Satellite Program–Operational Linescan System Nighttime Lights Data

The United States Air Force Defense Meteorological Satellite Program (DMSP) initiated nighttime visible imaging that recorded the intensity of Earth-based lights, with satellites installed with Operational Linescan System (OLS) sensors in 1970s. The data were archived in digital format from 1992 to 2013. As Table A1 shows, for several years, two satellites were flying simultaneously to collect imaging data. For this study, we utilized data from the following satellites: F10 for 1992, F12 for 1995, F15 for 2000, F16 for 2005, and F18 for 2010. When a year was covered by two satellites, the newer satellite was chosen because satellites tend to deteriorate over time, which affects the quality of information derived from them.

Table A1: Defense Meteorological Satellite Program–Operational Linescan System Satellites and Their Coverage Period

DMSP-OLS SATELLITES					
F10	F12	F14	F15	F16	F18
1992	–	–	–	–	–
1993	–	–	–	–	–
1994	1994	–	–	–	–
–	1995	–	–	–	–
–	1996	–	–	–	–
–	1997	1997	–	–	–
–	1998	1998	–	–	–
–	1999	1999	–	–	–
–	–	2000	2000	–	–
–	–	2001	2001	–	–
–	–	2002	2002	–	–
–	–	2003	2003	–	–
–	–	–	2004	2004	–
–	–	–	2005	2005	–
–	–	–	2006	2006	–
–	–	–	2007	2007	–
–	–	–	–	2008	–
–	–	–	–	2009	–
–	–	–	–	–	2010
–	–	–	–	–	2011
–	–	–	–	–	2012
–	–	–	–	–	2013

DMSP = Defense Meteorological Satellite Program, OLS = Operational Linescan System.
Source: National Oceanic and Atmospheric Administration, National Centers for Environmental Information.
<https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>.

Scientists from the National Geophysical Data Center of the National Oceanic and Atmospheric Administration developed an automatic algorithm to filter the nighttime light (NTL) observations and remove unwanted lights, noises, and cloud presence from the raw DMSP-OLS NTL data (Hsu et al. 2015). They have released several products to the public that contain annual global cloud-free nighttime lights.

The first product is the raw average visible lights, which show all visible band digital number values from 0 to 63 and areas with zero cloud-free observation, but with no further filtering done. Average visible imagery has a censored digital number value for urban core, but is not bottom-coded. Therefore, it will show areas that have dimmer lights, such as rural or suburban areas, but will have a more generalized urban core. The second product is the stable lights, which are screened average visible lights wherein background noise is set to zero and undesired properties, such as ephemeral lighting and presence of clouds, are removed. A third product is the radiance-calibrated nighttime light imagery, which are available for limited years only. These images are unsaturated on the urban core, but suffer from bottom-coding, which will likely cause loss of rural and suburban areas that have dimmer lights.

We used the raw average visible lights due to the completeness of data for all years that we analyzed, and because these lights can give an accurate extent of the urban areas.

DMSP-OLS NTL images suffer from significant blurring known as “blooming” or “overglowing.” According to Small, Pozzi, and Elvidge (2005), blooming is the result of the relatively coarse spatial resolution of the OLS sensor, the large overlap in the footprints of adjacent OLS pixels, and the accumulation of geolocation errors in the compositing process. With these blurring in the NTL images, the digital number values of pixels outside the actual illuminated areas are still positive. This is a major issue in terms of urban area delineation since it makes it hard to identify urban areas from nonurban areas.

To resolve the issue, several methodologies were developed, and these can be classified mainly into two types: (i) thresholding based (Small, Pozzi, and Elvidge. 2005; Abrahams, Oram, and Lozano-Garcia 2018), and (ii) classification based (Goldblatt et al. 2018). The thresholding method is done by applying frequency thresholds to the image to reduce the spatial exaggeration on the lighted area, while the classification method uses other auxiliary data such as daytime satellite images.

In this study, we employed the latest method developed in Abrahams, Oram, and Lozano-Garcia (2018), which is independent from auxiliary data and coded through MATLAB. The method deblurs the images by applying two filters. The first filter tries to invert the blurring process in a noise sensitive manner using a standard Weiner deconvolution. In this filter, it is assumed that the light was blurred via a symmetric Gaussian point-spread function. The second filter uses percent frequency of light detection (PCT) image wherein the pixel at which a light source is located will always be a local maximum in the PCT image. It is implied that when a pixel is not a local maximum in the PCT image, the light recorded for those pixels are considered erroneous. The authors showed that when a 20% PCT threshold is applied, most areas that are infrequently lit are removed and the remaining pixels approximate urban area light well. According to the assessment in Abrahams, Oram, and Lozano-Garcia (2018), the deblurred NTL data are comparable to the stable lights version of the National Oceanic and Atmospheric Administration products, but they have much less exaggeration on urban areas.

B. Visible Infrared Imaging Radiometer Suite Nighttime Lights Data

Nighttime lights from Visible Infrared Imaging Radiometer Suite (VIIRS) succeeded the widely used DMSP-OLS NTL data in 2012. This dataset has a finer resolution of 15 arc-seconds, and gives the actual radiance captured by the sensor. The monthly data are available from 2012 to the present day, and annual composite data are available for 2015 and 2016. We use the "vcm-orm-ntl" (VIIRS Cloud Mask - Outlier Removed - Nighttime Lights) version of the 2016 annual composite VIIRS NTL data, since it contains cloud-free average radiance values that have undergone an outlier removal process to filter out fires and other ephemeral lights, with background (nonlights) set to zero.

Unlike the DMSP-OLS NTL images, the VIIRS NTL data does not need deblurring and intercalibration, since it already has an onboard calibration and the movement of the VIIRS satellite prevents overlaps in the images, which mainly caused the blurs in the DMSP-OLS product.

C. Construction of Natural City Sample

We took the following steps to construct the natural city sample for this study, with the deblurred DMSP-OLS NTL images for data up to 2010 and VIIRS NTL images for 2016.

- (i) **Delineating extents of human settlements.** We implemented a practical definition of human settlements, that is, pixels with a digital number greater than zero are settlements, while those with a digital number equal to zero are not. The delineated polygons with small gaps between them (1 km for DMSP-OLS NTLs and 0.5 km for VIIRS NTLs) are joined together.
- (ii) **Identifying urban agglomerations (natural cities).** We used Global Rural Urban Mapping Project (GRUMP) settlement point data and matched the selected settlement points, which had populations above 100,000 in 2000, to the generated NTL polygons. These polygons, referred to as natural cities, may contain one or more GRUMP settlements as early as 1992.
- (iii) **Including major cities.** We included the major cities from the countries with no city meeting the criteria in (ii) and polygons greater than 100 square kilometers (km²) in 2000, although the associated administrative cities do not meet the criteria in (ii).
- (iv) **Estimating population of the natural cities.** We overlaid the natural cities to the grid population from LandScan, and aggregated the population count per polygon.

We performed further data validation on the following cases: (i) city polygons with an area less than 2 km²; (ii) natural cities with less than 100,000 population count for at least one of the analysis years; (iii) natural cities with extraordinary area growth or shrinking from 1992 to 2016; and (iv) natural cities with extraordinary population growth or shrinking from 2010 to 2016.

Some issues were detected and fixed including:

- (i) Some GRUMP settlement points may have been placed incorrectly and therefore the wrong polygons were tagged.

- (ii) When the settlement points did not fall in any NTL polygons, we assigned the nearest polygons to them. This is not always correct, and we fixed it by visually checking the luminosity, population agglomerations, and daytime satellite imagery.
- (iii) Several urban areas in 2016 included roads connecting to other settlements, which caused extraordinary area and population growth as opposed to 2010. We redefined the urban area in 2016 by cutting roads and the connected settlements from the city if the roads were obvious and/or the connected settlements were at least 20 km away from the city being analyzed.
- (iv) Some polygons with area greater than 100 km² were found to be oil fields with no administrative city located in them. We eliminated these from the sample.

We obtained a more logical dataset after conducting all these manual checks and revisions.

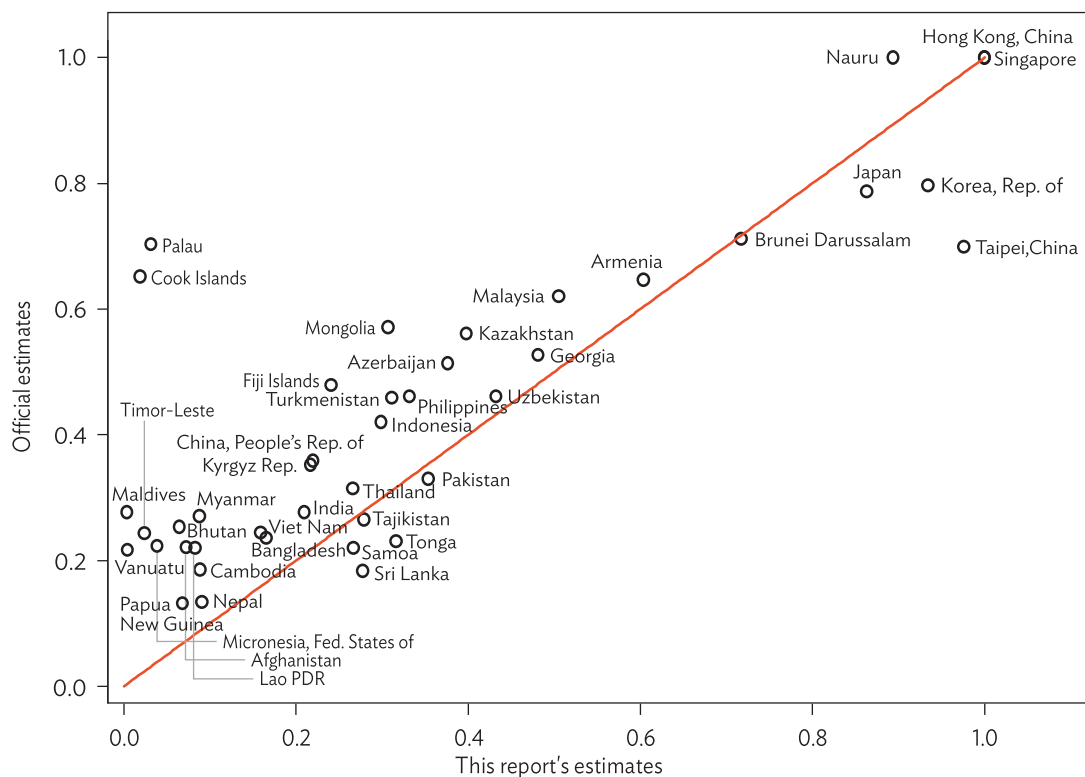
While extraordinary caution was exercised in processing the data, our dataset may not be free of errors. Measurement errors are likely to arise from imperfection of the NTL data and the grid population data. For instance, the quality of satellite imagery deteriorates with the length of the satellite's service years, which adversely affects the accuracy of the NTL data. The LandScan data is taken as given because validating LandScan's algorithm for the project purpose is beyond the capacity of the team.

APPENDIX 2: COMPARISON OF ESTIMATES OF URBANIZATION RATES

The criteria and methodologies used to define urban areas and estimate urban populations in this paper are distinct from those adopted by country governments to generate official statistics, which vary considerably across countries. Therefore, strictly speaking, our estimates of urbanization rates are not directly comparable with those derived from official data including the urbanization rates reported in the UN's World Urbanization Prospects (WUP), which rely on data provided by each country.

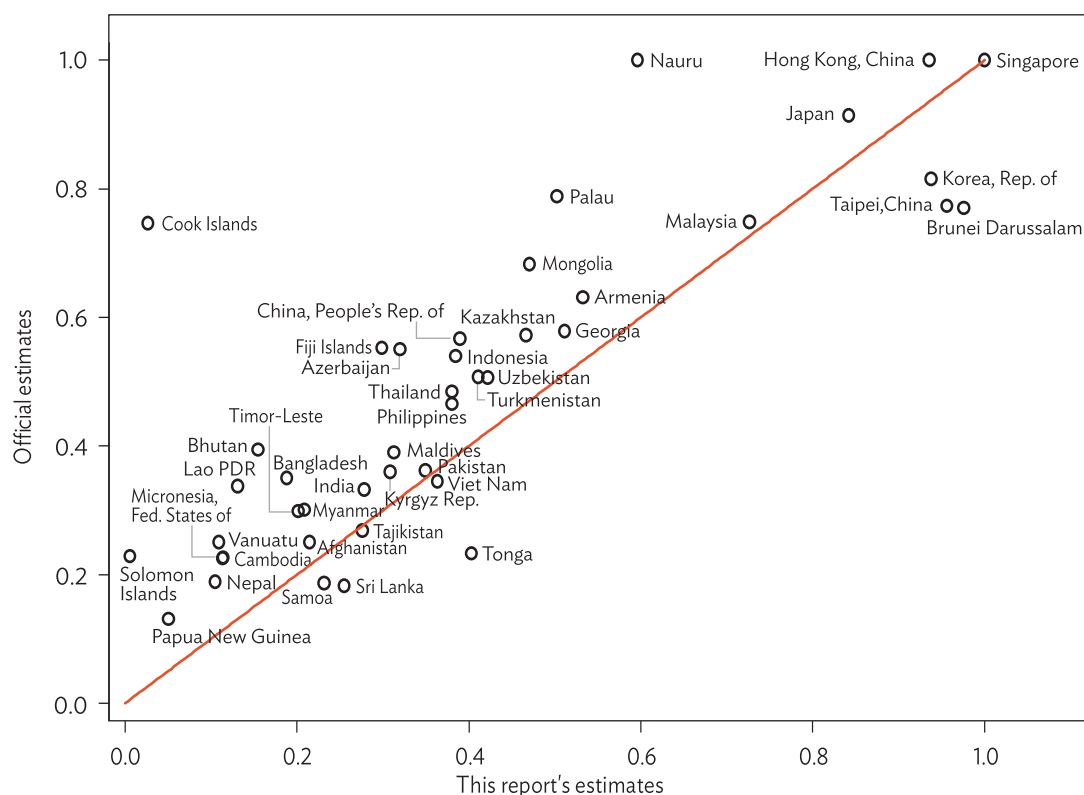
Having said this, it could be informative to observe the two estimates side by side. Figures A2.1 and A2.2 plot the urbanization rates estimated using our natural city data against WUP estimates for 2000 and 2016, respectively. First, the two estimates are positively correlated, with correlations equal to 0.83 for 2000 and 0.82 for 2016. Both are statistically significant at 1% level. On the other hand, our estimates tend to be lower than the WUP estimates for a number of countries. We suspect that one important reason is that one of the key criteria we used in constructing the dataset—only selecting the nighttime light-based settlements that covered the units in the Global Rural Urban Mapping Project with population greater than 100,000 in 2000—could have left many small cities out of the natural city dataset.

Figure A2.1: Estimates of Urbanization Rates based on Natural Cities versus World Urbanization Prospects Estimates, 2000



Source: Author's estimates.

Figure A2.2: Estimates of Urbanization Rates based on Natural Cities versus World Urbanization Prospects Estimates, 2016



Source: Author's estimates.

To further assess this possibility, we compared the total population sum across cities with inhabitants above 300,000 from our dataset and from WUP. The correlations increase substantially to 0.97 for 2000 and 0.98 for 2016. Moreover, the total populations calculated using natural city data generally exceed those based on the WUP data. For the regional total, the ratio between the two is 1.15 for 2000 and 1.21 for 2016. For the top five countries in the region, the ratios are 1.42 for India, 2.16 for Indonesia, 1.10 for Japan, 1.35 for Pakistan, and 0.96 for the PRC in 2016. This suggests that the natural city dataset captures relatively large urban agglomerations well, but may omit a host of small ones.

REFERENCES

- Abrahams, Alexei, Christopher Oram, and Nancy Lozano-Gracia. 2018. "Deblurring DMSP Nighttime Lights: A New Method Using Gaussian Filters and Frequencies of Illumination." *Remote Sensing of Environment* 210: 242–58.
- Ades, Alberto F., and Edward L. Glaeser. 1995. "Trade and Circuses: Explaining Urban Giants." *The Quarterly Journal of Economics* 110 (1): 195–227.
- Chauvin, Juan Pablo, Edward Glaeser, Yueran Ma, and Kristina Tobio. 2017. "What is Different about Urbanization in Rich and Poor Countries? Cities in Brazil, China, India and the United States." *Journal of Urban Economics* 98: 17–49.
- Chen, Xi, and William D. Nordhaus. 2011. "Using Luminosity Data as a Proxy for Economic Statistics." *Proceedings of the National Academy of Sciences* 108 (21): 8589–94.
- Colmer, Jonathan. 2016. "Urbanisation, Growth, and Development: Evidence from India." Unpublished Review Paper.
- da Mata, Daniel, U. Deichmann, J. Vernon Henderson, Somik V. Lall, and H.G. Wang. 2007. "Determinants of City Growth in Brazil." *Journal of Urban Economics* 62 (2): 252–72.
- Duranton, Gilles. 2007. "Urban Evolutions: The Fast, the Slow, and the Still." *American Economic Review* 97 (1): 197–221.
- Duranton, Gilles. 2008. "From Cities to Productivity and Growth in Developing Countries." *Canadian Journal of Economics/Revue canadienne d'économie* 41 (3): 689–736.
- Duranton, Gilles. 2016. "Determinants of City Growth in Colombia." *Papers in Regional Science* 95 (1): 101–31.
- Duranton, Gilles, and Diego Puga. 2001. "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products." *American Economic Review* 91 (5): 1454–77.
- Eeckhout, Jan. 2004. "Gibrat's Law for (All) Cities." *American Economic Review* 94 (5): 1429–51.
- Ellis, Peter, and Mark Roberts. 2015. *Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability*. Washington, DC: World Bank.
- Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." *The Quarterly Journal of Economics* 114 (3): 739–67.
- Gabaix, Xavier, and Rustam Ibragimov. 2011. "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents." *Journal of Business & Economic Statistics* 29 (1): 24–39.
- Goldblatt, Ran, Michelle F. Stuhlmacher, Beth Tellman, Nicholas Clinton, Gordon Hanson, Matei Georgescu, Chuyuan Wang et al. 2018. "Using Landsat and Nighttime Lights for Supervised Pixel-Based Image Classification of Urban Land Cover." *Remote Sensing of Environment* 205: 253–75.

Harari, Mariaflavia. 2016. "Cities in Bad Shape: Urban Geometry in India." *Processed, Wharton School of the University of Pennsylvania*.

Hasan, Rana, Yi Jiang, and Debolina Kundu. 2018. "The Growth of Indian Cities and "Good" Jobs: Evidence from the 2000s." In *Indian Policy Forum (Vol 2017/18)*, edited by S. Shah, B. Bosworth, and K. Muralidharan. SAGE.

Henderson, J. Vernon. 2005. "Urbanization and Growth." In *Handbook of Economic Growth Vol. 1*, edited by Philippe Aghion and Steven Durlauf. Elsevier.

Henderson, J. Vernon, John Quigley, and Edwin Lim. 2009. "Urbanization in China: Policy Issues and Options." Unpublished manuscript, Brown University.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028.

Hsu, Feng-Chi, Kimberly Baugh, Tilottama Ghosh, Mikhail Zhizhin, and Christopher Elvidge. 2015. "DMSP-OLS Radiance Calibrated Nighttime Lights Time Series with Intercalibration." *Remote Sensing* 7 (2): 1855–76.

Kohli, Harinder S., Ashok Sharma, and Anil Sood, eds. 2011. *Asia 2050: Realizing the Asian Century*. India: SAGE Publications.

Martin, Philippe, and Gianmarco IP Ottaviano. 1999. "Growing Locations: Industry Location in a Model of Endogenous Growth." *European Economic Review* 43 (2): 281–302.

Mellander, Charlotta, José Lobo, Kevin Stolarick, and Zara Matheson. 2015. "Night-Time Light Data: A Good Proxy Measure for Economic Activity?." *PloS One* 10 (10): e0139779.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *The Quarterly Journal of Economics* 129 (1): 151–213.

Rosen, Kenneth T., and Mitchel Resnick. 1980. "The Size Distribution of Cities: An Examination of the Pareto Law and Primacy." *Journal of Urban Economics* 8 (2): 165–86.

Schaffar, Alexandra, and Michel Dimou. 2012. "Rank-Size City Dynamics in China and India, 1981–2004." *Regional Studies* 46 (6): 707–21.

Small, Christopher, Francesca Pozzi, and Christopher D. Elvidge. 2005. "Spatial Analysis of Global Urban Extent from DMSP-OLS Night Lights." *Remote Sensing of Environment* 96 (3–4): 277–91.

Soo, Kwok Tong. 2005. "Zipf's Law for Cities: A Cross-Country Investigation." *Regional Science and Urban Economics* 35 (3): 239–63.

Soo, Kwok Tong. 2014. "Zipf, Gibrat and Geography: Evidence from China, India and Brazil." *Papers in Regional Science* 93 (1): 159–81.

Tewari, Meenu, Simon Alder, and Mark Roberts. 2017. "India's Urban and Spatial Development in the Post-Reform Period: An Empirical Analysis Based on Nightlight Data." Unpublished.

United Nations Department of Economic and Social Affairs, Population Division. 2018. "World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)." New York: United Nations.

Zhang, Qingling, and Karen C. Seto. 2011. "Mapping Urbanization Dynamics at Regional and Global Scales Using Multi-temporal DMSP/OLS Nighttime Light Data." *Remote Sensing of Environment* 115 (9): 2320–29.

Zhou, Naijun, Klaus Hubacek, and Mark Roberts. 2015. "Analysis of Spatial Patterns of Urban Growth across South Asia Using DMSP-OLS Nighttime Lights Data." *Applied Geography* 63: 292–303.

Spatial Dynamics and Driving Forces of Asian Cities

This study develops a new city-level panel dataset using satellite nighttime light imagery and grid population data for over 1,500 cities in 43 economies in Asia and the Pacific from 1992 to 2016. With the dataset, the author performs a variety of analyses and found novel evidence on inter-related topics including urbanization status and patterns, relations between urbanization and economic growth, evolution of urban systems, primate cities, the drivers of city growth, and emergence of city clusters. The analyses cover Asia and the Pacific as a whole as well as individually for five countries in the region—India, Indonesia, Japan, Pakistan, and the People’s Republic of China.

About the Asian Development Bank

ADB is committed to achieving a prosperous, inclusive, resilient, and sustainable Asia and the Pacific, while sustaining its efforts to eradicate extreme poverty. Established in 1966, it is owned by 68 members—49 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.



ASIAN DEVELOPMENT BANK

6 ADB Avenue, Mandaluyong City

1550 Metro Manila, Philippines

www.adb.org