

**E c o n o m i c &  
S o c i a l  
A f f a i r s**

# **World Urbanization Prospects The 2018 Revision**

**Methodology**



United Nations  
New York, 2018

This page is intentionally left blank

**Department of Economic and Social Affairs**  
Population Division

# **World Urbanization Prospects The 2018 Revision**

## **Methodology**



United Nations  
New York, 2018

The Department of Economic and Social Affairs of the United Nations Secretariat is a vital interface between global policies in the economic, social and environmental spheres and national action. The Department works in three main interlinked areas: (i) it compiles, generates and analyses a wide range of economic, social and environmental data and information on which States Members of the United Nations draw to review common problems and take stock of policy options; (ii) it facilitates the negotiations of Member States in many intergovernmental bodies on joint courses of action to address ongoing or emerging global challenges; and (iii) it advises interested Governments on the ways and means of translating policy frameworks developed in United Nations conferences and summits into programmes at the country level and, through technical assistance, helps build national capacities.

The Population Division of the Department of Economic and Social Affairs provides the international community with timely and accessible population data and analysis of population trends and development outcomes for all countries and areas of the world. To this end, the Division undertakes regular studies of population size and characteristics and of all three components of population change (fertility, mortality and migration). Founded in 1946, the Population Division provides substantive support on population and development issues to the United Nations General Assembly, the Economic and Social Council and the Commission on Population and Development. It also leads or participates in various interagency coordination mechanisms of the United Nations system. The work of the Division also contributes to strengthening the capacity of Member States to monitor population trends and to address current and emerging population issues.

## Notes

The designations employed in this report and the material presented in it do not imply the expression of any opinions whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The term “country” as used in this publication also refers, as appropriate, to territories or areas.

This report is available in electronic format on the Division’s website at [www.unpopulation.org](http://www.unpopulation.org). For further information about this report, please contact the Population Division, Department of Economic and Social Affairs, Two United Nations Plaza, DC2-1950, New York, 10017, USA; phone: +1 212-963-3209; e-mail: [population@un.org](mailto:population@un.org).

### Suggested citation:

United Nations, Department of Economic and Social Affairs, Population Division (2018). *World Urbanization Prospects: The 2018 Revision, Methodology*. Working Paper No. ESA/P/WP.252. New York: United Nations.

Official symbols of United Nations documents are composed of capital letters combined with numbers, as illustrated in the above citation.

## PREFACE

This report provides an overview of the methodology used to produce the official United Nations estimates and projections of urban and rural populations for 233 countries and areas of the world and for close to 1,900 urban settlements with 300,000 inhabitants or more in 2018, as published in *World Urbanization Prospects: The 2018 Revision*. This report and related products have been prepared by the Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. The data in this *Revision* are consistent with data on total population size from the latest revision of the United Nations global population estimates and projections (medium variant), as published in *World Population Prospects: The 2017 Revision*. For information on urbanization and the population size of cities, this *Revision* updates and supersedes all data published previously by the United Nations.

In preparing estimates and projections of the urban population, the United Nations relies on data from national sources, reflecting the definitions and criteria established by national authorities for both the level of urbanization, expressed as the percentage urban, and the population size of cities. For the *2018 Revision*, projected future levels of urbanization for countries and areas until 2050 and of population size for cities until 2035 were derived using the most recent data available combined with information from past trends. An analysis of historical trends confirmed the existence of an inverse relationship between the pace and the level of urbanization: with an increasing percentage urban, the pace of further increase tends to slow down. The method applied here to produce the projections assumes that future changes in the percentage urban of countries and areas will converge over time to the historical average observed in relevant populations at comparable levels of urbanization. Similarly, the pace of growth of city populations tends to decline with increasing population size. Projections of future city sizes were made based on this assumption, calibrated using historical data.

In preparing the *2018 Revision*, the Population Division relied on the collaboration of the Statistics Division of the Department of Economic and Social Affairs of the United Nations Secretariat, which produces the *United Nations Demographic Yearbook* and accompanying databases, official population statistics used in the preparation of the *2018 Revision*. The Population Division also acknowledges the use of information from the City Population website<sup>1</sup> and is grateful for support received directly from national statistical offices, many of which made available additional data and reports from recent censuses and surveys to assist in the preparation of the *2018 Revision*.

Within the Population Division, the *2018 Revision of World Urbanization Prospects* was prepared by a team led by François Pelletier that included Lina Bassarsky, Danan Gu, Neena Koshy, Mun Sim Lai, Igor Ribeiro, Cheryl Sawyer, Thomas Spoorenberg and Guangyu Zhang.

This publication is accessible on the website of the Population Division at [www.unpopulation.org](http://www.unpopulation.org). For further information about the *2018 Revision*, please contact the Population Division, DESA, Two United Nations Plaza, DC2-1950, New York 10017, USA; tel.: +1 212-963-3209; email: [population@un.org](mailto:population@un.org).

---

<sup>1</sup> Thomas Brinkhoff: City Population, <http://www.citypopulation.de>

## PROCEDURES TO ESTIMATE AND PROJECT THE POPULATION OF URBAN AREAS AND URBAN AGGLOMERATIONS

Within the *World Urbanization Prospects*, the estimation and projection of the urban population is based on observed changes in the proportion of the population living in urban areas. Consequently, the quality of the estimates and projections depends on the quality of the basic information enabling the calculation of the proportion urban. Such information normally consists of complete counts of both the total population in a country and the total population living in urban areas. Censuses or population registers are the most common sources of those counts. To be accurate, the proportion urban should be based on counts of the total and the urban population that achieve similar levels of coverage and that reflect correctly the division of territory into urban and rural areas. Because of the complexity and variety of situations in which the urbanization process occurs, it is not always straightforward to divide the inhabited territory into urban and rural areas. Indeed, the criteria used to identify urban areas vary from country to country and may not be consistent even between different data sources within a given country. Furthermore, as the process of urbanization proceeds, the number and extension of the areal units qualifying as urban generally expand. Therefore, keeping an urban versus rural division of territory constant over time would be misleading and would likely result in a major underestimation of the actual proportion of the population living in areas with urban characteristics.

In preparing estimates and projections of the urban population, the United Nations relies on the data produced by national sources, which reflect the definitions and criteria established by national authorities. Given the variety of situations in the world, it is currently not possible while using reported national data from sources such as censuses or population registers to adopt uniform criteria to distinguish urban areas from rural areas (see, for instance, United Nations, 1967 and 1969). For example, stipulating that any areal unit with at least 3,000 inhabitants should be considered urban may not be appropriate in populous countries such as India or Nigeria where rural settlements with none of the characteristics typical of urban areas often have large numbers of inhabitants. Overall, national statistical offices are often in the best position to establish the most appropriate criteria to characterize urban areas in their respective countries or areas.

No attempts have been made to impose consistency in definitions across countries. However, several efforts are underway at different institutions to produce globally comparable estimates of the urban population with uniform criteria to define urban areas based on satellite imagery of land cover or night-time lights. Nonetheless, to date, these approaches have not generated the long historical time series of urbanization estimates required for this report.

The urban and city estimates presented in this report are based on the definitions used for statistical purposes by the countries and areas themselves – except for cases lacking clear definitions, or historical changes that prevent reconstruction of consistent time series (e.g. Netherlands, Kenya). One hundred and twenty-one of the 233 countries or areas considered use administrative criteria to distinguish between urban and rural areas. Among these, 59 countries or areas use administrative designations as the sole criterion (table 1). In 108 cases, the criteria used to characterize urban areas include population size or population density, and in 37 cases such demographic characteristics are the sole criterion. However, the lower limit above which a settlement is considered to be urban varies considerably, ranging between 200 and 50,000 inhabitants. Economic characteristics were part of the criteria used to identify urban areas in 38 countries or areas. Criteria related to functional characteristics of urban areas, such as the existence of paved streets, water-supply systems, sewerage systems or electric lighting, were part of the definition of urban in 69 cases, but only in eight cases were such criteria used alone. Lastly, in 12 cases there was no definition or an unclear definition of what constitutes the urban environment and in 12 cases the entire population of a country or area was considered to be urban.

TABLE 1. NUMBER OF COUNTRIES OR AREAS ACCORDING TO THE CRITERIA USED IN DEFINING URBAN AREAS, 2018 REVISION

<i>Number and type of criteria</i>		<i>Number of countries or areas using criteria</i>	<i>Percentage (n=233)</i>	<i>Number of countries or areas using criteria in combination with additional criteria</i>	<i>Percentage (n=233)</i>
One criteria	Administrative	59	25.3	121	51.9
	Economic	—	—	38	16.3
	Population size/density	37	15.9	108	46.4
	Urban characteristics	8	3.4	69	29.6
Two criteria	Administrative and economic	—	—		
	Administrative and population size/density	17	7.3		
	Administrative and urban characteristics	20	8.6		
	Economic and population size/density	9	3.9		
	Economic and urban characteristics	—	—		
	Population size/density and urban characteristics	20	8.6		
Three criteria	Administrative, economic and population size/density	4	1.7		
	Administrative, economic and urban characteristics	—	—		
	Administrative, urban characteristics and population size/density	10	4.3		
	Economic, urban characteristics and population size/density	14	6.0		
Four criteria	Administrative, economic, population size/density and urban characteristics	11	4.7		
	Entire population is urban	12	5.2		
	No definition or unclear definition	12	5.2		
	Total number of countries or areas	233	100.0		

Despite the variety of criteria used to distinguish urban from rural areas and the resulting heterogeneity, no independent adjustment of national statistics was undertaken in this revision unless it was clear that the definitions used by a given country had changed over time in ways that would have led to inconsistencies. When applied, such adjustments typically eliminated the erratic peaks and troughs in urban growth resulting from changes in definition. Despite efforts to avoid inconsistencies within countries, it was not always possible to adjust the available data in ways that ensure consistency. In some cases, inconsistencies remain because the data needed to make the necessary adjustment were lacking. In cases where adjustment was possible, every effort was made to adjust earlier data so that they conformed to the most recent definition.

In the case of cities, population statistics are often reported in terms of the territory delimited by administrative boundaries that do not necessarily coincide with the extent of the urbanized territory as delimited by other standards. Thus, the “city proper” as defined by administrative boundaries may not include suburban areas where an important proportion of the population working or studying in the city resides. Furthermore, in some cases, two or more adjacent cities may be separately administered, although they might form jointly a single urbanized region. Alternatively, administrative boundaries of some cities

may cover large tracts of land primarily devoted to agriculture. Because of these problems it is advisable to base the measurement of a city's population on territorial boundaries that may differ from those established by administrative decisions. Since they are less affected by changes in administrative boundaries, two auxiliary concepts have been used to improve the comparability of measurements of city populations across countries and over time. The first is the concept of an urban agglomeration, which refers to the population contained within the contours of a contiguous territory inhabited at urban levels of residential density. The second is the concept of the metropolitan area, which includes both the contiguous territory inhabited at urban levels of residential density and additional surrounding areas with lower settlement density that are under the direct influence of the city (for example, through established transport networks, road linkages or commuting patterns).

In compiling information on city population size for this revision, the Population Division endeavoured to use data or estimates based on the concept of urban agglomeration or metropolitan area. When those data were not consistently available, population data that refer to the city as defined by its administrative boundaries were used. However, where administrative boundaries remain fixed for long periods of time, reliance on them can often result in an underestimation of the actual growth of a city with respect to both its territory and its population.

For a number of cities, the data available refer to two concepts: the city proper, as defined by administrative boundaries, and its metropolitan area. In those instances, the data referring to the metropolitan area were usually preferred because they are thought to approximate better the territory associated with the urban agglomeration than the data based on administrative boundaries. However, the population of the metropolitan area is also likely to be larger than that of the urban agglomeration associated with it, so an upward bias may have been introduced in specific cases. For many cities, an effort was made to ensure that the time series of population estimates derived from national sources conformed to the same definition over time. Adjustments were made when necessary to achieve internal consistency. In some cases, the availability of data determined that the criterion on which the population of a city is based should be changed. That was the case, for example, when data on a city in terms of the urban agglomeration were available for only one or two points in time, while a longer and more consistent series of data on the population of the city proper was readily available.

TABLE 2. NUMBER OF COUNTRIES OR AREAS ACCORDING TO THE CRITERIA USED IN DEFINING CITY POPULATIONS, 2018 REVISION

<i>Criterion</i>	<i>Number of countries or areas using criteria</i>	<i>Percentage (n=232)</i>	<i>Number of countries or areas using criteria in combination with additional criteria</i>	<i>Percentage (n=232)</i>
City proper .....	91	39.2	48	20.7
Urban agglomeration.....	74	31.9	45	19.4
Metropolitan area .....	13	5.6	19	8.2
Capital is urban agglomeration; other cities are city proper, urban agglomerations or metropolitan areas .....	27	11.6		
Capital is city proper; other cities are city proper, urban agglomerations or metropolitan areas .....	10	4.3		
Capital is metropolitan area; other cities are city proper, urban agglomerations or metropolitan areas .....	15	6.5		
Not defined.....	2	0.9	2	0.9
Total number of countries or areas	232	100.0		

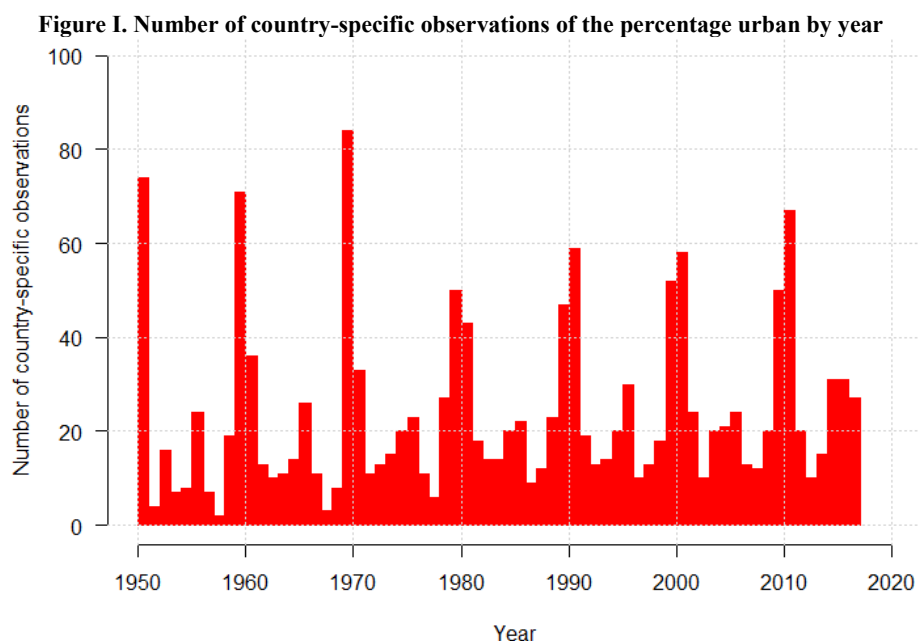


In the *2018 Revision*, the city data for 74 of the 232 countries or areas considered were based on the concept of urban agglomeration (table 2, above). In a further 27 countries or areas, data for the capital city were reported in terms of urban agglomeration, whereas data for other cities referred to city proper, urban agglomeration or metropolitan area. For an additional 91 countries or areas the city data available reflected a definition of a city proper. In 52 countries or areas different definitions were used for different cities. In several cases, adjustment of city data was carried out when information for a particular city had changed over time. Where possible, the urban agglomeration concept was used. However, when recent data were based on the concept of city proper and there was insufficient information to adjust the data to reflect the population in the urban agglomeration, a time series based on the city proper definition was used.

#### A. THE ESTIMATION OF URBAN INDICATORS OVER THE ESTIMATION PERIOD

In addition to variations in the definition of what constitutes an urban area, the data available for different countries or areas may also vary in terms of their time reference. Because census dates are not the same for all countries, estimates of the proportion urban or of city populations derived from census data typically refer to different points in time and are not directly comparable among countries. Similarly, there is no consistency among countries with respect to the reference dates of official estimates of urban or city populations. Consequently, in order to complete the 2018 revision of *World Urbanization Prospects*, estimates for specific points in time had to be derived. Interpolation or extrapolation based on the data available was used to produce estimates of the proportion urban or of city populations that referred to 1 July of the relevant calendar year, starting in 1950. The most recent mid-year estimate referred to the year that preceded the reference date of the most recent data available. From that point on, the projection procedure was used to complete the time series until 2050 for the proportion urban and until 2035 for city populations.

Among the 233 countries or areas considered in this *Revision*, the proportion urban was analysed for over 1,580 observations with 7 observations per country on average. As seen on figure I, the proportion urban is most often available in the years around the start of each decade, due to the availability of new census data.



For 216 countries or areas (93 per cent), the most recent data available refer to 2005 and later (table 3). For three countries (Democratic Republic of the Congo, Lebanon and Seychelles) the most recent data refer prior to 1990. When recent information on the proportion urban is available for a given country or area, it is more likely that the projections over the short-term approximate true trends.

TABLE 3. NUMBER OF COUNTRIES OR AREAS BY DATE OF MOST RECENT COUNTRY-SPECIFIC OBSERVATIONS OF THE PERCENTAGE URBAN

<i>Date of most recent information</i>	<i>Number of countries or areas</i>	<i>Percentage (n=232)</i>
Before 1990.....	3	1.3
1990-1994 .....	1	0.4
1995-1999 .....	1	0.4
2000-2004 .....	11	4.7
2005-2009 .....	26	11.2
2010-2014 .....	103	44.4
2015 and later.....	87	37.5
Total	232	100.0

The proportion of the population living in urban areas was estimated and projected by country or area for the period 1950-2050 in single-year intervals. Once values of the proportion urban at the national level were established for the 1950-2050 period, they were applied to the estimates and projections of the total national population of each country or area derived from *World Population Prospects: The 2017 Revision* (United Nations, 2017) so as to obtain the corresponding urban population for 1950 to 2050. At a later stage, country-level estimates and projections were aggregated to obtain figures corresponding to regions, sub-regions and the world as a whole.

Calculation of the proportion urban during the estimation period involved interpolation between recorded figures, and extrapolation back to 1 July 1950 when the earliest recorded figures referred to a later date. Such interpolation or extrapolation to 1950 is based on the *urban-rural ratio* ( $URR_t$ ), defined as the ratio of the urban to the rural population, that is:

$$URR_t = \frac{U_t}{R_t} \quad (1)$$

where  $U_t$  and  $R_t$  denote the size at time  $t$  of the urban and the rural populations, respectively. The urban-rural ratio at time  $t$  is directly related to the percentage urban ( $PU_t$ ) because:

$$PU_t = \frac{U_t}{(R_t + U_t)} = \frac{URR_t}{(1 + URR_t)} \quad (2)$$

Let  ${}_n r_{ur_t}$  denote the growth rate<sup>1</sup> of the urban-rural ratio between times  $t$  and  $t+n$ :

$${}_n r_{ur_t} = \frac{1}{n} \times \ln \left( \frac{URR_{t+n}}{URR_t} \right) \quad (3)$$

<sup>1</sup> The term "growth rate" is used in this report interchangeably with the term "rate of change" that is neutral in respect to either growth or decline.

Then, substituting  $URR_t$  for its equivalent according to (1), it follows that:

$$\begin{aligned}
 {}_n r_{ur_t} &= \frac{1}{n} \times \left[ \ln \left( \frac{U_{t+n}}{R_{t+n}} \right) - \ln \left( \frac{U_t}{R_t} \right) \right] \\
 &= \frac{1}{n} \times \left[ \ln(U_{t+n}) - \ln(R_{t+n}) - \ln(U_t) + \ln(R_t) \right] \\
 &= \frac{1}{n} \times \left[ \ln \left( \frac{U_{t+n}}{U_t} \right) - \ln \left( \frac{R_{t+n}}{R_t} \right) \right] \\
 &= \left[ \frac{1}{n} \times \ln \left( \frac{U_{t+n}}{U_t} \right) \right] - \left[ \frac{1}{n} \times \ln \left( \frac{R_{t+n}}{R_t} \right) \right] \\
 &= {}_n u_t - {}_n r_t
 \end{aligned} \tag{4}$$

where  ${}_n u_t$  denotes the growth rate of the urban population between times  $t$  and  $t+n$ , and  ${}_n r_t$  is the growth rate of the rural population over the same interval. That is, the growth rate of the urban-rural ratio is equivalent to the difference between the growth rates of the urban and the rural populations. Therefore,  ${}_n r_{ur_t}$  is known as the urban-rural growth difference and is the basis for the interpolation and extrapolation of the proportion urban.

Assuming a constant growth rate, if  $T$  is any time point within the intercensal period ( $t, t+n$ ), it follows that:

$$URR_T = URR_t \times e^{({}_n r_{ur_t} \times (T-t))} \tag{5}$$

The same equation (based on the same assumption) can be applied to obtain extrapolated values of  $URR$  when  $T$  is outside the intercensal period and ( $t, t+n$ ) is the intercensal period closest to it.

The use of equation (5) for interpolation and extrapolation purposes relies on an assumption that  $rur$  remains constant during each intercensal period, and during the period from 1950 to the reference date of the first observation available. Once an estimate of  $URR_T$  is available, it can be converted to  $PU_T$  with equation (2).

## B. PROJECTION OF THE PROPORTION URBAN AT THE NATIONAL LEVEL

The United Nations has developed a parsimonious and straightforward method to project the proportion urban. The method was first used in the 1970s (United Nations, 1974 and 1980) and, although it has undergone some revisions since then, the basic approach has not changed. The method projects the most recently observed urban-rural growth difference by assuming that the proportion urban follows a logistic path that attains a maximum growth rate when the proportion urban reaches 50 per cent and whose asymptotic value is 100 per cent.

Normally, an extrapolation based on a simple logistic curve would imply that the urban-rural growth difference would remain constant over the projection period. Yet empirical evidence shows that the urban-rural growth difference declines as the proportion urban increases because the pool of potential rural-urban migrants contributing to urban growth decreases as a fraction of the urban population, while it increases as a fraction of the rural population. Consequently, a model for the evolution of the urban-rural growth difference was developed so that it would evolve over the projection period, passing from the last observed value to a world norm of a hypothetical urban-rural growth difference consistent with historical experience. The norm is expressed in terms of a hypothetical urban-rural growth difference, denoted by  $hrur$ . For the 2018 revision, the hypothetical urban-rural growth difference, was obtained by regressing the urban-rural growth difference during any given time interval on the percentage urban at the mid-point of the

corresponding time interval, for the 149 countries or areas with 2 million or more inhabitants in 2018. The resulting regression equation<sup>2</sup> estimated on 1,117 observations (with 968 observations covering the period 1950-2018 and 149 observations for the period 1800-1949) is as follows:

$${}_n hrur_t = 0.032561 - 0.0153693 \times PU_{(t+\frac{n}{2})} \quad (6)$$

where  $PU_{(t+\frac{n}{2})}$  is the proportion urban for the mid-point of the intercensal period between times  $t$  and  $t+n$ .

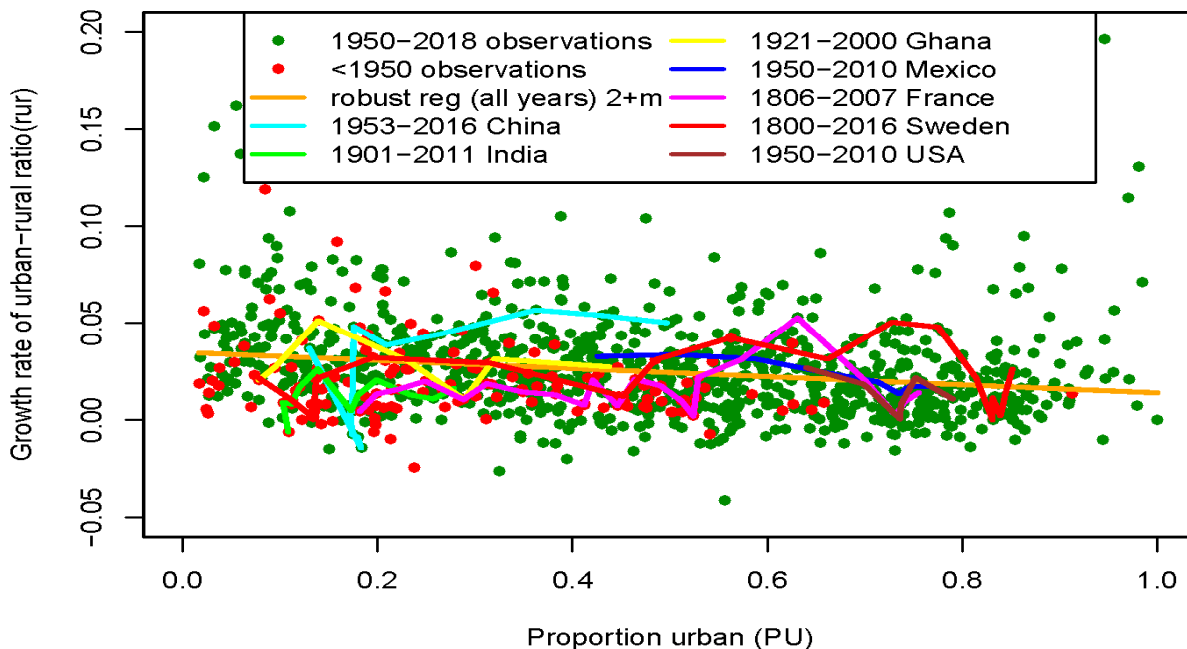
Figure II shows all the observations considered (with green dots representing observations since 1950 and red dots representing observations before 1950) together with the fitted regression line in orange. The empirical experience of several countries through the last one or two centuries illustrates how the observed urban-rural growth difference, denoted by  $rur$ , has changed over time as populations became more urban ( $PU$ ).

Equation (6) implies that, as the initial level of urbanization increases,  $hrur$  decreases. When the initial proportion urban is zero, an urban-rural growth difference of 0.032561 can be expected; when the proportion urban is 0.5, a value of  $hrur$  of 0.024876 can be expected. When the proportion urban reaches 1, no further growth in the percentage urban can occur; therefore, at that point the  $hrur$  is set equal to zero.

For the 82 countries or areas with less than 2 million inhabitants in 2018 (many of them are islands or enclaves), another regression equation<sup>3</sup> was estimated based on 563 observations (with 511 observations covering the period 1950-2018 and 52 observations for the period 1812-1949), and reads as follows:

$${}_n hrur_t = 0.0192479 - 0.0033084 \times PU_{(t+\frac{n}{2})} \quad (7)$$

**Figure II. Annual growth rate of the urban-rural ratio ( $rur$ ) by proportion urban ( $PU$ ) for countries or areas with 2 million or more inhabitants in 2018: selected country trajectories and fitted robust regression model**

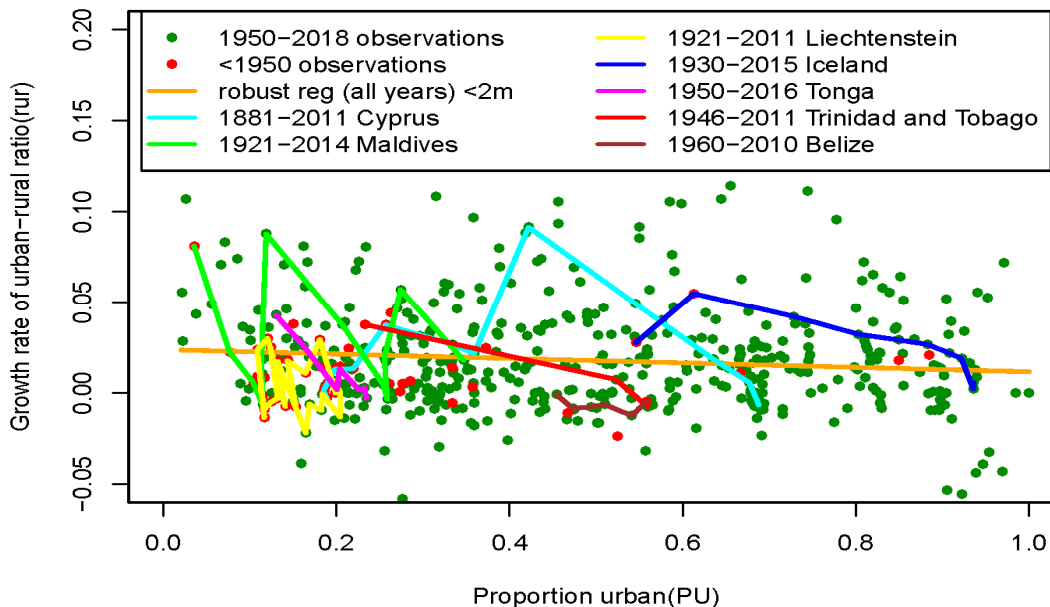


<sup>2</sup> Robust regression estimated using R package for Windows. Five extreme outliers associated with exceptional historical events where  $|rur| > 0.2$  were excluded from this analysis.  $R^2 = 0.02$ .

<sup>3</sup> Robust regression estimated using R package for Windows. Three extreme outliers associated with exceptional historical events where  $|rur| > 0.2$  were excluded from this analysis.  $R^2 = 0.001$ .

Figure III shows all the observations analysed and the fitted regression line (in orange). The trajectories for selected countries illustrate how *rur* has changed over time with greater urbanization (*PU*).

**Figure III. Annual growth rate of the urban-rural ratio by proportion urban for countries or areas with less than 2 million inhabitants in 2018: selected country trajectories and fitted robust regression model**



The projection of the proportion urban is based on the projection of the urban-rural growth differential. It is assumed that the most recently observed urban-rural growth difference in a given country converges to the hypothetical urban-rural growth difference, or world norm, over a period of 25 years. The projection is set to start at time  $t_0$ , the time at the most recently observed *PU*, and is carried forward until the year 2050.

The projection method has been modified slightly in two ways in this revision. First, because the world norm is a decreasing function of the percentage urban, the convergence of the projected urban-rural growth differential to the world norm was not assumed to be exactly linear. Rather, the projected urban-rural growth difference, denoted by  $rur^*$ , was obtained recursively as follows:

$$rur_{t+1}^* = rur_t^* + \left[ \frac{hrur(PU_t) - rur_t^*}{25 - (t - t_0)} \right] \quad \text{for all } t < t_0 + 25 \quad (8)$$

This ensures that the projected  $rur^*$  approaches the world norm smoothly in consonance with the contemporary level of the percentage urban, and that at  $t = t_0 + 25$ , the projected  $rur^*$  attains the world norm evaluated at the corresponding level of urbanization.

Second, for  $t > t_0 + 25$ , rather than keeping  $rur^*$  constant at the level of the world norm,  $rur^*$  was set to track the *hrur* that corresponds to the projected level of *PU*. In other words,

$$rur_t^* = hrur(PU_t) \quad \text{for } t_0 + 25 \leq t \leq \omega \quad (9)$$

where  $\omega$  is the last year of the projection, 2050.

Then, the urban-rural ratio was calculated as:

$$URR_{t+1} = URR_t \times e^{rur_t^*} \quad (10)$$

Each projected value of  $URR_t$  was then converted into a proportion urban  $PU_t$  using equation (2).

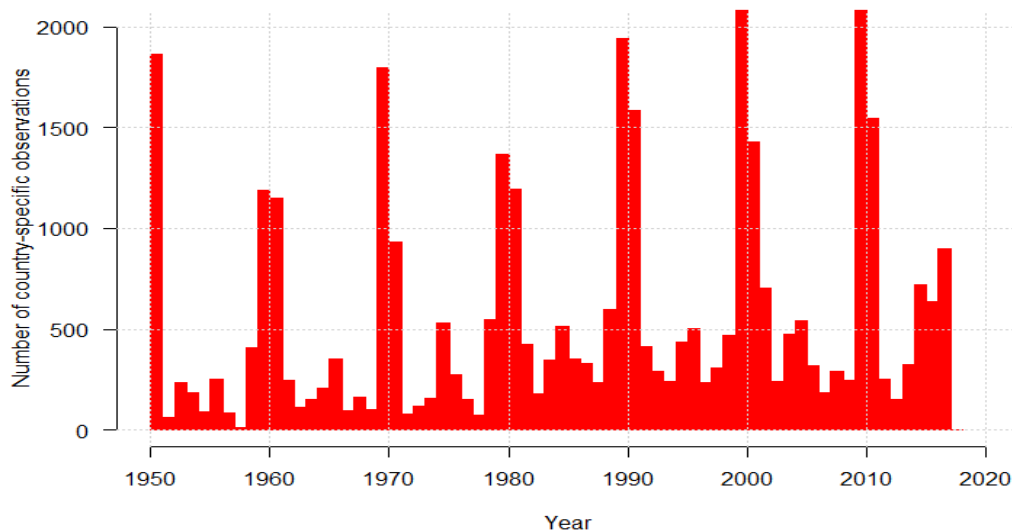
In order to derive the urban population at the national level, the proportion urban was multiplied by the total population of each country or area, obtained from the independent estimates and projections published in *World Population Prospects: The 2017 Revision* (United Nations, 2017). With respect to the estimates and projections of the urban population at the aggregate level, the urban populations of all countries or areas in the respective regions were cumulated. Lastly, regional totals were aggregated to derive estimates and projections at the world level.

### C. PAST ESTIMATES OF CITY POPULATIONS

Estimates and projections of the population of cities with an estimated population of 300,000 inhabitants or more in 2018 were calculated and published for the period 1950-2035. Overall, these comprise close to 1,900 urban settlements. However, in order to carry out a more comprehensive monitoring of population growth in cities, urban settlements reaching a population of 100,000 or more within the 1950-2018 period were also considered, provided that data on their population size was available from a census or population register<sup>4</sup>. Cities of 300,000 inhabitants or more in 2018 are illustrated on Map I by size class of urban settlement.

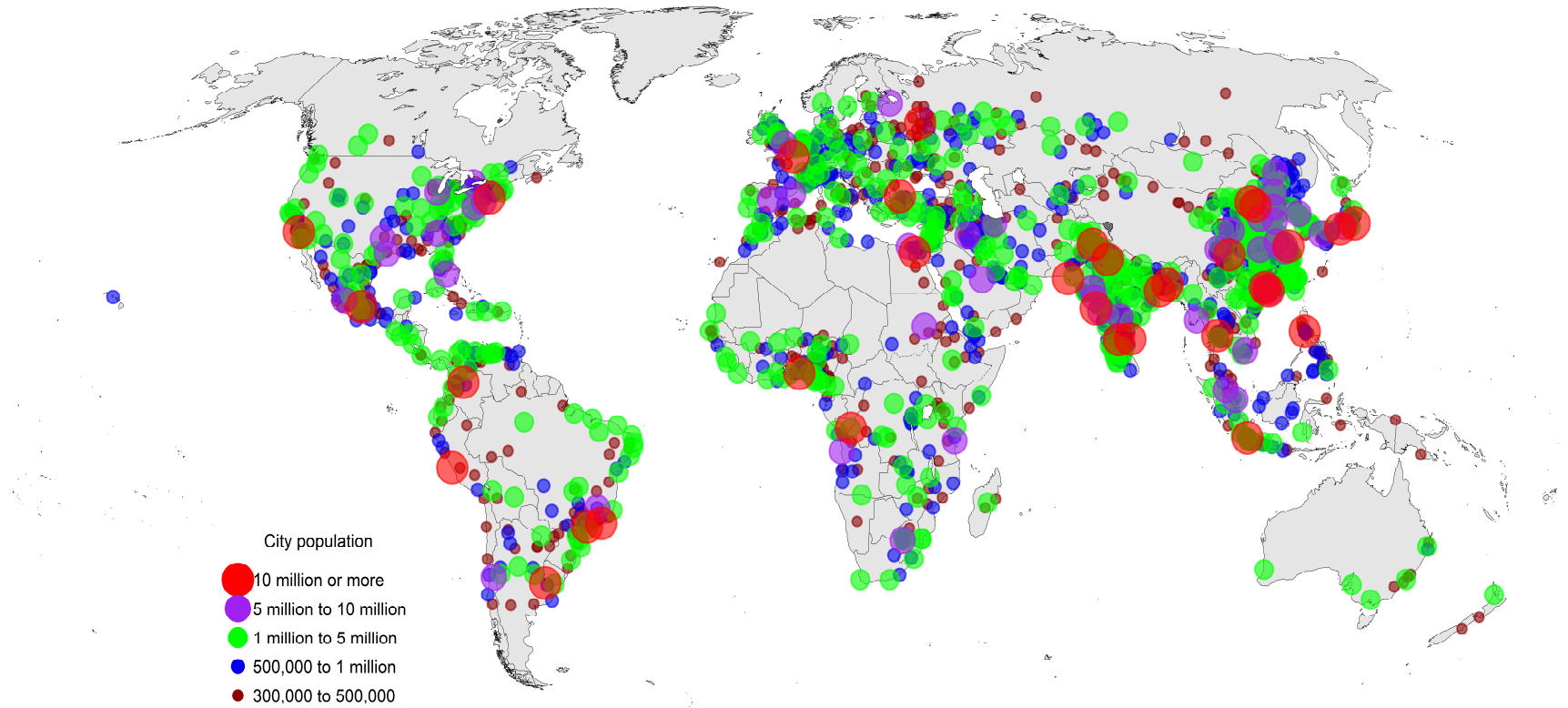
City population data were analysed for over 36,805 observations with on average 7 observations per city (and a standard deviation of 4 observations). As seen in figure IV, city population data are most often available in the years around the start of each decade, reflecting the importance of census data as the primary source of information on the size of city populations.

**Figure IV. Number of city-specific observations of city population by year**



<sup>4</sup> Furthermore, once a city has reached 100,000 inhabitants, its population size continues to be monitored by the Population Division even if the population subsequently fell below that threshold, provided that national statistical sources continue to report data. For the *2018 Revision*, 5,507 cities or urban agglomerations were considered, which also included 1,398 cities of less than 100,000 persons in 2018.

Map I. Population of cities by size class, 2018



Data source: World Urbanization Prospects: The 2018 Revision

The designations employed and the presentation of material on this map do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted line represents approximately the Line of Control in Jammu and Kashmir agreed upon by India and Pakistan. The final status of Jammu and Kashmir has not yet been agreed upon by the parties. Final boundary between the Republic of Sudan and the Republic of South Sudan has not yet been determined. A dispute exists between the Governments of Argentina and the United Kingdom of Great Britain and Northern Ireland concerning sovereignty over the Falkland Islands (Malvinas).

For 5,221 cities or urban agglomerations (95 per cent), the most recent data available referred to 2005 and later (table 4). Because countries take population censuses at different points in time, the actual dates of observation vary from city to city, although they are usually identical for cities within a particular country. Just as with the estimates of the proportion urban, while preparing estimates of city populations it was necessary to estimate the population size of all cities for the same reference dates in the past.

TABLE 4. CITIES OR URBAN AGGLOMERATIONS BY DATE OF MOST RECENT CITY-SPECIFIC OBSERVATIONS OF CITY POPULATION

<i>Date of most recent information</i>	<i>Number of cities or urban agglomerations</i>	<i>Percentage (n=5,507)</i>
Before 1990 .....	16	0.3
1990-1994 .....	11	0.2
1995-1999 .....	24	0.4
2000-2004 .....	235	4.3
2005-2009 .....	359	6.5
2010-2014 .....	2,688	48.8
2015 and later.....	2,174	39.5
Total	5,507	100.0

To estimate the population of cities on 1 July of any year in the past, the procedure used is similar to the one described above for the proportion urban. However, in this case, instead of using the urban-rural growth difference, the interpolation or extrapolation is based on the difference between the growth rate of a city and the growth rate of the population of the rest of the country ( $CRR_t$ ). Specifically, one considers the ratio of the city population at time  $t$ ,  $C_t$ , to the population of the rest of the country,  $RES_t$ :

$$CRR_t = \frac{C_t}{RES_t} \quad (11)$$

where

$$RES_t = P_t - C_t \quad (12)$$

and  $P_t$  is the total population of the country at time  $t$ . The growth rate of  $CRR$  between  $t$  and  $t+n$ , denoted by  ${}_nrcr_t$ , is

$${}_nrcr_t = \frac{1}{n} \times [\ln(CRR_{t+n}) - \ln(CRR_t)] \quad (13)$$

which is equivalent to

$${}_nrcr_t = {}_nc_t - {}_nres_t \quad (14)$$

where  ${}_nc_t$  is the growth rate of the city's population between  $t$  and  $t+n$ , and  ${}_nres_t$  is the growth rate of the rest of the country's population between  $t$  and  $t+n$ . Assuming constant growth rates, the value of  $CRR$  for any time  $T$  within the period  $(t, t+n)$  is given by:

$$CRR_T = CRR_t \times e^{({}_nrcr_t \times (T-t))} \quad (15)$$



The same equation can be applied to obtain extrapolated values of  $CRR$  when  $T$  is outside the intercensal period  $(t, t+n)$  and that period is the closest to  $T$ . The ratio of a city's population to the total population of a country at time  $T$ ,  $PC_T$ , is equivalent to:

$$PC_T = \frac{CRR_T}{1 + CRR_T} \quad (16)$$

To obtain interpolated or extrapolated estimates of city population size, the proportion  $PC_t$  was calculated for time  $T$  and multiplied by an independent estimate of the country's population to obtain the population of the city at time  $T$ . Such independent estimate was obtained from the country-level estimates published in *World Population Prospects: The 2017 Revision* (United Nations, 2017).

#### D. THE PROJECTION OF CITY POPULATIONS

The method used for projecting city populations is similar to that used for urban populations. The city growth rate over the most recent intercensal period is modified over the projection period so that it approaches linearly an expected value that is based on the city population and on the growth rate of the urban population as a whole. First, if  $(t, t+n)$  is the most recent intercensal period for a given country or, more specifically, the period between the two most recent sets of observed city populations, the growth rate of the city-urban ratio or the growth difference, denoted by  $rcu$  is calculated as:

$${}_nrcu_t = {}_nC_t - {}_nU_t \quad (17)$$

that is, the difference between the rate of population growth for the city and that for the total urban population. To project a city's population, it is necessary to establish future values of  $rcu$  using a model.

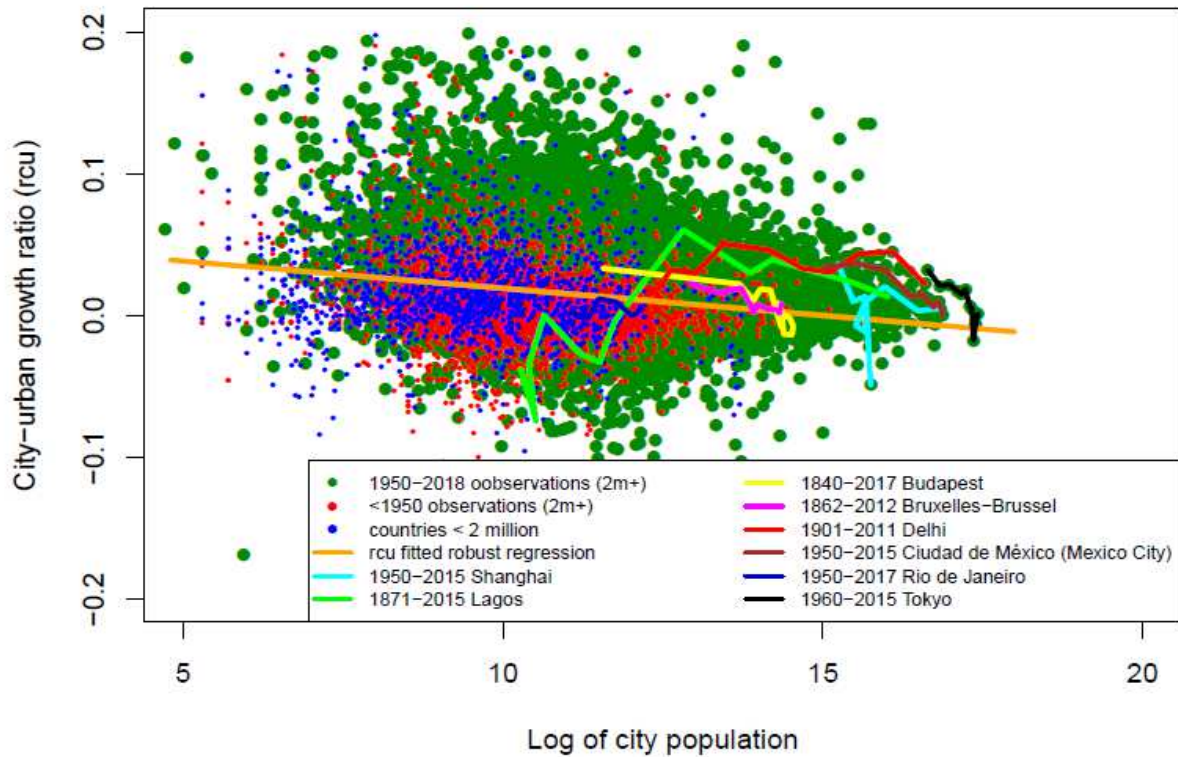
The model used to project  $rcu$  was developed by regressing the observed values of  $rcu$  for the most recent period for which data were available for each city on the logarithm of the city's population at the beginning of that period. The regression equation was fitted<sup>1</sup> to the data pertaining to 5,501 cities located in the 232 countries or areas analysed (41,132 observations including 36,307 observations covering the period 1950-2018 and 4,825 observations for the period 1815-1949).

Figure V shows all the observations for cities considered (green dots indicate observations since 1950 and red dots indicate observations before 1950 for countries or areas with at least 2 million inhabitants in 2018; blue dots indicate observations from countries or areas with less than 2 million inhabitants in 2018) together with the fitted regression line in orange. The trajectories of several cities illustrate how  $rcu$  has changed over time with the growth of the population of these cities.

---

<sup>1</sup> Robust regression estimated using R package. One hundred and nine extreme outliers associated to exceptional historical events with  $rcu > |0.20|$ , and 31 observations referring to city population less than 100 inhabitants were excluded from this analysis.  $R^2 = 0.0387$

Figure V. Annual growth rate of the city-urban ratio (*rcu*) by city population (log) for all countries or areas, selected city trajectories and fitted robust regression model



Although the correlation between the city-urban growth difference and the logarithm of the initial population size of each city is low ( $R^2=0.0476$ ), taking account of the influence of population size on city growth dampened the growth of the largest cities in a manner that seems plausible. The fitted model is the following:

$${}_nrcu_t = 0.0315719 - 0.0034566 \times \ln(C_t) - 0.0401943 \times {}_n u_t - 0.0343759 \times (\ln(C_t) \times {}_n u_t) \quad (18)$$

where  $C_t$  is the population of the city at time  $t$ ,  ${}_n u_t$  is the growth rate of total urban population, and  $(\ln(C_t) \times {}_n u_t)$  is the interaction term between these two variables. That is, at a given urban population growth rate, *rcu* decreases as the population of the city increases. Equation (18) was used recursively to calculate  ${}_nrcu_t$  over the projection period. Thus, if the projection period started at time  $T$  and the hypothetical growth rate of a city's population over the period  $(T-n, T)$  was denoted by  ${}_n hc_{T-n}$ , it was calculated as follows:

$${}_n hc_{T-n} = {}_n u_{T-n} + {}_nrcu_{T-n} \quad (19)$$

To deal with more specific constraints experienced by cities located in islands, enclaves and other small areas, regression equation (18) was also fitted<sup>1</sup> on the subset of the data relative to 208 cities located in the 82 countries or areas that had less than 2 million inhabitants in 2018 (1,492 observations including 1,225 observations covering the period 1950-2018 and 267 observations for the period 1815-1949), which yielded the following model:

<sup>1</sup> Robust regression estimated using R package for Windows. Nineteen extreme outliers associated to exceptional historical events with *rcu* > 0.20, and 7 observations referring to city population less than 100 inhabitants were excluded from this analysis.  $R^2=0.2626$

$${}_nrcu_t = 0.0250165 + 0.0027261 \times \ln(C_t) - 0.0623834 \times {}_nu_t - 0.2281567 \times (\ln(C_t) \times {}_nu_t) \quad (20)$$

where  $C_t$  is the population of the city at time  $t$ ,  ${}_nu_t$  is the growth rate of total urban population, and  $(\ln(C_t) \times {}_nu_t)$  is the interaction term between these two variables.

Regression equation (20) instead of (18) was used in equation (19) for cities located in countries or areas that had less than 2 million inhabitants in 2018.

Figure VI shows all the observations for cities considered (green dots represent observations since 1950 and red dots represent observations before 1950 for countries or areas with less than 2 million inhabitants in 2018) together with the fitted regression line in orange. In addition, the trajectories of several cities illustrate how  $rcu$  has changed over time with the growth of the population of these cities.

A city's growth was then projected according to equation (21). As in the approach for projecting the growth rates for the proportion urban (see equation 7), equation (21) includes weights to let the estimated growth rate converge to the hypothetical growth rate:

$${}_nc_T = (W_t \times {}_nc_{T-n}) + ((1-W_t) \times hc_{T-n}) \quad (21)$$

The weights  $W_t$  were calculated as:

$$W_t = \begin{cases} 1 - \frac{1}{d} \times (t - t_0) & (t - t_0) \leq d \\ 0 & (t - t_0) > d \end{cases} \quad (22)$$

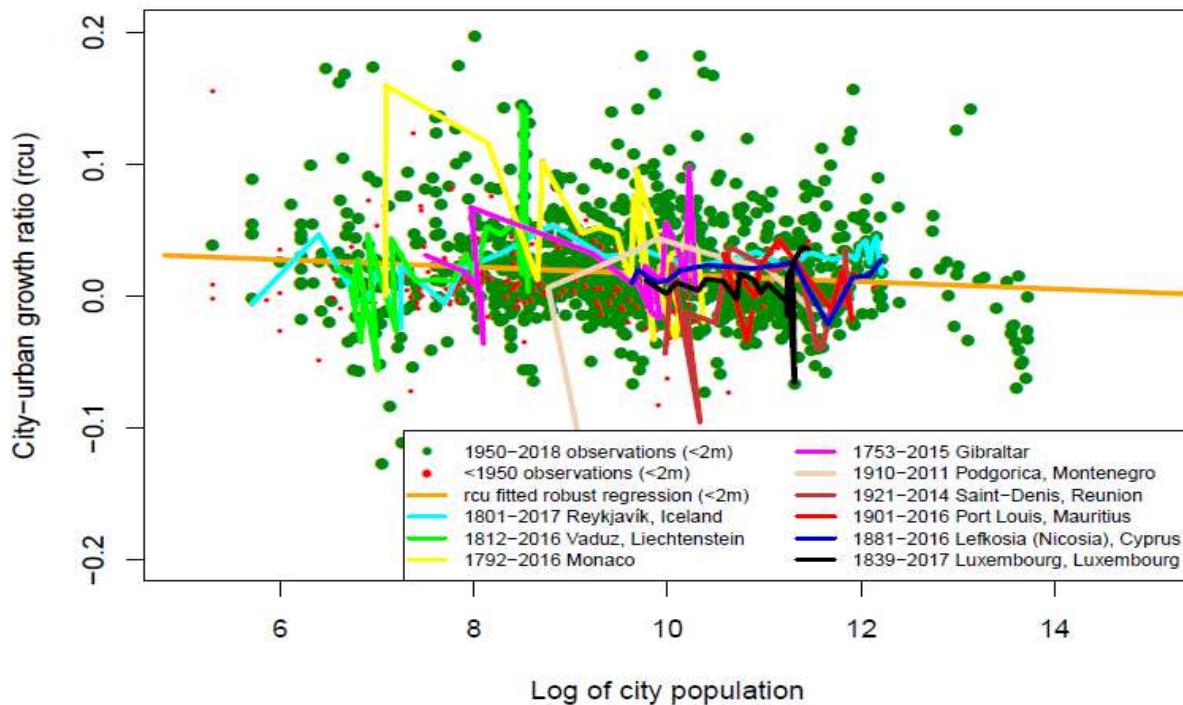
Projection calculations were carried out independently for each city within a country, but a further adjustment sometimes had to be made once the projected populations of all cities were available. If the aggregated projected values of the city populations of a country grew more rapidly than the total urban population of the country, a further dampening factor was imposed on the city growth rates. When this situation arose, the growth rate of each city was reduced by the following quantity:

$${}_n\delta_t = \frac{({}_nrtc_t - {}_nu_t) \times TC_t}{U_t} \quad (23)$$

where  $TC_t$  is the aggregated population of cities whose populations are being projected at time  $t$ ,  $U_t$  is the total urban population,  ${}_nrtc_t$  is the growth rate of the aggregated population of cities and  ${}_nu_t$  is the growth rate of the urban population. That is, the projected growth rate of the city was changed to:

$${}_nc_t^* = {}_nc_t - {}_n\delta_t \quad (24)$$

Figure VI. Annual growth rate of the city-urban ratio (*rcu*) by city population (log) for countries or areas with less than 2 million inhabitants in 2018, selected city trajectories and fitted robust regression model



This reduction assures that the total population of cities will not exceed the total urban population, while maintaining the differences in the growth rates among cities.

Adjustments were also made to the projected growth rates of cities when the future city-urban growth ratio (*rcu*) from equation (18) was less than zero; in such cases, *rcu* was set to zero. Similarly, if the hypothetical growth rate for a city (equation 19) was less than or equal to zero, the hypothetical growth rate was also set to zero. However, this second constraint was not enforced when the urban population was declining and the city population was growing (e.g., some Eastern European countries).

In all cases, equation (21) was used to project the city growth rate. However, in instances when both the urban population and the city population were declining, only the projected city growth was used (i.e.,  $n\hat{C}_t^*$  with  $W_t=1$  for all  $t$ ). In addition, the projection using equation (21) started from 2018 onward, except when the urban population and the city population were both growing; in those instances, the projection started from the last observed time point.

Lastly, it should be noted that for many countries, population figures were not compiled for all existing urban settlements. Therefore, in order to ensure greater consistency with the projected ratio between the sum of the urban settlements for which time-series of population estimates were compiled and the overall urban population, an additional step was undertaken. In that regard, the population of all other urban settlements for which estimates in a given country were unavailable or not compiled was projected as a “residual urban unit” as of 2018. This was done by applying the future annual growth rate of the urban population of each country to the estimated residual urban population, that is the difference between the total urban population and the sum of the cities for which city-level information was available.

## E. PRINCIPAL SYMBOLS USED

$C_t$	Population of a particular city at time $t$
${}^n c_t$	Growth rate of a city's population between times $t$ and $t+n$
${}^n \hat{c}_t^*$	Projected growth rate of a city between times $t$ and $t+n$ (adjusted to insure consistency with total urban population)
$CRR_t$	City-rest of the country ratio at time $t$ (ratio of a city's population to the population of the rest of the country)
$d$	Duration of convergence period
${}^n h c_t$	Hypothetical growth rate of a city's population between times $t$ and $t+n$
${}^n h r u r_t$	Hypothetical growth rate of the urban-rural ratio (hypothetical urban-rural growth difference) between time $t$ and $t+n$
$P_t$	Total population of a country at time $t$
$P C_t$	Proportion of a city's population to total population of the country at time $t$
$P U_t$	Percentage or proportion urban at time $t$
$R_t$	Rural population at time $t$
${}^n r_t$	Growth rate of the rural population between times $t$ and $t+n$
${}^n r c r_t$	Growth rate of CRR between times $t$ and $t+n$
${}^n r c u_t$	Growth rate of the city-urban ratio (the city-urban growth difference) between times $t$ and $t+n$
$RES_t$	Population of the rest of the country (total population minus a particular city) at time $t$
${}^n r e s_t$	Growth rate of the population of the rest of the country between times $t$ and $t+n$ (total population minus a particular city's population)
${}^n r t c_t$	Growth rate of the aggregated population of cities between times $t$ and $t+n$
${}^n r u r_t$	Growth rate of the urban-rural ratio (the urban-rural growth difference) between times $t$ and $t+n$
${}^n r u r_t^*$	Projected growth rate of the urban-rural ratio (the urban-rural growth difference) between times $t$ and $t+n$
$T C_t$	Aggregated population of cities at time $t$ ,
$U_t$	Urban population at time $t$
${}^n u_t$	Growth rate of the urban population between times $t$ and $t+n$
$URR_t$	Urban-rural ratio at time $t$ (ratio of the urban to the rural population)
$W_t$	Weight at time $t$
${}^n \delta_t$	Dampening factor for city growth rates between times $t$ and $t+n$
$t$	Index for time point
$n$	Length of time period

### General:

A subscript on the lower left means the length of an interval, here in years.

A subscript on the lower right stands for time.

Superscripts are used to assign additional attributes to a symbol.

## REFERENCES

United Nations, Department of Economic and Social Affairs, Population Division (2017). *World Population Prospects: The 2017 Revision*, DVD Edition.

\_\_\_\_\_ (1980). *Patterns of Urban and Rural Growth*. Population Studies, No. 68 (United Nations publications, Sales no E.79.XIII.9).

\_\_\_\_\_ (1974). *Methods for the Projection of Urban and Rural Population*. Population Studies, No. 55 (United Nations publications, Sales no E.74.XIII.3).

\_\_\_\_\_ (1969). *Growth of the World's Urban and Rural Population, 1920-2000*. Population Studies, No. 44 (United Nations publications, Sales no E.69.XIII.3).

\_\_\_\_\_ (1967). *Principles and Recommendations for the 1970 Population Censuses* (United Nations publications, Sales no E.67.XIII.3).