Forthcoming *Urban Empires: Cities as Global Rulers in the New Urban World*
(Edited by Edward Glaeser, Peter Nijkamp and Karima Kourtit)

The Size of Cities

Michael Batty*

Centre for Advanced Spatial Analysis (CASA)
University College London,
90 Tottenham Court Road,
London W1T 4TJ

m.batty@ucl.ac.uk

June 8, 2018

Abstract

In this chapter, we discuss ways of measuring the size of cities using population counts and building heights. We focus on the top world cities with respect to these measures and examine the shape of their distributions over time and space, time being from the early- to mid-20th century and space pertaining to different cities above a certain size threshold. We examine a data set of the biggest 589 cities globally produced by Nordpil in 2014 from UN Population data which is available in five year periods from 1950 and then a much larger data set of all 13,845 urban agglomerations over 50,000 population produced by the EU Joint Research Centre in 2016. These data throw up various definitional problems with respect to what constitutes the physical location of a city and because they are bounded by a threshold from below, they are inevitably incomplete. We then present a data set from the Skyscraper Center for 2018 of the world's highest buildings, over 200 meters in height, and from these 1,310 highest buildings, we explore how heights have increased during the 20th century but at a much slower rate than the growth of the largest cities. In all these data sets, we explore their shape in terms of the frequency of cities and buildings, showing that they follow remarkably consistent power laws in their upper tails which for the most part, is the entirety of the data sets used. We conclude by commenting on the fact that we need measures of size based on flows rather than locations, concluding that in thinking of cities in this way, the notion of size changes to one of connectivity and strength of relations.

* Tel 44-(0)7768 423 656; http://www.spatialcomplexity.info/ t@jmichaelbatty

In his book *Sapiens*, Yuval Harari (2016) paints a very convincing picture that as soon as cities emerged from their nomadic and agricultural past some 5000 years ago, they became organized into empires. To an extent until the advent of the modern era[†], cities were judged to be synonymous with the idea of empire. The number of such empires, however, has varied dramatically during this time with respect to population size and geographical extent, while many have been isolated from one another, often existing in parallel in very different regions of the world. Such empires have usually consisted of a hierarchy of towns of different sizes, often with a distinct capital city which provided the mechanisms and motive power for combining many rural hinterlands, their urban places serving as foci for the provision of services to a largely agrarian population. Only in the late modern era[†] has there been any sense in which empires have begun to merge, cooperate, and replace previous contact systems that were invariably based on competition for resources, predicated on warfare, family rivalries, and occasionally on political alliances.

The idea of empire is a quintessential principle of human organization in that it provides a framework for extending our individual economic capabilities through the spatial division of labor. Empires appear to be natural organizations of human power and intellect but for the last 5000 years, empires have risen and fallen with strong regularity, never quite mobilizing enough resources internally to continue holding themselves together (Tainter, 1990). But as technology has progressed and as we have perfected our organizational abilities more keenly, empires have become somewhat more robust as well as more flexible, while transforming their structure in response to increasingly global forces. We now stand at a threshold where we are about to, if we have not already, merged the world's great empires – North America, China and East Asia, possibly Europe. Until we are able to propel ourselves off the planet, there will thus be one great Empire, likely to dominate the rest of this century if not beyond, which will be characterized by the increasing fluidity of town and cities, as they merge into one another and capture ever wider global flows. In this sense then, a global empire which will largely consist of cities, is very different from a merger of nations or city states, and in this context, we will argue that this constitutes a truly global urban system.

In the wider context of the emergent global empire, this chapter explores what we know about the size of cities, as measured by their population, buildings, and contact systems. We will speculate that by the end of this century everybody will be living in cities, that all the world will be a 'city', or 'cities' or at least be 'urban' which is somewhat more expansive a concept (Batty, 2011). Two key forces are driving this transformation. First, there is globalization in that everyone is rapidly becoming connected to everyone else through the electronic skin that is being put in place of which the internet is currently the essential infrastructure. In fact, globalization began long before the digital age. As soon as cities and empires emerged, the notion of empire itself reflected a kind of local globalization, and since classical times, empires have both proliferated and become shorter in length, beginning to merge and bunch up together as cities have become more and more connected and as human interactions have become truly global. The second force involves diffusion. Cities are growing together, fusing as they both spread out and compact. New forms of planetary urbanization are emerging (Brenner and Schmid, 2012). McGee (2009) in exploring urbanization in China refers to the emergent urban form in places like the Shanghai-Suzhou-Nanjing and the Hong Kong-Shenzhen-Guangzhou-Zuhai-Macau regions, for example, as 'desakota'. This is a mixture of urban and rural with the

[†] The modern era dates from around 1500, the height of the Renaissance in Europe

functions becoming blurred as populations engage in both kinds of pursuit and as the mark of the big city in the form of high towers comes to dominate the countryside.

In this chapter, we will attempt to define the size of cities, beginning with some ideas about the extent to which empires define unique groupings of cities but using this mainly as a prelude to identifying the key features that we can use in measuring their size. Our thesis will be that as cities grow from individuals clustering together in locations where there has never been any settlement or where there is no settlement existing there, the lower bound on city size is always unclear. From the place where a lone explorer decides to settle – the archetypal log cabin on the frontier, to clusters of households forming a hamlet, thence a village that grows into a small market town, the definition of what constitutes a city is always problematic. Is the lower bound tens of households or hundreds or thousands? What is clear is that unless we define such a lower bound for a city size in terms of its population, there is no way we can measure the total number of cities. There is now the prospect of defining our own physical units from population data sets that define numbers of persons or households down to their individual address points (Arcaute et al., 2014) but this is only very recent and historically all we have are counts at administrative unit level which tend to be rather crude spatial aggregates.

In fact, here we will take our minimum size of city from various data sets as 50,000 persons. This means that we will not be able to measure the total numbers of cities globally or regionally at any point in time, and such statistics are somewhat strangely absent but understandably in most discussion of these questions. We will also define cities by two estimates of size; first by the numbers of persons within a certain physical area, notwithstanding our commentary that physical area is increasingly the most problematic of measures, and second by an estimate of the size of their highest building. Population of course varies through time and we have some continuity of evolution in this sense but high buildings are only fixed for an instant of time before they are demolished, or redeveloped. In this sense, we will focus on changes in city size only during the last 150 years from the time when truly high buildings were first developed (as skyscrapers greater than about 6 stories in height). These two kinds of measure – population numbers and building height – show the transition to a world where cities are ever more difficult to define, where cities begin to fuse together, in fact to a world where everybody will be living in one kind of city or another by the end of this century. In that sense, our thesis here is that we are rapidly heading towards a world of cities where there is one global empire. In this sense, we will imply at least that the concept of empire is becoming passé.

After a further discussion of empire in the next section, we will define city size from two data sets: first from the middle of the 20th century for large cities that currently have reached some 750,000 persons, and then for a cruder range of time periods from 1975 for a much larger number of urban agglomerations with current populations greater than 50,000 persons. These serve to define the growth of different sizes of city and we will show the well-known form that such city size distributions are highly skewed, following power laws in their upper tail. We then explore city size in terms of building heights, examining a smaller set of the world's truly high buildings greater than 200 meters in height (~650 feet). This switch in focus enables us to examine how the size of cities changes in a discrete way with new building appearing in contrast to the more disaggregate ways where we have continuously changing population sizes. Indeed, we need to note that when we explore high buildings in this way, none of the buildings that we examine have ever been demolished. Although, skyscrapers are periodically demolished, these are only a handful amongst the buildings we examine and thus the two data sets – population size and building heights – are remarkably different in what they are able to tell us about city size. Once we have presented these ideas, we will speculate on how other data

based on relationships between cities enable us to say something significant about the move to a world where cities in the global empire need to be defined in terms of flows, networks, and linkages rather than sets of locations.

Empires and Cities

Both empires and cities are based on ill-defined concepts which present a level of ambiguity that makes any measurement of their shape, size, and scale problematic. In fact, empires are historically constructed, often in retrospect, and although there is reasonable agreement amongst scholars on the dates at which different empires existed as well as the component states and cities that made up their composition, we still lack a definitive historical record that charts how complete is our knowledge of the world's empires to date. Empires in history are still being discovered, the ones we know of being continually reinterpreted in terms of their size and influence. Cities, on the other hand, are easier to define because their spatial location is clearer but they still require some conceptual agreement about where a city begins and ends in physical terms. Density and the networks which tie their populations together constitute major determinants of their physical extent and thus place limits on the area over which their population can be measured. But in a world where relationships between populations are key to defining agglomeration, cities are no longer simply physical entities; indeed, they never have been but increasingly it is difficult to see adjacent configurations as the physical elements that exclusively define a city. In fact, since the modern era, cities have no longer been exclusively physical in their definition.

Insofar as we can define empires, the focus is not upon the populations that are captured within the physical space defining the clustering of cities and regions that compose their form but on their geographical extent measured in terms of area. In fact, there are very few scholars who have sought to define the rise and fall of empires over recorded human history, from the agricultural revolution which began some 12,000 years ago to the urban revolution when cities first emerged about 5000 years ago through to contemporary times marked by the industrial revolution. Because historical scholars have tended to concentrate on particular periods of time with respect to the development of empires, there is no standard source cataloguing empires, and the further back we go in the historical record, the murkier the picture becomes. The best we can do is take the list of empires from Wikipedia and explore what this tells us (https://en.wikipedia.org/wiki/List_of_empires). The list was started in 2006 and has been continually updated since then. It is not attributable to any particular person for like all Wikipedia entries it is crowd-sourced. There are in fact 199 different empires listed with the earliest date as 2300 BCE when the Akkadian Empire in Sumeria began, which was succeeded by the Assyrian Empire in 2025 BCE. The average length of duration for all empires is 269 years with a standard deviation of 291.

To get some sense of this longevity and how this has changed over the last 5000 years, we plot duration against the start date of the empire which we show in Figure 1. Despite the huge variation in the longevity of empires, the general impression from this plot is that empires have got shorter in duration. Arguably, this might be due to the fact that we start 5000 years ago and as we approach the present, there is not enough time for an empire to establish real longevity. But our casual knowledge of world history is such that we know that empires are getting shorter as the world hurtles towards a completely global empire which is almost upon us. In Figure 1, we have also plotted the cumulative average duration as a continuous line which drops from the first empires in 2300 BCE and fits the curve quite well using a linear regression (with an r^2

= 0.658). This linear trend enables us to predict that in about 1000 years, the average duration will be close to zero, while the actual points plotted since the modern era began, suggest that it will be during this century that the idea of empire will disappear in terms of its economic, if not political logic. Much the same might be said of cities as we will imply below.

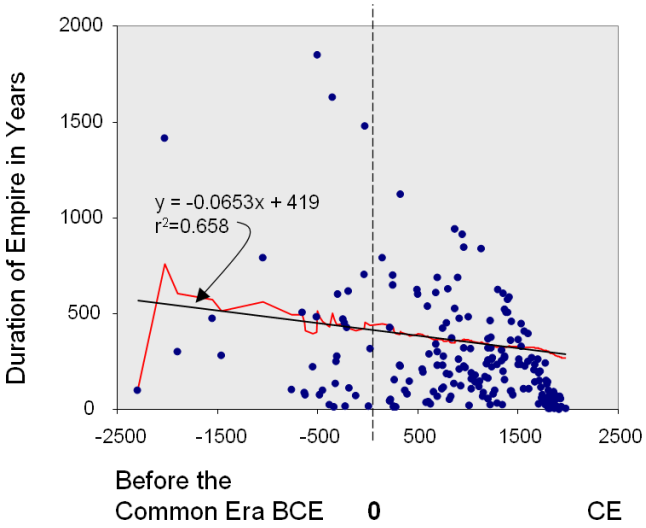


Figure 1: The Duration of Empires

We can also correlate duration with the growth of world population. Without providing such a plot with a much smaller number of observed populations from 2300 BCE, it is quite clear that the data is highly positively skewed following an approximate power law which explains about 55% of the variance in the data (r^2 = 0.546). In fact, we do not have any estimates of the population which is under the influence of each empire, for all we have – again from Wikipedia (https://en.wikipedia.org/wiki/List_of_largest_empires) – is the maximum land area which these empires cover. When we correlate land area with the duration of its empire, the correlation is almost zero which must be partly attributable to the fact that the earliest empires – prior to the Copernican Revolution and the discovery of the new world – were limited in their global extent. In fact, the smallest of these empires were the longest in duration in classical times and before. In short, we can say really very little about the size of empires for all that we know is anecdotal and no one, as far as are aware, has estimated the size of empires in population terms. None the less, our intuition and casual knowledge associates city building and the development of primate cities with each empire as well as more recently with nation states.

The Distribution of the Largest Cities by Population

As we have noted, the physical definition of cities is arbitrary. Locations at which a city becomes a rural area depend on the density of settlement, the extent to which persons living in different parts of the wider hinterland of the city actually interact with the core city through commuting, and other relationships involving flows of money, energy, and information, as well as the types of land uses and activities that define what goes on in a city. A rule of thumb for British cities is that areas with greater than 14 persons per hectare and with more than 20 percent of their population commuting to the core city provides a robust definition. However,

we will not invoke such rules here for the cities that we examine are defined by national and international agencies and thus we are dependent on their definitions.

Some general properties of city size distributions have been well known since the mid-20th century. Such frequency distributions are positively skewed, with higher frequencies for smaller sizes, their best approximation being the lognormal probability distribution. A particularly convenient way of representing such distributions is in their counter-cumulative form where the distribution is rank-ordered from the largest sizes to the smallest. This sort of representation is often similar to an inverse power function in their upper tail, although over their whole range they usually resemble a lognormal. A good deal of effort has been on simulating how such distributions occur using minimal assumptions defining a generative process which results in this form. This competitive process produces a pattern where the bigger objects win out against the smaller but where the distribution is continually refreshed from below, small being prior to becoming big when new objects are introduced. In this case, big cities are always small first, while such distributions are often referred to as rank-size distributions after Zipf (1949) who popularized this form. The first observations that cities and incomes and much else besides follow these forms go back, at least, to Auerbach (1913) and Pareto (1896).

These distributions have an intuitive logic that suggests that they are the outcome of a competitive process under the general constraint that resources are limited and that the bidding process leads to the most successful winning out and achieving an ever greater proportion of the total pie. The way this occurs under conditions of minimal random growth is explained quite intelligibly by Krugman (1996) who also demonstrates that such a process leads to a rank-size rule that approximates the observed distributions of city sizes from one time period to the next with unerring similarity. In fact, Krugman (1996) says: "The size distribution of cities in the United States is startlingly well described by a … power law: the number of cities whose population exceeds S is proportional to $1/S$. This simple regularity is puzzling; even more puzzling is the fact that it has apparently remained true for at least the past century" (p. 399). In short, over time the distributions appear to be quite stable but when the detailed movement of cities within the rank-size envelope is examined, there is considerable volatility. A way of visualizing this is through the rank clock which is a graphic intended to demonstrate the relative stability or otherwise of objects comprising such a distribution from one time period to the next (Batty, 2006). If we now examine the global distribution of city sizes for the largest cities, we do not see any primacy. No one city or even a group dominates but if we break the world into discrete regions, particularly nation states or even continental land masses, we do find some primacy where the biggest city or cities take a more than proportionate share of the population (and its resources); that is, more than proportionate with respect to the pure rank-size rule which is based on what the inverse power law might imply. In short, such a power law implies that the population of a place j called S_j is inversely proportional to its rank r_j where the rank is raised to the power of unity – that is, $S_j \sim 1/r_j = r_j^{-1}$.

Apart from the key definitional issues as to what constitutes a city, the problem we have in examining the size of cities through the city size distribution is the fact that we have to break into the distribution at some point, as much because there is no authoritative source of data on all cities. If we thus examine let us say the top 100 cities at time t and then we examine the top 100 at times $t+1, t+2, ...$, in each time period the cities will be different in that cities rise and fall in the rank-size hierarchy. Thus some cities join the list of top 100 while some leave. We can only avoid this problem if we have all cities for all time, but cities grow from small seeds and thus we will never have a complete set; by definition, it is impossible to know all

cities at all times, for many have not yet been invented for populations that do not yet exist. This might seem trite but it is key to exploring the size of cities relative to each another. For example, looking at the top 100 cities in the US from the Census of 1790, Houston did not enter the list until 1920, while Phoenix only entered in 1950. Charleston, SC which was in the top 10 in 1790 had fallen out of the list by 1900 as the colonial influence waned while cities in the rustbelt like Detroit grew to the 4th rank by 1960 but by 2000 had fallen to 11th. This kind of volatility is a puzzle given the relative stability of the rank-size envelope as is confirmed again from various data sets used here for large cities across the globe.

The best historical data set is that produced by Chandler (1987) which was used 6o first define the rank-clock noted above (Batty, 2006). A more considered version of this data set which has been consistently merged with that from Modelski (2003) by Reba, Reitsma, and Seto (2016) is now available for the last 6000 years and although we will not use this here, this gives a crude but long historical times series for which city sizes can be associated with empires at least until the modern era. However the real changes that we will focus on in this chapter are changes during the 20th and 21st centuries, during the great transition that is taking place as the world shifts from one based on energy to one based on information, from a world dominated by cities where the basic mode of transport was walking to one where automated kinds of transport enable populations to reach distant places much more quickly and where information enables populations to maintain intense contacts with no intervening physical distance whatsoever. In this world, the notion of city size will be increasingly problematic but before we speculate on this future, we will look at various data pertaining to large cities.

Our first data set was compiled initially by the United Nations and then cleaned, modified and extended in 2014 by Nordpil at the Stockholm Resilience Center. The data is for all 589 cities which had reached a population of 750,000 by 2015 and is based on the populations of those same 589 cities every 5 years from 1950. Thus it is obvious that cities that had not reached 750,000 in 2015 would not be included in the set even had they reached that population before 2015 and then declined below the threshold. In fact, it is most unlikely there are any such cities in the list. More likely there are cities that do not reach the threshold in 2015 that could well be greater than the populations of other cities in the data set at earlier points in time. The biggest city in the list in 1950 was New York City with 12 million persons but Tokyo took over in 1955 and remained at the top of the list throughout the period, reaching 36 million in 2015. The smallest city in the list in 1950 is Karaj in Iran with a population of 10,000 which grew to 1.6 million by the end of the period while the smallest at the end of the period is Lodz in Poland at the minimum threshold of 750,000.

The shape of the distribution, which we have argued can be approximated in its upper tail as a power law, provides an indication of the relative distribution of city sizes but the real power of this data relates to the comparisons of these distributions over time. To begin, we will plot the largest city in the data set (which we will set as Tokyo) at each of the 14 time instants from 1950 to 2015, and compare this with the total population which we show in Figure 2(a). It is quite clear that the total population of these cities grows much faster, almost exponentially during this period, than Tokyo, and this implies that the larger the city, the slower the growth rate, a fact that is borne out in general for the growth of many city populations during the last 50 to 100 years. We can show these relative growth rates as a more direct comparison. In Figure 2(b), we normalise three trajectories – that for the two populations in Figure 2(a), that is, the Tokyo population and the total urban population from this data set, and compare this with the total world population. Total urban population grows to 4.35 times its 1950 value, while total world population grows to 2.85 times its 1950 value with Tokyo population growing somewhat

faster than world population at first but then ending up at much the same relative growth, some 2.92 times its 1950 value. In this sense, we suggest that the largest city populations are growing at rates somewhat less than total city (urban) and world populations. We will explore this in the fuller JRC data set in the next section.

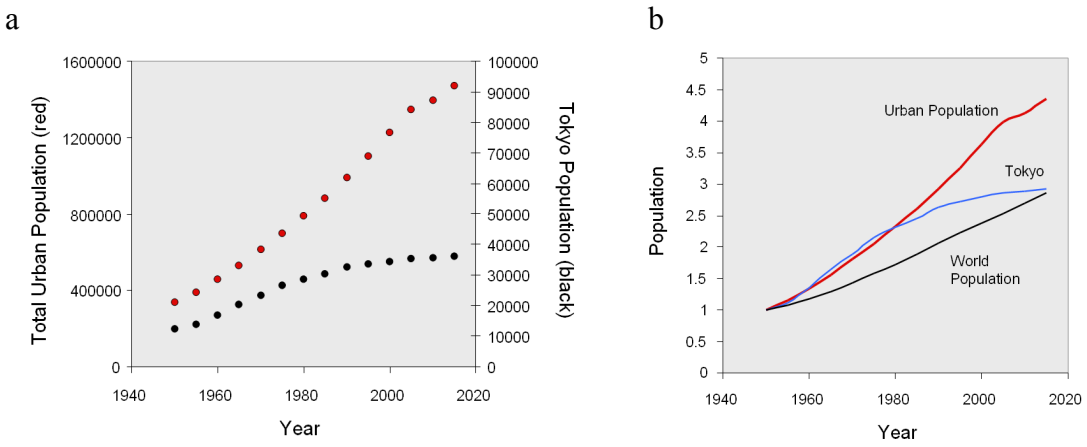


Figure 2: The Growth of Urban, World and Largest City Populations

There is no sense of primacy in this data with respect to the biggest city population, nor is this the case for the top 5 cities at each time period, in fact quite the contrary with the top five cities declining from 6.2% of the total to 3.9%. This is quite clear from the rank-size plots that are presented in Figure 3. Moreover, these plots also show that the slope of the distribution becomes a little lower (flatter in visual terms) as the population of these cities grow, again reinforcing the notion that it is the smaller cities that are growing somewhat faster than the largest. In fact, the current view, which is by no means borne out from a considered examination of the data, is that that big cities are getting bigger, the rich are getting richer, all in an age when big cities are no longer perceived as being 'evil' but are now 'cool'. But small, it seems, is still beautiful when it comes to city growth, or at least with respect to mega-cities!

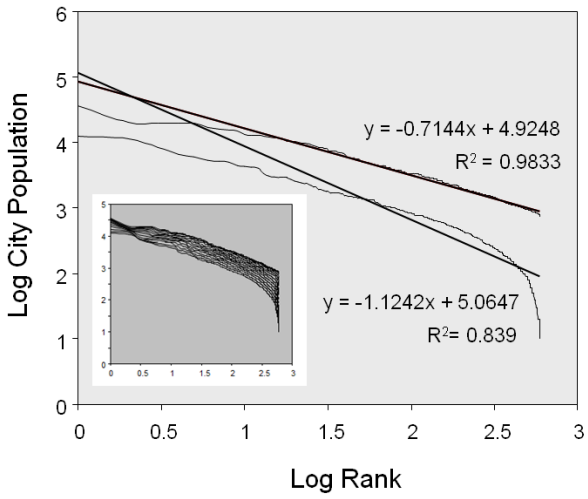


Figure 3: Rank-Size Relations for the Largest City Populations 1950 to 2015

City Size and Urban Agglomerations

To test our tentative conclusions, we need a much fuller database that is structured around more meaningful boundaries and enlarges the number of cities to the point where smaller towns are included. One way of achieving this is to use definitions of cities based on density and connectivity and in this context, it is more useful to talk of urban agglomerations that do not necessarily coincide with administrative city definitions used in the Nordpil (2014) database from the United Nations in the previous section. We have constructed such a data base from the smallest administrative units in Great Britain from which we have extracted contiguous city areas which meet our thresholds of a minimum density of 14 persons per hectare and a 20 percent commuting field (see Arcaute et al., 2014) and we have done the same for France (Cottineau et al., 2017). But this is not possible for the all cities, and thus we resort to the next best available data which is produced by the European Union's Joint Research Council (JRC). The JRC have recently developed such a database of all urban agglomerations greater than 50,000 persons which generates some 13,845 city clusters for the latest year which is 2015. The number of clusters at 1975, 1990 and 2000, is less than this at 13,622, 13,421, and 13,017 respectively where it is necessary to further slice off the smallest clusters at these years to preserve the shape of the distributions. This database is constructed using density and contiguity thresholds as well as remotely-sensed data that is taken from base imagery at 250 square meter resolution from which urban areas are extracted using state-of-the-art techniques. This is referred to as the Global Human Settlement Layer (http://ghsl.jrc.ec.europa.eu/) and it is constructed from Landsat and CIESIN GPW v4-Multitemporal data. Details of the data and its construction are available from Dijkstra and Poelman (2014), and Pesaresi and Freire (2016).

In terms of the growth of the population in this data, the total increases from 1.918 billion in 1975, to 2.655, 3.131, and 3.823 billion in 1990, 2000, and 2015 respectively. In terms of the largest city which is the Greater Bay Area (of Hong-Kong-Shenzhen-Guangzhou-Zuhai-Macau) with 46,038,520 million in 2015, this proportion actually falls from 1.2% of the total urban population in 1975 to 1.1% in 1990 and then 1.0% in 2000. It rises in 2015 back to 1.2% but this does not really imply that the biggest agglomeration is getting bigger as a percentage of all urban population. If we widen this cluster to the top 5 cities, the proportions range from 4.4% in 1975, 4.4% in 1990, up to 4.5% in 2000 and then to 4.6% in 2015 and this does show that the biggest cities are taking a slightly increasing proportion of the urban total. But the numbers are so dependent on the total numbers of cities which are defined relative to only one threshold at one point in time, that we do not have any confidence in stating that the biggest cities are becoming more than proportionately bigger as the world system continues to expand. In fact, when we examine these distributions as rank-size relationships below, it is very clear that the biggest cities are less in size than one might expect in terms of their overall distributions at each of these four dates and that this does not seem to be correcting itself through time. To anticipate our conclusion, there is no sense of primacy for the largest cities in this global data.

The best way to get some sense of the shape of these four distributions is to plot them by their rank-size and we illustrate this in Figure 4(a) for the 2015 distribution. This is the baseline for this data set because all the clusters are greater than 50,000 persons in size and there is no long tail effect. This distribution follows a classic power law, in fact one so extreme that it is much easier to show the strength of this form in the way we did for the Nordpil data in Figure 3 on logarithmic scales. We show the 1975 and the 2015 distributions in Figure 4(b) where it is quite clear that in the upper tail (where the largest clusters are located) these are less than what the best fit regressions lines imply from which we can speculate that the largest cities in the data

are in fact 'too small' and all the lesser cities too big (or more perversely, the other way around). Whatever the 'true' distribution might be, the estimated functions are remarkably consistent and extremely well-fitting and from this data it is extremely hard to argue that the city size distribution is anything other than a power law, thus echoing Krugman's (1996) observations noted earlier. These figures indicate a lack of primacy but contradict those of the Nordpil data in that the slope of the rank-size function increases very slightly over this 40-year period from 1975 to 2015, thus implying that big cities are becoming slightly bigger more than proportionately as they appear to grab more of the world's increasing population. But the evidence and the conclusion is weak and a more confident hypothesis would be one that suggested that the relative competitive effects in the system that generate such power laws are fairly stable.

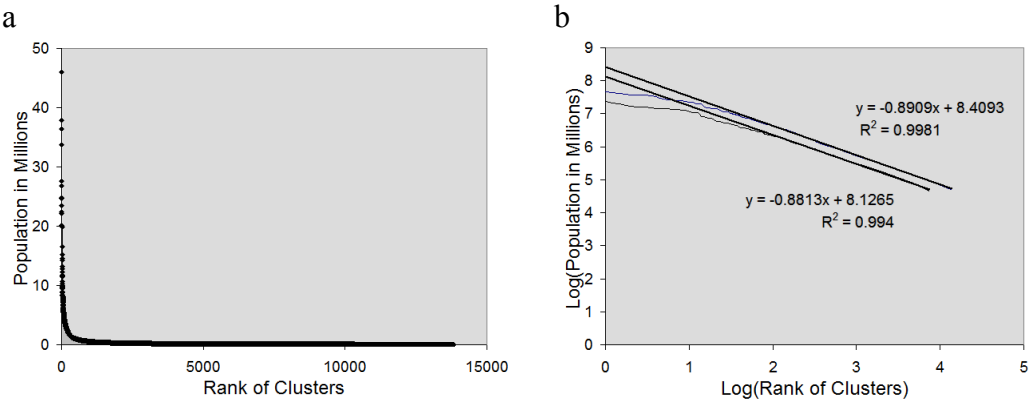


Figure 4: The Distribution of Urban Agglomerations than Greater than 50,000 Population
a) The Rank-Size Distribution for 2015 b) The Linearized Rank-size Distributions for 1975 and 2015

If we now ask the question, what will the population of the largest cities be if the world population continues to grow logistically, stabilizing at something like 10 billion by the end of this century? From this analysis, we can hazard some guesses as to how big the biggest cities might then be. We need to make the assumption that the entire city system will remain completely stable in that the largest cities do not grow any faster or slower than the smaller cities and that each city gains the same proportion of any new growth as any other. Then with these very severe constraints, to get an idea of what the largest city might be (and of course every city in the distribution), all we need to do is to assume the regression line shown for 2015 in Figure 4 (b) continues to have the same slope as the population increases. We need to assume that the distribution is closed and that the same share of population that goes to the cities continues to go to them. Currently there are 3.82 billion persons residing in the 13,845 agglomerations in 2015 where the world's population is 7.38 billion. Thus those living in urban clusters comprise some 52% and if this remains the same, when the population reaches 10 billion, some additional 1.35 billion persons would be part of these same clusters. In fact, we probably need to assume that all 10 billion might be urbanized and if this were the case then some 6.12 billion persons would be added to the existing urban agglomerations. All we need to do is scale the equations proportionately or in fact scale the basic data if we want to preserve the original structure of the city system. Scaling the data in the case where 1.35 billion persons are added, the largest cluster/city would be about 65 million while when all the world is living in cities, the largest would be 120 million. Brave assumptions, foolhardy predictions perhaps but food for thought as to what a future urban world might really look like!

Measuring City Size from Tall Buildings

When we look at size, there are several different but related measures. Population is a continuously changing variable that is highly mobile and needs to be aggregated into discrete spatial units for it to show any stability over space and time. Populations are usually counted from individuals, sometimes household groupings but occasionally the count is normalized by area to produce a density. In contrast, more stable variables that change very slowly over long time periods can be attributes of buildings such as heights, volumes, stories and so on. The advantage of using these measures is that there is much more comparability over time and if one concentrates on very high buildings, then over the last century or more, the highest buildings have remained intact. It is unusual for high buildings to be changed by the addition or deletion of stories and their demolition and replacement is only a minor part of their evolution. This has been the case for the last century although with new construction techniques, this notion of longevity of high buildings is likely to change, particularly as we move towards construction that is more organic, in that buildings are ever more likely to be 'grown' rather than engineered in traditional manufactured ways. Thus our focus here on high buildings over the last century could well be very short-lived.

Currently then, the best immobile attributes of a city are in its buildings and to make the comparison with a city's population, we will choose the highest buildings from a global database of high buildings – skyscrapers, so-called – all of which are greater than 200 meters (or 656 feet) in height. These are buildings with greater than about 40 stories with the oldest in the data set being constructed in 1909 and the latest being topped out in the year of writing, 2018. The tallest building in the set is the Burj Khalifa in Dubai which was completed in 2010 and which stands at 2,717 feet or 828 meters (163 stories). There are 1,310 buildings in the dataset and one consequence of examining the size of cities in this way is that there can be many high buildings for a single city constructed at a single point in time in contrast to only one population count for a city. The database we have used is available online from The Skyscraper Center which is run by the CTBUH (Council on Tall Buildings and Urban Habitat) and can be accessed at http://www.skyscrapercenter.com/. One feature of our use of this data is that it is truly dynamic in that we have a complete space-time series of high buildings which differs from the more restricted temporal city-size data sets that we have shown previously particularly for the Nordpil data where city size is shown over time in Figure 2.

In Figure 5(a), we plot the data with respect to the height of each building greater than 200 meters by year of construction (that is, when the building was finally completed or 'topped out') and it is clear from this that the frequency of building heights at any single year or over any period of time consists of many more smaller buildings than larger; that is to say, high buildings are harder and more expensive to construct than smaller ones but one obvious consequence of technological development is that maximum building heights are increasing with time, as clearly illustrated in the data. Figure 5(a) however shows that the first building in the data set constructed in 1909 was 700 feet in height but that more and more buildings at this height and below down to the 656 foot threshold have been constructed over time. To explore heights in this way, we have in fact focused on the highest building in each year since 1900 for it appears that this is a good measure of city size although other measures based on the distribution of heights for different cities do provide a different perspective. Therefore, in Figure 5(b), we plot the highest building in each year and we have also computed the cumulative average for each year which is the heavier 'flat' red line. This average is remarkably stable, varying from a minimum of 656 feet to 871 feet, and from 700 in 1909 to 793 in 2018. Somewhat incredibly the average peaked at 871 feet in 1931 which says a good deal about the

frenetic development of skyscrapers in New York City prior to the great depression (Barr, 2016). And it also casts doubt on using any kind of average to measure the size of cities in this way, making it essential to capture size using the highest building rather than the average height.

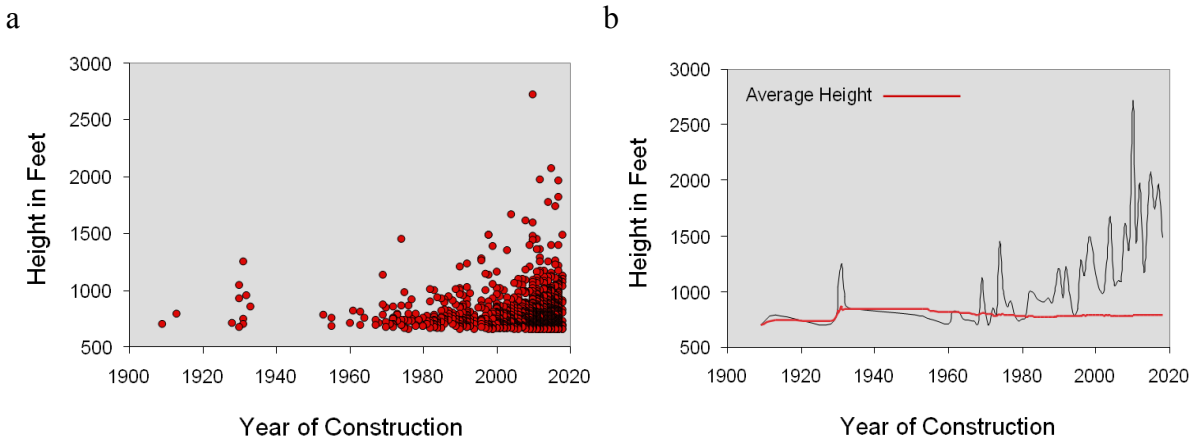


Figure 5: The Highest Buildings from 1900

a) All Buildings Greater than 200 meters (656 feet) b) The Highest Building & Average Height each Year

Before we look at the distribution of the highest buildings in each year, it is worth looking at the single highest building year by year as shown in Figure 5(b). What is very clear is the volatility that is very strongly correlated with boom and bust cycles in the building industry but massively confounded by the key economic cycles pertaining to depressions and recessions over the last 120 years. The 1930s recession is stark in terms of this graph and one can also recognize somewhat more muted recessions such as those in the period of rapid de-industrialization in the west – in the 1970s of which the oil crisis in 1974 was key. The recession in 2008 or thereabouts is also very significant but since the 1990s, the volatility has picked up and it would appear that this is partly due to the nature of the business cycles in the Middle East and Asia-Pacific which now dominate the construction of the tallest buildings globally. Technological change too in building construction is ever faster and the periodicity of these cycles seems to be getting shorter while their amplitude is getting bigger. These skyscraper cycles as they have been called may well be short-lived for it would appear that the economic forces that ended the 1920s boom in the west are quite different now for oil rich nation states and rapidly growing Chinese cities. No longer can one assume that the growth of high buildings is a close correlate to economic booms and busts as they have happened in the past (Lawrence, 1999; Thornton, 2005).

When we examine the shape of the distribution of the highest buildings where the rank order is from the most recent year 2018 to the earliest, in this case to 1909, there is a clear power law in the upper tail with a similar performance to the power laws for Nordpil and JRC databases shown in Figures 3 and 4(b). There is no sign of lognormality in the long tail for the relationship in logarithmic terms is entirely linear and fits the rank-size frequency with a slope of 0.174 and an r^2 of some 0.997. In terms of the variance explained, this is comparable to the the Nordpil data where the slope is 0.714 ($r^2 = 0.983$) and to the JRC data where the slope is 0.891 ($r^2 = 0.998$). In fact, the slope is much lower – meaning that the function is much flatter than those associated with the city populations – and this implies that the degree of competition is much more ferocious with respect to the creation of high buildings than that for city populations. In

short, bigger city populations appear to find it easier to outbid lesser city populations, thus ensuring that there is a greater range of city sizes than in the case of skyscrapers where the clustering is much more bunched. Unlike the city populations, Figure 6 illustrates an almost perfect power law from which we might speculate that such laws only emerge where there is severe competition and where cities continually strive for the tallest buildings. In fact, a somewhat legendary example of this competition is the example of the race for the tallest building in Manhattan in 1929: "In the era of the skyscraper boom, John Jacob Raskob determined that the Empire State Building should soar above the Chrysler Building, which was under construction at the same time. Architects had originally designed an 80-story building, but had to alter plans several times to keep up with the Chrysler Building. Eventually, a mooring mast was devised to add additional height. Intended as a docking station for air blimps, the idea was abandoned after a few failed attempts. Today, the tower of glass, steel, and aluminum is used as the base of the TV tower." (Penafiel, 2006).

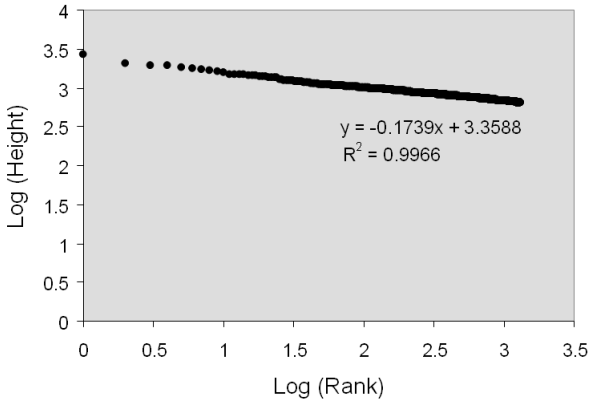


Figure 6: The Rank-Size Relationship for the Highest Buildings (Greater than 200 Meters Since 1909)

New Measures of City Size

Cities do not exist in isolation and although our measures of population are constructed with cities defined in such a way that commuting thresholds and densities determine what we consider the most appropriate form of physical self-containment, these are crude approximations to the multitude of flows which define cities in a global world. Elsewhere we have argued that cities in this century and beyond will be part of a global network that is hard to fracture and bound in the way we have categorized our cities in the past and the way we have measured their size in this paper (Batty, 2017). What we have described here is but a starting point in thinking about the shape of the global city system – the global empire in the focus of this book – and what we urgently need are other measures of city size, not the most obvious such as population and buildings but measures of income, information, energy, services, education, productivity and so on which serve to define the flows that provide the cement that binds the global system together. Only when we have good data concerning such linkages and connections will we be in a position to develop a somewhat more subtle and insightful view of how big the individual cities that form this global hierarchy actually are. Beaverstock, Smith, and Taylor (2000) who have done much to progress these ideas, are fairly pessimistic in their view when they say: "… we know about the nodes but not the links in this new metageography" (p.127). Our ability to unpack locations – nodes – into flows that define

networks is still fairly limited and represents a hard barrier to making advances on defining cities other than in the most obvious physical terms that we have employed here.

We do have good data on physical movements of populations between different cities, particularly in the west but these do not enable us to say much about the routine, day-to-day connectivity between cities that drives the world economy. Migration of populations is much more of an index that relates to the growth of cities and although such volumes can tell us about the relative positioning of cities in the growth hierarchy, they say little about functional connections. Research is slowly beginning to construct data sets that measure such connections between cities such as those produced by economic geographers commenting on world cities, notwithstanding the difficulties of unravelling such linkages which can be largely invisible to traditional methods of measurement. The problem however is generating a single measure comprehensive enough to provide an acceptable measure of city size. In fact, different flows are hard to compare and in any case, when we reduce our thinking about city size to flows between cities, the notion of size itself changes to one related to the strength of relationships. If we have different flow systems that can be reconciled to the size of the affected population, then it is possible to begin to construct measures of size but most of our flow data cannot be grounded in this way. For example, if we had flows of email traffic, between cities, we need some absolute arbiter to relate the amount of traffic to its significance. For example, a very large proportion of the population now use email and if we could measure this between cities, then we can get a good idea of relative connectivity. But in terms of measuring the size of cities in terms of email traffic, then we need to know how much of the population engages in this activity and how frequently; and to really produce meaningful conclusions, we need to know how significant the email traffic is with respect to other activities such as education, production and so on.

A good example of this difficulty is contained in the excellent but rudimentary research on connections between world cities produced by the Loughborough Globalization and World Cities Research Network. The links between world cities developed by Beaverstock, Smith, and Taylor (2000) are based on the coincidence of the same key office-services in different cities. They construct a matrix of probabilities of an office existing in each of 10 world cities if that office exists in any one of these cities. The problem with this data in measuring anything other than relative connectivity is that there is no way to ground the measure in any absolute strength of relation. We would need to know how big a proportion of the economy these offices constitute and even if we were to get this, we would need to compare this to other measures of size that were connected with activities not directly associated with the economy such as entertainment, education, culture and so on. To begin to construct a single measures based on these kinds of flows, we would need to look at a host of measures pertaining to other flow systems and their significance. Our tentative conclusion at this point is that measures of connectivity and their importance exist in parallel to the city size measures that we have focused upon here; in short, our analysis implies both are important and that they cannot be easily merged.

There are other single measures of size for individual cities that need to be considered alongside the measures presented here. Clearly income and GDP stand out as being significant in terms of the economy of cities but transformations of these measures into densities and per capita indices are important. It might seem strange that after at least half a century of sustained effort in thinking about the form and function of cities, their scale and shape as well as their prosperity and economy, we have to embark on working with some very simple measures pertaining to size. But cities are complex systems which are getting ever more complex. Our understanding

of them is changing and the role of physicality in their definition, is now under severe scrutiny as we head towards an age of the global city, where everyone will be living in one kind of city or another by the end of this century (Batty, 2011). The quest is still to find good and better measures that enable us to make comparisons between cities of different sizes in the effort to understand how cities will develop during the rest of this century.

References

Arcaute, E., Hatna, E., Ferguson, P., Youn, H., Johansson, A., and Batty, M. (2014) Constructing Cities, Deconstructing Scaling Laws, *Journal of the Royal Society Interface 12*, 20140745. http://dx.doi.org/10.1098/rsif.2014.0745

Auerbach, F. (1913) Das Gesetz der Bevölkerungskonzentration, *Petermann's Geographische Mitteilungen*, *59*–I, 74–77.

Barr. J. M. (2016) *Building the Skyline: The Birth and Growth of Manhattan's Skyscrapers*, New York: Oxford University Press.

Batty, M. (2006) Rank Clocks, *Nature*, *444*, 592-596.

Batty, M. (2011) Commentary: When all the World's a City, *Environment and Planning A 43*(4) 765–772.

Batty, M. (2017) Cities as Systems of Networks and Flows, in T. Haas and H. Westlund (Editors) *In The Post-Urban World: Emergent Transformation of Cities and Regions in the Innovative Global Economy*, London: Routledge, 56-69.

Beaverstock, J. V., Smith, R. G. and Taylor, P. J. (2000) World City Network: A New Meta-Geography? *Annals, Association of American Geographers*, *90* (1), 123-134.

Brenner, N. and Schmid, C. (2012) Planetary Urbanization, in M. Gandy (Editor) *Urban Constellations*, Berlin: Jovis, 10-13.

Chandler, T. (1987) *Four Thousand Years of Urban Growth: An Historical Census*, Lampeter, Wales, UK: Edward Mellon.

Cottineau, C., Hata, E., Arcaute, E. and Batty, M. (2017) Diverse Cities or the Systematic Paradox of Urban Scaling Laws, *Computers, Environment and Urban Systems*, *63*, 80-94.

Dijkstra, L. and Poelman, H. (2014) A Harmonised Definition of Cities and Rural Areas: The New Degree of Urbanisation, *Regional Working Paper 01/2014*, Brussels BE: European Commission Directorate-General for Regional and Urban Policy (DG REGIO).

Harari, Y. N. (2015) *Sapiens: A Brief History of Humankind*, London: Vintage Books.

Krugman, P. (1996) Confronting the Mystery of Urban Hierarchy, *Journal of the Japanese and International Economie*s, *10*, 399–418.

Lawrence, A. (1999) *The Curse Bites: Skyscraper Index Strikes*, Property Report, London:

Dresdner Kleinwort Benson Research.

McGee, T. G. (2009) The Spatiality of Urbanization: The Policy Challenges of Mega-Urban and Desakota Regions of Southeast Asia (PDF), *UNU-IAS Working Paper 161*, Tokyo, Japan: United Nations University Institute of Advanced Studies.

Modelski, G. (2003) *World Cities: -3000 to 2000*, Washington, DC: Faros 2000

Nordpil (2014) *World Database of Large Urban Areas, 1950-2050*, Stockholm, SE-10691: Stockholm Resilience Center, Stockholm University, accessed 4/06/18 at https://nordpil.com/resources/world-database-of-large-cities/

Pareto, V. (1896, 1897) *Cours d'Économie Politique,* Professé a l'Université de Lausanne, Volume I and Volume II, Lausanne, CH: F. Rouge.

Penafiel, K. (2006) The Empire State Building: An Innovative Skyscraper, *Buildings*, https://www.buildings.com/article-details/articleid/3180/title/the-empire-state-building-an-innovative-skyscraper

Pesaresi, M. and Freire S. (2016) GHS Settlement Grid following the REGIO Model 2014 in Application to GHSL Landsat and CIESIN GPW v4-Multitemporal 1975-1990-2000-2015. Ispra, Italy: Joint Research Centre JRC, European Commission, see PID, http://data.europa.eu/89h/jrc-ghsl-ghs_smod_pop_globe_r2016a

Reba, M., Reitsma, F., and Seto, K. C. (2016) Spatializing 6,000 Years of Global Urbanization from 3700 BC to AD 2000, *Scientific Data*, *3*, 160034, DOI: 10.1038/sdata.2016.34.

Tainter, J. (1990) *The Collapse of Complex Societies*, Cambridge, UK: Cambridge University Press.

Thornton, M. (2005) Skyscrapers and Business Cycles, *The Quarterly Journal of Austrian Economics*, *8* (1), 51–74

Zipf, G.K. (1949) *Human Behavior and the Principle of Least Effort*, Reading, MA: Addison Wesley.