Visualizing Space–Time Dynamics in Scaling Systems

MICHAEL BATTY
Centre for Advanced Spatial Analysis (CASA), University College London (UCL), London WC1E 6BT, United Kingdom

Received May 6, 2010; accepted June 27, 2010

The signature of scaling in human systems is the well-known power law whose key characteristic is that the size distributions of their objects display self-similarity in space and time. In many systems such as cities, firms, and high buildings used here as examples, power laws represent an approximation to the fat or heavy tails of their rank-size distributions, appearing stable in time with little change in their scaling over tens or even hundreds of years. However, when the detailed dynamics of how their ranks shift in time is examined, there is considerable volatility in such distributions. To explore this microvolatility, we introduce measures of rank shift over space and time and visualize size distributions using the idea of the "rank clock." We illustrate this for populations of Italian towns between 1300 and 1861 and then compare this analysis with city-size distributions for the world from 430 B.C.E., the United States from 1790, Great Britain (England, Scotland, Wales) from 1901, and Israel from 1950. When we extend this analysis to the distribution of US firms from 1955 and high buildings in New York City and the world from 1909, we generate a rich portfolio of space-time dynamics that adds to our understanding of how different systems can display stability and regularity at the macro level in the face of considerable volatility at the micro.

Key Words: scaling, power laws, Zipf's law, rank-size, rank clocks, city-size, firm-size, skyscrapers

SCALING IN COMPLEX SYSTEMS

Objects or entities that define many complex systems often scale with respect to the frequency at which they occur in space and/or in time. Such scaling reveals an order in the system manifest in the fact that patterns recur over different scales, revealing what is called in fractal geometry, self-similarity. This is best visualized as some configuration of system entities that appear the same, at least statistically, from one scale to another, good exemplars being dendrites whose branches mirror the way rivers drain a landscape, crystals solidify, and liquids of different viscosity penetrate one another, all the way to how energy is delivered to the human body and how organizations arrange themselves in overlapping hierarchies [1].

Formally, the most general scaling, which captures the frequency $f(x)$ with which elements of different size x

Corresponding author: Michael Batty; Centre for Advanced Spatial Analysis (CASA), University College London (UCL), London WC1E 6BT, United Kingdom. (e-mail: m.batty@ucl.ac.uk)

DOI 10.1002/cplx.20342
Published online in Wiley Online Library
(wileyonlinelibrary.com)

recur, is signified by a power law defined as $f(x) \sim x^{-\alpha}$, where α is the parameter of the distribution. Such frequencies scale in that if x is multiplied by λ then $f(\lambda x) \sim (\lambda x)^{-\alpha} = \lambda^{-\alpha}x^{-\alpha} \sim f(x)$. One essential feature of such power laws is that fact that they have no length scale, being unbounded with respect to the size of an event or its frequency. Systems so defined have the potential to generate events of any size on any scale, extreme events, and in some contexts, display the potential for uncapacitated growth and extreme competition. In systems that scale, the focus can either be upon the frequency of events of different size or on the size of the events themselves. In one sense, this is only a matter of changing perspective but it does lead to confusion in terminology with respect to the shape of distributions where $f(x) \sim x^{-\alpha}$; as events get larger, their frequency gets smaller and vice versa. In popular usage, those events that are less frequent form the "long tail" of the distribution with events of a smaller size being ever more frequent. In these systems, which range from word frequencies to the sale of popular music, size is less important for each event is considered as unique [2]. In this sense, frequency is often taken as size.

In other systems such as those we deal with here, the larger the size of the event, the less frequent this is, as for example, in city-size or firm-size distributions. Thus, the usual practice is to transform the frequency distribution into its counter cumulative, which gives the rank of the object in question according to its size, the largest being rank 1, the second largest rank 2, and so on. Integrating $f(x)$ from some x to $x_{\max}$, we define the rank as $r(x) = F_{x \to x_{\max}}(x) \sim x^{-\alpha+1}$, which can then be transformed with respect to size x as $x \sim r(x)^{1/1-\alpha}$ [3]. Zipf [4] called this the rank-size rule, which is in its purest form when $\alpha = 2$, $x \sim r^{-1}$ implies that when $r = 1$, $x = x_{\max}$, and thus $x = x_{\max}r^{-1}$. Here, the long tail is sometimes referred to as the tail characterizing the smallest sizes, with the heavy or fat tail the one describing the largest (and often most significant) objects or events of which there are far fewer. We will adopt this definition henceforth.

The reason for referring to these distributions with respect to their tails is due to the fact that power laws are often used to best approximate certain portions of the distribution such as their fat tails. Although such scaling distributions represent our starting point, many distributions that define complex systems can only be so approximated and at the present time, it appears that lognormal distributions represent a wider generic class describing the way such systems are structured. These emerge directly from extremely simple models of competitive growth involving growth and decline as well as births and deaths of new system elements, where under certain conditions, power laws are generated [5–7].

In the sequel, we first define growth mechanisms based on models of competition that lead to systems whose elements change in size. These types of generic mechanism define population systems such as cities, buildings, and firms, which, in general, change in size over decades or centuries but are subject to a more volatile dynamics over very long time periods such as millennia. Our analysis shows that although the frequencies defining these distributions are extremely stable which with respect to their upper or heavy tails, are essentially scaling, there is considerably more volatility over shorter periods of time such as decades. To provide some sense of these dynamics, we introduce a powerful method for visualizing how these elements change relative to each other over time in the form of the rank clock [8]. After introducing the generic clock for different city-size distributions, we look at nonspatial systems such as firm sizes, and thence, one-off growth regimes as characterized in the construction of high buildings. Finally, we draw together the implications that these morphologies have for different growth regimes, thence, defining an agenda for further research.

COMPETITION AND THE SPACE–TIME DYNAMICS OF CITIES

The simplest growth process assumes no births or deaths but starts at time t with a fixed number n of elements of size $x_i(t)$. It grows each element (i) by applying a positive random growth ε_i to their existing size so that $x_i(t+1) = (1+\varepsilon_i)x_i(t)$. Simple experimentation for a fixed number of objects such as firms or cities, indicates that it is increasingly unlikely for all the objects to grow at the greatest rate and that ultimately one of these will dominate. In this process, objects will go up and down in size, but eventually one will dominate while some objects will ultimately become too small and will be deemed to have died. Generalizing the process to random proportionate growth generates distributions that are essentially lognormal as first illustrated by Gibrat [9]. The long tail of the lognormal might be considered the fat tail of the rank-size distribution and it is this that is often approximated by a power function. If this model is further generalized to repel objects from becoming too small, various researchers [6,7,10] have shown that a power law distribution can be generated over the whole range of sizes. This is one of the most parsimonious ways of generating a size distribution that scales and when adapted to a network where the size of the node grows using proportionate effect and adding network links from which such effects emanate, similar scaling distributions for the size of the hubs emerge [11]. Indeed, Barabasi's [12] network models of preferential attachment can be seen as a generalization of these growth models, all falling under the umbrella of a wider class of models that emphasize cumulative mutual advantage.

There are several types of growth process that can be modeled using this generic mechanism. In terms of cities, populations that are measured by their size can grow or decline. Populations grow from the smallest settlements to the point where they "become" cities and their growth is, thus, asymmet-

DOI 10.1002/cplx

ric; to be a large city, one must first be a small city. Because of difficulties of defining what a city is at the lowest population level but more because of the lack of data, we will only work with the largest cities, taking the top 100, 200, and so on, thus establishing immediately that the analysis and visualization is relative to the size of the system in terms of number of objects chosen. In such cases, "births" of cities occur when they enter the top ranks—say the top 100—and cities "die" when they leave this ranking. The random proportionate growth model is consistent with the widely established spatial hierarchy of cities, which was first established by Christaller [13], a contemporary statement of which is provided by Fujita et al. [14].

There are of course difficulties in defining the spatial extent of cities but generally these can be dealt with. It is much harder to ensure that firms are defined appropriately as mergers and acquisitions can destroy any size grouping. In terms of buildings, although these have generally got taller as cities have grown, buildings do not usually grow per se. They are usually constructed afresh and thence demolished, and only small fraction of buildings are modified and grown or reduced in size in situ. It is easier to see how firms and cities compete to increase in size than buildings. Nevertheless, bigger buildings usually do occur in places where there are already high buildings, and it is easy to consider the random proportionate growth model being adapted to take account of copycat-like behaviors as ever higher buildings are constructed. Then in the case of buildings, we might expect a rather different dynamics but in the period in question—the last 100 years or so—we will disregard the very small number of tall buildings that have been demolished.

We begin with a modest example involving the growth of cities in the Italian peninsula from the 14th to the 19th century [15,16].[1] The space–time dynamics is relatively uncomplicated in that the main cities—Bologna, Firenze, Genova, Milano, Napoli, Padova, Palermo, Roma, Venezia, and Verona—were established by the 14th century, and the period until Italian unification when our analysis ends (in 1861) did not see dramatic growth or radical shifts in the ranking of the key centers. We have a manageable seven time instants—1300, 1400, 1500, 1600, 1700, 1800, and 1861—for which we have population size data on the 555 towns that existed during the period, and we can thus rank the towns by population size and examine any changes in rank rather easily. This core of towns has remained in the top 15 by population size since the beginning of the Renaissance. In the subsequent analysis, we will only examine the rankings of the top 100 towns by size at each of the seven time instants and over the temporal period, only 195 distinct towns enter the analysis. Of course, some enter the top 100 and then leave, only to enter again by the end of the period, and our analysis cannot account for towns that move up and down the rankings during the intertemporal periods for which we do not have data. A total of 360 towns that exist in the dataset in fact never enter the top 100 and are thus really villages or at least settlements that never become significant. Although we are dealing with the top 100 towns by population size in each of the seven time periods, in the second period ($t = 1400$), the total number of towns falls to 95, whereas in all other periods, the number is slightly greater than 100 due to ties in the rank orders. In terms of the relative stability of population during this period, the total population of all 195 towns, which is 2.1 million in 1300 declines to 1.71 million by 1500 (probably due to the Black Death and war). It only recovers to 2.22 million in 1700, rising to 3.97 by 1861, indicating the early beginnings of political unification and industrialization.

This is a good case as it lets us introduce a number of tools for analyzing and visualizing space–time dynamics using a particularly tractable and easy-to-understand example. We first graph the seven rank-size relations as logarithmic transformations of $x_i \sim r(x_i)^{1/1-\alpha}$ as $\log x_i = \Phi - \beta\, r(x_i)$ where the slope β is equal to the scaling parameter $1/(1-\alpha)$. We show these distributions in Figure 1(a) where their similarity is even clearer when they are collapsed onto one another. There are various ways of estimating the scaling parameter. The most traditional that has most bias is to estimate β using ordinary least squares (OLS) from which we can then compute the scaling parameter as $\hat{\alpha}_{\mathrm{OLS}} = 1 + \beta^{-1}$. A less biased way is to use maximum likelihood adapted to power functions (see Newman, Ref. [17]), and this consists of solving the likelihood equation for each distribution as

$$\hat{\alpha}_{ml} = 1 + n\left[\sum_i \log \frac{x_i}{x_{\min}}\right]^{-1} \quad (1)$$

where $\hat{\alpha}_{ml}$ is an estimate of the scaling parameter, n the number of observations in the city dataset (which can vary for each time period), and $x_{\min}$ the minimum population defining the lower bound of each city-size distribution. Because such size distributions are more likely to be in their steady state in the fat or upper tail, then the lower tail should be truncated at some minimum [18]. We explore this in the software available for these visualizations and having estimated $\hat{\alpha}_{ml}$ for K varying minimum values x_k, we form the average as

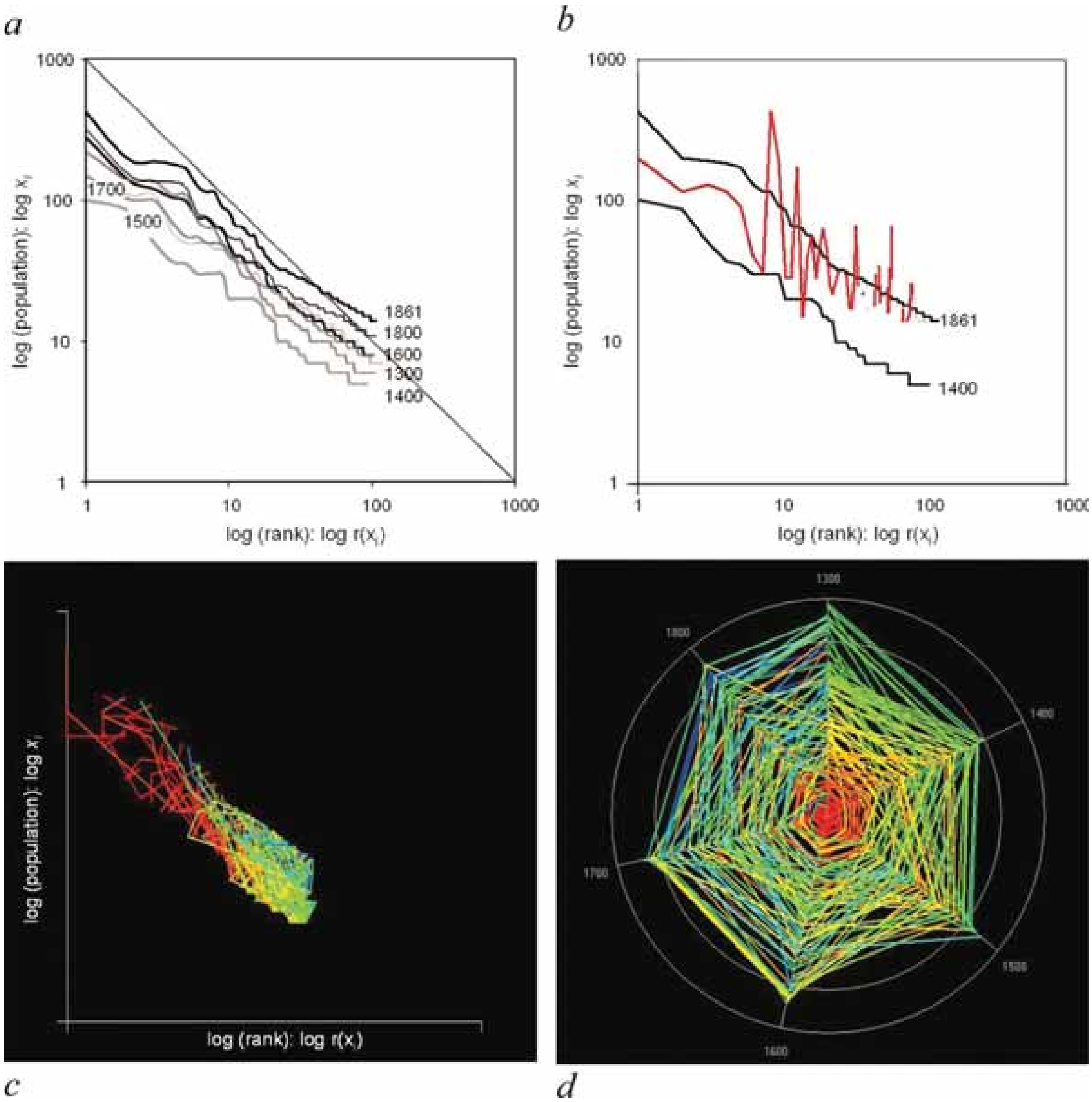
$$\overline{\alpha} = \frac{1}{K}\sum_k \left\{1 + n\left[\sum_i \log \frac{x_i}{x_k}\right]^{-1}\right\}. \quad (2)$$

In Table 1, we present these estimates from which we speculate that these values are close enough to suggest that there have been no major transitions in top-ranked cities during this 500-year period.

[1] *The data has been compiled by Paolo Malanima and is available at http://www.issm.cnr.it/.*

DOI 10.1002/cplx

FIGURE 1



Visualizing space–time dynamics in terms of rank shift. (a) Zipf plots for the seven time instants, (b) Shifts in ranks between 1400 and 1861, (c) Shifts in ranks of all towns throughout the seven time periods in rank space, and (d) the Rank Clock. The colors are chosen from red through blue according to the chronology of appearance of towns from 1300 and according to their rank. The first town, the top rank at 1300, is colored red; the last town is the last rank to enter at the latest time period and is colored blue. Animations of (c) and (d) can be seen at http://www.casa.ucl.ac.uk/complexity/.

VISUALIZING SCALING: THE RANK CLOCK

There are two obvious ways to further explore this space-time dynamics. First, we can examine the shift in ranks between two periods. For any two different time instants, the rank shift can be visualized by plotting one of the size distributions using the ranks associated with the other distribution. In Figure 1(b), we show this shift for the Italian city sizes, plotting the distributions in 1400 and 1861, where shift is based on plotting the 1400 sizes using the 1861 ranks. The picture is one of greater volatility than we

DOI 10.1002/cplx

TABLE 1

Estimates of Rank Size

Year	$\hat{\alpha}_{ml}$	$\overline{\alpha}$	$\hat{\alpha}_{OLS}$	r^2
1300	2.493	2.699	2.447	0.886
1400	2.836	2.696	2.393	0.851
1500	2.606	2.522	2.283	0.858
1600	2.500	2.438	2.262	0.873
1700	2.631	2.506	2.277	0.871
1800	2.698	2.594	2.403	0.893
1861	2.582	2.562	2.365	0.899

have seen so far. We can plot all the shifts for every time instant in rank space if we trace the rank and size of each city through their evolution as we do in Figure 1(c). Here, the colors are set as follows: the largest and most ancient town is colored red throughout and towns that are smaller and/or enter the top rank later are colored according to the spectrum red–orange–yellow–green to blue. We also use this color map for the rank clocks that we use to visualize these and other distributions introduced later. However, this visualization in rank space confuses the picture even more, whereas Figure 1(b) implies considerable shifts in rank and Figure 1(c) tends to play these down for the balance of color shows that the oldest and largest ranks from early in the evolution of the city system tends to remain in place throughout. What we need is something a little more visually intrusive to pull out any deviations that are significant, and to this end, we introduce the idea of the rank clock, which we show in Figure 1(d).

The rank clock focuses entirely on changes in the ranks with the trajectory of each object—in this example, a town or city—representing its rank order in a circular space. At any time, the top-ranked object is always at the center, the lowest rank at the circumference or the furthest point to the edge if the circumference has not been reached. Time is arrayed in a clockwise direction from the usual north-noon point of the clock. The time period over which the analysis takes place is marked by the circumference whose length is 2π, in short the complete clock. The clock is composed of distinct trajectories, five of which we visualize in Figure 2, although our more general quest is to compile different examples of these morphologies.

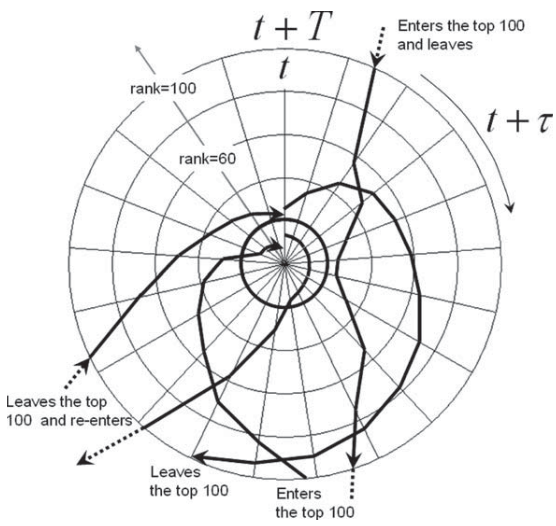
Objects that remain at the same rank are marked by exact circles around the clock, which close on themselves. Objects that rise inexorably in their rank, spiral into the clock while objects that decline systematically spiral out. Objects can, of course, enter and leave the clock as many times as possible, and there are obvious extreme cases: if an object were to enter at t, leave at $t + 1$, enter again at $t + 2$, and so on, then this pattern would simply appear as a dot every other time period at the rank where it entered. Note that, in general, we do not have the rank of the object before it enters the top ranks, so objects quite literally appear on the clock when they enter. If we did have these ranks, then we could compute their actual trajectories. One way of doing this would be to construct the clock for a much bigger system of objects and then reduce this to a lesser number of top ranks. In Figure 2, we now show an object that enters and then leaves to re-enter again and its opposite, one that leaves, re-enters, and leaves again. The pole of the clock is significant for in many scaling systems, one of the objects often remains the biggest. In city systems, this is quite common: for example, since 1500, Napoli has been at rank number 1 for the Italian urban system, and for the last 200 years, New York City has been at rank 1 in the United States, London at rank 1 in Great Britain, and so on. To compare two rank clocks, it must be assumed that the temporal circularity is directly comparable between examples and it is thus only relevant to do this, if one assumes that the temporal behaviors of two or more different systems are comparable within the sweep of the circumference. Moreover, systems with different number of objects—are usually not directly comparable.

From the clock in Figure 1(d), two features are immediately clear. First, the towns that are top ranked in the late Middle Ages and early Renaissance retain their dominance to the middle of the 19th century. This is clearly seen in the concentration of the color red around the pole of the clock. Second, volatility where colors diverge markedly from circularity occurs toward the edge of the clock. Here, we see little evidence of towns that are large and become small leaving the pole of the clock, spiraling out to the edge, or of towns that rise in rank dramatically spiraling into the clock. Unlike our later cases, there are only a handful of such examples in the Italian city system. Torino does not enter the picture (the top 100 ranks) until 1500 and then it rises swiftly to rank 4 by 1861. Siena, on the other hand, occupies rank 7 in 1300 and drops to rank 45 in 1861. It is hard to make these out on the clock in Figure 1(d), but the software that is available enables this to be illustrated quite easily. Readers are referred to http://www.casa.ucl.ac.uk/complexity/ where all these figures are illustrated in color, where animations of the order in which the towns emerge in time around the clock are presented, and where the software to explore these ideas can be downloaded.

The clock also focuses on other measures of change. Changes in rank from time period to time period are clearly seen as changes in the distance traveled around the clock as this kind of morphology is rooted in a circular geometry. For the rank $r_i(t)$ of each object i at time t, we can define an individual distance—a first-order difference—which can be plotted on the clock. We do not show this distance clock here [8] but distance is defined as

DOI 10.1002/cplx

FIGURE 2



Possible trajectories defining the morphology of the rank clock.

$$d_i(t) = |r_i(t) - r_i(t-1)|, \quad (3)$$

and aggregate distances can be defined over all cities for each time period as

$$d(t) = \sum_i |r_i(t) - r_i(t-1)|/n(t). \quad (4)$$

$n(t)$ is the number of cities in the distribution at each time period. If cities are unique in terms of measures of their size, then this might be set as $n(t) = 100$, $\forall t$ although in the Italian example because of ties in size, the number varies slightly at each time. This gives a measure of the average changes in rank that occur for all cities that remain in the set. In fact, second, third, fourth, and greater order changes can be computed if so required as can related measures of cumulative change such as

$$d(\tau) = \sum_{i,t=1,\ldots,\tau} |r_i(t) - r_i(t-1)|/n(t). \quad (5)$$

The average switch in ranks over all cities and all time periods, in this example, from 1300 to 1861, and over all 195 cities that appear in the top ranks, is

$$d = \sum_t d(t)/T. \quad (6)$$

where T is the total time over which the change takes place; in this example, it is 561 years. We show these measures for the Italian system in Table 2 where it is clear that on an average, a typical city shifts approximately 15 ranks (d) over 100 years, consistent with the shifts [$d(t)$] taking place over each 100-year interval, which range from 12 to 17.

Our last measure of space–time dynamics involves the speed at which cities enter or leave the top-ranked set of cities. It is easy enough to count the number of cities comprising the top set of ranks at time t, which are still in the new top set at some time later or earlier than t, say $t + \tau$ or $t - \tau$. We define this number as $L(t, t+\tau)$ for the cities that existed at time t in the top ranks and are still in the top ranks at $t + \tau$; and $L(t-\tau, t)$, the number of cities that were in the top ranks at time $t - \tau$ and are still in the top ranks at time t. A little reflection suggests symmetry of these counts, that is $L(t, t+\tau) = L(t-\tau, t)$. We can work out the average number of cities that perpetuate with respect to a given time period t as

$$L(t) = \sum_{m \neq t} L(t, m)/(n-1), \quad (7)$$

whereas the average number of cities that perpetuate from the top ranks at any time over all other time periods is

$$L = \sum_{\ell \neq t, m \neq \tau} L(\ell, m)/(n-1)^2. \quad (8)$$

These measures assume that the number of cities is fixed and need to be modified if these vary by time period as in the Italian example.

The half life is defined with respect to the entire time period T. As we do not have explicit functions that describe the process by which cities persist in the top ranks, we need to figure out these half lives by inspection and interpolation where the time intervals are usually not fine enough to compute these half lives exactly. We can do this for the whole series, of course, or we can do it for each time instant where it will vary. Essentially for each time t, we need to solve the equation $L(t,t)/2 = L(\tau, t) = L(t, \tau)$. Because no generic formal function for this persistence exists, we need to examine the matrix $L(t, t+\tau)$, but assuming exponential decay, it is likely that the number of cities in the top ranks will decline through time from the

TABLE 2

Distance Measures: Changes in Rank for All Cities over All Time Periods

Time Period	$d(t)$	$d(\tau)$
1300–1400	16.949	16.949
1400–1500	16.797	21.612
1500–1600	17.064	23.775
1600–1700	11.780	25.371
1700–1800	13.063	23.940
1800–1861	14.518	21.446
Average distance d	15.519	

DOI 10.1002/cplx

TABLE 3

Number of Cities $L(\ell,m)$ that Persist from Times ℓ to m

Time ℓ, m	1300	1400	1500	1600	1700	1800	1861
1300	*118*	81	80	71	70	67	65
1400	81	*94*	81	65	62	64	59
1500	80	81	*106*	82	76	73	70
1600	71	65	82	*104*	85	80	76
1700	70	62	76	85	*103*	89	79
1800	67	64	73	80	89	*107*	*86*
1861	65	59	70	76	79	86	109
$L(t)$	72	69	77	77	77	77	73
L	74						

point at which they are considered. If they do not, then this is evidence of extreme persistence and regularity in the system and the degree to which the half life approaches the maximum number of ranks is a measure of this stability.

Our Italian example is useful for the number of time periods is small enough to enable a casual examination of the matrix elements $L(t, t + \tau)$, which we present in Table 3. Although the maximum number of towns in any one distribution is 118, 195 different towns appear in the matrix from 1300 to 1861. In fact, there is a remarkable persistence of the core set of towns with an average of approximately 74 appearing to perpetuate throughout the period. In fact, for the remaining set of towns, there is considerable volatility in the rankings with substantial movements amongst the smaller towns, which continually and aggressively compete for pride of place. From Table 3, it is clear that the half life in all cases is perhaps a little longer than the 561 years over which these distributions of towns are observed, implying very little spatial restructuring and rather low levels of growth. For the 118 towns in 1300, some 65 or 55% of these still exist in the top 109 in 1861. An average half life of approximately 600 years would be a good guess, which is much larger than the other examples of city systems that we will now examine.

CITY SYSTEMS AT DIFFERENT SCALES

We now have a small arsenal of tools to explore the space-time dynamics of scaling systems and a brief summary is worthwhile. First, a good measure of stability with respect to size is the Zipf plot, specifically the scaling parameter α or its equivalent β, the degree of competition. When $\beta < 1$ or $\alpha > 2$, the larger objects in the distribution are closer in size to the smaller implying less competition and vice versa. Second, the best graphic is the rank clock for individual objects that can be traced in terms of their trajectories and the whole set of objects can be visualized as a morphology. The two measures of difference, the first based on distance and the second based on guesstimates of the half life, are useful in measuring actual shifts in rank. The distance of course depends on the time periods in question and represents the shift, which takes place in number of ranks over one standard time period. For the Italian data, this is 100 years (with 61 years, appropriately adjusted for the statistics, in the seventh period). Examining Table 2, we see that the average rank shift is approximately 15 over 100 years, and this suggests that over 500 years, the shift would be approximately 75. The second statistic is the half life, which we guessed at about 600 years where approximately 50–60 would shift out of the top ranks entirely during this period. As cities enter and leave the rankings, simply taking the average distance d and scaling it by 5 or 6 will give an overestimate of the shift but it suggests that judicious use of half lives and distance measures provide a rich picture of these dynamics The cumulative distance $d(\tau)$ is probably a better measure of shift because this takes into account cities that enter and re-enter the rankings as the system evolves.

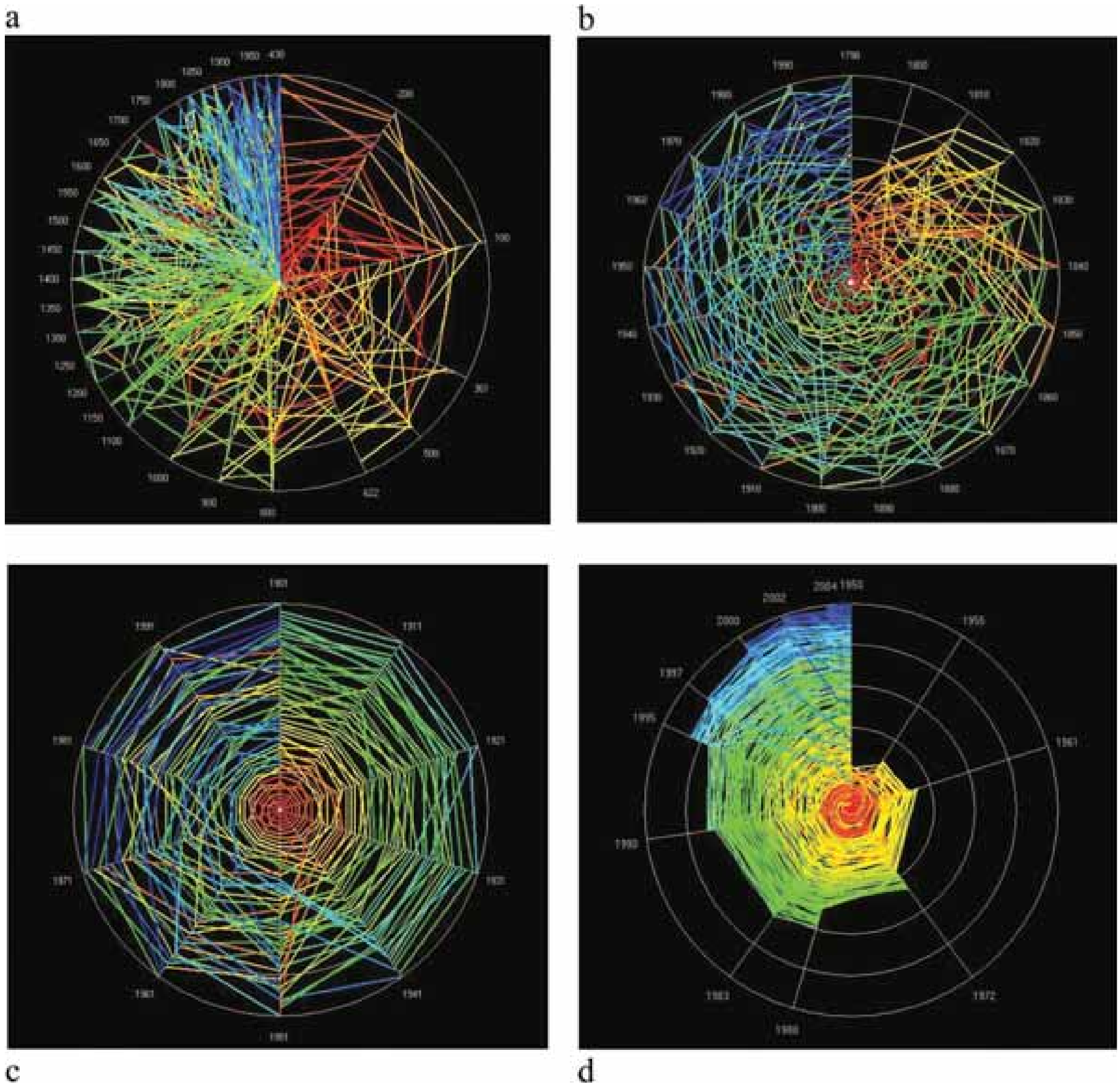
We are now in a position to select and track some very different city systems, which exist at very different scales. We have analyzed city-size distributions for the top 50 cities in the world from 430 B.C.E. to 2000 A. D. with variable time intervals using Chandler's [19] database, which is the largest scale and longest time period that we have dealt with. At the next scale down, the continental, we have examined changes in rank size for the top 100 cities in the United States from 1790 to 2000 at 10-year intervals (from the US Census Bureau [20]), then for the country level, the top 50 in Great Britain from 1901 to 2001 also at 10-year intervals,[2] and finally, for a much smaller region, the top 172 towns in Israel from 1950 to 2005 with variable time intervals [21, 22]. Israel is more like a metropolitan region with respect to scale although each of these examples has very specific geopolitical and cultural characteristics, which add to the variety of this selection.

For each of these systems, we need to choose either a fixed set of top ranks or simply let the number of cities in the wider database condition the number of cities being ranked at any time instant. For example, in the Israeli example, there are 172 towns, which exist at some time over the 55-year period from 1950 to 2005. In 1950, there are 34 towns defined, while by 2005, there are 164. However, eight of the towns appear at some point in the rankings for the 13 time instants, entering and then leaving the space by 2005. In this sense, we do not constrain the Israeli example whereas for the GB example, we have 458 towns that do not change throughout the entire 100-year

[2] *Census Dissemination Unit (CDU) http://census.ac.uk/cdu/, accessed on 06/01/10.*

DOI 10.1002/cplx

FIGURE 3



Rank clocks for the world (a), United States (b), Great Britain (c), and Israel (d).

period. However, we only examine the top 50 towns at any of the 10 time instants, and in this case, 70 of the 458 towns are considered. We choose this number so that we can clarify the dynamics, notwithstanding the obvious point that for the dynamics will change as we select more towns from the total set. In fact, in illustrating these dynamics using the static clock on the printed page, visualization is clearest when the number of towns is restricted to about 100 or less. For visual analysis of many more objects, zoom and pan facilities as well as animation are required (see http://www.casa.ucl.ac.uk/complexity/).

The world system of cities is the most volatile. There are no cities in the top 50 in 430 B.C.E. that are still there by 2000, and there are only six in the top 50 from the Fall of Constantinople in 1453. The half life of the original set of cities is approximately 200 years, which reduces to about 100 years by the 20th century. The rise and fall of civilizations, particularly Greece and Rome, the coming of the Dark Ages, the parallel growth and decline and growth again of China, and the explosion of cities in the developing world can all be gleaned from the trajectories of cities in this database. The morphology of the rank clock is shown in Figure 3(a) and it is quite clear that

DOI 10.1002/cplx

TABLE 4

Comparative Measures of Rank Shift for Five City Systems

City Systems	Number n	Time, T (yrs)	$t \rightarrow t+1$ (yrs)	Min $\hat{\alpha}_{ml}$	Max $\hat{\alpha}_{ml}$	Min r^2	Max r^2	Min $d(t)$	Max $d(t)$	d	L
World cities	47–50 (390)	2430	80 (25–260)	2.127	3.236	0.906	0.945	4.914	10.973	7.785	14
US cities	24–100 (266)	210	10	1.952	2.674	0.911	0.944	2.242	7.559	4.667	25
GB cities	50 (458)	100	10	2.746	4.062	0.948	0.951	1.941	9.857	4.220	41
Israeli cities	34–164 (172)	55	4 (1–11)	1.748	2.035	0.782	0.850	1.578	5.917	4.032	83
Italian cities	94–118 (555)	561	80 (61–100)	2.493	2.836	0.851	0.899	11.780	17.064	15.519	74

there is no sense in which there is a group of persistent core cities as in the Italian example. The longest lived example of a city in the top 50 is Suzhou in China, which exists for 2158 years of the 2430 years covered and even this city is no longer in the top 50 (although it is growing fast and could re-enter the list unless absorbed in Greater Shanghai). The half life of cities in the world system is clearly reducing fast and is now no more than 75 years, falling at the rate of 20 years for every additional 25 years of time. The system nearest this rate of change is the US system, which we plot in Figure 3(b), that displays all the features of the generic example in Figure 2. New York City remains at the center of the clock unchanging since 1790, while cities like Chicago, Los Angeles, and Houston spiral into the top ranks as the population has diffused to the Midwest, California, and the South West over the last 150 years. Cities, such as Charleston, in the old colonial east spiral out of the clock as the United States begins to industrialize from the mid 19th century on.

Our two other examples are as different again. In GB, the half life is about twice that of the United States and the current world system, at approximately 150 years, and it shows little sign of changing. This is because by 1901, all the key settlements were established—the core cities had developed during the 19th century, and suburbanization in the 20th has not made much impression on the overall pattern of urban development. The clock is shown in Figure 3(c) where the core stands out as in the Italian system. However, the Israeli system is quite different in that it is almost impossible to guess the half life as new settlements have been established and grown continually during the last half century during the time when the country has developed and consolidated. Thus, the half life is effectively an order of magnitude longer than the time during which development has taken place. A crude guess would be between 50 and 70 years but in a growth situation where all 34 of the original settlements in 1950 are still in the top ranks in 2005, this is hard to assess for the number of settlements dropping out of these ranks is very small. The half life is really a forward looking measure notwithstanding the symmetry of the flow matrix $L(\ell, m)$. Better measures involve distances, which we will now discuss. The rank clock for the Israeli system is shown in Figure 3(d).

The fit of the rank-size functions to the data for all city systems over all time periods is shown in Table 4, and the proportion of the variance explained ranges from 85 to 95%. The scaling parameters all tend to be greater than two suggesting that the largest cities tend to be less dominant than in the pure Zipf case ($\alpha \sim 2$). The average distance shift in ranks must be interpreted in the light of the average time periods over which the size trajectories are observed. Interestingly, since industrialization for the US, GB and Israeli city systems, these show shifts of rank between 4 and 8 over 10-year periods, whereas the shift in rank for the world system is approximately eight over periods of some 80 years. This distance measure is much less reliable as the time periods vary so massively, whereas in the case of the Italian system, the shift is about 16 over 80 years, which pro rata gives a shift of some five ranks for every 10 years, similar in fact to the other National City systems.

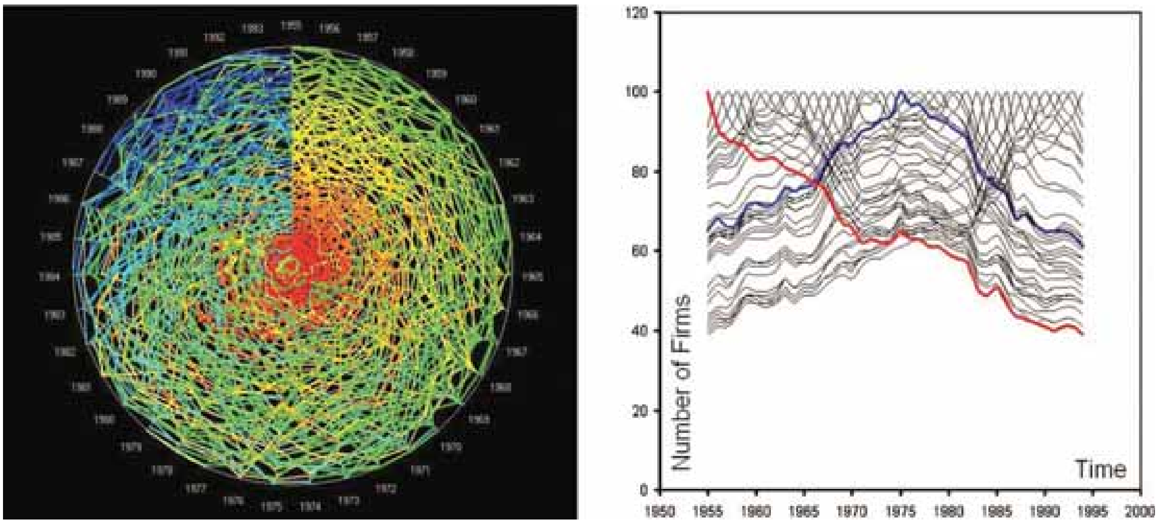
The last point we need to make is based on a visual comparison of the rank clocks. Clearly, the US and Israeli clocks begin with less than 100 top cities as these systems do not have 100 cities at their start. The US clock then reaches 100 in 1840 and then we keep these top ranks stable. The Israeli clock illustrates the entire growth trajectory with no constraints on numbers. Both these clocks show growth with the Israeli example showing the persistence of core cities and massive growth of new ones. The US system has much less of a focus on the core cities as is consistent with the rapid diffusion of large cities in the west and south. The British picture is one of more classic slow growth with many trajectories showing circularity, that is, ranks remaining similar over time and the core being maintained. The world system is by far the least stable in that cities of the ancient era are clearly quite different from those of the middle ages, the Chinese empire, and the modern world.

DISAGGREGATE POPULATIONS: FIRM SIZES AND BUILDING HEIGHTS

Our last two examples involve rather different human systems. The size distributions of firms from 1955 to 1994 for the United States are taken from the Fortune 500 data and the size distributions for tall buildings (skyscrapers) for the

DOI 10.1002/cplx

FIGURE 4



(a) The Fortune 100 rank clock and (b) the persistence decline of firms by rank 1955–1994.

world and for New York City from the Emporis buildings database.[3] In terms of firms, we have focused not on all 500 but on 100 as by now it must be clear that as the number of objects increases, the rank clock is more and more impressionistic in that it requires zoom capability to see individual trajectories. Firms can be ranked, as can cities, by any measure of their size, and in this case, we have revenue and the profit/earnings ratios, which give measures of how well the firm is doing. We will not show this latter set here (however, see Ref. [22]), but in Figure 4(a), we plot the revenue clock, and it is immediately clear that there is large but regular volatility in the rankings of firms.

Essentially, a core of firms does stand out but there is considerable erosion in their ranks. The rate of this erosion is rather regular and this is best seen in Figure 4(b), which is a visual plot of $L(\ell, m)$ showing the number of firms that stay in the top 100 rankings as time elapses. We can measure the half life from this, which is about 25 years, and Figure 4(b) marks out the firms that stay in the top rankings from the beginning year 1955 and the firms that enter or leave the rankings pivoted around the mid-year 1975. Figure 5 shows the trajectories of six of these firms where quite clearly steel and heavy industry such as rubber spiral out of the rankings while high tech (e.g., IBM) spiral in. American car manufacturers like American Motors have mixed fortunes but the largest such as Chrysler more or less maintain their rank with General Motors (not shown) ranked as number one throughout the 40 years. In Table5, the rank-size relations show extremely good fits with r^2 near 99% for all distributions.

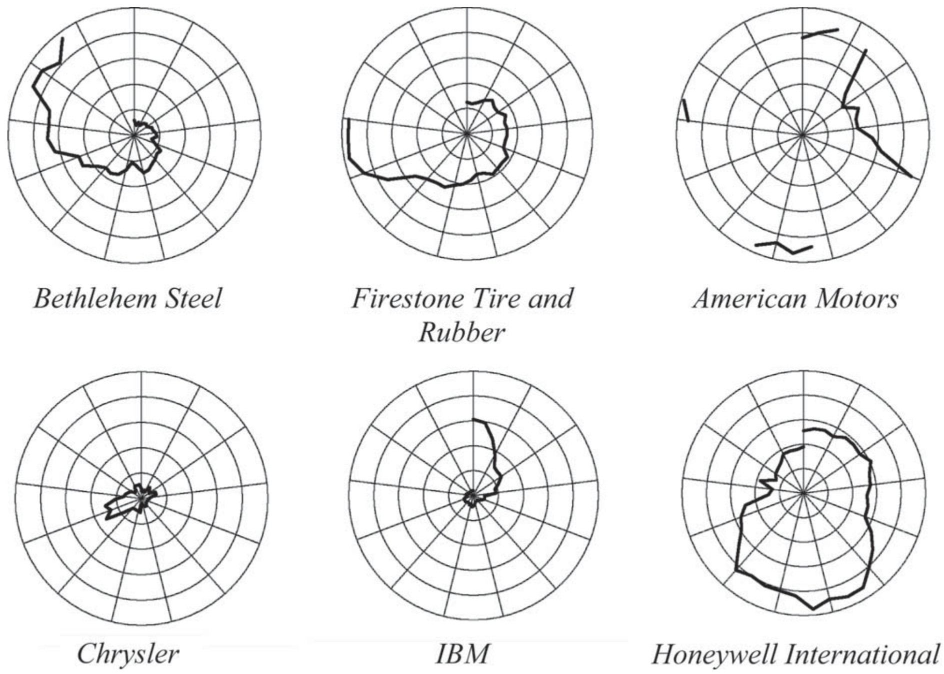
Our second example is dramatically different. All the tall buildings greater than 12 stories or 40 metres in height in our datasets have not changed in height since their construction. Only in places of very rapid growth such as Hong Kong are tall buildings now being systematically demolished and rebuilt and it might be argued that such instant changes in rank, which such demolition and reconstruction occasions is tantamount to the construction of new buildings. This means that buildings do not rise in rank, and they simply appear at a particular rank and never get any higher than their original rank.

The first high building in each dataset is taken as being constructed in 1909 although there are skyscrapers built before then. In New York City, there is a rapid increase in the number of such buildings being constructed in the 1920s and 1930s. Because of ties in height, the top 100 contain some 119 in 1916, which means that the particular rank clock shown in Figure 6(a) reaches out to 120 on its radius rather than 100 before falling back to 100 or 101. The clock

[3] *Fortune 500 data from 1955 to 2005 is available for each year from the CNN Money website. http://money.cnn.com/magazines/fortune/fortune500_archive/full/1955/index.html. We only use the data from 1955 to 1994 because the series was redefined in 1995. The high buildings data is from the Emporis global database http://www.emporis.com/en/bu/sk/.*

DOI 10.1002/cplx

FIGURE 5



Individual rank trajectories for selected Fortune 500 firms 1955–1994.

is based on a complete spiralling out and downward of buildings from the time they enter the top ranks. The early buildings before the late 1920s are colored red, orange, and yellow, and these are blanked out by the flurry of construction shown in green color in the early 1930s. Some of these buildings stay in the top 100 for the rest of the time period, but in general, later buildings are greater in height and to produce a clearer picture, it is necessary to zoom into the clock (see at http://www.casa.ucl.ac.uk/complexity/) and explore these trajectories. Figure 6(b) shows the clock for the world data where it is quite clear that globally, some high buildings constructed in earlier eras persist at high ranks into recent times. This picture of growth is considerably later than in the New York City case due to high building diffusing globally as countries in the developing world have grown. If one takes a cross section in these clocks at any time, this gives a picture of the time when a building was constructed and its rank, and this needs to be compared with the persistence matrix $L(\ell, m)$.

Numerical statistics are presented for these two examples in Table 5. The rank-size functions for New York City fit with reasonably good approximations as the r^2 statistics show with later distributions having better fits. The same is true of the world data although the performance earlier in the time series is not as good. One issue that makes these distributions very different is the degree of competition as reflected in the scal-

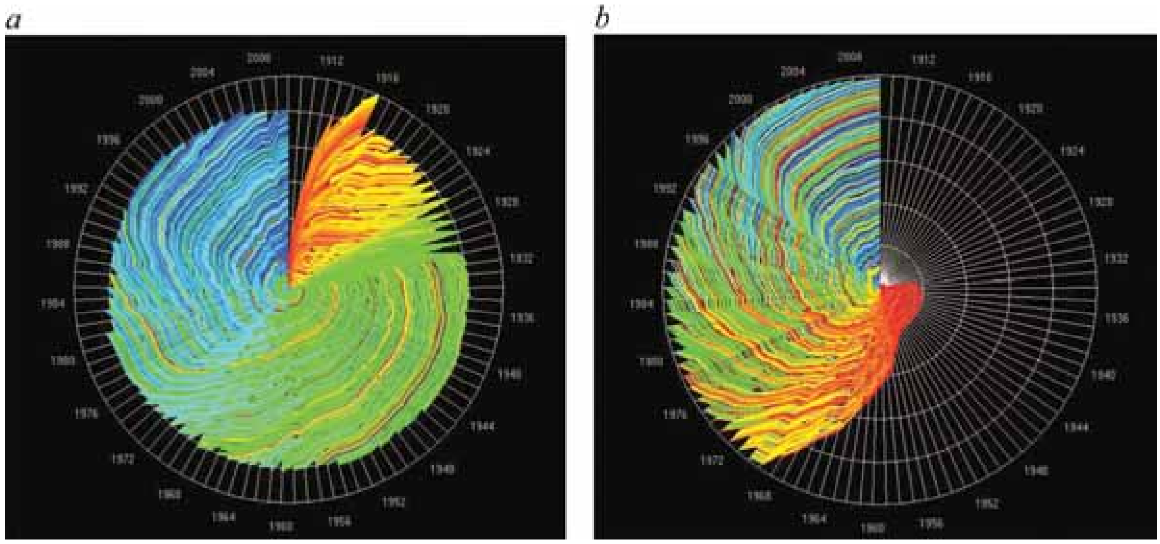
TABLE 5

Comparative Measures of Rank Shift for Firms and Tall Buildings

Scaling Systems	Number n	Time, T (yrs)	$t \rightarrow t+1$ (yrs)	Min $\hat{\alpha}_{ml}$	Max $\hat{\alpha}_{ml}$	Min r^2	Max r^2	Min $d(t)$	Max $d(t)$	d	L
Fortune 100 Firms	100 (343)	41	Yearly	2.205	2.739	0.987	0.995	3.397	7.988	5.158	56
NY City Skyscrapers	12–119 (516)	101	Yearly	3.048	6.259	0.873	0.959	–	–	–	39
World Skyscrapers	1–101 (500)	101	Yearly	3.997	7.462	0.627	0.962	–	–	–	19

DOI 10.1002/cplx

FIGURE 6



Rank clocks of the top 100 high buildings in the New York City (a) and the world (b) from 1909 until 2010.

ing parameter α. The values are much higher than two meaning that the slope of the rank-size curves are considerably flatter than the pure Zipf case where $\beta = -1$. This is of interest not only that it suggests much less competition between the construction of high buildings in an intraurban context but also that the growth dynamics is quite different from that which characterizes cities and firms, which grow in the biological sense due to accretion. In terms of the distance statistics, these are tricky to compute due to the fact that all the time periods are not distinct; in earlier periods, buildings persist but new ones are not always constructed at later time periods, and thus, ranks can remain the same. Moreover, the distance measures all point in one direction—downward in terms of rank shift. This suggests again that we need some modifications to these visualizations to account for systems where objects do not grow per se but do lose their position in the rankings due to the appearance (growth) of other objects.

NEXT STEPS AND FUTURE RESEARCH

Scaling systems involving human populations clearly display a form of regularity at the macro level, which masks dynamic volatility at the micro. We measure this by the extent to which objects change their position in relation to one another through their rank-size distributions. Yet, we do not have a detailed understanding of the way these dynamics play out but we do know that competition between these various elements is intrinsic to the way they capture growth from one another. In fact, in city systems, traditional theories where central places grow from the bottom up, gradually deriving their functions as they get larger and serving ever large populations, go a long way to explaining how cities scale at the macro level [14]. When it comes to the growth of firms, the logic is more complicated by the process of mergers and acquisitions whereas in the case of systems that are manufactured such as buildings, a very different substrata of dynamics is at work, dictated by the way developers define the need for high buildings, much complicated by the construction process and by investment decisions.

Our visualization of these changes suggests that different scaling growth regimes display different morphologies, and we have made a start at classifying these as rank clocks. However, we still need much better methods for making statistical comparisons between such systems. We need to pursue at least three lines of inquiry: first, to develop a more robust set of space−time statistics for scaling systems, which establish comparisons between different sizes of system; second, to generate network equivalents of such scaling systems that will add a richness to the analysis; and last but not least, we need more examples at different spatial scales where we can draw clear links between cities, buildings, and firms, which are all manifestations of how populations agglomerate in achieving economies of scale.

DOI 10.1002/cplx

References

1. Pumain, D., Ed. Hierarchy in the Natural and Social Sciences; Springer: Dordrecht, The Netherlands, 2006.
2. Anderson, C. The Long Tail: How Endless Choice is Creating Unlimited Demand; Random House: New York, 2006.
3. Adamic, L.A. Zipf, power-laws, and Pareto—A ranking tutorial. Information Dynamics Lab, HP Labs, Palo Alto, CA. Available at http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html, 2002. Accessed on April 1, 2010.
4. Zipf, G.K. Human Behavior and the Principle of Least Effort; Addison-Wesley: Cambridge, MA, 1949.
5. Simon, H.A. On a class of skew distribution functions. Biometrika 1955, 42, 425–440.
6. Gabaix, X. Zipf's law for cities: An explanation. Q J Econ 1999, 114, 739–767.
7. Blank, A.; Solomon, S. Power laws in cities population, financial markets and internet sites: Scaling and systems with a variable number of components. Physica A 2000, 287, 279–288.
8. Batty, M. Rank clocks. Nature 2006, 444, 592–596.
9. Gibrat, R. Les Inegalites Economiques; Librarie du Recueil: Sirey, Paris, 1931.
10. Sornette, D.; Cont, R. Convergent multiplicative processes repelled from zero: Power laws and truncated power laws. J Phys I (Paris) 1997, 7, 431–444.
11. Batty, M. Hierarchy in cities and city systems. In: Pumain, D., Ed. Hierarchy in the Natural and Social Sciences; Springer: Dordrecht, The Netherlands, 2006, pp 143–168.
12. Barabasi, A-L. The origin of bursts and heavy tails in human dynamics. Nature 2005, 435, 207–211.
13. Christaller, W. Central Places in Southern Germany; Prentice Hall: Englewood Cliffs, NJ, 1933; 1966 [translation].
14. Fujita, M.; Krugman, P.; Venables, A.J. The Spatial Economy: Cities, Regions, and International Trade; The MIT Press: Cambridge, MA, 1999.
15. Bosker, M.; Brakman, S.; Garretsen, H.; De Jong, H.; Schramm, M. The development of cities in Italy, 1300–1861. CESifo Working Paper No. 1893, 2007, Munich: DE. Available at http://www.cesifo.de. Accessed on April 1, 2010.
16. Malanima, P. Italian cities 1300–1800: A quantitative approach. Rivista di Storia Economica 1998, 14, 91–126.
17. Newman, M.E.J. Power laws, Pareto distributions and Zipf's law. Contemp Phys 2005, 46, 323–351.
18. Clauset, A.; Shalizi, C.R.; Newman, M.E.J. Power-law distributions in empirical data. SIAM Review 2009, 51, 661–703.
19. Chandler, T. Four Thousand Years of Urban Growth: An Historical Census; Edward Mellon: Lampeter, UK, 1987.
20. Gibson, C. Population of the 100 largest cities and other urban places in the United States: 1790 to 1990. Population Division Paper 27, 1998; US Bureau of the Census: Washington DC.
21. Benguigui, L.; Blumenfeld-Lieberthal, E.; Batty, M. Macro and micro dynamics of the city-size distribution: The case of Israel. In: Reggiani, A.; Nijkamp, P., Eds. Complexity and Spatial Networks: In Search of Simplicity; Springer-Verlag: Heidelberg, DE, pp 33–49, 2009.
22. Batty, M. Visualizing Creative Destruction, CASA Working Paper 112, 2007. University College London. Available at http://www.casa.ucl.ac.uk/working_papers/paper112.pdf. Accessed on April 1, 2010.

DOI 10.1002/cplx