Generative Historical Model of City Size Hierarchies: 430 BCE – 2005 *

Douglas R. White[1,2] Nataša Kejžar[2,5] Constantino Tsallis[2,3]
Céline Rozenblat[4]

December 4, 2005

[1] *Institute of Mathematical Behavioral Sciences, UC Irvine*
[2] *Santa Fe Institute – 1399 Hyde Park Road, Santa Fe, NM 87501, USA*
[3] *Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro – Rua Xavier Sigaud 150 22290-180 Rio de Janeiro-RJ, Brazil*
[4] *Université de Montpelier, Maison de la Géographie*
[5] *Faculty of Social Sciences, University of Ljubljana – Kardeljeva ploščad 5, 1000 Ljubljana, Slovenia*

1 Introduction

Despite the enormous changes in city size and differentiation of functions over time and space, and discontinuities in growth processes, key spatio-temporal and distributional processes shaping city sizes are often assumed to remain invariant. We demonstrate in this article that most of the facile assumptions about such invariance, that is, constancy over long historical periods, are unsupported when comparisons are made concerning city size distributions. We examine four aspects of the problem: the portion of the size distributions that are thought to be largely invariant (with exceptions for abnormal growth); how the city samples are constructed; the shapes of the mathematical functions that fit these distributions; and the historical variation in fitted constants of these functions. First, many geographers dispute Zipf's rank-size "law" empirically because the Zipfian is only satisfied in some cases for the tails of urban distributions, and there is a notorious deviation from an urban size power law when cities with smaller populations are considered (Malacarne, Mendes, and Lenzi 2002:2). Cities in centrally planned polities such as Russia and China also fail to follow the Zipfian (Marsili and Zang 2004:1). Second, these distributions vary as the sampling boundaries are changed, so that some cities might appear unusually large simply because the sampling region is drawn too narrowly. Third, others dispute Zipf's and other scale-free laws of the distributions because they are not explanations but merely mathematical descriptions, and Rapoport (1978:847) cautions that just about any group of objects arranged according to size will fit some monotonically decreasing curve. Carroll (1979) notes that in addition to the power-law distribution, the lognormal, Pareto and Yule distributions all have long tails, like the Zipfian.[1] Simon (1995) believes that the rank-size portion of a skew distribution results from any growth process at equilibrium in which size is initially random and growth is proportional to current size (see also Gabaix 1999), and Carroll also notes that this result also requires an assumption of a constant and steady rate of new entrants into the urban system. None of these conditions can be assumed historically. In order to address the fourth aspect of questions of distributional invariance, we examine all of the most

*This paper is the outcome of collaborations involving the Santa Fe Institute and its International Program, and the University of Ljubljana, Faculty of Social Sciences, and EU project ISCOM, Information Society as Complex System.

[1] The CDF for probability $P(N \geq n) \sim n^{-\beta}$, of N cities having n or more inhabitants as a Pareto city size distribution with an exponent β, translates to the rank $r(N = n)$ of the Nth city having n inhabitants, plotted with x and y axes reversed, where the convergence of $r(N = n) = \int_{-n}^{+\infty} p(n')\delta n' = n^{-1/\alpha}$ for binned intervals n' as they become infinitesimal, and $\beta = a + 1$. The special case of Zipf's law applies where $\alpha \sim 1$, in which case $\beta \sim 2$ in the equivalent Pareto law.

populated world cities in each of 29 historical periods and we independently fit the size distributions in each period to a single mathematical function that is motivated to fit, unlike the Zipfian, the full range of the size distribution, from smallest to largest cities. Our choice of function incorporates the intuition that, because city growth is more likely to accelerate in the upper part of the distribution, the function will more closely approximate a power law in the region of the larger sizes, even if the slope is not invariant. This choice of a two-parameter function (like the Zipfian or power-law distributions), if it is well fitted to each body of data from distinct historical periods, ought to address the question of whether there is invariance, on a scale that is not affected by the sampling problem, of city size distributions over time. If there is not invariance, the question to examine is whether the variation is random or autocorrelated, and if the latter, whether the historical variations are indicative of slow fluctuations or sharper breaks. Finally, if historical periodization is indicated by our results, we will want to examine whether there are clear historical correlates associated with these periods.

Physicists Marsili and Zhang (2004) are among those who have looked at city size as a function of interacting individuals and features that organize interaction. They start from a model of a reservoir of population with a fixed transition probability to becoming a city dweller and a small fixed probability that a new city will be formed, and model the case in which city growth is linear rather than power-law (superlinear). They then introduce a variant assumption (Zanette and Manrubia 1997) that processes governing urban growth are not independent, and show that modeling the effects of individual decisions that influence one another through network interaction generates a Zipfian effect (the transition-to-city rates for cities of size m now shift from linear in m to a mixture of linear and quadratic effects, reflecting a mix of individual decisions some of which are independent and some nonindependent due to network interaction). In this model there is a city transition size m_0 at which Zipf's law begins to hold: for a city with $m \ll m_0$ growth is constant while for $m \gg m_0$ growth is proportional to size, and follows a Zipfian. Their intuition for this model accords with Zipf (1949), for whom the scaling law is valid only for large cities.

Still, as Krugman (1996) notes, there has been no agreed-upon explanation for the rank-size law of cities. Various models deal with hierarchical organization in the fractal geometry of area distributions around cities (Batty and Longley 1994), how cities partition the plane (Malescio and Stanley 2004), how transport corridors are organized (Carvalho, Iida and Penn 2003), and city perimeter or area distribution of systems of cities as aspects of city morphology (Makse, Havlin and Stanley 1995). Lobos (2005) finds 17 indicators of the employment functionality of U.S. metropolitan areas in 2000 (e.g., construction, trade, education) that scale by a power law with city size.

We present a generative historical model of city size hierarchies that follows the approach of Malacarne, Mendes and Lenzi (2002), who use Mandelbrot's (1977) generalization of Zipf's distribution to deal with fractality by a measure of the extent to which the long tails of various distributions show anomalous decay. This model deals with a broad class of hierarchically organized and self-scaling processes and satisfies our requirements for a suitably general function with which to model city size distributions historically. It also subsumes the mathematical model of anomalous decay and fractality under a q-exponential distribution (Tsallis 1988, Boon and Tsallis 2005, White, Kejžar, Tsallis and Rozenblat 2006).

Malacarne et al. use the method of generalized monolog plots based on the generalized q-logarithmic function $\ln_q$ to estimate the parameter q, which measures an asymptotic slope $\alpha = 1/(q-1)$, as the function approaches a power law, for the larger city sizes. In this sense the q-exponential function subsumes a Pareto distribution for the upper city sizes, which in the special case of $\beta \sim 2$ also corresponds to the Zipfian. The generalized logarithmic function $\ln_q$ is well suited for studying the distributions of city sizes because it is somewhere between and inclusive of the distributions typical of small towns (constant growth) and those of large cities (proportional growth). In the derivation of Malacarne et al., q represents the mix between growth processes that do not depend on interaction (i.e., probabilistically independent, as with individual decisions in the Marsili and Zhang city growth model), and those that do and which will thereby generate power-law distributions. In the first case, $q = 1$, and because the constant rates of growth (or decline) are the result of statistically independent processes, this case fits a distribution that corresponds to a Boltzmann-Gibbs entropy distribution. If this were the case over successive historical periods, the constant rate of city growth would be reflected in changes through time in the scale parameter κ introduced shortly in the full statement of the q-exponential distribution. The second case is approximated as $q \gg 1$, or more generally, as q approaches 2 or even goes to infinity in the case of a gap between ordinary and a "primate" city. When $q > 1$, we are in the domain of a mathematical model of generalized entropy designed for mixes of independent and nonindependent processes – q-entropy (Tsallis 1988) – that may include network-dependent interactions (White, Kejžar, Tsallis and Rozenblat 2006). Thus, the Zipfian distribution for large cities becomes a special case subsumed under a more general distribution for cities of any size with a parameter q, relevant to urban hierarchy theory and not sensitive to size but to region or historical period, along with a parameter κ that is relevant to historical change in the scale or total populations of differently sized cities in an urban hierarchy. The latter parameter is not sensitive to different size cutoffs for the set of largest cities examined.

As a test of the applicability of q-entropic properties of city sizes, Malacarne, Mendes and Lenzi (2002:2) show that both the U.S. and Brazilian city sizes, down to the smallest cities, show a linear fit in the cumulative population distribution for city size rankings for which $q = 1.7$ and $R^2 = 0.99$ in both cases, and thus the asymptotic Pareto or CDF power-law slope $\beta = 1/(q-1) = 1.4$ with only three deviant "primate" cities (New York, Rio, and São Paolo) and no deviant smaller cities. These results (like $\alpha = 0.4$ in a PDF power-law) represent a mix of Zipfian and constant growth even at the upper city sizes.

2 Data

Data on largest city sizes for each of 29 periods from 430 BCE to 1950 from Chandler and Fox (1974) were homogeneited and georeferenced by Céline Rozenblat.[2] More large cities appear over time in these data but some of the smaller city-size bins were dropped in Chandler's coding so as to keep relatively constant the number of largest cities coded in each time period. Rozenblat also provided U.N. data for the largest cities for each of 12 periods from 1950 to 2005. For cities over a half-million, comparisons of the 1950 data from the two datasets differed by a scalar multiple of 1.182 (indicating that Chandler was consistently somewhat more conservative than the U.N. in estimating larger Metropolitan area populations), with $R^2 = .9985$. Below a half-million size for 1950, up to the size of two million for 2005, we deleted smaller cities from the U.N. data because they were undersampled. Cities were sorted by size and the sorted stacks of city sizes compiled in a single file prior to log-binning by size.

We calculated cumulative populations in cities of 50K up to 25.6 million in multiples of $\sqrt[3]{2}$. By taking cumulative populations we increase the accuracy of curve-fitting and make our graphs easier to interpret visually because the city-size curves tend to rise with time. The binning procedure resulted in distributions of number of people in cities above the minimum for the each of the 28 size bins. In addition, we counted the number of cities in each log-binned interval. If the largest bin containing a city was not full of cities with sizes up to the upper limit of the bin then the bin was dropped from analysis to avoid the effect of unequal bin sampling. Similarly, we dropped cells at the lower end of the binning scale for a given time period if they were not filled to their lower size limit. Because only those sizes from 200K upward were coded in all the files from 430 BCE to 1950, city-size averages as a measure of population fluctuation were computed for sizes above 200K, with missing data after 1950 in the lower size bins estimated from the q-exponential model.

Issues of data quality include the changing reliance on archaeological, historical, and census data. The earliest census data come from Han China. Censuses were taken for many world cities by 1700, with data quality increasing after 1800. Even in the earliest of our time periods, however, there are estimates considered to be reasonably reliable from archaeological as well as historical sources.

3 Methods and results

As described in the introduction we follow the modelling of Malacarne, Mendes and Lenzi, supplemented by other q-exponential methods but we model the Pareto CDF rather than the Zipfian rank-size CDF. We used

[2] Data were provided in text file with columns for city names, empires, city population, latitude and longitude.

the q-exponential function to fit the data of binned city size distributions from the coding procedures:

$$y(s) = y(0)e_q^{-s/\kappa} \tag{1}$$

$y(s)$ denotes cumulative number of people in cities of size s or larger. Sizes of the cities are expressed in 10^5. The q-exponential function is defined as follows (Gell-Mann, Tsallis 2004):

$$e_q^x = \Big[1 + (1-q)x\Big]^{1/(1-q)} \qquad (e_1^x = e^x) \tag{2}$$

if $1 + (1-q)x > 0$, and zero otherwise. This reduces to the usual exponential function when $q = 1$, but in the limit $x \to \infty$ it asymptotically approaches a power law when $q > 1$.

We used two different approaches to fit the data.

- The Gauss-Newton iterative algorithm for nonlinear least-squares (NLS) estimates model parameters q and κ and determines the extrapolated value $y(0)$ – the characteristic number of people. [3] Due to noise and lack of the data in some of the earlier time periods the Gauss-Newton algorithm did not always converge or the convergence was not appropriate because the estimate of q was smaller than 1. [4]

- For the second approach we used the $y(0)$ values from the NLS approach. For the data in the earlier time periods with no convergence of Gauss-Newton algorithm we manually added the extrapolated values of $y(0)$. Then we calculated the $\ln_q$:

$$\ln_q(s) = \frac{\left(\frac{y(s)}{y(0)}\right)^{(1-q)} - 1}{1-q} \tag{3}$$

  for various values of q. Linear regression for points $[s, \ln_q(s)]$ that gives the highest R^2 provides the best estimate of q and the slope of the regression line provides the $(-1/\kappa)$ estimate.

  The q and κ estimates from this approach overemphasize the number of people of the largest cities and therefore give poorer fits than NLS approach. In order to improve the fit we used weighted linear regression using number of cities in each bin for the weights. The highest R^2 varied from 0.92 to 0.99.

Figure **??** shows the city size distributions resulting from our coding procedures with the fitted q-exponential curves. The x axis represents the binned size and the y axis the cumulative population from each bin size up to the largest size. The axes are logarithmic with the y-axis labelled in 10^3 units of city population. In the first two graphs, (a) and (b), the lines were fitted by the NLS approach. For the data where appropriate convergence of Gauss-Newton algorithm did not occur no curves are plotted. For the curves in the graphs (c) and (d) we used estimates q and κ that we obtained from weighted linear regression (WLR). We can see that WLR estimates give very good fit to the data. There are two distributions for the year 1950 due to two different datasets. The data "1950 a" are taken from the U.N. dataset for 1950-2005 and the data for "1950 b" from the Chandler dataset for -430-1950. Because the q value for the dataset of the year 361 deviated highly from all other estimates, and had by far the lowest R^2 of .92, we excluded the data for this period from further analysis.

The dotted black diagonal lines in graphs (a) and (b) represent the line $x = y$ that cannot be crossed by data points according to data binning. The dotted red lines were plotted in the way to connect data points with similar cumulative number of cities in each data period. Cumulative number of cities that was used for plotting the dotted red lines follow the multiples of two (hence $2, 4, 8 \ldots$). It is the city numbers that are usually plotted in size bins for estimating city size distributions, but we use these dotted lines to inspect for the possibility of irregular historical changes between successive periods. If the dotted city-number lines

[3] MLS procedures were also used. They showed....

[4] Might $q < 1$ mean that smaller cities are changing size faster than larger ones?

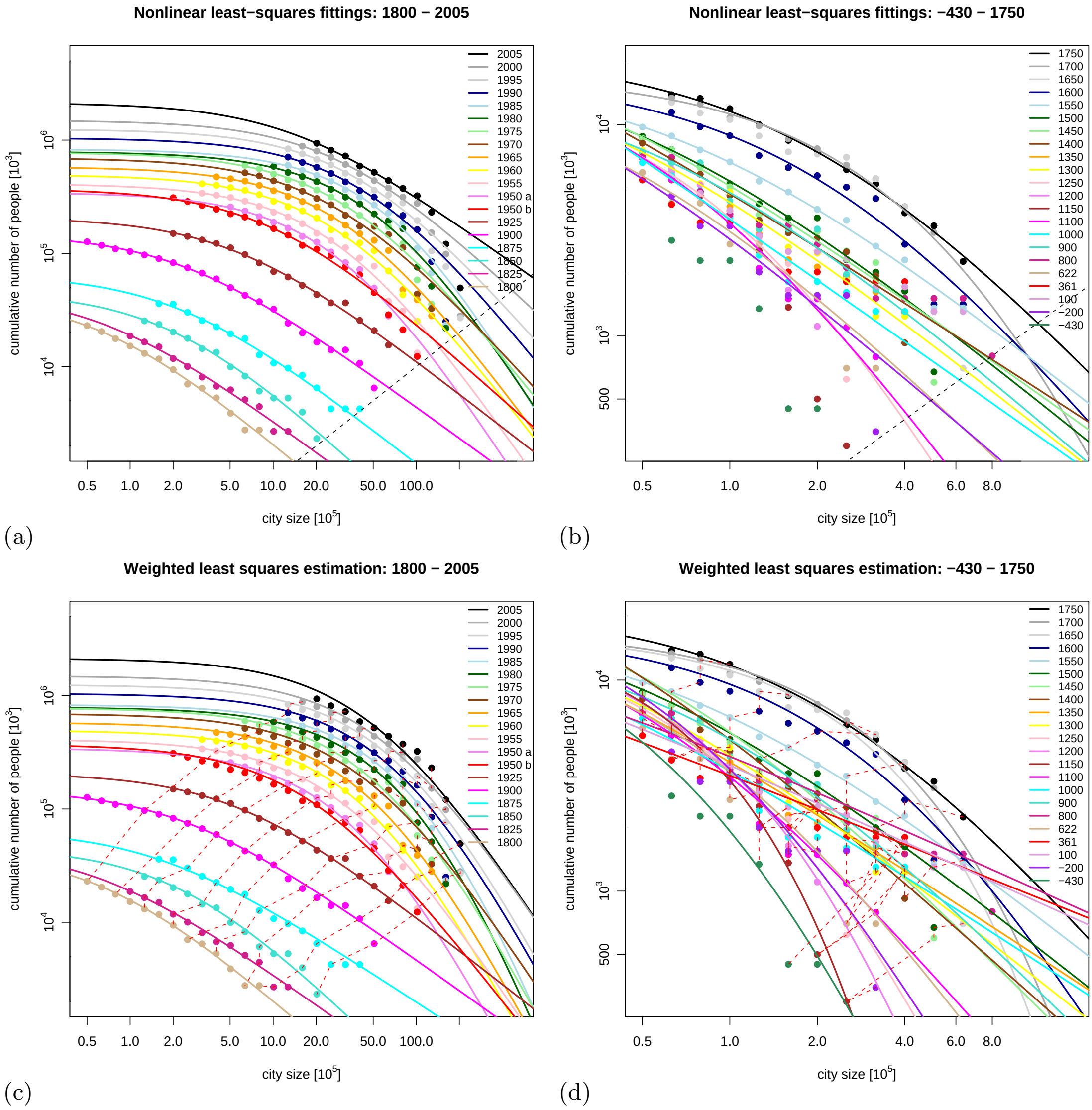

Figure 1: The points represent the data (different colors for different years) and the lines represent the fits. Note the log-log scale. The fittings were done using Gauss-Newton algorithm for NLS estimates (graphs (a) and (b)). For the data where the algorithm did not converge, the line is not plotted. The lines in graphs (c) and (d) are made using weighted linear regression estimates of q and κ. Data for the year 1950 are plotted twice (1950 a – from dataset 1950-2005 and 1950 b – from dataset -430-1950). The black dotted line represents the $x = y$ line, which is never exceeded by data points (due to the binnings). Red dotted lines on graphs (c) and (d) roughly connect same cumulative number of cities in data periods. It can be seen that data are noisier in the earlier time periods (the dotted lines cross).

were in parallel for bins of the same size from period to period, it would show that the q-exponential curves (with Zipfian or power-law as a special case) were associated with continuous growth-rates following some city size proportionality function. This is not the case in figure **??** (c) for such transitions as tho se for 1800-1825-1850-1900 during the industrial revolution, for example. Parallelism emerges between 1900 and 1960, but then is frequently disrupted after 1960 when populations limits are reached. Similarly, there are in figure **??** (d) city-number discontinuities such as those for the fifty-year transitions following incorporation of the new world into a global economy, severe new-world depopulation, and new bullion influx into Europe.

The downward tipping of the fat tails for the q-exponential curves of these distributions after 1950, as compared to flatter or mixed curves before 1950, do not appear from city-number comparisons to reflect differential winnowing of undersampled bins in the upper tails but might be due to a change in the distributions resulting from Chandler's estimates of sizes of urban agglomerations as opposed to the U.N. estimates of population within metropolitan limits is a way that tends to truncate the fat tails. Variations in q and other parameters are more consistent within the data series for the earlier and later periods, breaking at 1950, and as shown by a direct comparison of 1950 a (U.N.) and 1950 b (Chandler). The clustering of q-values after 1950, as seen in left part of figure **??** and variations in q before 1950, insofar as we rely on the city-number fluctuations for judgment, may well reflect historical fluctuations rather than random errors or measurement biases. Whether these historical variations are autocorrelated or indicative of slow fluctuations or sharper breaks requires testing.

4 Tests of Historical Periodicity

The 28 data points that could be estimated by NLS fitting give a runs test with p = .001 for the probability of no runs in values above and below the mean of q and an estimated 8 ± 2 such runs; we found 3 to 4 adjacent time periods for each run. A good example of how q varies by historical periods is the comparison of the period of 1800-1925, where $q \sim 1.84 \pm .09$ (s.d.), with that of 1950-1985, where $q \sim 1.55 \pm .09$ (s.d.) and significance p = .000001 for difference in means given the standard deviations in the two periods. The conclusion in this comparison is that the upper-bin urban hierarchy was more inegalitarian post-WW II than in the period of the industrial revolution. The trend line for q through all time periods is flat. The κ values, reflecting urban scale and an overall pattern of growth, show a temporal power-law trend toward a singularity in the current century. A runs test for the detrended values of κ shows 10 ± 3 runs and a probability p = .0003 of no runs. Moreover, q and κ are negatively correlated ($R^2 = .60$ between 900 and 1900), and the runs of high and low values show discrete breakpoints between historical periods. As shown in figure 2, the three runs of low q values indicate more inegalitarian periods, which correlate with κ values well above the κ trend line, indicating that these are periods of high urban population. Conversely, three periods of high q values and κ values below their trend line indicate more egalitarian periods with lower urban population.

Eurasia, from Europe to China, contributes over 95% of the large world cities for all time periods. Two historical correlates of q are shown for Europe in figure 3: approximate periods of land-route trade safety (relative absence of international wars and brigandage) and alternating economic periods dominated by financial capital operating through banks and corporate organization versus commercial capital operating largely through trading diaspora (see Arrighi 1994). Land route safety tends to lead to periods of financial capital accompanied by q inequality, while conflict interruptive of trade, raising the cost of land-route trading, leads to more decentralized or commercial forms of trading. An alternation of financial and commercial capital along with fluctuations in q does not hold for China, but has parallels in forms of empire, with integrative corporate extractive regimes in the medieval (Liao-Jin) and early modern (Qing) periods, alternating with the more dynamic but inward-turning Ming commercialism (1368-1644), in which maritime trade is completely abandoned, and finally, the disruptive period of the late Qing dynasty and aftermath of the Opium Wars. Communist China returns to a more corporate and centralized organizational form. The more general interpretation for Eurasia, then, is that the more eqalitarian q periods could be more conflictive or decentralized, alternating with expansive and centrally coordinated economies, whether through banks and corporations or through states and empires.

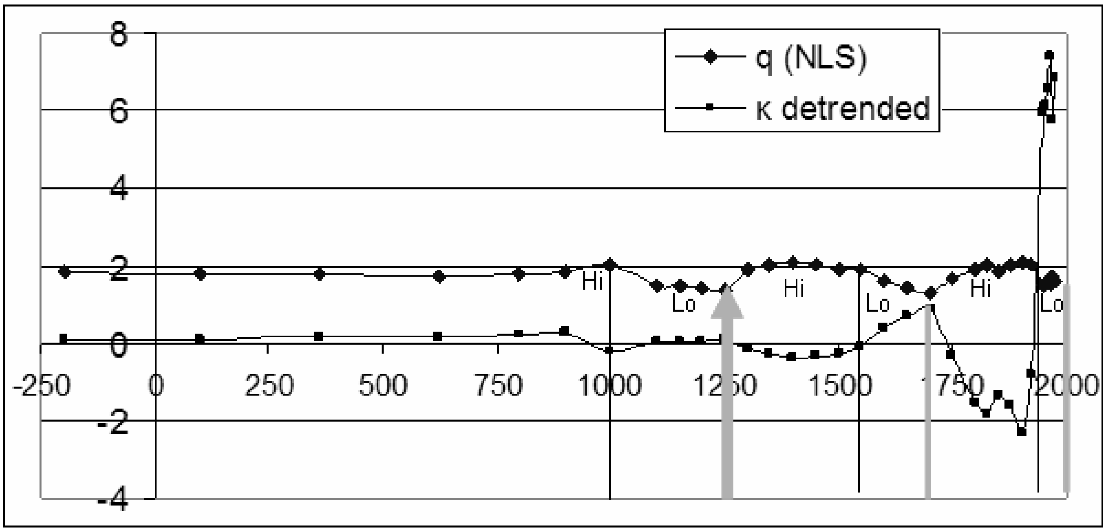

Figure 2: Correlation between q equality and detrended κ resulting in periodization (vertical lines).

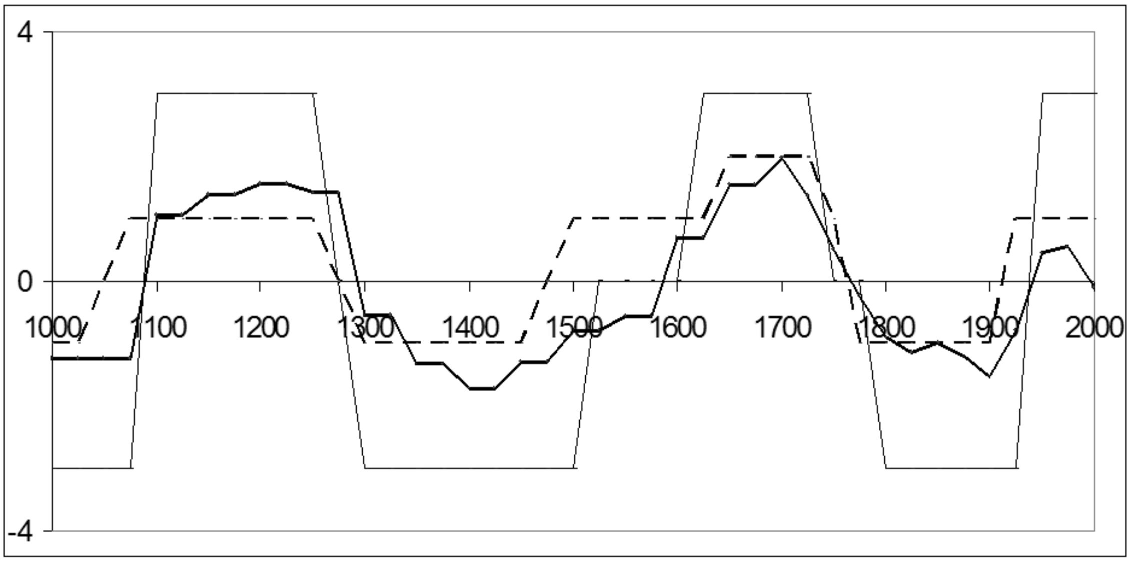

Figure 3: The heavy line indexes reversed variations in q (processes generating urban size hierarchy inequality) for world cities plotted by historical time along with two European historical correlates: dotted line shows approximate periods land trade route safety (relative absense of international wars) and light line shows alternating economic periods dominated by financial capital operating through banks and corporate organization (upper values) versus commercial capital operating largely through trading diaspora (lower values). Land route safety tends to lead changes in periods of financial capital accompanied by q inequality.

5 World Population and City Sizes

To complete our exploration of q-exponential models, we examined how the characteristic number of people (extrapolated value of $y(0)$) is correlated with estimated world population size. The latter data were obtained according to Kremer (up to 1990) and the U.S. Census Bureau (after 1990). The correlation between the two is seen in Fig 4. [Natasa: the 10^c coefficients are still incorrect on these graphs] On the left graph the world population (fitted by a line) and extrapolated $y(0)$ values (darker nodes) are plotted in time along with an inset that shows percentage world urbanization. City scale $y(0)$ spikes above world population in the classical period, again in the 13th century (medieval) renaissance, and then again with the shift in slope of percentage urbanization (ca. 1830) that accompanies the European industrial revolution. The right graph, with relative population measures on both axes of the plot, details the relationships between world population and $y(0)$ values for the three periods after 1900: κ lags up to 1945, then accelerates to 1975, lags to 2000, and leads again after 2000 (these figures finish at 2003). The $y(0)$ values up to 1980 (where $y(0) = 0.8$ million and world population equals 4.448 billion) follow a logistic curve as if urban population were leveling off (as did world population after 1962) but then turn sharply upward to the endpoint in 2003, indicating continued growth of urban population. The dotted line corresponds to $x = y$ as a reference line between the two population axes. The light solid curve represents $y(0)$ as a best fit exponential regression $\simeq 71315 \cdot 1.0005^w$ function of w, world population, ignoring their fluctuating relationship.

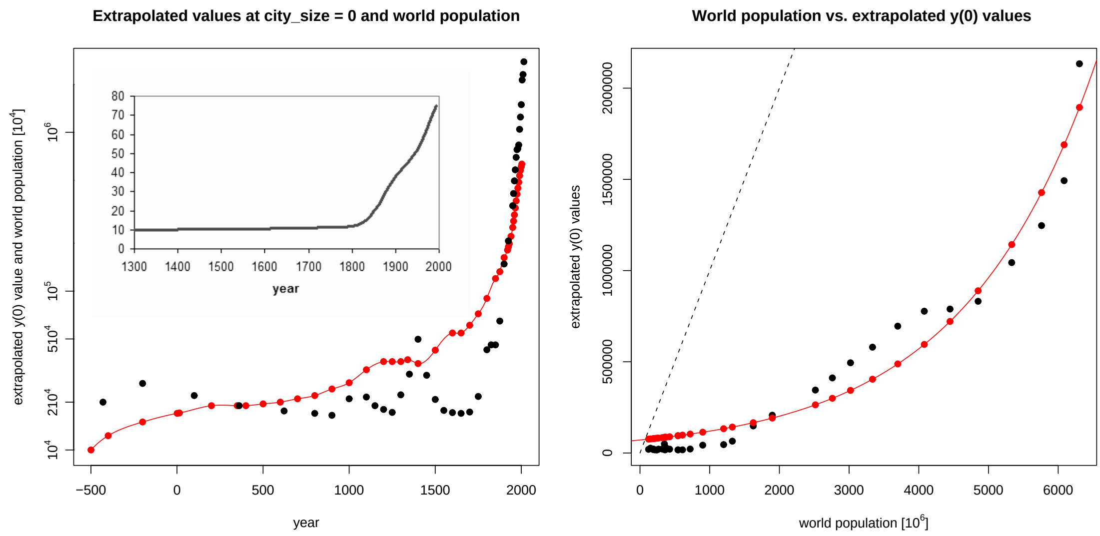

Figure 4: The estimates of $y(0)$ are plotted according to the NLS approach. Eight estimates (19%) had to be manually added due to non-convergence of estimation procedure. On the left graph black dots represent extrapolated $y(0)$ values and the curve with dots the data on world population size. The inset shows world urbanization percentage, 1300-2000. On the right graph black dots represent the $y(0)$ data and the curve is $y(0) \simeq 71315 \cdot 1.0005^w$ where w represents world population. The dotted black line in right graph is of the form $x = y$ and can be used as a reference line.

6 Conclusions

Fitted q-exponentials give a detailed description not of a Zipfian "law" of urban population but one highly relevant to urban and population processes. Contrary to the Zipfian model, our finding for the upper slopes of city size distributions is that they exhibit regular shifts in the structure of urban hierarchies, ones that are autocorrelated over characteristic time spans that change in discrete jumps preceded by and correlated with other historical processes. Consistent with prior experiments fitting q-exponentials to city size distributions (Malacarne, Mendes, and Lenzi 2002), our weighted linear regression results for q and κ show excellent fit, except the time period 361 AD, to the entire city size spectrum, unlike the Zipfian and Pareto models for city-size distributions. The R^2 for our fitted q-exponential models mostly varies from .956 to .998, and the R^2 fits after 1700 are no lower than .97.[5]

Our results lead us to conclude that Zipf's "law" must to be rejected as an invariant, and replaced by a more careful investigation of the historical processes that are associated with the observed structural changes in urban hierarchies. Because our analysis used comparable world inventories of largest cities in different historical periods, our findings are not affected by changes in the samples and boundary conditions for aggregating city data to test the characteristics of city size distributions. Because these inventories are dominated in every period by cities in Eurasia, the regularities in the historical shifts we observe must necessarily reflect some degree of synchronization that has operated at least in the last 2.5 thousand years within the city networks of Eurasia. Some of this synchronization could be endogenous, such as the effects of the interdependices of periods of rise and fall of interurban trade and conflict, or could be the result of exogenous shocks such as major continent- or world-wide climate changes or catastrophic depopulation due to disease. Such shocks tend to synchronize temporal parallels in processes dependant on demography.

Looked at historically, the distributions that did not converge to a q-exponential (and by definition, neither to an exponential nor a power-law distribution or its Zipfian counterpart) are also highly patterned. Rates of nonconvergence are 6 percent after 875 CE and 50 percent before ($p < .01$). Explanations for nonconvergence after 900 CE may be due to climate change or hemispheric disease pandemics: The three cases are nonconvergence in 1150 and 1200 following the medieval warming after 1100, nonconvergence in 1350 after the Black Death and population collapse in Eurasian cities. Nonconvergence in 1650 followed catastrophic urban depopulation after the spread of European diseases in the New World. These were hemispheric or world events that altered population structure, especially in cities. Before 875, in contrast, convergence of estimation for 200 BCE is associated with a stable period of relative world prosperity (although the fit is the second worst after 361), and 622 CE is a q-exponential convergence period accompanied by population growth after recovery from the collapse of Han Chinese and Roman Empires. The relationship of q and detrended κ, up to 1900, is one of inverse correlation, but with slopes rising at two discrete time periods by 4.5 (1500/1550) and 22 (1850/1875). After 1950, q tends to stabilize between 1.3 and 1.5, as reflected in figure **??**. This stabilization is halfway between an exponential and a power-law distribution, and another reason to prefer the q-exponential over the Zipfian or power-law distribution to describe city size heirarchies.

We have suggested that the possible organizational variables implicated in synchronies between Chinese and European alternations between more hierarchical city distributions and more egalitarian heterrarchies lack similarity in detail but have some generic similarities in terms of centralization and decentralization, or between alternatives such as openness of easy trade accessability through land routes versus closure through conflict. We have not argued for an endogenous dynamic of major variables that affect city networks and attribute oscillations, but the possibility of identifying such dynamics is not remote. Not only are variations in q clearly related to historical fluctuations in external variables related to trade and conflict or degree of centralization of urban and economic functions but q covaries with fluctuations κ, the scale coefficient, in discrete temporal intervals.

Because q fits an entire city-size distribution, not just the larger cities, we consider q-exponentials a far more useful measurement step towards describing and explaining urban hierarchies than the Zipfian, or binning numbers of cities by size and fitting power-law coefficients. κ also varies relative to significant historical changes, notably leading lagged changes in population growth rates. We consider it an advantage to have historically sensitive measures of differential city size distributions.

The q-exponential parameters appear to reflect historical contingency rather than a scale-free distribu-

[5] Because the heavy tails of these distributions for larger cities show noise typical in the tails of binned distributions, and some lower-size bins were incompletely sampled, we removed values for lower and upper bins for incomplete bins.

tional law independent of historical processes. q-exponential methods appear to be ideally suited to studying both regularity and change in nonindependent processes. These results support the conclusions that variations in the scaling coefficients of urban size hierarchies ought to be examined more closely in relation to changing historical configurations, and in terms of possible dynamic interactions among the variables we have examined as well as with others. As such, changes in q parameters may reflect a variety of processes, including the various processes of human decision-making, network interaction and the fractal properties of urban hierarchies and geometries discussed by earlier authors as reviewed in our introduction. Our results are consistent with those earlier studies.

References

[1] Arrighi, G. 1994. *The Long Twentieth Century*. London: Verso.

[2] Batty, M., and Longley, P. 1994. *Fractal Cities*. London, Academic Press.

[3] Boon, J. P., and C. Tsallis, eds. 2005. Special issue on nonextensive statistical mechanics. *Europhysicsnews* 36(6). `http://http://www.europhysicsnews.com`

[4] Butler, R. 2005. World's Largest Cities [Ranked by City Population]. `http://www.mongabay.com/cities_pop_01.htm`

[5] Carvalho, R., S. Iida and A. Penn. 2004. Scaling and universality in the micro-structure of urban space. *Physica A*, 32: 539-547.

[6] Chandler T., and Fox G. 1974. 3000 years of urban growth. New York, London, Adademic Press, 431 p.

[7] Gabaix, X. 1999. Zipf's law for cities: an explanation. *The Quarterly Journal of Economics* 114(3): 739-767.

[8] M. Gell-Mann and C. Tsallis, eds., 2004. *Nonextensive Entropy – Interdisciplinary Applications*, Oxford University Press, New York.

[9] Lobos, J. 2005. Scaling Relationship between Metropolitan Sectoral Employment and Metropolitan Population in the U.S. ISCOM working paper, Arizona State University.

[10] Malacarne, L. C., R. S. Mendes E. K. Lenzi. 2002. q-exponential distribution in urban agglomeration. *Physical Review E* 65 (017106):1-3.

[11] Makse, H. A., S. Havlin, and H. E. Stanley, 1995. Modelling Urban Growth Patterns, *Nature* (London): 377: 608-612. `http://xxx.lanl.gov/abs/cond-mat/9809431`

[12] Malescio, G., N. V. Dokholyan, S. V. Buldyrev, and H. E. Stanley, 2004. Hierarchical organization of cities and nations. *Physica A* 332: 539-547. `http://arxiv.org/abs/cond-mat/0005178`

[13] Mandelbrot, B. B. 1977. *The Fractal Geometry of Nature*. New York: Freeman.

[14] Marsili, M., and Y.-C. Zhang, 2004. Interacting Individuals Leading to Zipf's Law. *Phys. Rev. Lett.* 80, 2741-2744. oai:arXiv.org:cond-mat/9801289 (1998).

[15] Rapoport, A. 1978. Rank Size Relations. *International Encyclopedia of Statistics* v. 2, p. 847, eds. W. H. Kruskal and J. M. Tanur.

[16] Tsallis, C. 1988. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* 52: 479-478.

[17] White, D. R., N. Kejžar, C. Tsallis, D. Farmer, and S. D. White. 2006. Generative model for feedback networks. *Physical Review E.*

[18] Zanette, D. H., and S. C. Manrubia, 1997. Role of Intermittency in Urban Development: A Model of Large-Scale City Formation. *Phys. Rev. Lett.* 79, 523-526. http://prola.aps.org/abstract/PRL/v79/i3/p523_1

[Tsallis: Re bottom of p. 1 and top of 2. I find it counterintuitive that as q goes to 1 more independent processes cancelling each other out that the city hierarchy should be more hierarchical. Could you check through the logic and the scaling in the paper?