Can There Be a Science to Cities?
Workshop 12th-14th July 2012

The Growth, Scale and Size of Cities:

What Do We Know after
One Hundred Years of Effort?

Michael Batty
University College London
m.batty@ucl.ac.uk
@jmichaelbatty

http://www.complexcity.info/
http://www.casa.ucl.ac.uk/

Centre for Advanced Spatial Analysis UCL

My Key Topics

- What Is Scaling? What Is Growth?
- Space – Distance, Size – Frequency
- Seven Laws of Urban Scaling
- Three Exemplars
  - a) Allometry and Agglomeration
  - b) Size Distributions
  - c) Gravitational Interactions
- A First Attempt at Integrating Size, Scale & Interaction
- Open Questions: Defining Size, Choosing Scale, Measuring Time

Centre for Advanced Spatial Analysis UCL

What Is Scaling? What Is Growth?

I am going to assume that you all know what scaling is but I will nevertheless introduce my definitions of these ideas as I think there can be differences.

Scale is central to the way we organise our knowledge about cities – as we define them spatially and hierarchically from the region, even the nation state, to the neighbourhood and even below to clusters around streets.

The range of scales is thus bounded – and this means we must be wary of models that presume invariance to change over all scales. I tend to see scale as space related but I realise this is only one conception

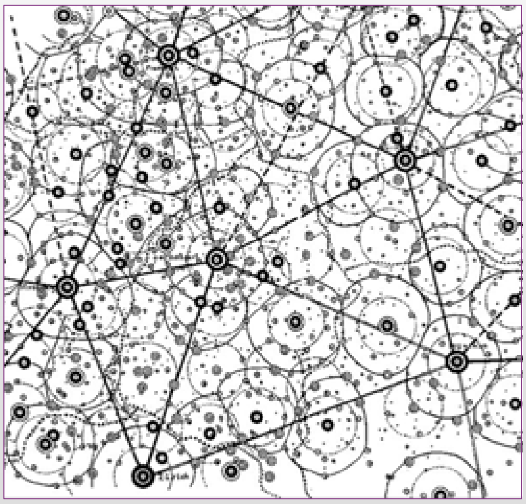
The biggest cities for example are in the order of millions – 25m Tokyo, Mexico City – and the smallest probably are in the order of hundreds, possibly less so the range is 5-6 orders of magnitude.

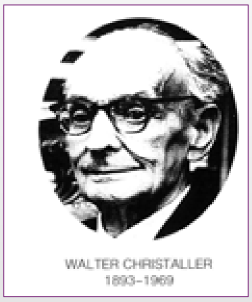
If we look at physical form in cities, this scaling is reflected in fractal structures, statistically self-similar forms that repeat themselves across these scales with again no more than 6 orders of magnitude.

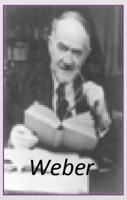
The same is more or less true with respect to the hierarchy of road systems, then.

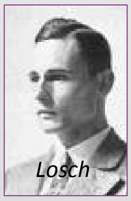
Long before ideas about scaling and fractals surfaced, urbanists recognised this scaling in structure.

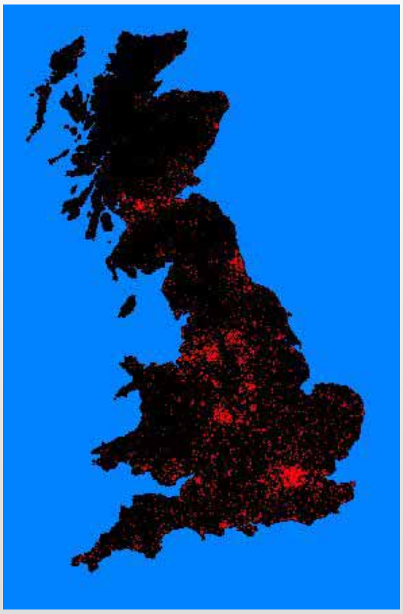
Let me break the ice and show some pictures of this before I become a little more abstract

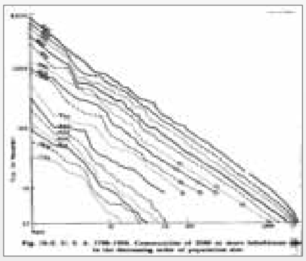




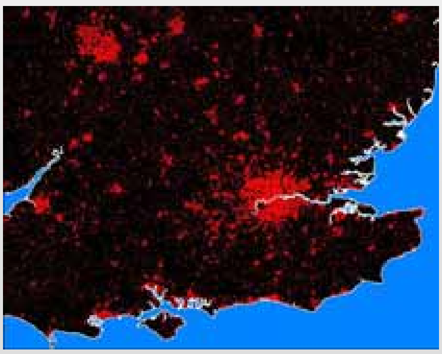






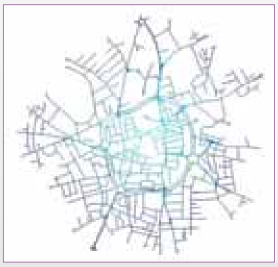


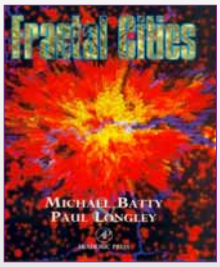
ZIPF

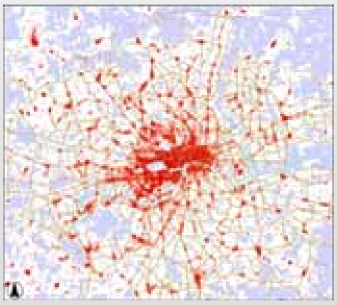


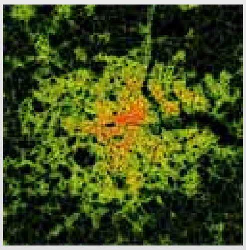
Centre for Advanced Spatial Analysis UCL

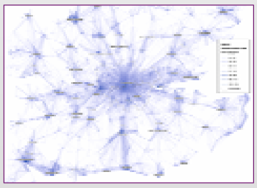
A lot of these patterns are illustrated in our 1994 book which can be downloaded at www.fractalcities.org











Let me say something briefly about growth. Our bias in thinking of cities is to growth

A big city has to be a small city first and there is thus a natural asymmetry

Scales tend to increase as cities grow. This means that new features at larger scales emerge as cities grow

So like all systems that grow from the bottom up, then scale changes as they grow. This has an immediate consequence for their shape and for the way we function within cities of different sizes.

This is central to the notion of changing shape with size – allometry which is central to this workshop

Space – Distance, Size – Frequency

Let me summarise what we know as the scaling properties of cities are reflected in morphology

a) Cities change shape as they change in size – this is allometry, and it means that we tend to move differently in cities of different sizes

b) There are many more small cities than big cities, and this scaling reflects competition for resources: to be a big city you must be a little city first

c) Cities are distributed by size in such a way that little cities are nested in the hinterlands of bigger cities. Big cities are spaced more widely than little cities.

d) People interact with each other more intensely in bigger than smaller cities. This is due to the fact that the no. of potential interactions in a population P is P^2. In fact Dunbar's number suggests than the number of potential interactions has an upper bound of about 250 but the pressure to interact is greater in bigger cities.

e) People interact with one another less with increasing distance between them: this is the gravitational law.

f) Other kinds of interaction that diffuse over space, fall off with distance from their source. This tends to reduce the potential interaction effects of bigger cities.

I could go on informally with this list of properties and examine more detailed activities which comprise cities but the point I am making is that many things scale with the geometry of cities.

And there are many aspects to this geometry. So it follows there are many aspects to this scaling.

As far as I am aware, there is no good discussion yet of these scaling relations and the way they interact.

It would be nice to think that someone might produce a decent synthesis of these ideas but currently they are entangled with one another in ways that are hard to unravel. Let me summarise possible laws.

The Laws of Urban Scaling

Let me try and formalise a little more how these scaling laws might be and have been developed. A word of warning. They may not be laws in the accepted sense of the term in the physical sciences but they are regularities that seem to persist in time and space.

All others things being equal, ceteris paribus…….we can state the following about cities

- As they grow, the number of 'potential connections' increases as the square of the population (*Metcalfe's Law*, the network equivalent of *Moore's Law*)
- As they grow, the average time to travel increases

- As they grow, the 'density' in their central cores tends to increase and in their peripheries to fall
- As they grow, more people travel by public transport
- As they get bigger, their average real income (and wealth?) increases (the *Bettencourt-West Law*) – this is allometry. It might also be called *Marshall's Law*
- As they get bigger, they get 'greener' (*Brand's Law*)
- As they get bigger, there are less of them (*Zipf's Law*) – this is city size – rank size

Let me look briefly at the third of these observations: that is, as cities grow, the density in their central cores tends to increase and in their peripheries to fall

In fact in urban economics, there is a long tradition of generating monocentric city models where rents scale inversely with distance (or travel cost) from the core. This can probably be called *von Thunen's Law* after the German Count who first observed this on his estate in Saxony in 1826.

As a power law, this is central to spatial interaction, so we really need a law of scaling that says that densities and rents decline as an inverse power or exponential function of distance from their cores. In fact I am going to call this *Alonso's Law* after his work in the early 1960s in resurrecting von Thunen and applying this theory to cities.

Three Exemplars

a) Allometry and Agglomeration

As cities get larger, as they grow, they change in shape. Strictly speaking, no one has quite measured such qualitative change in terms of morphology as yet but a proxy for this is in terms of how their 'attributes' change with respect to population size.

An example relates income or wages or some measure of wealth Y to population P as

$$Y = KP^{\alpha}$$

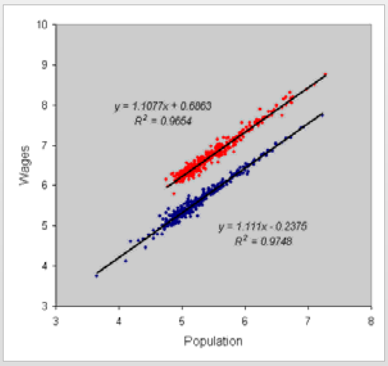
where K is a constant of proportionality and α is the scaling parameter

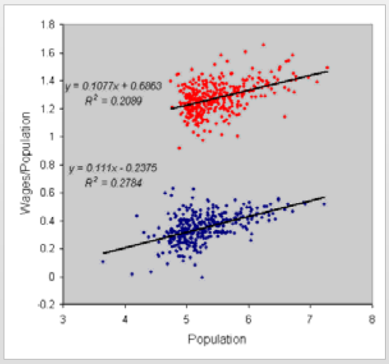
The scaling parameter can be greater than 1 $\alpha > 1$ which is positive allometry, less than 1 $\alpha < 1$, negative allometry, or equal to 1 $\alpha = 1$ which is isometry.

Greater than 1 is referred to as superlinearity and less than 1 sublinearity.

Geoff, Luis, Jose et al. from Santa Fe have done most in this in recent years, particularly following Geoff's work in biology on scaling and allometry. Much of this discussion is now about how big cities might be more wealthy, greener, more efficient and more divided can be predicated in these terms. But let me illustrate with an example from the US data.

I, following the Santa Fe group, used the US Bureau of Economic Analysis on SMSAs from which I simply took their 366 regions for which population and income/wages data available from 1969 to 2008. For the first and last years in this data set





Centre for Advanced Spatial Analysis UCL

This shows distinct superlinearity, distinct economies of scale, notwithstanding a debate which is beginning as to how strong these relationships are.

This work produces extremely plausible evidence that the things that scale sublinearly in cities tend to be physical objects such as infrastructures, while things that scale superlinearly are attributes of populations that are highly specialist.

As we will see a little later, there are considerable problems in wrestling with the data for these kinds of problems and probably the way the national space economy has developed is significant in this. I will come back to this at the end

Centre for Advanced Spatial Analysis UCL

b) Size Distributions

Our next relation involves the frequency $f(P)$ of different city sizes P and this of course is Zipf's Law which we state as

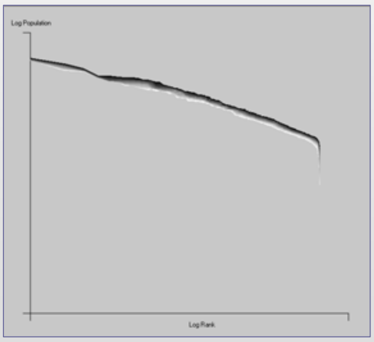
$$f(P) = KP^{-\beta}$$

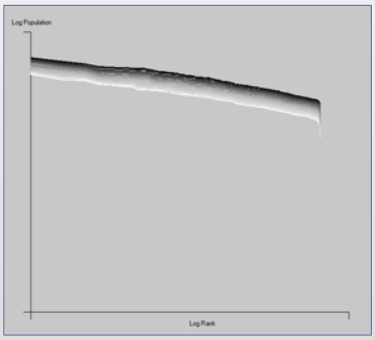
K is still a constant of proportionality and β is now the scaling parameter. Zipf's Law is usually presented in its counter cumulative form as the rank size rule and this can be stated from

$$F(P) = KP^{-\beta+1} = KP^{-\lambda}$$

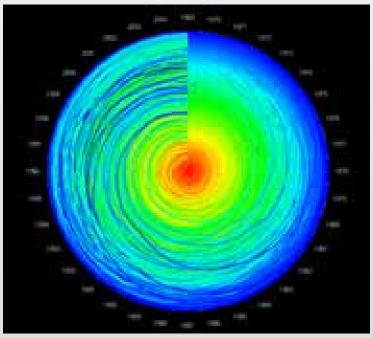
We get the strict form of Zipf's Law when $\beta = 2$, hence $\lambda = 1$

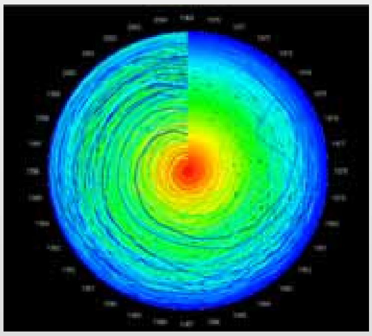
There are many illustrations of rank size that I might give but keeping to the SMSA data set used for exploring the allometry previously, then the ordering of population size and size of wages by rank is as follows





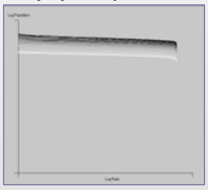
These are extremely regular and show little qualitative change over time. The volatility of the cities composing these relations is also quite muted compared to longer times, that is the cities do not change their ranks that much; they do a bit though

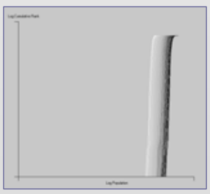


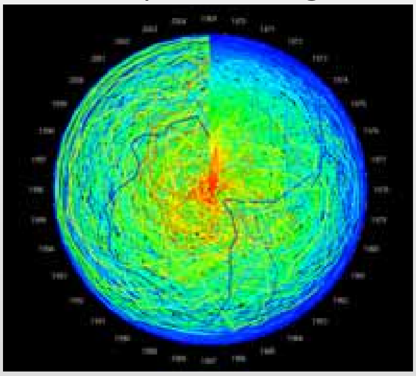


Centre for Advanced Spatial Analysis UCL

However when we look at their ratio, that is wages per capita, the volatility of this change is considerable. And represents a further puzzle. This however is a digression that I don't want to take any further here for my purpose is not to explore changes in rank







Centre for Advanced Spatial Analysis UCL

c) Gravitational Interactions

The last law that has been widely applied involves the analogy with gravitation that pertains to the interaction $T(P_i, P_j)$ between two populations which we can state as

$$T(P_i, P_j) = K \frac{P_i P_j}{d_{ij}^{\phi}}$$

where d_{ij} is some deterrence, intervening opportunity or often distance and ϕ is the scaling parameter

Sometimes $\phi = 2$, the inverse square law, but often as in Zipf's Law, it is different from its theoretical equivalent value.

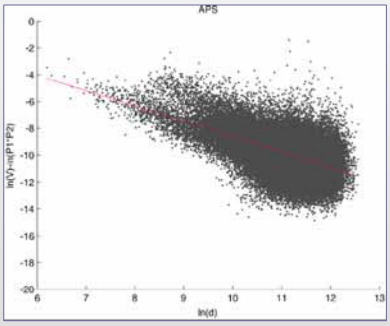
I should also say that our gravity model can be generalised to a population density model is we consider only one origin – such as the centre of the city and many destinations. Then we get the following sort of model which is very widely used and whose initial statement was by Colin Clark in 1951.

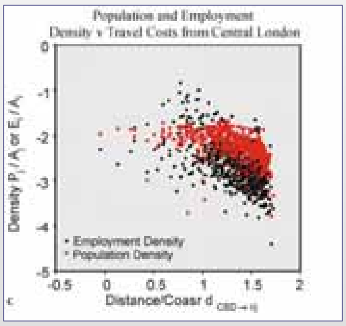
$$\rho_j = \frac{T(P_0, P_j)}{P_0 P_j} = K d_{ij}^{-\phi}$$

And this can also be generalised to a kind of rank-size where rank is unit distance from the CBD

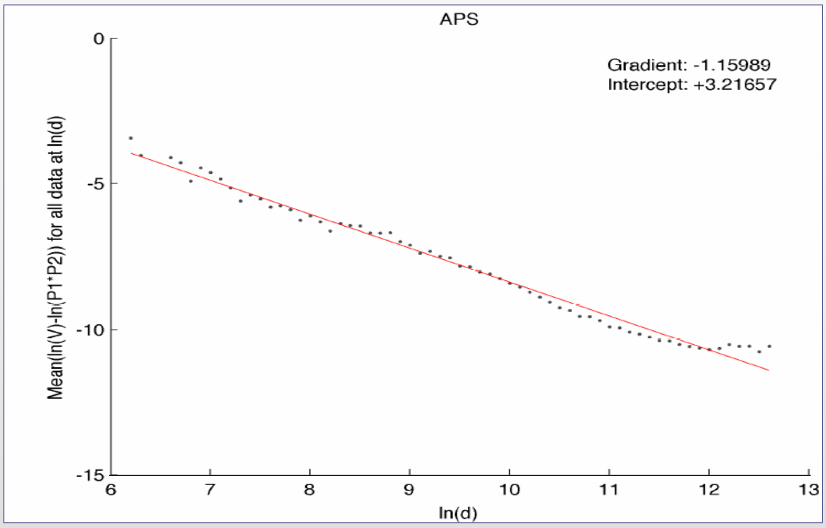
There are many many illustrations of these kinds of distance decay relations.

Jon Reades has some great distance decay curves for London for telephone calls and here is one – I also show some census data by the side of this taken from the 2001 Population Census.

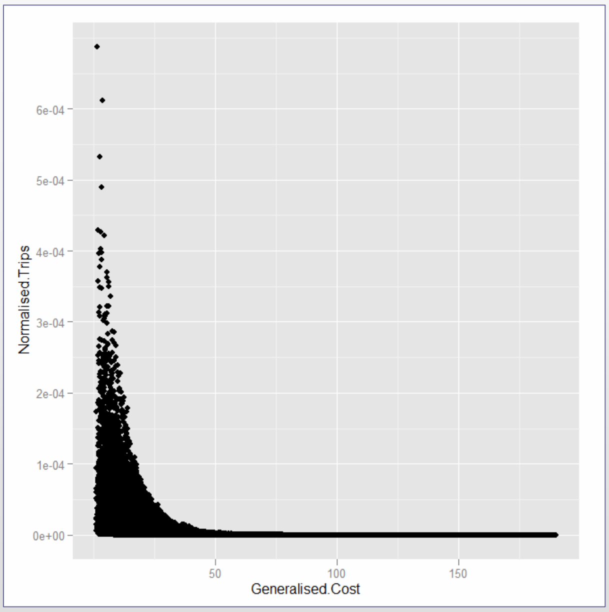




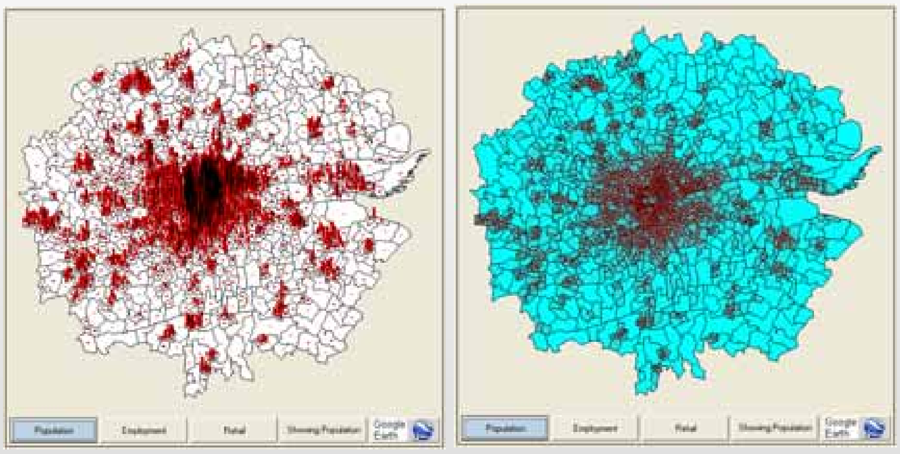
And for then for wider region, we can see this as



And for then for wider region, we can see this as



Plotting all normalised trips with respect to generalised cost of travel for the *LONDON* region



Centre for Advanced Spatial Analysis

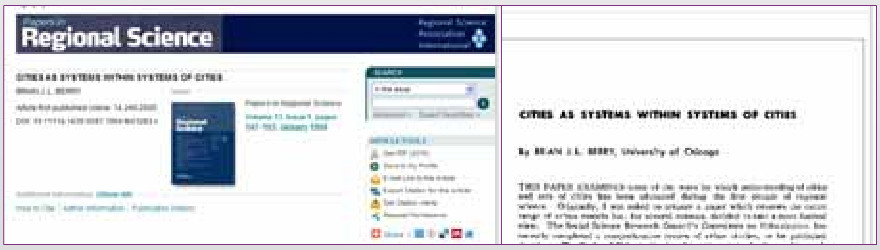
There are some trenchant debates surrounding these three types of relationship

Whether they are negative exponential, stretched exponential, or inverse power has preoccupied us a lot and substantial effort has put into ways in which power laws can be generated using simple models –

The heritage in this area is long and distinguished from Pareto, Yule, Lotka, Simon to Gabaix and Sornette et al. fusing urban growth theory with random stochastic models in the Gibrat tradition,

And in terms of allometry from Huxley, Haldane and so on through to the Santa Fe group.

In spatial interaction from von Thunen to Alonso to Wilson etc. Last but not least, much of this work on scaling and self-similarity came out of the quantitative revolution in geography from the mid 1950s onwards from Garrison and Berry to Tobler, Getis, Nysteun, to Woldenberg and many others. I offer a glimpse of this world in Berry's 1964 paper



A First Attempt at Integrating Size, Scale and Interaction

We can define locations that relate to one another in terms of how populations relate. Locations intensify as people demand to be together to exchange in markets and it is usual for there to be a limited number of points where this takes place.

The density around these points is highest and the population then distributes itself around such points usually following some sort of inverse distance law as implied by urban density scaling.

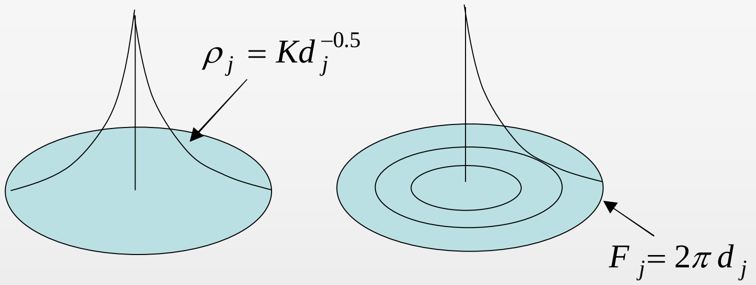
Centre for Advanced Spatial Analysis UCL

Assume that everyone interacts with a market C. Then the distance from a point j to the market is d_j and we assume the density ρ_j follows an inverse square law– a power law (often a negative exponential) – $\rho_j = Kd_j^{-2}$

Now we can plot a density cone in familiar form around the market centre C and note also that the number of points where people live around C varies according to the circumference of the circle at distance d from the centre, i.e. the no of locations is $F_j = 2\pi d_j$

The size of each point is the density $\rho_j = Kd_j^{-2} = P_j$

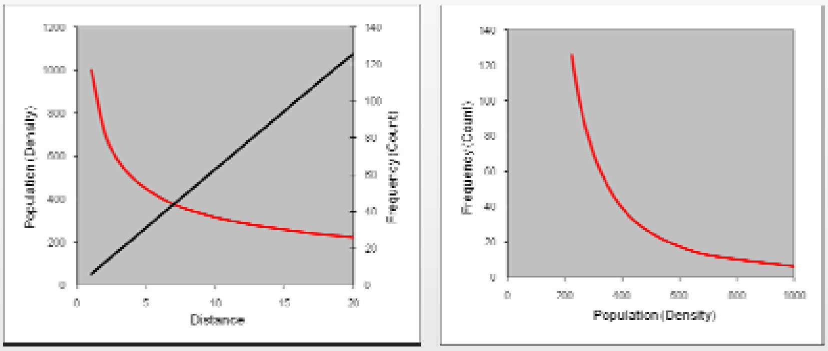
Centre for Advanced Spatial Analysis UCL



We have changed the power of distance to 0.5 because this gives us a better result and is arguably relevant because we are really doing this on a line not on a space. Now let us see if this satisfies our basic scaling relations – let us count the frequency of different locations and compare these against different sizes.

Centre for Advanced Spatial Analysis

UCL

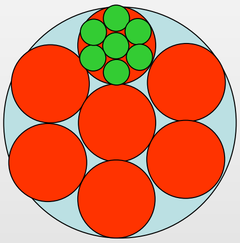


We thus compare F_j against ρ_j or P_j for a simple numerical example later. It is easy to show that the relation is dead simple and is (by construction) a power law, that is $F_j = GP_j^{-2}$ which leads directly to the rank size rule. Note the fact that distance could be unit rank –

Centre for Advanced Spatial Analysis

UCL

– albeit using a power of distance very different from the inverse square law. So we have established that this functional form leads to scaling.



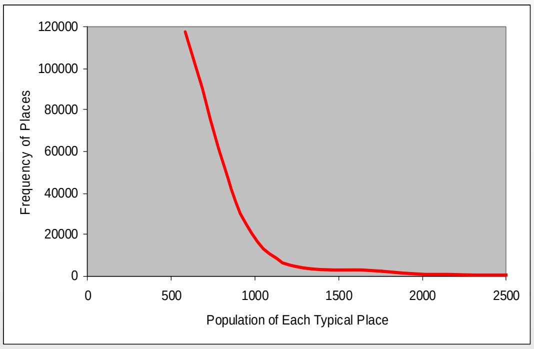
Now we can do exactly the same kind of exercise for a large space divided into a hierarchy of central places – we can assume a radius around the largest centre and calculate the total population, and then for successively smaller centres with smaller hinter-lands, we produce populations and then compare these against areas

which are frequencies and which generate the same kind of rule. Our lattice is as shown above, and we can forget the spaces in between – Applying the same logic as for each circular town at each level and computing total populations in the hierarchy, we derive the same sort of scaling as follows.

First we assume a maximum radius d=1000 for the biggest all embracing central place – the blue circle and this gives the following total population as the integral of the density up to d=1000; the population is approximately 15811

Then at the next level down we divide the area of the largest circle into say, 7 red sub-circles each with radius 1000/3 and each of gives a population of 9128. We then get 49 areas at the next level down – the green circles each with a population of 5270 and so on, down to where we fix the lowest level at 40,353,607 circular areas, each with a population of 113.

If we then graph the frequency of this hierarchy against typical population size and plot the following graph which is clearly scaling.



This relation is linear on a log-log scale with the power of the size around 0.28, dependent of course on the assumptions we have made about how many levels of hierarchy there are and how each successive size is subdivided.

Allometry of course relates to how these population scale with other phenomena and almost trivially we have defined them to scale with distance from C: and more roads are needed the smaller the populations are in this sort of monocentric city which is the sort of sublinearity we see in real cities. The equation is something like this

$$D^{0.5} = \int d(x)^{-0.5} dx = Z \int P(x) dx = P$$

In terms of income, we have not extended the model to income so we can say nothing about this.

This is the gist of an argument that might relate these various scaling laws.

Open Questions: Defining Size, Choosing Scale, Measuring Time

I will finish this rather general sweep through scaling, growth and cities with a three open questions which are absolutely crucial to empirical work in this area.

To deciding what kinds of relationships we have and even to deciding whether there is the kind of regularity that we suggest for city systems

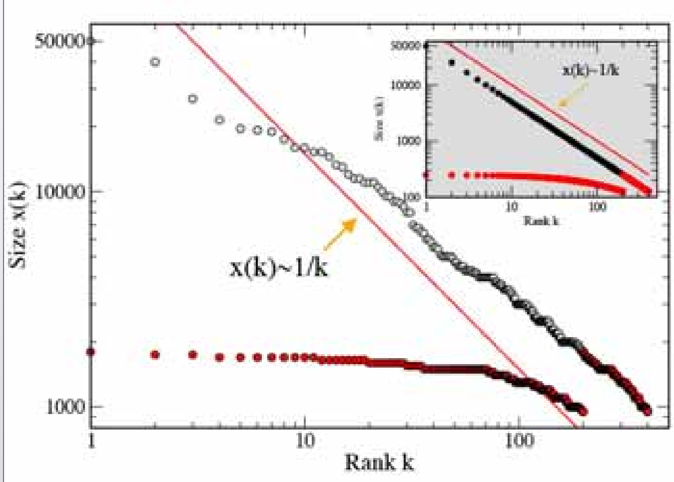
These are both to do with how we define our systems of interests first in terms of whether or not we have a 'complete' set of such objects, and second, whether or not we have the 'right size' of object.

First then whether or not we have a complete set of objects and I will illustrate this for the rank size rule.

If we have a pure rank size distribution and we then omit the first half of the set of ranks – let us say 1 to n/2 in a set of n objects, and we then reorder that ranks of the set n/2 to n as a new set 1 to n/2, then this second set no longer follows a rank size distribution.

The distribution no longer has the coherence of the initial set.

Here we can show this for the following data which is the 390 US billionaires from the Forbes List in 2010.



There is a very simple message here. If we miss out any objects in our set where we know or assume scaling, then we will never be able to demonstrate scaling.

It is particularly crucial for city size distributions (and firm size too) because we often do not have decent control over how we pick our cities.

But more to the point, we often have to use cities (or firms) that pertain to national boundaries and we may want to examine size distributions that cross boundaries. This is a veritable minefield of problems for if we go the other way and merge two sets which are Zipfian, we do not get Zipf's Law

My second example pertains to how we define cities – in terms of their extent. This has been a major problem from time immemorial in that

a) the concept of a city has changed through time

b) Merging of cities into one another complicates the picture – Geddes' the father of town planning in Britain at least defined the term conurbation for this kind of polycentric structure

c) But in an urbanised world, where do towns begin and end

d) And last but not least, in a global world, cities merge into one another virtually or rather parts of cities do

Centre for Advanced Spatial Analysis UCL

We are engaged in testing the Bettencourt-West ideas on the UK – attempting to look at the question of proportionality and scaling in terms of physical and socio-economic attributes of towns of different size and our results are confusing.

I am going to leave this for Else Arcaute to explain to you in the panel discussion I think

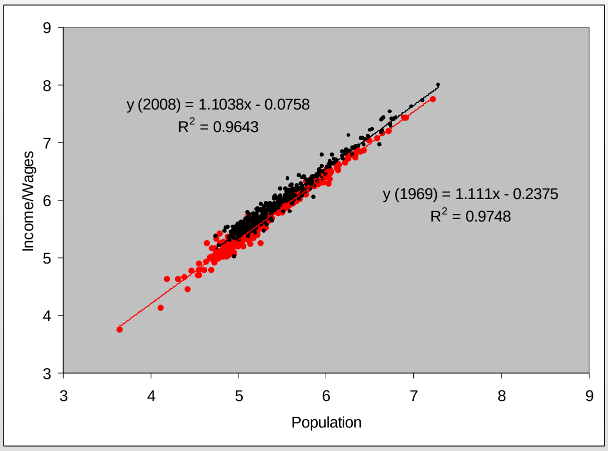
In essence we have some rather different results for cities in the UK that do not seem to demonstrate superlinear scaling for incomes, but more from Else later in the meeting.

My third and last example relates to how we test for allometry in systems of cities. Do we test across the range of cities at any one time, or do we test for an individual city at different times.

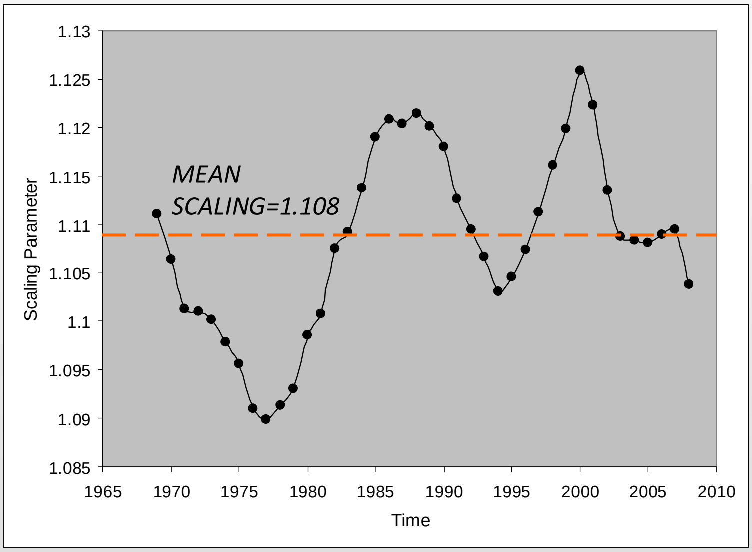
As different cities get bigger we should expect them to get more than proportionately richer; but should we expect a particular city to get richer as it gets bigger, that is more than proportionately richer?

Probably the answer is yes. From the SMSA data set, we have in fact now normalised income by cost of living (inflation index), and our new data is P_{it} and Y_{it} where *i* is a city and *t* is time

What we have been doing is looking at P_{it} and Y_{it} over each city i for different time intervals; we find the same scaling at 1969 and 2008 and also in between



Centre for Advanced Spatial Analysis UCL



Centre for Advanced Spatial Analysis UCL

If we now look at the top five and lowest five cities by population change – i.e. the greatest population change in absolute terms from 1969 to 2008 for the top and bottom 5 cities, these are

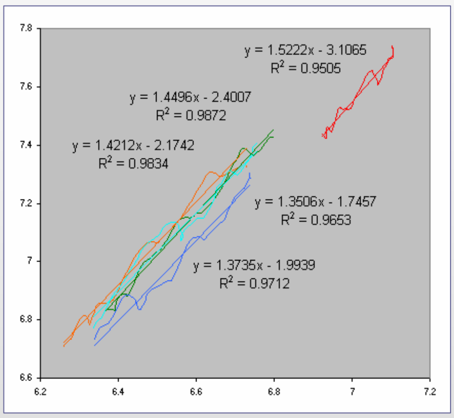
Top five + pop

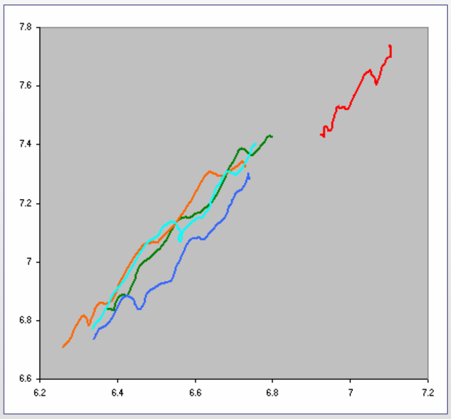
- Los Angeles-Long Beach-Santa Ana, CA (MSA)
- Dallas-Fort Worth-Arlington, TX (MSA)
- Atlanta-Sandy Springs-Marietta, GA (MSA)
- Houston-Sugar Land-Baytown, TX (MSA)
- Miami-Fort Lauderdale-Pompano Beach, FL (MSA)

Bottom five - pop

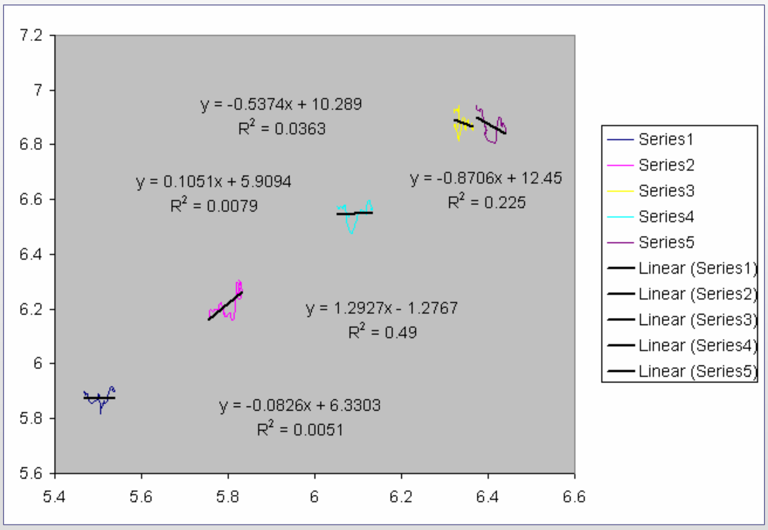
- Utica-Rome, NY (MSA)
- Youngstown-Warren-Boardman, OH-PA (MSA)
- Cleveland-Elyria-Mentor, OH (MSA)
- Buffalo-Niagara Falls, NY (MSA)
- Pittsburgh, PA (MSA)

This is what we get Top five – positive pop





This is what we get Bottom five – negative pop



I am not going to conclude in any more depth than this as I am well over time but I hope some of these ideas will resonate with the talks during the rest of the meeting and generate some good discussion.

I will post a version of the pdf of this powerpoint on the Presentation Pages of my web site. This is based on a version in Oxford last year but this variant today is posted on my odds and ends blog site under the post called Santa Fe, Complexity, Cities

http://www.spatialcomplexity.info/

Thanks