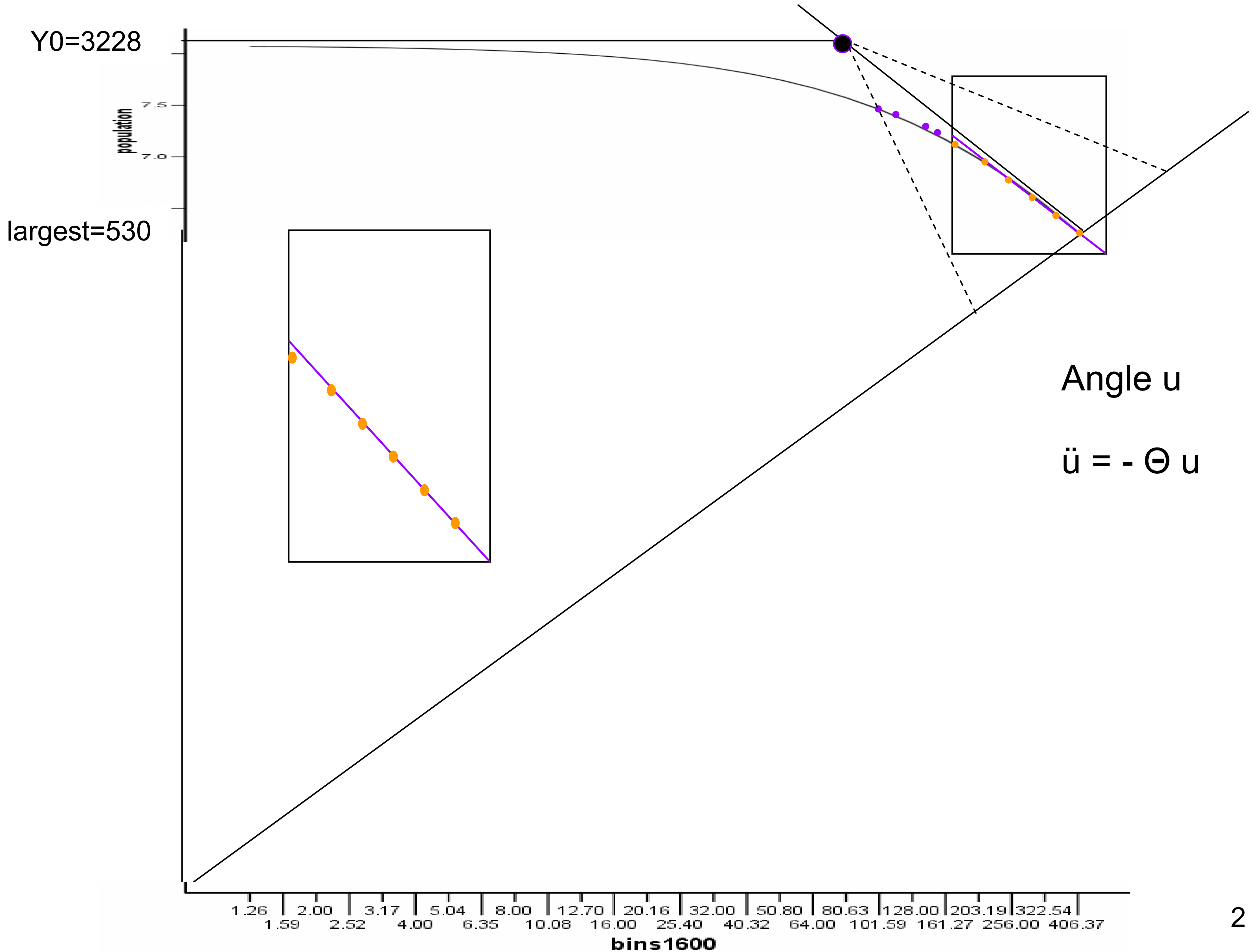
Modeling City Size Data with a Double-Asymptotic Model (Tsallis *q*-entropy)

Deriving the two Asymptotic Coefficients (*q*,Y0)
and the crossover parameter (kappa: ĸ)
for 24 historical periods, 900-1970
from Chandler's data in the largest world cities in each
checking that variations in the parameters
for adjacent periods entail real urban system variation
and that these variations characterize historical periods
then testing hypotheses about how these variations
tie in to what is known about
World system interaction dynamics

good lord, man, why would you want to do all this?

That will be the story

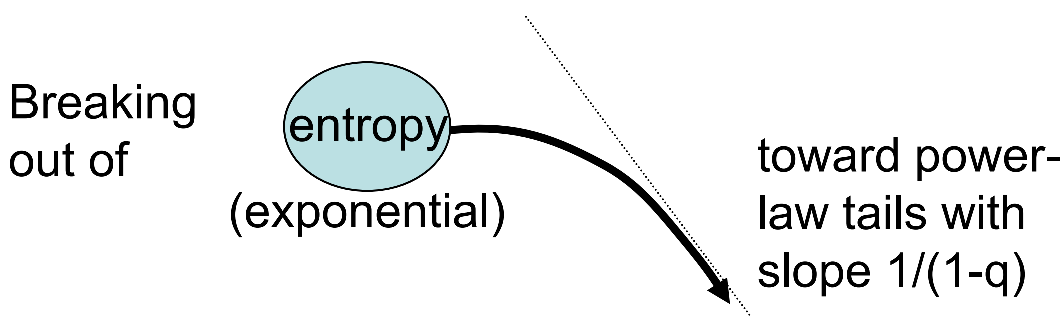

Y0=3228
largest=530
population
7.5
7.0
1.26
1.59
2.00
2.52
3.17
4.00
5.04
6.35
8.00
10.08
12.70
16.00
20.16
25.40
32.00
40.32
50.80
64.00
80.63
101.59
128.00
161.27
203.19
256.00
322.54
406.37
bins1600
Angle u
ü = - Θ u

Why Tsallis *q*-entropy?
That part of the story comes out of network analysis
there is a new kid on the block
beside scale-free and small-world models of networks
which are not very realistic
Tsallis *q*-entropy is realistic (more later)
but does it apply to social phenomena
as a general probabilistic model?
The bet was, with Tsallis,
that a generalized social circles network model would
not only fit but help to explain *q*-entropy
in terms of multiplicative effects
that occur in networks
when you have feedback

That's the history
of the paper in Physical Review E by DW, CTsallis, NKejzar, et al.
and we won the bet

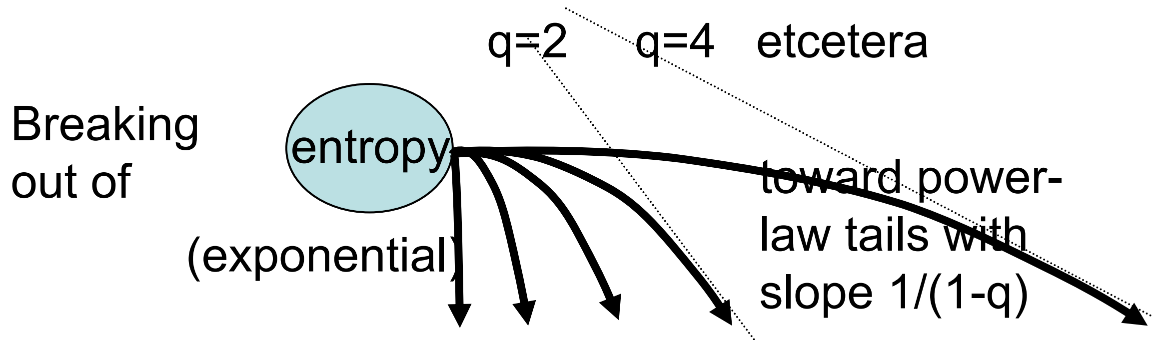
So what is Tsallis *q*-entropy?

It is a physical theory and mathematical model (of) how physical phenomena depart from randomness (entropy) but also fall back toward entropy at sufficiently small scale but that's only one side of the story, played out between: q=1 (entropy) and q>1, multiplicative effects as observed in power-law tendencies



That story
Is in Physical Review E 2006 by DW, CTsallis, NKejzar, et al.
for simulated feedback networks

So what's the other side of the story?



In the first part we had breakout from q=1 with *q increases* that *lower the slope*

Ok, now you have figured out that as q → 1 toward an infinite slope the q-entropy function converges to pure entropy, as measured by Boltzmann-Gibbs

But that's not all because there is another ordered state on the other side of entropy, where q (always ≥ 0) is less that 1! While $q > 1$ tends to power-law and q=1 converges to exponential (appropriate for BG entropy), $q < 1$ as it goes to 0 tends toward a simple linear function.

That story
is told in the Tsallis q-entropy equation

$$Yq \equiv Y0\ [1-(1-q)\ x/\kappa]^{1/(1-q)}$$

Ok, so given x, the variable sizes of cities, then Yq ≡ the q-exponential fitted to real data Y(x) by parameters Y0, κ, and q. *And* the q-exponential is simply the $e_q^{x'} \equiv x[1-(1-q)\,x']^{1/(1-q)}$ part of the function where it can be proven that $e_{q=1}^{x} \equiv e^{x}$ ≡ the measure of entropy. Then q is the metric measure of departure from entropy, in our two directions, above or below 1.

The story

is told in the Tsallis q-entropy equation

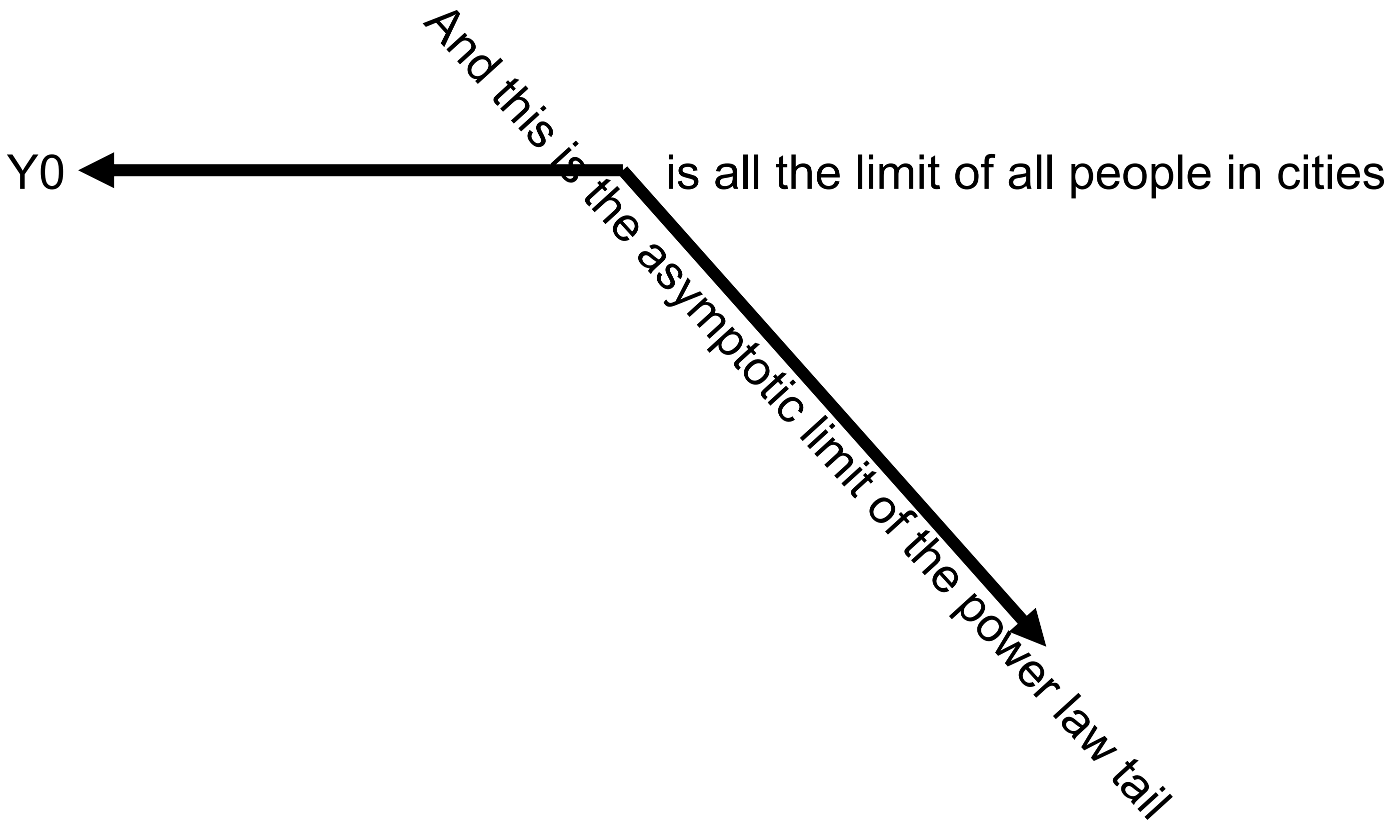
$$Yq \equiv Y0\,[1-(1-q)\,x/\kappa]^{1/(1-q)}$$

Ok, so now we know what *q* means, but what the parameters Y0 and κ? Well, remember: there are two asymptotes here, not just the asymptote to the power-law tail, but the asymptote to the smallness of scale at which the phenomena, such as "city of size x" no longer interacts with multiplier effects and may even cease to exist (are there cities with 10 people?)

This story
is told in the Tsallis q-entropy equation

$$Y_q \equiv Y_0 \, [1-(1-q)\, x/\kappa]^{1/(1-q)}$$

So, now let's look at the two asymptotes in the context of a cumulative distribution:

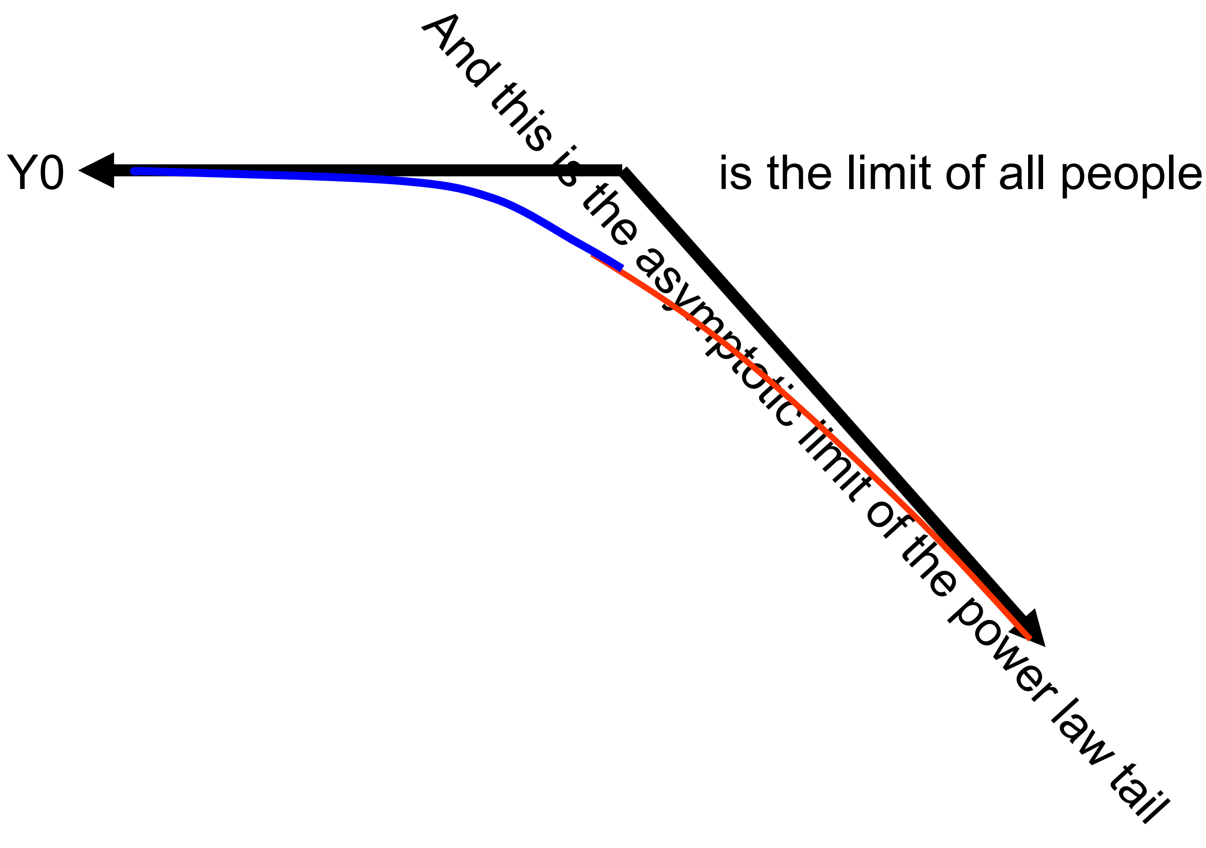


This story

is told in the Tsallis q-entropy equation

$$Y\mathit{q} \equiv Y0\ [1-(1-q)\ x/\kappa]^{1/(1-q)}$$

Here is a curve that fits these two asymptotes:

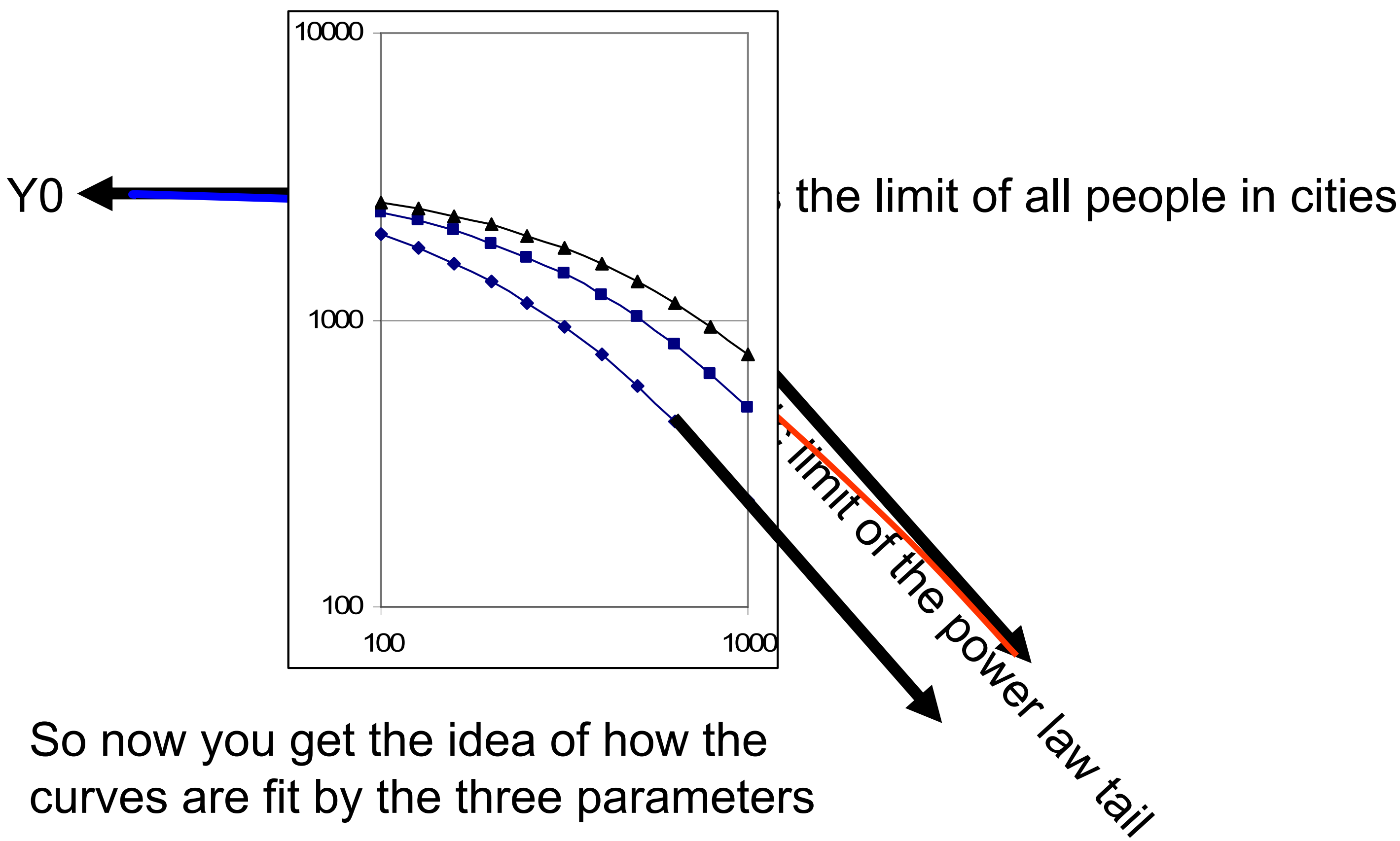


This story

is told in the Tsallis q-entropy equation

$$Yq \equiv Y0\,[1-(1-q)\,x/\kappa]^{1/(1-q)}$$

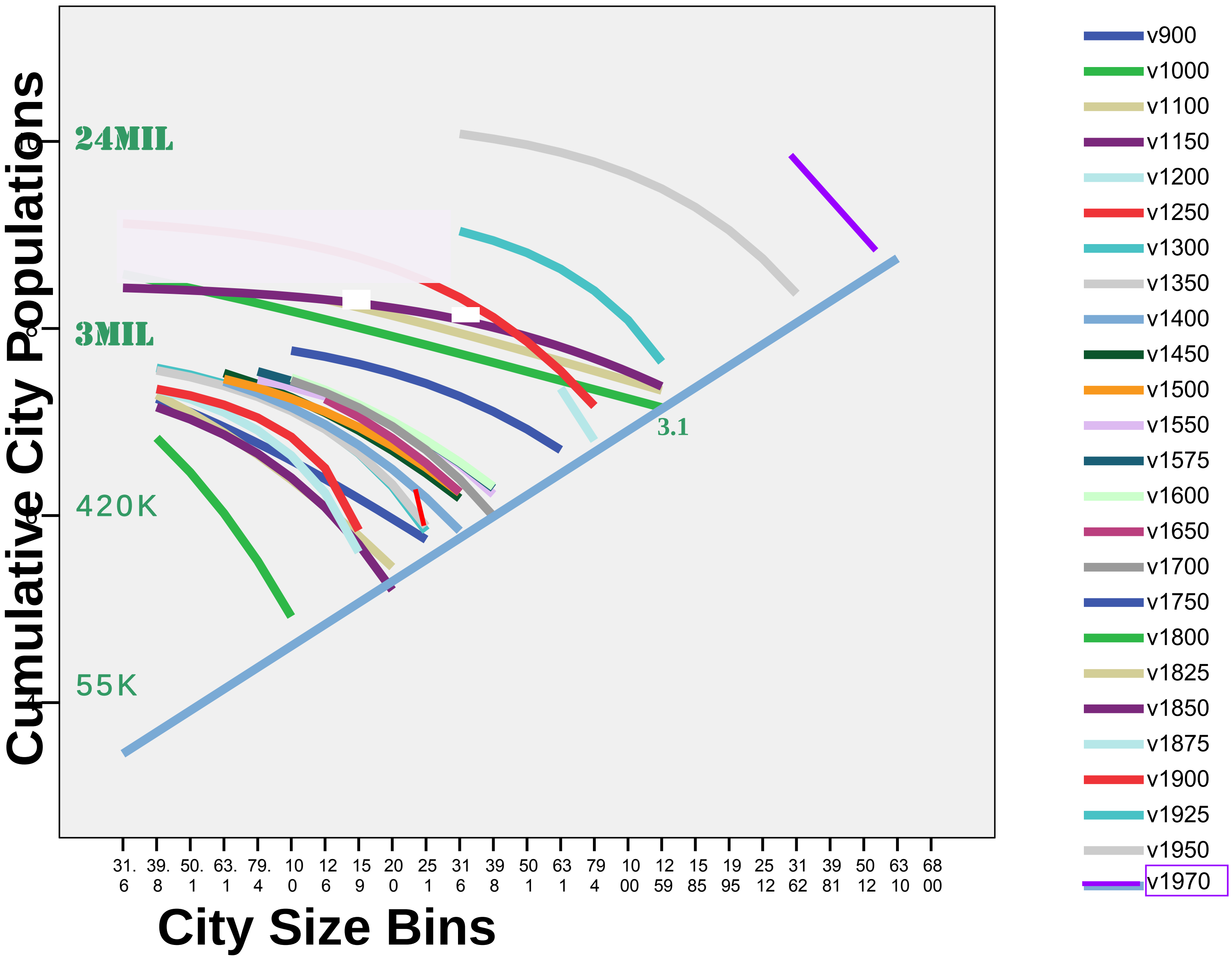
Here are three curves with the same Y0 and *q but different k*



So now you get the idea of how the curves are fit by the three parameters

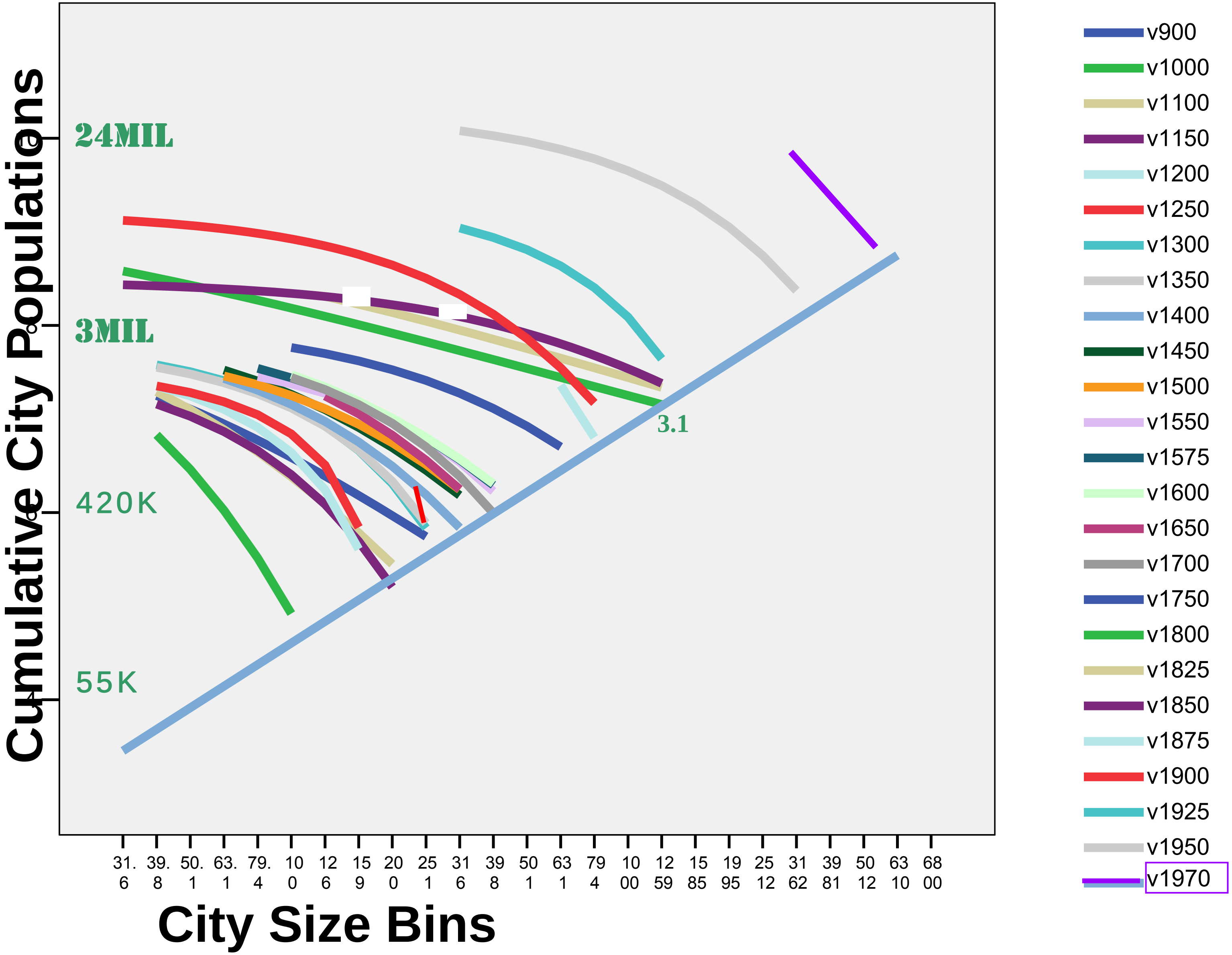
This story
is told in the Tsallis q-entropy equation

$$Yq \equiv Y0\ [1-(1-q)\ x/\kappa]^{1/(1-q)}$$



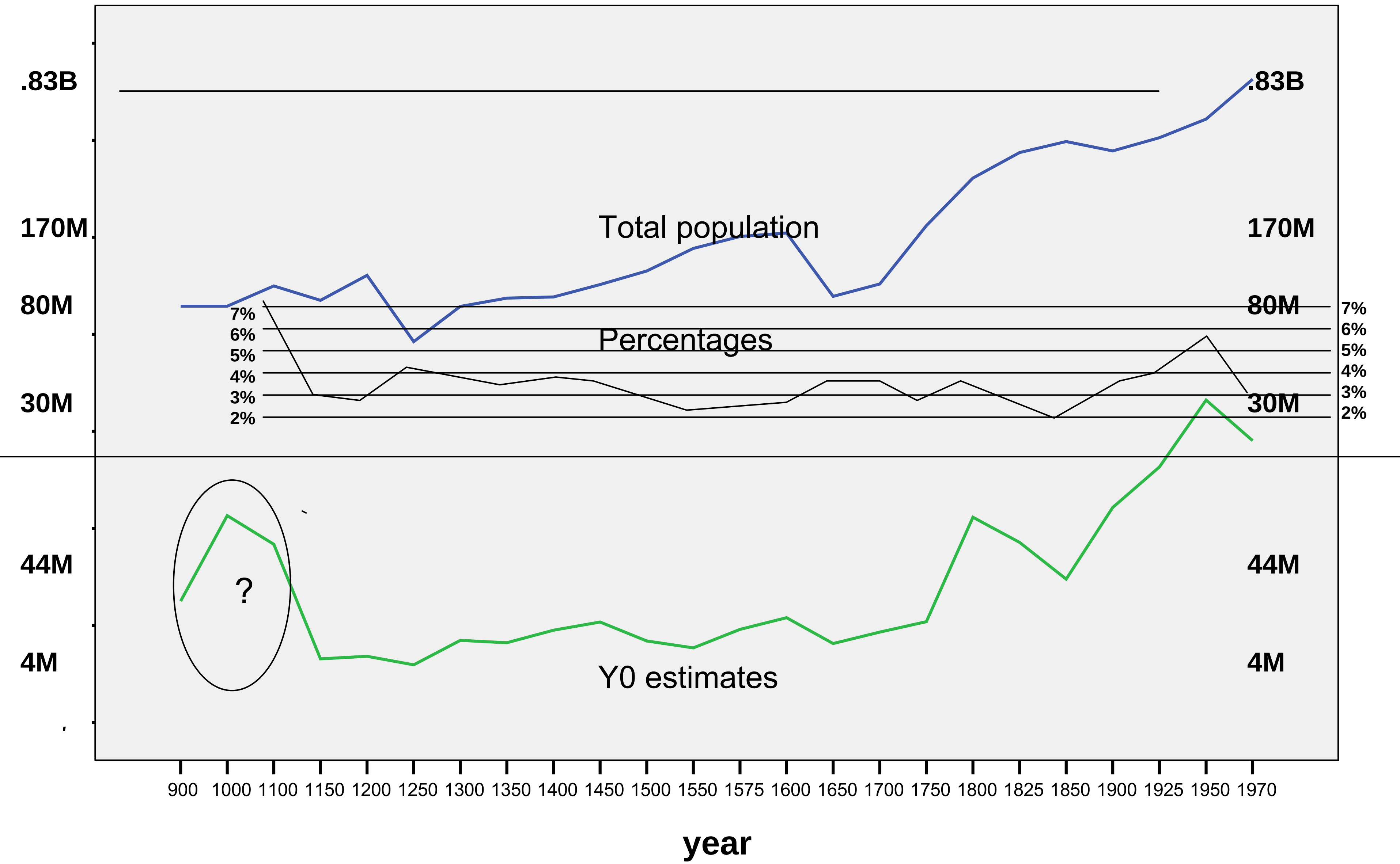
24MIL
3MIL
420K
55K
3.1
Cumulative City Populations
31.6
39.8
50.1
63.1
79.4
100
126
159
200
251
316
398
501
631
794
1000
1259
1585
1995
2512
3162
3981
5012
6310
6800
City Size Bins
v900
v1000
v1100
v1150
v1200
v1250
v1300
v1350
v1400
v1450
v1500
v1550
v1575
v1600
v1650
v1700
v1750
v1800
v1825
v1850
v1875
v1900
v1925
v1950
v1970

One feature in these fits is the estimate of Y0 (total urban populations)



China log population, log estimate Y0: urban population, and estimated % urban

(the estimates of Y0 are in exactly the right ratios to total population and %ages)



q *runs test: 8 Q-periods* (p=.06)

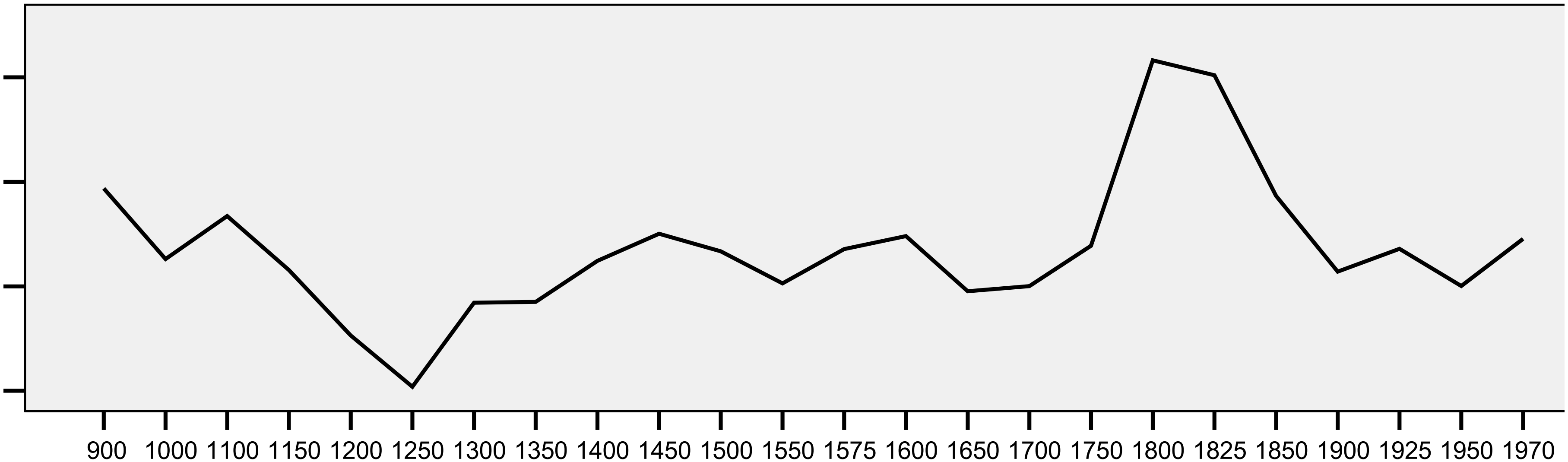


Table 1: Example of bootstrapped parameter estimates for 1650

Parameter Estimates

			95% Confidence Interval		95% Trimmed Range	
Parameter	Estimate	Std. Error	Lower Bound	Upper Bound	Lower Bound	Upper Bound
q	.795	.094	.608	.983	.795	.795
k	229.307	6.854	215.592	243.022	229.307	229.307
Y	2471.785	3.307	2465.167	2478.403	2471.785	2471.785

a. Based on 60 samples.

b. Loss function value equals 4161.644.

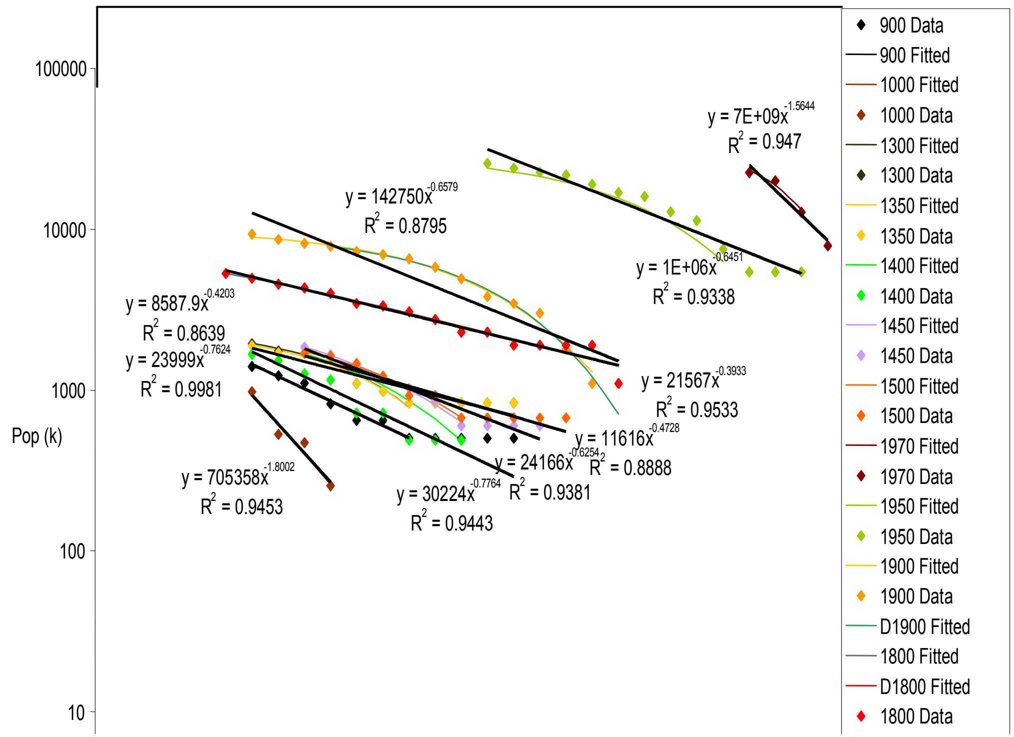
Average R^2

Power law fits

.93

q entropy fits

.984



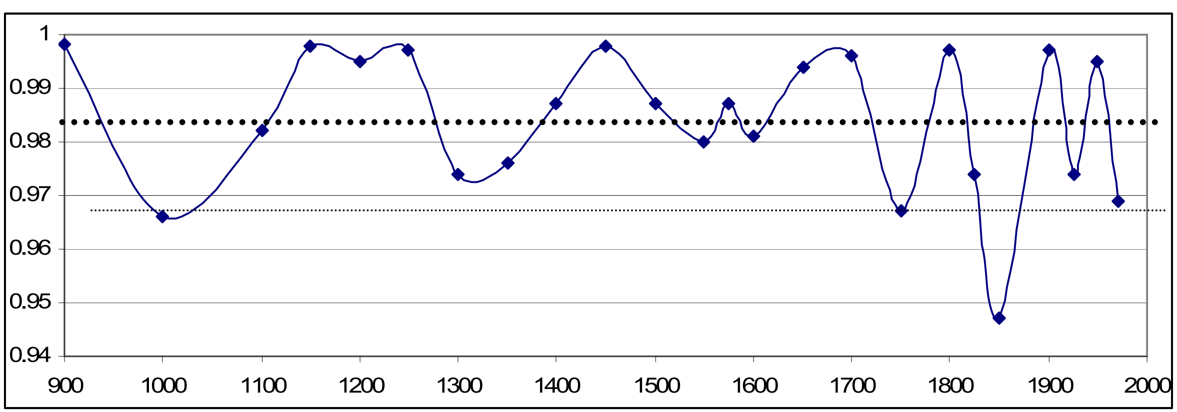


Figure 4: Variation in R^2 fit for *q* to the *q*-entropy model – China 900-1970

Key: Mean value for runs test shown by dotted line.

commensurability & lowest bin convergence to Y0

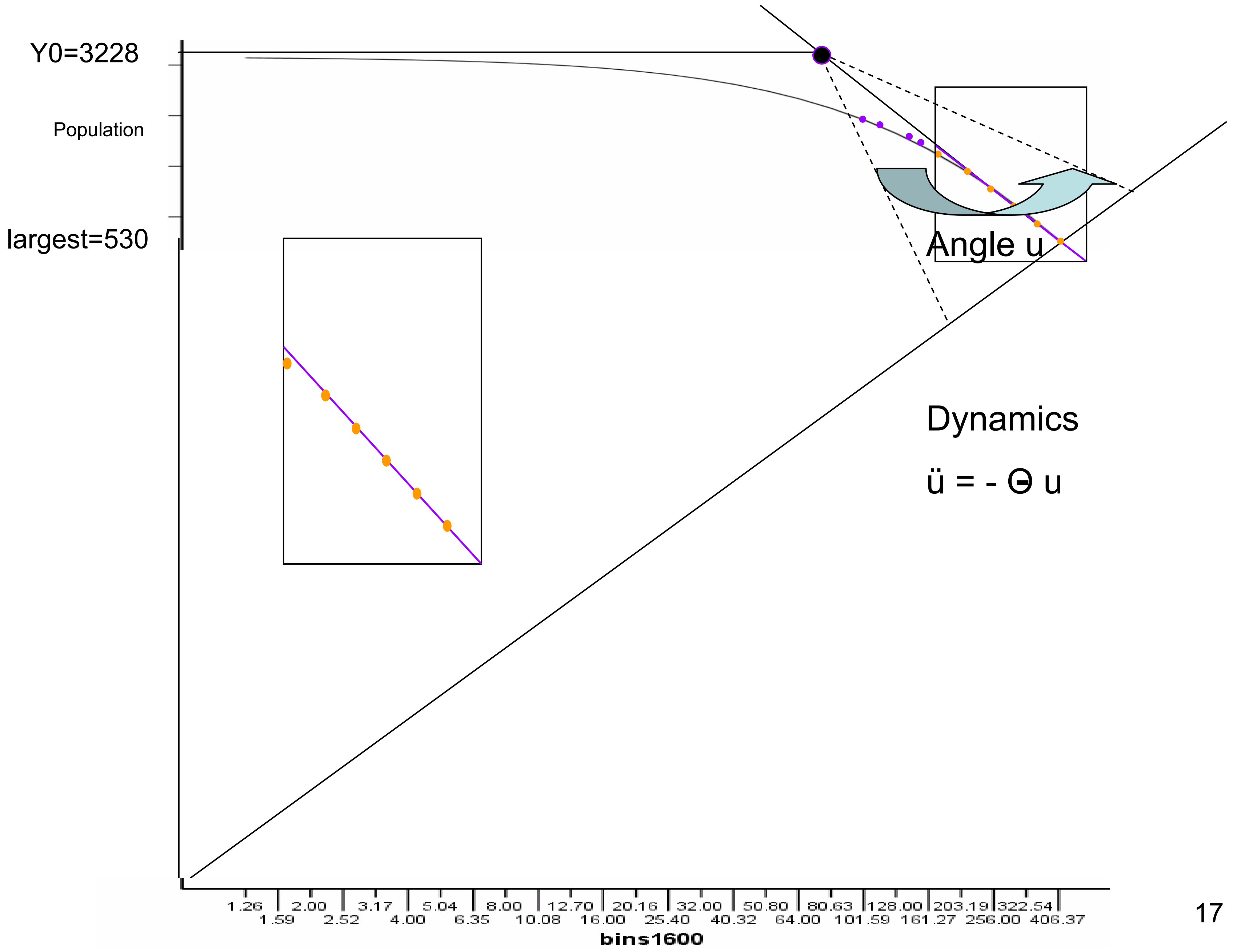
Table 2: Correlations among the commensurate-ordering variables in Table 3

	Pop	Y0	31.6K	Communalities
Total Chinese Population				.88
Y0 Estimate	.75**			.95
Bin Estimate at 31.6K	.81**	.96**		.97
K	.70**	.81**	.90**	.91

* p <.05 ** p < .01

:

At city bin size 10±2 thousand or greater 95% of the city distribution (as a fraction of Y0) is present: this is the effective (smallest) city sizes for all periods



Y0=3228
Population
largest=530
Angle u
Dynamics
ü = - Θ u
1.26 2.00 3.17 5.04 8.00 12.70 20.16 32.00 50.80 80.63 128.00 203.19 322.54
1.59 2.52 4.00 6.35 10.08 16.00 25.40 40.32 64.00 101.59 161.27 256.00 406.37
bins1600

Table 6: Total Chinese population oscillations and *q*

q ranges	Endogenous secular population cycle				Exceptions	
	'Early' pop. rise	'Late' pop. rise	Population Maximum	Crash	Economy Captured	Exception *deurbanized*
q~3 'abnormal'					*1800 2.77* *1825 2.99*	
q~1.7 'rigid'			*1100 1.72* *1850 1.85*			
q~1.5 Zipfian		*1000 1.37* *1450 1.50* *1500 1.34* *1925 1.39*	*1575 1.35* *1600 1.48*	*1150 1.4*		*1970 1.49*
q~1 'random'	*1300 0.85* *1350 0.85* *1400 1.24* *1700 1.00* *1750 1.29* *1900 1.14*		*1550 1.04* *1950 1.06*			
q~.5 - .8 'chaotic'			*1200 0.54*	*1650 0.8* *1875 <1?*		
q~0 'flee the cities'				*1250 0.02*		

Turchin's secular cycle dynamic-China

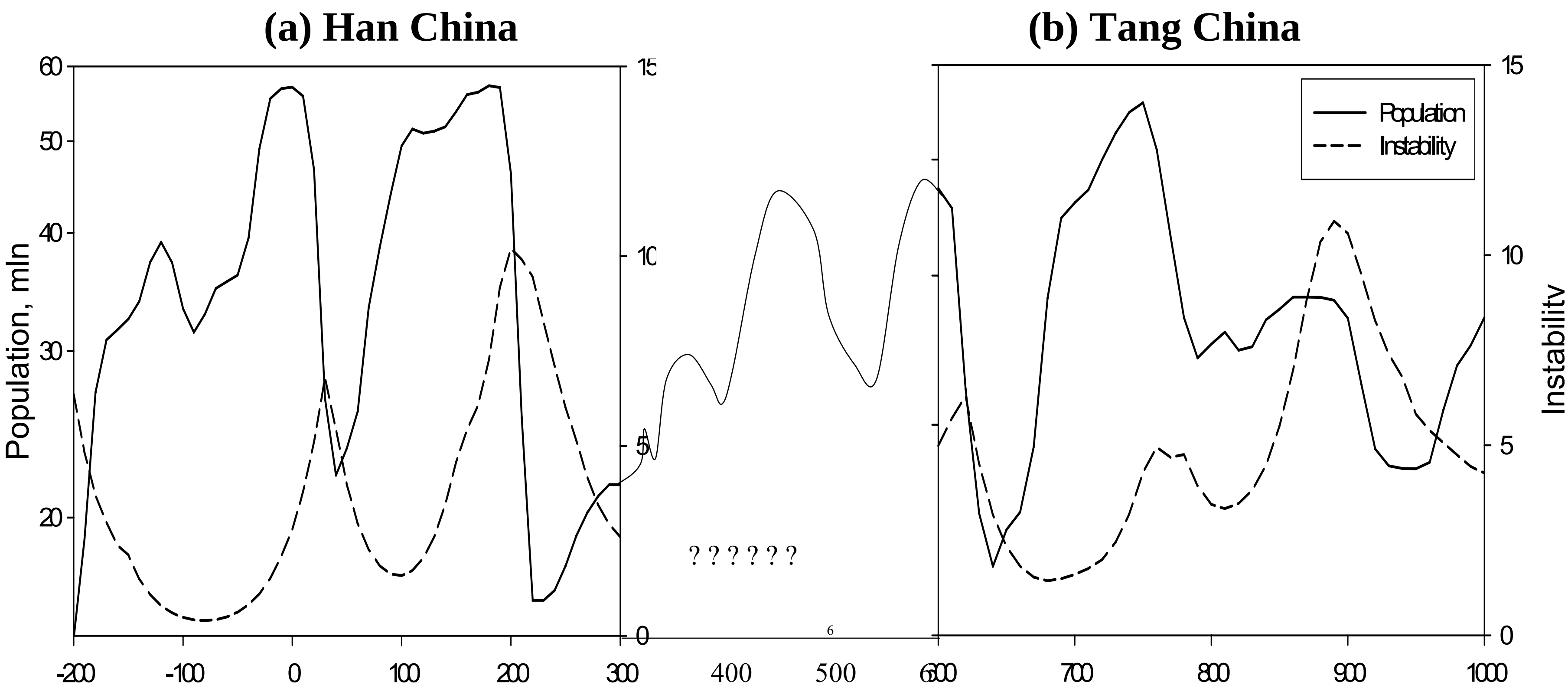


Figure 8: Turchin secular cycles graphs for China up to 1100

Note: (a) and (b) are from Turchin (2005), with population numbers between the Han and Tang Dynasties filled in. Sociopolitical instability in the gap between Turchin's Han and Tang graphs has not been measured.

Example: Kohler on Chaco

Kohler, et al. (2006) have replicated such cycles for pre-state Southwestern Colorado for the pre-Chacoan, Chacoan, and post-Chacoan, CE 600–1300, for which they have “one of the most accurate and precise demographic datasets for any prehistoric society in the world.” Secular oscillation correctly models those periods “when this area is a more or less closed system,” but, just as Turchin would have it, not in the “open-systems” period, where it “fits poorly during the time [a 200 year period] when this area is heavily influenced first by the spread of the Chacoan system, and then by its collapse and the local political reorganization that follows.”

Relative regional closure is a precondition of the applicability of the model of endogenous oscillation.

Kohler et al. note that their findings support Turchin’s model in terms of being “helpful in isolating periods in which the relationship between violence and population size is not as expected.

City Systems
China – Middle Asia – Europe
World system interaction dynamics

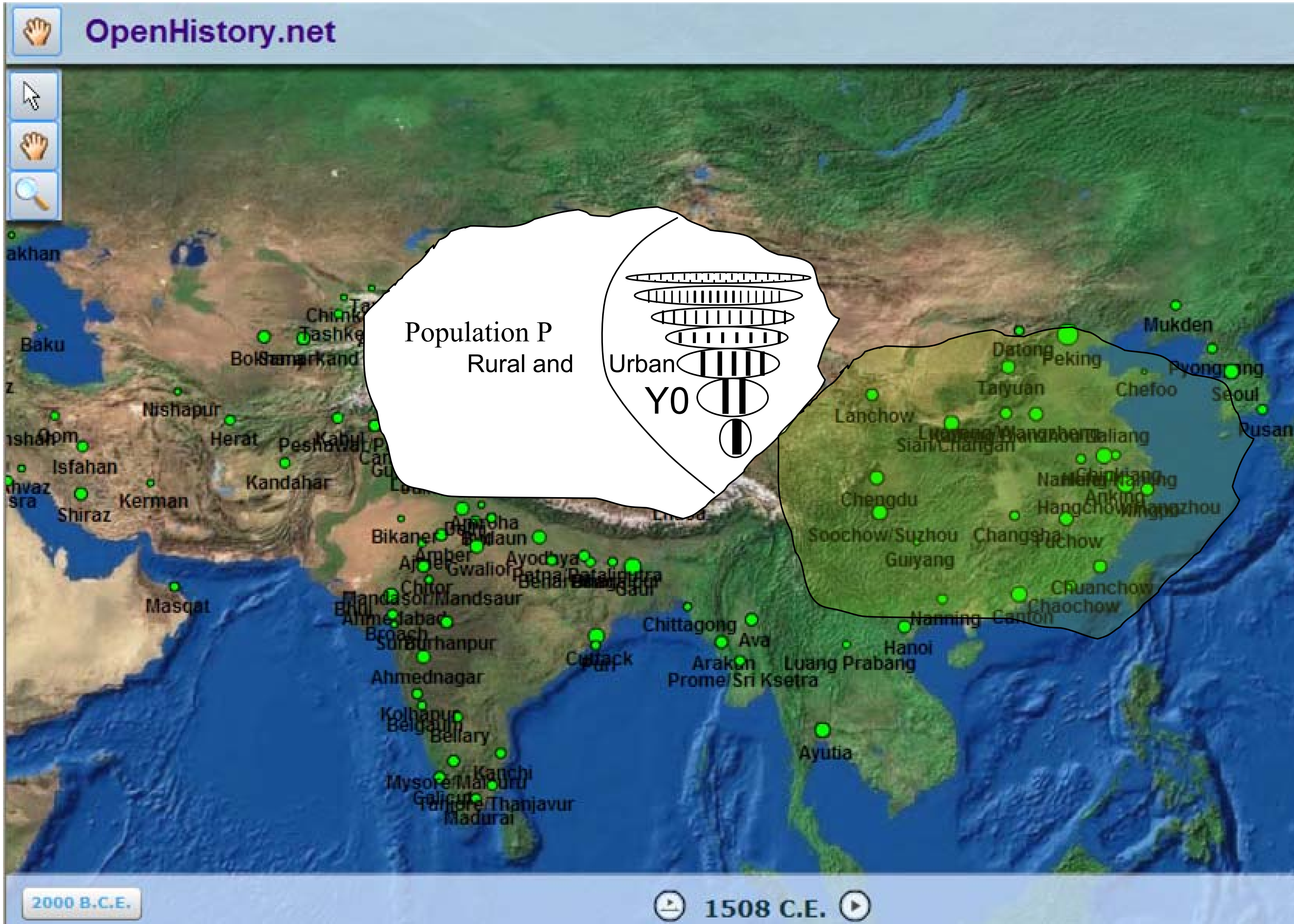
The basic idea of this series is to look at rise and fall of cities embedded in networks of exchange in different regions over the last millennium… and

How innovation or decline in one region affects the other

How cityrise and cityfall periods relate to the cycles of population and sociopolitical instability described by Turchin (endogenous dynamics in periods of relative closure)

How to expand models of historical dynamics from closed-period endogenous dynamics to economic relationships and conflict between regions or polities, i.e., world system interaction dynamics

Sufficient statistics to include population and *q* parameters plus spatial distribution and network configurations of transport links among cities of different sizes and functions.

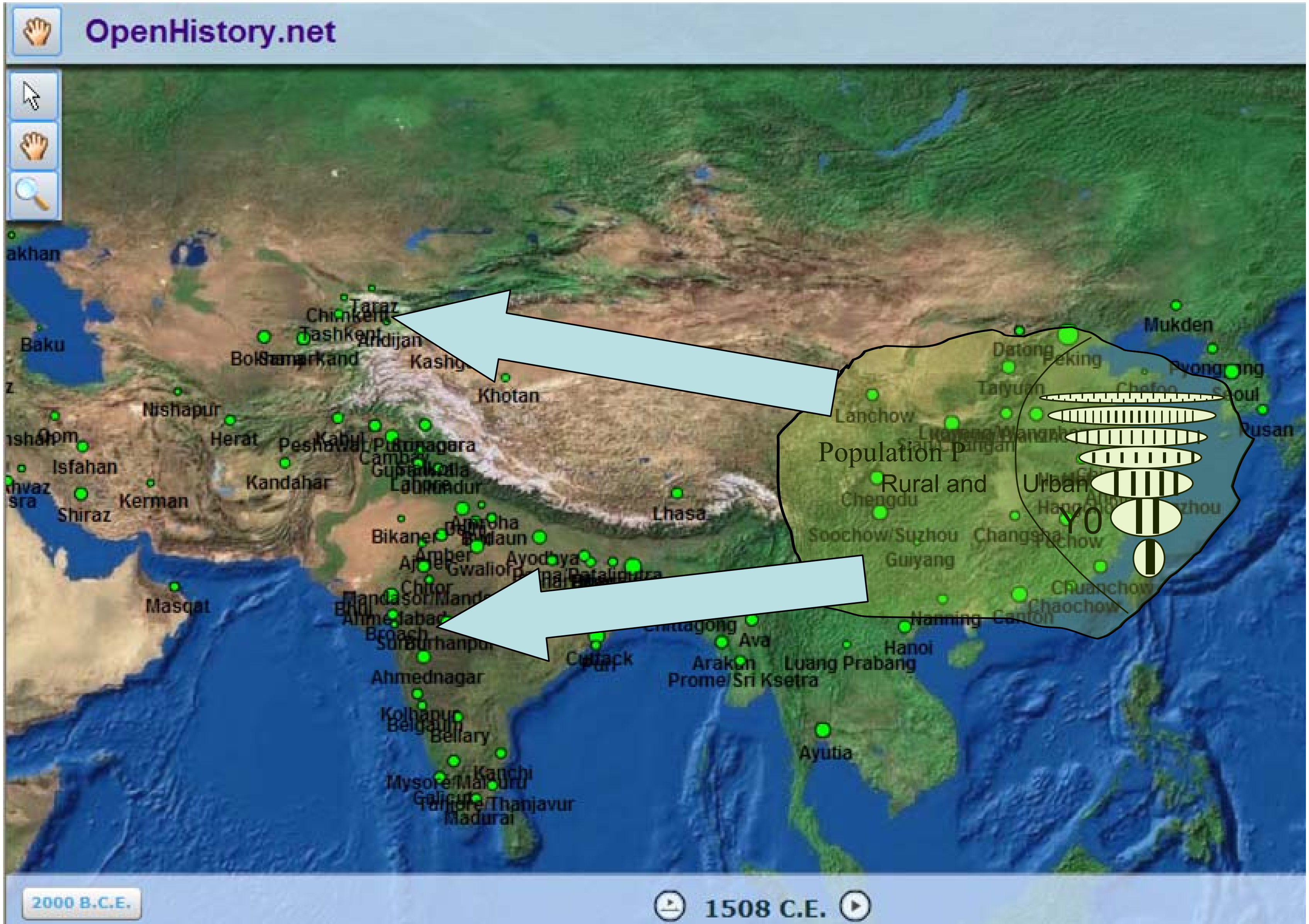


China – Middle Asia - Europe

The basic idea of the next series will be to measure the time lag correlation between variations of q in China and those in the Middle East/India, and Europe.

This will provide evidence that q provides a measure of city topology that relates to city function and to city growth, and that diffusions from regions of innovation to regions of borrowing

Sufficient statistics to include population and *q* parameters plus spatial distribution and network configurations of transport links among cities of different sizes and functions.



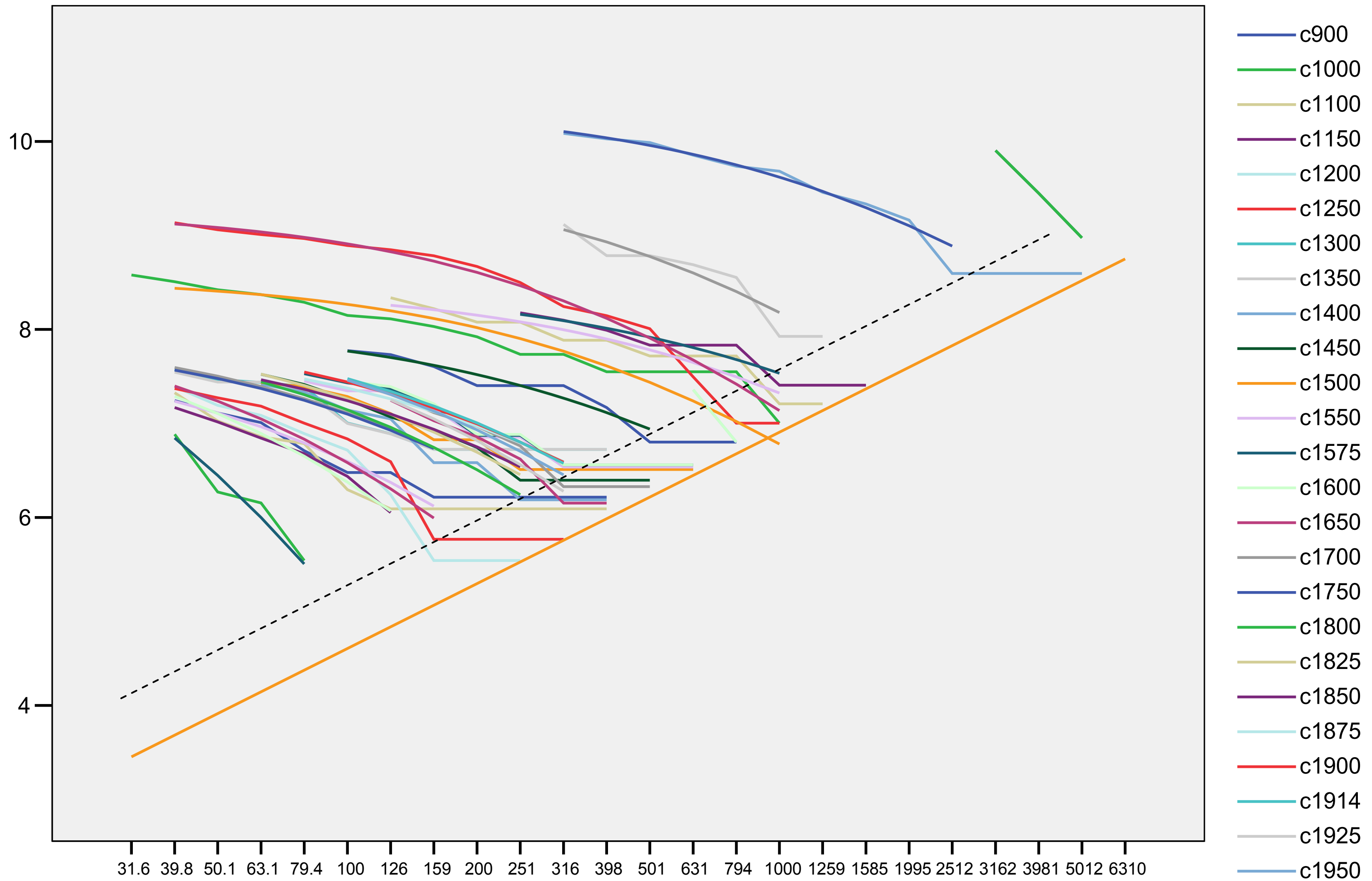


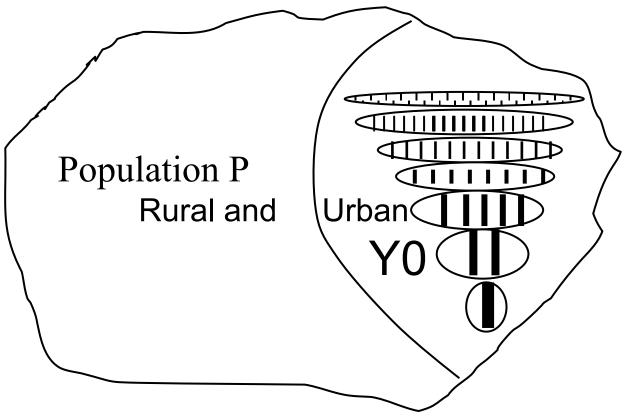
Figure 5: Chinese Cities, fitted *q*-lines and actual population size data

Table 6: Total Chinese population oscillations and *q*

q ranges	Endogenous secular population cycle				Exceptions	
	‘Late’ pop. rise	Population Maximum	Crash	‘Early’ pop. rise	Economy Captured	Exception *deurbanized*
q~3 ‘abnormal’					*1800 2.77* *1825 2.99*	
q~1.7 ‘rigid’		*1100 1.72* *1850 1.85*				
q~1.5 Zipfian	*1925 1.39*	*1600 1.48*	*1150 1.4*			*1970 1.49*
q~1 ‘random’		*1550 1.04* *1950 1.06*		*1300 0.85* *1350 0.85* *1400 1.24* *1700 1.00* *1750 1.29* *1900 1.14*		
q~.5 - .8 ‘chaotic’		*1200 0.54*	*1650 0.8* *1875 <1?*			
q~0 ‘flee the cities’			*1250 0.02*			

China – Middle Asia - Europe

The basic idea of this series of



Population P
Rural and Urban
Y0

q and κ

Modeling City Size Data with a Double-Asymptotic Model (Tsallis *q*-entropy)
Deriving the two Asymptotic Coefficients (*q*,Y0)
and the crossover parameter (kappa: κ)
for 24 historical periods, 900-1970
from Chandler’s data in the largest world cities in each
checking that variations in the parameters
for adjacent periods entail real urban system variation
and that these variations characterize historical periods
then testing hypotheses about how these variations
tie in to what is known about
World system interaction dynamics

good lord, man, why would you want to do all this?

That will be the story