Cities and Countries

Andrew K Rose

UC Berkeley, CEPR and NBER

Motivation

- Cities are a basic unit of urban economics
- Countries are a basic unit of international economics
- There is no obvious reason why the distribution of one should look the distribution of the other

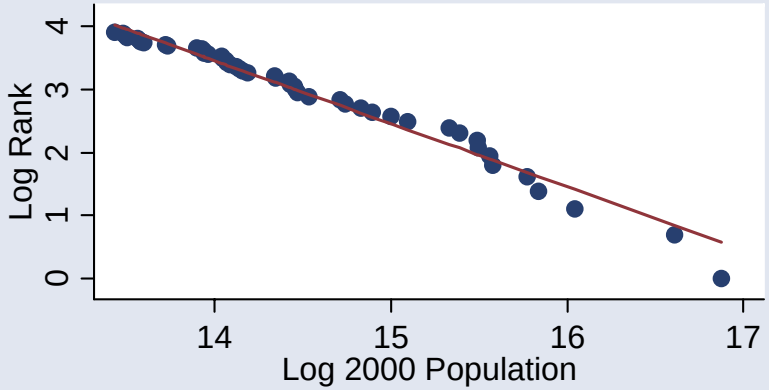
The Issue

- *But they do!*

Cities and Countries

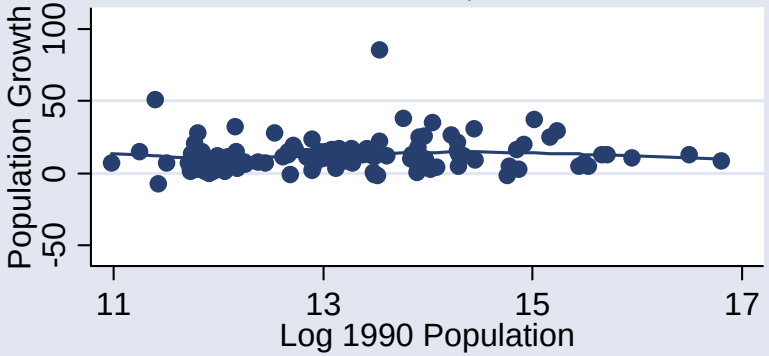
Zipf's Law: Log Size Rank

50 largest US Cities, 2000

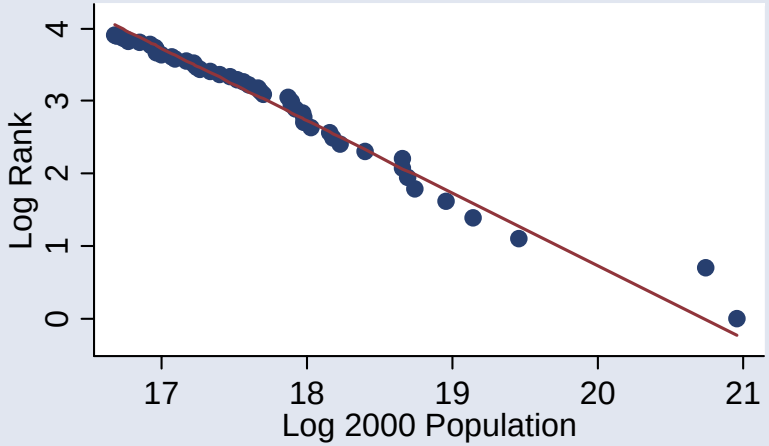


Gibrat's Law: Population Growth

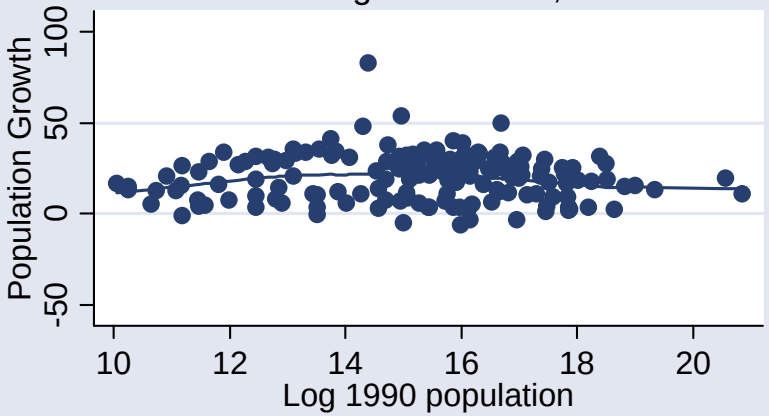
All 113 US Cities, 1990-2000



50 largest Countries, 2000



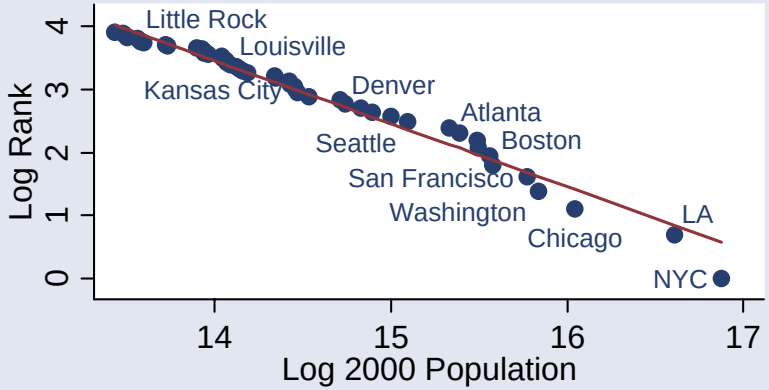
All 163 Sovereign Countries, 1990-2000



Cities and Countries

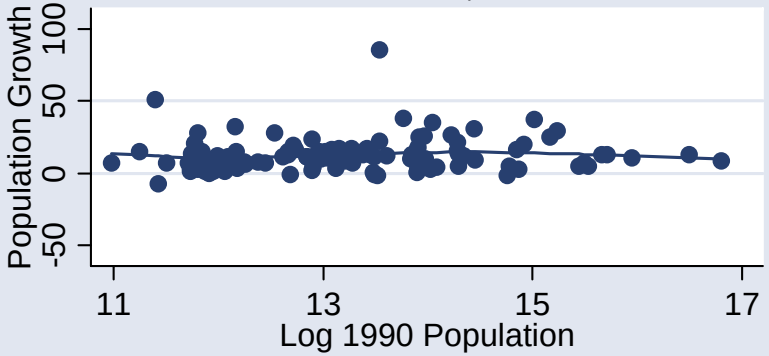
Zipf's Law: Log Size Rank

50 largest US Cities, 2000

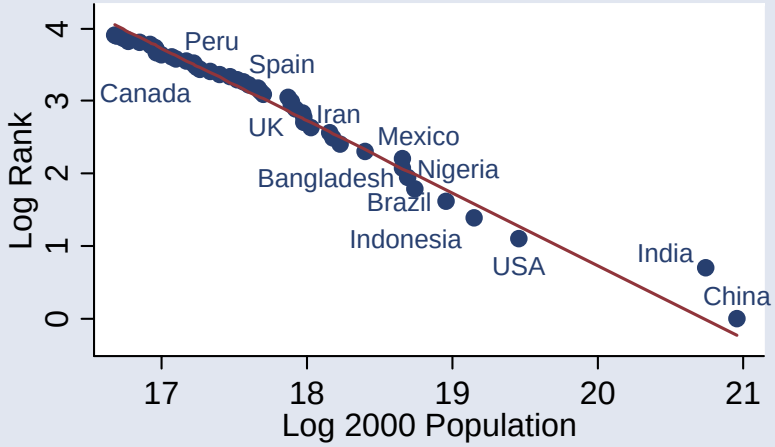


Gibrat's Law: Population Growth

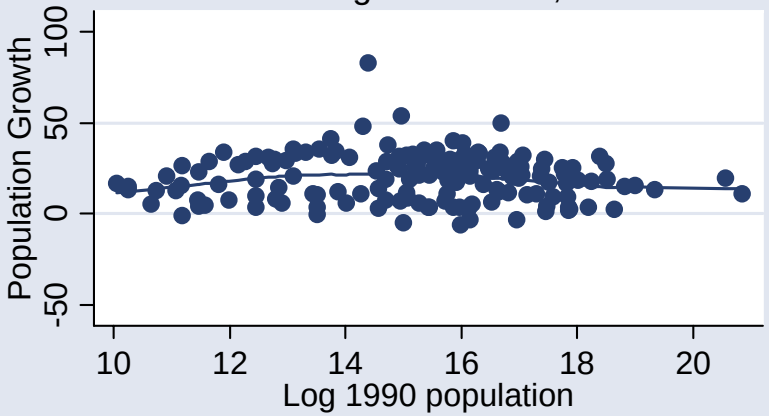
All 113 US Cities, 1990-2000



50 largest Countries, 2000



All 163 Sovereign Countries, 1990-2000



I focus on two aspects of size distribution:

1) "Zipf's Law [for Cities]"
   - The rank (by size) of a city is almost perfectly inversely correlated with its size
2) "Gibrat's Law [for Cities]"
   - The growth of a city is basically independent of its size

- Both are well-known, widely documented, undisputed
- Both have prompted much theorizing

Zipf's Law [for Cities]

- Rank cities by size: $S_1 > S_2 > \ldots > S_N$.
- Zipf: $P(Size > S) \approx \alpha S^{-\beta}$ where: α a constant, $\beta \approx 1$
- Much recent work: Eeckhout (2004), Gabaix (1999), Krugman (1996), Rossi-Hansberg and Wright (2004); Gabaix and Ioannides (2004).
- Can check via:
  1. Graphs
  2. Regressions of ln(i) on constant and $\ln(S_i)$

Smaller Factoids on Zipf's Law [for Cities]:

1. Broader Sample lowers slope (too few small cities)
2. Narrowing Definition raises slope
3. Works across time and countries

Gibrat's Law [for Cities]

- Expected Growth of City independent of initial size
- Again, much recent work: Eeckhout (2004), Gabaix (1999), Gabaix and Ioannides (2004).

My Strategy

- Replicate stylized facts for cities
- Check analogues for countries

Small Issue: What is a "Country"?

- No standard definition of "country" exists
- Two economic definitions:
  1. Ricardo: area within which factors are mobile, between which factors are immobile
  2. Political: area controlled by government with monopoly of legal coercion

- In practice I use areas considered by *WDI* and also check results on independent sovereign national states

Unimportant in Practice

- Most questionable "countries" are small
- Still, interesting to note existence of:
  1. SARs (e.g., Hong Kong)
  2. Associated States (Puerto Rico)
  3. Colonies (Cayman Islands)
  4. Overseas Possessions (Reunion)
  5. ? (West Bank and Gaza)

Data Sources

American Cities from Census

- Combined Statistical Areas (CSAs)
- Metropolitan/Micropolitan Statistical Areas (MSAs)
- Census Designated Places (CDPs)

Countries

- 1900 from *Statesman's Yearbook 1901*
- 1950 from Census
- 1960-2000 from *World Development Indicators*
- 2004 from CIAs *World Factbook*
- 2050 projection from Census

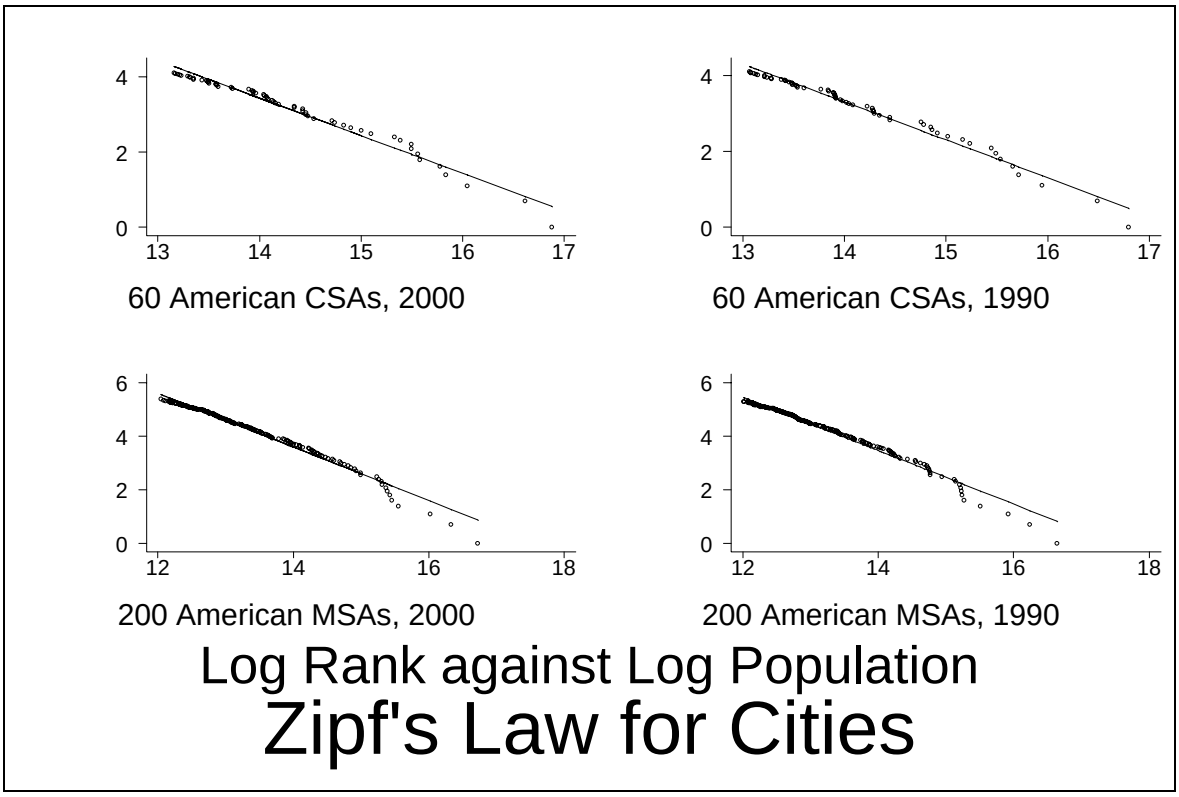


Figure 1: Size Distribution of Cities

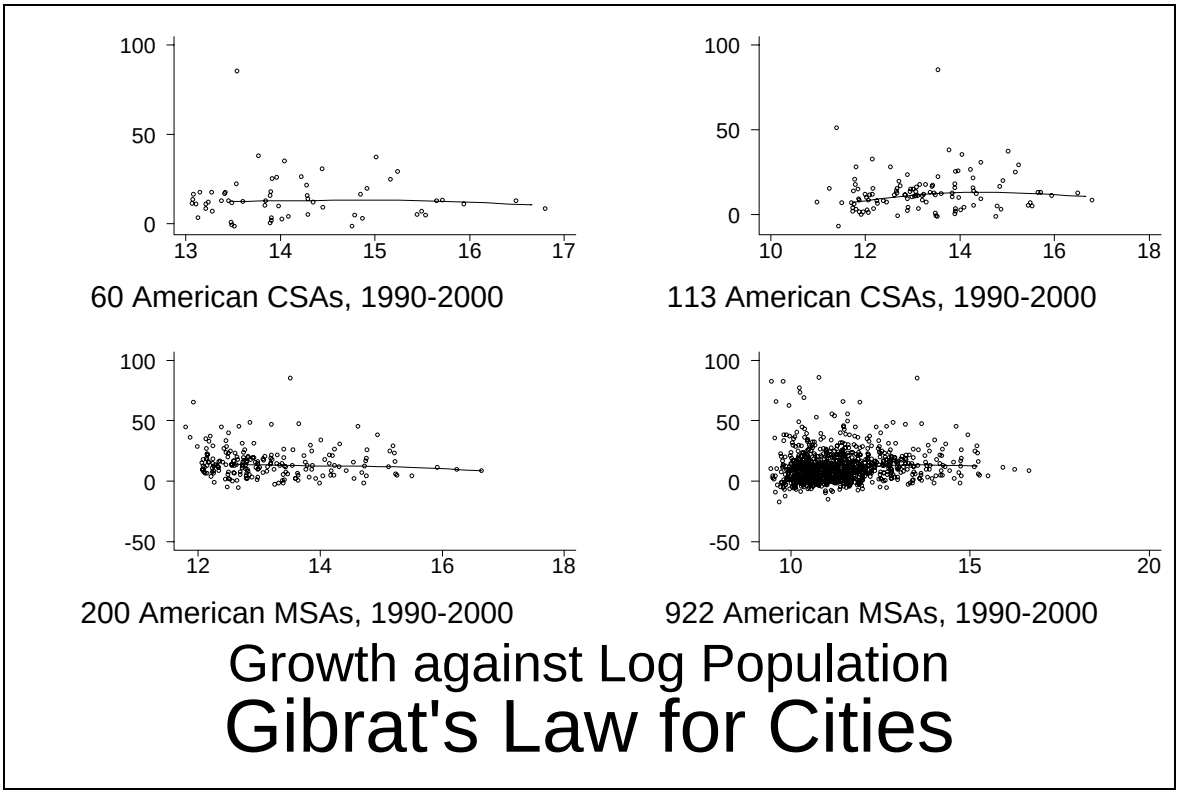


Figure 2: City Population Growth Rates

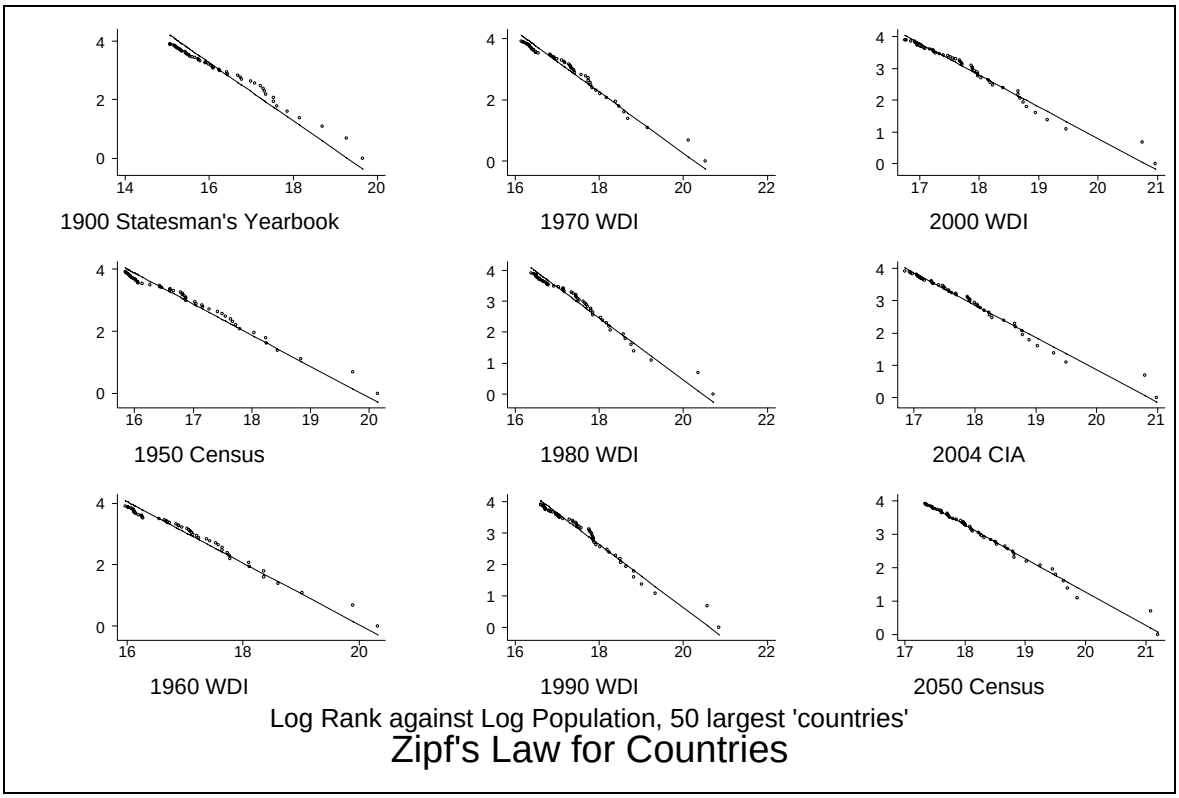


Figure 3: Size Distribution of Countries

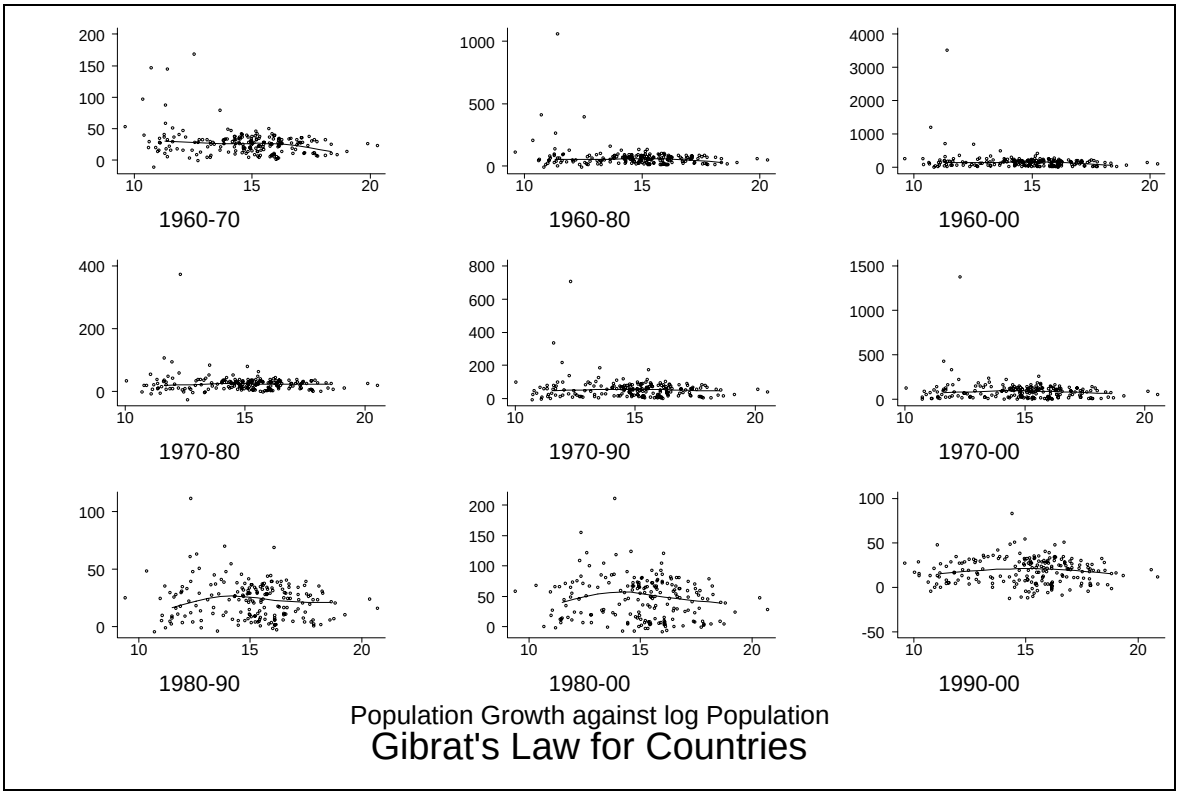


Figure 4: Country Population Growth Rates

Table 1: Zipf City Coefficients

Year	City Measure	Sample	Slope (se)	R^2
2000	CSAs	50	-1.03 (.21)	.98
1990	CSAs	50	-1.03 (.21)	.98
2000	MSAs	200	-1.01 (.1)	.98
1990	MSAs	200	-1.02 (.1)	.98
2000	CSAs	113	-.73 (.10)	.93
1990	CSAs	113	-.74 (.10)	.93
2000	MSAs	922	-.82 (.04)	.98
1990	MSAs	922	-.83 (.04)	.98
2000	CDPs	601	-1.34 (.08)	.998

Coefficients are slopes from OLS regressions of log rank on log population.

Table 2: Gibrat City Coefficients

City Measure	Sample	Slope (se)	R^2
CSAs	50	-1.48 (2.08)	.01
MSAs	200	-.01 (.78)	.00
CSAs	113	.97 (.82)	.01
MSAs	922	1.07** (.39)	.01

Table 3: Zipf Country Coefficients

Year	Slope (se)	R^2
1900	-.78 (.16)	.99
1950	-.87 (.17)	.99
1960	-.88 (.18)	.98
1970	-.89 (.18)	.98
1980	-.91 (.18)	.98
1990	-.93 (.19)	.98
2000	-.95 (.19)	.98
2004	-.96 (.19)	.98
2050	-.99 (.20)	.99

Table 4: Gibrat Country Coefficients

		All	All	Sov's	Sov's	Top 50	Top 50
Initial Year	Final Year	Slope (se)	R^2	Slope (se)	R^2	Slope (se)	R^2
1960	1970	-3** (1.0)	.08	-.7 (.8)	.01	-1.1 (1.4)	.01
1960	1980	-9.3* (4.6)	.05	-5.0* (2.0)	.02	-3.0 (3.3)	.01
1960	1990	-17* (8.4)	.05	-10** (3.5)	.07	-5.1 (5.7)	.01
1960	2000	-26.6 (14.4)	.04	-20** (5.6)	.11	-8.9 (8.5)	.02
1970	1980	-1.8 (1.36)	.01	-1.3 (.9)	.02	-2.7 (1.5)	.05
1970	1990	-4.3 (2.7)	.02	-2.9 (1.8)	.02	-6.1 (3.5)	.05
1970	2000	-7.8 (4.9)	.02	-7.3* (3.0)	.02	-11.2 (5.8)	.06
1980	1990	-.8 (.6)	.01	-.9 (.7)	.01	-2.4 (1.4)	.04
1980	2000	-1.7 (1.1)	.01	-3.6* (1.5)	.04	-6.1 (3.2)	.05
1990	2000	-.1 (.4)	.00	-1.2 (.7)	.02	-2.5 (1.6)	.04

Cities work as usual

- Zipf's Law (and deviations) works well
- Gibrat's Law works well

Countries work basically as well

- Zipf's Law (and deviations) works well
  - Slopes close to -1, insignificantly different
  - High goodness of fit
    - 1900 the biggest exception
    - Broader sample lowers slope
    - Exact definition of "country" unimportant
    - Works across time

- Gibrat's Law works pretty well too
  - Some signs of negative significant relationships
  - Poor Fit

Tangent: Log-Normality describes Country Populations

- Little Kurtosis
- Some skewness (too fit small countries)
- Can't reject statistically 1960-2000

So what?

- An empirical regularity: size distribution of cities similar to size distribution of countries
- Theoretical Explanation?
  - Little work on size of countries (except Alesina-Spolaore, who don't study distribution)

Much theoretical work on city-size distributions

- Ex: Eeckhout (2004), Krugman (1996), Rossi-Hansberg and Wright (2004)
- All balance agglomeration benefits (knowledge spillovers, scale economies …) with negative externalities (congestion, commuting costs, land prices, …)
- Need to have and balance both externalities to induce mobile labor to migrate between cities appropriately

City-Size theory not easily applicable to countries!

- Countries control policies, institutions more
- Mobility higher between cities inside country than between countries
- Externalities, agglomeration effects, amenity shocks, congestion costs, scale economies … all more plausible at local than national level

But common empirical regularity makes common theoretical explanation natural.

Conclusion

- Cities and Countries both adhere reasonably well to:
  - a) Zipf's (size-rank) Law; and
  - b) Gibrat's (growth) law
- Common empirical resemblance cries out for common theoretical explanation