EC1410 Topic #8

Systems of Cities

Matthew A. Turner
Brown University
Spring 2023

(Updated January 6, 2023)

Outline

1. Some stylized facts
2. Zipf's law
3. Systems of Cities

Stylized facts about cities

- What does the size distribution of cities look like? How does it change over time? Do cities change their position?
- What are patterns of sectoral specialization? Does this change over time?
- Can we explain these patterns as a consequence of spatial equilibrium?

This discussion draws on Duranton and Puga (2000).

To begin, we need to define 'specialized' and 'diversified' cities. Let

$$s_{ij} \sim \text{Share of industry } j \text{ employment in city } i$$
$$ZI_i \equiv \max_j(s_{ij}) \sim \text{ specialization}$$

Tthe specialization of a city is the share of employment in its largest sector. Its also useful to think about relative specialization,

$$s_j \sim \text{Share of industry } j \text{ national employment.}$$
$$RZI_i \equiv \max_j(s_{ij}/s_j) \sim \text{relative specialization}$$

Providence is relatively specialized in the manufacture of submarines, even though it is a small share of overall employment.

Measuring diversity is a little trickier. What does it mean to say that one set of jobs is ‘less different than another’?

Consider, the two sets of jobs,

$$\{\text{2 mechanics, 1 brain surgeon}\}$$
$$\{\text{1 Banker, 1 Baker, 1 Carpenter}\}$$

Which is more diverse?

The industry standard for answering this question is the ‘Herfindahl Index’,

$$DI_i = \left(\sum_j s_{ij}^2\right)^{-1}$$

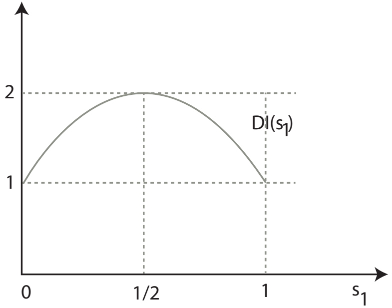
For s_{ij} the share of industry j in city i’s employment.

To see how this works, suppose $j = 1, 2$ and drop the i subscript.

Then we have $s_2 = 1 - s_2$ and

$$DI_i = (s_1^2 + s_2^2)^{-1} = (s_1^2 + (1 - s_1)^2)^{-1}$$
$$= (s_1^2 + (1 - 2s_1 + s_1^2))^{-1} = (2s_1^2 + 1 - 2s_1)^{-1}$$

If we plot this as s_1 ranges from 0 to 1, we get something like this,



So 'most diverse' means employment is uniformly distributed across all alternatives.

The corresponding relative diversity index is also sometimes useful,

$$RDI_i = \left(\sum_j (s_{ij} - s_j)^2 \right)^{-1}$$

- Most of the literature is based on indexes like those defined here.
- A city can be relatively specialized and relatively diverse!

With definitions of specialization and diversity in place, we can state our first two stylized facts.

FACT 1: *specialized and diversified cities coexist*.

Table 1. Most and least specialised and diversified US cities in 1992

	Specialisation		Diversity	
Rank	City (sector)	RZI	City	RDI
1	Richmond, VA (tobacco)	64.4	Cincinnati, OH	166.6
2	Macon, GA (tobacco)	55.0	Oakland, CA	161.2
3	Lewiston, ME (leather)	49.6	Atlanta, GA	159.4
4	Galveston, TX (petroleum)	49.1	Philadelphia, PA	151.4
5	Bangor, ME (leather)	45.6	Salt Lake City, UT	120.8
6	Owensboro, KY (tobacco)	44.4	Buffalo, NY	110.1
7	Corpus Christi, TX (petroleum)	37.6	Columbus, OH	108.3
8	Cheyenne, WY (petroleum)	33.4	Portland, OR	94.1
315	Buffalo, NY (rubber and plastics)	1.6	Lawton, OK	2.4
316	Cincinnati, OH (chemicals)	1.5	Richland, WA	2.4
317	Chicago, IL (metal products)	1.5	Steubenville, OH	2.4

Source: Black and Henderson data set.

Duranton and Puga (2000)

Fact 2: *Larger cities are more diversified than small cities.*

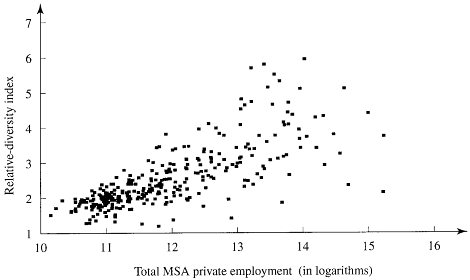


Figure 1. The size–diversity relation for US cities in 1992. *Source:* Black and Henderson (1998) data set.

Duranton and Puga (2000)

- All cities have a large share of employment is non-traded sectors, e.g., services, so all cities are pretty diversified.

- Cities > 500k are more specialized in business services; Finance, Insurance, Real Estate (FIRE), less in manufacturing.
- Cities 50-500k are more specialized in 'mature industries' like 'textiles', less in 'new industries', like 'instruments'.
- Big cities are relatively specialized in 'new industries'.

All together, there is evidence for cities as nurseries of new industrial processes. New processes start in large diverse congested places and once the process is established, migrate to smaller, less congested and less diverse cities.

Fact 3: *The distribution of city sizes is stable over time (more soon), the ranking of cities by size is less stable, and the sectoral specialization is still less stable.*

TABLE 1—RANKINGS AND CHANGES FOR THE TEN LARGEST US METROPOLITAN AREAS, 1977 AND 1997

Rank in 1977 (and change in rank between 1977 and 1997)	Total population	Apparel	Transportation equipment	Instruments
New York	1 (0)	1 (+1)	3 (+20)	2 (+2)
Los Angeles	2 (0)	2 (−1)	2 (0)	5 (−3)
Chicago	3 (0)	11 (−3)	8 (+4)	4 (+5)
Washington	4 (0)	13 (+7)	29 (+8)	16 (−6)
Philadelphia	5 (+1)	3 (+3)	13 (−3)	7 (+4)
Boston	6 (+1)	4 (0)	17 (+24)	3 (−2)
Detroit	7 (+1)	10 (0)	1 (0)	20 (−1)
San Francisco	8 (−3)	15 (−12)	9 (0)	6 (−3)
Cleveland	9 (+5)	18 (+26)	6 (+15)	12 (+6)
Dallas	10 (−1)	5 (+4)	7 (−2)	13 (−5)
Total rank variation	(12)	(57)	(76)	(37)

Notes: The first number in each column is the 1977 ranking (among 272 metropolitan areas). It is followed (in parentheses) by the change in ranking between 1977 and 1997 (where an increase is a decline). The final row sums the absolute value of all the changes above. All calculations use the year 2000 definition of US metropolitan areas. Columns 3 to 5 give the ranking (and the 1977–1997 change) for total employment in SIC23 (apparel and other textile products), SIC37 (transportation equipment), and SIC38 (instruments and related products).

Source: US Census, County Business Patterns, and author's calculations.

Duranton (2007). When thinking about sectoral stability, it matters how aggregated are the sectors. Employment in 'furniture making' is less stable that employment in 'manufacturing'.

Fact 4: *City employment and population are related to specialization and diversity.*

Fact 5abc: *Plants/firms have short half-lives, about 5 years. Most innovations occur in big diversified cities. Most firm relocations are from big diversified cities to smaller less diversified cities.*

A subset of the literature on agglomeration investigates the relationship between diversity, specialization and productivity. Short answer: knowledge intensive activities tend to better in diverse cities, established processes do better in specialized cities (i.e., factory towns).

Zipf's Law/Rank-Size Rule, Gibrat's law

Zipf's law is an extraordinary feature of the size distribution of cities.

Formally, Zipf's law is that the size distribution of cities follows a Pareto distribution with exponent equal to 1.

Less obscurely, it implies a rank size rule (in expectation). If city size is N and the rank of the city in the set of cities under consideration (usually a country) is $r(N)$, then

$$\ln r(N) = \ln A - \zeta \ln N$$
$$\zeta = 1$$

Reorganizing and using $\zeta = 1$, we have

$$\begin{aligned} \ln r(N) &= \ln (A/N) \\ \implies r &= A/N \\ \implies N &= A/r \end{aligned}$$

Consider the first and second and third ranked cities, N_1, N_2, N_3. Then

$$\frac{N_2}{N_1} = \frac{A/2}{A/1} = 1/2 \text{ and } \frac{N_3}{N_1} = \frac{A/3}{A/1} = 1/3.$$

That is, the largest city is twice as large as the second, three times as large as the third, and so on.

This is an odd property, and not one that we would expect to be true. We can test it with the regression

$$\ln r(N) = \ln A - \zeta \ln N + \varepsilon$$

and see if we get $\zeta = 1$.

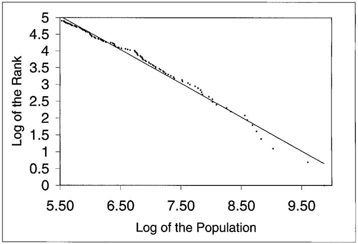
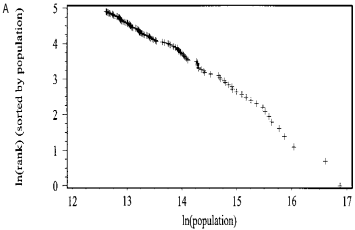


FIGURE I
Log Size versus Log Rank of the 135 largest U. S. Metropolitan Areas in 1991
Source: Statistical Abstract of the United States [1993].

(1) $\ln \text{Rank} = 10.53 - \underset{(.010)}{1.005} \ln \text{Size},$

Gabaix (1999), 135 US MSAs in 1990



Holmes and Lee (2010), 135 US MSAs in 2000

- This looks pretty good for Zipf's law.
- These figures are showing the same data, 10 years apart, but the x-axes are a lot different. Can you explain this? Hint $ln(1000) \approx 6.9$.

This figure is remarkable, and it must be telling us something important about how cities grow.

Since urban growth is pretty clearly central to the process of economic growth and development, this means that Zipf's law is telling us something important about the process of economic growth and development.

Another way to say this, we can reject any model of the development process that does not give us a city size distribution that satisfies Zipf's law. Since Zipf's law is such a specific prediction, this lets us eliminate a lot of models.

This has led to two questions.

- Does Zipf's law really hold?
- Why might we expect Zipf's law to hold?

Why should Zipf's law hold? I

In a remarkable paper, Gabaix (1999) shows that Zipf's law is 'like' the central limit theorem.

The Central Limit Theorem is one of the most important results in statistics. It says that if you take many draws from an arbitrary distribution and average them, and repeat this many times, then the distribution of the *averages* will be Normal, that is, a 'bell curve'.

Here is an animation demonstrating this, `https://www.youtube.com/watch?v=XAuMfxWg6eI`.

Gabaix's argument rests on two assumptions and a theorem.

The assumptions are that,

- Cities grow by a random multiplicative share each year. This share can be drawn from more-or-less any distribution.
- The size of cities has a strictly positive lower bound (with multiplicative growth, once a city hits zero, it stays there. This assumption stops that.)

To state the theorem, let N be an arbitrary size threshold, and N' be the size of a city drawn at random from the sample of cities, then

$$Pr(N' > N) = A/N^{\zeta}.$$

This is called a Pareto distribution. The theorem Gabaix appeals to is that such a Pareto distribution implies the rank size rule (in expectation).

Recalling the Central limit theorem, what should happen if the growth rate of cities is random in each year?

Over time, the distribution of city sizes should approach a Normal distribution (really log Normal because the shocks are multiplicative).

What Gabaix shows is that with the lower bound on city sizes, the city size distribution converges to a Pareto distribution with parameter one, which implies the rank size rule the rank size rule in expectation.

This means that if city growth rates are unrelated to city size, a claim known as 'Gibrat's law', then Zipf's law is implied.

Not that the opposite implication probably does not hold. That is, Zipf's law probably does not imply Gibrat's law.

Does Zipf’s law really hold? I

There has been a lot of research checking whether Zipf’s law holds. This research has taken three basic approaches.

- More careful analysis of the 135 largest US cities shown earlier
- Expansion of the set of cities.
- Extension to other countries.

As Gabaix (1999) points out, even in the set of 135 cities, the data deviates slightly from a perfectly straight line.

- There is often an outlier, often a capital.
- Also, the middle size cities often decrease in size 'too fast'. This is visible in our figure as the plot being a little flatter than 1 as it approaches zero. Looking at the figure very closely, the plot of points looks slightly concave.

Holmes and Lee (2010) investigate whether Zipf's law holds more broadly in the US.

The city size distribution depends in large part on how the boundaries of MSAs are drawn. Thus, the dramatic figures we have seen above may reflect the rules used for drawing MSA boundaries rather than important economic fundamentals.

To address this issue, Holmes and Lee (2010) check whether the rank size rule applies to 85,287 6×6 mile squares drawn on a regular grid covering the continental US in 2000.

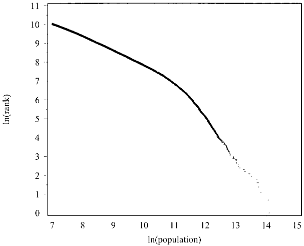


Fig. 3.5 Square-level Zipf plot for continental United States (all 23,974 squares with population at least 1,000)

Table 3.5 Six-by-six-square-level Zipf regression results (squares with population 1,000 and above)

		Piecewise linear				Linear	
Sample of squares	N	Kink	Slope1	Slope2	R^2	Slope	R^2
All squares with population ≥ 1,000	23,974	10.89	.747	1.937	.998	.833	.969
By Census division							
New England	1,027	9.96	.569	1.521	.996	.763	.930
Middle Atlantic	2,184	10.28	.669	1.249	.997	.759	.965
East North Central	4,313	10.92	.784	1.982	.999	.861	.975
West North Central	2,337	11.04	.886	2.607	.999	.941	.984
South Atlantic	4,977	10.72	.756	2.175	.995	.857	.959
East South Central	2,898	10.48	1.010	2.357	.997	1.072	.983
West South Central	3,078	11.17	.786	2.834	.997	.857	.969
Mountain	1,383	11.55	.723	3.662	.997	.791	.964
Pacific	1,777	11.21	.521	1.872	.992	.646	.922

Here, the 'piecewise linear regression allows the slope of the regression line to change at log population 11.

This does not look as good for Zipf's law.

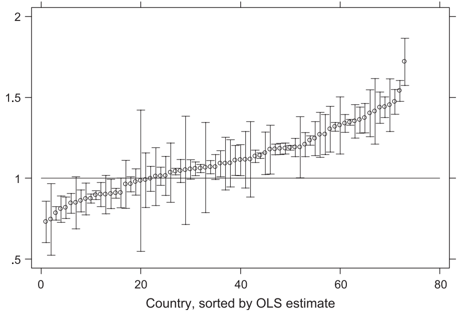
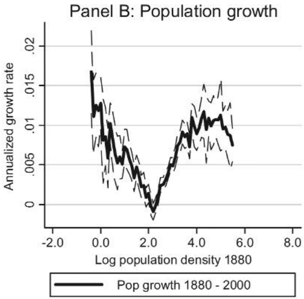


Fig. 1. Values of the OLS estimate of the Pareto exponent with the 95% confidence interval, for the full sample of 73 countries for the latest available period, sorted according to the Pareto exponent.

Soo (2005) estimates Zipf's law for a sample of 73 countries.

y-axis in figure is Zipf coefficient and confidence interval. There is clearly some noise around a slope of one in the rank-size rule.

Since Zipf's law is implied by a model of random growth and scale invariant growth rates, it is interesting to check if city growth rates are invariant to city size.



Michaels et al. (2012) do exactly this using 120 years of US data. This does not look like scale invariant growth rates, though in the upper range of densities, i.e. in the big cities, it looks OK.

Summing up,

- The Zipf's law figures are very dramatic, and surely tell us something about the process governing city growth.
- There has been a lot of research on this.
- It's pretty clear that Zipf's law applies pretty well to the upper tail of city size distributions, and not very well otherwise.
- Gabaix's result that Zipf's law is a consequence of a random growth process is a big step forward, but...
- Gibrat's law is more problematic than Zipf's law empirically, and probably holds only in the upper tail of the size distribution, if at all.
- The 'random growth model' is not very satisfying as an explanation for how the world works.

Probably the most definitive statement we can make is that Zipf’s law describes the upper tail of the city size distribution, and that in this region, Gibrat’s law looks OK, too, and so we should take seriously the model of random growth for large cities.

Costs and Benefits of Cities

We now have several facts about cities. One relating to the overall distribution, Zipf's law, a number describing patterns of diversity and specialization, and a few describing changes in sizes and sectoral composition.

We'd like to develop a theory that let's us organize all of this. The first step is a model of 'systems of cities', which lets us think about how a population chooses cities. One of the first formal statements of this problem is due to Henderson (1974), and this discussion loosely follows this paper.

To start, we need a simple description of how the costs and benefits of cities vary with their size.

For this purpose, ‘benefits’ are ‘output’ and are subject to increasing returns to scale in production, ‘costs’ are only commuting costs and the opportunity cost of labor.

Use the same notation as for agglomeration effects; Production, Y; labor/population, N. Also following the earlier analysis, let

$$Y_i = AN^{1+\sigma}$$

be total output, and

$$w_i = (AN^{\sigma})$$

be the market wage.

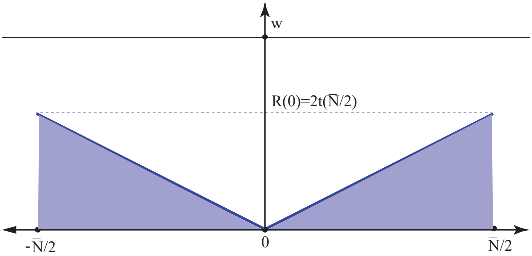
Recall, agents ignore their effect on aggregate output and firms are small, so a competitive labor market leads to w_i being the average product of labor, not its marginal product.

To describe commuting costs, we recall the monocentric city model and make the following assumptions:

- No housing, everyone consumes one unit of land.
- Commute costs are $2t$ per unit distance.
- Agricultural land rent is zero.

This means that the length of the city is equal to its population, $N = 2\overline{x}$.

We want to calculate total and average commuting costs.



Total commute cost is the shaded area. Calculate this by,

- Integrating,

$$TC(N) = 2\int_0^{N/2} 2tx dx = \frac{t}{2}N^2$$

- or recalling that the area of a triangle is $1/2\times$ width $\times$ height. This gives $TC(N) = 2 \times \frac{1}{2} \times \frac{N}{2} \times 2t(\frac{N}{2}) = \frac{t}{2}N^2$.

It follows that average commuting cost is

$$AC(N) = TC(N)/N = \frac{t}{2}N$$

We want to consider two processes for assigning people to cities.

- Spatial equilibrium: (i) no one wants to move (ii) wages are determined competitively.
- Planner's solution: The planner chooses people's locations.

Start with spatial equilibrium. Suppose

- Reservation consumption is $\overline{c}$ and reservation utility is $\overline{u} = u(\overline{c})$. This is what a rural household gets.
- All urban households have the average commute cost for their city.
- Land rents are divided uniformly between city residents (instead of going to absentee landlords).

The first of these means that there is a large pool of rural locations that are not subject to crowding. The next two relieve us of having to keep track of individual incomes, rents and commute distances. They are simplifying assumptions.

Using the second two assumptions, we calculate household consumption as a function of city size for a city resident. This is the difference between wages and average commute costs.

$$
\begin{aligned}
c_E(N) &= w(N) - \frac{TC(N)}{N} \\
&= AN^{\sigma} - \frac{1}{N}\frac{tN^2}{2} \\
&= AN^{\sigma} - \frac{t}{2}N.
\end{aligned}
$$

That is, average product of labor minus average commute cost. Note that land rent nets out.

In a spatial equilibrium, we are going to need to choose N so that

$$\begin{aligned} c_E(N) &= \overline{c} \qquad (1) \\ \implies AN^{\sigma} - \frac{t}{2}N &= \overline{c} \end{aligned}$$

We'll see below that $c_E(N)$ is concave and this equation can have two solutions. Call the smaller one $\underline{N}_E$ and the larger one N_E^*. We will mainly be interested in N_E^*.

A candidate for 'optimal city size' is going to be the N_E^{**} that maximizes $c_E(N)$. To find this, solve,

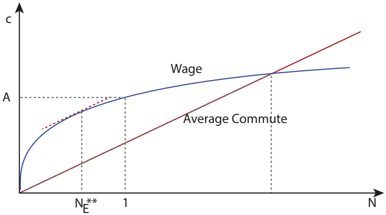
$$\begin{aligned} 0 &= \frac{dc_E(N)}{dN} \\ \implies 0 &= \sigma A N^{\sigma-1} - \frac{t}{2} \\ \implies N_E^{**} &= \left(\frac{t}{2\sigma A}\right)^{1/(\sigma-1)} \quad . \end{aligned}$$

N_E^{**} is the city size that maximizes *equilibrium* land rent (why?).

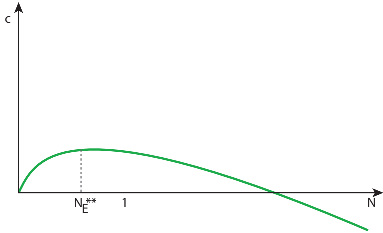
To check that this is a maximum, note that

$$\begin{aligned}
\frac{d^2 c^*(N)}{dN^2} &= \frac{d}{dN}\left(\frac{dc_E(N)}{dN} \right) \\
\frac{d^2 c^*(N)}{dN^2} &= \frac{d}{dN}\left(\sigma A N^{\sigma-1} - \frac{t}{2} \right) \\
&= (\sigma - 1)\sigma A N^{\sigma-2} \\
&< 0.
\end{aligned}$$

Where the inequality follows because the data indicates that σ is in the neighborhood of 0.05 and much smaller than 1.

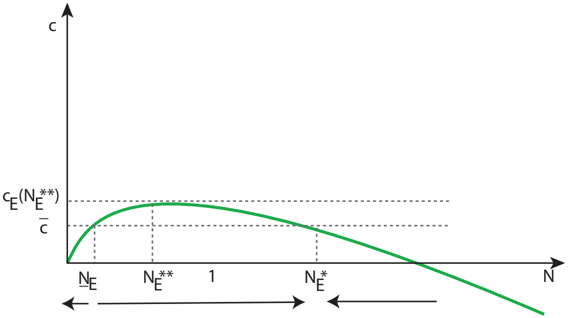


Components of $c_E(N)$



$c_E(N)$

Figure: Illustration of spatial equilibrium for a single city



What happens in equilibrium?

- For any given city, there are two levels of population where urban consumption equals rural consumption, $\underline{N}_E$ and N_E^*.
- $\underline{N}_E$ is 'unstable'. For any small displacement $\varepsilon > 0$,
  - $c_E(\underline{N}_E - \varepsilon) < \overline{c}$. If the city experiences such a displacement, everyone will want to emmigrate and the city will depopulate.
  - $c_E(\underline{N}_E + \varepsilon) > \overline{c}$. If the city experiences such a displacement, everyone will want to immigrate to the city until $N = N_E^*$.

  We rule out the unstable equilibrium. By a similar argument, N_E^* is stable.
- Except in the special case where $\overline{c} = c_E(N_E^{**})$, this is going to lead us to equilibrium city sizes that are 'too large' in the sense that they are bigger than N_E^{**}.

This argument requires a number of comments.

- The reliance on ‘stability’ to resolve problems with multiple equilibria is common and problematic. Implicitly, notions of stability are dynamic and our agents are not forward looking. Justifying this requires logical gymnastics.
- In this context, it is common to describe commute costs as ‘congestion’. As the city grows, access to the center becomes congested and it requires longer average commutes for the city to function.
- Recall that in the competitive labor market, agents ignore their impact on aggregate productivity. This should lead to cities that are ‘too small’. It is striking that this is not what happens. People also ignore their effect on other people’s commute costs, and this effect leads to cities that are too big. The second effect dominates.

- This model gives us way to think about the size and number of cities, but the model is plagued by multiple equilibrium. That is, it is missing some mechanism to winnow the set of possible spatial equilibria.
  - If there are initially zero cities, what does this model suggest will happen? Nothing. The first individual who migrates to a city receives zero wage, so no individual can start a city. We need some form of collective action to get cities.
  - If the pool of rural people is arbitrarily large, then any number of cities can exist in equilibrium.
  - If the pool of rural people is finite, then any number of cities can exist such that all rural people urbanize, the size of all cities is the same, and this size is in the interval $[N_E^{**}, N_E^{*}]$.
- So far, this is really a model of one city and one sector, so it is not useful for explaining the stylized facts presented earlier. More on this later.

Now consider the planner's problem. The planner would like to maximize the 'surplus' created by the city. This is the value of output minus the cost of commuting and the opportunity cost of labor. In math,

$$\begin{aligned} W(N) &= Y(N) - TC(N) - \overline{c}N \\ &= AN^{1+\sigma} - \frac{tN^2}{2} - \overline{c}N \end{aligned}$$

To optimize, the planner wants to choose N_P^* to satisfy

$$
\begin{aligned}
0 &= \frac{dW(N)}{dN} \\
&= \frac{d}{dN}(AN^{1+\sigma} - \frac{tN^2}{2} - \overline{c}N) \\
&= (1+\sigma)AN^{\sigma} - tN - \overline{c} \\
\implies \overline{c} &= (1+\sigma)AN_P^{\sigma} - tN \equiv c_P(N)
\end{aligned}
$$

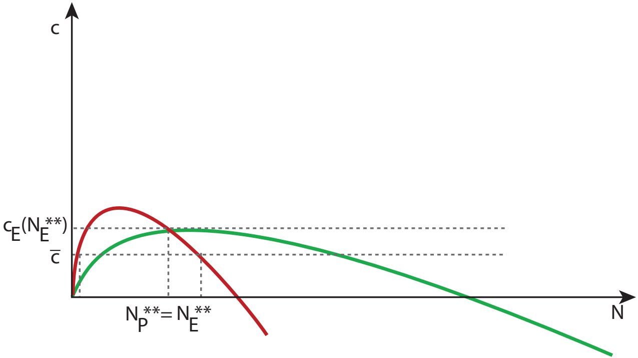
By inspection, $c_P(N)$ is positive for small N and concave. Comparing it to $c_E(N)$ it is steeper at zero.

Is the equilibrium city ever the same size as the optimal city? A natural guess would be 'no'. In the equilibrium city, households choose their location without thinking about how their decision affects the productivity or commute of other residents, while the optimal decision requires making these marginal trade-offs.

In fact, this intution is not quite right, to see this, check if we can find any values of N where the levels of consumption in the optimum and the equilibrium coincide. That is, solve the following for N,

$$
\begin{aligned}
c_E(N) &= c_P(N) \\
\implies AN^{\sigma} - \frac{t}{2}N &= (1+\sigma)AN^{\sigma} - tN \\
\implies AN^{\sigma-1} - \frac{t}{2} &= (1+\sigma)AN^{\sigma-1} - t \\
\implies N &= \left(\frac{t}{2\sigma A}\right)^{1/(\sigma-1)} = N_E^{**}
\end{aligned}
$$

What is going one here? The optimum and the equilibrium agree at a single point, when the size of the equilibrium city is N_E^{**}. This is much easier with a picture ...



- There are generally two values of N that solve the planner's problem. The smaller will be 'unstable' and so we ignore it.
- In the special case where $\overline{c} = c_E(N_E^{**})$ equilibrium and optimum coincide.

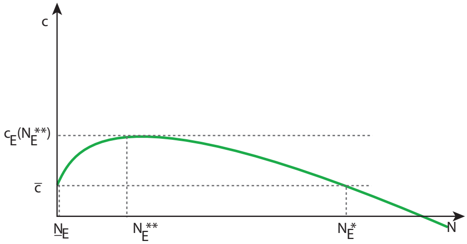
- Otherwise, equilibrium cities are always larger than the optimum, but the gap between them decreases as the rural area is more attractive.

Extension: Allowing cities of size zero

Suppose that $c_E(N)$ is shifted up so that $c_E(0) = \overline{c}$. This seems pretty sensible. A single person city gets the same payoff as the countryside, not less, as we have implicitly assumed up to now.

In this case, the equilibrium picture looks like this,

Figure: Planner's problem for a single city



- This seems a little more realistic.
- It suggests dynamics for growth of a system of cities. Once a city reaches size N_E^*, the next person starts a new city. Once this happens, the existing city spilts and people divide themselves between the two cites. This requires that more than half the city population live to the right of N_E^{**}, as drawn.
- This process repeats when both cities are full.
- Earlier caveat about dynamics applies here, too.

Extension: Real estate developers

- It is common to think of cities as being created by real estate developers. These agents are entrepreneurs who start cities, collect all land rent, and can exclude residents. They will solve the planner's problem and lead to cities of the optimal size.
- This actually happens, sometimes. For example, Irvine Ca., with population almost 300k, was planned by the Irvine company.
- City councils sometimes seem to act in much the same way. They do lots of things to stop people from moving into their jurisdiction, mainly by making it hard to build things. If this model is right, should we discourage this sort of activity? Recall, there can be big wage differences across cities.

Conclusion I

- We know a lot about patterns of diversity and specialization that occur in any population of cities. These things are pretty easy to observe. Similarly, about patterns of change.
- We also observe striking and almost conclusive evidence for the rank size rule. It holds pretty well for large cities, not very well for small cities.
- We have a number of theories to explain the rank size rule. In particular, Gabaix (1999) proposes a theory of random growth.

Conclusion II

- The basic intuition for the systems of cities problem was first laid out in Henderson (1974). This paper remains relevant 50 years after it was published. It provides a framework for thinking about the problem, but there are multiple equilibria and no very satisfactory way of choosing between them. It also does not provide much basis for thinking about patterns of sectoral specialization.
- There has been important work on the second set of problems since, Rossi-Hansberg and Wright (2007) and Duranton and Puga (2001) are noteworthy, but the problem of multiple equilibria is still unresolved. That is, we still lack a complete model of systems of cities.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic Review*, 97(1):197–221.

Duranton, G. and Puga, D. (2000). Diversity and specialisation in cities: why, where and when does it matter? *Urban studies*, 37(3):533–555.

Duranton, G. and Puga, D. (2001). Nursery cities: urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5):1454–1477.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly journal of economics*, 114(3):739–767.

Henderson, J. V. (1974). The sizes and types of cities. *The American Economic Review*, pages 640–656.

Holmes, T. J. and Lee, S. (2010). 3. cities as six-by-six-mile squares: Zipf's law? In *Agglomeration economics*, pages 105–132. University of Chicago Press.

Michaels, G., Rauch, F., and Redding, S. J. (2012). Urbanization and structural transformation. *The Quarterly Journal of Economics*, 127(2):535–586.

Rossi-Hansberg, E. and Wright, M. L. (2007). Urban structure and growth. *The Review of Economic Studies*, 74(2):597–624.

Soo, K. T. (2005). Zipf's law for cities: a cross-country investigation. *Regional science and urban Economics*, 35(3):239–263.