A physicist's approach to city modelling

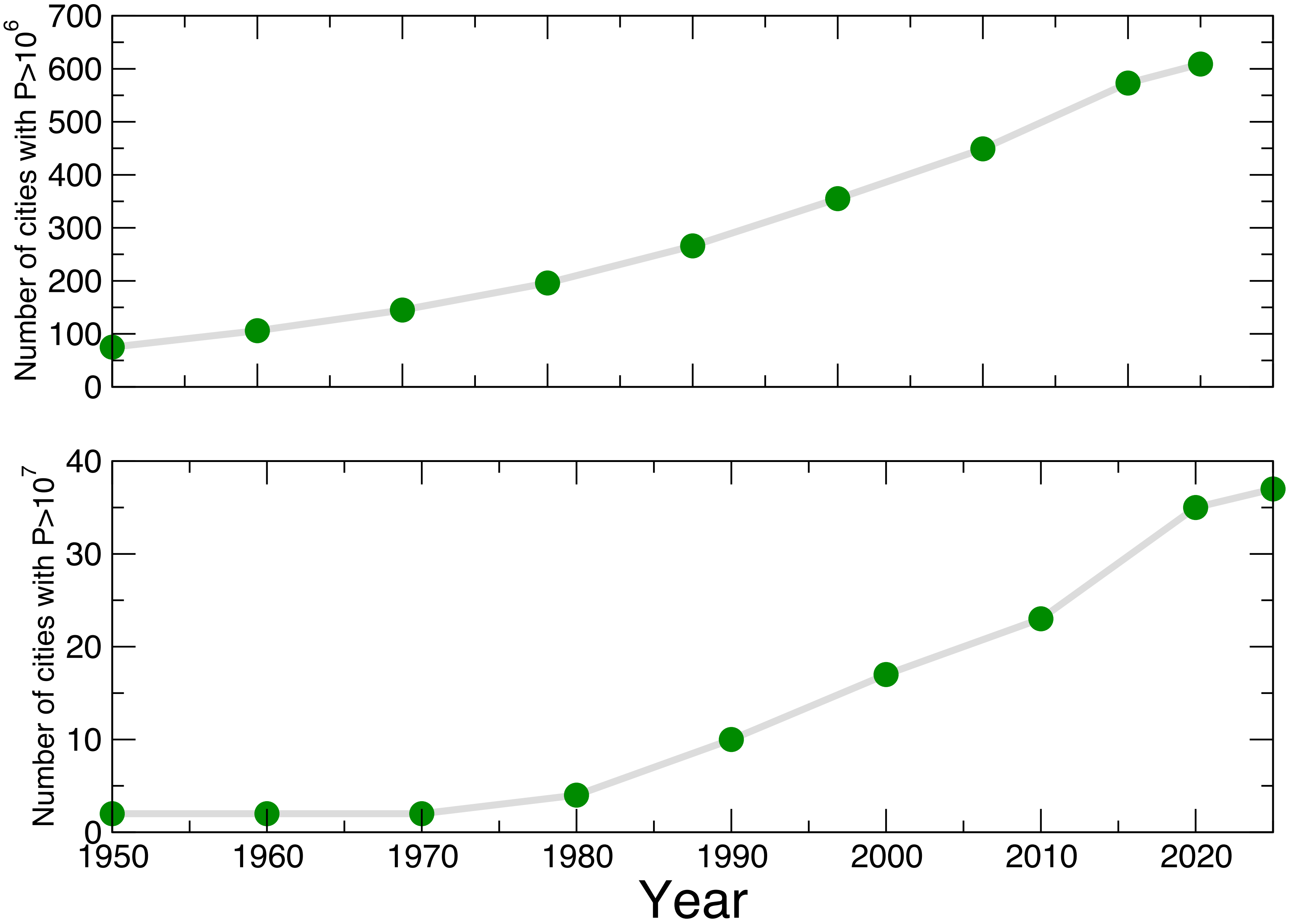
Marc Barthelemy

CEA, Institut de Physique Théorique, Saclay, France

EHESS, Centre d' Analyse et de Mathématiques sociales, Paris, France

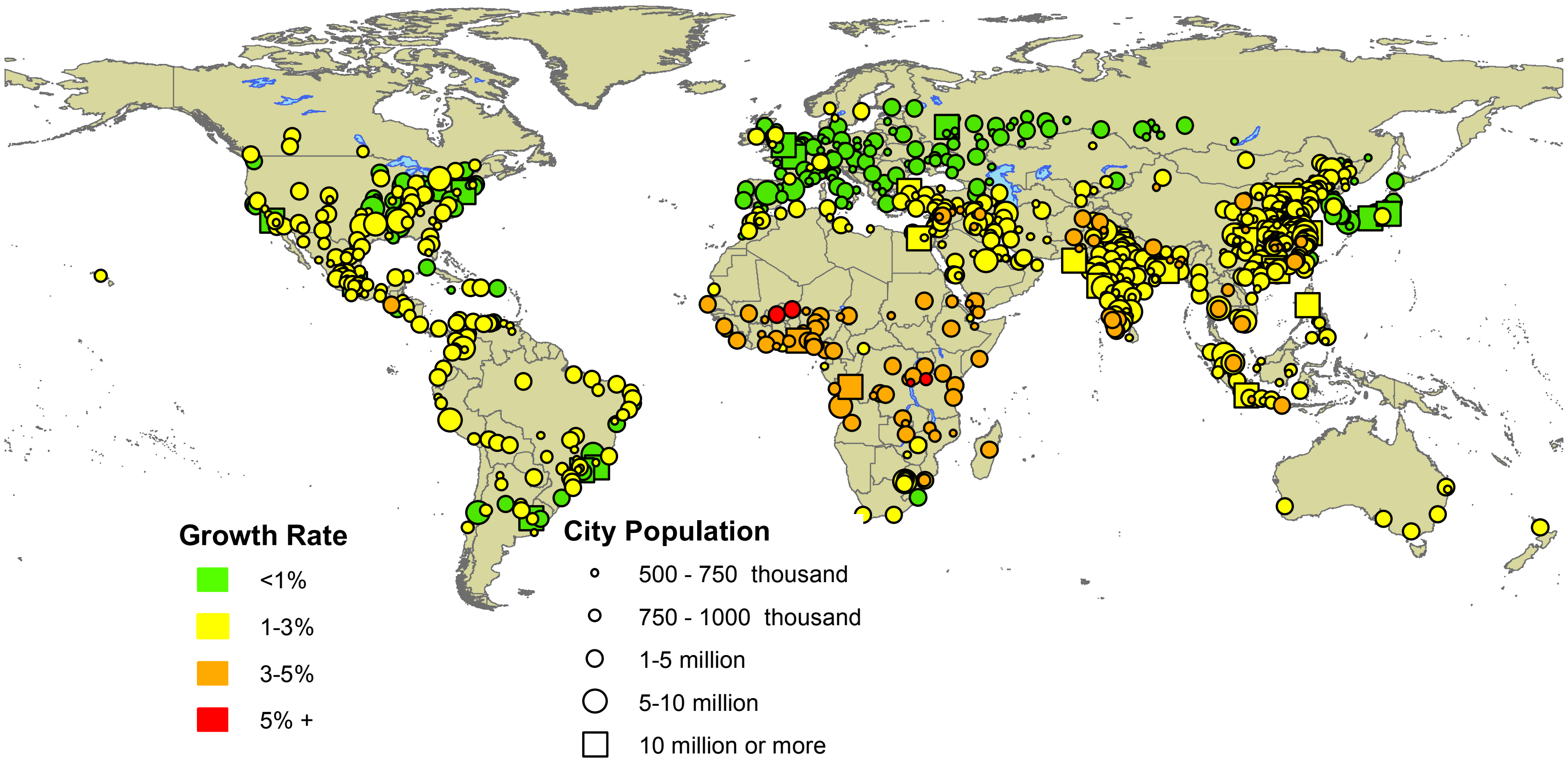
marc.barthelemy@ipht.fr
www.quanturb.com

Importance of cities: megacities



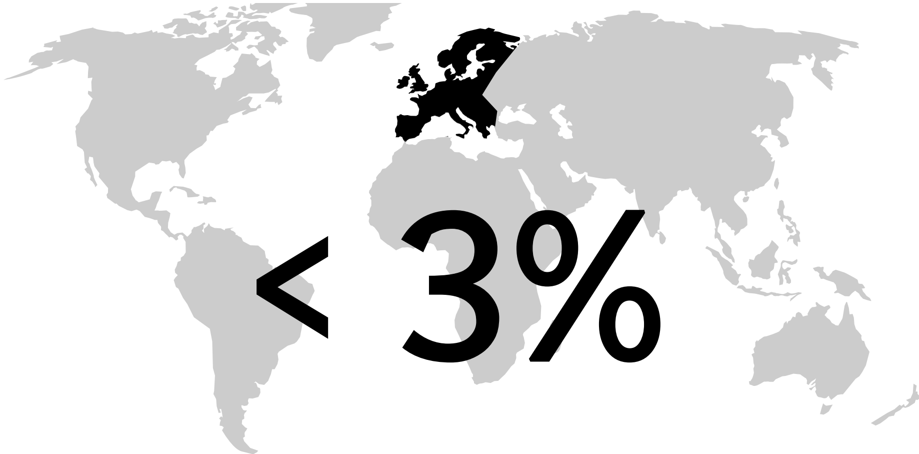
Data from www.geohive.com (UN data)

Heterogeneous distribution of growth rates



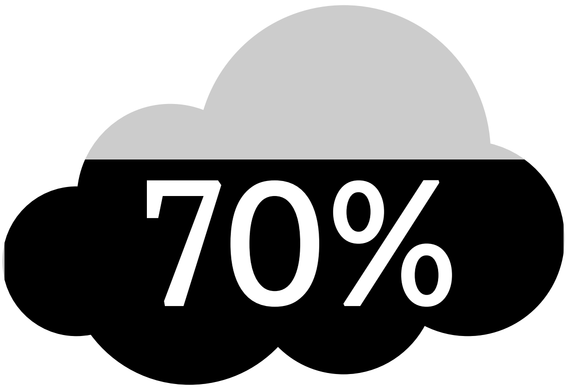
Cities are about concentration

Urbanized area



NASA 2001

CO2 emissions



ONU-HABITAT 2011

GDP

80%

ONU 2011

Urbanism: a lot of "theory" .. But many practical problems !

- Social and economical problems (spatial segregation, crime, accessibility, etc.)

- Mobility: congestion, pollution, …

- Sustainability of urban structures ?

- We need robust models, and a better understanding of the "physics" of cities
- Build a "science of cities" validated by data

Science and cities: state of the art

Number of parameters

Urban economics (and physics):
Very abstract models, empirical tests ?

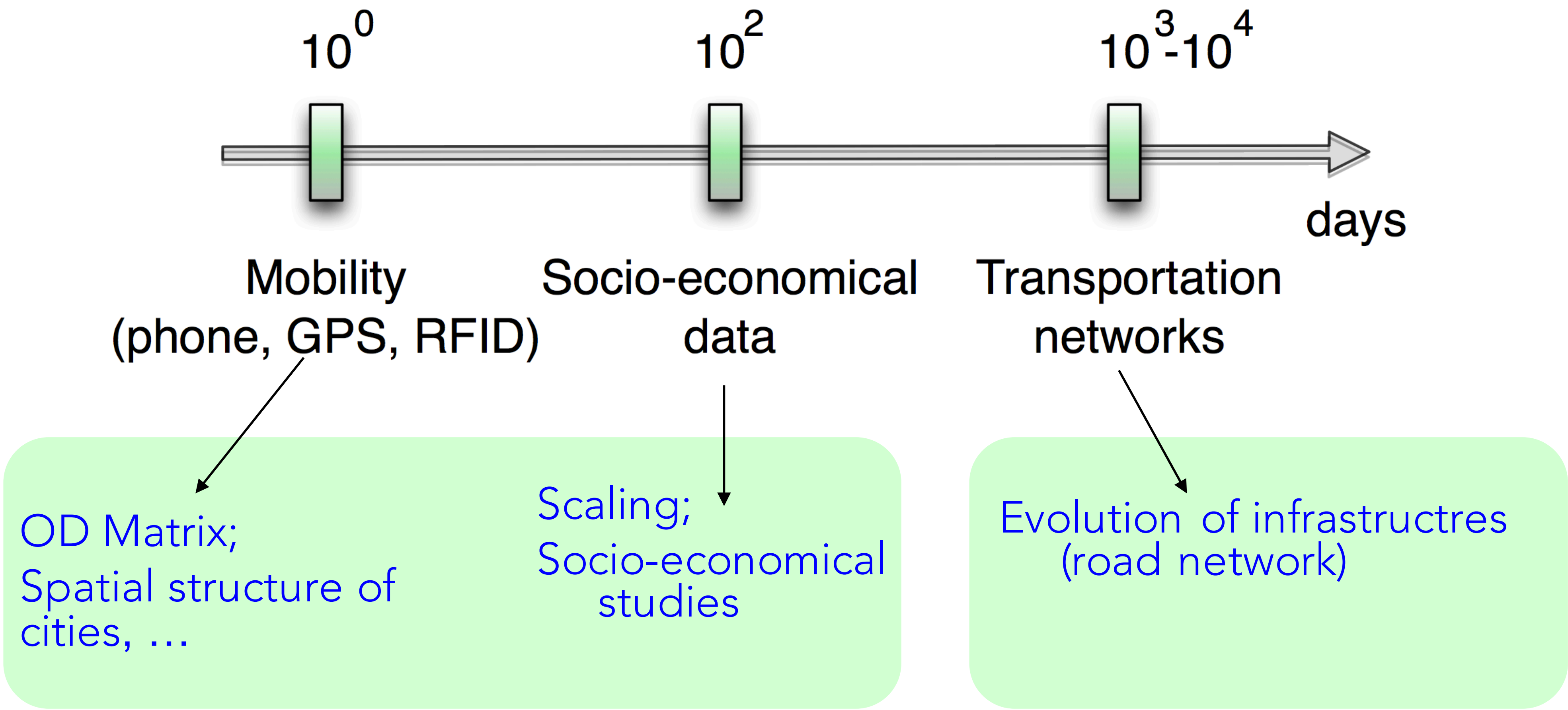
Simulations, agent-based models (LUTI models): Validity ? Strong perturbation ?
Machine learning ?

Minimal model: the smallest number of parameters and the largerst number of verified predictions

- Loop theory-observation necessary
- "Machine learning": black box, output difficult to interpret…

A new science of cities

- Game changer ? Urban data !
- Different scales (and different processes)

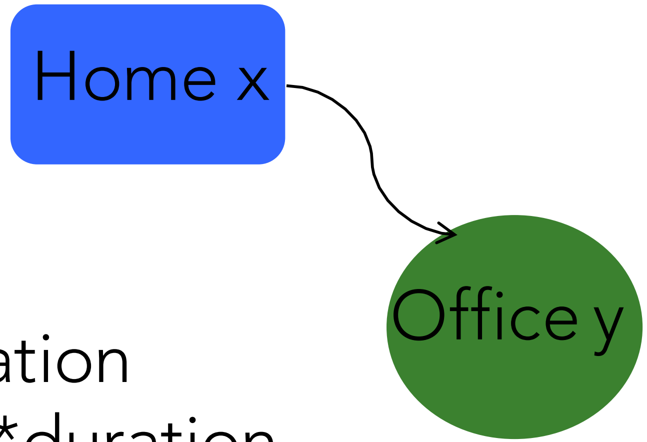


Revisiting Spatial economics: the Fujita-Ogawa model (1982)

- A model for the spatial structure of cities: an agent will choose to live in *x* and work in *y* such that

$$Z_0(x, y) = W(y) - C_R(x) - G_T(x, y)$$

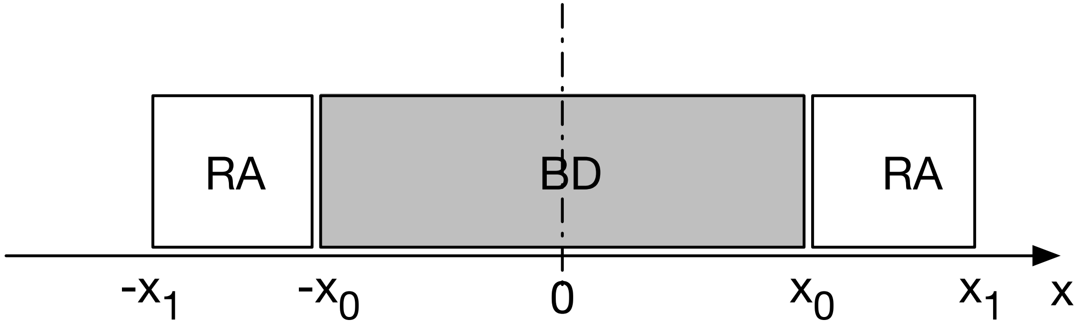
is maximum



- W(y) is the wage at y
- C_R(x) is the rent at x
- G_T(x,y) is the generalized transportation cost from x to y =monetary cost+V*duration

- …and a similar equation for companies (maximum profit)

Spatial economics: Fujita-Ogawa (1982)



- Main result: monocentric configuration stable if

$$\frac{t}{k} \leq \alpha$$

- t: transport cost
- 1/α interaction distance between firms

- Effect of congestion: larger cost t

Spatial economics: Fujita-Ogawa (1982)

- There are many problems with this model:
  - Not dynamical: optimization. We want an out-of-equilibrium model
  - No congestion (!) We want to include congestion (for car traffic). Only one transport mode – we want to include mode choice
  - No empirical test. Extract testable predictions (see the book: Spatial Economics, by Fujita, Krugman, Venables)

Spatial economics: Fujita-Ogawa (1982)

- This model is unable to predict the spatial structure of cities in general
- We will "simplify" the problem and discuss two phenomena:

  - (1) the evolution of car traffic with population
  - (2) the number of activity centers….(if time allows)

I. Modeling car traffic in cities

Car traffic: Newman-Kenworthy 1989

- Many problems:
  - Data availability ?
  - Reproductibility ?
  - Interpretation and use ?
  - Theoretical foundation ?

$$Q_{gas} \approx \frac{1}{\rho^{\sigma}} \quad ?$$

Newman & Kenworthy.
Gasoline consumption and cities: a comparison of US cities with a global survey.
J. Am. Plann. Assoc 1989

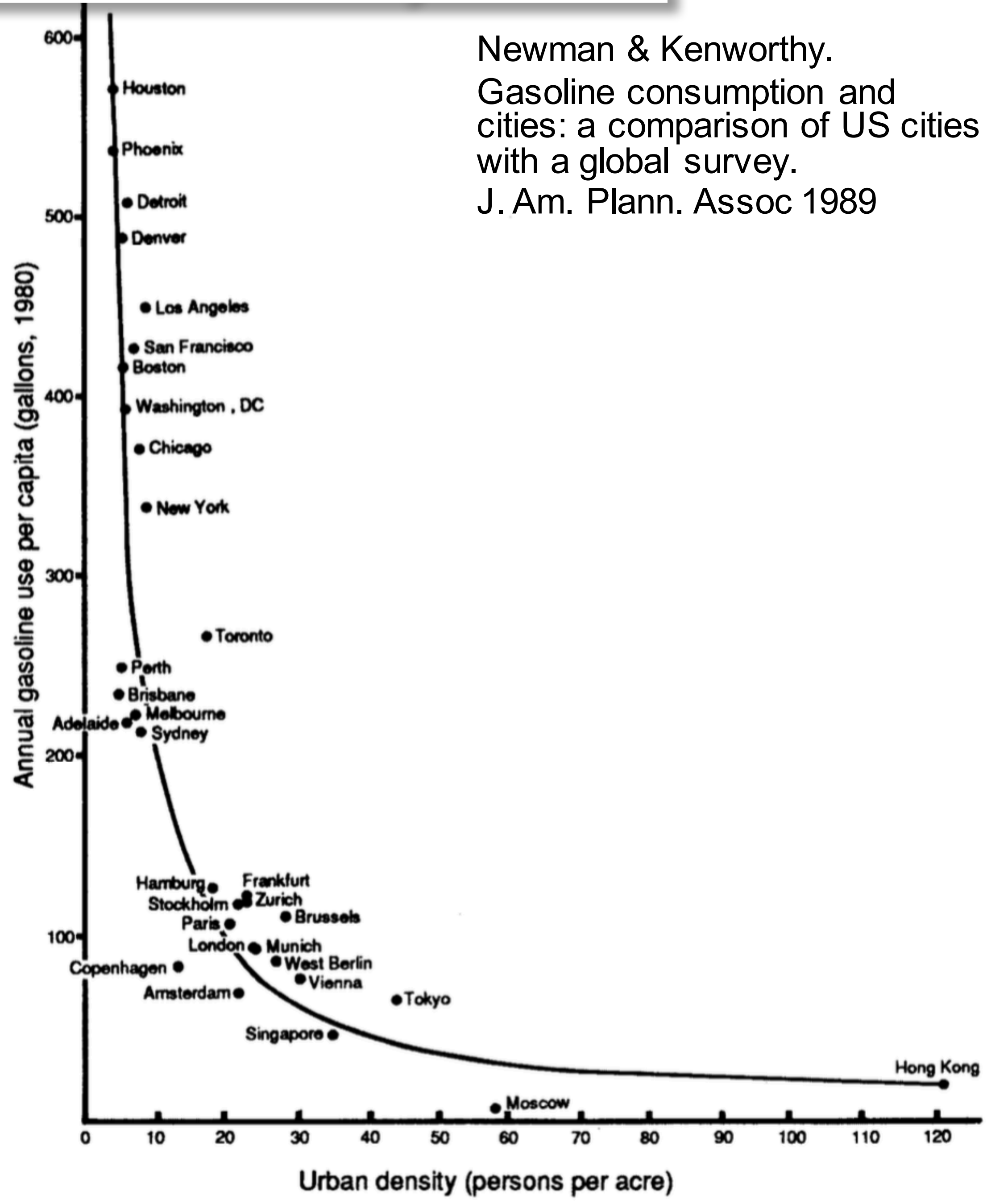


Figure 1. Gasoline use per capita versus population density (1980)

Modeling car traffic in cities

- Theoretical approach with testable predictions ?

- Ingredients:
  - Budget optimization: maximize (Fujita & Ogawa)

$$\max(Z_0 = \cancel{W(y)} - \cancel{C_R(x)} - G_T(x,y))$$

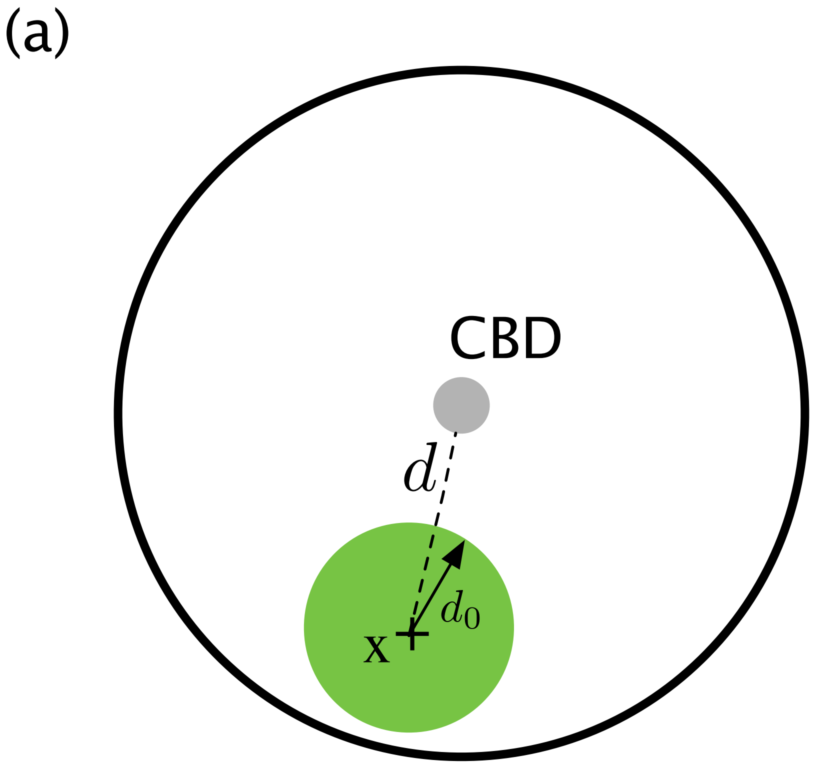
  - Individuals randomly located across the city
  - Monocentric case: same wage at the CBD

$$\max(Z_0) \Rightarrow \min G_T$$

  - Minimum computed over the different modes: mass rapid transit (subway) and private car

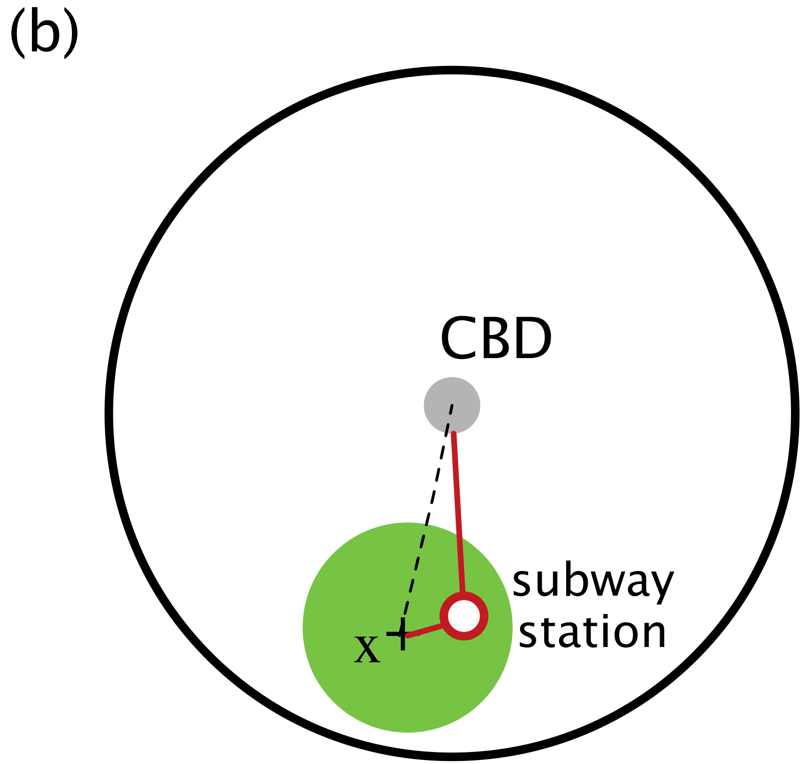
Modeling car traffic in cities

- Probability p to have access (<1km) to a MRT station



Probability 1–p:
no subway station

$$\text{Cost} = G_{car}(x)$$



Probability p:
subway station

$$\text{Cost} = \min(G_{car}(x), G_{MRT}(x))$$

Generalized cost

- Generalized costs: **monetary cost+V*time**

- C_c Monetary cost: price of the car, insurance, etc…

- V value of time=amount of money willing to pay in order to save one hour of time; increasing with income (typically a fraction of the income).

- Large V => time is the most important
- Small V => money is the most important

Generalized cost: car

- The time to go from x to y separated by d(x,y) is (Bureau of Public Roads function)

$$\tau(x,y) = \frac{d(x,y)}{\overline{v}}\left[1+\left(\frac{T(x,y)}{C}\right)^{\mu}\right]$$

where:

$\overline{v}$ is the average free flow velocity

T(x,y) is the traffic

C is the capacity of the road system

μ is an exponent >1 characterizing the sensitivity to congestion

- With congestion time=d/v is not valid anymore !

Generalized cost: car

- The time for a trip of length d, on road of capacity C and free flow velocity v_c, and with traffic T is then

$$\tau(d) = \frac{d}{v_c}\left[1 + \left(\frac{T}{C}\right)^{\mu}\right]$$

- The generalized cost for the car is then

$$G_{car} = C_c + V\frac{d}{v_c}\left(1 + T^{\mu}\right)$$

Generalized cost: MRT (subway)

- We neglect monetary costs (compared to C_c)
- V value of time, for a distance d, trip duration

$$\tau(d) = f + \frac{d}{v_m}$$

- Driving is faster: $v_c > v_m$ (typically 40 vs 30km/h) but more expensive
- f: walking+waiting time

- Generalized cost for the MRT

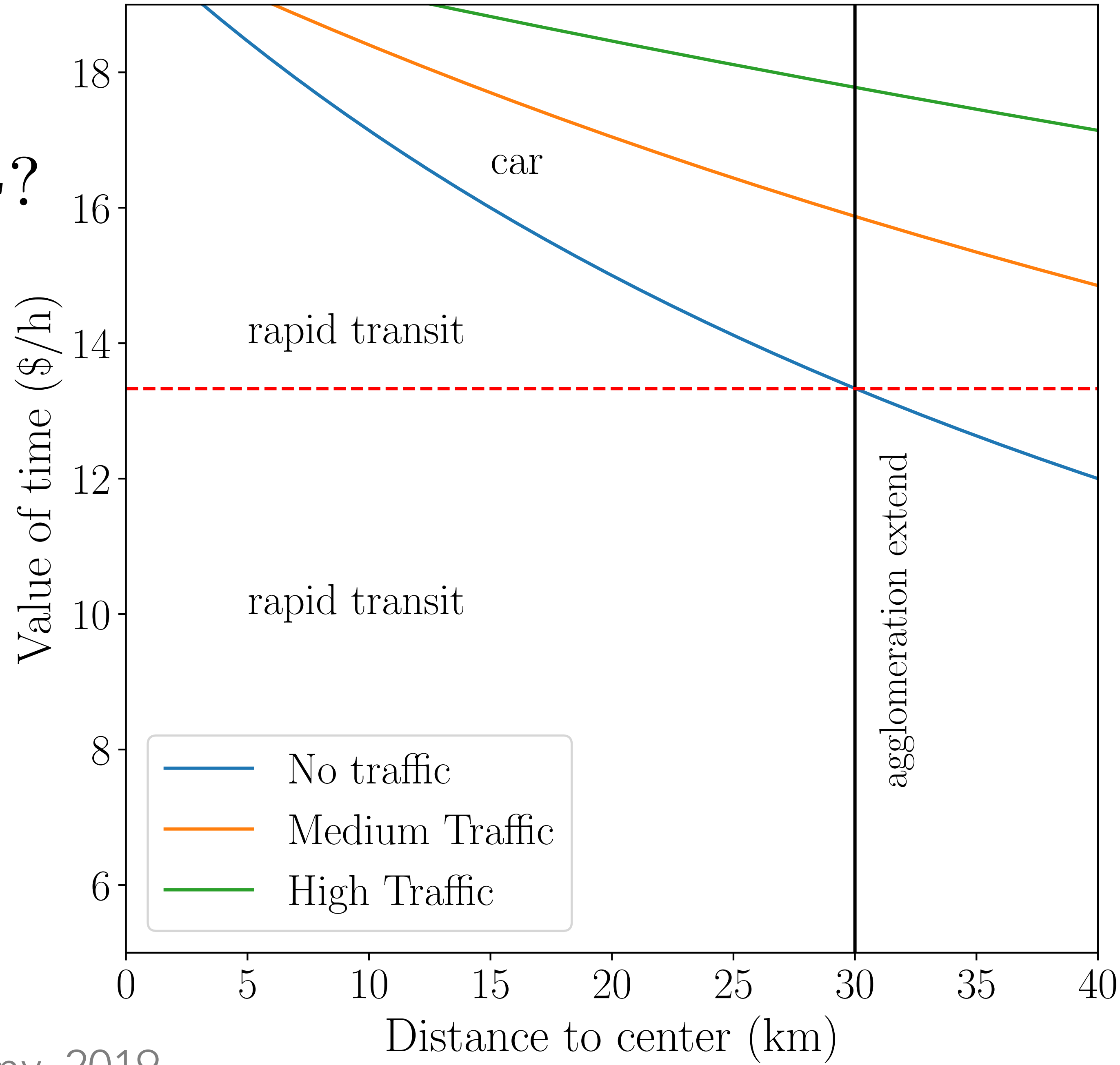
$$G_{MRT} = V\left(f + \frac{d}{v_m}\right)$$

Modeling car traffic in cities

- Mode choice:

$$G_{car} > G_{MRT}?$$

- Choice of mode depends on V



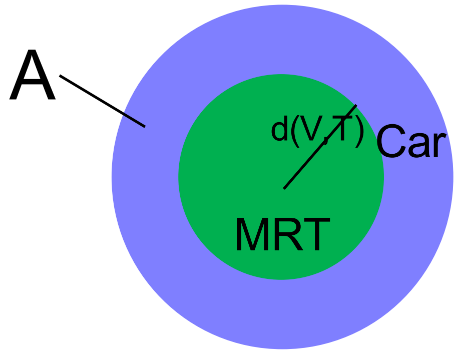
Verbavatz & Barthelemy, 2019

Critical distance and traffic

- Writing $G_{MRT}=G_{car}$ gives a critical distance ($L \sim A^{1/2}$, A area of the city)

$$d(V,T) = \min\left(L, \frac{\frac{C_c}{V} - f}{\frac{1}{v_m} - \frac{1}{v_c}}\left(1 + \left(\frac{T}{C}\right)^{\mu}\right)\right)$$

- If d<d(V,T) => MRT
- If d>d(V,T) => car



- Writing d(V,T*)=L gives the critical maximum traffic T* above which MRT is beneficial in the whole city

$$T^* = CF\left(v_m, v_c, \frac{1}{L}\left(\frac{C_c}{V} - f\right)\right)$$

Evolution equation for the car traffic T

- Population P increases
- Probability p to have access (<1km) to a MRT station

- For T<T*:

$$\frac{\mathrm{d}T}{\mathrm{d}P} = 1 - p + p\left(1 - \frac{\pi d(V,T)^2}{A}\right)$$

No access to MRT

Access to MRT and in the « car regime »

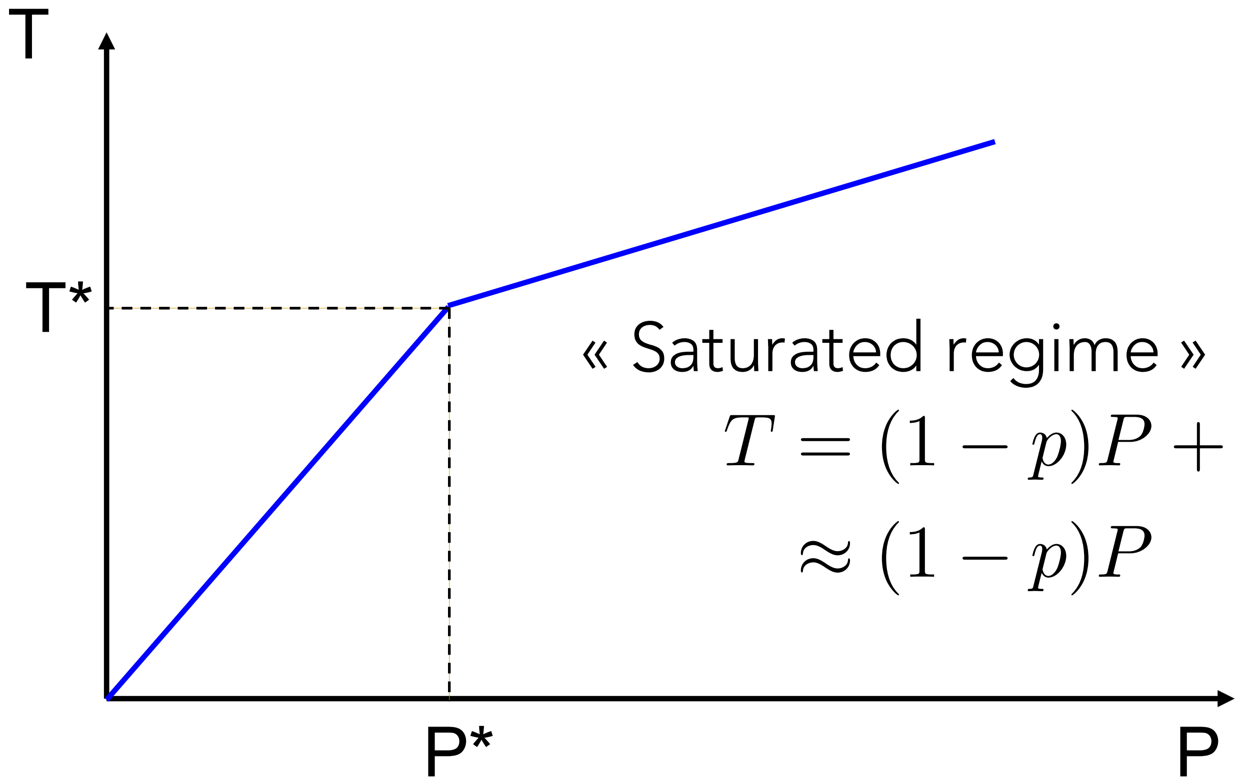
- For T>T* and P>P*:

$$\frac{\mathrm{d}T}{\mathrm{d}P} = 1 - p$$

Verbavatz & Barthelemy, 2019

Modeling car traffic in cities

- Evolution equation for the traffic



« Saturated regime »

$$T = (1-p)P + \text{const}$$
$$\approx (1-p)P$$

Results of the model

- For P>P* the only source of car traffic comes from individuals who do not have access to MRT

$$T \approx (1-p)P$$

where p is the proba to have access to MRT

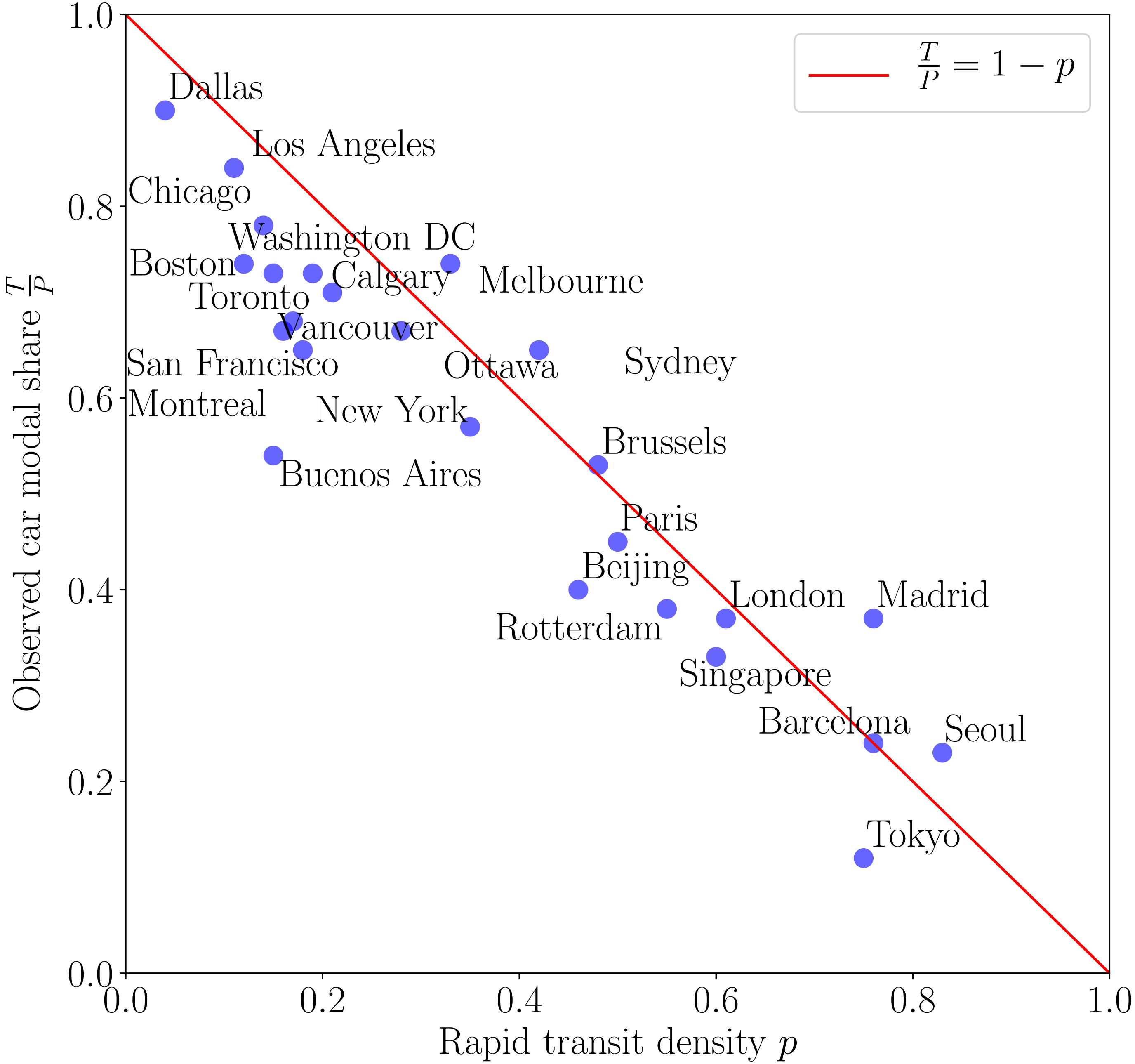
- P* depends on the details of the city and individuals, usually small).

- For most large cities, the traffic is « saturated »: the only car drivers do not have an alternative

- We got the data for 25 cities in the world (bottleneck: p)

Verbavatz & Barthelemy, 2019

Modeling car traffic in cities



Verbavatz & Barthelemy, 2019

CO_2 emissions

- CO_2 emissions proportional to the time spent driving

$$Q_{CO_2} \propto \sum_{drivers\ i} d(x_i) \left[1 + \left(\frac{T}{C}\right)^{\mu}\right]$$
$$\propto g\sqrt{A}(1-p)P[1+\tau]$$

where τ the average delay due to traffic jams (data available:TomTom)

Marseille 40%

- We then obtain

$$\boxed{Q_{CO_2}/P \approx (1-p)\sqrt{A}(1+\tau)}$$

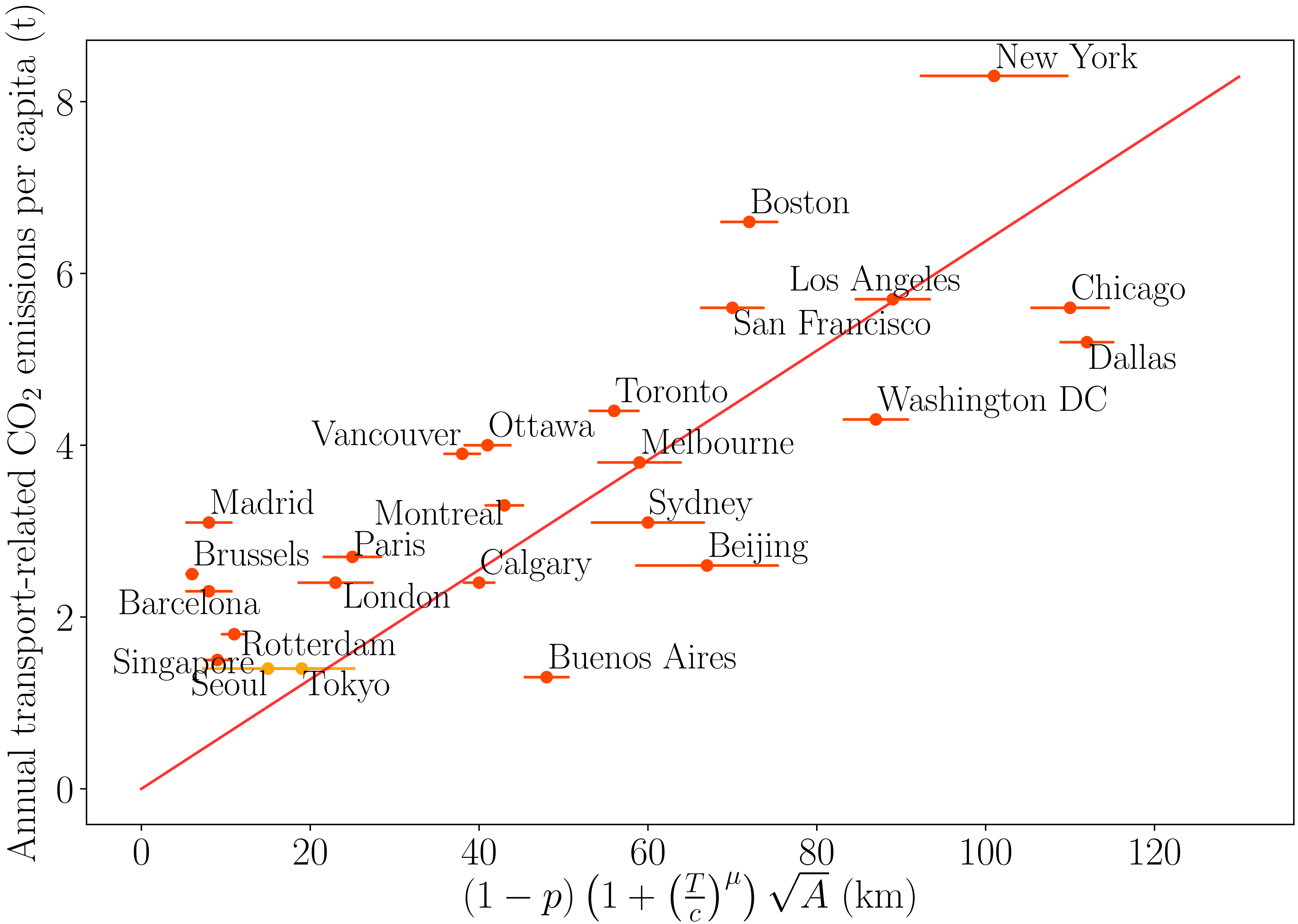
Public transport density

Size of city

Effect of congestion

Verbavatz & Barthelemy, 2019

CO_2 emissions in cities



Verbavatz & Barthelemy, 2019

Modelling car traffic in cities

- We assume $Q_{gas} \propto Q_{CO_2}$
- From $Q_{gas}/P \approx (1-p)\sqrt{A}(1+\tau)$ we obtain

$$Q_{gas}/P \approx \frac{\sqrt{P}}{\sqrt{\rho}} \approx \frac{1}{\sqrt{\rho}}$$

where ρ=P/A is the average urban density

- We « understand » here the result of Newman and Kenworthy

Verbavatz & Barthelemy, 2019

Discussion

- The model predicts that it is not the density that controls gasoline consumption (and CO_2 emissions due to transport) but:
  - the density of public transport
  - car congestion
  - the area size of the city

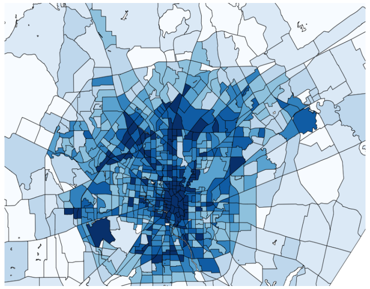
- In general increasing the density in order to decrease CO_2 emissions is.. wrong !
  - If P increases (at A fixed)=>Q_{CO2} increases (congestion)
  - We have to decrease A or more realistically increase p or density at MRT stations

This simple model helped us to point to the relevant parameters…

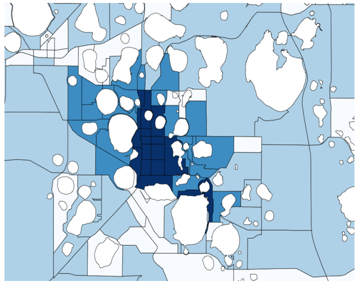
Verbavatz & Barthelemy, 2019

II. Polycentric structure of cities

Polycentric structure: empirical result

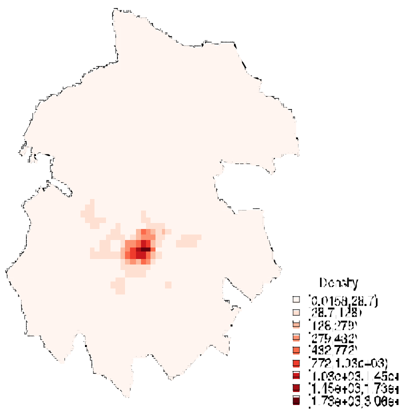


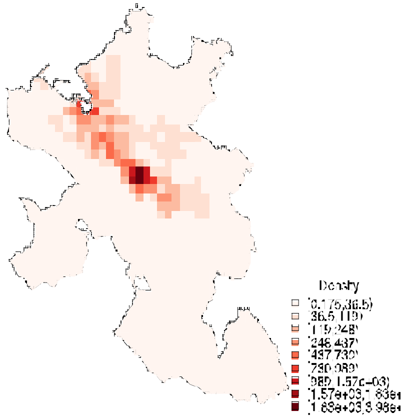
San Antonio (TX), USA



Winter Haven (FL), USA

# of employees per zip code, USA (9,000 cities)





Mobile phone data: density of users (urban areas in spain)

Number H of activity centers

$$H \sim P^{\beta}$$

- Exponent **β**~0.5-0.6

The spatial structure of activity in cities

- We have a polycentric structure, evolving with P
- We can count the number H of centers

$$H \sim P^{\beta} \quad \beta \approx 0.5 - 0.6$$

Computing β ?

- Mobility is the key: we need to model how individuals choose their home and work place
- Problem largely studied in geography, and in spatial economics: Edge City model (Krugman 1996), Fujita-Ogawa model (1982)
- Revisiting Fujita-Ogawa: predicting the value of β

Spatial economics: the edge city model (Krugman 1996)

- The important ingredient is the 'market potential'

$$\Pi(x) = \int K(x-z)\rho(z)\mathrm{d}z$$

- Describes the spillovers due to the density in z
- Specifically

$$K(u) = A(u) - B(u)$$

- The average market potential is

$$\overline{\Pi} = \frac{1}{\Omega}\int \Pi(x)\rho(x)\mathrm{d}x$$

Spatial economics: the edge city model (Krugman 1996)

- The equation for the evolution of business density is

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \gamma \left(\Pi(x) - \overline{\Pi}\right)$$

- Linearize around flat situation $\rho(x) = \rho_0 + \delta\rho(x)$

$$\delta\rho(k) \sim \mathrm{e}^{\gamma K(k)t}$$

- At least one maximum at k=k*; the number of hotspots is then:

$$H \sim \Omega k_*^2$$

- Scaling with the population ?
- Link micromotives-macrobehavior ?

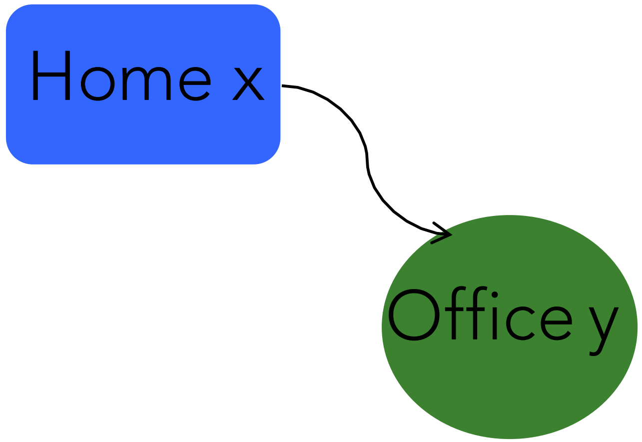
Spatial economics: Fujita-Ogawa (1982)

- A model for the spatial structure of cities: an agent will choose to live in *x* and work in *y* such that

$$Z_0(x,y) = W(y) - \cancel{C_R}(x) - C_T(x,y)$$

is maximum

- $W(y)$ is the wage at y
- $C_R(x)$ is the rent at x
- $C_T(x,y)$ is the transportation cost



- Assumptions and simplifications:
  - Assume that home is uniformly distributed (x): find a job !
  - We have now to discuss W(y) and C_T

Estimating the wage

- The wage results from a large number of interactions: complex quantity !

- In physics (heavy ions) replacing a complex quantity by a random variable is useful and sometimes accurate (Wigner 55) ! (=> theory of random matrices)

- We then choose:

$$W(y) = s\eta(y)$$

where s sets the salary scale and η is a random variable

- Note: the disorder is quenched here

Louf, MB, PRL 2013

Generalized cost: car

- The time to go from x to y separated by d(x,y) is (Bureau of Public Roads function)

$$\tau(x,y) = \frac{d(x,y)}{\overline{v}}\left[1 + \left(\frac{T(x,y)}{C}\right)^{\mu}\right]$$

where:

$\overline{v}$ is the average free flow velocity
T(x,y) is the traffic
C is the capacity of the road system
μ is an exponent >1 characterizing the sensitivity to congestion

- We write

$$C_T(x,y) \propto \tau(x,y)$$

Summary: the model

- Every time step, add a new individual at a random x

- The individual will choose to work in *y* (among N_c possible centers) such that

$$Z(x,y) = \eta(y) - \frac{d(x,y)}{\ell}\left[1 + \left(\frac{T(x,y)}{C}\right)^{\mu}\right]$$

is maximum

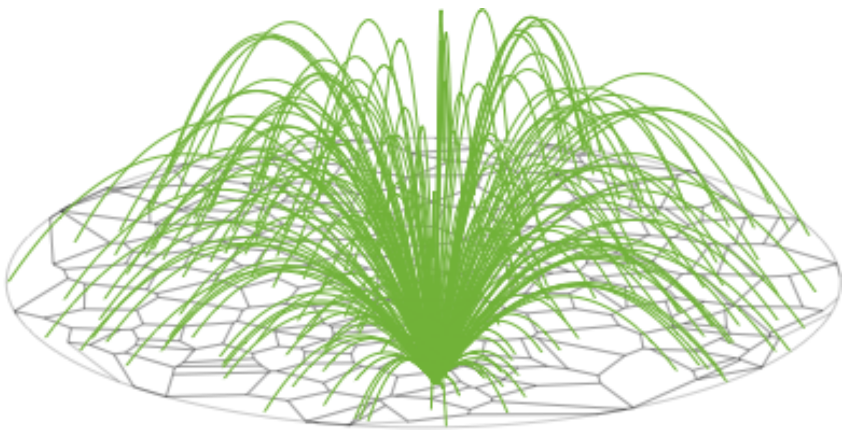
- $\eta(y)$ is the wage at y --> random

- $C_T(x,y)$ is the transportation cost from x to y: depends on the traffic from x to y --> congestion effects

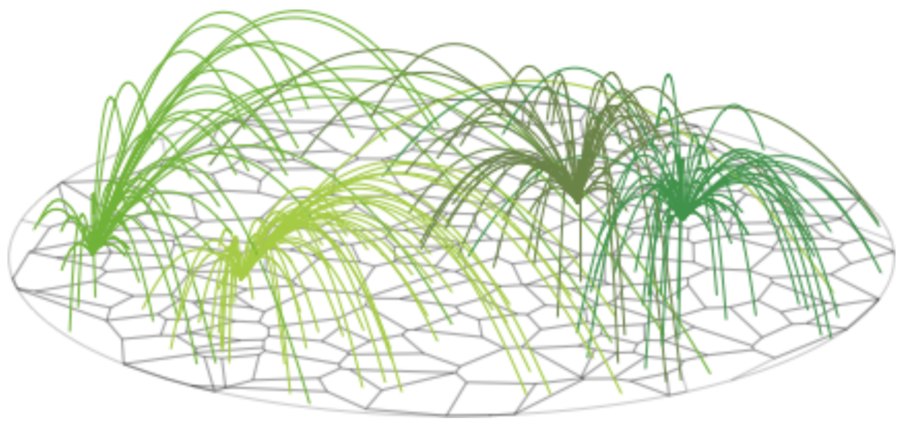
Louf, MB, PRL 2013

Results

- Depending on the values of parameters, we see two types of mobility patterns: Monocentric vs. polycentric



Monocentric



Polycentric

Monocentric-polycentric transition

- Start with one center $\eta_1 > \eta_j$
- All other subcenters have a zero traffic T(j)=0
- The number of individuals P increases, T(1) increases and at a certain point there is another j such that:

$$Z(i,j) > Z(i,1)$$

Or:

$$\eta_j - \frac{d_{ij}}{\ell} > \eta_1 - \frac{d_{i1}}{\ell}\left[1 + \left(\frac{P}{c}\right)^{\mu}\right]$$

Monocentric-polycentric transition

- Critical value for the population: effect of congestion !

$$P > P^* = C \left(\frac{\ell}{\sqrt{A} N_c} \right)^{1/\mu}$$

- C: capacity of the road system sets the scale

- If ℓ is too small, $P^* < 1$ and the monocentric regime is never stable

Results: scaling for the number of centers

- We obtain the average population for which a kth subcenter appears is:

$$\overline{P}_k = P^*(k-1)^{\frac{\mu+1}{\mu}}$$

which implies:

$$\boxed{H \sim \left(\frac{P}{P*}\right)^{\frac{\mu}{\mu+1}}}$$

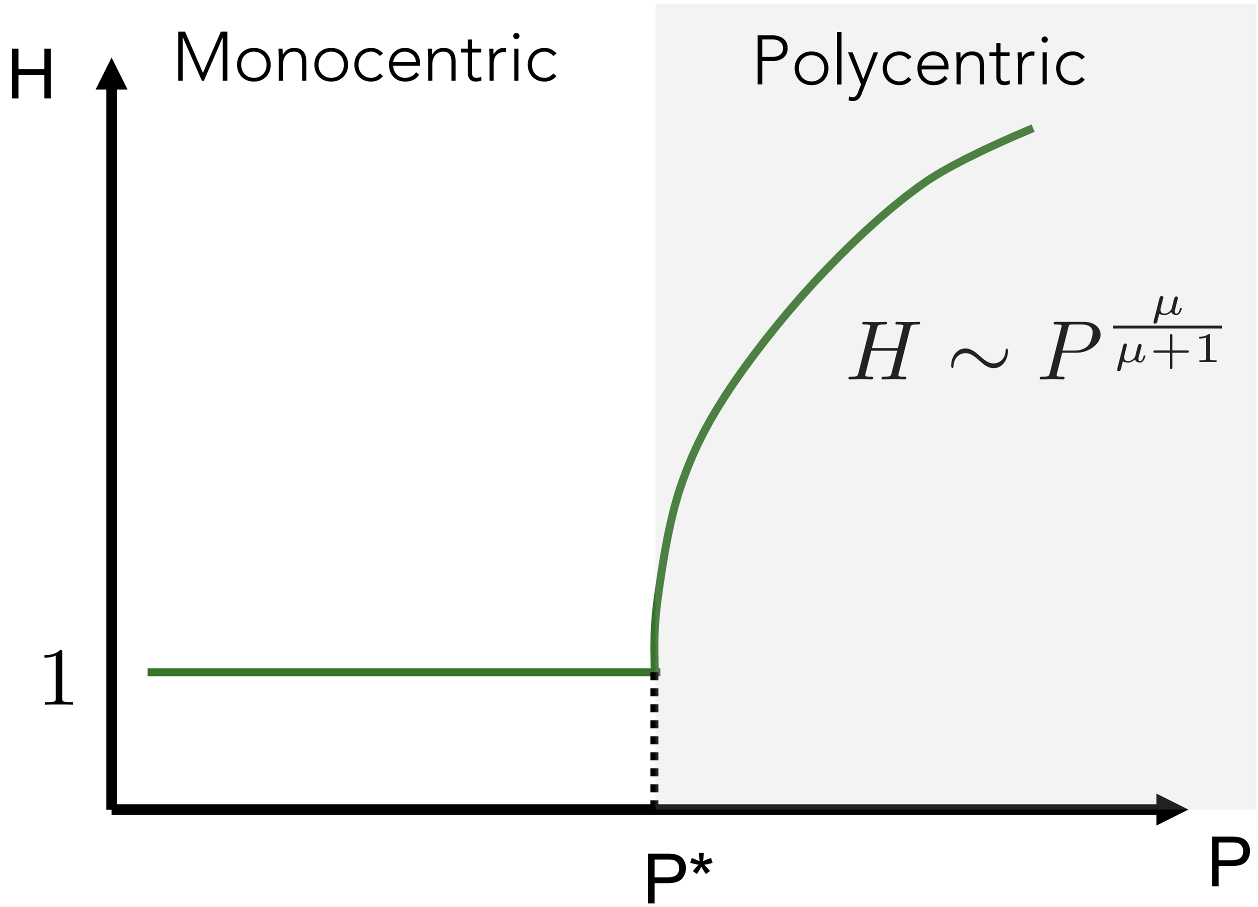
Sublinear relation !

- From US employment data (9000 cities)

$$H \sim P^{0.64} \quad (\Rightarrow \mu \simeq 2)$$

'Urban transition: Phase diagram'

Number of hotspots H versus population P

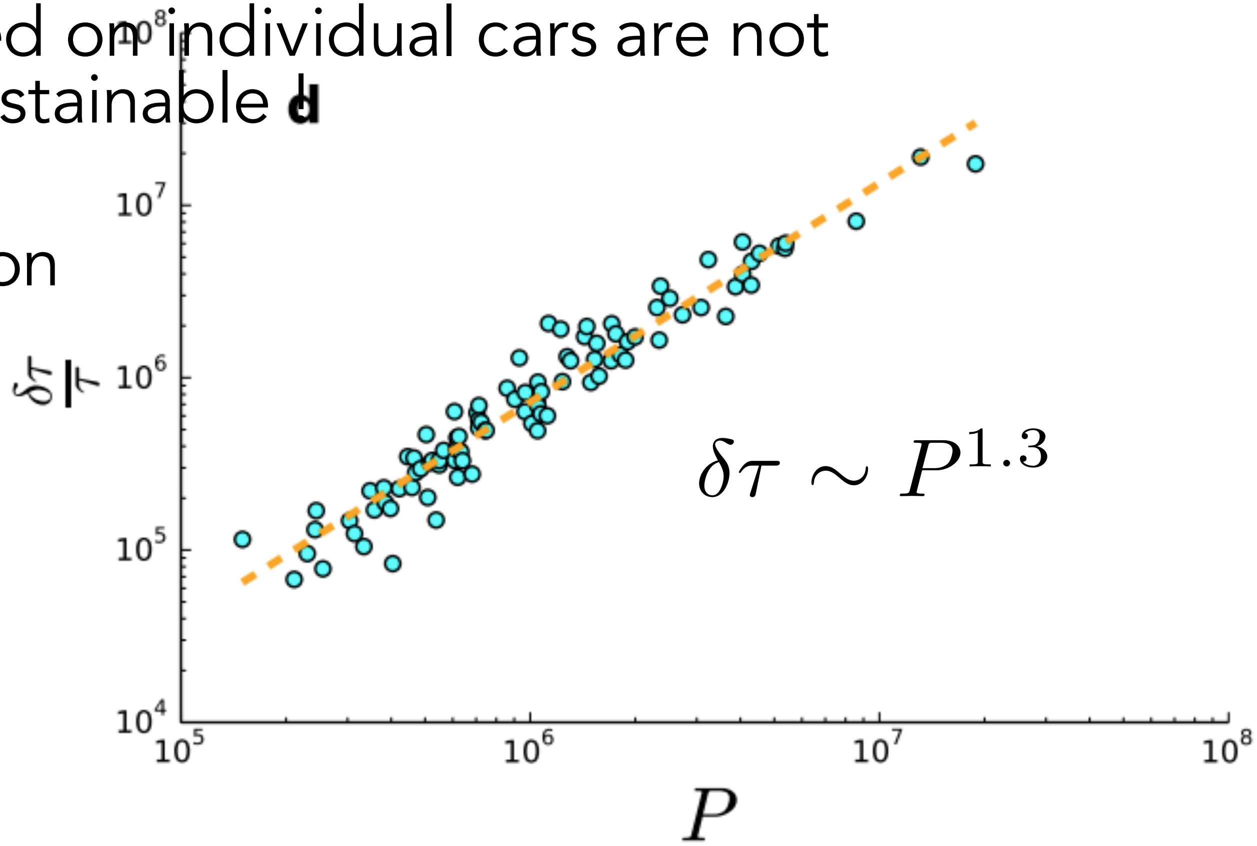


Important cause of polycentricity: traffic congestion

Discussion

- We observe a scaling of H with P which can be explained by a simple model integrating congestion
- Polycentrism is the natural response of cities to congestion, but not enough !
- For large P: Effect of congestion becomes very large
  => large cities based on individual cars are not
  economically sustainable !

Delay due to congestion
(US cities)



Louf, MB (2013, 2014)

Perspectives

- Science needs data ! Availability of data is critical for Science and also for improving our societies

- Parsimonious models translate for us the information hidden in large datasets, and provide guides to explore data, phenomena and to identify critical factors

- Future: modeling of complex systems…
  - Machine learning is useful for practical applications but do not improve so far (!) our knowledge
  - Mathematical modeling assisted by artificial intelligence ?
  - Or is this the end of mathematical modeling ?

http://www.quanturb.com
marc.barthelemy@ipht.fr

(Former and current) Students and Postdocs:

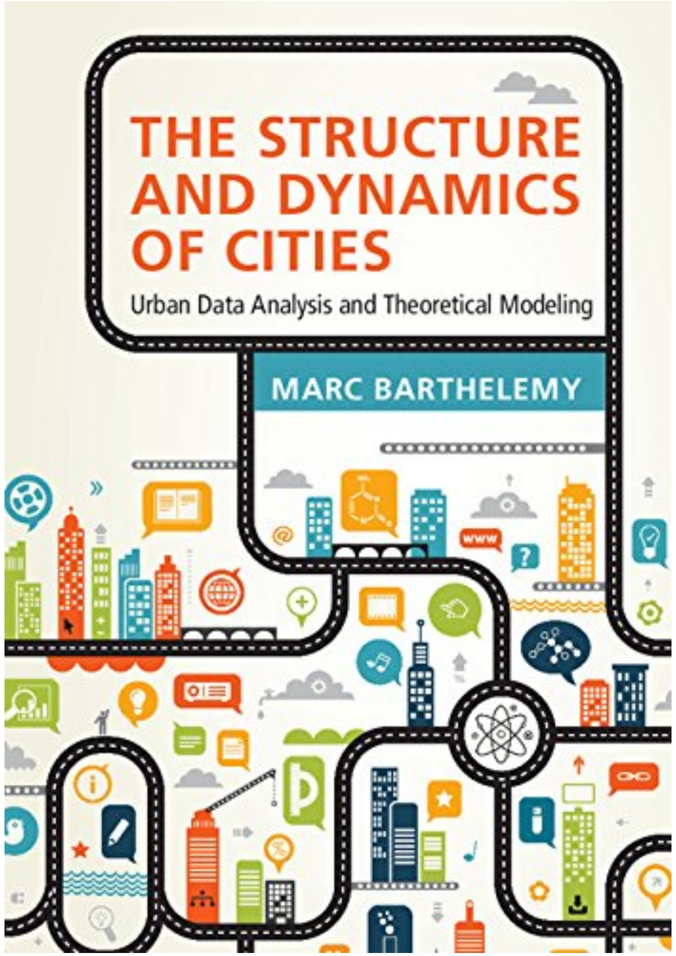
Giulia Carra (PhD student)	Riccardo Gallotti (Postdoc)
Remi Louf (PhD student)	A. Bourges (M2)
J. Depersin (M2)	R. Morris (Postdoc)
Alexandre Diet (M2)	Vincent Verbavatz (M2)
Valerio Volpati (Postdoc)	

Collaborators:

A. Arenas	M. Batty	A. Bazzani	H. Berestycki
G. Bianconi	P. Bordin	M. Breuillé	S. Dobson
M. Fosgerau	M. Gribaudi	J. Le Gallo	J. Gleeson
P. Jensen	M. Kivela	M. Lenormand	T. Louail
Y. Moreno	I. Mulalic	JP. Nadal	V. Nicosia
V. Latora	J. Perret	S. Porta	MA. Porter
JJ. Ramasco	S. Rambaldi	C. Roth	M. San Miguel
S. Shay	E. Strano	MP. Viana	

Mathematicians, computer scientists (27%)
Geographers, urbanists, GIS experts, historian (27%)
Economists (13%)
Physicists (33%)

Thank you for your attention.



Cambridge Univ Press 2017

http://www.quanturb.com
marc.barthelemy@ipht.fr