Urban Mobility Scaling: Lessons from 'Little Data'

(and implications for planning and climate change)

Galen J. Wilkerson
Technische Universität - Berlin
gjwilkerson@gmail.com

Overview

- Background
  - Complex Systems (whirlwind tour)
  - Urban Scaling, Mobility Scaling, 'Big' Mobility Data
- Our preliminary findings
  - Our data
  - Categories matter!
  - Urban Scale
  - Mobility scaling confirms previous results
  - Distance vs. purpose mechanisms
  - Time matters!
- Summary
- Future work
- Relevance to Climate Change and Planning

Background

- Complex systems (a whirlwind tour)
  - Self-organizing 'big' systems
    - bacteria colony, crowd of people
    - (> ~1000 = 10^3 parts) - described statistically
  - Formalism comes from Statistical Physics

    (e.g. Thermodynamics of glass of water – Ludwig Boltzmann, late 19th cent.)
  - Random processes – Brownian motion, diffusion

Background

- Complex systems (a whirlwind tour)
  - *Emergence* of large-scale (macroscopic) 'state' from states of small-scale (microscopic) parts. (e.g. 'solid';
  - *Phase transitions,* as macroscopic state changes (e.g. liquid to solid; peace to war; locusts swarming)
  - *Universality classes*, different types of large-scale phenomena that seem to have similar characteristics. (e.g. rate of diffusion; number of friends)
  - 'scale-free' = self-similar = fractal - seems to be found everywhere in self-organizing systems – (new 'laws of nature'?)

Background

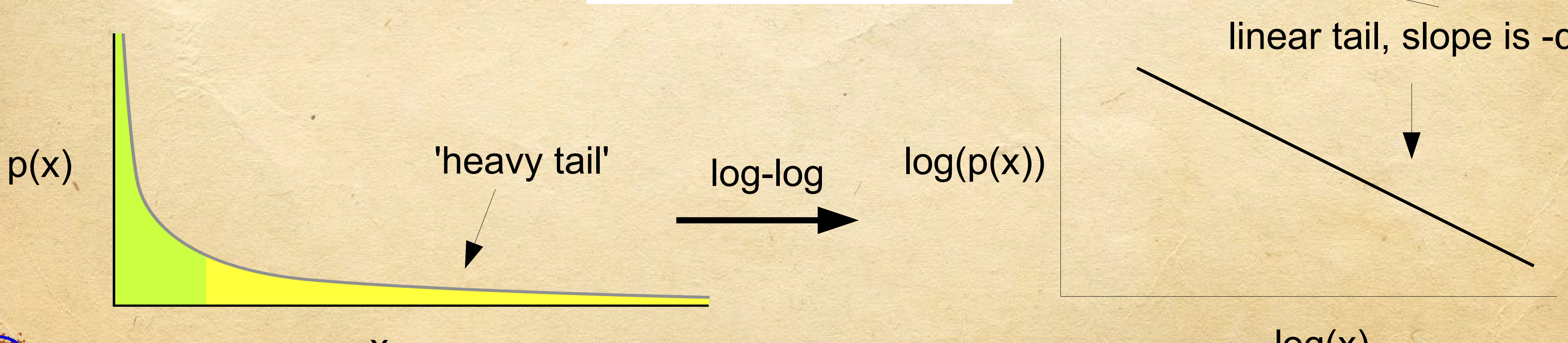
- Complex systems (a little math!)
  - power laws (one of many 'heavy-tailed' distributions)

$$p(x) = Cx^{-\alpha}$$

scaling exponent

$$\implies \log p(x) = -\alpha \log x + \log C$$

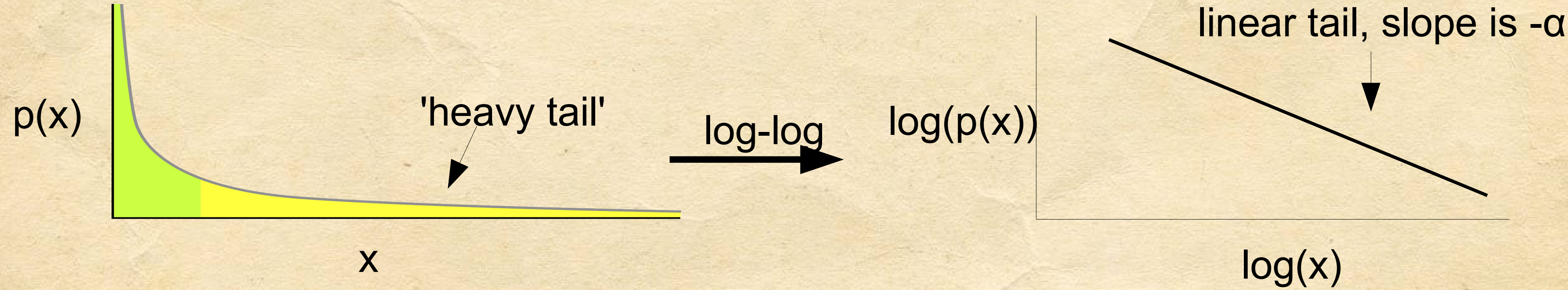
$$y = mx + b$$



[wikipedia.org]

Background

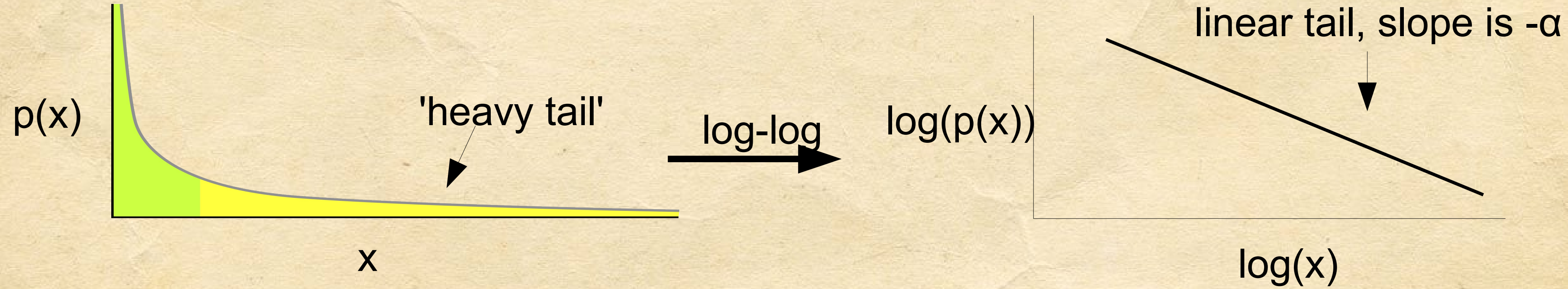
- Complex systems
  - power laws



- 'heavy tail' - rare events ***<u>do actually happen</u>***, and 'more often' than we may expect.
  - E.g. There are a few HUGE airports, VERY rich people, a few words that are used VERY OFTEN, etc.
- Inverse linear relationship between orders of magnitude of size and frequency (Zipf's law – to describe word frequency)
  - 'scale-free' = fixed rate of change between orders of magnitude = self-similar = fractal = ratios found at all scales!

Background

- Complex systems
  - power laws

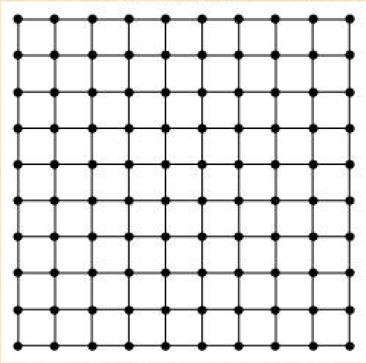


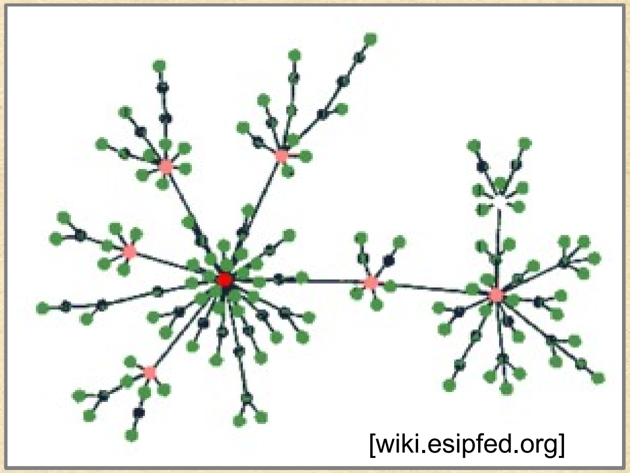
- Slope α describes *universality class* of phenomenon being measured (captures certain salient large-scale features of system)
- i.e. system with $2 < \alpha < 3$ is somehow fundamentally (and mathematically) different from system with

  $1 < \alpha < 2$, or $3 < \alpha < 4$ in terms of mean and variance, as well as underlying causal mechanisms.

Background

- Complex Networks





  – Another way of seeing complex systems, since there are (almost always) some kinds of interactions between parts.
    - Also big, described statistically (e.g. number of connections per node)
    - ***Many*** examples of such *real-world networks*:
      Social networks, lattices, metabolic networks, co-author networks, etc.
    - Some commonality with theoretical computer science (graphs)
    - Often number of neighbors (degree) follows a heavy-tail
    - (large literature and recent work!)

Background

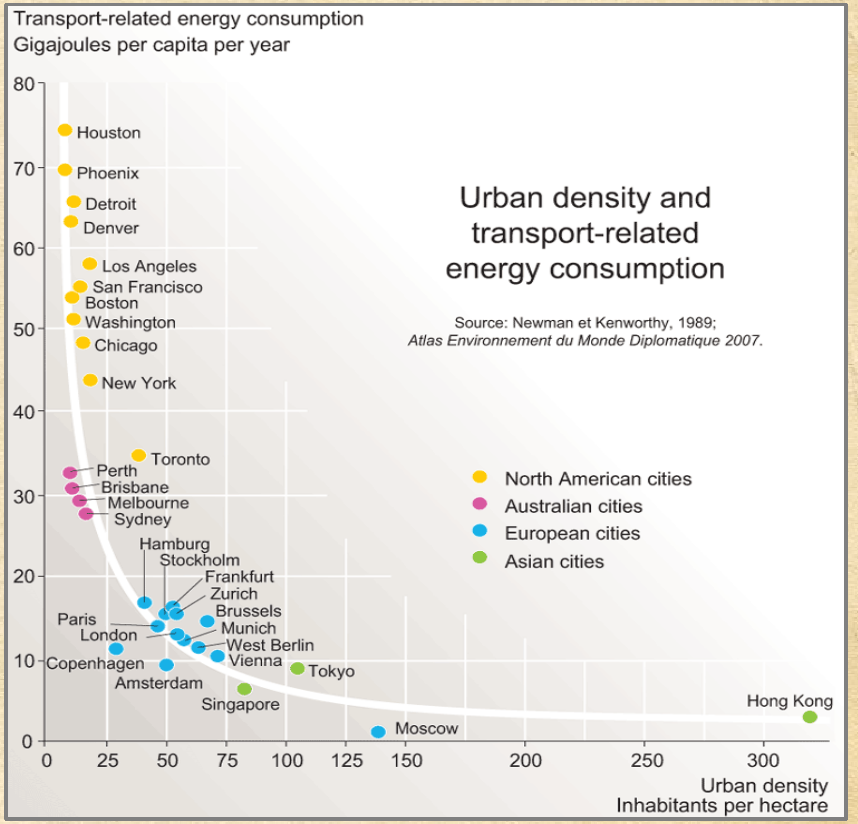
Caveat Emptor!

- Complex systems - potentially very useful and powerful, but we must be careful! (e.g. rare events outliers of normal dist. or tail of 'heavy-tail' dist.?)
  - Depends on random process underlying phenomenon.
- 'Big Data' is exciting, but we must understand what it is (and isn't) able to provide.
  - A “problem looking for a set of tools” → ~ statistical physics
- Humans are not molecules (or atoms, or ants)!
- Need to integrate this new perspective with urban & transportation planning German DLR, US RITA (DOT), European Union, IPCC, Think Tanks, etc.

- (...lots to say about each of these!!)

Background

- Urban Scaling
  - Newman and Kenworthy (1980) (transportation and urban planning, environmental science)
    - Per capita energy vs. density
  - Michael Batty et al. (many publications since 1970s) (urban planning/geography)
    - Work on spatial distribution of populations in cities (fractal) and related topics (spatial entropy), more recently, scaling as well.



[Newman, Kenworthy]

[M. Batty]

Roads colored by connectivity

Background

- Urban Scaling
  - Luis Bettencourt, Dirk Helbing, Geoffrey West (2007) (statistical physics)
    - "allometric scaling" of urban phenomena with city population

$$Y(t) = Y_0 N(t)^{\beta}$$

      - Similar to non-linear scaling of metabolism with mass of animal (mouse vs. whale)
      - Compare patents to gasoline sales
  - Sante Fe Institute Sustainability, ETHZ (Helbing et al.) FuturICT

Y	β
New patents	1.27
Inventors	1.25
Private R&D employment	1.34
"Supercreative" employment	1.15
R&D establishments	1.19
R&D employment	1.26
Total wages	1.12
Total bank deposits	1.08
GDP	1.15
GDP	1.26
GDP	1.13
Total electrical consumption	1.07
New AIDS cases	1.23
Serious crimes	1.16
Total housing	1.00
Total employment	1.01
Household electrical consumption	1.00
Household electrical consumption	1.05
Household water consumption	1.01
Gasoline stations	0.77
Gasoline sales	0.79
Length of electrical cables	0.87
Road surface	0.83

[PNAS 2007]

Background

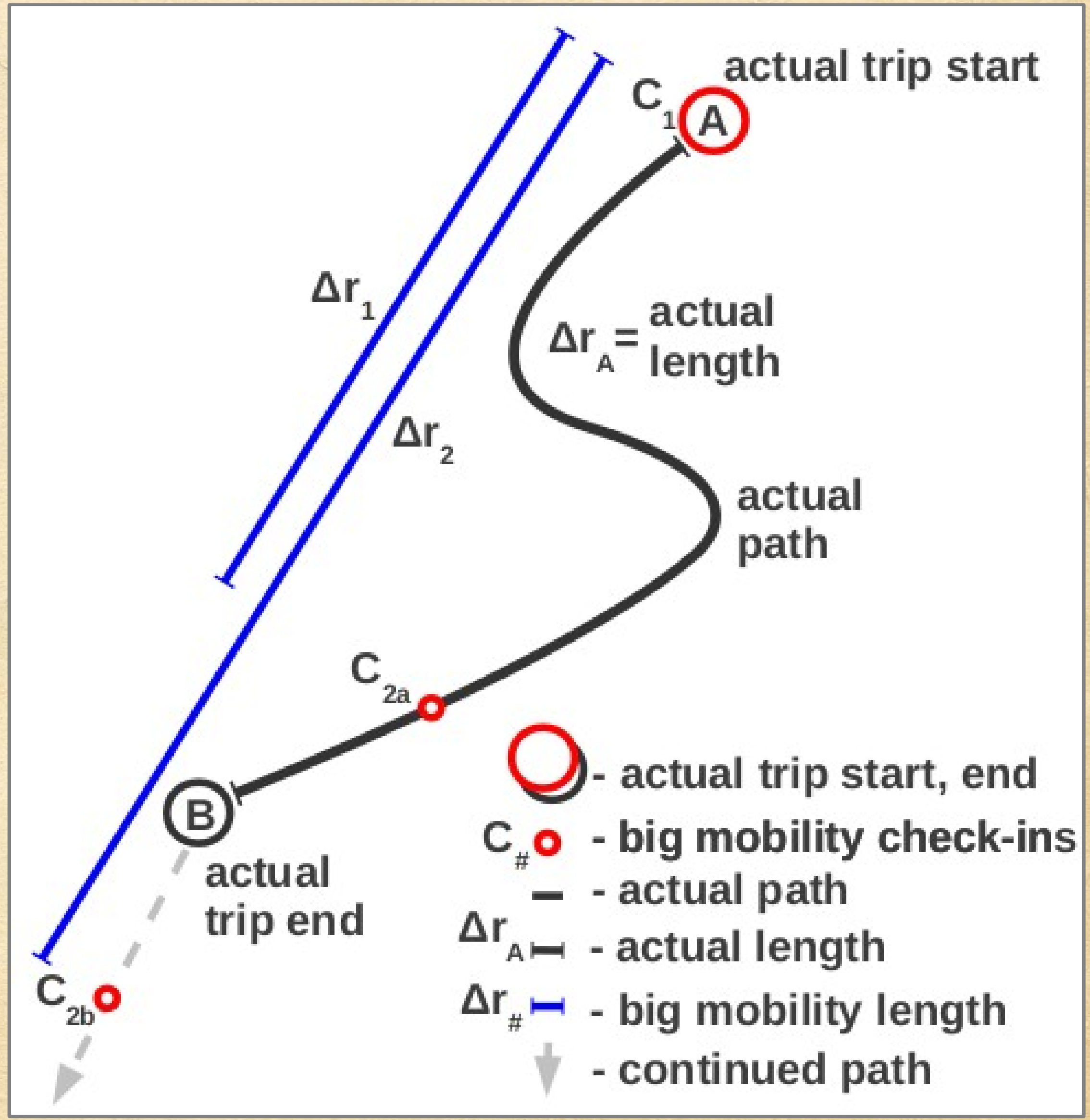
- Mobility Scaling and 'Big' mobility data
  - 'big' is not only defining feature
  - Collected as side-effect of some other activity
    - Not much survey design
  - 'check-ins'
    - (calls, tweets, posts, where's george submissions)
  - Low-dimensional (little known about trips or displacements)
  - Sometimes low-density (you need a lot of it)
    - Aggregated in space and time →
      mean-field approximations: "If we get enough data, things average out" (do they?)
  - Exciting/enticing/revolutionary (be careful!)
  - Allows large-scale system perspective

Background

- Mobility Scaling

  'check-ins'

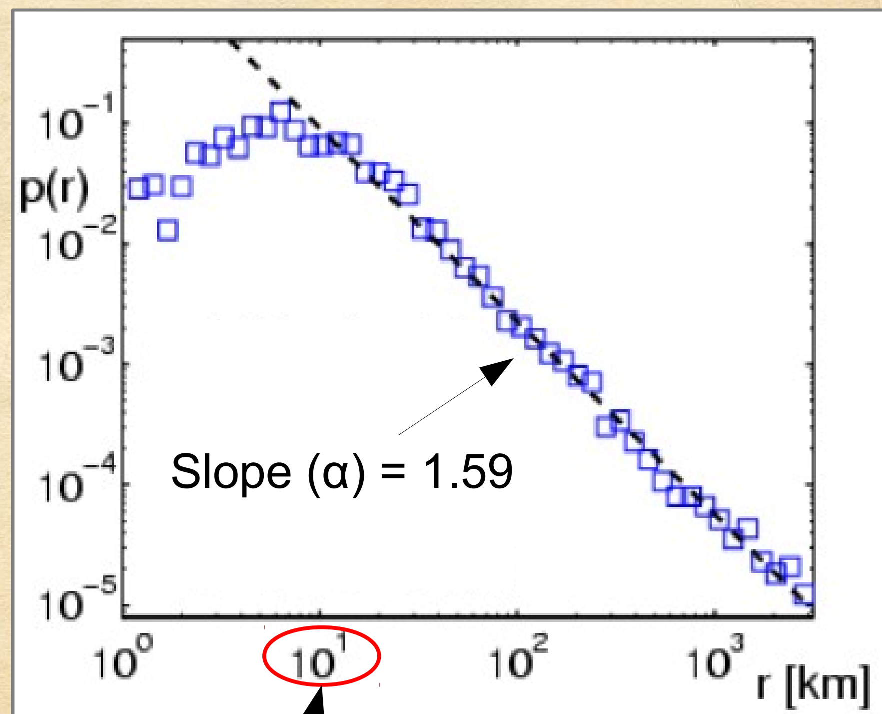
  - Assumes linear displacements
  - Assumes check-ins occur at trip start and end
  - Does not always account for sampling rate variability (e.g. walking vs. driving)



[G.Wilkerson 2014]

Background

- Mobility Scaling
  - Dirk Brockmann et al. (2006)
    (statistical physics – *now at HU Berlin!*)
  - pioneer of 'Big' mobility data for scaling
    (wheresgeorge.com)
    - Movement data as 'side effect' of experiment
    - Low-dimensional
  - Extension of animal-foraging literature
    - Monkeys, fish movements
  - Dollar-bill movement to estimate human mobility, from diffusion/random walk perspective (displacement over time)
  - Heavy-tailed long-distance displacements! (i.e. long-distance trips occur more than we might expect, since movement of money is *super-diffusive*):   scaling exponent α = 1.59

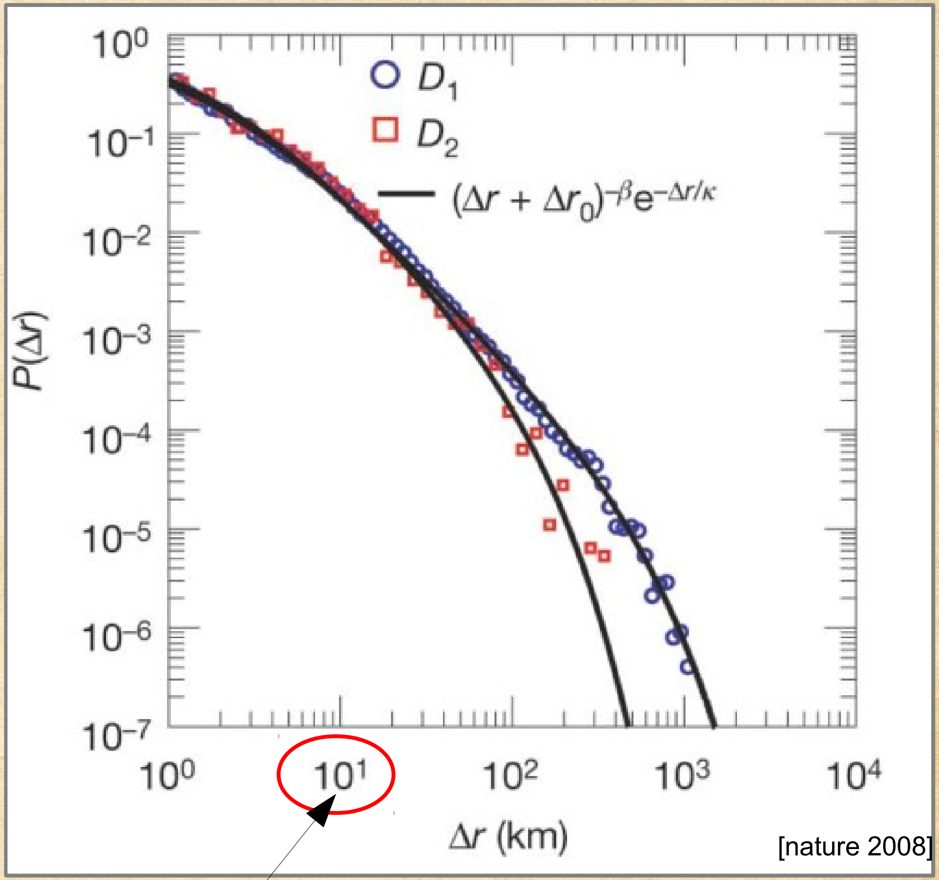


[nature 2006]

urban scale

Background

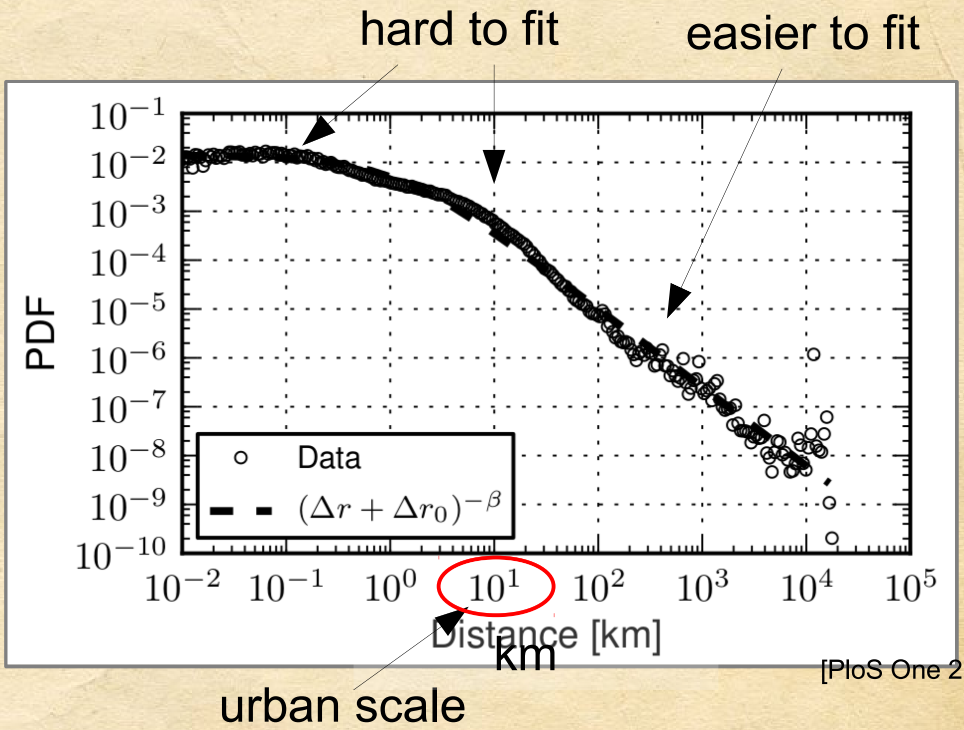
- Mobility Scaling
  - Gonzalez, Hidalgo, Barabasi (2008) (statistical physics)
  - 'Big' mobility data (phone calls: D1 = 16,264,308 displacements, constant locations (cell towers) D2 = 10,407 displacements)
  - Slightly more complex function due to fitting truncation
  - Perhaps challenging to characterize urban-scale mobility due to
    - sampling method (for D1, only during phone calls)
    - spatial resolution: Cell-phone tower resolution (3km)
  - Found scaling exponent of 1.75 (here called β)



urban scale

Background

- Mobility Scaling
  - Noulas et al. (2012)
    - Foursquare across planet
    - For long trips, measured scaling exponent of 1.50
    - Claims difficult to characterize mobility at urban scale from big data and distance mechanism alone
    - Intervening opportunity mechanisms seem to determine urban mobility patterns
      (e.g. we go to ~nearest
      grocery store)



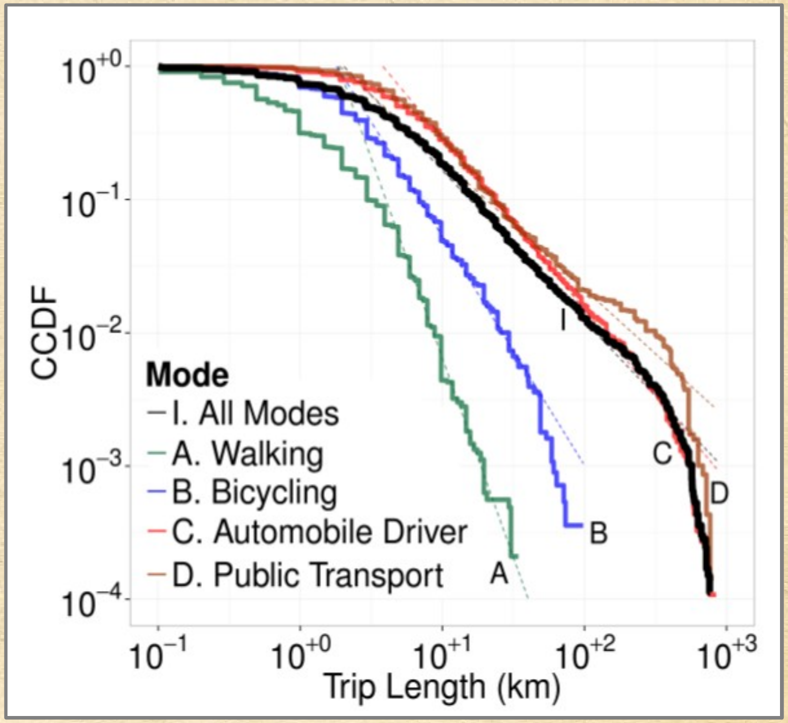
[PloS One 2012]

Our preliminary findings

- Our Data
  - Mobilität in Deutschland 2008 (MID 2008)
  - Conventional ('little' – not really) transportation survey
    - 25,922 households, 60,713 individuals, **193,290 trips and 36,182 travel** (w/overnight stay) events
    - Each trip has ~140 columns (dimensions) of information (weather, purpose, transportation mode, gender, age, etc.)
      - Allows investigation of many relationships
    - *Intentional Survey* designed to be large, balanced, and have resolution down to 100m
    - Our data describes actual trip lengths,
      - not linear displacements,
      - not 'check-ins' - all trips on survey day
- Methods
  - We just look at relative scaling of trip lengths, without fitting (for now).

Our preliminary findings

- Categories matter!
  - Our trips are high-dimensional (mode, purpose, gender, weather, etc.)
    - It is very obvious that trip lengths should be sensitive to mode!
    - Scaling exponents indicate different universality classes for modes (esp. walking, bicycling, vs. motorized)

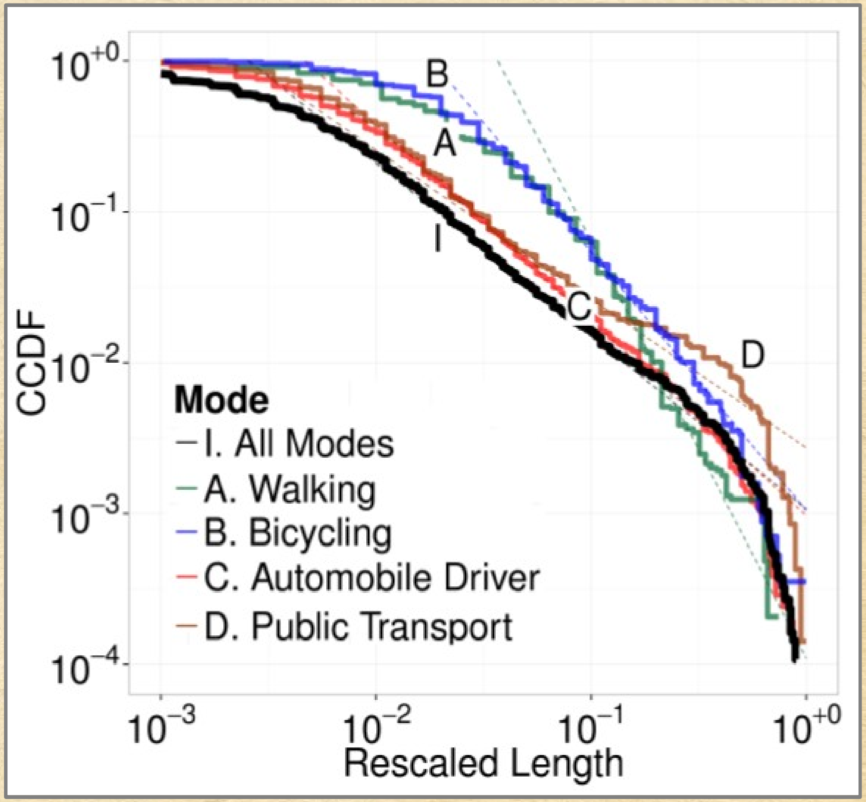


Mode	Count	α	ℓ_0 (km)	$\bar{\ell}$ (km)	σ^2
I. All Modes	52973	2.13	29.40	9.99	1313.79
A. Walk	14303	3.99	6.37	1.37	3.77
B. Bicycle	5581	2.72	6.37	3.47	30.06
C. Auto. Driver	18484	2.29	39.90	13.06	1331.84
D. Public Trans.	6944	1.97	27.98	16.34	2875.92
Auto. Passenger	7658	2.00	24.32	17.69	2949.11

[G.Wilkerson 2014]

Our preliminary findings

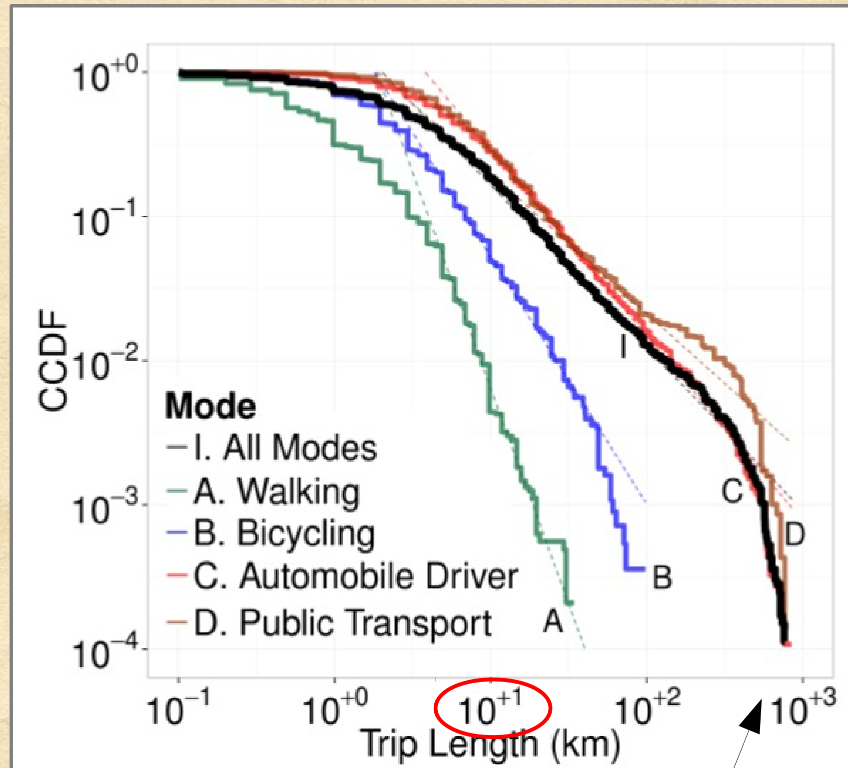
- Categories matter!
  - Trip lengths re-scaled by fraction of maximum are revealing!
  - Walking, bicycling may have different underlying process vs. fossil-fuel modes! (perhaps exponential? - i.e. almost no rare events – seem not heavy- tailed)
  - May agree with very recent research on human mobility in cities



[G.Wilkerson 2014]

Our preliminary findings

- Urban scale
  - In Germany, cities over 100,000 pop. have avg. 'diameter' of 14.89km ($\approx 10^{1.17}$ km)
  - Contrary to Noulas et al., mobility scaling exponents can be distinguished (and probably statistically fit) well within the urban scale!
  - Arguably, urban trips are complicated, not in straight line, influenced by many factors (including street geometry) vs. straight flights



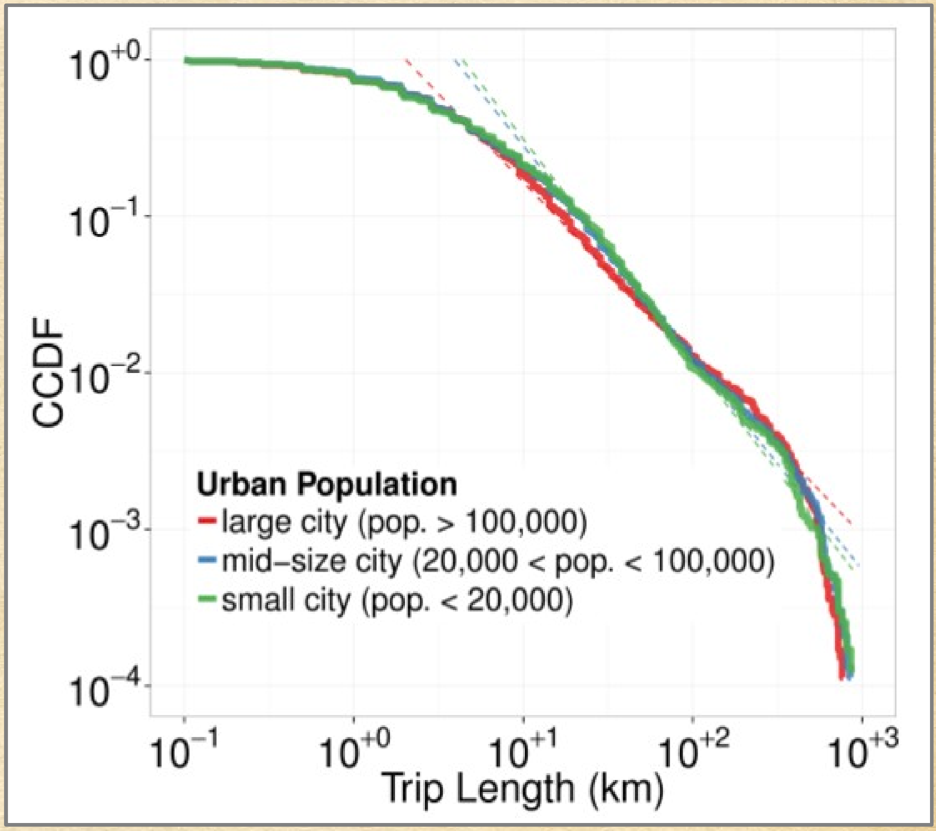
urban scale ($\approx 10^{1.17}$ km)

'diameter' of Germany ($\sim 10^{2.83}$ km)

[G.Wilkerson 2014]

Our preliminary findings

- Trip length scaling with population
  - It is possible, but not entirely certain that trip lengths scale allometrically with population



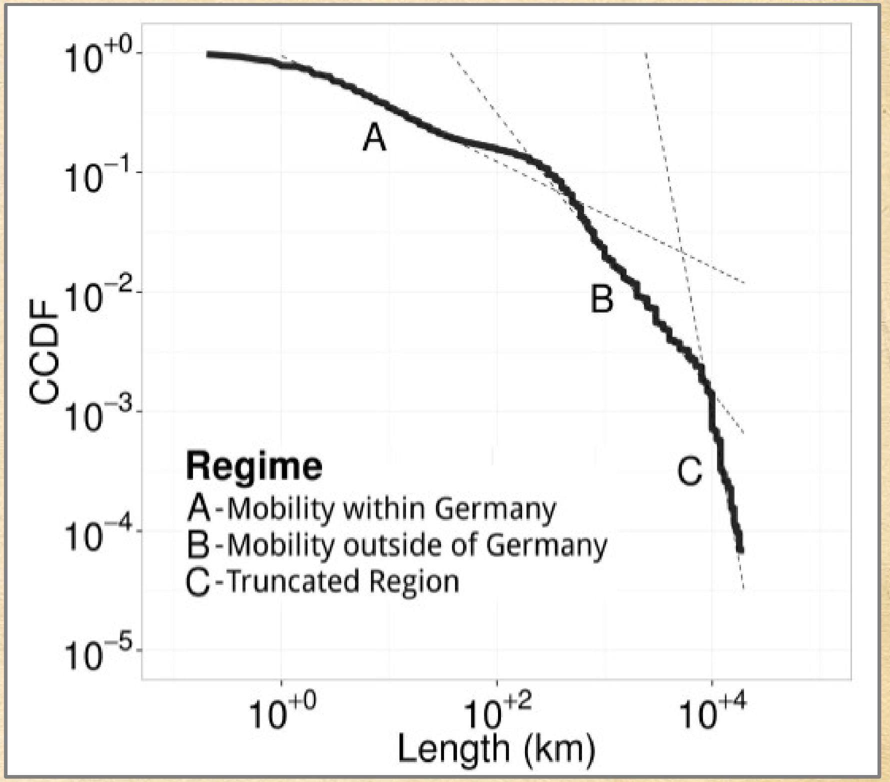
Urban Population	Count	α	ℓ_0 (km)	$\bar{\ell}$ (km)	σ^2
small ($< 20k$)	23433	2.41	43.32	10.52	1202.28
medium ($20k$-$100k$)	53038	2.35	30.38	10.62	1329.72
large ($> 100k$)	53011	2.13	29.40	9.99	1312.92

[G.Wilkerson 2014]

Our preliminary findings

- Mobility scaling results seems to agree with previous studies when trips and travel taken together (daily & overnight mobility)
  - We have α = 1.44, previous values ranged from α = 1.50 to 1.75

- Truncation at 'diameter' of Germany (~$10^{2.83}$ km) and distance of Thailand (~$10^{3.94}$ km)

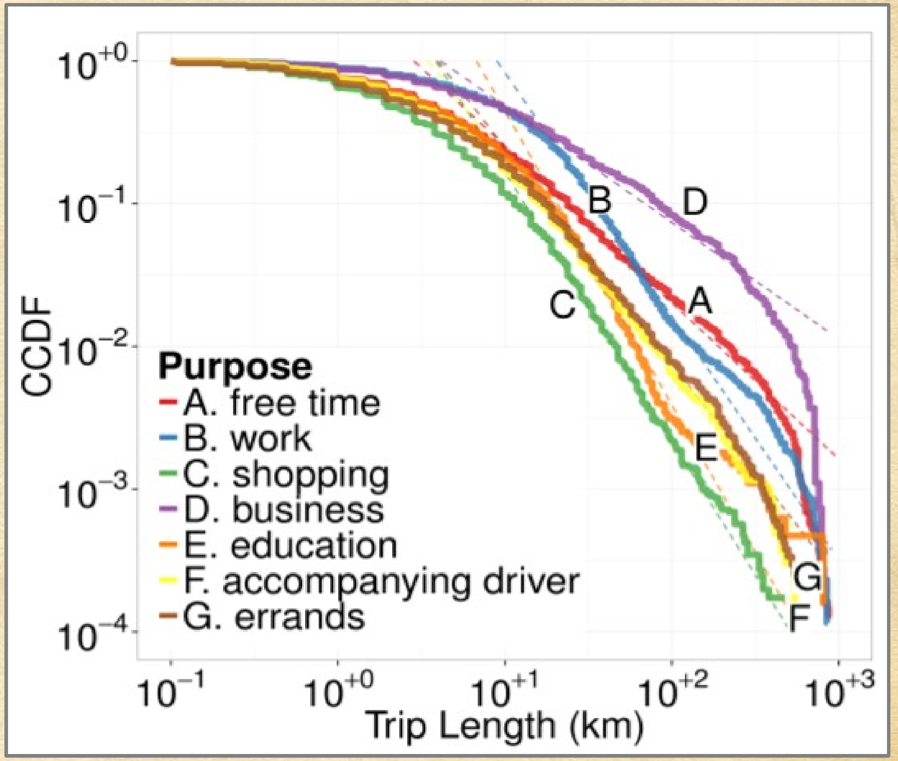


Regime	Count	α	ℓ_0 (km)	$\bar{\ell}$ (km)	σ^2
A	209,045	1.44	1.81	48.97	14,727.00
B	8,055	2.17	816.00	1,670.36	2,172,741.49
C	380	5.91	11,000.00	11,312.92	7,047,781.70

[G.Wilkerson 2014]

Our preliminary findings

- Purpose matters
  - Education vs. business, vs. others
  - Distance vs. intervening opportunity mechanisms
    - Response to purpose implies a possible connection between them



Purpose	Count	α	ℓ_0 (km)	$\bar{\ell}$ (km)	σ^2
education	12704	3.06	31.07	8.15	574.29
shopping	40322	2.88	35.15	5.19	196.73
work	25808	2.71	38.95	17.40	1654.51
errands	23716	2.51	45.13	8.06	593.81
accompanying driver	16447	2.50	32.30	7.74	476.70
free time	61152	2.10	30.38	13.55	2209.65
business	2706	1.82	12.35	36.58	8011.00

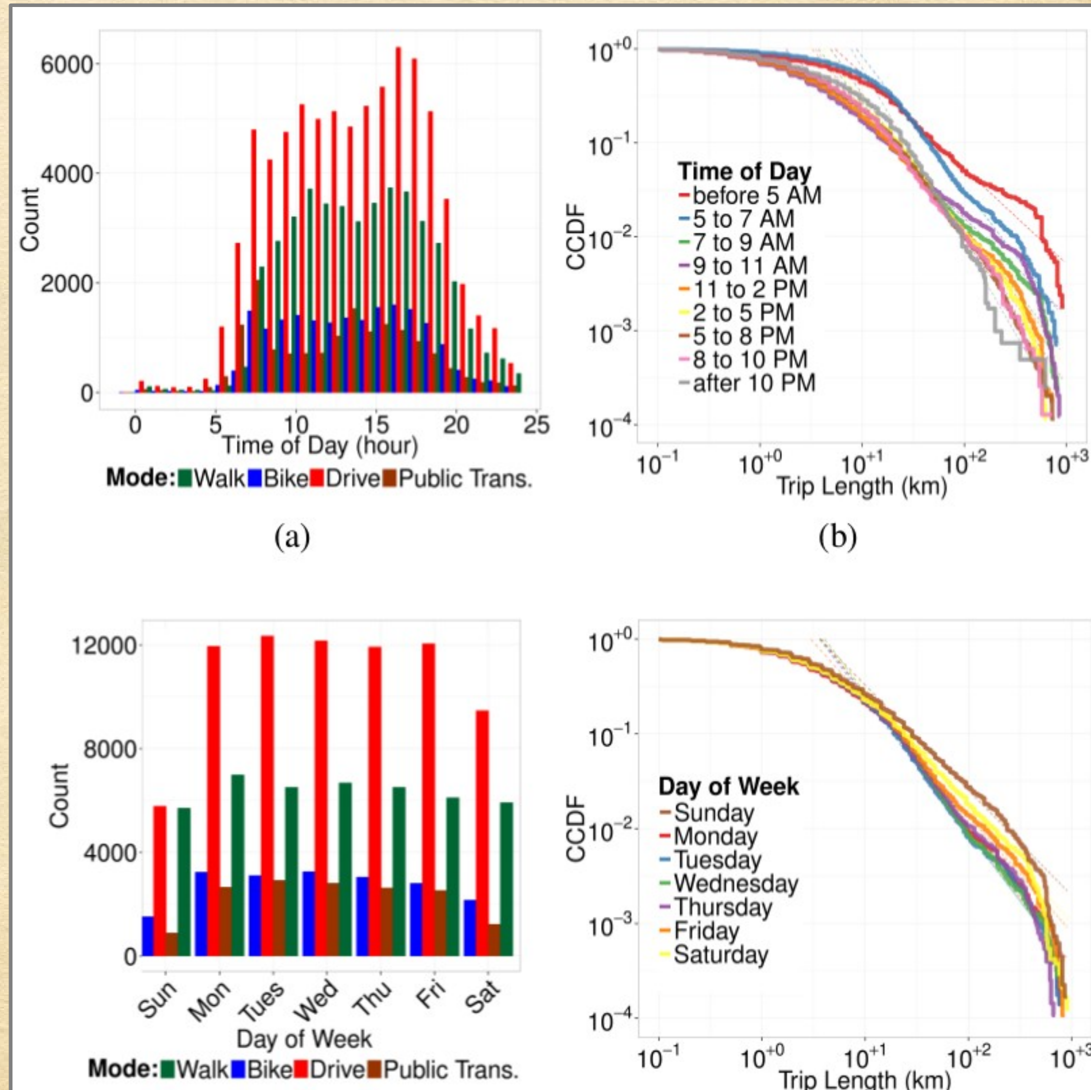
[G.Wilkerson 2014]

Our preliminary findings

- Time matters!

Time of Day	Count	α	ℓ_0 (km)	$\bar{\ell}$ (km)	σ^2
before 5 AM	1670	2.01	25.27	32.92	10,123.99
5 to 7 AM	7026	2.41	20.58	23.71	3,268.39
7 to 9 AM	21991	2.33	16.15	11.29	1,717.64
9 to 11 AM	24511	2.03	10.45	11.31	2,159.07
11 to 2 PM	37693	2.32	31.36	9.49	1,046.26
2 to 5 PM	43375	2.43	51.30	10.34	868.00
5 to 8 PM	34742	2.55	31.36	9.68	705.90
8 to 10 PM	7819	2.39	34.30	9.58	684.20
after 10 PM	4060	2.89	30.40	10.94	550.62

Day of Week	Count	α	ℓ_0 (km)	$\bar{\ell}$ (km)	σ^2
Sunday	17768	2.11	32.34	15.84	2,652.07
Monday	28476	2.42	34.20	9.66	1,026.79
Tuesday	28449	2.42	38.81	9.47	919.62
Wednesday	28649	2.46	48.45	9.86	966.94
Thursday	27787	2.38	38.95	10.07	943.46
Friday	27878	2.22	43.23	11.46	1,507.96
Saturday	23880	2.23	32.30	12.60	1,789.28



[G.Wilkerson 2014]

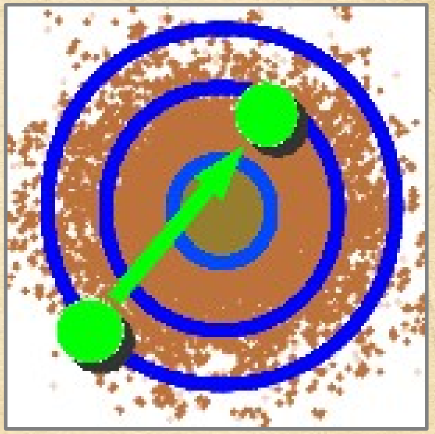
Summary

- Systems – scaling approach may be useful to characterize most salient urban features
  - Scaling exponent and shape of curve reflect underlying large-scale processes
  - Statistical laws may govern interaction of large-scale urban parameters
- 'Big' check-in mobility data aggregation must be handled carefully.
- Conventional transportation data can also contribute significantly to urban mobility scaling research
  - Categories matter
  - Urban-scale mobility can be subject to categories.
  - There may be allometric scaling of trip lengths.
  - Human-powered modes may be subject to different underlying processes than fossil-fuel powered modes.
  - Mode, purpose, and other parameters may link distance- and intervening opportunity mechanisms for trip length.

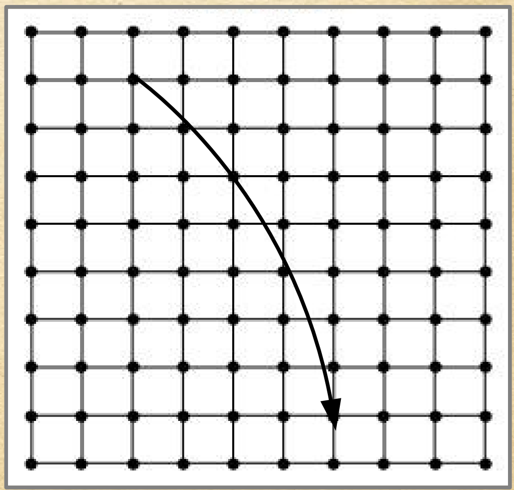
Future work

- Better connection to planning literature
- Careful statistical fitting and understanding of underlying processes
- 'Dimensionality reduction'
  (~140 parameters, which determine trip length, duration, etc.?)
- Trip duration, velocity, energy (!!) distributions
- Exponential vs. heavy tails (human-powered vs. fossil-fuel-powered)
    - cf. Animal foraging and recent findings on urban-scale exponential trip lengths
- Intervening-opportunity vs. purely distance-based arguments

Future work



- Basic geometric mechanisms
  - Distance distribution between points A, B chosen randomly in a disc?
- Using other data sources (carefully - big data)



- Networks (social and infrastructure)
  - Highways as 'small-world' connections vs. 'lattice' of city
  - Requires energy to build and use them (autos) !
  - Relation to mode share?

Relevance to Climate Change and Planning

- Useful to understand city as a whole, complex, interacting, self-organizing system, with statistical 'laws'. (What are 'possible cities'?)
  - Gedankenexperiment – probability that no-one drives tomorrow? Why?
  - certain policies will work, others ***impossible***, due to laws relating population, area, energy, mode share, etc.
- Small change in scaling exponent (α) for fossil-fuel modes can imply a large change in total energy and CO_2 emissions   (logs = orders of magnitude!)
- Be smart! – use response of ***most salient*** system features (scaling exponent) to understand how policy mechanisms can influence well-being and sustainability. (time, mode, purpose, …)
- Be smarter! - Develop a science of sustainability - complete with formal definitions, theories and laws! (scaling-based sustainability index - Muneepeerakul, R., & Qubbaj, M. R. (2012))
- Be smartest! - Develop systems to not only collect, but disseminate large-scale information to help us share resources and have more well-being and sustainability.
  - Larger population = more sharing?
  - Higher density telecommunications = phase transition in cooperation?
  - “Smart Mobility for the 21st Century”, Winner - Transportation Efficiency, MIT Climate Colab (2013)

Thanks! and References

gjwilkerson@gmail.com

Wilkerson, G., Khalili, R., and Schmid, S., “Urban Mobility Scaling: Lessons from 'Little' Data”, Netscicom 2014 (IEEE Infocom) http://arxiv.org/abs/1401.0207

General background on Complex Systems and Complex Networks:

Mitchell, Melanie. *Complexity: A guided tour.* Oxford University Press, 2009.

Publications by Mark Newman, Schlomo Havlin, Eugene Stanley, Per Bak, Laszlo Barabasi, and many others.

Urban scaling:

Bettencourt, Luís MA, et al. "Growth, innovation, scaling, and the pace of life in cities." Proceedings of the National Academy of Sciences 104.17 (2007): 7301-7306.

Batty, Michael. Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals. The MIT press, 2007.

Muneepeerakul, R., & Qubbaj, M. R. (2012). The effect of scaling and connection on the sustainability of a socio-economic resource system. Ecological Economics, 77(C)

Mobility scaling:

Brockmann, Dirk, Lars Hufnagel, and Theo Geisel. "The scaling laws of human travel." Nature 439.7075 (2006): 462-465.

Gonzalez, Marta C., Cesar A. Hidalgo, and Albert-Laszlo Barabasi. "Understanding individual human mobility patterns." Nature 453.7196 (2008): 779-782.

Noulas, Anastasios, et al. "A tale of many cities: universal patterns in human urban mobility." PloS one 7.5 (2012): e37027.

Cho, Eunjoon, Seth A. Myers, and Jure Leskovec. "Friendship and mobility: user movement in location-based social networks." Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011.