Revisiting urban economics in light of data

Marc Barthelemy

CEA, Institut de Physique Théorique, Saclay, France

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

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

ENPC-2017

Outline

- Modelling mobility
  - Mobility: gravity law
  - The radiation model
- Relation between commuting distance and income
  - Empirical results
  - Testing the McCall model of job search
  - The 'closest opportunity' model
- Mobility: statistical properties

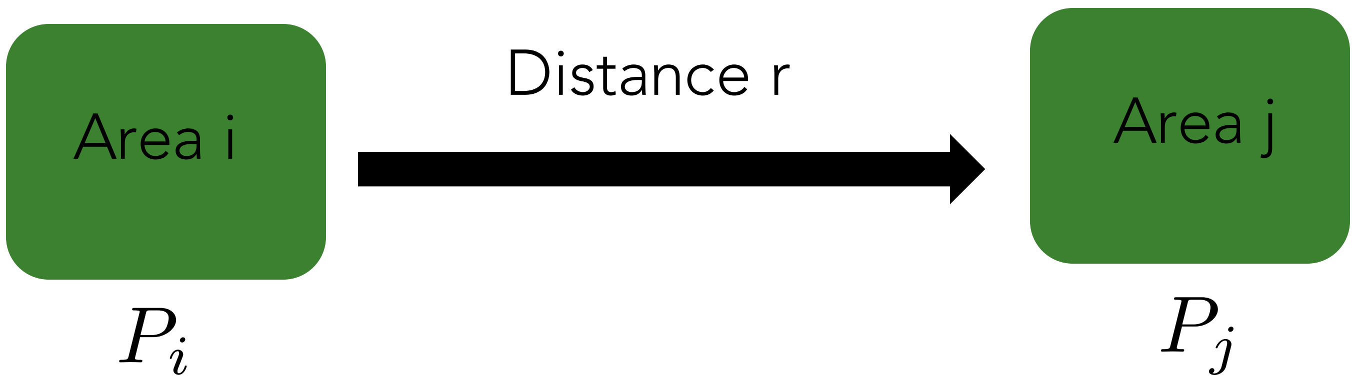
Motivations: understanding mobility

- Structure of cities: spatial distribution of residences and activity centers

- Useful for modelling many practical applications:
  - Urban planning (transport planning)
  - epidemic spread
  - …

General models: gravity and radiation

The gravity model



- Number of trips between i and j ?

$$T_{ij} = K \frac{P_i P_j}{r^\sigma}$$

- Gravity law (Reilly 1929, Zipf 1946)

The gravity model

- Problems with the gravity model
  - Congestion ?
  - Theoretical derivation ?

- A derivation proposal (Wilson, 1967): number of ways to construct a configuration $\{T_{ij}\}$

$$\Omega = \frac{T!}{\prod_{ij} T_{ij}}$$

- Maximize Ω with constraints:

$$\sum_i T_{ij} = T_j, \quad \sum_j T_{ij} = T_i$$

$$\sum_{ij} T_{ij} = T, \quad \sum_{ij} T_{ij} C_{ij} = C$$

The gravity model

- We then obtain

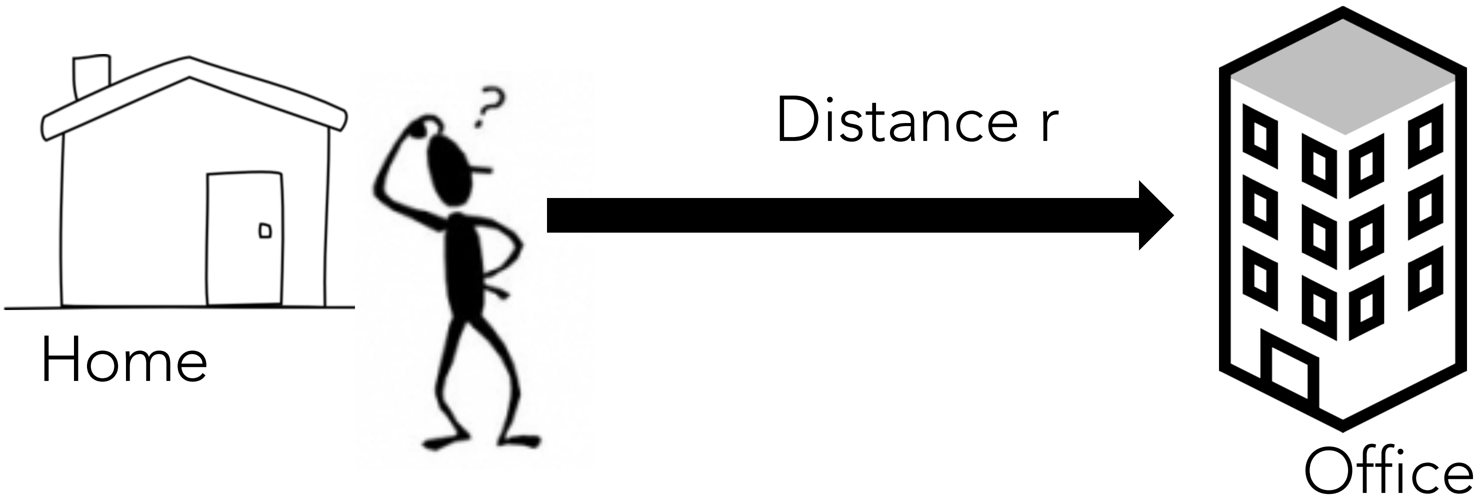
$$T_{ij} \propto T_i T_j \mathrm{e}^{-\beta C_{ij}}$$

- But the cost needs to be given

$$C_{ij} \propto \begin{cases} d_{ij} \Rightarrow T_{ij} \sim \mathrm{e}^{-d(i,j)} \\ \log d(i,j) \Rightarrow T_{ij} \sim 1/d(i,j)^{\sigma} \end{cases}$$

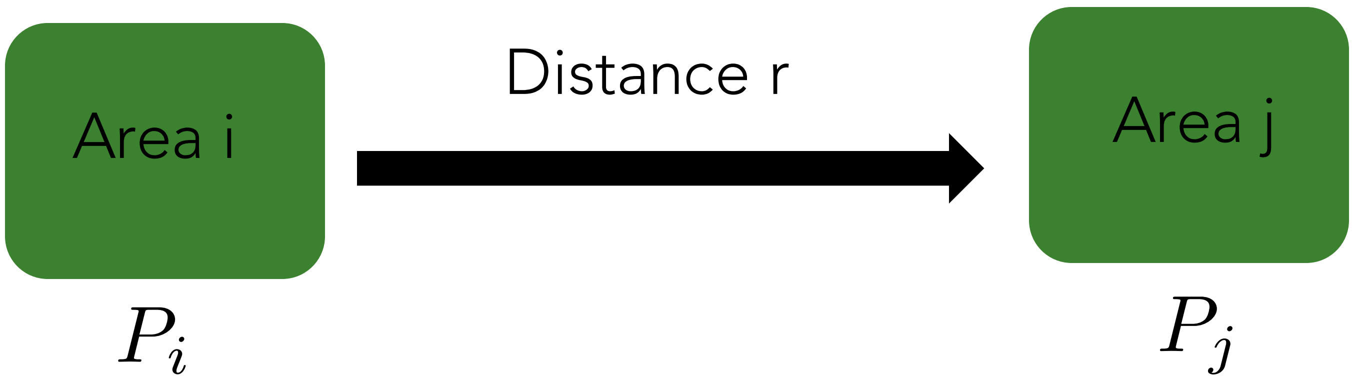
The radiation model (Simini et al, 2012)

- Alternative to the Gravity model



- How to choose a job

The radiation model (Simini et al, 2012)



- Each individual has his threshold z_i
- And looks for the closest job $\tilde{z}_j$ such that

$$\tilde{z}_j > z_i$$

- A choice $z_i = \max\{X_1, X_2, \ldots, X_{P_i}\}$

$$\tilde{z}_j = \max\{X_1, X_2, \ldots, X_{P_j}\}$$

with $X \sim p(X)$ (and cumulative F)

The radiation model (Simini et al, 2012)

- The probability of being emitted at i and absorbed at j is:

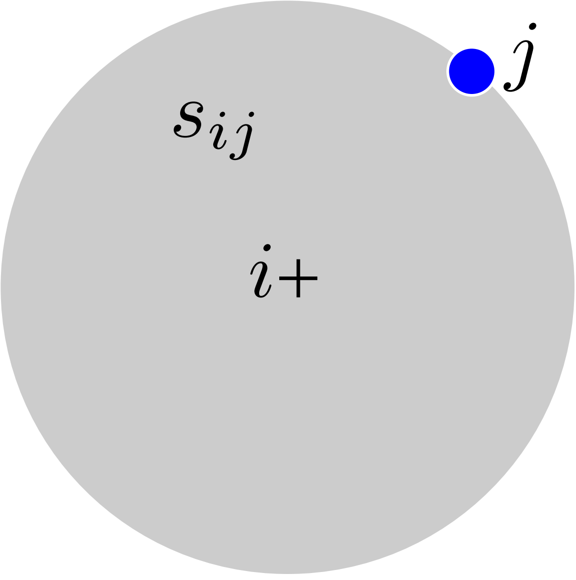
$$P(i \to j) = \int \mathrm{d}z P_{P_i}(z) P_{s_{ij}}(< z) P_{P_j}(> z)$$

where

$$P_{P_i}(z) = P_i F(z)^{P_i - 1} \frac{\mathrm{d}F}{\mathrm{d}z}$$

$$P_{s_{ij}}(< z) = F(z)^{s_{ij}}$$

$$P_{P_j}(> z) = 1 - F(z)^{P_j}$$



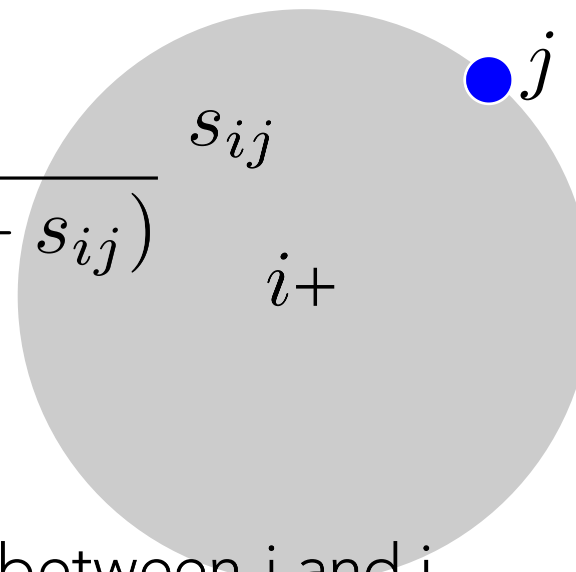
The radiation model (Simini et al, 2012)

- We then have:

$$P(i \to j) = P_i \int \mathrm{d}z F(z)^{P_i - 1} \frac{\mathrm{d}F}{\mathrm{d}z} F(z)^{s_{ij}} (1 - F(z)^{P_j})$$

a change of variables then gives (u=F)

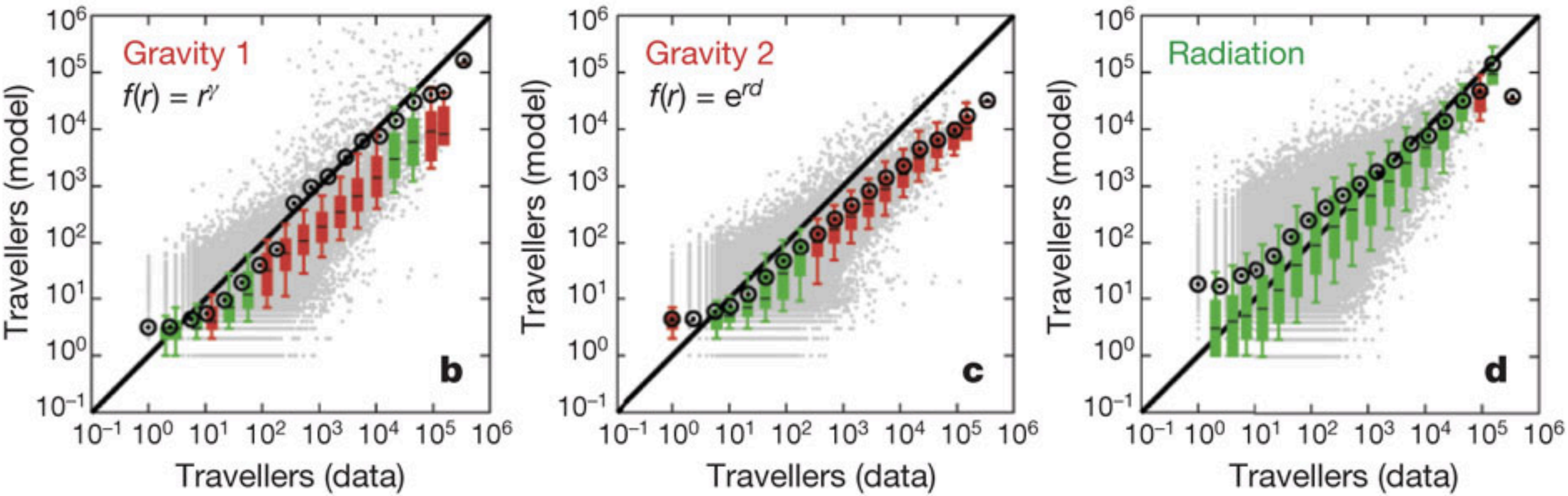
$$P(i \to j) = \frac{P_i P_j}{(P_i + s_{ij})(P_i + P_j + s_{ij})}$$



Note: link with the rank = # individuals between i and j
(ie. such that d(i,w)<d(i,j))

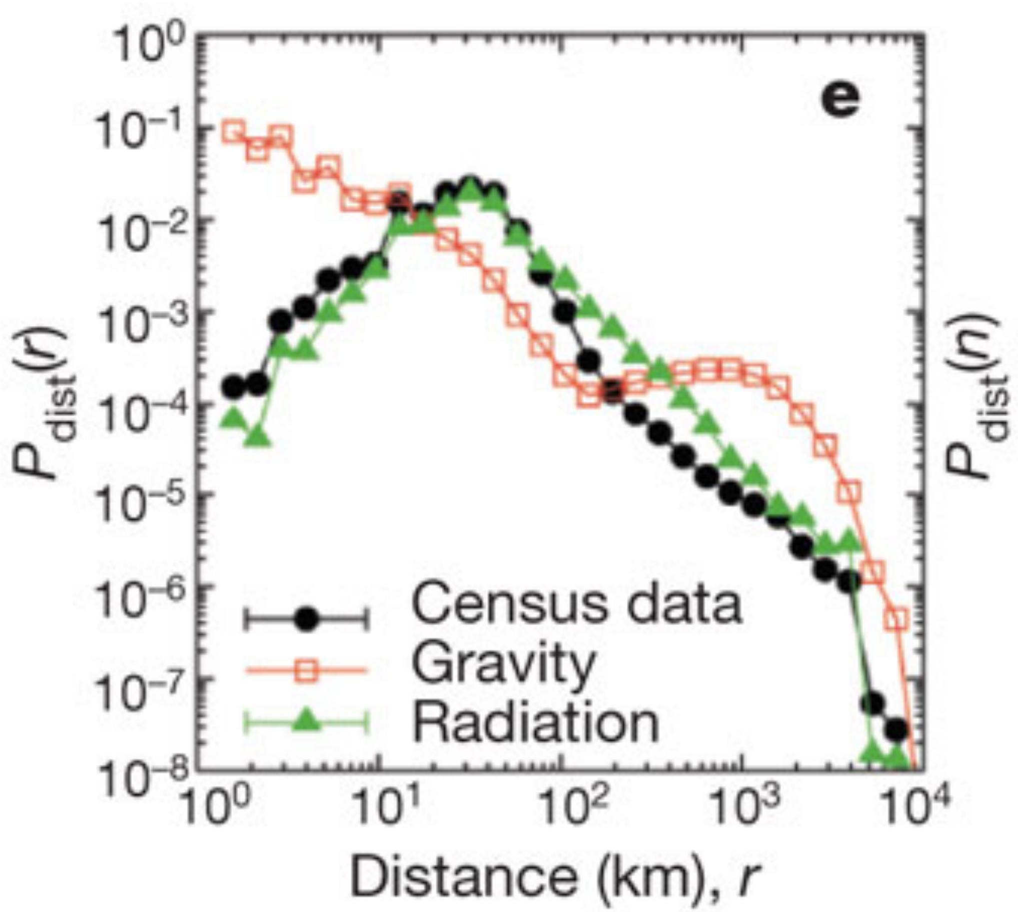
Results

- Comparison with empirical data (US data):



Results

- Comparison with empirical data (US data):

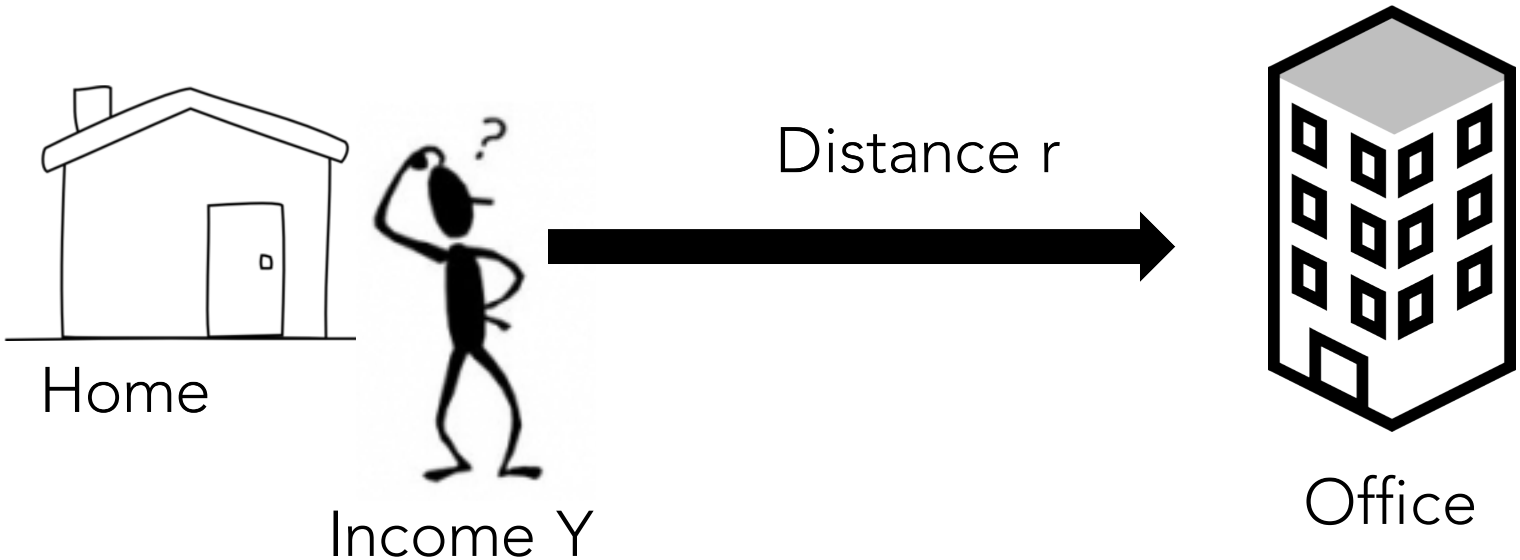


Mobility and socio-economics: Commuting distance and income

What is the relation between income and commuting distance ?

with G. Carra, I. Mulalic, M. Fosgerau

Income and commuting distance

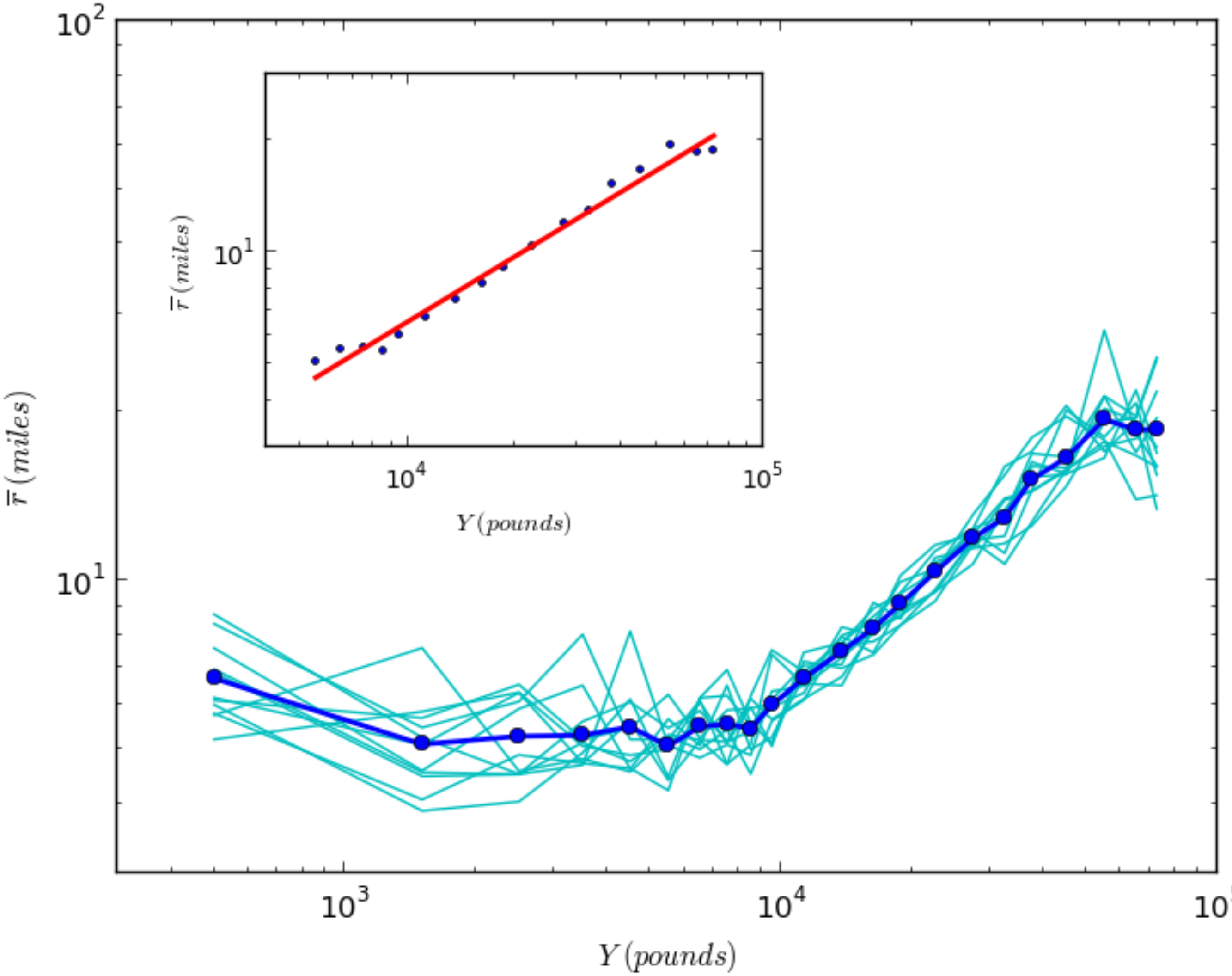


- Finding/choosing a job
- Questions:
  - Average commuting distance r versus Y
  - Distribution P(r|Y) ?

Income and commuting distance: UK data

$$\bar{r} \sim Y^{\beta}$$

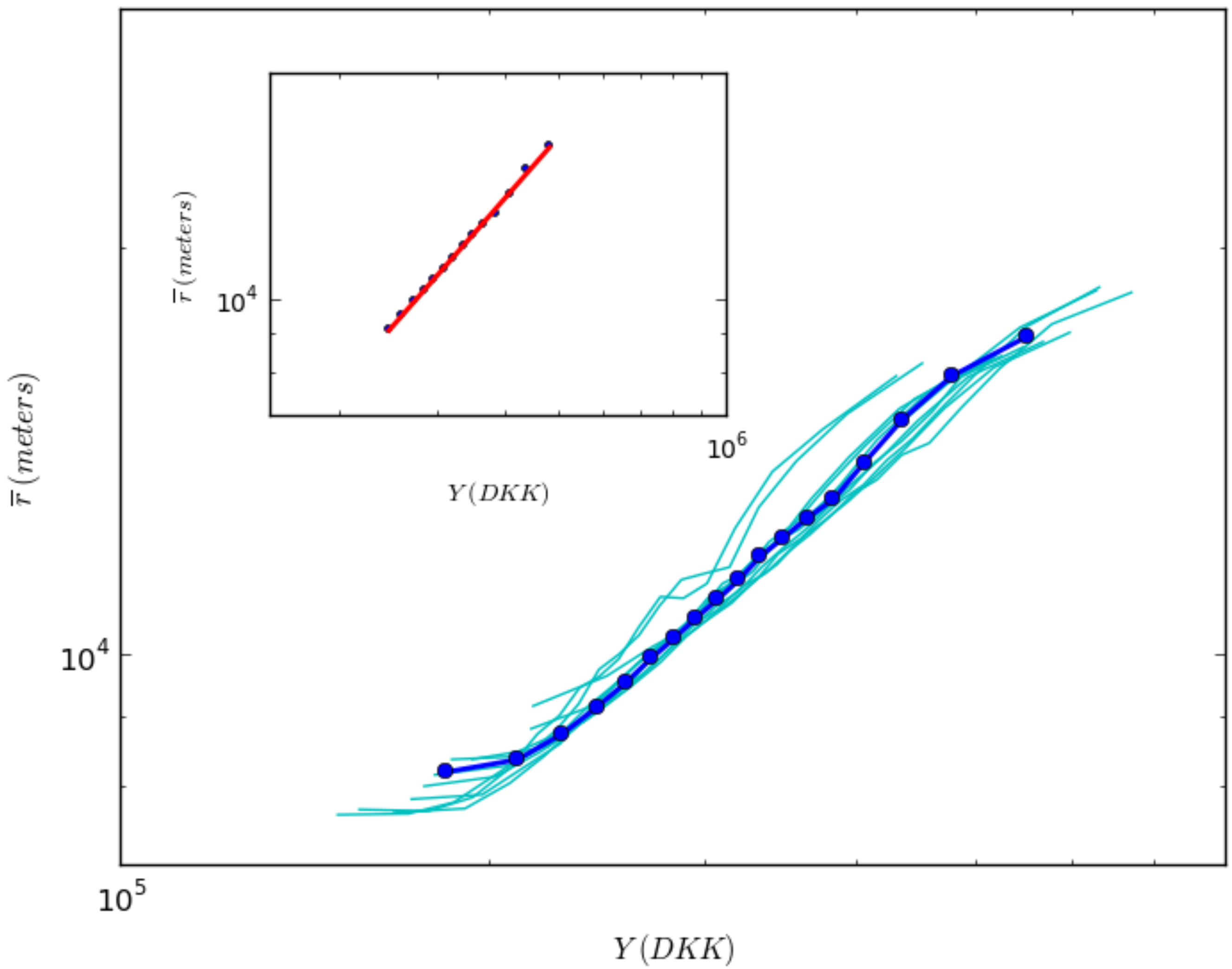
$$\beta \approx 0.5$$



Income and commuting distance: DK data

$$\overline{r} \sim Y^{\beta}$$

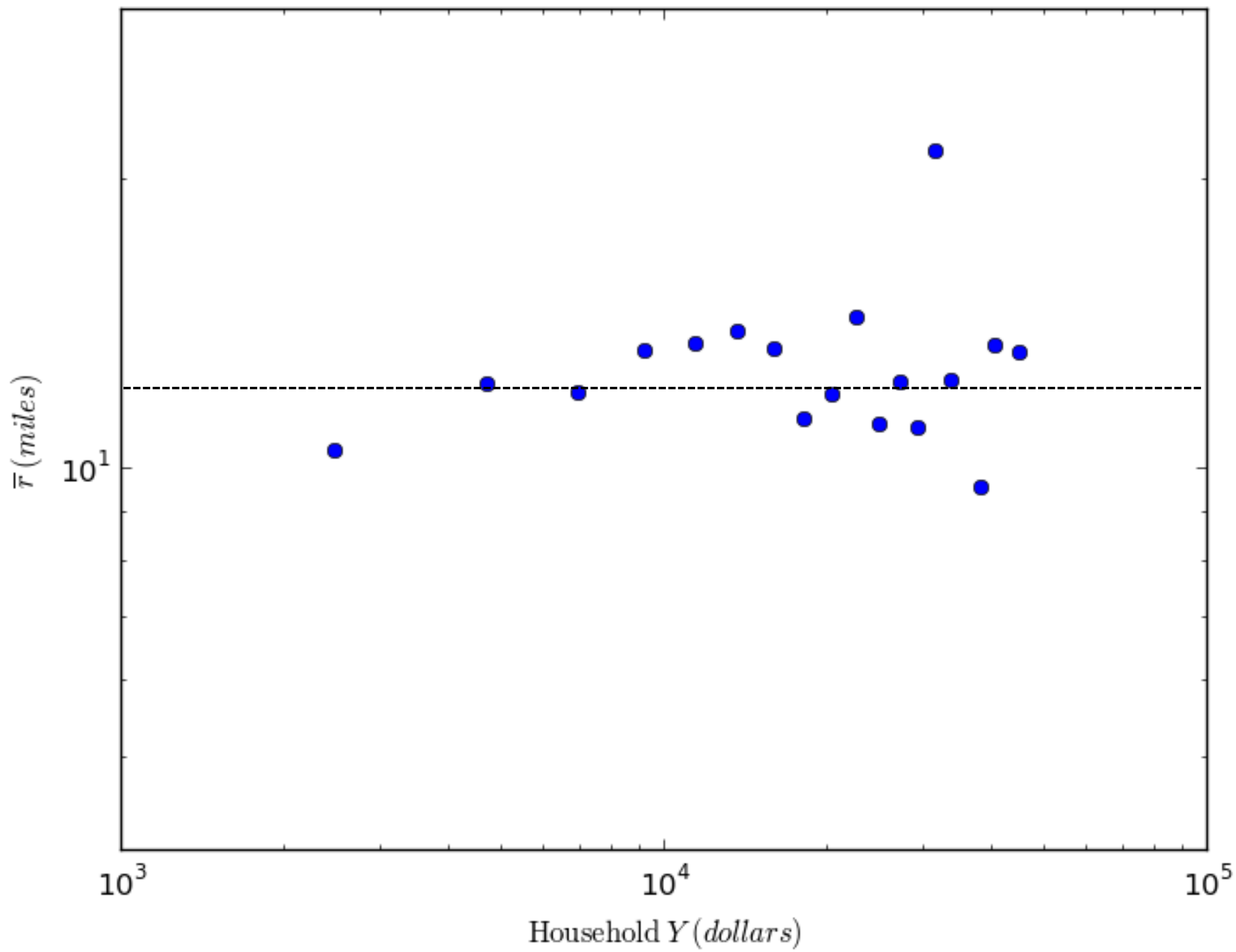
$$\beta \approx 0.8$$



Income and commuting distance: US data

$$\bar{r} \sim Y^{\beta}$$

$$\beta \approx 0.0$$

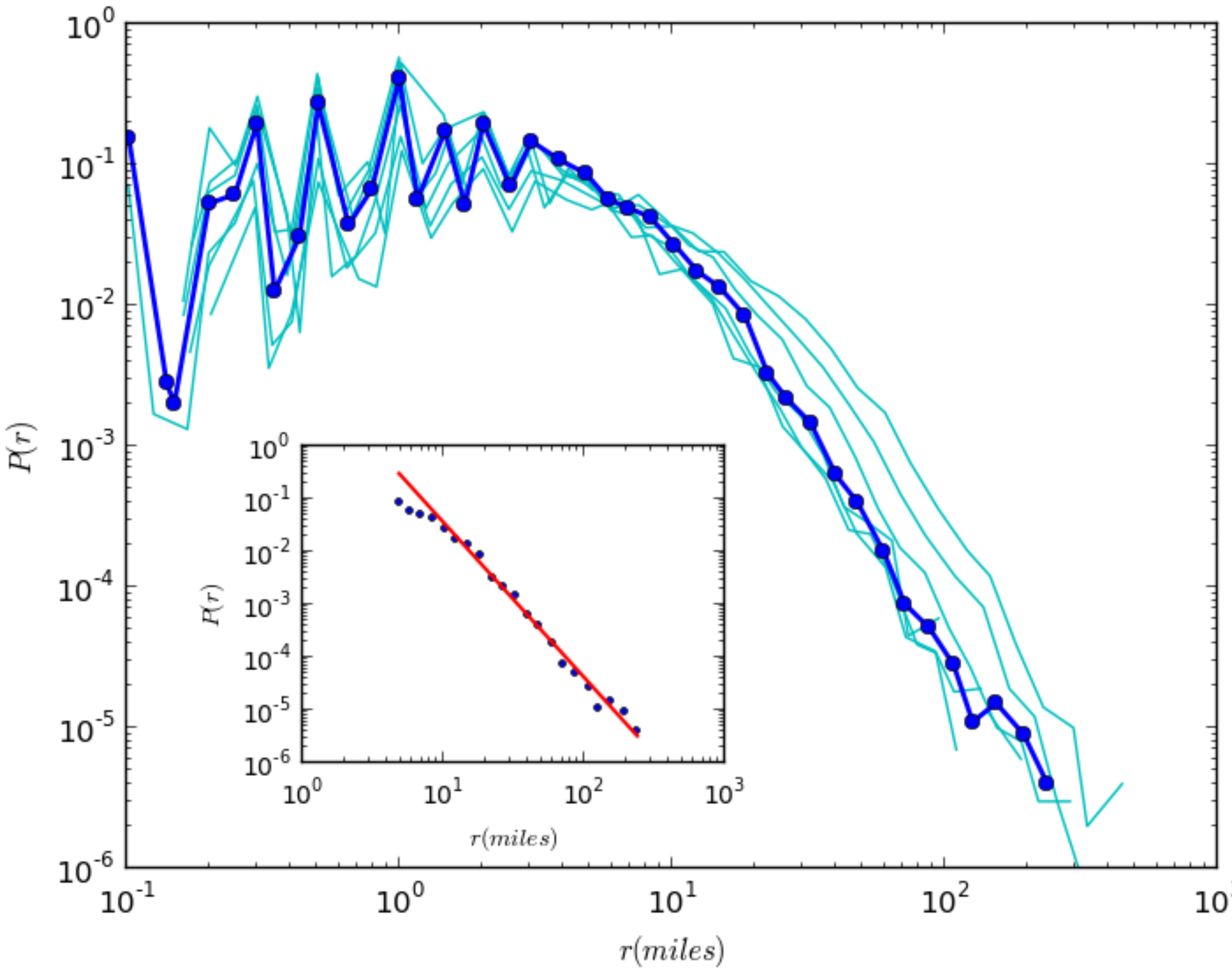


Distribution of commuting distance (UK)

- Distribution

$$P(r|Y) \sim r^{-\gamma}$$

$$\gamma \in [2.65, 3.0]$$

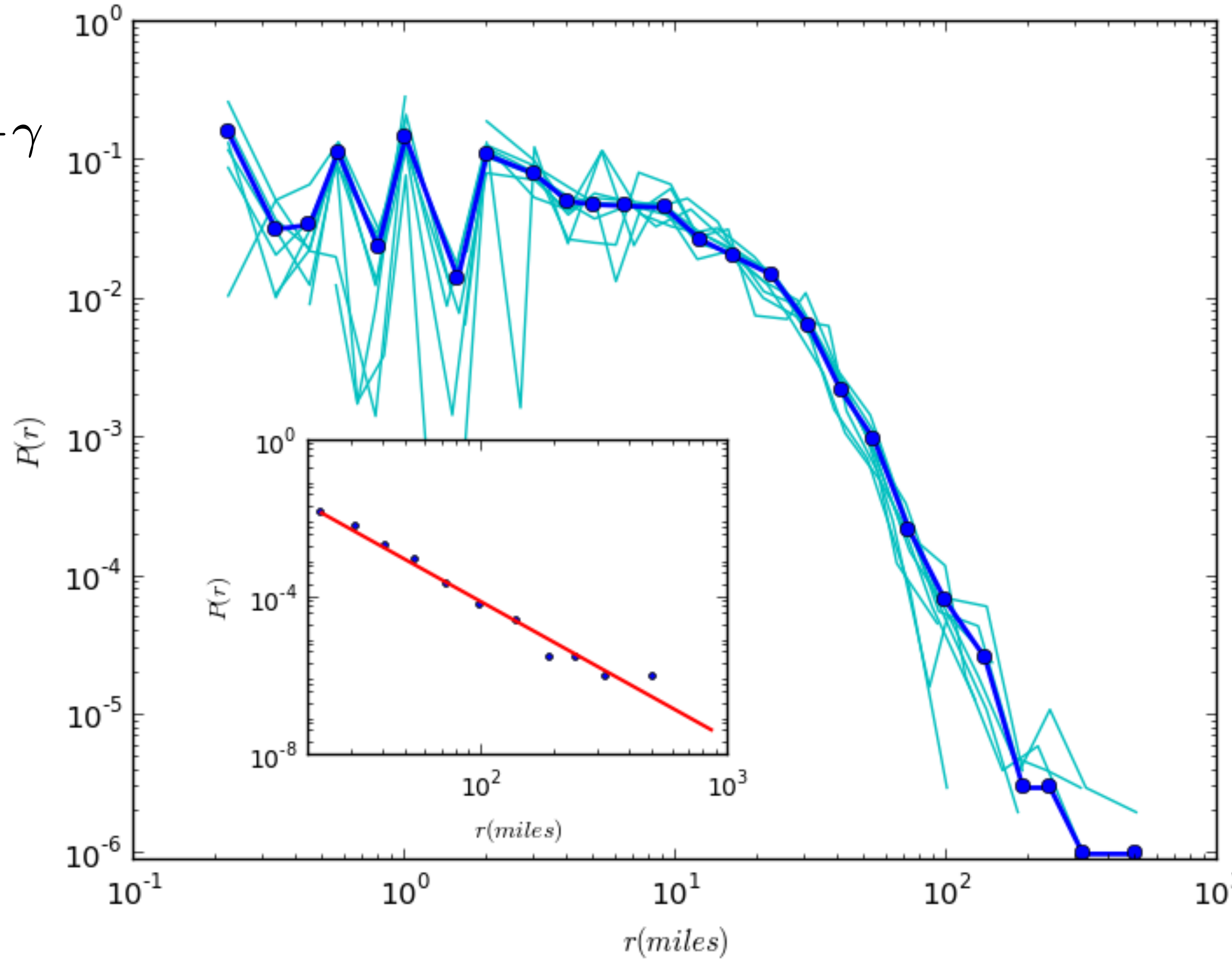


Distribution of commuting distance (US)

- Distribution

$$P(r|Y) \sim r^{-\gamma}$$

$$\gamma \in [2.8, 3.3]$$



Summary: empirical results

- Average commuting distance

$$\overline{r} \sim Y^{\beta}$$

where β depends on the country

- Distribution: broad law

$$P(r|Y) \sim r^{-\gamma}$$

where $\gamma \approx 3$

Classical model: McCall (1971)

- Optimal strategy (stopping problem)
  - Offers drawn from cumulative distribution F(x)
  - Waiting time cost c
  - Goal: maximize expected value v(w) for a given offer w in hand

$$v(w) = \langle \sum_{t=0}^{\infty} \beta^t y(t) | \text{offer} = w \rangle$$

$$where:$$

$$y(t) = \begin{cases} w & \text{if accepts offer} \\ -c & \text{if refuses offer} \end{cases}$$

Classical model: McCall (1971)

- Bellman equation

$$v(w) = \max\left\{\frac{w}{1-\beta}, -c + \beta \int v(w')\mathrm{d}F(w')\right\}$$

  - First term: accepts offer w
  - Second term: refuses; average over all possible offers w'

- Solution of the Bellman equation: optimal strategy with a reservation wage τ

  - If w< τ continue search
  - If w> τ accept offer

- Probability to accept an offer:

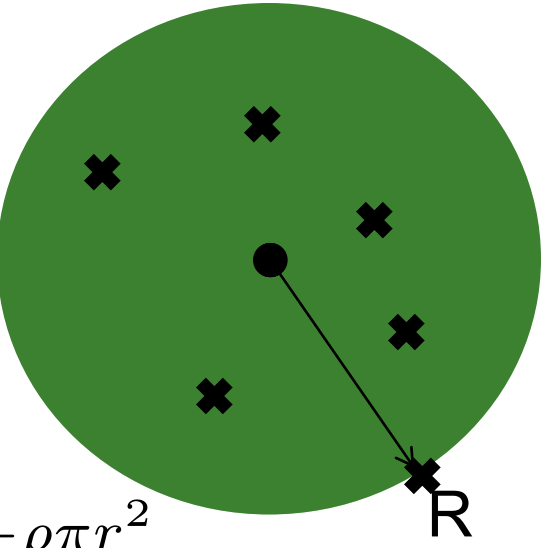
$$p = \int_{\tau}^{\infty} \mathrm{d}F(w') = 1 - F(\tau)$$

Adding space to the McCall model

- We assume that the offers are distributed uniformly in 2d space with density ρ. Individuals are starting at r=0 (home).

- Each time we encounter an offer, there is a probability p to accept it. The probability to accept the N^{th} offer is:

$$P(N) = (1-p)^{N-1}p$$



- Probability to be at distance r with N points (uniform distribution):

$$P(R=r|N) = \frac{2}{(N-1)!}\frac{1}{r}(\rho\pi r^2)^N \mathrm{e}^{-\rho\pi r^2}$$

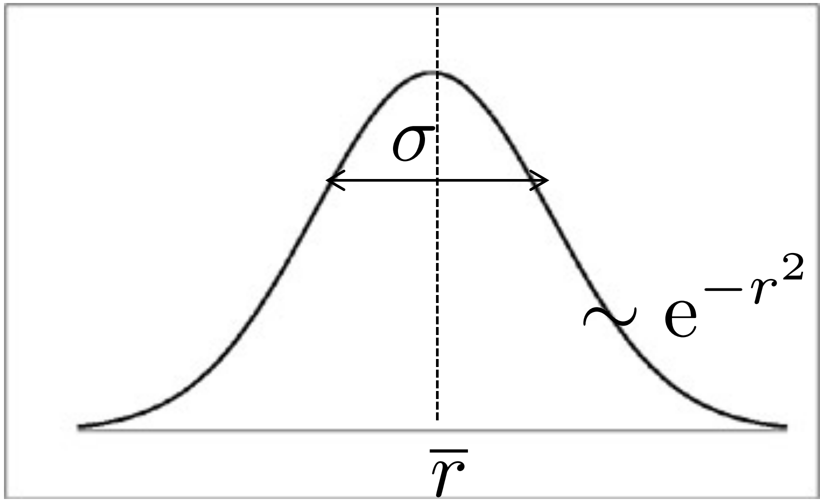
Adding space to the McCall model

- P(R=r|N) and P(N) combined together lead to:

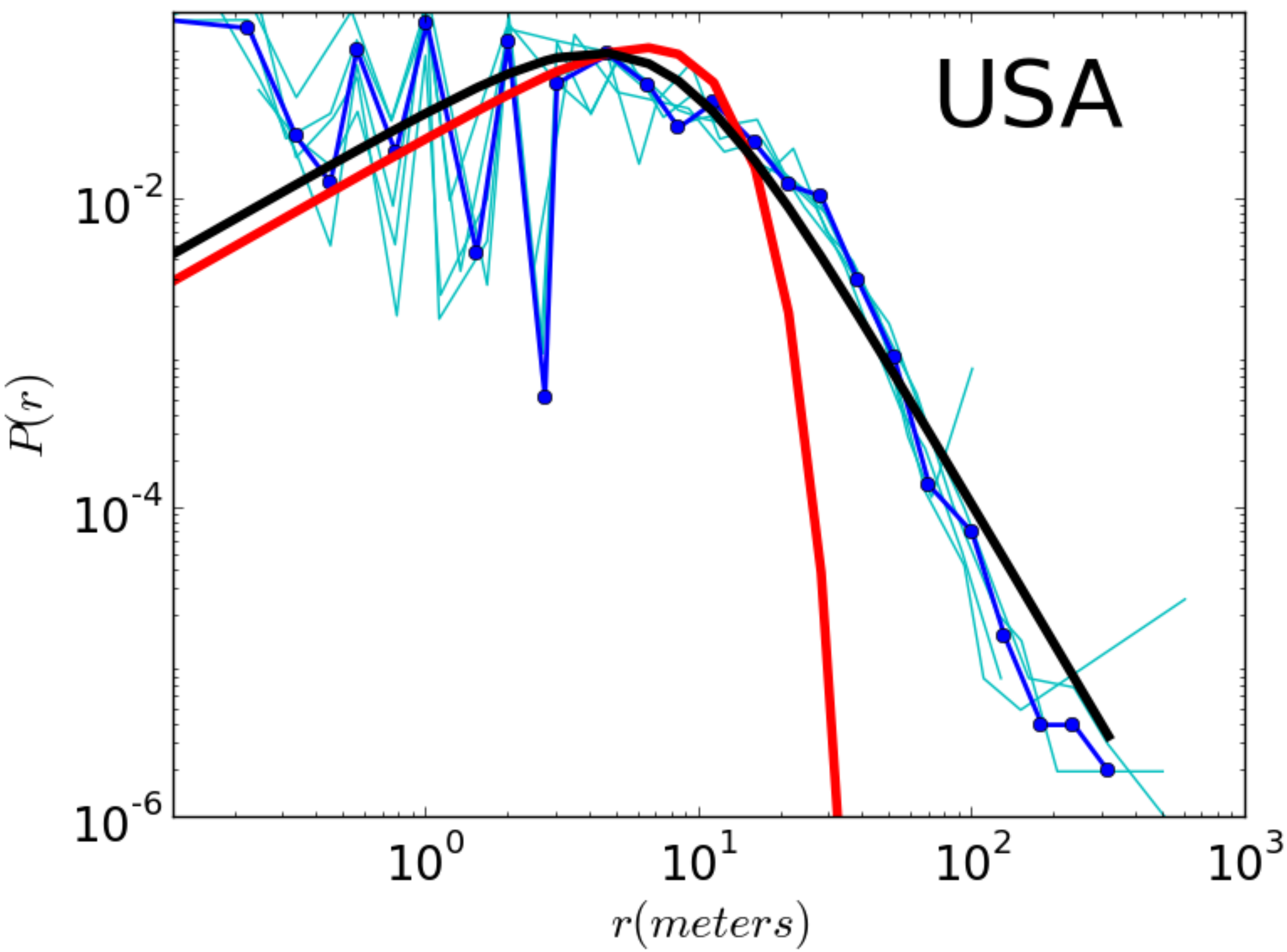
$$P(r|Y) = 2\pi\rho r p \mathrm{e}^{-p\rho\pi r^2}$$

-> Exponentially decreasing function !
Centered at $\bar{r} \sim 1/\sqrt{p\rho}$

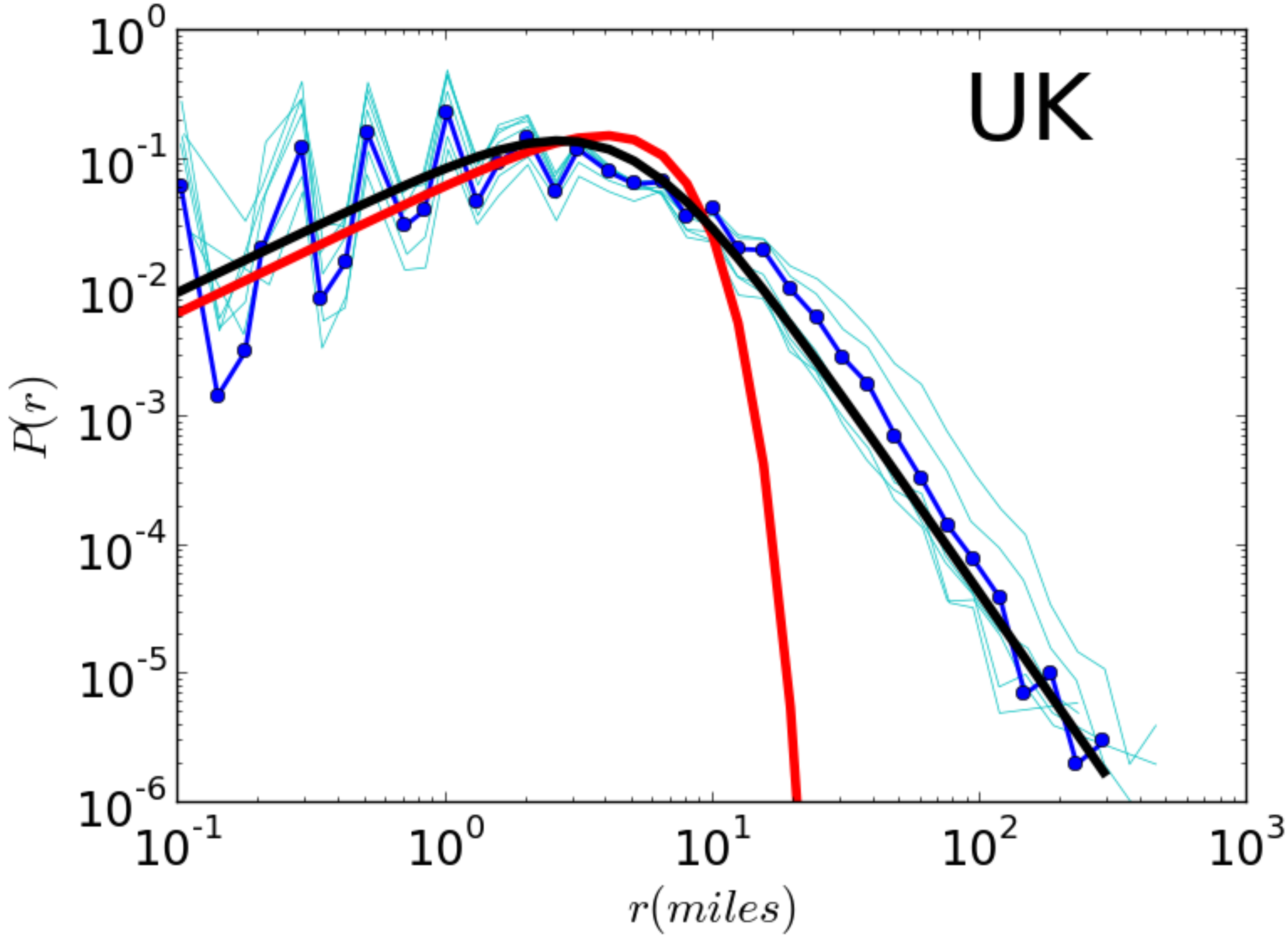
with finite width $\sigma \sim 1/\sqrt{p\rho}$

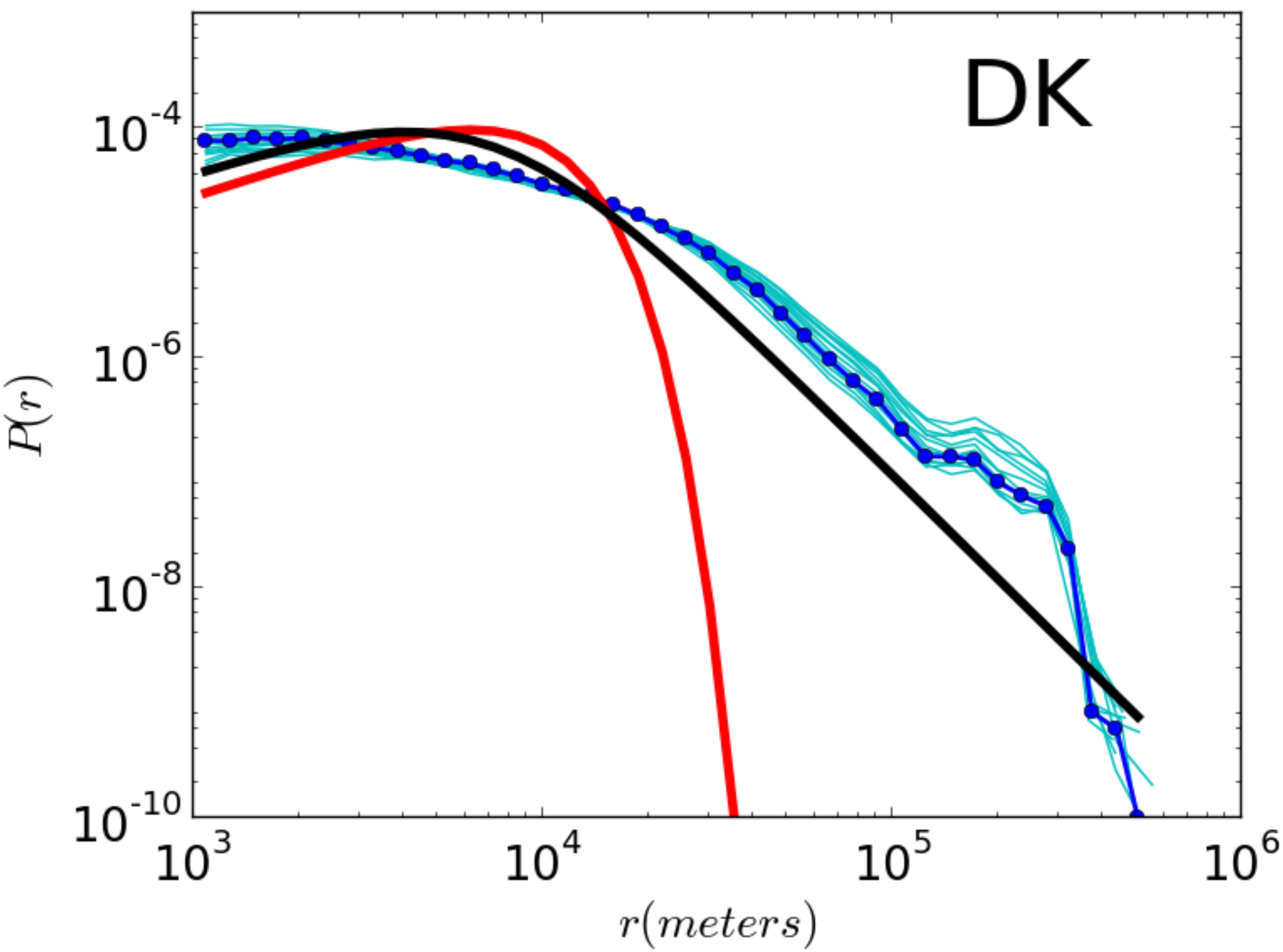


- The McCall model is not in agreement with data !
- Optimal strategy assumption ??



USA
10^{-2}
10^{-4}
10^{-6}
$P(r)$
10^{0}
10^{1}
10^{2}
10^{3}
$r(meters)$


UK
$P(r)$
$r(miles)$
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{1}
10^{2}
10^{3}


DK
10^{-4}
10^{-6}
10^{-8}
10^{-10}
$P(r)$
10^{3}
10^{4}
10^{5}
10^{6}
$r(meters)$

The "closest opportunity" model

- **Random reservation wage** distributed according to the same distribution ϕ (cumulative F) as offers
- Start from home, explore space with increasing r, accept the first job search such that the offer>reservation wage
- Density of job offers depends on the income: higher skills jobs less dense than lower skills jobs (skills <=> income)

$$\rho = \frac{\rho_0}{Y^\alpha}$$

where α characterizes the job market in the country

with: G. Carra, I. Mulalic, M. Fosgerau

The closest opportunity model

- Random reservation wage distributed according to ϕ
- We then have

$$P(r)\mathrm{d}r = \int \phi(\tau) P(x<\tau)^{\rho\pi r^2} \left[1 - P(x<\tau)^{2\pi r \rho \mathrm{d}r}\right] \mathrm{d}\tau$$

Probability to have value τ (annotating $\phi(\tau)$)

Probability that there are no interesting offer in the disk <r (annotating $P(x<\tau)^{\rho\pi r^2}$)

Probability that there is at least one interesting offer in the ring [r,r+dr] (annotating $\left[1 - P(x<\tau)^{2\pi r \rho \mathrm{d}r}\right]$)

The closest opportunity model

- The expression

$$P(r)\mathrm{d}r = \int \phi(\tau) P(x<\tau)^{\rho\pi r^2} \left[1 - P(x<\tau)^{2\pi r\rho \mathrm{d}r}\right] \mathrm{d}\tau$$

can be rewritten as

$$P(R=r) = -2\pi\rho r \int \phi(\tau) F(\tau)^{\rho\pi r^2} \log F(\tau) \mathrm{d}\tau$$

which is independent from F and is equal to

$$P(r) = \frac{2\pi\rho r}{(1+\rho\pi r^2)^2}$$

The closest opportunity model

- Consequences of

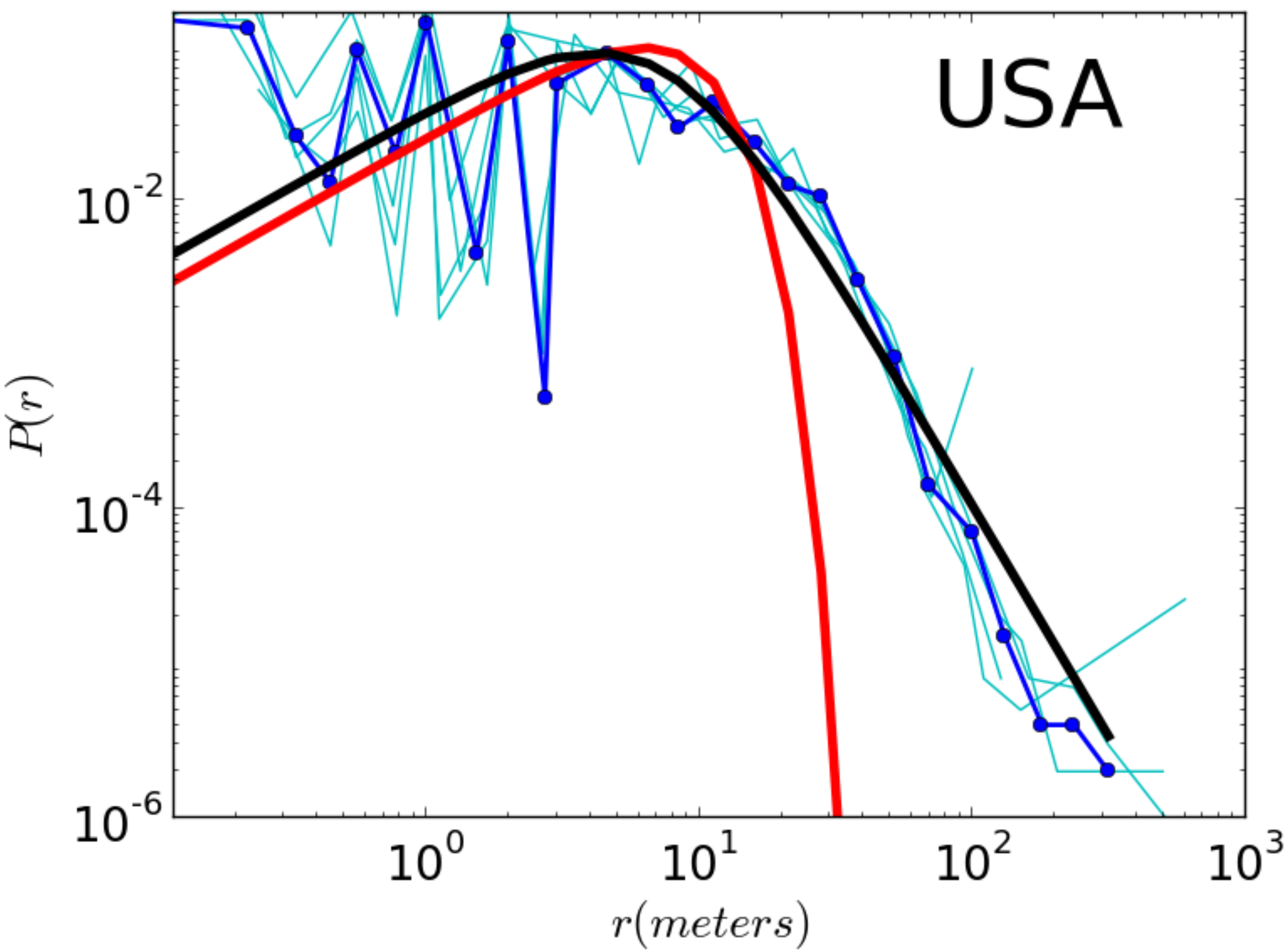
$$P(r) = \frac{2\pi\rho r}{(1+\rho\pi r^2)^2}$$

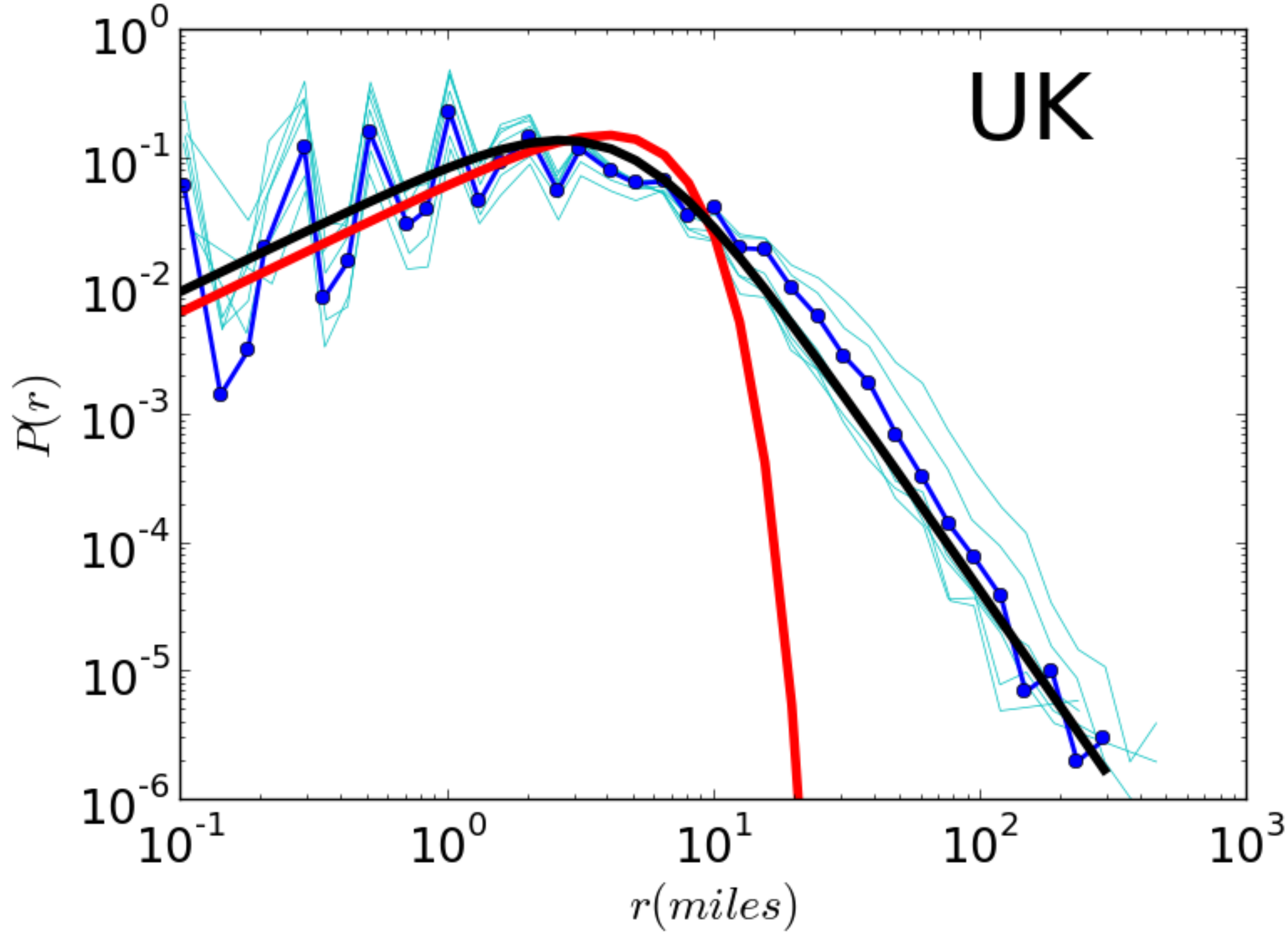
- Average distance: $\overline{r}(Y) = \frac{1}{2}\sqrt{\frac{\pi}{\rho_0}}Y^{\alpha/2}$
- Distribution decays as

$$P(r) \sim 1/r^3$$

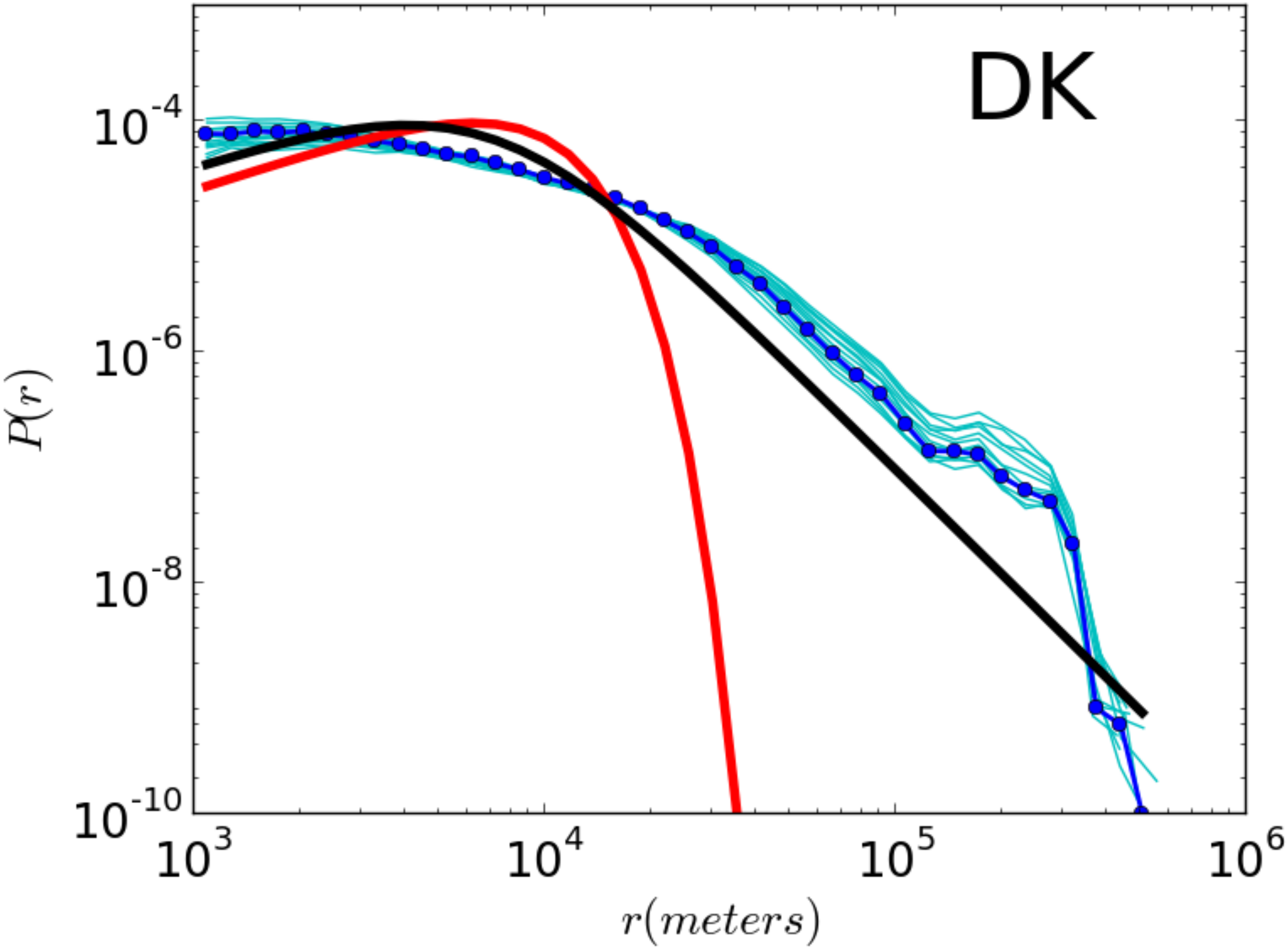
- Data collapse

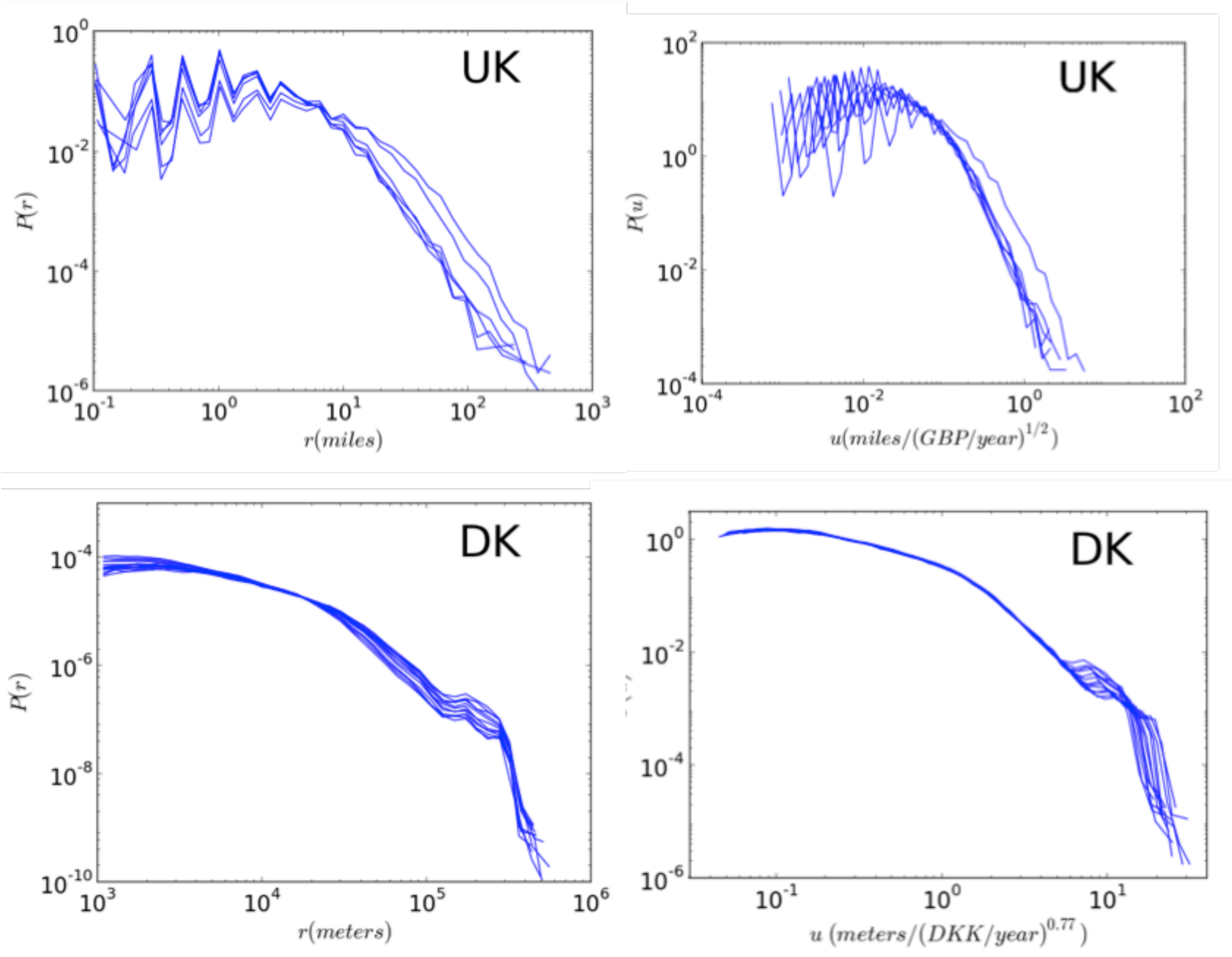
$$r \to r/Y^{\alpha/2} \qquad P(u) = \frac{2\pi\rho_0 u}{(1+\pi\rho_0 u^2)^2}$$


USA
10^{-2}
10^{-4}
10^{-6}
$P(r)$
10^{0}
10^{1}
10^{2}
10^{3}
$r(meters)$



UK
$P(r)$
$r(miles)$
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}


DK
10^-4
10^-6
10^-8
10^-10
P(r)
10^3
10^4
10^5
10^6
r(meters)



UK
P(r)
r(miles)
UK
P(u)
u(miles/(GBP/year)^{1/2})
DK
P(r)
r(meters)
DK
u (meters/(DKK/year)^{0.77})

Discussion

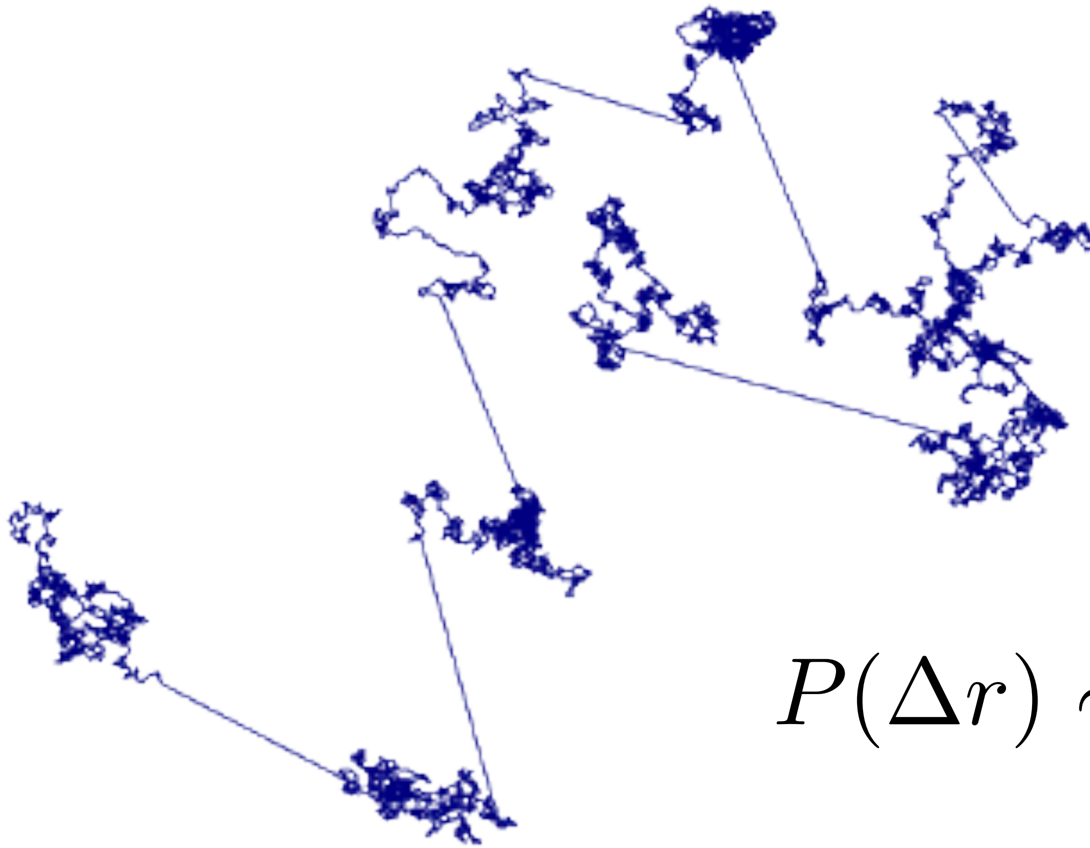
- Optimal strategy does not seem to be realistic for the job search problem
- Empirical results: broad tail for the commuting distance distribution
- A simple stochastic model
  - Predicts the 'universal' broad tail
  - Shows the importance of the relation between the density of jobs and skills (not significative for the US, strong for the UK and very strong for DK).

Mobility and statistical physics: a multilayer view

ENPC-2017

Human mobility: Levy flight ?

- Many small jumps and some rare long jumps



$$P(\Delta r) \sim \frac{1}{\Delta r^{\beta}}$$

Human mobility: Levy flight ?

- Empirical studies on the displacement distribution

Data Source	Trajectories	β	κ	Δr_0
Dollar Bills [8]	464K	1.59	∞	0
Mobile Phones [10]	100K	1.75	400 km	1.5 km
Mobile Phones [10]	206	1.75	80 km	1.5 km
Mobile Phones [11]	3M	1.55	100 km	0
Location Sharing [12]	220K	1.88	∞	0
GPS tracks [32]	101	[1.16,1.82]	∞	0
Location Sharing [20]	900K	1.50	∞	2.87 km
Location Sharing [20]	900K	4.67	∞	18.42 km
Taxis [18]	12K	0	4.29 km	-
Taxis [19]	7K	1.2	10 km	0.31 km
Mobile Phones [22]	7K	0	[2, 5.8] km	-
Travel Diaries [34]	230	1.05	50 km	0
Tweets [13]	13M	1.62	∞	0
Location Sharing [25]	521K	0	300 km	-
Taxis [21]	30K	0	[2, 4.6] km	-
Taxis [46]	1100	[0.50,1.17]	[4.5, 6.5] km	0

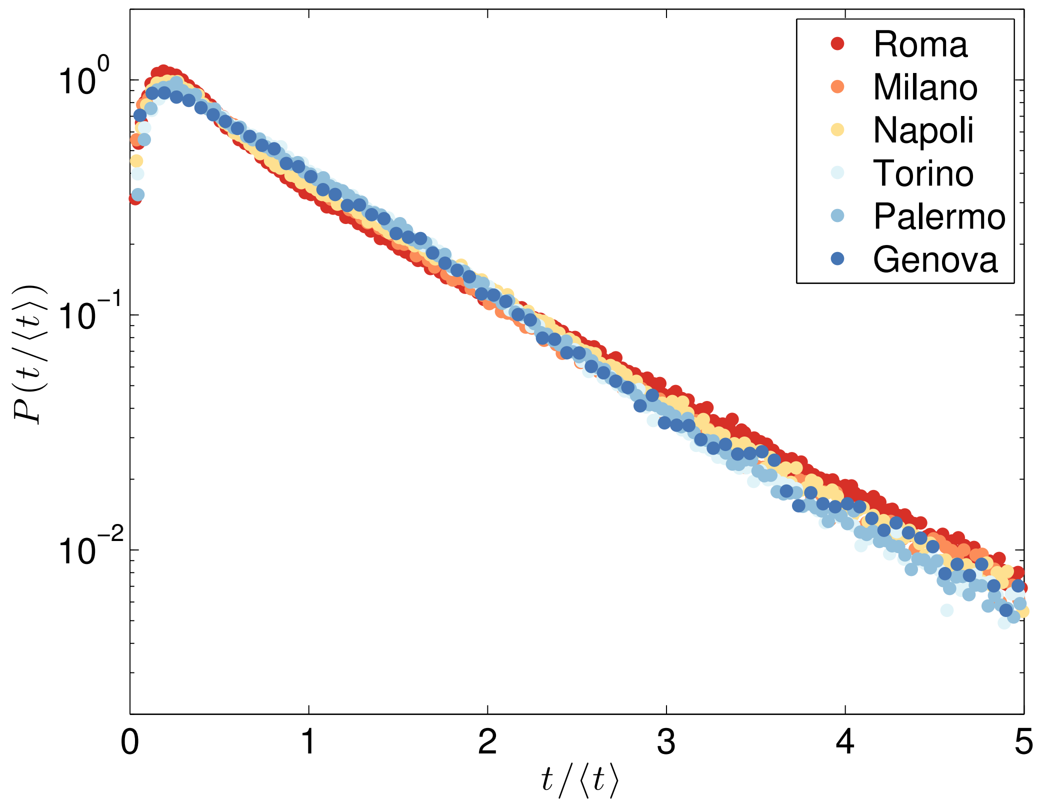
$$P(\Delta r) \sim \frac{1}{(\Delta r + \Delta r_0)^\beta} \mathrm{e}^{-\Delta r/\kappa}$$

- (Truncated) Levy flight ? Very empirical: model, mechanism ?

(Brockmann et al, Nature 2006; Gonzalez et al, Nature 2008)

Human mobility: empirical results

- Empirical results on the trip duration (GPS, 800,000 private cars in Italy)

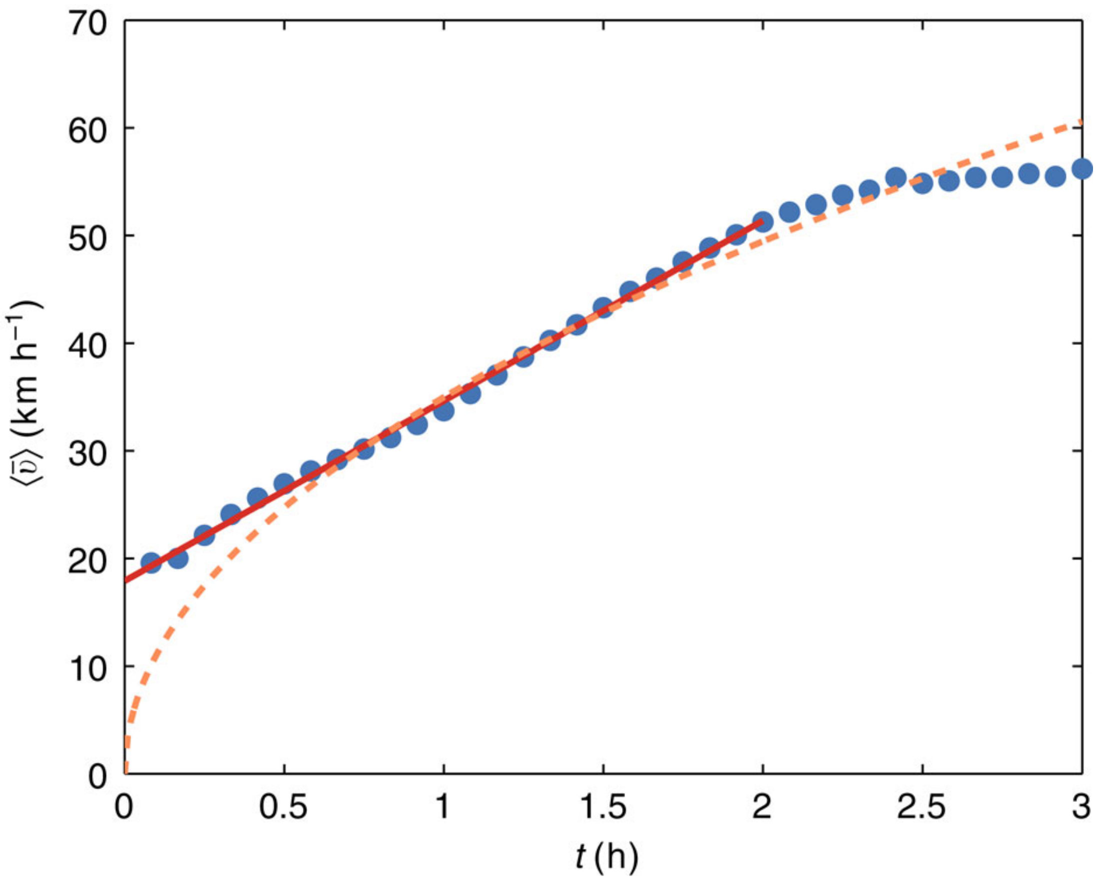


$$P(t) \sim \mathrm{e}^{-t/\langle t\rangle}$$

Gallotti, Bazzani, Rambaldi, MB, Nature Comms 2016

Human mobility: the further, the faster

- Average velocity with the trip duration



$$\langle \overline{v} \rangle \simeq v_0 + at$$

Cars :

$v_0 \simeq 17.9 km/h$

$a \simeq 16.7 km/h^2$

Public transport :

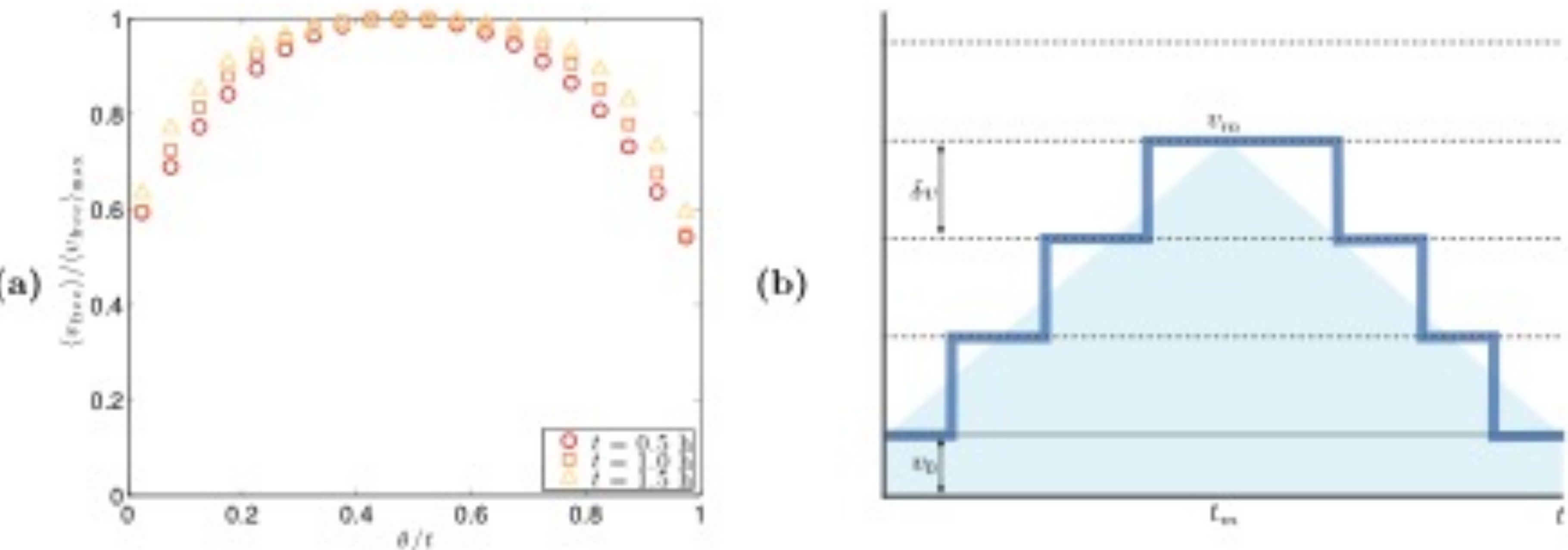
$v_0 \simeq 7.4 km/h$

$a \simeq 12 km/h^2$

Gallotti, Bazzani, Rambaldi, MB, Nature Comms 2016

Human mobility: a simple model

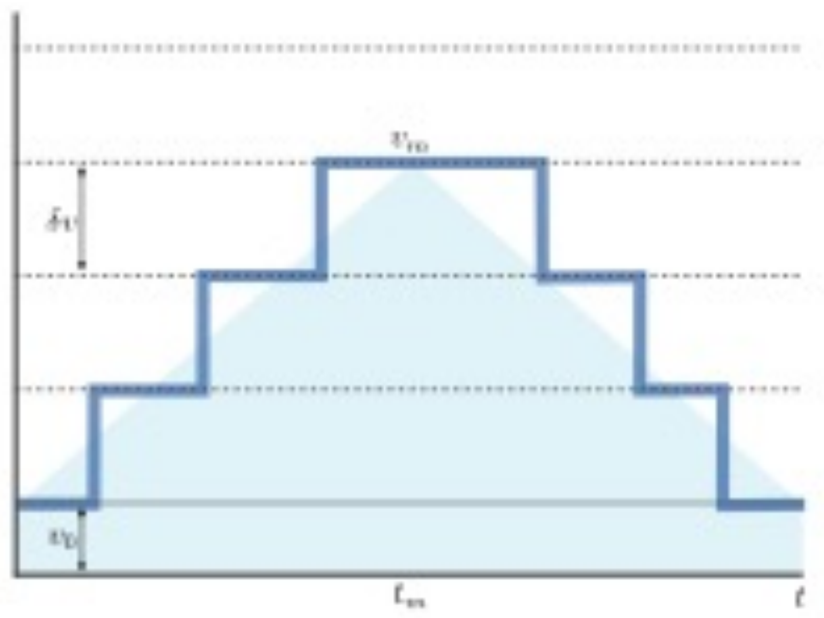
- Average trip: two phases (acceleration and decceleration).



Gallotti, Bazzani, Rambaldi, MB, Nature Comms 2016

Human mobility: a simple model

- Acceleration phase: Random acceleration kicks



$$v = v_0 + k(t)\delta v$$

$$v_m = v_0 + k(t_m)\delta v$$

with :

$$P(k) = \mathrm{e}^{-\lambda}\lambda^k/k!$$

$$\lambda = pt/2$$

Gallotti, Bazzani, Rambaldi, MB, Nature Comms 2016

Human mobility: a simple model

- This model predicts:

$$P(v|t) \sim \frac{e^{-pt/2+\frac{v-v_0}{\delta v'}\log(pt/2)}}{\Gamma(1+\frac{v-v_0}{\delta v'})}$$

- Determine p and δv by fitting all P(v|t) for all durations v (from t=5mns to 180mns)
- p=2 jumps/hour and δv=40km/h (consistent with 50-90-130)

Gallotti, Bazzani, Rambaldi, MB, Nature Comms 2016

Human mobility: recap

- Recap (all parameters determined):

$$P(t) \sim \mathrm{e}^{-t/\bar{t}},\ P(v|t)$$

- Which gives $P(\Delta r)$ for $\Delta r = vt$:

$$P(\Delta r) = \int P(v|t)P(t)\delta(\Delta r - vt)\mathrm{d}v\mathrm{d}t$$

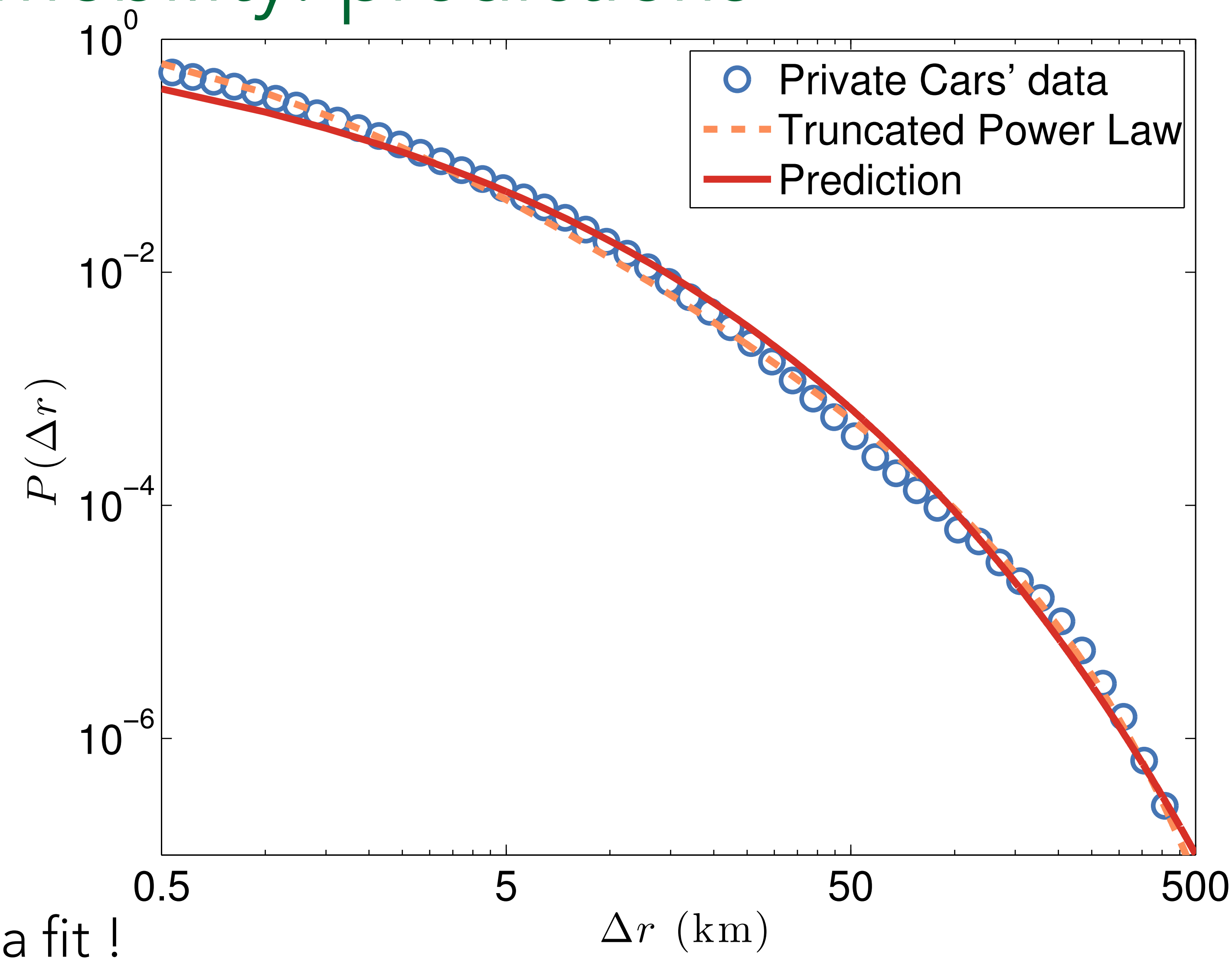
$$P(\Delta r) \sim \frac{1}{\Delta r^{\gamma}}\mathrm{e}^{-C\Delta r^{\delta}}$$

- with $\delta=1/2$ (γ depends on the parameters)

Gallotti et al, Nature Comms 2016

Human mobility: predictions

- Result:



- This is not a fit !
- This is not a Levy walk !

Gallotti et al, Nature Comms 2016

Summary and Perspectives

- The multilayer view allows:

  - to describe and understand important features due to the coupling of layers

  - to characterize them and their efficiency (new tools needed).

- Helpful for comparing systems, and testing and finding specific optimization strategies.

- Simple models allow to understand the essential mechanisms of mobility