HAL
open science

Modeling the relation between income and commuting distance

Giulia Carra, Ismir Mulalic, Mogens Fosgerau, Marc Barthelemy

To cite this version:

Giulia Carra, Ismir Mulalic, Mogens Fosgerau, Marc Barthelemy. Modeling the relation between income and commuting distance. Journal of the Royal Society Interface, 2017, 13 (119), pp.20160306. 10.1098/rsif.2016.0306 . cea-01568435

HAL Id: cea-01568435

https://cea.hal.science/cea-01568435v1

Submitted on 25 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1602.01578v2 [physics.soc-ph] 19 May 2016

Modeling the relation between income and commuting distance

Giulia Carra
Institut de Physique Théorique, CEA, CNRS-URA 2306, F-91191, Gif-sur-Yvette, France

Ismir Mulalic
DTU Denmark and
Kraks Fond - Institute for Urban Economic Research

Mogens Fosgerau
DTU Denmark

Marc Barthelemy*
Institut de Physique Théorique, CEA, CNRS-URA 2306, F-91191, Gif-sur-Yvette, France and
Centre d'Analyse et de Mathématique Sociales, (CNRS/EHESS) 190-198,
avenue de France, 75244 Paris Cedex 13, France

We discuss the distribution of commuting distances and its relation to income. Using data from Denmark, the UK, and the US, we show that the commuting distance is (i) broadly distributed with a slow decaying tail that can be fitted by a power law with exponent $\gamma \approx 3$ and (ii) an average growing slowly as a power law with an exponent less than one that depends on the country considered. The classical theory for job search is based on the idea that workers evaluate the wage of potential jobs as they arrive sequentially through time, and extending this model with space, we obtain predictions that are strongly contradicted by our empirical findings. We propose an alternative model that is based on the idea that workers evaluate potential jobs based on a quality aspect and that workers search for jobs sequentially across space. We also assume that the density of potential jobs depends on the skills of the worker and decreases with the wage. The predicted distribution of commuting distances decays as $1/r^3$ and is independent of the distribution of the quality of jobs. We find our alternative model to be in agreement with our data. This type of approach opens new perspectives for the modeling of mobility.

Keywords: Statistical Physics — Urban economics — Mobility — Job-search — Modeling

INTRODUCTION

Cities are growing and the majority of individuals in the world now live in urban areas [1]. Understanding what governs the evolution and the organization of urban systems is thus of primary interest for policy makers and planners. The availability of large scale data about almost all aspects of cities has opened the possibility of a new interdisciplinary science of cities with solid foundations [2]. In particular, understanding mobility patterns is a central problem in this field and is related to the labor market, a fundamental area of interest in economics, where the choice of work and residential locations determines the commuting. We focus here on a part of this area, namely the job search process which has a direct impact on the spatial distribution of commuting trips. The seminal contributions on job search theory in economics [3–5] rely on the central assumption that individuals choose among different job offers that arrive sequentially in time, by maximizing their expected discounted net wage, while waiting to accept a job offer is costly. These are clearly very strong assumptions that should be tested against empirical data.

Surprisingly, the standard model of job search [4] does not integrate space (some labour market studies do take into account, see for example [6]). We introduce here a spatial component in this model and derive the consequences for the distribution of the commuting distance. In particular, we show that the basic McCall model [4] does not explain some fundamental statistical features observed in empirical data. We therefore propose a new stochastic model that does not rely on the assumption of optimal control in search through time, but instead on the idea that workers search through space, accepting an offer if it has a certain level of 'quality'. The quality of a job is random and unobserved by the researcher and it may integrate any number of quality aspects specific to each individual. We find excellent agreement between this new model and empirical data for Denmark, the UK, and the US.

Beyond the prediction of the distribution of commuting distances and their relation with income, our model provides a search-based microfoundation for models of spatial patterns that can be found in the mobility literature [7]. More generally, we question here the relevance of optimal control theory as the main framework to explain mobility and the behavior of living organisms. Optimal control theory is a mathematical optimization method used to find the policies that optimize the outcome of a given process. This method has been applied to many

different problems in areas such as biology, economics and finance, ecology, and management [8–11]. Here, we propose an alternative framework to study human or animal behavior, closer to theories developed about foraging [12] and in which actions are taken not on the basis of an optimal strategy but on the first opportunity that is judged to be good enough.

In the first part of this paper, we present an empirical analysis of the distribution of commuting distances for Denmark, the UK, and the US, exploring how the average commuting distance scales with income. In a second theoretical section, we derive the probability distribution for the commuting distance from the spatial extension of the standard job search model. We compare this theoretical prediction with our empirical results, and show that the standard theoretical framework is not in agreement with data. We then propose a new stochastic model which does not rely on the optimal strategy assumption, and where workers evaluate potential jobs sequentially across space and based on a quality aspect. We then show that this new model is in excellent agreement with our data.

EMPIRICAL RESULTS

In this section, we investigate the distribution of commuting distances and its relation to individual income using datasets for three different countries: Denmark, the UK [13], and the US [14]. These datasets are produced by national agencies and national household surveys (see Materials and Methods for details) and record the commuting distance and the income range at the individual level. The datasets cover the whole of each country and take into account all transportation modes. For the UK the data is for the years 2002-2012, for the US three different years are available (1995, 2001, 2009), and for Denmark we have access to 10 years (2001-2010).

The average commuting distance

We first focus on the simplest quantity, the average commuting distance, and how it varies with income. The results for the three countries studied here are shown in Fig. 1 (left column). The basic equilibrium models of urban economics [15–17] predict, within a single city, that workers with higher incomes will have longer commuting distances. This prediction is confirmed for Denmark and the UK, while no particular trend can be detected for the US.

For Denmark, we observe an increasing range and a saturation at large income values, while for the UK we observe a plateau at low income values. In the range where the increase is observed we can fit the data by a power law of the form

$$\overline{r}(Y) \sim Y^{\beta} \tag{1}$$

where Y is the individual income and where the exponent β depends on the country considered. For the US, the fit gives an exponent $\beta \approx 0$ indicating that there is no clear trend. For the UK, the plateau around the commuting distance value $\overline{r} \approx 5$ miles occurs in the low income range $[10^2, 10^4]$ (GBP/year). The fit on UK data for incomes higher than $5,000$ GBP (for all modes and all years) gives an exponent value $\beta \approx 0.5$ (in the range $[0.53, 0.66]$ when considering different years). In contrast, we observe for the Danish data a strong dependence with a large exponent of order 0.8 for yearly incomes larger than $250,000$ DKK and smaller than $500,000$ DKK (for lower incomes we observe a small plateau). Depending on the year considered, the exponent β varies in this case in the range $[0.61, 0.88]$.

The distribution of commuting distance

We now consider the full distribution of the commuting distance, shown in Fig. 1 (right column) for different incomes for Denmark, the UK, and the US. There are two important facts that we can extract from these empirical observations. First, for all datasets studied here, the distribution is broad. This means that the variation range of commuting distances is extremely large. Indeed, we observe that with a non-negligible probability, individuals in Denmark, UK, and US are commuting on distances of the order of a few hundred kilometers. Second, the shape of the distribution and the large distance behavior are remarkably similar among the different countries we have studied here. These non-trivial features are very important as they provide an opportunity to test for any model that aims to describe spatial commuting patterns.

THEORETICAL MODELING

The three datasets observed here display a slow increase of the average commuting distance with income and, more importantly, a slowly decaying tail for large distances. We would like to understand these two characteristics theoretically. We begin with a discussion of the standard job search model of economics [3–5], and compare its predictions with our empirical observations. This will lead us to propose another model, the 'closest opportunity' model with predictions that are in much better agreement with the data at hand.

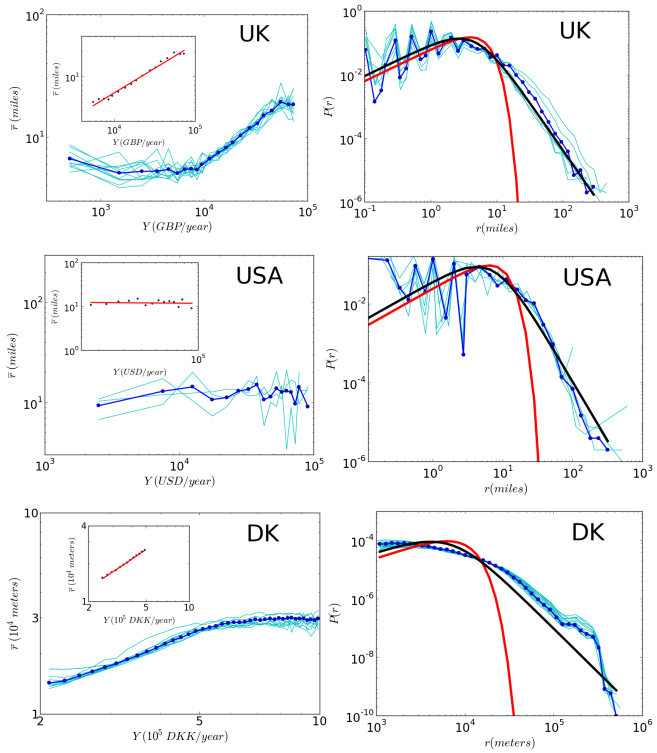

FIG. 1: **Left column: Average commuting distance versus income for different years.** In dark blue, the commuting distance is averaged over all years. (Top) UK data. This loglog plot displays a plateau for small values of income followed by a regime, when fitted by a power law (see inset), gives an exponent $\beta \approx 0.5$ ($[0.53, 0.66]$). In the inset the average commuting distance is averaged over all years and the power law fit gives an exponent $\beta = 0.58$. (Middle) US data. In this loglog plot we do not observe an income dependence. Indeed, a power law fit gives an exponent $\beta \approx 0$. (Bottom) Danish data. The power law fit on the commuting distance averaged over all years (in the inset) gives an exponent $\beta = 0.77$. **Right column: Commuting distance distribution for different income classes.** The probability distribution is shown for different income classes. In dark blue we show the distribution for a particular value of the income for which fits have been performed. In red, we show the one parameter fit with the analytical function predicted by the extended McCall model (Eq. 9), and in blue, the one parameter fit with the analytical function predicted by the closest opportunity model (Eq. 16). (Top) UK data (averaged over all available years). (Middle) US data (averaged over all available years). (Bottom) Danish data (all years give the same result and we choose here to show the year 2008). In all cases we observe that the tail predicted by the extended McCall model (Eq. 9) decays too quickly and cannot fit the data for long distances. In contrast, the closest opportunity model is in excellent agreement with empirical observations.

The spatial optimal job search model

Optimal control theory is a well known mathematical optimization method used to find policies that maximize the benefit of a given process. An example of its application is the stopping problem [18], where one has to choose the optimal time to take an action based on successive observations of a random variable. Optimal control theory has been applied in many different areas [19–23], and to the job search problem in economics [3–5]. As a starting point, we will here consider the important McCall model [4] that has been used in many different forms and variants. We will study the implications of the McCall model for the spatial distribution of distances between residences and jobs depending on the income.

We begin by describing the McCall model in its simplest version. The job search process is sequential in time. A worker who is unemployed at time 0 reviews at every time step a random wage offer w drawn from a distribution with density f (and cumulative F). At each time step, the worker can either accept the current job offer and keep it forever, or she can pay a waiting cost c to discard the offer and wait for the next offer. The worker's income y_t at time t will thus be $y_t = w$ if she accepts the offer or $y_t = -c$ if she refuses it. The actual value of her total returns is the discounted sum of her future payoffs

$$\sum_{t=0}^{\infty} \mu^t y_t \ ,$$

where the discount factor $\mu < 1$ takes into account that the value of a given amount of money is higher the earlier it is received. In this model, with an offer w at hand, the worker maximizes the expected value of her total return $v(w)$

$$v(w) = \langle \sum_{t=0}^{\infty} \mu^t y_t \rangle \ , \tag{2}$$

where the brackets denote the average over the offer distribution. The classical way to solve this problem is to write the Bellman equation for this stopping process which reads [24]

$$v(w) = \max \left\{ \frac{w}{1-\mu}, -c + \mu \int v(w') f(w') \mathrm{d}w' \right\} \ . \tag{3}$$

This equation has a simple interpretation. The value of the current offer $v(w)$ is the maximum of two terms: the first term is the total return if the current job offer is accepted, and the second term is the expected value of rejecting the current offer and waiting for the next. In the latter case, the worker pays the waiting cost c and evaluates the expectation of the value $v(w')$ of the next random offer w'. The optimal strategy that solves this equation is to accept the current offer if it is larger than

a *reservation wage* τ and to refuse it if it is lower. The reservation wage satisfies the equation

$$\frac{\tau}{1-\mu} = -c + \frac{\mu}{1-\mu}\left[\tau F(\tau) + \int_\tau^\infty w' f(w')\mathrm{d}w'\right] , \quad (4)$$

so that the worker is indifferent between accepting the job for which $w = \tau$ or waiting for another offer. By solving this equation, we obtain a function τ that depends on the offer distribution. The probability p of accepting an offer is then

$$p = \int_\tau^\infty f(w)\mathrm{d}w , \quad (5)$$

and the number of trials N before accepting a job offer thus follows a geometric distribution

$$P(N) = (1-p)^{N-1}p . \quad (6)$$

Space is absent in the McCall model and we will now extend it in the simplest possible way. We assume now that the individual reviews the job offers sequentially in the order of increasing distance from home. The first offer reviewed is the closest to her residence, the second one is the second closest and the nth time step corresponds to the nth closest job to the seeker residence. Each random wage offer w is still drawn from a distribution with density f (and cumulative F)and thus the probability that the individual accepts an offer is still given by Eq. (5). This means that the worker, starting from home, will examine the offer and will choose the first one that is above her reservation wage. We will also assume that jobs are uniformly distributed in space with density ρ. If a worker has accepted the Nth offer, the probability that she has moved a distance r from its residence is given by a classical result for the Nth nearest neighbors in dimension $d = 2$ for uniformly distributed points [25]

$$P(R = r|N) = \frac{2}{(N-1)!}\frac{1}{r}(\rho\pi r^2)^N \mathrm{e}^{-\rho\pi r^2} . \quad (7)$$

The distribution of the commuting distance R is then given by

$$P(R = r) = \sum_{N\geq 1} P(r|N)P(N) \quad (8)$$

and since the distribution of N is geometric (Eq. (6)), we obtain

$$P(R = r) = 2p\rho\pi r\mathrm{e}^{-p\rho\pi r^2} . \quad (9)$$

This distribution decreases as a gaussian over a scale of order $\sim 1/\sqrt{\rho p}$ where $1/\sqrt{\rho}$ corresponds to a typical interdistance between different offers (τ and therefore p depend on the income Y and so does this distance too). We also note that the average commuting distance decreases if the spatial density of opportunities ρ increases. A decrease in the number of job openings during economic downturns then leads to increasing commuting distances.

To test the consistency of these result with empirical data, we fit in Fig. 1 (right column) empirical data using the prediction Eq. (9) of the extended McCall model. We observe that the best (one parameter) fit is reasonable for the short distance regime but is unable to reproduce the slow decay observed for large distances. In addition, we have also considered another generalization of the McCall model with transport costs, and showed that it also cannot reproduce a slow decaying tail such as a power law (see the general argument presented in the Materials and Methods section). It thus seems that the McCall model is not consistent with our data. We therefore seek an alternative model that does predict the empirical findings just outlined. We will propose such a model in the next section and compare its predictions with data.

The closest opportunity model

In this new model proposed here, we change three important assumptions of the McCall model. First, we assume that workers evaluate offers sequentially across space, whereas in the original McCall model the evaluation was performed through time. Second, jobs are chosen based on some 'quality' aspect that could take into account many factors and not only on the wage (see for instance [26, 27]). Finally, we change the framework used to study human behavior, and the reservation wage of the McCall model, which is the result of an optimal strategy, is replaced by a *reservation quality* representing the minimal job quality that meets worker expectations.

We still consider the problem of a worker who looks for a job starting from her residence (that we assume to be located at $r = 0$). Job offers are uniformly distributed across space with density ρ. The density of jobs ρ relevant for the worker depends on the income level Y and we assume that it is simply

$$\rho = \frac{\rho_0}{Y^\alpha} \quad (10)$$

such that higher income jobs are less dense than lower income jobs. The exponent α depends on the country under consideration and reflects many exogenous factors concerning job offers at a certain income level [26, 27]. We remark that the job density ρ is the only parameter that discerns here different types of workers. We also note that the framework introduced here for the income allows for many generalizations to other quantities such as the skill level for example.

The McCall model assumes that jobs are primarily characterized by the wage they offer. We depart from this and assume instead that each job is characterized by a random 'quality' X that encodes many factors. The job

quality is distributed according to f (with corresponding cumulative distribution F) and job qualities are independent. We further assume that a given worker has a *reservation quality* value τ (in the same spirit as the reservation wage), and she will keep expanding her search radius until this threshold is met. We denote by R the commuting distance and its cumulative thus reads

$$P(R \leq r|\tau) = P(X_{[0,r]} \geq \tau) = 1 - F(\tau)^{\rho\pi r^2}. \quad (11)$$

We now take into account that workers have different search costs and different expectations for a future job, which leads them to have different reservation qualities. We consider the reservation quality as random, distributed according to a density $g(\tau)$, and obtain the cumulative distribution of commute distances

$$P(R \leq r) = \int g(\tau) P(R \leq r|\tau) \mathrm{d}\tau, \quad (12)$$

with corresponding density

$$\begin{aligned} P(R = r) &= \frac{dP(R \leq r)}{dr} \\ &= -2\rho\pi r \int g(\tau) F(\tau)^{\rho\pi r^2} \log F(\tau) \mathrm{d}\tau. \quad (13) \end{aligned}$$

The first term in this integral is the probability that a worker has reservation quality τ, the second term is the probability that all offers are below τ in the disk of radius r, and the last term (the logarithm) corresponds to the probability that at least one offer is above τ in the circular band $[r, r + \mathrm{d}r]$ (see Fig. 2 for a simple illustration of this process). A simple and natural assumption for the distribution of the reservation quality τ is that it is the same as the distribution of job quality F: $g(\tau) \equiv f(\tau)$. Then Eq. (13) simplifies in a remarkable way as follows

$$\begin{aligned} P(R = r) &= -2\rho\pi r \int f(\tau) F(\tau)^{\rho\pi r^2} \log(F(\tau)) \mathrm{d}\tau \\ &= -2\rho\pi r \int_0^1 x^{\rho\pi r^2} \log x \, \mathrm{d}x \\ &= \frac{2\rho\pi r}{(1 + \rho\pi r^2)^2}. \quad (14) \end{aligned}$$

Under these assumptions, the distribution of commuting distances does not depend on the distribution of job quality, an effect that was already observed in the specific case discussed in [7], and the model proposed here can then be considered as a microfoundation for this type of process. This also means that we may generalize the interpretation of the model: we may allow the distribution of job quality to be specific to each worker, since this has no consequence for the distribution of commuting distances.

In contrast to the McCall job-search model of the previous section that displayed a rapid gaussian decaying tail, we observe here that the distribution is slowly decaying as $P(R = r) \sim r^{-3}$ for large r. The average commuting distance is easily computed within the closest opportunity model and we find

FIG. 2: Illustration of the argument leading to Eq. (13).

$$\overline{r} = \frac{1}{2}\sqrt{\frac{\pi}{\rho}}. \quad (15)$$

Replacing ρ by ρ_0/Y^α, we find that the distribution of commute distance conditional on income is

$$P(R = r|Y) = \frac{2\rho_0 \pi r Y^\alpha}{(Y^\alpha + \rho_0 \pi r^2)^2} \quad (16)$$

and that the average commute distance is

$$\overline{r}(I) = \frac{1}{2}\sqrt{\frac{\pi}{\rho_0}} Y^{\alpha/2}, \quad (17)$$

which is a power law with exponent $\beta = \alpha/2$.

The theoretical result Eq. (16) also implies a simple scaling that can be checked empirically. Indeed, if we rescale the commuting distance by $Y^{\alpha/2}$, $u = r/\sqrt{Y^\alpha}$, all the curves for different incomes should collapse on the unique curve that depends on only one parameter and is given by

$$P(u) = \frac{2\pi\rho_0 u}{(1 + \rho_0 \pi u^2)^2}. \quad (18)$$

In the next section, we evaluate these theoretical predictions against our data.

Comparison with empirical results

The closest opportunity model predicts that the average commuting distance varies with income as $\overline{r} \sim Y^{\alpha/2}$, where α depends on the country considered. We will interpret our empirical results in terms of this relationship. For the US, we observe an exponent $\beta_{US} \approx 0$ indicating that the density of jobs is independent from the skill level in the US. For the UK and Denmark, we observe a non-zero exponent with $\beta_{UK} \approx 1/2$ for the UK and a larger value for Denmark $\beta_{DK} \approx 0.8$. These results indicate that the density of jobs decreases with the skill level, more in Denmark than in the UK. The observed difference between the US and two European countries in the

spatial density of jobs at different income levels suggests a more general difference between Europe and the US (for a discussion in equilibrium theory about the spatial distribution of workers and skill levels, see for example [28]). It is interesting to note that there seems to be a correlation between the value of the exponent β and the size of the country. Further studies are however needed in order to confirm this observation.

The crucial prediction allowing us to distinguish between models is the distribution of commuting distances and how it depends on income. Indeed, for the simple spatial extension of the McCall model presented here, the distribution of r decreases very quickly (Eq. (9)) and is not a broad distribution (extending the McCall model with transport costs can lead to a broad distribution such as a power law but this requires fine-tuning of parameters, see the material and methods section). In sharp contrast, in the closest opportunity model, we have a broad distribution of the form given by Eq. (16) and in Fig. 1 we display the one parameter fit with this form for a given income category. The agreement with data is very good for the UK and the US, but there are some discrepancies in the Danish case. It seems that for this Danish case there are other heterogeneities that are not taken into account in our model. In particular, Denmark is a small country with a large proportion of the population living in islands, imposing important constraints on commuting patterns.

An additional and very strong test of the validity of Eq. (16) is provided by the data collapse on the curve given by Eq. (18). In Fig. 3 (right column) we plot the rescaled commuting distance distribution for different income categories and we observe a very good collapse, except for the lower income category in the UK for which the square root behavior is not applicable. We remark that for the US $\beta = 0$, which implies that the probability distribution Eq. (16) does not depend on the income category so that the curves are automatically collapsed. We furthermore note that the agreement between the data and the closest opportunity model for Denmark is strongly reinforced by the data collapse predicted by our model and observed in the data (shown in figure 3).

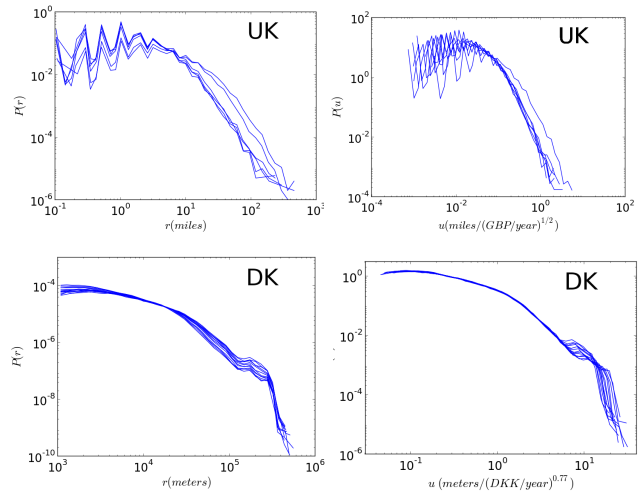

FIG. 3: **Left column: Commuting distance distribution for different income classes.** The probability distribution is shown for different income classes. (Top) UK data (averaged over all available years). (Bottom) Danish data (all years give the same result and we choose here to show the year 2008). **Right column: rescaled probability distribution $P(u)$ for different income classes.** We observe a very good data collapse for both UK data (top), with $\beta = 0.5$ and averaged over all available years, and the Danish data (bottom) for $\beta = 0.77$ and for the year 2008.

DISCUSSION AND PERSPECTIVES

With the increasing availability of ever more precise and comprehensive data we can test a number of predictions of models for the urban structure and its processes. In this paper we predict the distribution of commuting distances and discuss its relation with income. We showed that the empirical data do not support the standard McCall model (based on optimal control) for the job search process. Instead, we have proposed a model based on the closest opportunity that meets the expectation of each individual is able to predict correctly the behavior of the average commuting distance with income in terms of the density of jobs offers. More importantly, this model is able to correctly predict the form of the commuting distance distribution, its broad tail, and the data collapse predicted by its form.

Stated succinctly, previous models relied on the idea that workers wait for a job that pays enough, while in the new closest opportunity model, workers search space for a job that is good enough. Although further studies on more countries are certainly needed, this stochastic model provides a microscopic foundation for a large class of mobility models and opens many interesting research directions in modeling mobility while leading to testable predictions. More generally, we proposed here an alternative framework to study human or animal behavior, in which actions are taken not on the basis of an optimal strategy but on the first opportunity that is good enough. This framework would potentially find some applications in our understanding of foraging for example and other applications in ecology or finance where optimal control might be a too strong assumption.

MATERIALS AND METHODS

Data description

As we describe below, for both the US and the UK dataset a weighting methodology has been developed in order to take into account non-responses, undercoverage, multiple telephones in a household (for the US dataset) and drops-off in the travel recording (for the UK dataset). This methodology has been developed in order to make data trustable and usable, but without any doubt, there is noise in the data (and probably self-reporting errors too). One can indeed note that there is a bias for low income values (for both for UK and USA data) which is very likely due to rounding. However, this bias does not change the order of magnitude of the commuting distance and thus does not substantially affect the results.

UK data

We used data from the UK National Travel Survey (NTS) for the years $2002-2012$ [13]. Each year's sample has a size of $15,048$ addresses and was designed to provide a representative sample of households in the UK. A weighting methodology was developed to adjust for non-responses and drop-offs in the travel recording. Data collection is obtained from face-to-face interviews and a seven day travel record of individual daily travel activity.

We specifically exploit the *individual* and the *trip* files of this dataset. The *individual* file is used to determine the income category of each individual (data provides 23 income bands). The *trip* file allows us to link individuals to their weekly commuting trips for which we know the distance. In order to compute the average commuting distance as a function of the income class, we first average the commuting distance of each individual, including all transportation modes, over the number of commuting trips undertaken during the week. We then average these quantities over all individuals for each income category. When we consider average values from these data, we do not distinguish between different transportation modes or the geographical locations of the origin and destination of the trip.

US data

We used data from the 1995, 2001 and 2009 national household travel survey (NHTS) [14], a survey of the civilian, non-institutionalized population of the United States. The NHTS datasets contain data for respectively $42,033$, $26,032$, and $150,147$ households (with approximately 40,000 add-on interviews for the latest version). Weighting factors are used in order to take into account nonresponses, undercoverage, and multiple telephones in a household.

These datasets allow us to associate an income category to each worker (this dataset indicates 18 different income bins) and the one-way distance to workplace. For the 2009 NHTS, the personal income is not provided, in this case we proxy personal income by the household income divided by the household size.

Danish data

The Danish data are derived from annual administrative register data from Statistics Denmark for the years $2001-2010$. We observe the full population of workers, and for each year, we have information on the workers annual income and their commuting distance. We used the post-tax income. Commuting distances have been calculated using information on exact residence and workplace addresses using the shortest route in between. Note that for these data, no weighting methodology is required as we observe the full population of workers in the country.

Including transport cost in the McCall model

We discuss here the general case for the McCall model where there is a transport cost associated with distance. The distance from the home of worker to a job offer is then a random variable R having density $2\pi\rho r$, which is independent of the wage W associated with the job. In order to link the probability of accepting a job to space, we assume a linear transport cost δR that is paid by the worker if she accepts a job. Ultimately, she cares about the net wage $W-\delta R$. The optimal strategy of the worker involves a reservation wage τ and the worker accepts the first offer that offers a net wage $W-\delta R>\tau$. These assumptions already imply that the commuting distance for the accepted job satisfies $R<\frac{W-\tau}{\delta}$. Then, the tail behavior of the commuting distance cannot follow a power law if W has a bounded distribution and we therefore allow W to have an unbounded distribution. The density of commuting distances is

$$\begin{aligned} P\left(R=r|W-\delta R>\tau\right) &= \frac{P\left(R=r\right)P\left(W-\delta r>\tau\right)}{P\left(W-\delta R>\tau\right)} \\ &= \frac{2\pi\rho r\left(1-F\left(\tau+\delta r\right)\right)}{\int_0^\infty 2\pi\rho s\left(1-F\left(\tau+\delta s\right)\right)ds}. \end{aligned} \tag{19}$$

From this, we can observe that

$$\frac{\partial \ln P}{\partial \ln r} = 1-\frac{f\left(\tau+\delta r\right)}{1-F\left(\tau+\delta r\right)}\delta r. \tag{20}$$

which shows that in general P does not decay as a power law, unless $\frac{f(\tau+\delta r)}{1-F(\tau+\delta r)}=\frac{\zeta}{r}$ for some $\zeta>1$. In the specific

case where W follows a power law with $F(W) = 1 - W^{-\zeta}, W > 1$, we obtain

$$\frac{\partial \ln P}{\partial \ln r} = 1 - \frac{\zeta r}{\tau + \delta r}, \tag{21}$$

which tends to $1 - \frac{\zeta}{\delta}$ as $r \to \infty$. This model thus leads to a power law for the distribution of commute distances, if the distribution of wage offers follows a power law. If we consider all wages, the Pareto law tells us that they can be broadly distributed, but this is not the quantity needed here. Indeed we are considering here the offer distribution for a given set of skills and it is very unlikely that a given individual will sample offers that range over the whole income distribution.

We can then compute the relationship between the average commuting distance and income in this model. For $w - \delta r > \tau$, we have

$$\begin{aligned} & P(R = r | W = w, W - \delta R > \tau) = \\ & = \frac{P(R = r, W = w, w - \delta r > \tau)}{P(W = w, \delta R < w - \tau)} \\ & = \frac{2r}{\left(\frac{w-\tau}{\delta}\right)^2}, \end{aligned} \tag{22}$$

which leads to the conditional expectation

$$\begin{aligned} E(R | W = w, W - \delta R > \tau) &= \int_0^{\frac{w-\tau}{\delta}} \frac{2r^2}{\left(\frac{w-\tau}{\delta}\right)^2} \mathrm{d}r \\ &= \frac{2}{3} \frac{w - \tau}{\delta}. \end{aligned} \tag{23}$$

This model thus predicts that for a linear transport cost, the expected commute distance is always linear in income which does not fit the empirical findings.

In any case, it seems that in order to predict results consistent with empirical observations (a broad law such as a power law with exponent close to 3 for the distribution, and a power law behavior for the average distance), this model needs fine-tuning of the parameters, in sharp contrast with the closest opportunity model.

Including transport costs in the closest opportunity model

Workers base their decisions on transport costs that depend not only on distance but also on monetary costs and travel time. We shall see how transport costs can be accommodated by the closest opportunity model proposed in this paper. This is useful as we get exact predictions regarding how the observables of the model are modified by transport costs. The model can then also be used for prediction in cases when transport costs change.

We let the variable r represent here the transport cost, the closest opportunity model predicts

$$\mathrm{d} \log P(R = r) / \mathrm{d} \log(r) = -3 \tag{24}$$

In general we may expect that the transport cost is an increasing and concave function of distance, since travelers switch to faster modes for longer trips. Denoting the physical distance by ℓ, we assume that $r \sim \ell^{\nu}$, where $0 < \nu < 1$. In terms of distance we then find that

$$\frac{\mathrm{d} \log P(R = r)}{\mathrm{d} \log \ell} = \frac{\mathrm{d} \log P(R = r)}{\mathrm{d} \log r} \frac{\mathrm{d} \log r}{\mathrm{d} \log \ell} = -3\nu. \tag{25}$$

For the income elasticity, the model predicts a relationship between transport cost and income that is $\beta = \mathrm{d} \log \overline{r(I)} / \mathrm{d} \log I = 1/2$. The elasticity of commuting distance with respect to income is then larger:

$$\frac{\mathrm{d} \log \overline{r^{1/\nu}(Y)}}{\mathrm{d} \log Y} = 1/2 + \frac{1}{2\nu}.$$

Observing the commuting distance rather than transport cost, we thus expect an exponent in the tail of the distribution smaller than 3 in absolute value and an income elasticity of the average commuting distance that is greater than 1/2. It is thus possible to back out the exponent ν from both observed exponents.

DATA ACCESSIBILITY

For UK data [13], a special licence access can be provided upon request to the Department for Transport, National Travel Survey, 2002-2012: Special Licence Access [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], December 2014. SN: 7553 , `http://dx.doi.org/10.5255/UKDA-SN-7553-2`.

USA data [14] are provided by the U.S. Department of Transportation, Federal Highway Administration, National Household Travel Survey. These can by download from URL address: `http://nhts.ornl.gov.`

The Danish original dataset is not in open access. Please contact the authors for further information about the accessibility of aggregated data.

AUTHORS' CONTRIBUTIONS

All authors (GC, IM, MF, MB) participated to the design of the study and drafted the manuscript. IM prepared the Danish data and GC carried out the data analysis.

ACKNOWLEDGEMENTS

We acknowledge the Departement of Transport, the National Centre for Social Research, the UK Data Archive and the Crown, for the UK dataset. We thank Statistics Denmark for providing us with the danish dataset. Giulia Carra thanks the Complex Systems Institute in Paris (ISC-PIF) for hosting her during part

of this work. Mogens Fosgerau and Ismir Mulalic have received funding from the Danish Innovation Fund under the Urban project. Marc Barthelemy thanks Jean-Philippe Bouchaud, Pablo Jensen, Clément Sire, and Guy Theraulaz for stimulating discussions.

* Electronic address: marc.barthelemy@cea.fr

[1] United Nations. http://unstats.un.org/.

[2] The New Science of Cities. Michael Batty MIT Press, 2013.

[3] Lippman S.A., McCall J.J. (1976). The economics of job search: a survey. *Economic Inquiry*, 14(2), 155-189.

[4] McCall J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 84(1), 113-126.

[5] Stigler G. J. (1961). The economics of information. *The journal of political economy*, 69(3), 213-225.

[6] Zenou, Y. (2009). Urban labor economics. Cambridge University Press.

[7] F. Simini, M. C. Gonzáles, A. Maritan and A.L. Barabási (2012). A universal model for mobility and migration patterns. *Nature*, 484, 96-100.

[8] S. Lenhart and J. T. Workman (2007). Optimal Control Applied to Biological Models, Mathematical and Computational Biology, Chapman and Hall, Boca Raton, Fla, USA ;CRC Press, London, UK.

[9] M. Athans and P. L. Falb (2006). Optimal Control: An Introduction to the Theory and Its Applications, Dover Publications.

[10] P. Chen and S. M. N. Islam (2005). Optimal Control Models in Finance: A New Computation Approach, Springer.

[11] S. Anita, V. Arnautu, and V. Capasso (2011). An Introduction to Optimal Control Problems in Life Sciences and Economics: From Mathematical Models to Numerical Simulation with MATLAB.Springer Science & Business Media (2011).

[12] G.M. Viswanathan, M.G. Da Luz, E.P. Raposo, H.E. Stanley. *The physics of foraging: an introduction to random searches and biological encounters*. Cambridge University Press (2011).

[13] Department for Transport, National Travel Survey, 2002-2012: Special Licence Access [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], December 2014. SN: 7553 , `http://dx.doi.org/10.5255/UKDA-SN-7553-2`.

[14] U.S. Department of Transportation, Federal Highway Administration, 2009 National Household Travel Survey. URL: `http://nhts.ornl.gov`.

[15] Alonso, W. (1964). Location and land use. Toward a general theory of land rent. Location and land use. Toward a general theory of land rent. Cambridge Mass., Harvard University Press.

[16] Muth, R. F. (1969) Cities and Housing. Chicago. University of Chicago Press.

[17] Brueckner, J. K. (2000). Urban sprawl: diagnosis and remedies. International regional science review, 23(2), 160-171.

[18] Chow Y. S., Robbins H., Siegmund D. (1971). Great expectations: the theory of optimal stopping. *Houghton Mifflin*.

[19] A. Wald (1947). Sequential Analysis. *John Wiley & Sons*.

[20] R. N. Bradt, S. M. Johnson, S. Karlin (1956). On sequential designs for maximizing the sum of n observations. *Ann. Math. Statist.* 27,1060-1074.

[21] A. N. Shiryaev (1963). On optimal methods in quickest detection problems. *Theory Prob. and Appl.* 8, 22-46.

[22] G. W. Haggstrom (1966). Optimal stopping and experimental design. *Ann. Math. Statist.* 37, 7-29.

[23] S. Rasmussen, N. Starr (1979). Optimal and adaptive search for a new species. *J. Amer. Statist. Assoc.* 74, 661-667.

[24] Bellman R. (1957). Dynamic programming. *Princeton University Press*.

[25] P. J. Diggle (1983). Statistical analysis of spatial point patterns. *Academic Press*, London.

[26] Hornstein, A., Krusell, P., Violante, G. L. (2007). Frictional wage dispersion in search models: A quantitative assessment (No. w13674). National Bureau of Economic Research.

[27] Hall, R. E., Mueller, A. I. (2013). Wage dispersion and search behavior. Unpublished manuscript `http://www.nber.org/papers/w21764#` (accessed 12.01.2015).

[28] Brueckner, J. K., Thisse, J. F., Zenou, Y. (2002). Local labor markets, job matching, and urban location. International Economic Review, 43(1), 155-171.

[29] Rupert P., Stancanelli E., Wasmer E. (2009). *Annales d'Economie et de Statistiques*, (95 - 96), 201 - 221.

[30] M. Belkin and P. Niyogi, *Using manifold structure for partially labelled classification*, Advances in NIPS, 15 (2003).