Modeling the Polycentric Transition of Cities

Rémi Louf* and Marc Barthelemy†
Institut de Physique Théorique, CEA, CNRS-URA 2306, F-91191 Gif-sur-Yvette, France
(Received 24 May 2013; published 6 November 2013)

Empirical evidence suggests that most urban systems experience a transition from a monocentric to a polycentric organization as they grow and expand. We propose here a stochastic, out-of-equilibrium model of the city, which explains the appearance of subcenters as an effect of traffic congestion. We show that congestion triggers the instability of the monocentric regime and that the number of subcenters and the total commuting distance within a city scale sublinearly with its population, predictions that are in agreement with data gathered for around 9000 U.S. cities between 1994 and 2010.

DOI: 10.1103/PhysRevLett.111.198702 PACS numbers: 89.65.Lm, 05.10.Gg, 89.75.Fb, 89.75.Hc

As cities grow, they evolve from monocentric organizations where all the activities are concentrated in the same geographical area—usually the central business district—to more distributed, polycentric organizations [1–8]. Traditional approaches in spatial economics have attempted to describe the phenomenon within the framework of equilibrium models of the city [9,10]. These models are traditionally based on the concept of agglomeration economies—to explain why economical activities tend to group—and the spatial distribution of wages and rents across the urban space. However, these approaches fail at giving a satisfactory quantitative account [11,12] of the polycentric transition of cities. First, they describe a city as being in an equilibrium characterized by static spatial distributions of households and business firms. However, the equilibrium assumption is unsupported as cities are out-of-equilibrium systems and their dynamics is of particular interest for practical applications [12]. Second, these models integrate so many interactions and variables that it is difficult to understand the hierarchy of processes governing the evolution of cities, which ones are fundamental and which ones are irrelevant. Yet, traffic congestion is not explicitly taken into account in the existing models, despite being mentioned in the economics literature as a possible reason for the polycentric transition [6]. Lastly, the models do not make any quantitative prediction and are therefore unsupported by data. We present in this Letter a stochastic, out-of-equilibrium model of the city which relies on the assumption that the polycentric structure of large cities might find its origin in congestion, irrespective of the particular local economic details. We are able to reproduce many stylized facts, and, most importantly, to derive a general relation between the number of activity centers of a city and its population. Finally, we verify this relation against the employment data from around 9000 cities in the U.S. between 1994 and 2010.

Following recent interdisciplinary efforts to construct a quantitative description of cities and their evolution [12–18], we deliberately omit certain details and focus instead on basic processes. We thereby aim at building a minimal model which captures the complexity of the system and is able to account for qualitative as well as quantitative stylized facts. The model we propose is by essence dynamical and describes the evolution of cities' organization as their populations increase. We focus on car congestion—mainly due to journey-to-work commutes—and its effect on the job location choice for individuals.

According to Fujita and Ogawa's classical model [9] in spatial economics, an individual living at location i will choose to work at the location j that maximizes the net income after the deduction of rent and commuting costs [9]

$$Z_0 = W(j) - C_R(i) - C_T(i, j), \qquad (1)$$

where $W(j)$ is the average wage paid by business firms located at j (and thus varies from one location to another), $C_R(i)$ is the land rent at i, and $C_T(i, j)$ is the commuting cost between i and j. The wage and the land rent result from the interplay between the households' and companies' locations, agglomeration effects being taken into account. The commuting cost, on the other hand, does not usually take congestion into account and is taken proportionally to the Euclidean distance $C_T(i, j) = td_{ij}$ (where t is the transportation cost per unit of distance) in most studies.

The time scales involved in the evolution of cities are usually such that the employment turnover rate is larger than the relocation rate of households. On a short time scale, we can thus focus on the process of job seeking alone, leaving aside the problem of the choice of residence. In other words, we assume the coupling between both processes to be negligible: we assume that each inhabitant newly added to the city has a random residence location and we concentrate on understanding how such an inhabitant chooses its job among a pool of N_c potential activity centers (which we suppose are also randomly distributed among the city). The active subcenters are then defined as the subset of potential centers, which have a nonzero incoming number of individuals. As a result of these assumptions, a worker living at i will choose to work at the center j such that the quantity

$$Z_{ij} = W(j) - C_T(i, j) \tag{2}$$

is maximum.

We now discuss the form of the two terms $W(j)$ and C_T. The problem of determining the (spatial) variations of the average wage $W(j)$ at location j is very reminiscent of some problems encountered in fundamental physics. Indeed, the wage depends on many different factors ranging from the type of company, the education level of the inhabitant, the level of agglomeration, etc., and in this respect is not too different from quantities that can be measured in a large atom made of a large number of interacting particles. In this situation, physicists found out that although it is possible to write down the corresponding equations, not only is it impossible to solve them, but it is also not really useful. In fact, they found out [19] that a statistical description of these systems relying on random matrices could lead to predictions that agree with experimental results. We wish to add in spatial economics this idea of replacing a complex quantity such as wages, which depends on so many factors and interactions, by a random one. We therefore decide to account for the interaction between activity centers and people by taking the wage as proportional to a random variable $\eta_j \in [0, 1]$ such that $W(j) = s\eta_j$, where s defines the maximum attainable average wage in the considered city.

As for the transportation cost $C_T(i, j)$, we choose it to be proportional to the commuting time between i and j. In a typical situation where passenger transportation is dominated by personal vehicles, this commuting time not only depends on the distance between the two places but also on the traffic between i and j, the vehicle capacity of the underlying network, and its resilience to congestion. The Bureau of Public Road formula [20] proposes a simple form taking all these factors into account. In our framework, it leads to the following expression for the commuting costs:

$$C_T(i, j) = td_{ij}\left[1 + \left(\frac{T_{ij}}{c}\right)^{\mu}\right], \tag{3}$$

where T_{ij} is the traffic per unit of time between i and j, and c is the typical capacity of a road (taken constant here). The quantity μ is a parameter quantifying the resilience of the transportation network to congestion. We further simplify the problem by assuming that the traffic T_{ij} is only a function of the subcenter j and therefore write $T_{ij} = T(j)$, the total traffic incoming in the subcenter j (see Supplemental Material [21] for a short discussion).

In summary, our model is defined as follows. At each time step, we add a new individual i located at random in the city, who will choose to work in the activity area j (among N_c possibilities located at random) such that the quantity

$$Z_{ij} = \eta_j - \frac{d_{ij}}{\ell}\left[1 + \left(\frac{T(j)}{c}\right)^{\mu}\right] \tag{4}$$

is maximum (we omitted irrelevant multiplicative factors). The quantity $\ell = s/t$ is interpreted as the maximum effective commuting distance that people can financially withstand.

Depending on the relative importance of wages, distance, and congestion, the model predicts the existence of three different regimes: the monocentric regime (top Fig. 1), the distance-driven polycentric (middle Fig. 1) regime, and the attractivity-driven polycentric (bottom Fig. 1) regime.

From now on, we will assume that ℓ is large enough so that a monocentric state exists for small values of the population. In this regime, the value of η prevails, and the monocentric state evolves to an attractivity-driven polycentric structure as the population increases (if ℓ is too small, the monocentric regime does not exist—see the Supplemental Material [21] for more details on these points). Starting from a small city with a monocentric organization, the traffic is negligible and $Z_{ij} \approx \eta_j$, which implies that all individuals are going to choose the most attractive center (with the largest value of η_j, say η_1). When the number P of households increases, the traffic will also increase and some initially less attractive centers (with smaller values of η) might become more attractive, leading to the appearance of new subcenters characterized

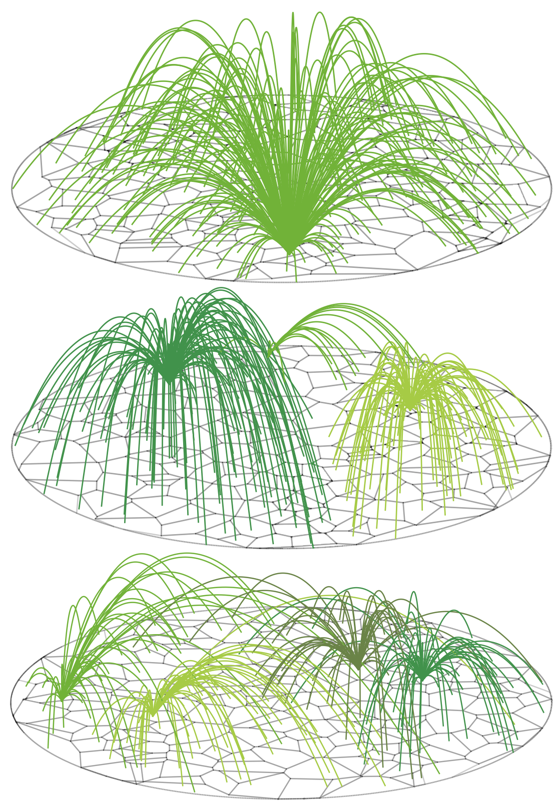

FIG. 1 (color online). The monocentric (top), distance-driven polycentric (middle), and attractivity-driven polycentric (bottom) regimes as produced by our model. Each link represents a commute to an activity center.

by a nonzero number of commuters. More precisely, a new subcenter j will appear, when for an individual i, we have $Z_{ij} > Z_{i1}$. The traffic so far is $T(1) = P$ and $T(j) = 0$, which leads to the equation

$$\eta_j - \frac{d_{ij}}{\ell} > \eta_1 - \frac{d_{i1}}{\ell}\left[1 + \left(\frac{P}{c}\right)^{\mu}\right]. \quad (5)$$

We assume that there are no spatial correlations in the subcenter distribution, so that we can make the approximation $d_{ij} \sim d_{i1} \sim L$, where L is the linear size of the city. The new subcenter will thus be such that $\eta_1 - \eta_j$ is minimum, implying that it will have the second largest value denoted by $\eta_j = \eta_2$. For a uniform distribution (details of the calculation can be found in the Supplemental Material [21], Section 2), on average $\overline{\eta_1 - \eta_2} \simeq 1/N_c$ leading to a critical value for the population

$$P^* = c\left(\frac{\ell}{LN_c}\right)^{1/\mu}. \quad (6)$$

Whichever system is considered, there will therefore always be a critical value of the population above which the city becomes polycentric (which can be smaller than 1, in which case there is no monocentric regime at all, see the Supplemental Material [21]). The monocentric regime is therefore fundamentally unstable with regards to population increase, which is in agreement with the fact that no major city in the world exhibits a monocentric structure. We note that the smaller the value of μ (or the larger the value of the capacity c), the larger the critical population value P^*, which means that cities with good road systems that are capable of absorbing large traffic show a monocentric structure for a longer period of time.

Having established that cities will eventually adopt a polycentric structure, we can wonder how the number of subcenters varies with the population. We compute the value of the population at which the kth center appears. We still assume that we are in the attractivity-driven regime and that, so far, $k-1$ centers have emerged with $\eta_1 \geq \eta_2 \geq \ldots \geq \eta_{k-1}$ [21], with a number of commuters $T(1), T(2), \ldots, T(k-1)$, respectively. The next worker i will choose the center k if

$$Z_{ik} > \max_{j\in[1,k-1]} Z_{ij}, \quad (7)$$

which reads

$$\eta_k - \frac{d_{ik}}{\ell} > \max_{j\in[1,k-1]}\left\{\eta_j - \frac{d_{ij}}{\ell}\left[1 + \left(\frac{T(j)}{c}\right)^{\mu}\right]\right\}. \quad (8)$$

The distribution of traffic $T(j)$ is narrow [21], which means that all the centers have roughly the same number of commuters $T(j) \sim P/(k-1)$. As above, we also assume that the distance between the workers' households and the activity centers is typically $d_{ij} \sim d_{ik} \sim L$. The previous expression now reads

$$\frac{L}{\ell}\left(\frac{P}{(k-1)c}\right)^{\mu} > \max_{j\in[1,k-1]}(\eta_j) - \eta_k. \quad (9)$$

Following our definitions, $\max_{j\in[1,k-1]}(\eta_j) = \eta_1$. According to order statistics, if the η_j's are uniformly distributed, we have on average $\overline{\eta_1 - \eta_k} = (k-1)/(N_c+1)$. It follows from these assumptions that (1) the kth center to appear is the kth most attractive one, and (2) the average value of the population $\bar{P}_k$ at which the kth center appears is given by

$$\bar{P}_k = P^*(k-1)^{(\mu+1)/\mu}. \quad (10)$$

Conversely, the number k of subcenters scales sublinearly with population as

$$k \sim \left(\frac{\bar{P}}{P^*}\right)^{\mu/(\mu+1)}. \quad (11)$$

It is interesting to note that this result is robust: the dependence is sublinear, whatever the distribution of the random variable η (see the Supplemental Material [21] for a discussion on this point). We can therefore conclude that, probably very generally and under mild assumptions, the number of activity subcenters in urban areas scales sublinearly with their populations where the prefactor and the exponent depend on the properties of the transportation network of the city under consideration.

A previous study [22] showed that the total miles driven daily in a city—the "total commuting distance"—scales with the population as $L_{\text{tot}} \sim P^{\gamma}$ where $\gamma \in [0.5, 1]$, which the authors interpreted as cities having neither totally centralized nor totally decentralized structures. We can discuss this result within the framework of our model in the following way. If the system was in the pure attractivity-driven regime, we would have $L_{\text{tot}} \sim P$. But, if we assume that we are in an intermediate regime where Eq. (10) holds and where the system exhibits spatial coherence [21], we can write the total length of the commutes as

$$L_{\text{tot}} \sim P\frac{L}{\sqrt{k}}. \quad (12)$$

Inverting the result from Eq. (10), we therefore get

$$L_{\text{tot}} \sim P^{1-(\beta/2)}, \quad (13)$$

where $\beta = (\mu/(\mu+1)) \in [0, 1]$. Our model is thus consistent with the fact that the total traveled miles scales with population with a nontrivial exponent comprised in [0.5, 1].

We now test the prediction given by Eq. (10). For that purpose, we obtained data on the number of employees per U.S. zip code that were collected over a span of 16 years (from 1994 to 2010) [23], as well as the population of all cities in the U.S. between 1994 and 2010 [24]. We estimate the number of subcenters by constructing the rank plot of the employment density ρ (number of employees per km^2) for each zip code of a given urban area [5,25]. These plots display a decay as fast as an exponential (Fig. 2), which

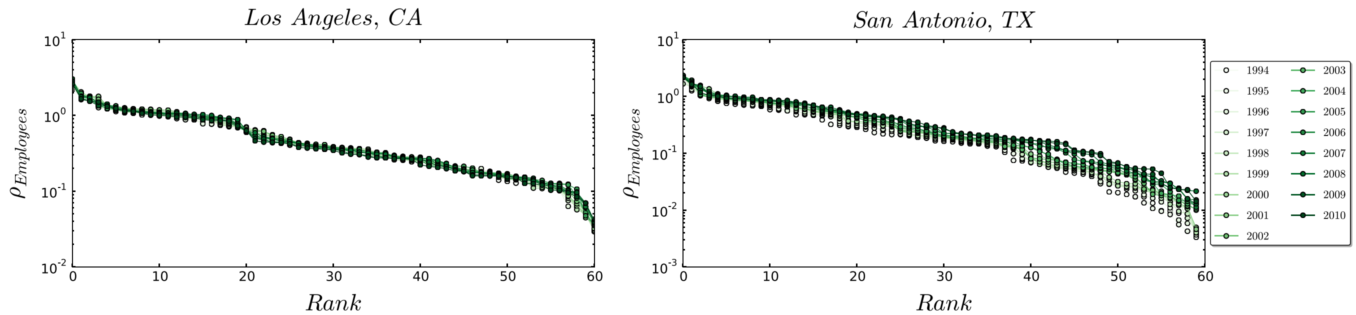

FIG. 2 (color online). Rank plot for the employment density (in employees per km^2) in Los Angeles, CA (left) and San Antonio, TX (right) between 1994 and 2010. See the Supplemental Material [21] for more details.

implies that there exists a natural scale for the rank that we interpret here as the typical number of activity centers. It also implies that any reasonable method should give an estimate of the number of subcenters of the same order of magnitude (which would not be the case for slowly decaying functions, such as power laws, for example). We first note that for some cities—typically large ones with stable populations (Fig. 2, left)—the employment spatial statistics remained stable over the period of study. For other cities, we observe large variation of the number of subcenters (Fig. 2, right). We then plot (Fig. 3) the population P of cities (with population $P > 100$) versus the estimated number of subcenters k (the dispersion in the scatter plot probably results from the fact that different cities have different resilience levels to congestion). On average, we observe a power law dependence with exponent $\delta = 1.56 \pm 0.15$ (the result is robust with regards to the estimate of k, see the Supplemental Material [21] for more

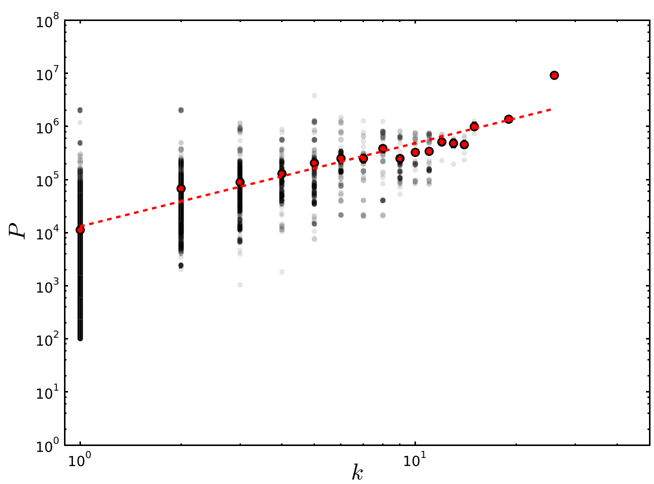

FIG. 3 (color online). Scatter plot of the estimated number of subcenters versus the population for about 9000 cities with populations over 100 people in the U.S. The red dots represent the average population for a given number of subcenters. We fit this average with a power-law dependence (represented by a red dashed line) giving an exponent $\delta = 1.56 \pm 0.15$ ($r^2 = 0.87$). See the Supplemental Material [21] for more details on the computation of k and the robustness of the results.

details). Inverting this relation gives us the number of subcenters as a function of the population

$$k \sim P^{\beta}, \tag{14}$$

with $\beta \sim 0.64$. This result is strikingly in agreement with the prediction given by our model: the number of subcenters in a city scales sublinearly with its population.

Using the measured value of β and Eq. (13), we can estimate the exponent of the scaling of L_{tot} with the population and find $L_{\mathrm{tot}} \sim P^{0.68}$, which agrees very well with the value 0.66 measured in Ref. [22] directly on the data of the daily total miles driven in more than 400 cities in the U.S.

While agglomeration economies seem to be the basic process explaining the existence of cities and their spectacular resilience, this study brings evidence that congestion is the driving force that tears them apart. The nontrivial spatial patterns observed in large cities can thus be understood as a result of the interplay between these competing processes. We believe that the present model represents an important step towards a quantitative, predictive science of cities. More generally, this microscopic approach is an interesting example of an out-of-equilibrium model: it is governed by local optimization with saturation effects, leads to different regimes, and is characterized by nontrivial dynamical exponents. In this respect, we believe that this discrete approach might be of use in the study of pattern formation in biology, which has been so far explored from a global optimization perspective [26], or as a discrete approach to reaction-diffusion processes with density-dependent diffusion coefficients [27] to compute quantities that are out of reach within the current methods.

M. B. thanks H. Berestycki, A. Flammini, and K. Mallick for stimulating discussions. M. B. acknowledges funding from the EU Commission through project EUNOIA (Project No. FP7-DG.Connect-318367).

*remi.louf@cea.fr
†marc.barthelemy@cea.fr

[1] P. Kemper and R. Schmenner, J. Urban Economics **1**, 410 (1974).

[2] J. Odland, Econ. Geogr. **54**, 234 (1978).
[3] E. S. Mills, *Studies in the Structure of the Urban Economy* (John Hopkins University Press, Baltimore, MD, 1972).
[4] D. A. Griffith, Prof. Geogr. **33**, 189 (1981).
[5] V. Dökmeci and L. Berköz, Eur. Plan. Stud. **2**, 193 (1994).
[6] D. P. McMillen and S. C. Smith, J. Urban Economics **53**, 321 (2003).
[7] R. H. M. Pereira, V. Nadalin, L. Monasterio, and P. H. M. Albuquerque, Geogr. Anal. **45**, 77 (2013).
[8] C. Roth, S. M. Kang, M. Batty, and M. Barthelemy, PLoS One **6**, e15923 (2011).
[9] M. Fujita and H. Ogawa, Regional Science Urban Economics **12**, 161 (1982).
[10] M. A. Fujita, A. Venables, and P. Krugman, *The Spatial Economy: Cities, Regions and International Trade* (MIT Press, Cambridge, MA, 1999).
[11] J.-P. Bouchaud, Nature (London) **455**, 1181 (2008).
[12] M. Batty, Science **319**, 769 (2008).
[13] H. A. Makse, S. Havlin, and H. E. Stanley, Nature (London) **377**, 608 (1995).
[14] D. H. Zanette and S. C. Manrubia, Phys. Rev. Lett. **79**, 523 (1997).
[15] M. Marsili, S. Maslov, and Y.-C. Zhang, Phys. Rev. Lett. **80**, 4830 (1998).
[16] M. Marsili and Y.-C. Zhang, Phys. Rev. Lett. **80**, 2741 (1998).
[17] M. Batty, *Cities and Complexity: Understanding Cities Through Cellular Automata, Agent-Based Models, and Fractals* (MIT Press, Cambridge, MA, 2005).
[18] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, Proc. Natl. Acad. Sci. U.S.A. **104**, 7301 (2007).
[19] F. J. Dyson, J. Math. Phys. (N.Y.) **3**, 140 (1962).
[20] D. Branston, Transportation research **10**, 223 (1976).
[21] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.111.198702 for more details on the data analysis and technical details of the calculations.
[22] H. Samaniego and M. E. Moses, J. Transp. Land Use **1**, 21 (2008).
[23] Zip business patterns data can be found online at http://www.census.gov/.
[24] Census data can be found online at http://www.census.gov/.
[25] D. A. Griffith, J. Urban Economics **9**, 298 (1981).
[26] D. J. Ashton, T. C. Jarrett, and N. F. Johnson, Phys. Rev. Lett. **94**, 058701 (2005)
[27] M. E. Cates, Rep. Prog. Phys. **75**, 042601 (2012)