A new economic geography model of central places

Takatoshi Tabuchi [a,*], Jacques-François Thisse [b,c,d]

[a] Faculty of Economics, University of Tokyo, Japan
[b] CORE, Université catholique de Louvain, Belgium
[c] PSE, France
[d] CEPR, United Kingdom

ABSTRACT

One of the most striking feature of the space-economy is that cities form a hierarchical system exhibiting some regularity in terms of their size and the array of goods they supply. In order to show how such a hierarchical system may emerge, we consider a model with monopolistically competitive markets for the industrial sectors. As transport costs steadily decrease from large values, the urban system formed by several small cities entails structural changes in that some cities expand at the expense of the others by attracting a growing number of industries. Beyond some threshold, some cities disappear from the space-economy. Such an evolution of the urban system describes fairly well what has been observed in various historical periods that have experienced major changes in transportation technologies and/ or political unification.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The main thrust of new economic geography is that steadily decreasing transport costs foster the agglomeration of economic activities in a small number of urban regions. Tackling the formation of the urban system from this angle is especially relevant because, ever since the beginning of the Industrial Revolution, the transport sector has undergone the most stunning changes. According to Bairoch (1997) "On the whole, between 1800 and 1910, it can be estimated that the lowering of the real average prices of transportation was on the order of 10 to 1." [our translation]. Transport costs have continued to decrease after World War I. For example, in the United States, Glaeser and Kohlhase (2004) observe that over the twentieth century, the costs of moving manufactured goods have declined by over 90% in real terms. The World Bank (1995) reports comparable drops from 10 to 3 and from 10 to 1.5 in maritime and air freights, respectively.

To a large degree, however, by focussing primarily on the two-location framework of international trade theory, the existing literature has failed to address its main objective. Using the dataset of World Urbanization Prospects 2009, we want to figure out whether the above prediction is supported by casual empirical evidence.

Specifically, we have computed the ratio of the population in the top five urban agglomerations to the total population in each OECD country having at least five agglomerations with more than 750,000 inhabitants over the period 1950–2005.[1] This leaves us with 10 countries. By focussing on the share, instead of the level, of urban population we control for the urban population growth, while retaining OECD countries allows us to avoid comparing countries having very different transport infrastructure. In order to determine the evolution of the population share in each country's five largest metros, we have computed the correlation between the value of this share and the corresponding year. Note that testing the significance of the correlation between the share and the year is equivalent to testing the significance of coefficients when regressing the share on the year. The correlation is positive and significant in 8 out of 10 countries (Australia, Canada, France,

---

* Corresponding author.
  E-mail address: ttabuchi@e.u-tokyo.ac.jp (T. Tabuchi).

[1] According to World Urbanization Prospects (United Nations, Department of Economic and Social Affairs, http://esa.un.org/wup2009/wup/source/country.aspx), the term "urban agglomeration" refers to the population contained within the contours of a contiguous territory inhabited at urban density levels without regard to administrative boundaries. It incorporates the population in a city plus that in the suburban areas lying outside of, but being adjacent to, the city boundaries. However, some countries do not produce data according to the concept of urban agglomeration but use instead that of metropolitan area or city proper. Whenever possible, these data are adjusted to match the concept urban agglomeration. When sufficient information is not available to permit such an adjustment, data based on the concept of city proper or metropolitan area are used.

Italy, Japan, Mexico, Republic of Korea, and Turkey), whereas the correlation is almost zero in the US and negative and significant only for the UK.[2] Thus, we find it fair to say that those results do not reject the main prediction of new economic geography.

Building on this, we want to study whether and how a system of central places may emerge in a multi-location space when transport costs keep decreasing. To achieve our goal, we consider a spatial economy endowed with different industrial sectors operating under monopolistic competition and increasing returns, in which the number, size, and location of cities are determined endogenously. More precisely, we focus on the size and location of cities (the *urban* aspect) as well as on the spatial distribution of each industry across cities (the *industrial* aspect) when workers are free to choose where to live and for which industry to work. The novelty of our analysis is that the hierarchical principle of central place theory (i.e. the number of goods supplied in a city rises with its size and the spacing of cities having the same size is equal) stems from a symmetry-breaking process, which is itself triggered by falling transport costs.

That said, we may summarize our main findings as follows. First of all, when transport costs are high, we show that there exist a large number of small and equidistant cities. That each city has the same size and industrial structure makes all goods accessible to its workers and to the farmers living in its hinterland. This takes the concrete form of consuming a limited range of varieties of each good as importing varieties from other cities is very expensive. Such a pattern agrees with the fact that, in pre-industrial economies characterized by high transport costs, a city's rural hinterland was often its main external market. As transport costs steadily decrease, the volume of trade grows and the urban system entails some structural changes: some cities expand at the expense of the others by attracting industries and workers. The reason is that bigger cities allow firms to better exploit scale economies, even though the goods produced in small cities are also produced in big cities. To put it bluntly, the urban system now involves *the coexistence of cities having different sizes and different industrial structures*, i.e. small towns and big cities. Moreover, in equilibrium, only adjacent cities are equidistant, which means that our setting allows for an equilibrium pattern in which city locations are asymmetric.

Following a ladder of thresholds, we show that the urban system displays a series of pitchfork bifurcations in which small cities disappear gradually from the space-economy, while *a shrinking number of cities accommodate a growing range of activities*. Another important distinguishing feature of our results is that our economic geography is of the putty-clay type: workers and firms are free to launch a city anywhere but, once it exists, a city has a well-defined location that does not change, even in the absence of durable infrastructure such as roads and public facilities. Last, *cities gaining primacy in the urban hierarchy retain their high rank during the whole process*, the main victims being the towns at the bottom. All of this is in accordance with the fact that "cities show remarkable resilience" (Hohenberg, 2004).

Related literature. The bulk of the research on central place theory has been directed towards identifying geometric conditions under which a superposition of regular structures is possible. These considerations are only interesting if they are based on microeconomic foundations. If there are no economic forces which lead firms of different types to cluster, it is hard to see why a central place system would be more likely to emerge than any other configuration. One of the first economic contributions to central place theory we are aware of is due to Eaton and Lipsey (1982),

who develop a spatial competition model of central places, and to Quinzii and Thisse (1990), who retain the same approach to show that the central place configuration is socially optimal. More recently, Hsu and Holmes (2009) have followed a similar approach and have extended it to the case of several sectors. Farmers are fixed and uniformly distributed along the real line, whereas workers are mobile. Consumers have perfectly inelastic demands up to some reservation prices; firms operate under increasing returns and price discriminate across consumers through goods delivery. Hsu (2009) then shows that the urban hierarchy principle holds once firms producing different goods face different fixed production costs.

All these papers build on spatial competition theory and focus on partial equilibrium. In contrast, Fujita et al. (1999) use the framework of new economic geography and deal with general equilibrium. As the population increases, they show that a more or less regular hierarchical central place system emerges within the economy. The urban hierarchy that emerges from their simulations is more involved than in Christaller: horizontal relations are superimposed onto the pyramidal structure of central place theory because cities supply differentiated products. Our setting belongs to the same strand of literature and supplements their analysis by investigating the role of another fundamental determinant of the spatial pattern of activities, i.e. transport costs. Before proceeding, we want to stress the fact that the urban hierarchical principle is different, though not independent, from the Zipf Law. As a result, our paper should *not* be viewed as a new attempt to provide microeconomic foundations to this law. Therefore, there is no need to discuss here the literature devoted to this lively research topic.

The rest of the paper is organized as follows. The model with several industries is described in the next section. In Section 3, we characterize the spatial–sectoral equilibria. Section 4 is devoted to the derivation of the hierarchical urban principle in the cases of one and several industries. Section 5 concludes.

2. The model

The spatial economy is described by a circumference $(0, 1]$ of length 1. There are two production factors, one being immobile and the other mobile across space. It is convenient to think of them as being farmers and workers. In the agricultural sector, a homogenous good (e.g., rice) is produced under constant returns, perfect competition and zero transport cost; this good is taken as the numéraire. The economy involves a given number $I \geqslant 1$ of manufacturing sectors, which differ according to consumer expenditure share and the elasticity of substitution across product varieties. Each sector produces a differentiated good under increasing returns, monopolistic competition and positive transport costs; the array of varieties supplied by an industry varies with the mass of workers in this particular industry. Note, however, that the above interpretation of production factors and sectors is not necessary for our results to hold; it is made for expositional convenience only.

One may wonder why shipping the agricultural good is assumed to be costless, while shipping manufactured goods is costly. Recall that our primary purpose is to investigate how decreases in the transport costs of manufactured goods produced in cities affects the structure of the urban system. In order to isolate this effect, we have chosen to work with a setting in which farmers' wages are equalized across space; this is guaranteed by the assumption of zero transport cost for the agricultural good. We acknowledge the fact that this assumption is restrictive since decreasing transport costs for manufactured goods below some threshold, while preserving those of agricultural goods, stops the concentration process and leads to the redispersion of manufacturing firms and population (Fujita and Mori, 2005). One should keep in mind, however, that both types of transport costs have actually

---

[2] Using the data gathered by Eaton and Eckstein (1997), the correlation is positive for France and Japan for much longer time spans, that is, 1876–1990 in France and 1925–1985 in Japan.

declined. This implies that the urban pattern is determined by the *relative* decline in these costs (Picard and Zeng, 2005). In such a context, the analysis becomes much more involved.

The size of the workforce is fixed and split according to given shares between farmers and workers. Each farmer/worker supplies inelastically one unit of labor. Both units of labor may then be chosen for the exogenously given masses of farmers and workers to be equal to $1 - \mu > 0$ and $\mu > 0$, respectively. Farmers are immobile and uniformly distributed across space with a density equal to $1 - \mu$. In other words, we do not allow for rural–urban migration. Admittedly, this is a fairly restrictive assumption as such migration was one the main engines of urban growth at the time of the Industrial Revolution. Our line of defense is that the spatial immobility of farmers provides a simple dispersion force preventing the agglomeration of all industries into a single giant city, a clearly unrealistic pattern. As discussed in our concluding section, allowing some farmers to move and to become workers would reinforce our results.

Manufacturing firms set up in cities. A city exists as long as it accommodates a positive mass of workers so that the number $C \geqslant 1$ of cities is variable. Furthermore, cities may differ in their size and industrial mix: the mass $\lambda_c^i \geqslant 0$ of workers in industry $i = 1, \ldots, I$ and city $c = 1, \ldots, C$ is endogenous. (Throughout the paper, industries are described by the superscript i and cities by the subscript c.) A firm supplies a single variety and a variety is produced by a single firm. This implies that each variety is produced in a single city. It is worth stressing, however, that this result does not run against the urban hierarchical principle because this one relies on the supply of *goods*, not of specific varieties, in a nested system of cities. Finally, the number C of cities, their location, denoted $x_c \in (0,1]$, and size, given by $\sum_{i=1}^{I} \lambda_c^i$, are all endogenous. Firms and workers do not use land, so that cities have no spatial extension. Last, since workers can migrate freely across cities and industries, the variables λ_c^i must satisfy the constraint $\sum_{i=1}^{I}\sum_{c=1}^{C}\lambda_c^i = \mu$.

Technologies are as follows. In order to produce one unit of the homogenous good, one unit of farmer's labor is needed. As the technology exhibits constant returns, the equilibrium wage equals 1 in the agricultural sector. The initial endowment $\overline{H}$ of the good is assumed to be sufficiently large for its equilibrium consumption to be positive for each consumer. In the manufacturing sector, to produce one unit of a variety of the differentiated good i, a firm needs a fixed requirement of $f^i > 0$ workers and a marginal requirement of $m^i > 0$ units of the numéraire (e.g., workers eat rice). If n_c^i denotes the mass of varieties of good i produced in city c, city labor market clearing thus implies:

$$f^i n_c^i = \lambda_c^i \quad \text{for all } c \text{ and } i. \tag{1}$$

As illustrated by the famous "Do people follow jobs or do jobs follow people?", there is no established theory about workers' residential and occupational choices. Yet, empirical evidence suggests that both are tied through an egg-and-chicken relation. This is why we consider a *simultaneous* choice process. However, because location and job decisions are often sticky, we find it reasonable to consider a two-stage setting. The second stage outcome is interpreted as a short-run equilibrium in which firms chooses prices, while workers and farmers choose their consumption of the differentiated and homogeneous goods. The outcome of the first stage may then be viewed as a long-run equilibrium, in which workers choose simultaneously where to live and which job to take. As usual, we determine the sub-game-perfect Nash equilibrium by backward induction.

2.1. Stage II: prices and profit maximization

Stage II is fairly standard. Firms select prices so as to maximize profits conditional upon consumers' demands whose distribution

across cities and sectors is given. Consumers's preferences are given by

$$U = \sum_{i=1}^{I} \alpha^i \log Q^i + H \tag{2}$$

with

$$Q^i \equiv \left[\int_0^{n^i} q^i(v)^{\frac{\sigma^i-1}{\sigma^i}} \mathrm{d}v \right]^{\frac{\sigma^i}{\sigma^i-1}}$$

and $\alpha^i > 0$. In (2), H is quantity of the homogenous good, Q^i the quantity of the composite good associated with industry i, $q_c^i(v)$ the consumption of variety v produced by a firm belonging to industry i and located in city c, whereas n^i is the mass of varieties supplied by industry i. The parameter $\sigma^i > 1$ is the elasticity of substitution between good i-varieties, and is a direct measure of the degree of competitiveness of the corresponding industry: the larger σ^i, the more competitive industry i. Furthermore, the fact that Q^i displays a CES-form means that consumers have a *preference for variety* for each differentiated good. Finally, the strict concavity of the utility with respect to variables Q^i implies that consumers also have a *preference for diversity* across the different goods.

Because a quasi-linear utility abstracts from the income effect in the demand for the differentiated good, our modeling strategy has a partial equilibrium flavor. However, it does not remove the interaction between product and labor markets, thus allowing us to develop a full-fledged general equilibrium model of central places. Note also that our assumption on preferences does not impose any restriction on the share of industries in consumers' expenditure.

The budget constraint of a consumer living at $x \in (0,1]$ is as follows:

$$\sum_{i=1}^{I} \sum_{c=1}^{C} \int_0^{n_c^i} p^i(v;x_c,x) q^i(v;x_c,x)\mathrm{d}v + H = w(x) + \overline{H}$$

where $p^i(v;x_c,x)$ is the delivered price at x of variety v of good i produced in city c at x_c, $q^i(v;x_c,x)$ is the consumption at x of variety v of good i produced in city c at x_c, and $w(x)$ the income she earns at location x. Due to symmetry in preferences, all i-firms set up in the same city choose the same price; thus, we may drop the variable v in what follows.

The maximization of utility (2) yields the following individual demand in location x for a variety of good i produced in location x_c:

$$q^i(x_c,x) = \alpha^i \frac{p^i(x_c,x)^{-\sigma^i}}{P^i(x)^{-(\sigma^i-1)}}$$

where $p^i(x_c,x)$ is the common delivered price of a variety of good i produced in city c at x_c and consumed at x and where

$$P^i(x) = \left[\sum_{c=1}^{C} n_c^i p^i(x_c,x)^{-(\sigma^i-1)} \right]^{\frac{-1}{\sigma^i-1}} \tag{3}$$

is the price index of good i that prevails at location x. Accordingly, the indirect utility of a consumer residing in location x and working for industry i is as follows:

$$V^i(x) = \sum_{j=1}^{I} \alpha^j \left[\log \alpha^j - 1 - \log P^j(x) \right] + w^i(x_c)$$

where $w^i(x_c)$ is the wage that i-firms set up in city c pay to their workers.

Let $\tau^i(x_c,x) \equiv m^i(\tau^i)^{\min\{|x_c-x|,1-|x_c-x|\}}$ be the number of units of the numéraire that a firm producing good i and located in city c at x_c has to bear to produce and ship one unit of its output to location

x. In this expression, $m^i > 0$ stands for the marginal production cost of good *i*. Without loss of generality, we may choose the unit of good *i* for m^i to be equal to 1. The parameter $\tau^i > 1$, which we call the *transport rate*, accounts for all the impediments to trading good *i* across locations. To capture the general decline in transport costs in a simple way, we assume that $\tau^i = \tau > 1$ for all $i = 1, \ldots, I$ and study how decreases in τ affect the spatial and sectoral distribution of activities.

Because workers have chosen where to live in stage I, the population of city *c*

$$\Lambda_c \equiv \sum_{j=1}^I \lambda_c^j$$

is given. The profit of a firm belonging to industry *i* and located in city *c* is then defined as follows:

$$\pi^i(x_c) = \sum_{d=1}^C \Lambda_d \left[p^i(x_c, x_d) - \tau(x_c, x_d) \right] q^i(x_c, x_d) + (1 - \mu)$$
$$\times \int_0^1 \left[p^i(x_c, x) - \tau(x_c, x) \right] q^i(x_c, x) \mathrm{d}x - f^i w^i(x_c) \quad (4)$$

where firms are price-makers on good *i*-markets. In contrast, because they have no monopsony power, firms are wage-takers on urban labor markets. The maximization of (4) yields the equilibrium price:

$$p^i(x_c, x) = \frac{\sigma^i}{\sigma^i - 1} \tau(x_c, x) \quad \text{for all } c \text{ and } i \quad (5)$$

Hence, the price index (3) becomes

$$P^i(x) = \left[\frac{\sigma^i}{\sigma^i - 1} \sum_{c=1}^C n_c^i \phi^i(x_c, x) \right]^{\frac{-1}{\sigma^i - 1}} \quad \text{for all } i \quad (6)$$

where

$$\phi^i(x_c, x) \equiv \tau(x_c, x)^{1-\sigma^i} \in (0, 1]$$

measures the accessibility of a consumer residing at *x* to any variety of good *i* produced in city *c*. The accessibility is maximized when transport costs are zero, that is, when $\tau = 1$, and minimized when $\tau \to \infty$. Having this in mind, the price index (6) may then be interpreted as the inverse of consumer's accessibility to the whole range of varieties of good *i* produced in all cities. Clearly, when more firms belonging to industry *i* are set up in cities situated in the vicinity of the consumer, the corresponding price index decreases because transport costs are lower.

Workers' wages are determined through a bidding process in which firms belonging to industry *i* and set up in city *c* compete to hire them. As a result, the wage bill of a firm is less than ($\lambda_c^i = 0$) or equal to ($\lambda_c^i \geqslant 0$) its operating profits. Plugging (5) into (4) and using the zero-profit condition together with the labor market equilibrium condition (1), we get the equilibrium wage paid by the *i*-firms at location *x*:

$$w^i(x) = \frac{\alpha^i}{\sigma^i} \left[\sum_{c=1}^C \frac{\Lambda_c \phi^i(x, x_c)}{\sum_{d=1}^C \lambda_d^i \phi^i(x_d, x_c)} + \int_0^1 \frac{(1-\mu)\phi^i(x, x_a)}{\sum_{d=1}^C \lambda_d^i \phi^i(x_d, x_a)} \mathrm{d}x_a \right] \quad (7)$$

where x_c denotes city *c*-workers' location, x_d city *d*-firms' location, and x_a the locations of farmers. Observe that $w^i(x)$ depends on the whole distribution of workers across cities (Λ_c) as well as on the distribution of industry *i*-workers among cities (λ_d^i). In addition, wages may vary across cities and industries because firms do not necessarily earn the same operating profits.

Substituting (6) and (7) for $P^i(x)$ and $w^i(x)$ in $V^i(x)$ allows us to rewrite the indirect utility of a consumer working in industry *i* and living in city *c* as a function of the spatial distribution of each industry λ_d^j weighted by the accessibility coefficients ϕ^j:

$$V^i(x_c) = \sum_{j=1}^I \frac{\alpha^j}{\sigma^j - 1} \log \left[\sum_{d=1}^C \lambda_d^j \phi^j(x_c, x_d) \right]$$
$$+ \frac{\alpha^i}{\sigma^i} \left[\sum_{e=1}^C \frac{\Lambda_e \phi^i(x_c, x_e)}{\sum_{d=1}^C \lambda_d^i \phi^i(x_d, x_e)} + \int_0^1 \frac{(1-\mu)\phi^i(x_c, x_a)}{\sum_{d=1}^C \lambda_d^i \phi^i(x_d, x_a)} \mathrm{d}x_a \right]$$
$$(8)$$

in which we have dismissed an additive constant which has no impact on workers' locational decisions. The indirect utility a worker would reach by working in an industry *i* firm located at *x*, $V^i(x)$, can be expressed by replacing x_c with *x* in (8).

The indirect utility (8) encapsulates the agglomeration and dispersion forces at work in our setting. The first term of (8) acts as an agglomeration force because it has the nature of a market potential: a higher accessibility to the whole array of goods makes city *c* a more attractive place to live in. The second term involves both an agglomeration and a dispersion force. Indeed, a more concentrated pattern of firms belonging to the same industry in city *c* leads to lower operating profits for these firms, thus leading them to pay lower wages to their workers (the dispersion force). However, because the agglomeration of firms belonging to different industries raises the expenditure $\Lambda_e - \lambda_e^i$, operating profits and wages in industry *i* tend to be higher (the agglomeration force). Last, the third term stems from the fact that farmers are immobile and uniformly distributed across places. This is the main dispersion force in our setting since the concentration of industries in a few big cities raises the cost of shipping the manufactured goods to all farmers.[3] Consequently, we may conclude that the expression (8) encapsulates the trade-off that a consumer faces when she chooses where to live and to work.

2.2. Stage I: industry and location choice

In the first stage, a worker has to choose the industry and the city that give her the highest utility level, anticipating her consumption of all goods. Because workers are identical, in the long-run equilibrium, they must reach the same utility level. Hence, a *spatial–sectoral equilibrium* is such that there exists a constant $\overline{V}$ for which

$$V^i(x) \leqslant \overline{V} \quad \text{and} \quad [V^i(x) - \overline{V}]\lambda_c^i = 0 \quad \text{for all } i, \ c \text{ and } x \in (0, 1] \quad (9)$$

Because workers are free to live in a new city located at $x \in (0, 1]$, it must be that $V^i(x) < \overline{V}$ for all $x \neq x_c$. The $C \cdot I$ Eqs. (9) then allow for the determination of the $C \cdot I$ shares $\lambda_c^i \geqslant 0$. Since (8) is continuous with respect to λ_c^i, it follows from Ginsburgh et al. (1985) that, for any given *C*, a spatial–sectoral equilibrium exists.

Because all workers living in city *c* face the same prices, the intersectoral mobility of workers implies that workers earn the same wage across all the industries established in city *c*. Stated differently, there exists a positive constant $\bar{w}_c$, the city-wage, such that

$$w_c^i \leqslant \bar{w}_c \quad \text{and} \quad (w_c^i - \bar{w}_c)\lambda_c^i = 0 \quad \text{for all } i \quad (10)$$

where $w_c^i \equiv w^i(x_c)$ to ease the burden of notation. Hence, industry *i* is not active in city *c* ($\lambda_c^i = 0$) if and only if the highest wage (i.e. the zero-profit wage (7)) that the corresponding firms are willing to pay is lower than the city-wage $\bar{w}_c$. The *industrial mix* of city *c* is then

---

[3] Note that this dispersion force could be replaced by a congestion cost that would increase with the mass of workers living in a city (for example, through commuting costs and workers' land consumption). This would not affect our main results, but would render the analysis more involved.

formed by all the industries for which $\lambda_c^i > 0$, that is, those having a zero-profit wage equal to the city-wage. That not all industries need be present in city c is a mere reflection of the fact that the costs and benefits of being located in that city are not the same across industries.

Like most settings involving increasing returns and monopolistic competition, ours involves multiple equilibria. This leads us to retain stability as an equilibrium refinement. Following a well-established tradition in economic geography, we focus on an adjustment process in which workers spread themselves among several industries and cities, being attracted (repelled) by industries and cities providing them with high (low) utility levels. This formulation aims at capturing the fundamental idea that workers' reactions are both sluggish and heterogenous because of differences in the perception of places and occupations. Therefore, the stability analysis is conducted by means of the replicator dynamics of the $C \cdot I$ variables:

$$\dot{\lambda}_c^i = \frac{\lambda_c^i}{\mu}\left(V_c^i - \sum_{j=1}^{I}\sum_{d=1}^{C}\frac{\lambda_d^j}{\mu}V_d^j\right) \equiv J_c^i \quad \text{for all } c \text{ and } i \tag{11}$$

where $\cdot$ denotes the time derivative of λ_c^i and $V_c^i \equiv V^i(x_c)$. In words, the growing cities and industries are those that give workers a high utility as compared to the average utility. The existence and stability of equilibrium are analyzed in the next section. On this occasion, we will encounter a major difficulty: both the number of cities and the number of industries present in a city vary with the transport cost level τ. The replicator holds for a *given* number C of cities and does not allow one to study the transition from a pattern involving C cities to another displaying a different number of cities, as in Fujita et al. (1999).[4] Therefore, it is important to keep in mind that we use (11) as a refinement of the equilibrium when the dimension of the dynamic system is given.

In what follows, we will focus on two special, but relevant, configurations that will be shown to be spatial–sectoral equilibria. The first one involves C equidistant cities, with $C = 2, 4, 8, 16, \ldots$, each having the same size and industrial mix.[5] All industries are, therefore, present in each city; however, each city is specialized in a different set of varieties, which leads to intra-industry and inter-city trade. The second configuration is described by a set of nested cities that have different sizes and industry structures. The biggest cities have all industries; as one moves down the hierarchy, cities get smaller and accommodate only a subset of the industries present in higher-level cities, thus implying the existence of inter-industry trade flows from the high-level to the low-level cities. It should be clear that such a configuration corresponds to the urban hierarchy principle discussed in the introduction.

3. The nature of spatial–sectoral equilibria

As in many nonlinear models, the equilibrium need not be symmetric, unique or stable, thus making the analysis much more involved than in the second stage. In addition, because our setting involves both increasing returns and monopolistic competition, its analysis cannot give unique-equilibrium comparative statics results of the kind we are used to. The results presented in the rest

of this section are valid for any even number of cities. However, in order to keep the analysis consistent with what will be done in the next section, we assume that the number of cities is such that $C = 2^K$. In the following three sections, we investigate changes in the number and sizes of cities when (i) the number of cities is divided by two, (ii) the symmetric equilibrium breaks, and (iii) city sizes differ.

3.1. The doubling point

Our first task is to show that there exists a critical level of transport costs such that the symmetric configuration involves $2C$ cities when τ is slightly above this threshold, but only C cities when it is slightly below. A *symmetric spatial–sectoral equilibrium* with 2^K cities, $K = 1, 2, \ldots$, is such that all cities are located equidistantly and have the same population size. It remains to show that the share of each industry is the same across cities. By solving the sectoral equilibrium condition $V^i = \overline{V}$ for all i, the equilibrium share is uniquely determined as follows:

$$\lambda_c^i = \lambda^i \equiv \theta^i \frac{\mu}{C} \quad \text{for all } c \text{ and } i \tag{12}$$

where

$$\theta^i \equiv \frac{\alpha^i/\sigma^i}{\sum_j \alpha^j/\sigma^j}$$

is the share of industry i in a city (because cities are identical, we have deleted the subscript c). The share of an industry decreases with the elasticity of substitution (σ^i) because the corresponding market is more fragmented, whereas it increases with its salience coefficient in preferences (α^i). Hence, all cities supply the whole array of goods and the same number $(\theta^i\mu)/(f^iC)$ of varieties of each good; they have the same size, the city population μ/C being distributed between the different industries according to the shares θ^i. At a symmetric equilibrium, cities are, therefore, not specialized. However, as varieties are differentiated, there is intra-industry and inter-city trade of each good i.

For (12) to be a spatial equilibrium with C cities, it must be that $\overline{V} \geqslant V^i(x)$ holds for all i and all x. To show this, we first prove the following result in Appendix A.

Lemma 1. *In the symmetric configuration with C cities, the indirect utility $V^i(x)$ is maximized either at city locations or at midpoints of two adjacent cities.*

Accordingly, (12) is a spatial equilibrium if and only if $V^i(0) \geqslant V^i(1/2C)$ holds for all i. This result also tells us something interesting for our purpose, that is, when the symmetric equilibrium pattern involves C cities, perturbing this equilibrium yields a pattern involving C or $2C$ cities.

Furthermore, we show in Appendix B that there exists a function $g(\cdot)$ such that the equilibrium condition $V^i(0) \geqslant V^i(1/2C)$ is equivalent to

$$g(\varphi^i; \mu, \sigma) \geqslant 0 \quad \text{for all } i$$

where

$$\varphi^i \equiv (\phi^i)^{\frac{1}{C}} \equiv (\tau^i)^{\frac{1-\sigma^i}{C}} \in (0, 1].$$

For any even number C and for each i, it is also shown in this appendix that $g(\varphi^i) = 0$ has at least one interior solution. Let $\tau_2^i(C)$ be the largest of these solutions and set

$$\tau_2(C) \equiv \min_i \tau_2^i(C)$$

which we call the *doubling point* for a reason that will become clear below. Thus, as long as $\tau < \tau_2(C)$, we have $V^i(0) > V^i(1/2C)$, which implies that no worker has an incentive to locate at the midpoint

of any two adjacent cities $x = 1/2C, x = 3/2C, \ldots, x = (2C − 1)/2C$. Conversely, when $\tau > \tau_2(C)$, the configuration with C cities is not an equilibrium because $V^i(1/2C)$ now exceeds $V^i(0)$.

To sum-up, we have:

Proposition 1. *For all $C = 2^K$, there exists a unique threshold $\tau_2(C) > 0$ such that the equilibrium involves $2C$ equidistant cities when the transport rate exceeds $\tau_2(C)$. Furthermore, when the transport rate is slightly below $\tau_2(C)$ the equilibrium number of cities at the symmetric configuration is equal to C.*

To put it differently, if there are 2^K equidistant cities for some value of τ, the number of cities is divided by 2 once τ has sufficiently decreased. Since this holds for any value of K, we may conclude that there exist equilibria involving $2^K, 2^{K-1}, \ldots$ cities when the cost of shipping manufactured goods keeps decreasing. Eventually, when τ is low enough, there is an equilibrium involving a single city accommodating all firms and workers. This explains why $\tau_2(C)$ has been called a "doubling point." We must stress, however, that Proposition 1 does not aim to describe the transition path at $\tau_2(C)$. Instead, we restrict ourselves to a comparison of the number and locations of cities above and below $\tau_2(C)$.

3.2. The break point

Our next task is to study the conditions under which the foregoing symmetric pattern is stable. One of the most common equilibrium refinements is asymptotic stability, the equilibrium (12) being stable if the real parts of all the eigenvalues of the Jacobian matrix of (11) are negative. This appears exceedingly difficult to handle for an arbitrary number of cities.

This is why we propose to use the weaker concept of asymptotic stability for the decomposed replicator dynamics with I variables and C variables, for which we can compute analytically the eigenvalues. Specifically, we decompose the dynamics (11) with $C \cdot I$ variables into two sub-systems:

$$\dot{\lambda}_c^i = \frac{\lambda_c^i}{\Lambda_c}\left(V_c^i - \sum_{j=1}^{I}\frac{\lambda_c^j}{\Lambda_c}V_c^j\right) \equiv J^i \quad \text{for all } i \qquad (13)$$

with I variables for any given city c and

$$\dot{\lambda}_c^i = \frac{\lambda_c^i}{\Lambda^i}\left(V_c^i - \sum_{d=1}^{C}\frac{\lambda_d^i}{\Lambda^i}V_d^i\right) \equiv J_c \quad \text{for all } c \qquad (14)$$

with C variables for any given industry i, $\Lambda^i \equiv \sum_{c=1}^{C}\lambda_c^i$ denoting the total mass of workers in industry i. We then study the asymptotic stability of each of these two dynamics. More precisely, when a spatial–sectoral equilibrium is perturbed, this is either by changing workers' occupation or place, but not both simultaneously. In such a context, the dynamics (13) describes intersectoral mobility within a city during a certain time period (e.g., overnight), whereas (14) describes inter-city migration of workers belonging to the same industry during another period (e.g., in daytime). Although it affects the stability conditions, this decomposition does not change the equilibrium outcome. Note also that, when the symmetric equilibrium is unstable for the decomposed replicator dynamics, it is also unstable for the simultaneous replicator dynamics since this one allows for a broader range of deviations. Recall that stability is used here an equilibrium refinement; it does not attempt to describe workers' off-equilibrium behavior. Thus, what we said above about (11) also applies to (13) and (14).

Regarding the dynamics (13), our main result is as follows (the proof is given in Appendix C).

Lemma 2. *When there is no inter-city migration, the intersectoral mobility described by (13) is globally stable.*

This implies that the spatial–sectoral equilibrium is always stable once we consider only changes in occupation.

Let us now focus on (14) whose analysis is less straightforward. If

$$\mu < \min_i \frac{\sigma^i - 1}{\theta^i(2\sigma^i - 1)} \quad \text{for all } i \qquad (15)$$

we show in Appendix D that there exists a critical level $\tau_1(C)$, called the *break point*, such that the following result holds.

Proposition 2. *Let $C = 2^K$ be the initial number of cities. If (15) holds, then there exists a unique threshold $\tau_1(C) > 0$ for which the symmetric equilibrium with C cities is stable (resp., unstable) when the transport rate is higher (resp., lower) than $\tau_1(C)$.*

Thus, as long as τ belongs to the interval $(\tau_1(C), \tau_2(C))$. Propositions 1 and 2 imply that *there exists a symmetric stable equilibrium with $C = 2^K$ cities*. However, for this to arise, it must be that $\tau_1(C) < \tau_2(C)$. Note that other stable equilibria may exist in this interval, but they are not studied here.

Note also that Appendices C and D imply that both $\tau_1(C)$ and $\tau_2(C)$ increase with C. Therefore, we may safely conclude that the symmetric equilibrium path (if any) must involve a decreasing number of cities when transport costs steadily decline.

It remains to explain the role played by condition (15). When the share $1 − \mu$ of the agricultural sector in the economy is small, the corresponding dispersion force becomes too weak to prevent the emergence of a unique metropolis involving all firms and workers, hardly a plausible pattern.[6] Therefore, the parameter μ to be bounded above, thus showing the role played by farmers in our setting.

3.3. When do city sizes differ?

At this stage, the following question naturally comes to mind: what happens when τ crosses the break point $\tau_1(C)$ from above? The symmetric pattern with $C = 2^K$ cities being now unstable, several scenarios are conceivable. In what follows, we focus on the case in which the equilibrium involves the same number of cities, but these ones now have different sizes. Such a *selection* is justified by the fact that cities, unlike the shanty towns of the gold rush, show an amazing resilience and seldom disappear suddenly. Such a research strategy strikes us as being much more realistic than the catastrophic changes stressed in the new economic geography literature (Krugman, 1991).

The solution to the equilibrium condition (9) then yields what we call an *alternating equilibrium*, which is defined as follows:

$$\lambda_{2c}^i = \lambda \quad \text{and} \quad \lambda_{2c-1}^i = 2\lambda^i - \lambda \quad \text{for } \lambda \in [0, 2\lambda^i] \quad \text{and}$$
$$c = 1, 2, \ldots, C/2 \qquad (16)$$

where λ^i is given by (12). At such a configuration, cities with different sizes and industrial mixes coexist (this is so when λ is equal to $2\lambda^i$ for some i), a small city alternating with a big city. Thus, *the same spatial pattern involving cities of different sizes is repeated*, as in the urban hierarchical principle. However, the market outcome always involves two types of cities.

The alternating configuration (16) corresponds to a pitchfork bifurcation if and only if the following conditions hold at the break point (Rasband, 1990, p. 31):

---

[6] Note that (15) is the counter-part of the no-black-hole condition derived in the standard core-periphery model (Krugman, 1991).

$$dV^i(\lambda^{i*} - \lambda) = -dV^i(\lambda^{i*} + \lambda)$$

$$h(\varphi^i; \mu, \sigma) = \frac{\partial dV^i(\lambda)}{\partial \lambda}\bigg|_{\lambda=\lambda^{i*}, \tau=\tau_1(C)} = 0$$

$$\frac{\partial^2 dV^i(\lambda)}{\partial \lambda \partial \tau}\bigg|_{\lambda=\lambda^{i*}, \tau=\tau_1(C)} > 0$$

$$l(\varphi^i; \mu, \sigma) \equiv \frac{\partial^3 dV^i(\lambda)}{\partial \lambda^3}\bigg|_{\lambda=\lambda^{i*}, \tau=\tau_1(C)} < 0 \qquad (17)$$

for all i, the functions $h(\cdot)$ and $l(\cdot)$ being defined in Appendices D and E, respectively. Alternating symmetry implies that the first and third conditions are always met. The last condition (17) ensures that the symmetric equilibrium gets unstable, whereas the asymmetric ones become stable when τ is slightly lower than $\tau_1(C)$. Hence, $h(\varphi^i; \mu, \sigma) = 0$ and $l(\varphi^i; \mu, \sigma) < 0$ appear to be the critical conditions for the existence of an alternating equilibrium. Whether or not they hold depends on the parameter values. This is the topic we investigate in the next section.

4. The urban hierarchy principle

In what follows, we have chosen to show that the urban hierarchical principle may be sustained as an equilibrium outcome. By recognizing the existence of multiple equilibria, our setting provides a broader perspective of the urban hierarchy than central place theory since it recognized explicitly that alternative spatial arrangements may emerge. However, characterizing the whole class of stable equilibria is beyond the scope of this paper.

4.1. The one-industry case

To start with, we consider the case of a single industry. Preferences become:

$$U = \log Q + H$$

where

$$Q \equiv \left[\int_0^n q(v)^{\frac{\sigma-1}{\sigma}} dv\right]^{\frac{\sigma}{\sigma-1}}$$

The symmetric equilibrium with C cities is given $\lambda = \mu/C$. As seen in the foregoing, this equilibrium turns into the alternating equilibrium when both the break condition $h(\varphi; \mu, \sigma) = 0$ and the bifurcation condition $l(\varphi; \mu, \sigma) < 0$ hold. Consider the two equations

$$h(\varphi; \mu, \sigma) = 0 \quad \text{and} \quad l(\varphi; \mu, \sigma) = 0 \qquad (18)$$

that delineate the corresponding parameter domain. The system of Eq. (18) has no explicit solutions. Nevertheless, we may characterize the domain we seek through implicit expressions.

Formally, we must distinguish between the two cases in which $C = 2^2, 2^3, \ldots$ and $C = 2$ because the system (18) takes two different forms. In both cases, each equation of the system (18) is bilinear in μ and σ so that this system has a unique solution. In the former case, we show in Appendix E that the solution is given by $\sigma_2(\varphi)$ and $\mu_2(\varphi)$. Because $\mu_2(\varphi)$ is shown to be monotone in the relevant range, we may construct the inverse mapping $\varphi_2^{-1}(\mu)$ of $\mu_2(\varphi)$ and define the new function $\sigma_2(\mu) = \sigma_2[\varphi_2^{-1}(\mu)]$. We construct the functions $\sigma_2(\varphi)$ and $\mu_2(\varphi)$ in a similar way for $C = 2$. The implicit function $\sigma_2(\mu)$ is drawn in Fig. 1 for $C = 2^2, 2^3, \ldots$ The function $\sigma_2(\mu)$ for $C = 2$ is slightly southwest of that for $C = 2^2, 2^3, \ldots$

Thus, we have the following result.

Lemma 3. *Assume that there is one industry. If $\sigma < \sigma_2(\mu)$, then the symmetric equilibrium with C cities is unstable for $\tau < \tau_1(C)$ and bifurcates to an alternating equilibrium at $\tau_1(C)$.*
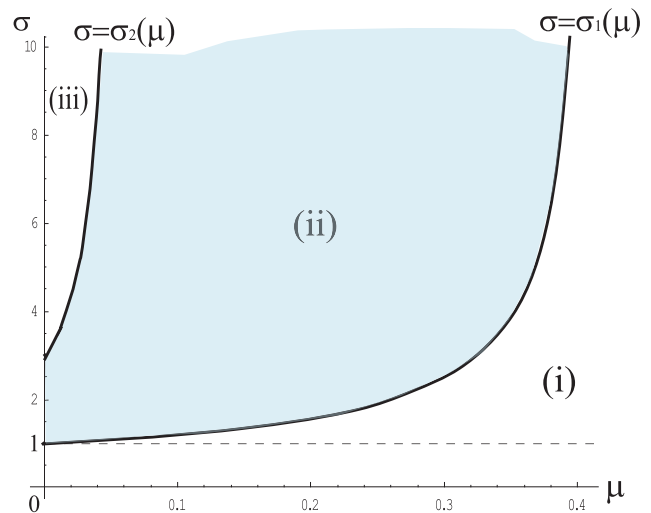

Fig. 1. Domains of (i) no symmetric equilibrium, (ii) pitchfork bifurcation, and (iii) tomahawk bifurcation.

Hence, when the value of transport costs steadily decreases and crosses the break point $\tau_1(C)$ from above, every other city grows, while the size of the others decreases.

Consider now the two equations $g(\varphi; \mu, \sigma) = 0$ and $h(\varphi; \mu, \sigma) = 0$ defining the domain for which the symmetric configuration with C cities is a stable spatial equilibrium, a city size being given $\lambda = \mu/C$. Repeating the foregoing argument, we come to the values $\mu = \mu_1(\varphi)$ and $\sigma = \sigma_1(\varphi)$. It is then readily verified that (15), in which $\theta = 1$ since there is a single industry, is always satisfied whenever $\sigma > \sigma_1(\mu)$. The implicit function $\sigma_1(\mu)$ is depicted in Fig. 1 for $C = 2^2, 2^3, \ldots$ The function $\sigma_1(\mu)$ for $C = 2$ is slightly southwest of that for $C = 2^2, 2^3, \ldots$ Hence, we have:

Lemma 4. *Assume that there is one industry and that $\tau_1(C) < \tau < \tau_2(C)$. If $\sigma > \sigma_1(\mu)$ (resp., $\sigma < \sigma_1(\mu)$), then a stable symmetric equilibrium with C cities exists (resp., does not exist).*

Lemmas 3 and 4 show under which conditions the symmetric equilibrium is stable over $\tau_1(C) < \tau < \tau_2(C)$, but becomes unstable at $\tau = \tau_1(C)$ where it bifurcates to an alternating equilibrium. Since $\sigma_1(\mu) < \sigma_2(\mu)$, the following three cases may arise. Case (i) involves no symmetric equilibrium: if $\sigma < \sigma_1(\mu)$, a stable symmetric equilibrium never exists regardless of the value of C. Case (ii) displays a pitchfork bifurcation: if $\sigma_1(\mu) < \sigma < \sigma_2(\mu)$, there exists a stable symmetric equilibrium involving C cities as long as $\tau_1(C) < \tau < \tau_2(C)$; this equilibrium smoothly bifurcates to an alternating equilibrium at $\tau_1(C)$. Case (iii) displays tomahawk bifurcation: if $\sigma > \sigma_2(\mu)$, there exists a stable symmetric equilibrium for $\tau_1(C) < \tau < \tau_2(C)$. This equilibrium breaks, but does not smoothly bifurcate to an alternating equilibrium at $\tau_1(C)$ for all value of C so that different outcomes may emerge. For example, when a symmetric equilibrium with $C = 8$ breaks, the new equilibrium may be a symmetric equilibrium with $C = 4$ or an asymmetric equilibrium with $C = 6$–8. Hence, we will focus on case (ii) because it guarantees smooth bifurcations. It is described by the shaded area of Fig. 1. Furthermore, simulations reveal that, at the alternating equilibrium, λ increases not just for τ slightly lower than $\tau_1(C)$, but also when τ decreases from $\tau_1(C)$ to $\tau_2(C/2)$.

We may state these results a way that fits better central place theory.

Proposition 3. *Assume that $\sigma_1(\mu) < \sigma < \sigma_2(\mu)$ and that τ is sufficiently large for the economy to have a stable symmetric equilibrium involving $C = 2^K$ cities with $K > 1$. Then, when the transport rate*

τ steadily decreases, the urban system is described by the sequence formed by the following two patterns with k decreasing from K to 2: (i) an alternating equilibrium with 2^k cities for $\tau_2(2^{k-1}) < \tau < \tau_1(2^k)$; (ii) a symmetric equilibrium with 2^{k-1} cities for $\tau_1(2^{k-1}) < \tau < \tau_2(2^{k-1})$. Finally, once the value k = 1 is reached, there is partial agglomeration in one city for $\tau_2(1) < \tau < \tau_1(2)$ and full agglomeration for $0 < \tau < \tau_2(1)$.

We must first clarify the meaning of the assumption $\sigma_1(\mu) < \sigma < \sigma_2(\mu)$ made in the proposition. The first restriction $\sigma > \sigma_1(\mu)$ allows us to eliminate the extreme and unrealistic situation in which all firms and workers are agglomerated in a single city regardless of the transport cost value. The second restriction $\sigma < \sigma_2(\mu)$ ensures the existence of a smooth transition at each symmetry break point, as explained in Section 3.3. So these two restrictions have clear and well-defined economic interpretation.

To illustrate how the evolution process works, consider a stable pattern with C = 16 symmetric cities and let the transport rate τ to decrease steadily. First, the urban system remains the same until the break point $\tau_1(16)$ is reached. When τ falls slightly below $\tau_1(16)$, the symmetric configuration becomes unstable, so that the urban system now involves 16 cities that alternate in size. As τ keep falling, the big cities grow while the small cities shrink. Once $\tau_2(16)$ is reached, the small cities disappear and the urban system involves only 8 bigger and identical cities; both the intercity distance and the city size have doubled. The same pattern is repeated with 4 and 2 cities, until τ reaches the threshold $\tau_2(1)$.

Case (ii) is especially relevant to us because it provides a foundation for the urban hierarchical principle. According to Fig. 1, case (ii) requires an upper and a lower bound on the degree of product differentiation. It is worth noting that the share μ of manufacturing employment is 0.24 in OECD countries. Estimations of the elasticity of substitution σ is between 4.9 and 7.6 in the US (Hanson, 2005), whereas σ would be around 3.9 for Germany and 1.9 for Italy (Brakman et al., 2004; Mion, 2004). Fig. 1 shows that all those values belong to the shaded area, thus suggesting that case (ii) is likely to be the empirically relevant one.

For completeness, it remains to discuss what happens when the condition $\sigma_1(\mu) < \sigma < \sigma_2(\mu)$ does not hold. Case (i) bears some resemblance with the so-called black-hole condition in the core-periphery model in that the manufacturing share μ is large and the elasticity of substitution σ low (Krugman, 1991). However, the manufacturing sector does not get agglomerated within a single city regardless of the value of the transport costs because the farming population is uniformly distributed across space. Since Case (i) does not satisfy the symmetric equilibrium condition, a stable equilibrium is necessarily asymmetric, involving big and small cities. Unfortunately, such asymmetric equilibria do not satisfy the urban hierarchical principle of equally-spaced, equal-size cities. Case (iii) is the opposite to Case (i) in that the manufacturing share μ is small and the elasticity of substitution σ is high, so that a symmetric equilibrium with a certain number C of cities is likely to be stable. When τ falls below $\tau_1(C)$, the symmetric equilibrium breaks but, as noted above, we no longer have a pitchfork bifurcation. The resulting multiplicity of equilibria prevents us to determine which stable equilibrium is likely to emerge after symmetry breaking.

4.2. The multi-industry case

Consider now the case of I > 1 industries in which $\sigma^1 < \sigma^2 < \cdots < \sigma^I$. As in the foregoing section, we assume that

$$\sigma_1(\mu) < \sigma^i < \sigma_2(\mu) \quad \text{for all } i$$

to ensure that a symmetric equilibrium exists for any value of $C = 2^K$ and any value of I. The analysis developed in the above sections shows how difficult it is to solve the model analytically with a sin-

gle industry. It is, therefore, no surprise that dealing with several industries is a very hard task. Yet, it is possible to gain insights about the evolution of the urban system. In particular, insofar as manufacturing sectors are not too heterogeneous, that is, the elasticities of substitution are not too different, they react in a similar way to decreasing transport costs. In this case, Proposition 3 proved for one industry holds true in the multi-industry case.

Consider first our main thought experiment in which transport costs decrease. Starting from an alternating equilibrium in which the C cities accommodate all industries, Proposition 2 implies that this equilibrium is stable as long as τ is just below $\tau_1(C)$. In this case, even though city sizes differ, the number of active industries is the same in each city. However, the size of each industry in large cities is bigger than that in small cities. In order to gain more insights about the evolution of the industrial mix, suppose that, at the pitchfork bifurcation arising at $\tau_1(C)$, the shares λ_c^j change and are no longer identical across cities, while the shares λ_c^i remain the same for all $i \neq j$. For all industries $i \neq j$, the value of the first and second terms of the indirect utility (8) in the big cities is strictly greater than the utility level prevailing in the small cities, whereas the last term is the same in both types of cities. Therefore, the utility differential between the big and small cities is positive for all industries $i \neq j$. This implies that workers move from the small to the big cities. In other words, as soon as τ is just below $\tau_1(C)$, the big cities capture a greater share of each industry at the expense of the small ones. Consequently, as transport costs fall, the small cities lose, first, one industry, then two, and so forth until their complete disappearance from the urban system.

Formally, we have the following result (the proof is given in Appendix F).

Proposition 4. Assume that elasticities of substitution are not too different and consider a 2^k-alternating equilibrium. If transport costs decrease from $\tau_1(2^k)$ to $\tau_2(2^{k-1})$, then the biggest cities accommodate all industries, whereas the other cities accommodate a steadily decreasing number of industries from I to 0. Furthermore, a more competitive industry (i.e. a larger elasticity of substitution) is present in a larger number of cities.

Hence, firms selling very differentiated goods leave small cities to set up in big cities where they benefit from growing local markets without having to face tough competition. Simultaneously, they can supply more easily distant consumers since transport costs are lower. In contrast, firms supplying poorly differentiated goods remain dispersed across all existing cities. In doing so, they maintain their profit margins through soft competition. Finally, when transport costs are very low, spatial separation ceases to protect these firms from competition. This leads firms belonging to competitive industries to locate as well in big cities where they enjoy a good access to local markets, the size of which grows through the inflows of workers.

We are now equipped to discuss the evolution of the urban system when transport costs steadily decrease. Because cities smoothly bifurcate at the break point $\tau_1(2^k)$, the spatial–sectoral equilibrium is continuous and interior. Accordingly, the value of each indirect utility is the same for all industries in big and small cities. As transport costs decrease further, workers's utility in the industry supplying the most differentiated good becomes strictly lower than that in all the others.[7] Eventually, the least competitive industries disappear one by one from small cities, whereas the most competitive industries remain dispersed across all cities. In other

---

[7] Using the terminology of Fujita et al. (1999), the potential function of the industry having the smallest σ^i is below the line at locations of small cities, but the potential functions of other $I - 1$ industries hit the horizontal line of 1 at locations of large cities.
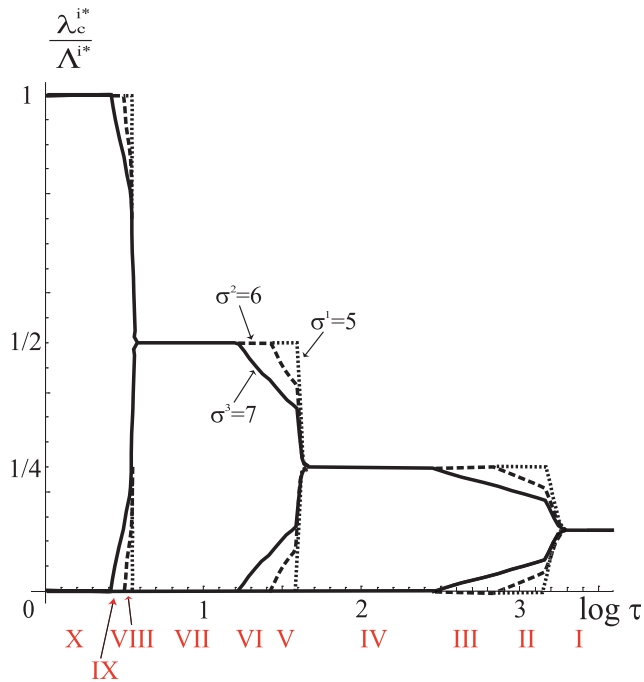
Fig. 2. Evolutionary process for transport cost changes with similar industries.

words, while 2^{k-1} big cities accommodate all industries, 2^{k-1} small cities have fewer industries.

To illustrate the foregoing argument, we resort to numerical analysis. Indeed, by using Propositions 1 and 2 as well as the argument that leads to Propositions 3 and 4, we may simulate the evolution of the urban system when there are $I > 1$ industries. To this end, we have computed the continuous equilibrium path for a wide domain of parameter values belonging to domain (ii) of Fig. 1.

Fig. 2 provides the results of a simulation for falling transport costs in the case of three industries ($I = 3$) with the initial number of cities $C = 8$. The parameter values are $\mu = 0.2$, $\alpha^1 = \alpha^2 = \alpha^3 = 1$, $\sigma^1 = 5$, $\sigma^2 = 6$, and $\sigma^3 = 7$. In order to obviate the difficulty generated by the possible existence of multiple equilibria, we have started the search from an initial distribution close to the fully symmetric distribution and have obtained a unique equilibrium for each value of τ. As seen above, the industry with the smallest elasticity of substitution is always the most concentrated (the dotted line), while the industry with the largest elasticity is the least concentrated (the bold line). In addition, the growth and decline of cities are gradual and smooth.

Our numerical findings may be summarized by means of the following steps: (**I**) a fully symmetric equilibrium involving eight cities that accommodate all industries and have the same industrial mix or an alternating equilibrium in which four big and four small cities, each supplying all goods, alternate; (**II**) an alternating equilibrium in which the four big cities keep
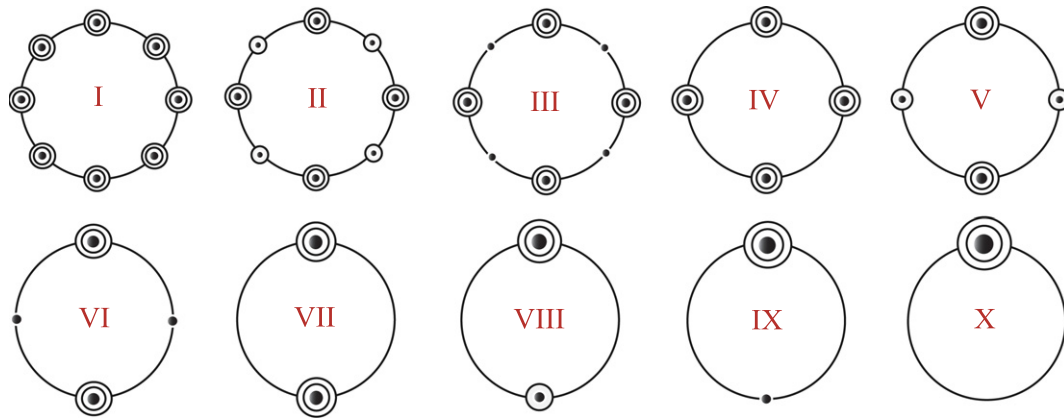

Fig. 3. Evolutionary process with similar industries ($\sigma^1 = 5$, $\sigma^2 = 6$, $\sigma^3 = 7$).
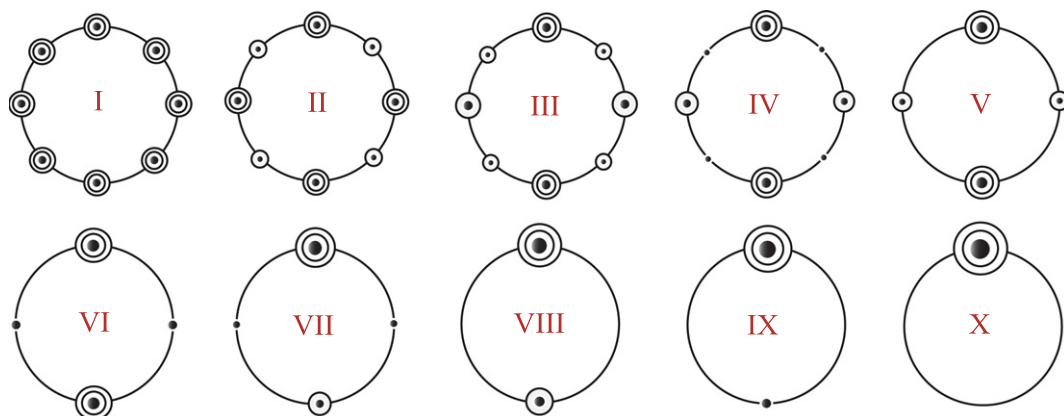

Fig. 4. Evolutionary process with different industries ($\sigma^1 = 3$, $\sigma^2 = 6$, $\sigma^3 = 9$).

accommodating all industries, whereas the most differentiated good 1 is no longer produced in the four small cities; (**III**) an alternating equilibrium in which four cities accommodate all industries, whereas goods 1 and 2 are no longer produced in the other 4 cities; steps **IV–VI** are the same as steps **I–III** with half the number of cities; steps **VII–IX** are the same as steps **IV–VI** with half the number of cities; and (**X**) a fully agglomerated equilibrium in which all industries are located within a single city. Those steps are illustrated in Fig. 3. For each city, the size of a circle is proportional to the size of the corresponding sector. The urban hierarchical principle appears in steps **II**, **III**, **V**, **VI**, **VIII** and **IX**. Full symmetry arises in steps **I**, **IV** and **VII**, and full agglomeration in step **X**.

Several remarks are in order. First, as transport costs decrease, the small cities lose one industry at a time. Industries gradually disappear from every other city in the reverse order of the competition toughness in the corresponding industry or, equivalently, the goods' degree of product differentiation. As transport costs fall further, half the cities have disappeared. More precisely, the above process may be subdivided in two steps: (i) *the number of industries decreases in half of the cities* and (ii) *the number of cities itself decreases by one half.* Clearly, the pattern obtained in Proposition 3 is a special case of the one described above.

Second, the number of cities shrinks and, when transport costs are sufficiently low, all industries end up being agglomerated in a small number of cities. In a nutshell, *an equilibrium with lower transport costs features fewer but bigger cities.* Such a process lends credence to the idea that, without the spectacular decline in transport costs observed in the aftermath of the Industrial Revolution, the emergence of very big cities in Europe and the United States would not have been possible (Bairoch, 1988). Another implication of our findings, which concurs with Fujita et al. (1999), is that there is both inter-industry and intra-industry trade within the urban system. Large cities exports more goods than small cities, whereas all cities trade differentiated varieties.

Last, the alternating symmetry obeys the urban hierarchical principle because the goods produced in small cities are always produced in large cities. Our setting thus provides new foundations to central place theory à la Christaller in the sense that the urban hierarchy emerges as the result of a process of economic integration: *big cities grow at the expense of small cities, which steadily lose jobs and industries.* In addition, big cities supply more goods as well as a wider range of varieties of the corresponding goods than small cities do as in the one-industry case.

The following issue remains to be solved: what happens when the elasticities of substitution are very different? Our results are illustrated in Fig. 4 for the values $\sigma^1 = 3$, $\sigma^2 = 6$ and $\sigma^3 = 9$. Observe that the evolutionary processes displayed in Figs. 3 and 4 are qualitatively similar. However, steps **III**, **IV** and **VII** for intermediate transport costs are different. This is because industries are now heterogeneous. For example, in step **IV** of Fig. 4, the most differentiated good 1 is produced only in 2 cities, good 2 is produced in 4 cities, and good 3 is produced in every city, which does not happen in Fig. 3 with similar industries Thus, when industries are heterogeneous enough, *the equilibrium pattern may involve more than two types of cities.* Furthermore, we observe that the number of goods and the number of varieties supplied in a city rise with its size, a result that mimics Christaller's urban hierarchy principle.

5. Conclusion

We have used a simple general equilibrium model with monopolistic competition and increasing returns in which the number, location and size of cities are determined as the outcome of the interplay between a few basic agglomeration and dispersion forces. Because they have a preference for diversity and variety, workers are attracted by places where there are many firms as different goods and varieties are available there, but they are repelled by places where they are many because this depresses the local labor market. Similarly, firms are attracted by places crowded by many consumers who provide them with large outlets, but they are repelled by places with many firms belonging to the same industry because local competition is tough.

Our main result is that a central place system emerges as a market outcome in which firms and workers are driven by their own interest. To the best of our knowledge, our paper is one of the very few economic contributions that provide an equilibrium approach to central place theory. As in Christaller (1933), cities supply the agricultural regions and the urban centers of lower ranks. However, unlike Christaller but very much as in Fujita et al. (1999), cities play two roles in the economy in the sense that they are goods suppliers to agricultural areas and cities of lower ranks, but also to the urban centers of higher ranks, the reason being that consumers have a preference for variety while firms produce differentiated goods.

As economic integration gets deeper and deeper, the relative size of cities changes in the sense that big cities attract more and more firms and workers, with small cities losing them. Structural changes in the urban system arise when one industry moves away from the small cities to set up in the big ones. Eventually, when transport costs are sufficiently low, small cities disappear from the space-economy, with the big cities becoming even bigger. Such a result shows that weaker and weaker spatial frictions tend to foster more and more spatial concentration of economic activity. More precisely, we find that economic integration fosters the emergence of big cities supplying (almost) all goods – the large and diversified metropolises of our time – which coexist with small and specialized cities in which only a handful of goods are produced. In such a context, trade occurs mainly across big cities where the demand for manufactured goods is concentrated.

One could consider the opposite thought experiment in which transport costs rise. When the bifurcation is a pitchfork, the rise of transport costs and/or tariff barriers would trigger a more dispersed pattern of activities. In such a context, the urban system would gradually collapse with the disappearance of big cities, while involving a very large number of small places. Trade would then occur mainly between those small urban centers and their rural hinterlands. This is what happened in Western Europe after the fall of the Roman Empire (Pirenne, 1925).

In closing, a word on rural–urban migration. As long as condition (15) remains valid, we may conjecture that the foregoing results hold true when the share μ of workers rises. Indeed, as productivity rises in the agricultural sector, farmers move into cities and become workers. As a result, the dispersion force gets weaker while the agglomeration force gets stronger since more varieties are produced. All of these fosters a more concentrated pattern of activities through the growth of big cities.

Appendix A. The argmax of $V^i(x)$ is 0 or 1/2C

When firms are distributed according to the symmetric equilibrium $\lambda_c^i = \mu\theta^i/C$, some long, but standard, calculations show that the indirect utility of an industry i-worker who resides at location $x \neq x_c$, $c = 1,\ldots,C$, is as follows:

$$V^i(x) = \frac{\alpha^i \theta^i A_1}{2\mu\sigma(\sigma-1)(1+\varphi^i)\log\varphi^i(\phi^i)^x} + \text{constant}$$

where

$$
A_1 \equiv 2\theta^i\mu\sigma^i(1+\varphi^i)(\phi^i)^x\log\frac{\mu\left[1-(\varphi^i)^{C/2}\right]\left[(\phi^i)^{2x}+\varphi^i\right]}{(1-\varphi^i)(\phi^i)^x}
$$
$$
- (1-\mu)(\sigma^i-1)\left(1-(\varphi^i)^2\right)\left[2\log\phi^i x(\phi^i)^{2x}\right.
$$
$$
+ \left((\phi^i)^{2x}+\varphi^i\right)\log\frac{1+\varphi^i}{(\phi^i)^{2x}+\varphi^i}\bigg]
$$
$$
- \varphi^i\log\phi^i x\Big[(\sigma^i-1)\left(1+(\phi^i)^{2x}-\mu+(\phi^i)^{2x}\mu\right.
$$
$$
+ \left(1+(\phi^i)^{2x}+\mu-(\phi^i)^{2x}\mu\right)\phi^i\Big)\Big]
$$

Because $V^i(x)$ is symmetric about $1/2C$, we may focus on the interval $[0, 1/2C]$.

Studying directly the behavior of this function appears to be especially hard. This is why we start with a high-order derivative that takes a much simpler form. Let

$$z \equiv (\phi^i)^x \in [(\phi^i)^{1/2C}, 1]$$

The zero-slope condition $dV^i(1/2C)/dz = 0$ can be shown to be equivalent to

$$G_1(z) \equiv \frac{G_2(z)}{z^2 - \varphi^i} = 0$$

where

$$
G_2(z) \equiv \Big\{(\sigma-1)(z^2+\varphi^i)\Big[1-\mu+\left(1-\mu+2\theta^i\mu\right)\varphi^i
$$
$$
- z^2\left(1-\mu+2\theta^i\mu+\varphi^i-\mu\varphi^i\right)\Big]
$$
$$
- 2\theta^i z\mu\sigma(1+\varphi^i)(z^2-\varphi^i)\varphi^i\log\varphi^i\Big\}
$$
$$
- 2z^2(1-\mu)(\sigma-1)(z^2+\varphi^i)\left[1-(\phi^i)^2\right]\log z
$$
$$
- (1-\mu)(\sigma-1)\left[1-(\phi^i)^2\right]\left[z^4-(\phi^i)^2\right]\log\frac{1+\varphi^i}{z^2+\varphi^i}
$$

and $\text{sgn}(G_1(z)) = \text{sgn}(dV^i(1/2C)/dz)$. Computing the fourth-order derivative of $G_1(z)$ with respect to z and setting

$$y \equiv z^2/\varphi^i = (\phi^i)^{2x-1/C} \in \left[1, \left(\phi^i\right)^{-1/C}\right]$$

we obtain:

$$F(y) = A_2\left[(y-1)(y^4+28y^3+38y^2+28y+1) - 12y(y+1)^3\log y\right]$$

where $A_2 > 0$. We show below that $F(y)$ has at most one an interior solution in the interval of $(1, (\phi^i)^{-1/C})$ and that this solution is a local minimizer of $V^i(x)$. Hence, the interval $(0, 1/2C)$ does not contain a local maximizer of $V^i(x)$. Since

$$\frac{d^5 F(y)}{dy^5} = -\frac{24}{y^4}\left(5y^4 - 12y^3 + 9y^2 - 6y + 3\right) \quad \text{and} \quad \frac{d^5 F(1)}{dy^5} < 0$$

we have $d^5F(y)/dy^5 < 0$ for $1 \leqslant y < 1.596$ and $d^5F(y)/dy^5 > 0$ for $1.596 < y \leqslant (\phi^i)^{-1/C}$. Because $d^\nu F(0)/dy^\nu = 0$ for $\nu = 0$–4, it must be

that $F(y)$ and $d^4 G_1(z)/dz^4$ are either (i) always decreasing in y or (ii) first decreasing and then increasing in y. However, since $d^3 G_1((\phi^i)^{1/2C})/dz^3 < 0$ and $d^3 G_1(1)/dz^3 < 0$ hold for all $0 < \mu < (\sigma^i - 1)/\theta^i(2\sigma^i - 1)$, we get $d^3 G_1(z)/dz^3 < 0$ for all $z \in [(\phi^i)^{1/2C}, 1]$.

Computing now the first- and second-order derivatives of $G_1(z)$, we have

$$\text{sgn}\left[dG_1\left((\phi^i)^{1/2C}\right)/dz\right] = \text{sgn}\left[d^2 G_1\left((\phi^i)^{1/2C}\right)/dz^2\right]$$
$$dV^i(1)/dz > 0$$

Therefore, the following two cases may arise.

(i) If $dG_1[(\phi^i)^{1/2C}]/dz \geqslant 0$ and $d^2 G_1[(\phi^i)^{1/2C}]/dz^2 \geqslant 0$, then $G_1(z) > 0$ holds over the interval $[(\phi^i)^{1/2C}, 1]$. Because there exists no interior solution satisfying $G_1(z) = 0$ for $z \in ((\phi^i)^{1/2C}, 1)$, there is no local maximizer of $V^i(x)$ over $(0, 1/2C)$.

(ii) If $G_1[(\phi^i)^{1/2C}]/dz < 0$ and $d^2 G_1[(\phi^i)^{1/2C}]/dz^2 < 0$, then there exists a unique interior solution $(\phi^i)^{1/2C} < z^* < 1$ satisfying $G_1(z^*) = 0$. Since z^* is a local minimizer of $G_1(z)$, $x^* \equiv \log z^*/\log \phi^i$ is also a local minimizer of $V^i(x)$, implying that there is no local maximizer of $V^i(x)$ in $(0, 1/2C)$.

Appendix B. Proof of Proposition 1

We rewrite the condition $V^i(0) \geqslant V^i(1/2C)$ as follows:

$$g(\varphi^i) \geqslant 0 \quad \text{for all } i$$

where

$$g(\varphi^i) \equiv 2\varphi^i\log\varphi^i\left[\theta^i\mu\sigma^i\log\varphi^i - 2(\sigma^i-1)\left(1-\mu-\mu\theta^i+\mu\theta^i\varphi^i\right)\right]$$
$$- (1-\mu)(\sigma^i-1)\left(1-(\varphi^i)^2\right)$$

for $C = 2$ and

$$g(\varphi^i) \equiv (\sigma^i-1)(1-\varphi^i)\Big[\mu\left(1+3(\varphi^i)^2\right)(1+\varphi^i)$$
$$- (1+3\varphi^i)\left(1+(\varphi^i)^2\right)\Big]\log\varphi^i$$
$$+ \left[2\theta^i\mu\sigma^i\varphi^i - (1-\mu)(\sigma^i-1)\left(1-(\varphi^i)^2\right)\right]\log\varphi^i$$
$$\times \left[1+(\varphi^i)^2\right]\log\frac{2}{\varphi^i+(\varphi^i)^3} + 4\theta^i\mu\sigma^i\varphi^i\left[1+(\varphi^i)^2\right](\log\varphi^i)^2$$

for $C = 2^2, 2^3, \ldots, 2^K$. It is readily verified that $\lim_{\varphi^i \to 0_+} g(\varphi^i) < 0$ and $g(1) = g'(1) = g''(1) = 0$ and $g'''(1) < 0$. Consequently, for any given value of C and for each i, the equation $g(\varphi^i) = 0$ has at least one interior solution. Let $\tau_2^i(C)$ be the largest of these solutions. $\square$

Appendix C. Proof of Lemma 2

We may assume that all industries are active in city c, for otherwise we restrict ourselves to the sole active industries. Differentiating (10) yields:

$$\frac{\partial w_c^i}{\partial \lambda_c^i} = -\frac{\alpha^i}{\sigma^i}\left[\sum_{e=1}^{C}\frac{\Lambda_e\left(\phi^i(x_c-x_e)\right)^2}{\left(\sum_{d=1}^{C}\phi^i(x_d-x_e)\lambda_d^i\right)^2}\right.$$
$$\left. + \int_0^1 \frac{(1-\mu)\left(\phi^i(x_c-x_a)\right)^2}{\left(\sum_{d=1}^{C}\phi^i(x_d-x_a)\lambda_d^i\right)^2}dx_a\right] < 0$$

and

$$\frac{\partial w_c^i}{\partial \lambda_c^j} = 0 \quad \text{for all } j \neq i$$

These conditions imply that the Jacobian matrix of $w_c^i - w_c^j$ is negative definite. It then follows from Section 2.4 in Hofbauer and Sandholm (2009) that (10) has a unique sectoral equilibrium within each city c. □

Appendix D. Proof of Proposition 2

It follows from Bellman (1970, pp. 242–243) that the real parts of the eigenvalues of the circulant Jacobian matrix of (14) are given by

$$z_c^i = \sum_{d=1}^{C} \frac{\partial J_d}{\partial \lambda_C^i} \cos \frac{2\pi cd}{C} \quad \text{for all } c$$

Since

$$\frac{\partial J_d}{\partial \lambda_C^i} = \frac{\mu}{C} \left(\frac{\partial V^i(d/C)}{\partial \lambda_C^i} - V^i(1) - \sum_e \frac{\partial V^i(e/C)}{\partial \lambda_C^i} \right) \Bigg|_{\lambda_c^i = \lambda^i}$$

at a symmetric equilibrium, we obtain

$$z_c^i = \frac{\mu}{C} \sum_d \frac{\partial V^i(d/C)}{\partial \lambda_C^i} \cos \frac{2\pi cd}{C} \Bigg|_{\lambda_c^i = \lambda^i} \tag{A.1}$$

Thus, the symmetric equilibrium (λ^i) is stable if (A.1) is negative for all c and all i. Setting

$$\Phi^i \equiv \sum_d \phi^i(x_c, x_d)$$

which is the same for all c at a symmetric equilibrium, it is straightforward to show that

$$\frac{\partial V^i(d/C)}{\partial \lambda_C^i} \Bigg|_{\lambda_c^i = \lambda^i} = \frac{\alpha^i}{\sigma^i \lambda^i} \left[\frac{(2\sigma^i - 1)\phi^i(x_d, 1)}{(\sigma^i - 1)\Phi^i} \right.$$
$$- \frac{\mu}{C(\Phi^i)^2 \lambda^i} \sum_e \phi^i(x_d, x_e)\phi^i(x_e, 1)$$
$$\left. - \frac{1 - \mu}{\lambda^i} \int_0^1 \frac{\phi^i(x_d, x)\phi^i(x, 1)}{\left(\sum_e \phi^i(x, x_e) \right)^2} dx \right]$$

The value of z_c^i evaluated at the symmetric equilibrium is then obtained by plugging the RHS of this expression in (A.1):

$$z_c^i = \begin{cases} \dfrac{C\alpha^i(1-\varphi^i)^2 \left(A_3 + A_4 \cos\frac{2c\pi}{C} \right)}{\log(\varphi^i)(\mu\theta^i)^2(\sigma^i-1)\sigma^i \left(1 + \varphi^{i2} - 2\varphi^i \cos\frac{2c\pi}{C}\right)^2} & \text{for } c=2,4,\ldots,C/2 \\[4mm] \dfrac{C\alpha^i(1-\varphi^i)^2 \left(1 + \varphi^{i\frac{C}{2}}\right)\left(A_5 + A_6 \cos\frac{2n\pi}{C} \right)}{\log(\varphi^i)\left(\mu\theta^i\right)^2(\sigma^i-1)\sigma^i \left(1-\varphi^{i\frac{C}{2}}\right)^2 \left(1 + \varphi^{i2} - 2\varphi^i \cos\frac{2n\pi}{C}\right)^2} & \text{for } c=1,3,\ldots,C/2-1 \end{cases}$$

where

$$A_3 \equiv \left[\theta^i \mu(2\sigma^i - 1)\left(1 + \varphi^{i2}\right) - (\sigma^i - 1)\left(1 - 2\mu\varphi^i + \varphi^{i2}\right) \right] \log(\varphi^i)$$
$$- (1 - \mu)(\sigma^i - 1)\left(1 - \varphi^{i2}\right)$$

$$A_4 \equiv 2\left[\sigma^i - 1 + \mu\left(1 + \theta^i - \sigma^i - 2\theta^i \sigma^i\right) \right] \varphi^i \log(\varphi^i)$$
$$+ (1 - \mu)(\sigma^i - 1)\left(1 - \varphi^{i2}\right)$$

$$A_5 \equiv \left[\theta^i \mu(2\sigma^i - 1)\left(1 + \varphi^{i2}\right)\left(1 - \varphi^{i\frac{C}{2}}\right) \right.$$
$$- (\sigma^i - 1)\left(1 - 2\mu\varphi^i + \varphi^{i2}\right)\left(1 + \varphi^{i\frac{C}{2}}\right) \Bigg] \log(\varphi^i)$$
$$- (1 - \mu)(\sigma^i - 1)\left(1 - \varphi^{i2}\right)\left(1 + \varphi^{i\frac{C}{2}}\right)$$

$$A_6 \equiv 2\varphi^i \left[-1 + \mu + \theta^i \mu + \sigma^i - \mu\sigma^i - 2\theta^i \mu\sigma^i \right.$$
$$+ \left(\sigma^i - 1 + \mu\left(1 - \theta^i - \sigma^i + 2\theta^i \sigma^i\right) \right) \varphi^{i\frac{C}{2}} \Bigg] \log(\varphi^i)$$
$$+ (1 - \mu)(\sigma^i - 1)\left(1 - \varphi^{i2}\right)\left(1 + \varphi^{i\frac{C}{2}}\right)$$

Because $z_c^i = z_{C-c}^i$ holds for all c due to the circulant Jacobian matrix of (14), we consider only for $c = 1, 2, \ldots, C/2$.

For $c = 2, 4, \ldots, C/2$, because the denominator of z_c^i is always negative, $z_c^i < 0$ implies $A_3 + A_4 \cos(2c\pi/C) > 0$. However, A_4 is shown to be positive and $\cos\frac{2c\pi}{C}$ is decreasing in $c = 2, 4, \ldots, C/2$. Hence, the condition $z_c^i < 0$ is replaced with $A_3 - A_4 > 0$ for $c = 2, 4, \ldots, C/2$. Likewise, for $c = 1, 3, \ldots, C/2 - 1$, because the denominator of z_c^i is negative, $z_c^i < 0$ implies $A_5 + A_6 \cos(2c\pi/C) > 0$. However, A_6 is shown to be positive. Hence, the condition $z_c^i < 0$ is replaced with $A_3 - A_4 > 0$ for $c = 1, 3, \ldots, C/2 - 1$.

Finally, it is readily verified that $z_{C/2}^i > z_{C/2-1}^i$ holds. Hence, for any given value of i, z_c^i is maximized at $c = C/2$. Thus, the symmetry breaking condition amounts to $\max_i z_{C/2}^i < 0$. Computing $z_{C/2}^i$ shows that

$$z_{C/2}^i = h(\varphi^i)$$

where

$$h(\varphi^i) \equiv 2(1 - \mu)(\sigma^i - 1)\left[1 - (\varphi^i)^2\right] + \left[(1 - \mu)(\sigma^i - 1)(1 + \varphi^i)^2 \right.$$
$$\left. - \mu\theta^i(1 - \varphi^i)(\sigma^i - 2\varphi^i + 3\sigma^i \varphi^i)\right] \log \varphi^i$$

for $C = 2$ and

$$h(\varphi^i) \equiv 2(1 - \mu)(\sigma^i - 1)\left[1 - (\varphi^i)^2\right]$$
$$+ \left[(\sigma^i - 1)\left[1 + 2\varphi^i - 4\mu\varphi^i + (\varphi^i)^2\right]\right.$$
$$\left. - \mu\theta^i(2\sigma^i - 1)(1 + \varphi^i)^2\right] \log \varphi^i$$

for $C = 2^2, 2^3, \ldots, 2^K$. Because $\partial^2 h(\varphi^i)/\partial(\varphi^i)^2 > 0$, h is strictly convex over $[0, 1]$.

Furthermore, $\lim_{\varphi^i \to 1} h(\varphi^i) = 0$ and

$$\text{sgn}\left[\lim_{\varphi^i \to 0_+} h(\varphi^i) \right] = \text{sgn}\left[\sigma^i - 1 - \theta^i(2\sigma^i - 1)\mu \right]$$

Therefore, if

$$\mu < \min_i \frac{\sigma^i - 1}{\theta^i(2\sigma^i - 1)}$$

it must be that, for each i, the equation $h(\varphi^i) = 0$ has a single positive solution $\tau_1^i(C)$. Hence, the condition that $z_{C/2}^i$ must be negative for all i is equivalent to $\tau < \tau_1(C) \equiv \max_i \tau_1^i(C)$. □

Appendix E. Implicit determination of boundaries

(i) While $g(\varphi) = 0$ and $h(\varphi) = 0$ are given in the above, $l(\varphi^i)$ is given by

$$l(\varphi^i) \equiv 8(1-\mu)(\sigma^i-1)\left[1-(\varphi^i)^2\right]\left[1+\varphi^i+(\varphi^i)^2\right]$$
$$+ \left[3(1-\mu)(\sigma^i-1)(1+\varphi^i)^4 \right.$$
$$\left. - \theta^i\mu(1-\varphi^i)^3(\sigma^i+7\sigma^i\varphi^i-6\varphi^i)\right]\log\varphi^i$$

for $C = 2$ and by

$$l(\varphi^i) \equiv 8(1-\mu)(\sigma^i-1)\left[1-(\varphi^i)^2\right]\left[1+\varphi^i+(\varphi^i)^2\right]$$
$$+ \left\{3(\sigma-1)\left[(1+\varphi^i)^4-8\mu\varphi^i\left[1+(\varphi^i)^2\right]\right]\right.$$
$$\left. - \theta^i\mu(4\sigma-3)\left[1-(\varphi^i)^2\right]^2\right\}\log\varphi^i$$

for $C = 2^2, 2^3, \ldots, 2^k$.

(ii) The boundary of the domain between bifurcation and symmetry breaking is obtained from the two bilinear equations $h(\varphi;\ \mu,\ \sigma) = 0$ and $l(\varphi;\ \mu,\ \sigma) = 0$:

$$\begin{cases} \mu_2(\varphi) = \dfrac{(1+\varphi)^2\left[3-3\varphi^2+(1+4\varphi+\varphi^2)\log\varphi\right]}{3(1-\varphi^2)(1+\varphi)^2+(1+8\varphi+4\varphi^2+12\varphi^3-\varphi^4)\log\varphi} \\ \sigma_2(\varphi) = \dfrac{2\varphi\left[(1-\varphi)\left(1+10\varphi+\varphi^2\right)+6\varphi(1+\varphi)\log\varphi\right]}{(1-\varphi)(3+11\varphi+29\varphi^2+5\varphi^3)+(1+\varphi)(1+5\varphi+17\varphi^2+\varphi^3)\log\varphi} \end{cases}$$

for $\varphi \in [0.533,1)$ when $C = 2$ and

$$\begin{cases} \mu_2(\varphi) = \dfrac{(1+\varphi)^2\left[3-3\varphi^2+(1+4\varphi+\varphi^2)\log\varphi\right]}{3(1-\varphi^2)(1+\varphi)^2+(1+\varphi)(1+10\varphi+\varphi^2)\log\varphi} \\ \sigma_2(\varphi) = \dfrac{4\varphi\left[(1-\varphi)\left(1+10\varphi+\varphi^2\right)+6\varphi(1+\varphi)\log\varphi\right]}{(1-\varphi)(3+16\varphi+58\varphi^2+16\varphi^3+3\varphi^4)+(1+\varphi)\left(1+6\varphi+34\varphi^2+6\varphi^3+\varphi^4\right)\log\varphi} \end{cases}$$

for $\varphi \in [0.336,1)$ when $C = 2^2, 2^3, \ldots, 2^k$. All of them are shown to be monotone in φ.

(iii) Similarly, the boundary of the domain for which there exists a stable symmetric equilibrium is obtained from the two bilinear equations $g(\varphi;\ \mu,\ \sigma) = 0$ and $h(\varphi;\ \mu,\ \sigma) = 0$ for $\varphi \in [0.00056, 0.00152)$ when $C = 2$, and for $\varphi \in [0.00056, 0.00150)$ when $C = 2^2, \ldots, 2^k$, all of which are shown to be monotone in φ.[8]

Appendix F. Proof of Proposition 4

Consider the utility differential

$$\Delta V^i(\tau) \equiv V^i_{1/2C} - V^i_0.$$

where the indirect utilities (8) is evaluated at the midpoint $x = 1/2C$ and the city location $x = 0$ of a symmetric equilibrium with $\lambda^i_c = \theta_i\mu/C$. Because $\Delta V^i(1) = 0 < \lim_{\tau\to\infty}\Delta V^i(\tau)$ and $\mathrm{d}\Delta V^i(1)/\mathrm{d}\tau > 0$, the equation $\Delta V^i(\tau) = 0$ has at least one solution belonging to the interval $(1,\infty)$. The smallest solution is, by definition, the doubling point $\tau_2(C)$.

At the midpoint $x = 1/2C$, the first term of (8) is the same across industries, while the second and third terms of (8) put together are as follows:

$$\frac{\alpha^i}{\sigma^i}\left[\sum_{e=1}^{C}\frac{\Lambda_e\phi^i(x_c-x_e)}{\sum_{d=1}^{C}\lambda^i_d\phi^i(x_d-x_e)}+\int_0^1\frac{(1-\mu)\phi^i(x_c-x_a)}{\sum_{d=1}^{C}\lambda^i_d\phi^i(x_d-x_a)}\mathrm{d}x_a\right]$$
$$= \sum_j\frac{\alpha^j}{\sigma^j}\left[\frac{2\sqrt{\varphi^i}}{(1+\varphi^i)}+\frac{1-\mu}{\mu}\frac{\varphi^i\log\varphi^i+(1-\varphi^i)\log\frac{1+\varphi^i}{2}}{\sqrt{\varphi^i}\log\varphi^i}\right]$$

This expression differs only with respect to $\varphi^i = \tau^{1-\sigma^i}$ across industries. If $\sigma^i > \sigma^j$ holds for all $j \neq i$, then $[\tau_2(C)]^{\frac{1-\sigma^i}{C}} < [\tau_2(C)]^{\frac{1-\sigma^j}{C}}$ also holds, so that φ^i reaches its smallest value at $\tau = \tau_2(C)$. This implies that, when transport costs increase, $\Delta V^i(\tau_2(C)) = 0$ holds for the industry i having the largest elasticity of substitution σ^i.

Although the terms of (8) change with the emergence of these small cities, the above argument can be repeated for these cities to accommodate two industries, and so on until a fully symmetric configuration emerges. Putting things on their head, when the transport costs decrease, industries disappear one by one according to the increasing order of σ^i. □

References

Bairoch, P., 1988. Cities and Economic Development: From the Dawn of History to the Present. University of Chicago Press, Chicago.

Bairoch, P., 1997. Victoires et déboires. Histoire é conomique et sociale du monde du XVIᵉ siècle à nos jours. Editions Gallimard, Paris.

Bellman, R., 1970. Introduction to Matrix Analysis, second ed. McGraw-Hill, New York.

Brakman, S., Garretsen, H., Schramm, M., 2004. The spatial distribution of wages and employment: estimating the Helpman–Hanson model for Germany. Journal of Regional Science 44, 437–466.

Christaller, W., 1933. Die Zentralen Orte in Sü ddeutschland. Jena: Gustav Fischer Verlag (English translation: The Central Places of Southern Germany. Prentice-Hall, Englewood Cliffs, NJ, 1966).

Eaton, J., Eckstein, Z., 1997. Cities and growth: theory and evidence from France and Japan. Regional Science and Urban Economics 27, 443–474.

Eaton, B.C., Lipsey, R.G., 1982. An economic theory of central places. Economic Journal 92, 56–72.

Fujita, M., Mori, T., 2005. Transport development and the evolution of economic geography. Portuguese Economic Journal 4, 129–159.

Fujita, M., Krugman, P., Mori, T., 1999. On the evolution of hierarchical urban systems. European Economic Review 43, 209–251.

Ginsburgh, V., Papageorgiou, Y.Y., Thisse, J.-F., 1985. On existence and stability of spatial equilibria and steady-states. Regional Science and Urban Economics 15, 149–158.

Glaeser, E.L., Kohlhase, J.E., 2004. Cities, regions and the decline of transport costs. Papers in Regional Science 83, 197–228.

Hanson, G., 2005. Market potential, increasing returns, and geographic concentration. Journal of International Economics 67, 1–24.

Hofbauer, J., Sandholm, W.H., 2009. Stable games and their dynamics. Journal of Economic Theory 144, 1665–1693.

Hohenberg, P.M., 2004. The historical geography of European cities. An interpretative essay. In: Henderson, J.V., Thisse, J.-F. (Eds.), Handbook of Regional and Urban Economics. Cities and Geography. North-Holland, Amsterdam, pp. 3021–3052.

Hsu, W.-T., 2009. Central place theory and city size distribution. Mimeograph. <http://ihome.cuhk.edu.hk/~b118917/city_hierarchy_1.pdf>.

Hsu, W.-T., Holmes, T.J., 2009. Optimal city hierarchy: a dynamic programming approach to central place theory. Mimeograph. <http://www.econ.umn.edu/~holmes/papers/city_optimal.pdf>.

Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99, 483–499.

Mion, G., 2004. Spatial externalities and empirical analysis: the case of Italy. Journal of Urban Economics 56, 97–118.

Picard, P., Zeng, D.-Z., 2005. Agricultural sector and industrial agglomeration. Journal of Development Economics 77, 75–106.

Pirenne, H., 1925. Medieval Cities. Princeton University Press, Princeton, NJ.

Quinzii, M., Thisse, J.-F., 1990. On the optimality of central places. Econometrica 58, 1101–1119.

Rasband, S.N., 1990. Chaotic Dynamics of Nonlinear Systems. Wiley, New York.

Scotchmer, S., 2002. Local public goods and clubs. In: Auerbach, A., Feldstein, M. (Eds.), Handbook of Public Economics, vol. 4. North-Holland, Amsterdam, pp. 1997–2042.

World Bank, 1995. World Development Report: Workers in an Integrating World. Oxford University Press, Oxford.

---

[8] The expressions of $\mu_1(\varphi)$ and $\sigma_1(\varphi)$ as well as proofs may be obtained from the authors.