How and where satellite cities form around a large city

Kiyohiro Ikeda,[1] Hiroki Aizawa,[2] José M. Gaspar[3]

Abstract

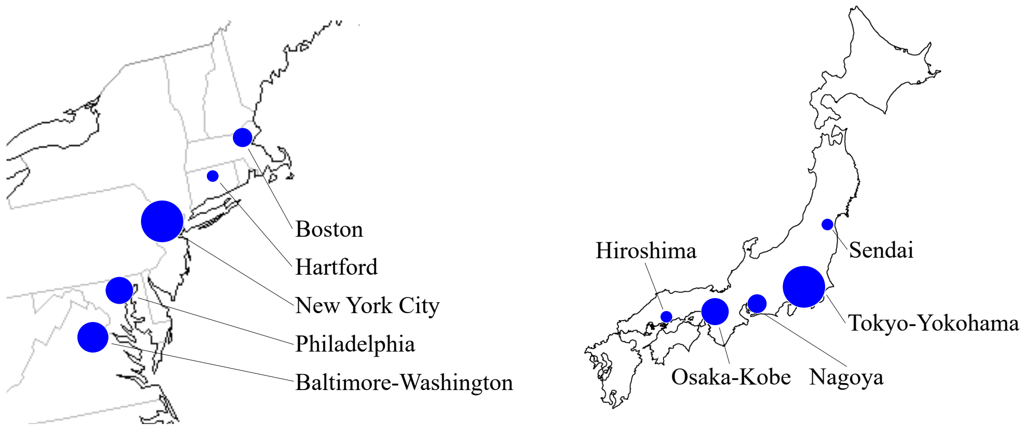
We investigate economic agglomerations in a long narrow economy, in which discrete locations are evenly spread over a line segment. The bifurcation mechanism of a monocentric city at the center is analyzed analytically to show how satellite cities form for a general spatial economic model. This is an important step to elucidate the mechanism of the competition between a large central city and satellite cities, which is taking place worldwide. By the analysis of the Forslid & Ottaviano (J Econ Geo, 2003) model, we show where satellite cities form. The larger the agglomeration forces, the farther away from the monocentric city satellite cities emerge. The transition of stable agglomeration patterns is observed and is compared with those in the real world.

Keywords: Bifurcation; economic geography; replicator dynamics; satellite cities; spatial period doubling.

[1]Address for correspondence: Kiyohiro Ikeda, Department of Civil and Environmental Engineering, Tohoku University, Aoba, Sendai 980-8579, Japan; kiyohiro.ikeda.b4@tohoku.ac.jp

[2]Department of Civil and Environmental Engineering, Tohoku University, Aoba, Sendai 980-8579, Japan; hiroki.aizawa.p3@dc.tohoku.ac.jp

[3]CEGE and Católica Porto Business School, Universidade Católica Portuguesa; jgaspar@porto.ucp.pt



(a) East Atlantic (b) Main Island in Japan

Figure 1: A chain of cities in the world

1. Introduction

Megalopolises consisting of clusters or chains of networked cities and towns that form large and densely populated areas or urban complexes (megaregions) have been continuously sprawling around the world ever since at least the early 20th century (Gottmann, 1957), both at the national and transnational scale. A chain of cities prospers in a closed narrow corridor between the Atlantic Ocean and the Appalachian Mountains (see Fig. 1(a)) and in the Main Island of Japan (see Fig. 1(b)). The mechanism of the growth of a megalopolis, such as New York City and Tokyo, among a chain of cities is of great interest in spatial economics.

Geography, amongst other factors, plays a big role in characterizing chains of cities according to their location and spatial topology. One such particular configuration is the line segment, whose analysis, despite its stylized geometry, is of great interest because: (i) it is both simple and generates asymmetries that confers advantages to some regions; and (ii) it is empirically relevant as it fits several real world examples of megalopolises or megaregions, as explained below. This paper elucidates the mechanism of economic

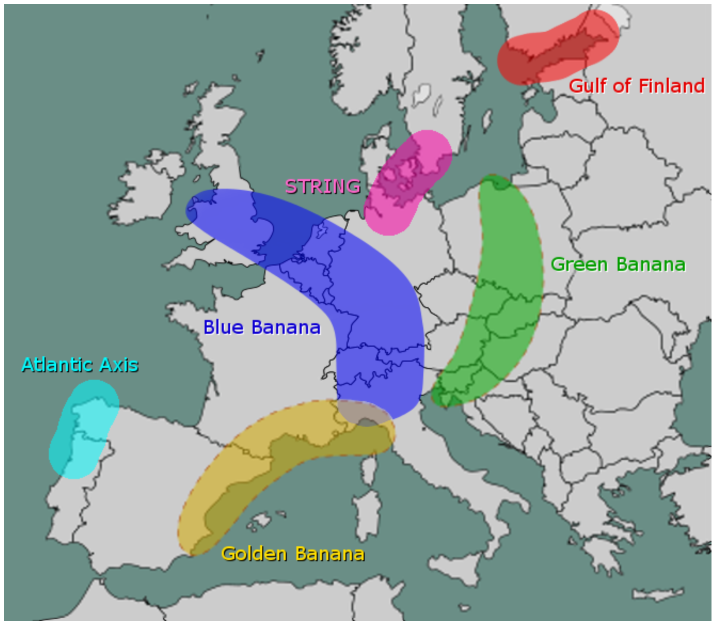


Figure 2: European transnational megalopolises. Source: Maps on the Web: https://mapsontheweb.zoom-maps.com.

agglomerations in a long narrow economy, in which discrete locations are evenly spread over a line segment.

Chains of cities along a narrow corridor can be found also at transnational scales, particularly in Europe (see Fig. 2), such as the Atlantic Axis (from Porto in Portugal to Coruña in Spain) and the STRING (from Hamburg in Germany to Oslo in Norway). Other famous megaregions in Europe are the so called "bananas" (blue, green and golden). The Golden Banana or "sun belt", for instance, is a term used to describe urbanisation in a European context. It denotes an area of higher population density lying between Valencia in the West and Genoa in the East along the coast of the Mediterranean Sea, defined by the "Europe 2000" report from the European Commission in 1995 to be analogous to the Blue Banana. The region is a centre for information technology and manufacturing. Despite the obvious shape that inherits their names, the golden banana, for instance, is the one that most closely resembles a narrow corridor (i.e., with a minimal curvature).

This paper models a chain of cities by a long narrow economy with equally spaced discrete places on a line segment. The literature reports several characteristic agglomeration patterns of this economy: the simplest core–satellite pattern for three places (Ago et al., 2006), a chain of spatially repeated core–periphery patterns *a la* Christaller and Lösch (e.g., Fujita and Mori, 1997), and a megalopolis which consists of large core cities that are connected by *an industrial belt*, i.e., *a continuum of cities* (Mori, 1997). These patterns were numerically observed by changing agglomeration forces and transport costs (Ikeda et al., 2017). Yet such patterns were investigated somewhat fragmentarily and in an ad hoc manner up to now.

That said, this paper aims to answer the question "How and where do satellite cities form around a large city?" in the framework of spatial economics in a long narrow economy. This is apparently a difficult mission as the associated agglomeration properties are dependent on spatial economic models and as well as on their microeconomic parameters. To tackle this mission, we elucidate the bifurcation/agglomeration mechanism of a long narrow economy in the following two steps:

1. The bifurcation mechanism for a general spatial economy.

2. The bifurcation mechanism for a particular well-known spatial economic model.

The results for the first step are more general and hold in a framework which encompasses several settings as particular cases, whereas those for the second step are more informative, albeit more restrictive as its predictions hinge on the particular assumptions of the model.

In the first step, as a novel theoretical contribution of the paper, we answer the question "How do satellite cities form around a large city?" in a manner applicable to a general economic geography model with an arbitrary number of places. A state of full agglomeration to a large single city at the center is shown to encounter a bifurcation at a critical level[4] of transport costs (freeness of trade) above (below) which it becomes economically unsustainable and leads to the emergence of satellite cities around the large central city. Nowadays it seems far more important to investigate the competition between central and satellite cities than to investigate the self-organization of cities in a flat land as envisaged by Central Place Theory (Christaller, 1933).

In the second step, we answer the question "Where do satellite cities form around a large city?". We resort to a many-region version of the model (FE model) by Forslid and Ottaviano (2003) in favor of its analytical tractability and close resemblance to Krugman's (1991) seminal Core-Periphery model.[5] We analyze analytically the existence and uniqueness of the sustain point for the state of the full agglomeration, and the existence and stability of bifurcating solutions from this point that engender satellite cities.

The location of satellite cities is demonstrated to be dependent on the agglomeration forces that are a consequence of: (i) the global size of the industrial sector relative to the traditional sector, and (ii) the degree of scale economies in the industrial sector. When agglomeration forces are very small, a large central place surrounded by two neighboring satellite places emerges, thus forming a hump-shaped megalopolis around the central city.

[4]This critical level is called the *sustain point* (Fujita et al., 1999).

[5]In fact, as shown by Robert-Nicoud (2005), the FE model is isomorphic to the Core-Periphery model in an economically meaningful state space.

When these forces are large, satellite cities appear far away from the primary city at the center. This would give an economic implication of *agglomeration shadow* (Arthur, 1990),[6] cast by cities with a large industry size over locations in vicinity, in which little or no settlement takes place because competition between neighboring regions is too intense to make them profitable for firms to settle. In contrast, sufficiently separated satellite cities and the central region can share industry.

Our paper relates with the the recent work by Turner et al (2021), who study agents' location as increasing returns to scale in production increase in an urban economics setting with costly commuting and heterogeneity in preferences for residential location. Their setup is akin to ours in that their spatial topology comprises a linear city with three discrete locations. Turner et al. (2021) find that, once increasing returns become sufficiently strong, stronger returns to scale disperse economic activity towards the peripheral locations. However, location patterns in the urban setting are shown to depend heavily on the dispersion of agents' heterogeneity, while in our setting no heterogeneity is required to reproduce some of their predictions.

The progress of stable and sustainable equilibria as the trade freeness increases is of great economic interest as it captures the historical process of increasing economic integration and globalization. To observe this progress, we conduct extensive comparative statics analyses for various number of cities. There appear three stages called (1) *Dawn* stage, (2) *Core–satellite* stage, and (3) *Full agglomeration* stage, in this order. In the Dawn stage, every other city grows, forming a chain of spatially repeated core–periphery

[6]See also Fujita et al. (1999), Ioannides and Overman (2004), and Fujita and Mori (2005).

patterns *a la* Christaller and Lösch.[7] The Core–satellite stage accommodates a central place with twin satellite place.[8] As the trade freeness increases further, the core place at the center grows and the twin satellite cities shrink, thereby leading to the Full agglomeration stage for a gigantic mono-center. Such progress captures the essence of historical agglomeration tendencies.

We take a step further by looking at the city sizes in terms of population across different megaregions (Atlantic Axis and Golden Banana) and find that, indeed, megalopolises with a huge industry sizes tend to have two large cities at the opposite ends of the line segments. Moreover, different megaregions along a narrow corridor can be classified under different stages (dawn, core-satellite, full agglomeration) of spatial development. We also look at population differences in Japan between 1950 and 2020 to find that the "central" region of Nagoya has increased in population relatively more than neighbouring regions, a process which may well be attributed to the huge development of the transportation system in Japan (in particular the high-speed Shinkansen trains that connect major cities), leading to an overall decrease in transport and commuting costs.

Noteworthy, the results regarding the general bifurcation mechanism obtained from the aforementioned first step (i.e., our first research question) potentially encompass several models that fall under the various sub-fields of spatial economics or economic geography, be it new economic geography, urban economics, or both, location theory, or the

[7]Such a spatial alternation of a core place with a large population and a peripheral place with zero population was observed and studied for the racetrack economy in Tabuchi and Thisse (2011), Ikeda et al. (2012), and Akamatsu et al. (2012). Mossay and Picard (2011) and Ikeda et al. (2017) conducted comparative studies of long narrow and racetrack economies.

[8]The existence of such pattern has come to be observed in the population data (Ikeda er al., 2019).

more recent quantitative spatial economics (see e.g. Redding and Rossi-Hansberg, 2017; Behrens and Murata, 2021), as long as they consider inter-regional mobility of at least some production factors, such as labour, that are susceptible to assignment of some sort of dynamics that will govern their location decisions. The plethora of available different settings can be employed in the future to investigate which results are model dependent and which predictions are more general regarding the formation of megaregions. This, however, falls out of scope of the present work.

This paper is organized as follows. The bifurcation mechanism of a long narrow economy for a general spatial economic model is described in Section 2. For the FE model, the bifurcation mechanism is studied in Section 3 and a comparative static analysis is conducted in Section 4. Comparison with agglomeration patterns in the real world is conducted in Section 5. Section 6 is left for concluding remarks.

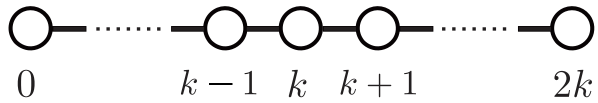


Figure 3: A long narrow economy

2. General bifurcation mechanism of a long narrow economy

We would like to answer the question "How and where do satellite cities form around a large city?" For this purpose, we investigate the bifurcation mechanism of the full agglomeration at the center that leads to the emergence of satellite cities. The results are general and applicable to general spatial economic models with a single scalar independent variable at each city.

2.1. Modeling of the spatial economy

The long narrow economy has $K = 2k + 1$ ($k \in \mathbb{Z} : k \geq 1$) cities labeled $i \in N = \{0, ..., k, ..., 2k\}$, which are equally spread on a line segment (Fig. 3). The kth city is located at the center, and a city $i \neq k$ is said to be $\delta \equiv |i - k|$ *steps* away from the center. In other words, $\delta \in \{0, ..., k\} = N_\delta$ is simply the integer expressing the number of step, or *number of cities* to the right of left of the central city.

There are inter-regionally mobile agents (footloose entrepreneurs), the number of which at city $i \in N$ is denoted by λ_i under the constraint $\sum_{i \in N} \lambda_i = 1$. We introduce a spatial equilibrium in which the footloose entrepreneurs migrate among cities and choose to live in the city that offers them the highest utility. A customary way of defining such an equilibrium is to consider the following problem: Find $(\boldsymbol{\lambda}^*, \hat{v})$ satisfying

$$(v_i - \hat{v})\lambda_i^* = 0, \qquad v_i - \hat{v} \leq 0, \qquad \lambda_i^* \geq 0, \qquad \sum_{i \in N} \lambda_i^* = 1, \tag{1}$$

where $\hat{v}$ is the highest (indirect) utility of the solution to this problem.

We consider the replicator dynamics (Sandholm, 2010): $\frac{d\boldsymbol{\lambda}}{dt} = \boldsymbol{F}(\boldsymbol{\lambda}, \phi)$, where $\boldsymbol{\lambda} = (\lambda_i \mid i \in N)$, $\boldsymbol{F}(\boldsymbol{\lambda}, \phi) = (F_i(\boldsymbol{\lambda}, \phi) \mid i \in N)$, and:

$$F_i(\boldsymbol{\lambda}, \phi) = (v_i(\boldsymbol{\lambda}, \phi) - \bar{v}(\boldsymbol{\lambda}, \phi))\lambda_i, \quad i \in N. \tag{2}$$

Here, $\bar{v} = \sum_{i \in N} \lambda_i v_i$ represents the weighted average utility and $\phi \in (0, 1)$ is the *trade freeness*, which is an inverse measure of transportation costs. We choose the freeness of trade as the bifurcation parameter in order to capture the historical tendency of falling/increasing transport costs, as is customary in (new) economic geography.[9]

A set of stable spatial equilibria is obtained as a set of stable and sustainable stationary points of the replicator dynamics (Sandholm, 2010). Stationary points (rest points) $(\boldsymbol{\lambda}, \phi)$ are defined as solutions of the static governing equation

$$\boldsymbol{F}(\boldsymbol{\lambda}, \phi) = \mathbf{0}. \tag{3}$$

2.2. Full agglomeration to a single and twin cities

We introduce two kinds of important patterns (see, e.g., these patterns for $K = 5$ cities in Fig. 4). As a candidate of a core place that accommodates satellite around it, we resort to the full agglomeration (FA) to a single place located δ steps away from the center, i.e.,

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\delta}^{\mathrm{FA}} \text{ with } \lambda_{k-\delta} = 1 \text{ for some } \delta \ (0 \le \delta \le k).$$

[9]We do not disregard the important role of other costs in determining the size and distribution of cities, such as congestion or commuting costs. For a generic spatial economic model, we thus think of ϕ as an index that captures integration between regions in the broadest sense possible (i.e., it may reflect export hurdles due to trade tariffs, the quality of transportation infrastructures, or any kind of institutional barrier.

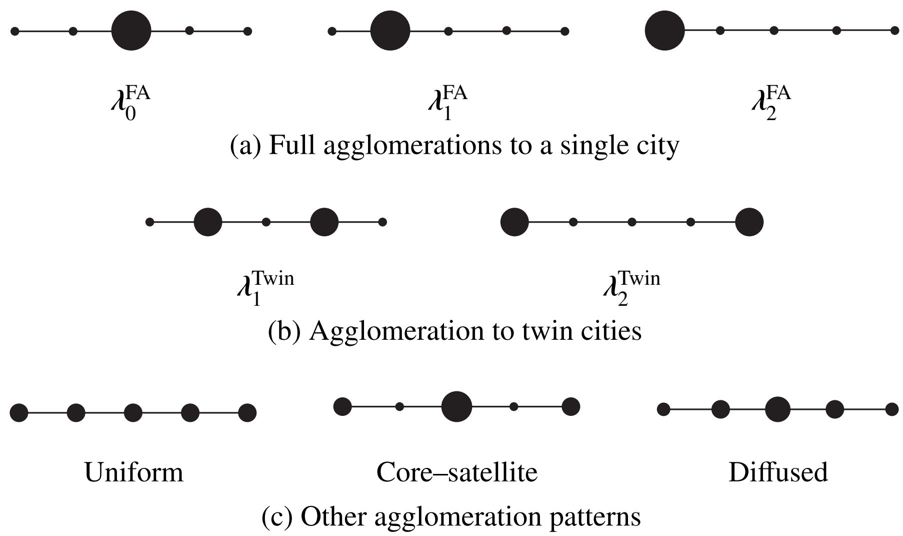


Figure 4: Agglomeration patterns in a long narrow economy

As a candidate of satellite cities around the core place, we resort to agglomeration to the twin cities located δ steps away from the center, i.e.,

$$\boldsymbol{\lambda}_\delta^{\mathrm{Twin}} \ \text{ with } \ \lambda_{k+\delta} = \lambda_{k-\delta} = 1/2 \ \text{ for some } \ \delta \ (1 \leq \delta \leq k).$$

Moreover, these two kinds of agglomeration patterns $\boldsymbol{\lambda}_\delta^{\mathrm{FA}}$ and $\boldsymbol{\lambda}_\delta^{\mathrm{Twin}}$ have special features (called invariant patterns in Ikeda et al., 2012, 2018, and Aizawa et al., 2020) as explained below.

Proposition 1. *$\boldsymbol{\lambda} = \boldsymbol{\lambda}_\delta^{\mathrm{FA}}$ and $\boldsymbol{\lambda}_\delta^{\mathrm{Twin}}$ are stationary points of the replicator dynamics for any values of the trade freeness ϕ (and any value of any other parameter).*

Proof. See Appendix C.1 for the proof. □

Other patterns, such as the uniform, core–satellite, and diffused patterns in Fig. 4(c), appear as stable and sustainable equilibria for the FE model in Section 4. Each of these patterns, however, is a stationary point only for some specific value of ϕ.

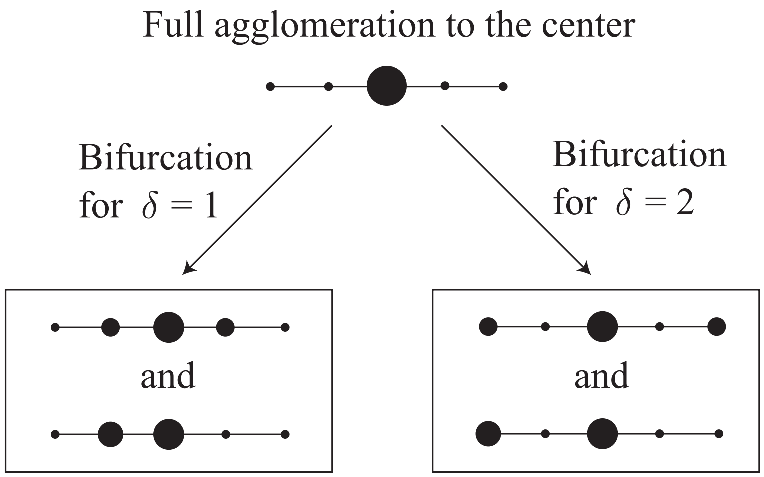


Figure 5: Possible bifurcations for $K = 5$ cities

2.3. Bifurcation from a full agglomeration at the center

To elucidate the mechanism of the emergence of satellite cities around a large core place, we investigate the bifurcation from a state of full agglomeration (FA) at the center $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0^{\mathrm{FA}}$. This state turns out to be much superior in sustainability to full agglomerations elsewhere (refer to Section 4). Since the state has the bilateral symmetry about the center, the indirect utility in cities $i = k \pm \delta$ satisfies $v_{k-\delta}(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi) = v_{k+\delta}(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi)$ $(\delta \in N_\delta)$.

The full agglomeration $\boldsymbol{\lambda}^{\mathrm{FA}}$ at the center has a critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$ (for some $\delta \in N_\delta$) where $v_{k\pm\delta} - v_k = 0$ is satisfied (see Appendix C.2). We can show the emergence of one or two satellite cities, δ steps away from the central region, branching from this critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$, as stated in the following Proposition and is illustrated for $K = 5$ cities in Fig. 5.

Proposition 2. *The critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$ is a corner bifurcation point with two kinds of bifurcating solutions that have either two satellite cities ($\lambda_i > 0$ at $i = k,\ k \pm \delta$) or one satellite city ($\lambda_i > 0$ at $i = k,\ k - \delta$ or $i = k,\ k + \delta$).*

Proof. See Lemma 2 in Appendix C.3 for the proof. □

Full agglomeration $\boldsymbol{\lambda}^{\mathrm{FA}}$ is sustainable if $v_{k-\delta} - \bar{v} = v_{k-\delta} - v_k < 0$ ($\forall\delta \in N_\delta$), that is, $(\max_{\delta \in N_\delta} v_{k-\delta}) - v_k < 0$. In other words, agglomeration at the central city is economically sustainable if the indirect utility there is higher than the highest indirect utility across all potential satellite cities with zero population (i.e., no mobile agents). We use the following assumption on sustainability, with reference to the behavior of the FE model (Proposition 5 in Section 3).

Assumption 1. *Among the corner bifurcation points, there is a sustain bifurcation point* $\phi^{\mathrm{s}} \equiv \max_{\delta \in N_\delta} \phi^{\mathrm{c}}_\delta$ *and* $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi)$ *is sustainable for* $\phi > \phi^{\mathrm{s}}$ *and is unsustainable for* $\phi < \phi^{\mathrm{s}}$.

This assumption turns out to be true for most early (new) economic geography models, namely the seminal Core–Periphery model (Krugman, 1991) and their "identical twins" (Robert-Nicoud, 2005), such as the class of footloose entrepreneur models (the FE model analysed here included), as it was conceived as a way to explain how large spatial imbalances (like the full agglomeration) have increased tremendously as a result of the historical sharp decline in transport costs.

How stable and sustainable bifurcating equilibria is engendered from the corner bifurcation points, which is of great interest, is described as follows.

Proposition 3. *(i) Just after bifurcation, the sustain bifurcation point has zero or one bifurcating path that is stable and sustainable, whereas other corner bifurcation points have no stable and sustainable bifurcating path.*

(ii) The stable and sustainable bifurcating path, if it exists, branches in the direction of decreasing trade freeness ($\phi < \phi^{\mathrm{c}}_\delta$).

Proof. See Appendix C.4 for the proof. □

For the FE model (Section 4.2), stable and sustainable bifurcating equilibria with the twin satellite cities are observed, but those with a single satellite city are never observed. By Proposition 3(ii), a stable and sustainable bifurcating equilibrium, just after bifurcation from the sustain bifurcation point, exists only in the direction of decreasing trade freeness ϕ, engendering a stable state of one or two satellite cities. If we observe this bifurcation behavior conversely, following a historical trend of increasing trade freeness ϕ, we see an emergence of a sustainable state of full agglomeration at the center by steadily absorbing and finally nullifying the (mobile) population of satellite cities.

3. Satellite city formation around full agglomeration for the FE model

The general bifurcation mechanism of full agglomeration $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0^{\mathrm{FA}}$ to the city at the center was presented in the previous section to elucidate the mechanism of *how* satellite cities form. In this section, to elucidate the mechanism of *where* satellite cities form, we investigate stability and sustainability of bifurcating equilibria in more detail for an analytically solvable Core–Periphery model, – the footloose entrepreneur (FE) model – proposed by Forslid and Ottaviano (2003).

3.1. Basic assumptions for the FE model

A multi-regional version of the FE model is briefly introduced, whereas details are given in Appendix B. There are two factors of production (skilled and unskilled labor), and two sectors (manufacturing, M, and agriculture, A). The H skilled and L unskilled workers consume final goods of two types: manufacturing sector goods and an agricultural sector good. Workers supply one unit of each type of labor inelastically. Skilled workers are mobile among cities. The number of skilled workers in city $i \in N$ is denoted by λ_i under the constraint $\sum_{i \in N} \lambda_i = 1$. Unskilled workers are immobile and distributed equally across all cities with $L_i = L/K$ for all $i \in N$.

Preferences U over the M-sector and A-sector goods are identical across individuals. The utility of an individual in city i is

$$U(C_i^{\mathrm{M}}, C_i^{\mathrm{A}}) = \mu \ln C_i^{\mathrm{M}} + (1 - \mu) \ln C_i^{\mathrm{A}} \qquad (0 < \mu < 1), \tag{4}$$

where μ is a constant parameter expressing the expenditure share of manufacturing sector goods, C_i^{A} stands for the consumption of the A-sector product in city i, and C_i^{M} represents the manufacturing aggregate in city i, defined as $C_i^{\mathrm{M}} \equiv \left(\sum_{j \in N} \int_0^{n_j} q_{ji}(\ell)^{(\sigma-1)/\sigma} d\ell\right)^{\sigma/(\sigma-1)}$, where $q_{ji}(\ell)$ represents the consumption in city $i \in N$ of a variety $\ell \in [0, n_j]$ produced in city $j \in N$, n_j stands for the number of produced varieties at city j, and $\sigma > 1$ denotes the constant elasticity of substitution between any two varieties.

The transportation costs for M-sector goods are assumed to take the iceberg form. That is, for each unit of M-sector goods transported from city i to city $j \neq i$, only a fraction $1/\tau_{ij} < 1$ actually arrives ($\tau_{ii} = 1$). It is assumed that $\tau_{ij} = \exp(\tau\, m(i, j)\, \tilde{L})$ is a function of a transport cost parameter $\tau > 0$, where $m(i, j)$ is an integer expressing the road distance between cities i and j and $\tilde{L}$ is the distance unit. As our bifurcation parameter, we introduce the trade freeness (a converse measure of transport costs)

$$\phi = \exp[-\tau(\sigma - 1)\tilde{L}] \in (0, 1). \tag{5}$$

The market equilibrium wage vector $\boldsymbol{w} = (w_i)$ can be obtained analytically ((B.10) in Appendix B). Indirect utility v_i is expressed in terms of w_i and $\Delta_i = \sum_{k \in N} d_{ki} \lambda_k$ as[10]

$$v_i = \frac{\mu}{\sigma - 1} \ln \Delta_i + \ln w_i. \tag{6}$$

By virtue of the analytical solvability of the FE model, the indirect utility at each city $i = k \pm \delta$ ($\delta \in N_\delta$) for $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0^{\mathrm{FA}}$ is expressed explicitly as (see Appendix D.1)[11]

$$v_k = \ln \frac{\theta}{1 - \theta}(2k + 1), \qquad \theta = \frac{\mu}{\sigma} \in (0, 1); \tag{7}$$

$$v_{k\pm\delta} = \ln \frac{\theta}{1 - \theta} + \frac{\delta\mu}{\sigma - 1} \ln \phi + \ln \left((\theta k + k + 1)\phi^{\delta} + (1 - \theta) \left[(k - \delta)\phi^{-\delta} + \sum_{p=1}^{\delta} \phi^{\delta - 2p} \right] \right). \tag{8}$$

[10]The spatial discounting factor $d_{ji} = \tau_{ji}^{1-\sigma} = \phi^{m(i,j)}$ represents friction between cities j and i that decays in proportion to the transportation distance. In our formulation, which relies on d_{ji}, the distance unit $\tilde{L}$ need not be specified.

[11]The choice of the total population L of low skilled workers is not influential on the results as the payoff is linear in L (see also Gaspar et al. (2019, pp. 9) for a detailed explanation). For simplicity, we set $L = K$.

By bilateral symmetry of the full agglomeration, we have $v_{k-\delta} = v_{k+\delta}$ (since $\lambda_{k-\delta} = \lambda_{k+\delta} = 0$). We, therefore, consider only the cities on the left hand side of the economy labeled by $i = \{0, \ldots, k\}$ in the discussion below. We hereafter assume the no-black-hole condition $\mu < \sigma - 1$ (Forslid and Ottaviano, 2003) since its violation is quite exceptional and empirically unrealistic.[12]

3.2. Corner and sustain bifurcation points

We march on to investigate the bifurcation from the state of full agglomeration $(\lambda_0^{\mathrm{FA}}, \phi)$. The existence of a corner bifurcation point engendering bifurcating solutions with one or two satellite cities can be shown by the following Proposition. The uniqueness of the corner bifurcation point is dependent on the steps (points on the line segment, or cities) away from the central city δ and is guaranteed for $\delta \leq 6$. For $\delta \geq 7$, we are yet to analytically prove the uniqueness.

Proposition 4. *(i) There is a corner bifurcation point satisfying* $v_{k-\delta} - v_k = 0$ *for each* $\delta \in N_\delta$. *(ii) There exists one unique corner bifurcation point* $\phi_\delta^{\mathrm{c}} \in (0, 1)$ *(possibly a sustain bifurcation point) for each* $\delta \in \{1, 2, \ldots, 6\}$ *and for any number of cities* $(K \geq 2\delta + 1)$.

Proof. See Appendix D.2 and Appendix E. □

Denote by ϕ_δ^{c} the largest ϕ satisfying $v_{k-\delta} - v_k = 0$ for each $\delta \in N_\delta$ and set: $\phi^{\mathrm{s}} = \max\limits_{\delta \in N^\delta} \phi_\delta^{\mathrm{c}}$. Based on Lemma 5 in Appendix D.3. and on the Intermediate Value Theorem, a sustain point ϕ^{s} always exists for the full agglomeration $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi)$, similarly to

[12] Anderson and Wincoop (2004), for instance, find that the elasticity of substitution σ is likely to range between 5 and 10.

the two-place economy for most spatial economic models. The next result establishes sustainability of full agglomeration in relation with the sustain point.

Proposition 5. *There exists a sustain point at ϕ^{s} on the full agglomeration at the center and the full agglomeration is sustainable for $\phi > \phi^{\mathrm{s}}$.*

Proof. See Appendix D.4 for the proof. □

This proposition underpins Assumption 1 in Section 2. Of course, for $K = 3$, the bifurcation point is unique and corresponds to the sustain point.

The sustain bifurcation point ϕ^{s} is dependent on σ and μ as explained below, displaying the same tendency as the two-place economy (e.g., Fujita et al., 1999).

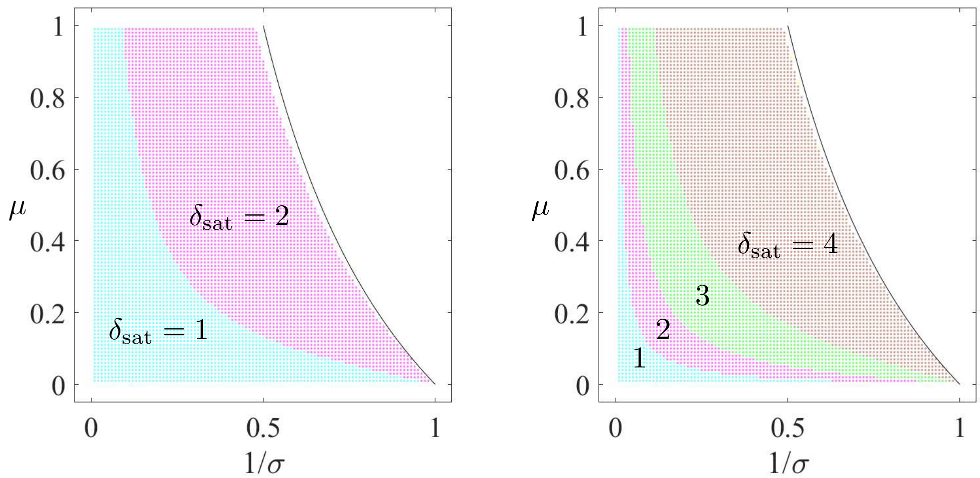
Proposition 6. *As the elasticity of substitution σ increases and/or the fraction of income spent on manufactures μ decreases, the sustain point increases.*

Proof. See Appendix D.5 for the proof. □

This is in line with the intuition that, for a larger σ, scale economies become weaker as goods become more substitutable, which mitigates the agglomeration forces that promote the full agglomeration of industry. On the other hand, an increase of the expenditure share μ on manufactured goods expands the relative size of the industrial sector, which favors full agglomeration.

3.3. Location of satellite cities

Recall that a major target of this paper is to answer the question "where do satellite cities form?" As an index for the location of a satellite city for the sustain bifurcation



$K = 5$ $\qquad$ $K = 11$

Figure 6: Dependence of the location of satellite cities emerging from the sustain point on the values of the parameters σ and μ (solid line: $\mu = \sigma - 1$)

point, we denote by $\delta_{\rm sat}$ the integer δ that maximizes $\phi^{\rm c}_{\delta}$; then the sustain point is associated with this location, i.e., $\phi^{\rm s} = \phi^{\rm c}_{\delta_{\rm sat}}$. This location is dependent on the microeconomic parameters σ and μ, as well as on the number K of cities as explained below.

First, we investigate the dependence of the location of satellite cities, given by the number of steps $\delta_{\rm sat}$ away from the center, on the values of the parameters σ and μ. For $K = 5$ and 11 cities, Fig. 6 depicts the contour of $\delta_{\rm sat}$ in the space of μ and $1/\sigma$ in the range $(0, 1) \times (0, 1)$ that is obtained numerically. There is a white zone ($\mu > \sigma - 1$) at the upper right corner, where the full agglomeration is always sustainable and no satellite city emerges. It is to be noted that, the parameter zone for the the border city ($\delta_{\rm sat} = 5$) is not discernible for $K = 11$. Thus, locations too far from the center are not suitable for the accommodation of satellite cities.

As $1/\sigma$ and/or μ increases, $\delta_{\rm sat}$ increases one by one from the smallest value of $\delta_{\rm sat} = 1$. That is, in association with an increase of agglomeration forces due to stronger scale

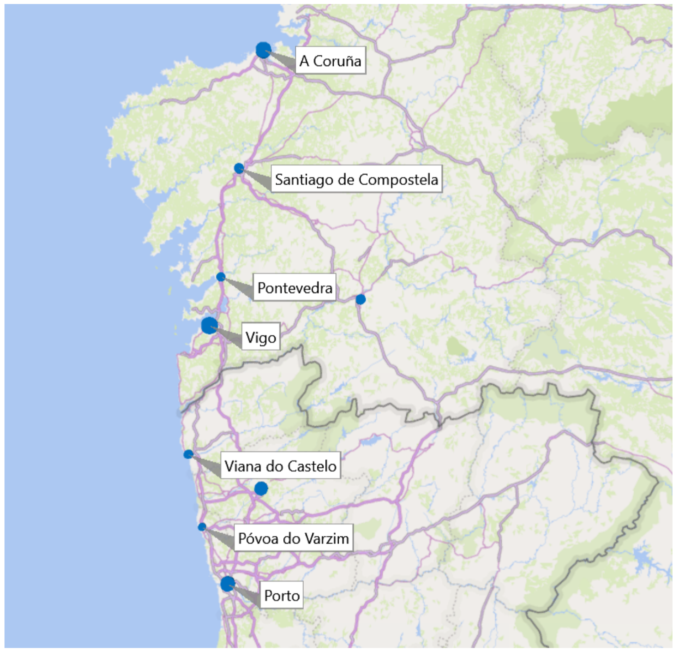


Figure 7: Atlantic Axis mega-region. Population from 2017, Eurostat.

economies or a larger size of the manufacturing sector (resp., a decrease in σ and/or an increase in μ), the satellite cities tend to form away from the primary city at the center, thereby forming an *agglomeration shadow* (Arthur, 1990; Ikeda et al., Fig. 5, 2017).

Empirically, this resembles the case of the Atlantic Axis, in Fig. 7 when restricted to 7 cities, from Porto southwards to La Coruña in the North, where we can see the latter at the borders of the corridor acting as large cities three steps away from the bigger central city (Vigo). Indeed, both in Portugal and Spain there is a large size of the business and industrial tissue concentrated in the northwestern coastal provinces, thus corroborating our predictions that higher agglomeration forces push satellite cities away from the center.

By contrast, as agglomeration forces decrease, the satellite cities tend to locate closer to the primary city, thereby forming a hump-shaped megalopolis around this city for $\delta_{\text{sat}} = 1$. Thus, we have observed the dependence of agglomeration patterns on the values of microeconomic parameters, which possibly are a source of the diversity of the population distribution of a chain of cities observed worldwide.

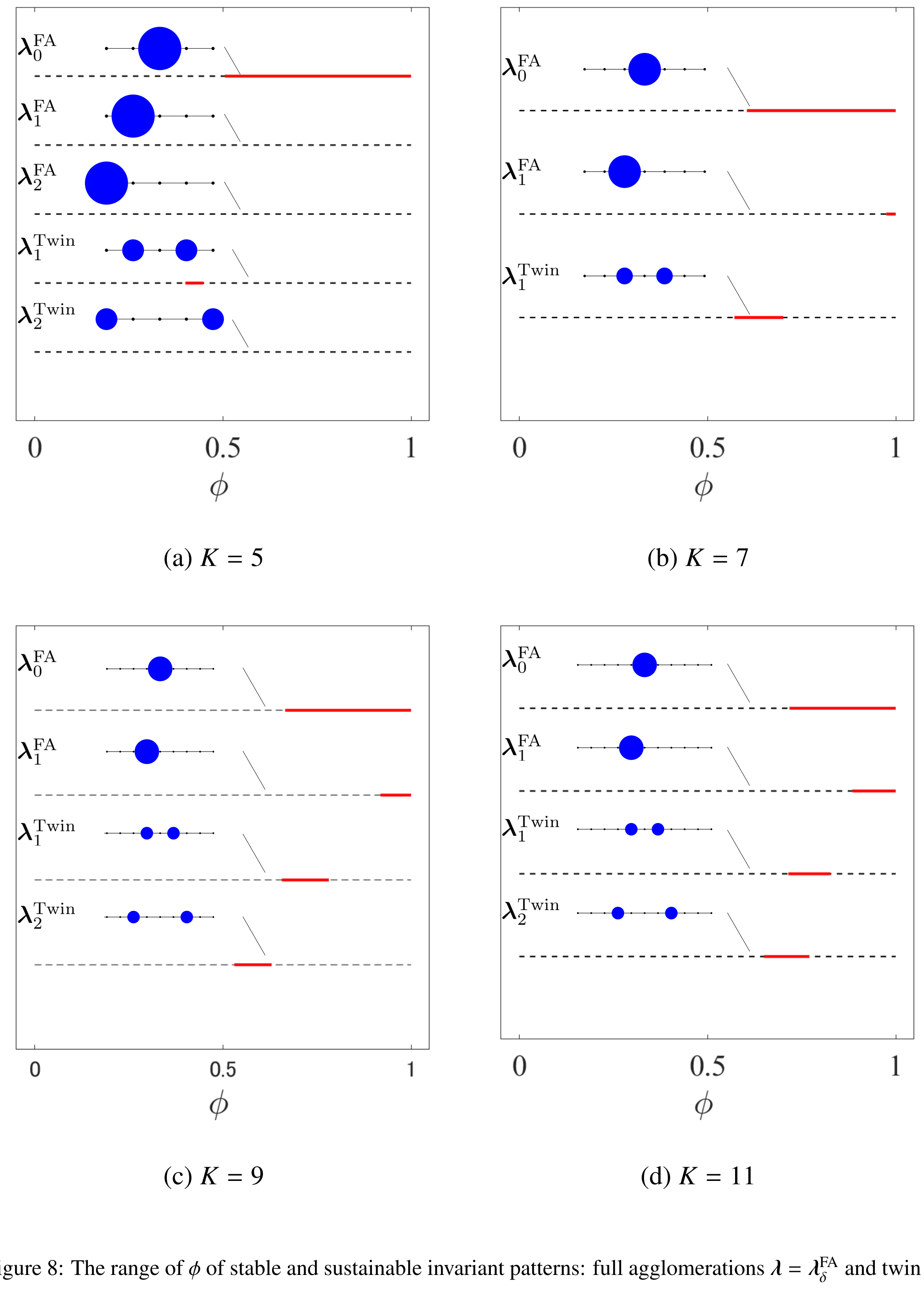
4. Transition of stable and sustainable equilibria for the FE model

We now study the agglomeration mechanism of the FE model in a long narrow economy as the trade freeness increases. The economy with five cities is employed as the standard model of a chain of cities, such as (1) Boston, Hartford, New York City, Philadelphia, and Baltimore–Washington in the East Atlantic and (2) Sendai, Tokyo–Yokohama, Nagoya, Osaka–Kobe, and Hiroshima in the Main Island of Japan. The former is closer to a full agglomeration at the center, while the latter to twin cities. The economy with more than five cities is also analyzed to provide insights on how megalopolises, along narrow corridors with an increased number of cities, behave.

We use $(\sigma, \mu) = (6.0, 0.4)$, which satisfies the no-black-hole condition ($\mu < \sigma - 1$), follows Footnote 12, and is often used as a benchmark case in economic geography models. It should be noted that the choice of benchmark parameter values in Sections 4.1–4.3, particularly an intermediate to high value of $\sigma = 6$, implies weaker returns to scale. As we have observed in Fig. 6 in Section 3.3, the parameter values $(\sigma, \mu) = (6, 0.4)$ imply that $\delta_{sat} = 1$ when $K = 5$; accordingly, weaker agglomeration forces engender satellite cities closer to the central region. Other choices of parameter values could well lead to different qualitative results. In particular, stronger agglomeration forces would engender satellite cities outwards (or towards the borders).

4.1. Stability and sustainability of full agglomerations and twin cities

As candidates of core places, we consider the states of full agglomerations $\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\delta}^{\mathrm{FA}}$, while the twin cities $\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\delta}^{\mathrm{Twin}}$ as candidates of satellite cities. These two kinds of states are stationary points for any ϕ (Proposition 1). Among these states, we investigate which ones are superior in stability and sustainability, and the transition of such superior ones as



(a) $K = 5$

(b) $K = 7$

(c) $K = 9$

(d) $K = 11$

Figure 8: The range of ϕ of stable and sustainable invariant patterns: full agglomerations $\lambda = \lambda_\delta^{\rm FA}$ and twin cities $\lambda = \lambda_\delta^{\rm Twin}$ (unstable and/or unsustainable patterns are included only for $K = 5$; $(\sigma, \mu) = (6.0, 0.4)$; red solid line: stable and sustainable; broken line: unstable and/or unsustainable)

the value of ϕ changes in $\phi \in (0, 1)$. Such transition is to be observed in the comparative static analysis in the next subsection.

As basic data for this investigation, the ranges of ϕ in which these patterns are stable and sustainable are depicted by red solid lines in Fig. 8 for $K = 5, 7, 9$, and 11 cities.

First, we consider the full agglomerations. For each city size K, the full agglomeration $\boldsymbol{\lambda} = \boldsymbol{\lambda}_0^{\mathrm{FA}}$ at the center has the longest range of sustainable state $\phi \in (\phi^{\mathrm{s}}_{\delta=0}, 1)$ and is the one which becomes sustainable first, when the trade freeness increases from a low value. In contrast, full agglomeration $\boldsymbol{\lambda}_\delta^{\mathrm{FA}}$ ($\delta \geq 1$) in the city away from the center is much inferior in stability and sustainability.[13] We, accordingly, specifically examine this full agglomeration to the center, which is the most advantageous location of economic activity compared to potential core cities placed anywhere else, in the following subsections.

Next, we investigate twin cities. For the economy with a relatively small number of cities ($K = 5, 7$ shown in Figs. 8(a) and (b)), only the twin cities $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1^{\mathrm{Twin}}$ located one step away from the center have stable equilibria with a short range of ϕ (near $\phi = 0.44$ for $K = 5$ and $\phi = 0.65$ for $K = 7$). For more cities ($K = 9, 11$ in Fig. 8(c), (d)), the location of stable and sustainable twin cities extends outwards ($1 \leq \delta \leq 2$). Although all of these twin cities have only short ranges of ϕ for stable equilibria, some of their ranges cover smaller values of ϕ that the full agglomeration $\boldsymbol{\lambda}_0^{\mathrm{FA}}$ at the center cannot cover. This demonstrates the importance of the state of twin cites for an intermediate

[13]The full agglomeration $\boldsymbol{\lambda}_1^{\mathrm{FA}}$ one step way from the center is unsustainable for $K = 5$ (Figs 8(a)), and is sustainable only for a shorter range $\phi \in (\phi^{\mathrm{s}}_{\delta=1}, 1)$ ($\phi^{\mathrm{s}}_{\delta=1} > \phi^{\mathrm{s}}_{\delta=0}$), for more cities $K = 7, 9, 11$ (Figs. 8(b), (c), (d)). Full agglomeration in the city furthest away $\delta = k$ is unsustainable for any ϕ.

value of ϕ, whereas the full agglomeration state dominates for a large ϕ. It implies an inevitable transition from the twin cities to the full agglomeration as ϕ increases from an intermediate to a large value, as we will actually see in the comparative static analysis in the next subsection.

4.2. Bifurcating equilibria from the full agglomeration

The paths of equilibria branching from the full agglomeration at the center were obtained by comparative static analysis for $K = \{5, 7, 9, 11\}$ cities following the analysis procedure in Appendix F. Figures 9 and 10 plot the paths for $K = 5$ and $K = 7$, respectively, whereas the paths for larger number of cities $K = \{9, 11\}$ are given in Appendix F and are referred to from time to time to support the discussion. The vertical axis is the population λ_k at the central region $i = k$ ($k = (K - 1)/2$) and the horizontal axis is the trade freeness ϕ ($0 < \phi < 1$). The stable and sustainable equilibria are shown by solid curves and unstable and/or unsustainable ones by broken lines.

To see how and where satellite cities form, we investigate the bifurcation from the state of the full agglomeration λ_0^{FA} at the center. This state corresponds to the horizontal line at $\lambda_k = 1$ in Figs. 9 and 10. To begin with, this state λ_0^{FA} has a unique critical (sustain or bifurcation) point for each δ as predicted by Proposition 4(ii) (e.g., $\delta = 1$ for the Point I and $\delta = 2$ for the Point K for an economy with $K = 5$ cities). For each K, one of these bifurcation points is the sustain bifurcation point I and the full agglomeration is sustainable during the Path IJ ($\phi \in (\phi^{\mathrm{s}}, 1)$) (Proposition 5).

We search for stable and sustainable bifurcating paths from the sustain bifurcation

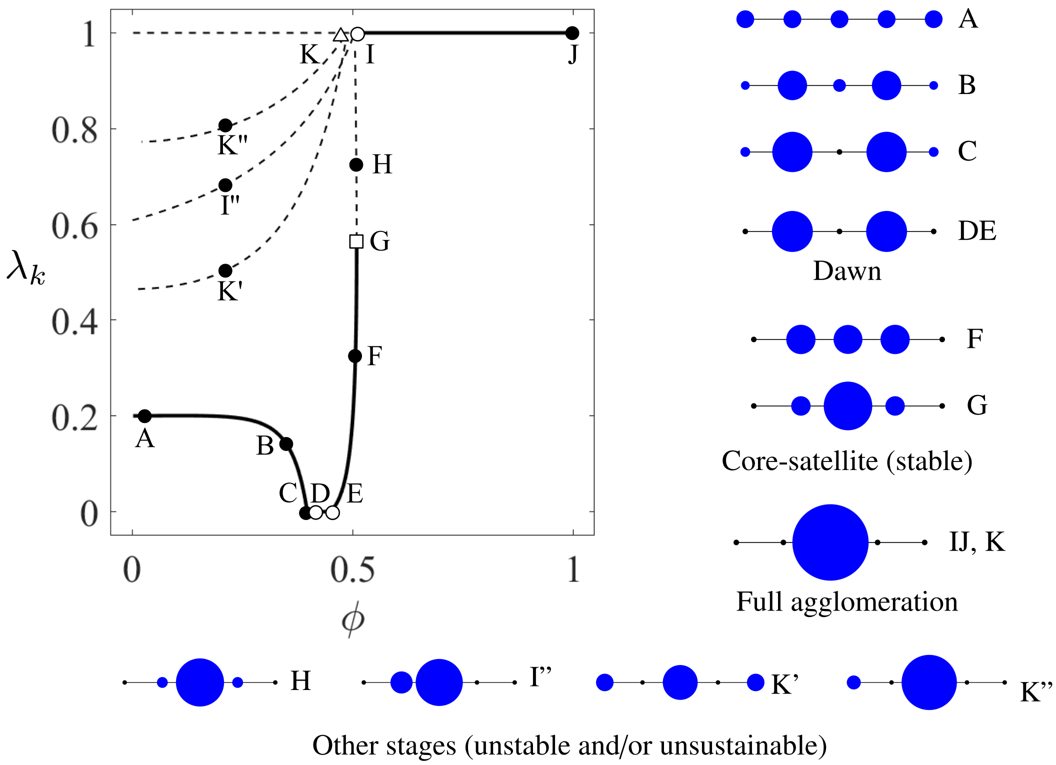


Figure 9: Paths of equilibria for $K = 5$ cities for $(\sigma, \mu) = (6.0, 0.4)$ (solid line: stable and sustainable; broken line: unstable and/or unsustainable; △: bifurcation point; ○: sustain point; □: maximum point of ϕ)

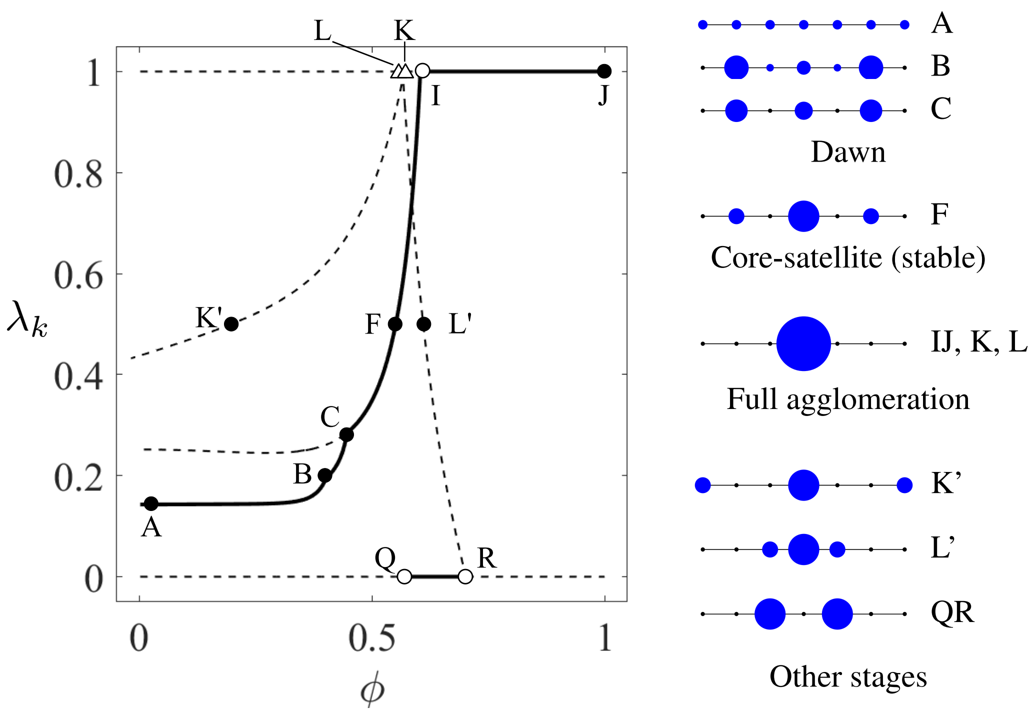


Figure 10: Paths of equilibria for $K = 7$ cities for $(\sigma, \mu) = (6.0, 0.4)$ (solid line: stable and sustainable; broken line: unstable and/or unsustainable; △: bifurcation point; ○: sustain point)

point I.[14] We found stable and sustainable bifurcating paths with twin satellite cities,[15] whereas bifurcating paths with one satellite city were all unstable and/or unsustainable. Thus, bifurcating paths with twin satellite cities are superior in stability and, accordingly, are of most economical importance. Such superior stability is due to the natural symmetry of the line segment around the central region, which realizes balanced economic activities in both sides of the economy.

4.3. Transition of stable equilibria as the trade freeness increases

We conduct comparative static analysis with respect to the trade freeness ϕ to observe the progress of stable equilibria as ϕ increases. As the trade freeness ϕ increases from a very low value, irrespective of the number K of cities, we can observe three characteristic and distinctive stages of stable equilibria: (1) *Dawn* stage, (2) *Core–satellite* stage, and (3) *Full agglomeration* stage, in this order.

In the Dawn stage, an almost uniform population distribution prevails for a very low value of ϕ. As ϕ increases, odd numbered cities ($i = 1, 3, \ldots, 2k-1$) grow, while even numbered cities ($i = 0, 2, \ldots, 2k$) that include border cities shrink. For instance, for the Path DE for $K = 5$ in Fig. 9, there appears an agglomeration to every other city at $i = \{1, 3\}$ and no population at another every other city at $i = \{0, 2, 4\}$. This looks like a chain of spatially repeated core–periphery patterns *a la* Christaller and Lösch (e.g.,

[14]All bifurcating paths other than those of the sustain point are unsustainable by Proposition 3(ii) and inferior in stability and sustainability.

[15]The bifurcating path IHG for $K = 5$ cities is unstable just after bifurcation but regains stability at the limit point of ϕ (Point G shown as □), where satellite cities grow to have significant population size (Fig. 9).

Fujita and Mori, 1997). This behavior is called the *spatial period doubling* since the spatial period between agglomerated places is doubled from 1 to 2 steps between cities.[16]

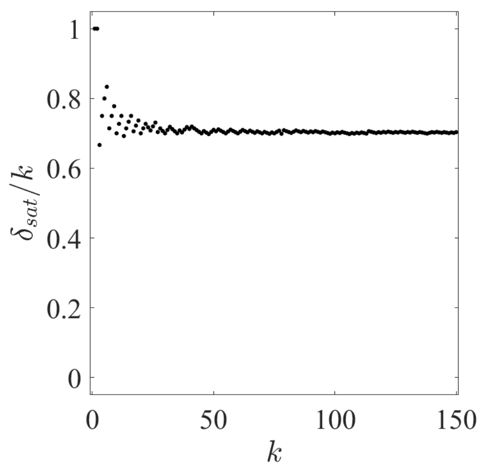
After the Dawn stage, the economy evolves to the Core–satellite stage with a large central city and twin satellite cities (Point F in Figs. 9 , 10 and F.1, F.2 in Appendix F). As ϕ increases further, the core city at the center grows and the twin satellite cities shrink, eventually leading to the Full agglomeration stage for sustainable λ^{FA} (Path IJ in each figure).

4.4. Location of satellite cities for a large number of cities

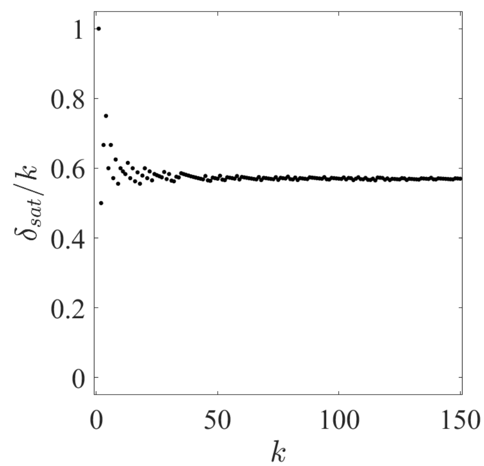
As we have seen in Figs. 9, 10, F.1, and F.2, the location of satellite cities for the sustain bifurcation point is $\delta_{\mathrm{sat}} = 1$ for $K = 5$ cities, $\delta_{\mathrm{sat}} = 2$ for $K = 7$ and $K = 9$, $\delta_{\mathrm{sat}} = 3$ for $K = 11$. Thus, satellite cities locate further away from the center as the number K of cities increase. As an index for the optimal location of the satellite cities, we introduce a normalized length from the center, being defined as δ_{sat}/k. Figure 11 plots the normalized length δ_{sat}/k for $\mu = 0.4$ and $\sigma = \{2.5, 6.0, 10.0\}$ against $k = (K-1)/2$. As k increases to a large value, such as $k = 50$, the location of satellite cities becomes convergent to $\delta_{\mathrm{sat}}/k \approx \{0.70, 0.58, 0.48\}$ for $\sigma = \{2.5, 6.0, 10.0\}$, respectively.

These results are in accordance with our findings so far. With a low σ, agglomeration forces are strong, and thus the optimal location location is closer to 1. For a higher σ, the optimal location is just above half the length from the center. Finally, an even higher σ implies an optimal location of satellites just below 1/2.

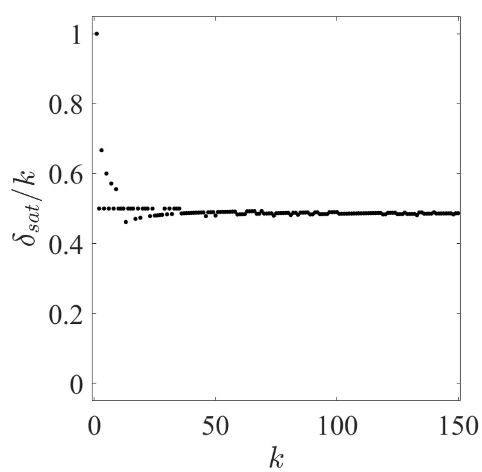
[16]Such doubling is studied for a long narrow economy (Ikeda et al., 2017) and for a racetrack economy (Tabuchi and Thisse, 2011; Ikeda et al., 2012; and Akamatsu et al., 2012).



(a) σ = 2.5



(b) σ = 6.0



(c) σ = 10.0

Figure 11: Convergence of normalized length δ_{sat}/k from the center for a large k.

5. Agglomeration patterns in the real world

We have seen several patterns of satellite cities around a large city at the center in Section 4. There patterns are compared with worldwide city size distributions.

Taking Japan in Fig. 1(b) as reference with $K = 5$ regions, one could infer that its configuration resembles that of the Dawn stage at pattern B (Fig. 9), with Nagoya at the centre being highly populated but scarcely if compared to the gigantic satellites of Osaka and Tokyo at $\delta = 1$, and with the border of regions of Hiroshima and Sendai at $\delta = 2$ a little smaller than Nagoya. However, when we compare the five cities' population in the 1950's (shown by blue arcs in Fig. 12) with the population in 2020 (shown by orange arcs),[17] we observe that Nagoya has grown relatively more compared to Tokyo and Osaka, thus suggesting that Japan may eventually shift to a Core–satellite stage (skipping the path CDE at the Dawn Stage in Fig. 9) and further progress en route to a partial agglomeration at Nagoya. While it seems very unlikely that the capital region of Tokyo may someday become a peripheral region, we remind that we are looking at a narrow corridor of $K = 5$ regions, when in fact a more realistic space economy that depicts Japan could be a line-segment comprised of several regions (and actually either Osaka or Tokyo could be considered as central regions). This is evident from Fig. 12.

The current state of the Golden Banana, when restricted to 5 main regions (Valencia, Barcelona, Marseille, Nice and Genoa), resembles the states F and G during the Core–satellite stage (Fig. 9), whereby Marseille (the central city) and the two neighboring cities have huge populations (see Fig. 13).[18] The exception is that the border regions (Valencia

[17]We have used data provided by the UN, Dept. Econ. & Social Affairs, Population Div. (2018).

[18]Although Nice has less population than Marseille, Barcelona is considerably more populated, which

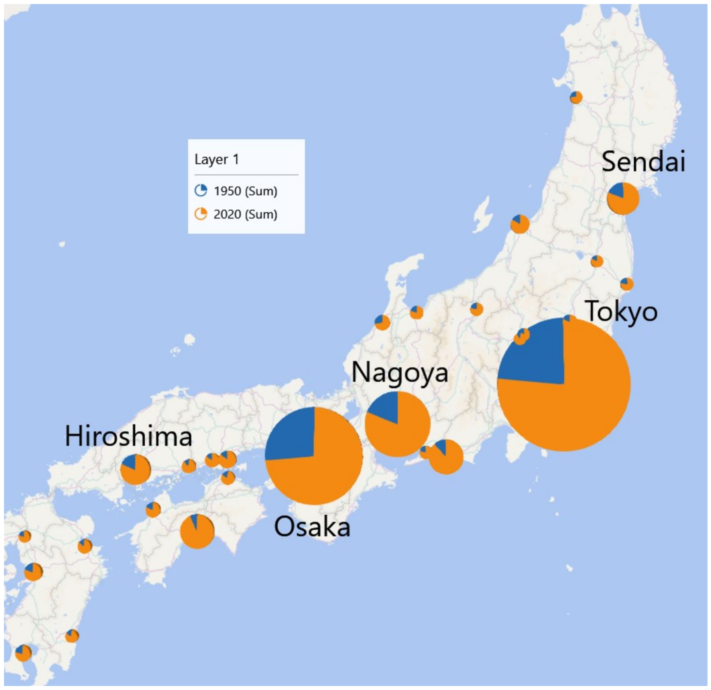


Figure 12: Population in Japanese cities. Blue is population in 1950 and orange is population in 2020. The data was provided by UN, Department of Economic and Social Affairs, Population Division (2018).

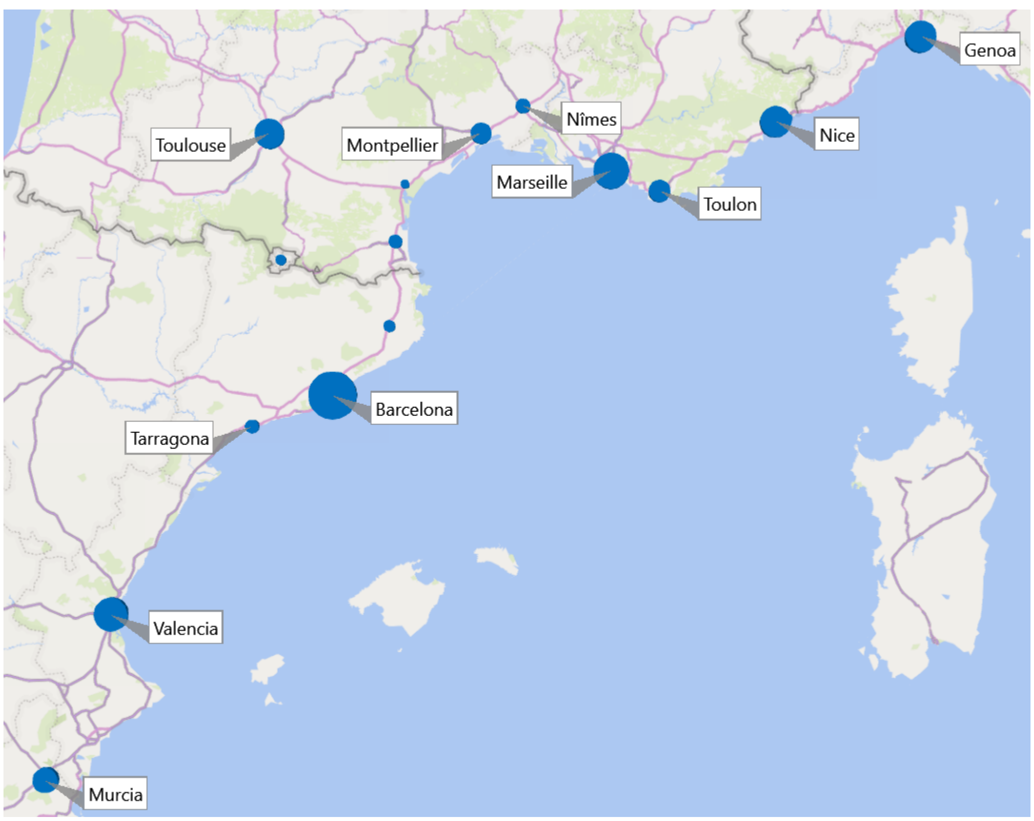


Figure 13: The Golden Banana or European Sunbelt mega-region. Population from 2017, Eurostat.

and Genoa) are also highly populated.[19] On the one hand, this is in accordance with twin cities right next to the central region; on the other hand, the Golden Banana is comprised of countless new industries, which could well imply stronger agglomeration forces in this particular mega-region resulting in higher population at the borders.

A few notes of caution are warranted at this point. First, the choice of parameter values for our simulations in Figs. 9 and 10 is plausible for illustration purposes but may be arbitrary from an empirical perspective. This means that obtaining specific estimates for the megaregions discussed throughout this paper would likely improve our understanding and predictive capacity regarding the location of satellite cities around a central city. This, however, is left for future research.

Second, a great part of the distribution and size of cities may well be related with activities regarding services instead of manufacturing goods, something which the FE model disregards completely.[20]

Third, for most of the empirically depicted cases in this work, particularly in Europe, the mega-regions seem to always have significant populations at the border regions. This could be due to an underestimation of agglomeration forces by our choice of parameter values. If agglomeration forces are instead very strong in these megaregions, then more population at the borders lies in accordance with our predictions.

---

arguably places the Golden Banana between a Dawn State in transition to a Core–Satellite stage. Note however that, since we are talking at a transnational scale, the central region chosen here is arbitrary and could easily correspond to Barcelona instead.

[19]However, as argued at the beginning of Section 4, this might be a result of choosing parameter values under which agglomeration forces are relatively weaker.

[20]We thank Ricardo Gonalves for pointing this out to us.

6. Conclusion

We have conducted a theoretical study on several characteristic agglomeration patterns, such as full agglomeration, twin cities, core–satellite patterns, and spatial period doubling patterns, as prototypes of diverse spatial agglomeration patterns of a chain of cities observed worldwide. We have elucidated the bifurcation mechanism for the full agglomeration at a single big city in a long narrow economy in a manner readily applicable to many spatial economic models. In particular, a sustain bifurcation from the full agglomeration is highlighted as a mechanism to engender a core–satellite pattern with twin satellite cities around a large city. There is a budding of a search of core–satellite patterns in the real population data in Western Germany and Eastern USA (Ikeda et al., 2019).

A remark is on the standpoint of this paper. While it is customary to start from the uniform state, we place emphasis on agglomeration patterns emanating from the completely agglomerated state. Nowadays it would be far more important to investigate the competition between a large central city and satellite cities than to investigate the self-organization of cities in a flat land envisaged in Central Place Theory. Future work will extend this theory to different spatial topologies, such as a "star economy", or regions in two-dimensional space.

A pertinent combination of model-independent general bifurcation mechanism with model-dependent properties, such as stability/sustainability and parameter dependency, is vital in the successful elucidation of the agglomeration mechanism. For the FE model, we have shown that the higher the expenditure share of manufactured goods on income is and/or the lower the elasticity of substitution is, the farther satellite cities emerge from the central city. Conversely, if the size of the industrial sector relative to the traditional sector is very low and/or scale economies are weak, there emerges a hump-shaped megalopolis with satellite cities located side-by-side with the primary central city.

It is pertinent to relate our results with the findings of Turner et al. (2021), who study urbanization patterns in a linear city with three discrete locations and heterogeneous agents as increasing returns to scale in production increase. Granted, the settings differ fundamentally in that their model falls under the class of urban economics whereas the FE model here belongs to the field of new economic geography and does not consider heterogeneous agents nor the internal dimension of regions/cities, i.e., commuting costs have no role in our paper.[21] However, what we wish to highlight is what makes both works relatable: the line segment with discrete locations.

Turner et al. (2021) find that, once increasing returns become sufficiently strong, stronger returns to scale may induce the dispersion of economic activities toward the peripheral (border regions). In their setting, the authors argue that their results hinge heavily on the heterogneity of agents regarding residential location preferences. Thus, to increase comparability between the models we could introduce heterogeneity in our setting as in Castro et al. (2021). However, Castro et al. (2021) also show that the common logit type heterogeneity need not change the predictions of the original FE model.

It would thus be interesting to check if this dispersion process for high increasing returns to scale also implies higher residence in outer regions with more locations in their setting. If the answer is positive, the results are similar to ours, not in the con-

[21]Note that this does not imply in any way that we consider NEG and urban economics competing theories. Rather, they can and should be very much complementary (see Gaspar, 2018; 2021).

ventional sense of dispersion (as in uniformity of the spatial distribution), but in the sense of increasing distance regarding the centre for locations that become more concentrated/populated. This could potentially add to a weak conjecture of geographical scale invariance which would suggest a fractal relationship between spatial configurations at low scales and very large scales.

The progress of stable equilibria when the trade freeness ϕ increases is observed for the FE model. This progress is of great economic interest as it captures the historical process of increasing economic integration and globalization. When the trade freeness ϕ increases from a small value to a large value, we have observed the following three characteristic stages regardless the number of cities. It starts with the Dawn stage with a chain of spatially repeated core–periphery patterns *a la* Christaller and Lösch. As the trade freeness increases, a central city with twin satellite cities emerges in the Core–satellite stage. Thereafter, the central city grows and the twin satellite cities shrink, en route to the Full agglomeration stage, in which the population is completely agglomerated in the central city. Admittedly only for a spatial economic model, this paper has demonstrated a scenario of historical progress of spatial agglomerations in a chain of cities. It will be a topic in the future to investigate the progress of satellite cities' formation for other spatial economic models based on the bifurcation mechanism proposed in this paper.

Among these, we have the early (new) economic geography models such as the original Core-Periphery model by Krugman (1991) and similar ones: the modified version with land instead of immobile workers by Puga (1999), the models with dispersive congestion effects by Helpman (1998) and Tabuchi (1998), or the quasi-linear upper tier utility footloose entrepreneur settings in Ottaviano et al. (2002), Pflüger (2004). More recently, we have the spatial economic models of Tabuchi et al. (2005), Murata and Thisse (2005), Redding and Sturm (2008), Allen and Arkolakis (2014), Redding and Rossi-Hansberg (2017), Behrens and Murata (2021) and Gaspar et al. (2021),[22] to name a few. We refer the reader to Akamatsu et al. (2020) for a complete taxonomy of these and other models in the literature.

We have thus observed diverse agglomeration patterns dependent on the values of trade freeness and on microeconomic parameters. Such dependence possibly is a source of the diversity of the population distribution of a chain of cities observed worldwide.

Acknowledgments

We are very grateful to Liliana Garrido-da-Silva for her very useful comments. We also thank Sofia B.S.D. Castro, João Correia-da-Silva, Ricardo Gonçalves, participants at the 18th Internal Conference, Católica Porto Business School, and participants at the 10th European Meeting of the Urban Economics Association. We gratefully acknowledge the Grant-in-Aid for JSPS 18K04380 and 18K18874. The last author acknowledges the support from Fundação para a Ciência e Tecnologia (through projects UIDB/04105/2020, UIDB/00731/2020, PTDC/EGE-ECO/30080/2017 and CEECIND/02741/2017).

[22] The latter is an extension of the Pflüger (2004) model to three (equidistant/racetrack setting) and more regions (equidistant setting).

Appendix A. Classification of stationary points

Stationary points $(\boldsymbol{\lambda}, \phi)$ of the replicator dynamics are classified into *interior solutions*, for which all cities have positive populations, and *corner solutions*, for which some cities have zero population (i.e., skilled workers). We can appropriately permute the components of $\boldsymbol{\lambda}$, without loss of generality, to arrive at $\hat{\boldsymbol{\lambda}} = (\boldsymbol{\lambda}_+, \boldsymbol{\lambda}_0)$ with $\boldsymbol{\lambda}_+ = \{\lambda_i > 0 \mid i = 0, 1, \ldots, m\}$ and $\boldsymbol{\lambda}_0 = \mathbf{0}$. Whereas $\boldsymbol{\lambda}_0$ is present for a corner solution and is absent for an interior solution, $\boldsymbol{\lambda}_+$ is present for both solutions. The static governing equation (3) and associated Jacobian matrix can be rearranged, respectively, as (Ikeda et al., 2012)

$$\hat{\boldsymbol{F}} = \begin{pmatrix} \boldsymbol{F}_+(\boldsymbol{\lambda}_+, \boldsymbol{\lambda}_0, \phi) \\ \boldsymbol{F}_0(\boldsymbol{\lambda}_+, \boldsymbol{\lambda}_0, \phi) \end{pmatrix}, \qquad \hat{J} = \frac{\partial \hat{\boldsymbol{F}}}{\partial \hat{\boldsymbol{\lambda}}} = \begin{pmatrix} J_+ & J_{+0} \\ O & J_0 \end{pmatrix}, \tag{A.1}$$

where $J_0 = \mathrm{diag}(v_{m+1} - \bar{v}, \ldots, v_{K-1} - \bar{v})$ and $\mathrm{diag}(\cdots)$ denotes a diagonal matrix with the entries in parentheses.

A stable spatial equilibrium is given by a stable and sustainable stationary solution, for which all eigenvalues of $\hat{J}$ are negative. We have the conditions:

$$\begin{cases} \text{Stability condition:} & \text{all eigenvalues of } J_+ \text{ are negative.} \\ \text{Sustainability condition:} & \text{all diagonal entries of } J_0 \text{ are negative.} \end{cases} \tag{A.2}$$

Critical points are those which have one or more zero eigenvalue(s) of the Jacobian matrix $\hat{J}$. Critical points are classified into a *break* bifurcation point with singular J_+, a *corner* bifurcation point with singular J_0, and a *limit point* of ϕ, with singular J_+.

Appendix B. Details of modeling of the spatial economy

The fundamental logic and the governing equation of a multi-regional version of the model by Forslid and Ottaviano (2003) are presented (Akamatsu et al., 2016). The budget constraint is given as

$$p_i^{\mathrm{A}} C_i^{\mathrm{A}} + \sum_{j \in N} \int_0^{n_j} p_{ji}(\ell) q_{ji}(\ell) d\ell = Y_i, \tag{B.1}$$

where p_i^{A} represents the price of the A-sector good in place i, C_i^{A} is the consumption of A-sector goods in place i, $N = \{0, 1, ..., K-1\}$, n_j is the number of varieties produced in region j, $p_{ji}(\ell)$ denotes the price of a variety ℓ in place i produced in place j, $q_{ji}(\ell)$ is the consumption of variety $\ell \in [0, n_j]$ in place i produced in place j, and Y_i is the income of an individual in place i. The incomes (wages) of skilled workers and unskilled workers are represented respectively by w_i and w_i^{L}.

An individual at place i maximizes the utility in (4) subject to the budget constraint in (B.1). This maximization yields the following demand functions

$$C_i^{\mathrm{A}} = (1-\mu)\frac{Y_i}{p_i^{\mathrm{A}}}, \qquad C_i^{\mathrm{M}} = \mu \frac{Y_i}{\rho_i}, \qquad q_{ji}(\ell) = \mu \frac{\rho_i^{\sigma-1} Y_i}{p_{ji}(\ell)^{\sigma}},$$

where ρ_i denotes the price index of the differentiated products in place i, and is given by

$$\rho_i = \left(\sum_{j \in N} \int_0^{n_j} p_{ji}(\ell)^{1-\sigma} d\ell \right)^{1/(1-\sigma)}. \tag{B.2}$$

Because the total income in place i is $w_i \lambda_i + w_i^{\mathrm{L}}$, the total demand $Q_{ji}(\ell)$ in place i for a variety ℓ produced in place j is given as

$$Q_{ji}(\ell) = \mu \frac{\rho_i^{\sigma-1}}{p_{ji}(\ell)^{\sigma}} (w_i \lambda_i + w_i^{\mathrm{L}}). \tag{B.3}$$

The A-sector is perfectly competitive and produces homogeneous goods under constant-returns-to-scale, and requires one unit of unskilled labor per unit of output. The A-sector good is traded freely across locations and is chosen as the numéraire. In equilibrium, $p_i^{\mathrm{A}} = w_i^{\mathrm{L}} = 1$ for each i.

The M-sector output is produced under increasing-returns-to-scale and Dixit–Stiglitz monopolistic competition. A firm incurs a fixed input requirement of α units of skilled labor and a marginal input requirement of β units of unskilled labor. An M-sector firm located in place i chooses $(p_{ij}(\ell) \mid j \in N)$ that maximizes its profit

$$\Pi_i(\ell) = \sum_{j \in N} p_{ij}(\ell) Q_{ij}(\ell) - (\alpha w_i + \beta x_i(\ell)), \tag{B.4}$$

where $x_i(\ell)$ denotes the total supply of variety ℓ produced in place i and $\alpha w_i + \beta x_i(\ell)$ signifies the cost function introduced by Flam and Helpman (1987).

With the use of the iceberg form of the transport cost, we have

$$x_i(\ell) = \sum_{j \in N} \tau_{ij} Q_{ij}(\ell). \tag{B.5}$$

Then the profit function of the M-sector firm in place i, given in (B.4), becomes

$$\Pi_i(\ell) = \sum_{j \in N} p_{ij}(\ell) Q_{ij}(\ell) - \left(\alpha w_i + \beta \sum_{j \in N} \tau_{ij} Q_{ij}(\ell) \right),$$

which is maximized by the firm. The first-order condition for this profit maximization yields the following optimal price

$$p_{ij}(\ell) = \frac{\sigma \beta}{\sigma - 1} \tau_{ij}. \tag{B.6}$$

This result implies that $p_{ij}(\ell)$, $Q_{ij}(\ell)$, and $x_i(\ell)$ are independent of ℓ. Therefore, the argument ℓ is suppressed subsequently.

In the short run, skilled workers are immobile between places, i.e., their spatial distribution $\boldsymbol{\lambda} = (\lambda_i \mid i \in N)$ is assumed to be given. The market equilibrium conditions consist of three conditions: the M-sector goods market clearing condition, the zero-profit condition attributable to the free entry and exit of firms, and the skilled labor market clearing condition. The first condition is written as (B.5) above. The second one requires that the operating profit of a firm, given in (B.4), be absorbed entirely by the wage bill of its skilled workers. This gives

$$w_i = \frac{1}{\alpha}\left\{\sum_{j\in N} p_{ij}Q_{ij} - \beta x_i\right\}. \tag{B.7}$$

The third condition is expressed as $\alpha n_i = \lambda_i$ and the price index ρ_i in (B.2) can be rewritten using (B.6) as

$$\rho_i = \frac{\sigma\beta}{\sigma-1}\left(\frac{1}{\alpha}\sum_{j\in N}\lambda_j d_{ji}\right)^{1/(1-\sigma)}. \tag{B.8}$$

The market equilibrium wage w_i in (B.7) can be represented as

$$w_i = \frac{\mu}{\sigma}\sum_{j\in N}\frac{d_{ij}}{\Delta_j}(w_j\lambda_j + 1) \tag{B.9}$$

using $d_{ji} = \tau_{ji}^{1-\sigma} = \phi^{m(i,j)}$, (B.3), (B.5), (B.6), and (B.8). Here, $\Delta_j = \sum_{k\in N} d_{kj}\lambda_k$. Equation (B.9) can be rewritten, using $\boldsymbol{w} = (w_i)$, as $\boldsymbol{w} = \frac{\mu}{\sigma} D\Delta^{-1}(\Lambda\boldsymbol{w} + \mathbf{1})$, which is solved for $\boldsymbol{w}$ as

$$\boldsymbol{w} = \frac{\mu}{\sigma}\left(I - \frac{\mu}{\sigma} D\Delta^{-1}\Lambda\right)^{-1} D\Delta^{-1}\mathbf{1} \tag{B.10}$$

with I being the identity matrix, $\mathbf{1} = (1,\dots,1)^\top$, $D = (d_{ij})$, $\Delta = \mathrm{diag}(\Delta_0,\dots,\Delta_{K-1})$, and $\Lambda = \mathrm{diag}(\lambda_0,\dots,\lambda_{K-1})$.

Appendix C. Theoretical details of Section 2

Appendix C.1. Proof of Proposition 1

For $\boldsymbol{\lambda} = \boldsymbol{\lambda}_\delta^{\mathrm{FA}}$, we have $\lambda_i = 0$ ($i \neq k-\delta$) and $v_{k-\delta} - \bar{v} = 0$ since $\bar{v} = v_{k-\delta}$; accordingly, the governing equation (3) with (2) is satisfied for any $i \in N$. For $\boldsymbol{\lambda}_\delta^{\mathrm{Twin}}$, we have $v_{k-\delta} = v_{k+\delta}$ by symmetry, $v_{k-\delta} - \bar{v} = v_{k-\delta} - \frac{1}{2}v_{k-\delta} - \frac{1}{2}v_{k+\delta} = 0$ and similarly $v_{k+\delta} - \bar{v} = 0$. We also have $\lambda_i = 0$ ($i \neq k \pm \delta$); accordingly, the governing equation is satisfied.

Appendix C.2. Critical points for the full agglomeration

Since a break bifurcation point is absent for the full agglomeration,[23] we focus hereafter on a corner bifurcation point, at which the matrix J_0 in (A.1) becomes singular (cf., Appendix A). The matrix J_0 at $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\mathrm{FA}}$ is a diagonal matrix with the diagonal entries: $\{v_{k\pm\delta} - \bar{v} \mid \delta \in N_\delta\}$, and each entry is repeated twice because $v_{k-\delta} - \bar{v} = v_{k+\delta} - \bar{v}$. Thus there possibly exists a series of critical points $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$ $(\delta \in N_\delta)$.

Appendix C.3. Proof of Proposition 2

In the analysis of bifurcating solutions at a critical point, the so-called bifurcation equation is employed (e.g., Golubitsky et al., 1988 and Ikeda and Murota, 2019). In the neighborhood of the present critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$, the governing equation $\boldsymbol{F}(\boldsymbol{\lambda}, \tau) = \boldsymbol{0}$ in (3) can be reduced to a two-dimensional bifurcation equation $\tilde{F}_i = 0$ $(i = k-\delta,\ k+\delta)$ in two independent variables $v_{k-\delta}$ and $v_{k+\delta}$ (see Lemma 1 in Appendix C.3).

We can derive a two-dimensional bifurcation equation in incremental variables $(x, y, \psi) = (\lambda_{k-\delta}, \lambda_{k+\delta}, \psi)$ at the critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$, using $\psi = \phi - \phi_\delta^{\mathrm{c}}$, as follows.

Lemma 1. *The bifurcation equation at the critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$ is expressed as*

$$\begin{aligned} \tilde{F}_{k-\delta}(x, y, \psi) &= x\,(a\psi + bx + cy + \text{higher order terms}) = 0, \\ \tilde{F}_{k+\delta}(x, y, \psi) &= y\,(a\psi + by + cx + \text{higher order terms}) = 0 \end{aligned} \tag{C.1}$$

[23] Since eigenvector associated with $J_+ = -v_k < 0$ for the full agglomeration is not in the $(K-1)$-dimensional simplex, there is no break bifurcation.

with the symmetry condition $\tilde{F}_{k+\delta}(x,y,\psi) = \tilde{F}_{k-\delta}(y,x,\psi)$ *and expansion coefficients:*

$$(a,b,c) = \left.\left(\frac{\partial g}{\partial \phi}, \frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)\right|_{(x,y,\psi)=(0,0,\phi_\delta^{\mathrm{c}})}, \qquad g(x,y,\psi) = v_{k-\delta}(\tilde{\boldsymbol{\lambda}}) - v_k(\tilde{\boldsymbol{\lambda}});$$

$$\tilde{\boldsymbol{\lambda}} = (\mathbf{0}_{k-\delta-1}, x, \mathbf{0}_\delta, 1-x-y, \mathbf{0}_\delta, y, \mathbf{0}_{k-\delta-1}, \phi_\delta^{\mathrm{c}} + \psi), \qquad \mathbf{0}_p = (\underbrace{0,\ldots,0}_{p\ \text{times}}).$$

Proof. In the neighborhood of the critical point $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$, $\boldsymbol{F}(\boldsymbol{\lambda},\tau) = \mathbf{0}$ in (3) reduces to three equations $F_j = 0$ with three variables v_j ($j = k, k\pm\delta$), while the other variables are equal to 0. Then $F_{k-\delta} + F_k + F_{k+\delta} = 0$ gives the conservation law: $\lambda_{k-\delta} + \lambda_k + \lambda_{k+\delta} = 0$. The variable λ_k can be eliminated from $F_{k-\delta}$ and $F_{k+\delta}$ to arrive at (C.1). The symmetry condition arises from the bilateral symmetry of the long narrow economy. $\square$

The bifurcation equation (C.1) with the symmetry condition has solutions $(x,y) = (\lambda_{k-\delta}, \lambda_{k+\delta}) = (0,0)$, $(w,0)$, $(0,w)$, and (w,w); $(x,y) = (0,0)$ corresponds to the pre-bifurcation solution $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi)$ and others to bifurcating solutions. Since the solutions $(w,0)$ and $(0,w)$ are identical, we hereafter consider only the former solution.

Lemma 2. *The critical point* $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$ *is a bifurcation point with two kinds of branches:*

$$(\boldsymbol{\lambda}, \phi) = (\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}}) + (\Delta\boldsymbol{\lambda}_p, \psi_p), \quad p = 1,2;$$

$$\Delta\boldsymbol{\lambda}_1 = w(\boldsymbol{e}_\delta^1, -2, \boldsymbol{e}_\delta^2), \quad \psi_1 \approx -(b+c)w/a; \quad \boldsymbol{e}_\delta^1 = (\mathbf{0}_{k-\delta-1}, 1, \mathbf{0}_\delta), \quad 0 < w \ll 1, \tag{C.2}$$

$$\Delta\boldsymbol{\lambda}_2 = w(\boldsymbol{e}_\delta^1, -1, \mathbf{0}_k), \quad \psi_2 \approx -bw/a; \qquad \boldsymbol{e}_\delta^2 = (\mathbf{0}_\delta, 1, \mathbf{0}_{k-\delta-1}). \tag{C.3}$$

Proof. We see that $(x,y) = (\lambda_{k-\delta}, \lambda_{k+\delta}) = (w,w)$ corresponds to $\Delta\boldsymbol{\lambda}_1 = w(\boldsymbol{e}_\delta^1, -2, \boldsymbol{e}_\delta^2)$ and satisfies (C.1) in Lemma 1 for $\psi = \psi_1 \approx -(b+c)w/a$. Also, $(x,y) = (w,0)$ corresponds to $\Delta\boldsymbol{\lambda}_2 = w(\boldsymbol{e}_\delta^1, -1, \mathbf{0}_k)$ and satisfies (C.1) for $\psi = \psi_2 \approx -bw/a$. $\square$

Appendix C.4. Proof of Proposition 3

The Jacobian matrix for the bifurcation equation (C.1) reads

$$\hat{J} \approx \begin{pmatrix} a\psi + 2bx + cy & cx \\ cy & a\psi + 2by + cx \end{pmatrix}.$$

The use of $(x, y) = w(1, 1)$ and $\psi = \psi_1 \approx -(b + c)w/a$ (cf., (C.2)) in $\hat{J}$ leads to $\hat{J}_1$ and the use of $(x, y) = w(1, 0)$ and $\psi = \psi_2 \approx -bw/a$ (cf, (C.3)) leads to $\hat{J}_2$ as follows:

$$\hat{J}_1 \approx w \begin{pmatrix} b & c \\ c & b \end{pmatrix}, \qquad \hat{J}_2 \approx w \begin{pmatrix} b & c \\ 0 & c - b \end{pmatrix}.$$

Lemma 3. *The bifurcating solution $(\Delta\boldsymbol{\lambda}_1, \psi_1)$ has the eigenvalues: $e_1 \approx (b+c)w$ and $e_2 \approx (b - c)w$. On the other hand, $(\Delta\boldsymbol{\lambda}_2, \psi_2)$ has the eigenvalues: $e_1 \approx bw$ and $e_2 \approx (c - b)w$.*

Lemma 4. *Under Assumption 1, there are three cases: (i) If $-b > |c|$, only the first bifurcating path $(\Delta\boldsymbol{\lambda}_1, \psi_1)$ is stable and sustainable. (ii) If $c < b < 0$, only the second bifurcating path $(\Delta\boldsymbol{\lambda}_2, \psi_2)$ is stable and sustainable. (iii) Otherwise, both paths are unstable and/or unsustainable. A stable and sustainable bifurcating path branches in the direction of $\psi < 0$.*

Proof. For the fully agglomerated state $(x, y) = (0, 0)$, we have $\hat{J} = a\psi I$ with the eigenvalue $a\psi$ (twice repeated). Since, by Assumption 1, this state is sustainable for $\psi > 0$, we have $a < 0$. (i) The first bifurcating solution $(\Delta\boldsymbol{\lambda}_1, \psi_1)$ with $e_1 \approx (b + c)w$ and $e_2 \approx (b - c)w$ (cf., Lemma 3) is stable if $-b > |c|$. Since $b + c < 0$, $a < 0$, and $w > 0$, $\psi = \psi_1 \approx -(b + c)w/a$ in (C.2) gives $\psi = \psi_1 < 0$. (ii) The second bifurcating solution $(\Delta\boldsymbol{\lambda}_2, \psi_2)$ with $e_1 \approx bw$ and $e_2 \approx (c-b)w$ $(w > 0)$ is stable if $c < b < 0$. Since $b < 0$, $a < 0$ and $w > 0$, $\psi = \psi_2 \approx -bw/a$ in (C.3) gives $\psi = \psi_2 < 0$. The two bifurcating solutions cannot be stable simultaneously since $-b > |c|$ and $c < b < 0$ are contradictory. □

Let a corner bifurcation point ϕ_δ^{c} not be the sustain point. Then there exists δ' ($\delta' \neq \delta$) such that $v_{k-\delta'} - v_k > 0$ at this point. By continuity of $v_{k-\delta'}$ and v_k as functions in ϕ, $v_{k-\delta'} - v_k > 0$ is satisfied in a neighborhood of $(\boldsymbol{\lambda}^{\mathrm{FA}}, \phi_\delta^{\mathrm{c}})$. Therefore, the bifurcation solution is unsustainable just after bifurcation.

Appendix D. Theoretical details of Section 3

Appendix D.1. Proof of (7) *and* (8)

For the full agglomeration $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{\mathrm{FA}}$, we rearrange the components of the variable $\boldsymbol{\lambda}$ using the permutation of place numbers:

$$\begin{pmatrix} 0 & \cdots & k-1 & k & k+1 & \cdots & 2k \\ k & \cdots & 1 & 0 & k+1 & \cdots & 2k \end{pmatrix}.$$

Then the variables used to define $\boldsymbol{w}$ in (B.10) are expressed using $\boldsymbol{d} = (\phi, \phi^2, \ldots, \phi^k)$ as

$$\Lambda = \mathrm{diag}(1, \mathbf{0}_{2k}), \qquad \Delta_i = \sum_{j=0}^{2k} d_{ji}\lambda_j = d_{0i}, \qquad (\Delta_0, \Delta_1, \ldots, \Delta_{2k}) = (1, \boldsymbol{d}, \boldsymbol{d}), \tag{D.1}$$

$$\Delta = \mathrm{diag}(\Delta_0, \ldots, \Delta_{2k}) = \mathrm{diag}\ (1, \boldsymbol{d}, \boldsymbol{d}), \qquad \Delta^{-1} = \mathrm{diag}(1, \Theta, \Theta), \qquad \Theta = \mathrm{diag}(\boldsymbol{d})^{-1},$$

$$D = \left(\begin{array}{c|c|c} 1 & \boldsymbol{d} & \boldsymbol{d} \\ \hline \boldsymbol{d}^\top & D_1 & D_2 \\ \hline \boldsymbol{d}^\top & D_2 & D_1 \end{array}\right), \quad D\Delta^{-1} = \left(\begin{array}{c|c|c} 1 & \boldsymbol{d}\Theta & \boldsymbol{d}\Theta \\ \hline \boldsymbol{d}^\top & D_1\Theta & D_2\Theta \\ \hline \boldsymbol{d}^\top & D_2\Theta & D_1\Theta \end{array}\right), \quad D\Delta^{-1}\Lambda = \left(\begin{array}{c|c|c} 1 & O & O \\ \hline \boldsymbol{d}^\top & O & O \\ \hline \boldsymbol{d}^\top & O & O \end{array}\right),$$

$$D\Delta^{-1}\mathbf{1} = (2k+1, \boldsymbol{g}, \boldsymbol{g})^\top\ ;\ D_1 = \{\phi^{|j-i|} \mid 1 \le i, j \le k\},\ D_2 = \{\phi^{i+j} \mid 1 \le i, j \le k\}, \tag{D.2}$$

$$\boldsymbol{d}\Theta = \mathbf{1}^\top; \qquad D_1\Theta = \{\phi^{|j-i|-j} \mid 1 \le i, j \le k\}, \qquad D_2\Theta = \{\phi^i \mid 1 \le i, j \le k\};$$

$$\boldsymbol{g} = \{g_i \mid 1 \le i \le k\}, \qquad g_i = \phi^i + \sum_{j=1}^{k}(\phi^i + \phi^{|i-j|-j}) = (k+1)\phi^i + (k-i)\phi^{-i} + \sum_{p=1}^{i}\phi^{i-2p}.$$

Hence we have (I_k being $k \times k$ identity matrix)

$$(I - \theta D\Delta^{-1}\Lambda)^{-1} = \left(\begin{array}{c|c|c} 1-\theta & & \\ \hline -\theta \boldsymbol{d}^\top & I_k & \\ \hline -\theta \boldsymbol{d}^\top & & I_k \end{array}\right)^{-1} = \frac{1}{1-\theta}\left(\begin{array}{c|c|c} 1 & & \\ \hline \theta \boldsymbol{d}^\top & (1-\theta)I_k & \\ \hline \theta \boldsymbol{d}^\top & & (1-\theta)I_k \end{array}\right),$$

$$(I - \theta D\Delta^{-1}\Lambda)^{-1} D\Delta^{-1}\mathbf{1} = \frac{\theta}{1-\theta}(2k+1, \boldsymbol{z}, \boldsymbol{z})^\top, \qquad \boldsymbol{z} = \theta(2k+1)\boldsymbol{d} + (1-\theta)\boldsymbol{g}.$$

The use of (D.2) and this equation in (B.10) leads to the expressions of the wage as

$$w_0 = \frac{\theta(2k+1)}{1-\theta}, \quad w_i = w_{i+k} = \frac{\theta}{1-\theta}\left\{(\theta k + k + 1)\phi^i + (1-\theta)\left[(k-i)\phi^{-i} + \sum_{p=1}^{i} \phi^{i-2p}\right]\right\}$$

($1 \leq i \leq k$). In the original place numbers $i \mapsto k - i = \delta$, these equations are rewritten as

$$w_k = \frac{\theta(2k+1)}{1-\theta}, \quad w_{k\pm\delta} = \frac{\theta}{1-\theta}\left\{(\theta k + k + 1)\phi^\delta + (1-\theta)\left[(k-\delta)\phi^{-\delta} + S_\delta\right]\right\} \quad (1 \leq \delta \leq k)$$

with $S_\delta = \sum_{p=1}^{\delta} \phi^{\delta-2p}$. The use of (D.1) and these expressions in (6) proves (7) and (8).

Appendix D.2. Proof of Proposition 4(i)

By Lemma 5, $v_{k-\delta} - v_k > 0$ as $\phi \to +0$ and $v_{k-\delta} - v_k < 0$ as $\phi \to 1$ for each $\delta \in N_\delta$. By the Intermediate Value Theorem, there is a critical point satisfying $v_{k-\delta} - v_k = 0$ for each $\delta \in N_\delta$. By Proposition 2, this critical point is a bifurcation point, at which two satellite cities emerges at the $(k \pm \delta)$th cities or a satellite city emerge at the $(k - \delta)$th city.

Appendix D.3. Limit behaviors when trade freeness is very low or very high

We consider the limit behaviors when the trade freeness ϕ is either very low or very high in the following Lemma. In an extreme case of $\phi \to 1$ with no transport costs, the central city has a locational advantage due to a higher market access and a wider array of varieties for consumers. Firms in the central region can avoid costly transportation while consumers consume more varieties and enjoy a lower cost of living (lower regional price index). Thus, the central city has a better trade environment and workers living there are endowed with a larger indirect utility. This is in line with the limit behavior of the Krugman model for a long narrow economy with three places (Ago et al., 2006).

In another extreme case of $\phi \to +0$, i.e. with prohibitively high transport costs, under the no-black-hole condition $\mu < \sigma - 1$, a city at an outer location has a larger indirect utility when transport costs are extremely high as explained in the Lemma below. This is because price competition in the central region is fiercer which induces firms to locate at cities near the border where competition is softer.

Lemma 5. *(i) As $\phi \to 1$, we have $v_k > v_{k-1} > \cdots > v_0$. (ii) As $\phi \to +0$, we have $v_k < v_{k-1} < \cdots < v_0$ under the no-black-hole condition $\mu < \sigma - 1$.*

Proof. To prove $v_i > v_{i-1}$ $(1 \le i \le k)$ for $\phi \to 1$, we put $\phi = 1 - \epsilon$ $(0 < \epsilon \ll 1)$ and consider a limit of $\epsilon \to +0$. Then $v_i = v_{k-\delta}$ in (8) $(0 \le i \le k-1)$ can be expanded as

$$\begin{aligned} v_i &= \ln \frac{\theta}{1-\theta} + \frac{\delta(i)\mu}{\sigma - 1} \ln\phi + \ln \hat{v}_i \\ &= \ln \frac{\theta}{1-\theta} + \left(-\frac{\delta(i)\mu}{\sigma - 1}\epsilon + \text{h.o.t.}\right) + \left((\ln \hat{v}_i)|_{\epsilon=0} + \left.\frac{\partial(\ln \hat{v}_i)}{\partial \epsilon}\right|_{\epsilon=0} \epsilon + \text{h.o.t.}\right) \\ &= \ln \frac{\theta}{1-\theta} - \frac{\delta(i)\mu}{\sigma - 1}\epsilon + \ln(2k+1) + \left.\frac{\partial(\ln \hat{v}_i)}{\partial \epsilon}\right|_{\epsilon=0} \epsilon + \text{h.o.t.} \end{aligned}$$

with $\delta = \delta(i) = k - i$ $(1 \le \delta \le k)$ and

$$\hat{v}_i = (\theta k + k + 1)(1-\epsilon)^{\delta} + (1-\theta)\left[(k-\delta)(1-\epsilon)^{-\delta} + \sum_{p=1}^{\delta}(1-\epsilon)^{\delta-2p}\right], \tag{D.3}$$

$$\frac{\partial(\ln \hat{v}_i)}{\partial \epsilon} = -\frac{1}{\hat{v}_i}\{\delta(\theta k + k + 1)(1-\epsilon)^{\delta-1} + (1-\theta)[\delta(\delta-k)(1-\epsilon)^{-(\delta+1)} + \hat{S}\,]\},$$

$$\left.\frac{\partial(\ln \hat{v}_i)}{\partial \epsilon}\right|_{\epsilon=0} = -\left(\theta\delta + \frac{(1-\theta)\delta^2}{2k+1}\right); \quad \hat{S} = \sum_{p=1}^{\delta}(\delta - 2p)(1-\epsilon)^{\delta-2p-1}.$$

We can express v_i $(i \neq k)$ asymptotically as

$$v_i \approx \ln \frac{\theta}{1-\theta} + \ln\,(2k+1) - \left[\left(\theta + \frac{\mu}{\sigma-1}\right)\delta + \frac{1-\theta}{2k+1}\delta^2\right]\epsilon$$
$$= v_k - \left[\left(\theta + \frac{\mu}{\sigma-1}\right)(k-i) + \frac{1-\theta}{2k+1}(k-i)^2\right]\epsilon. \tag{D.4}$$

We have $v_k > v_i$ $(i \neq k)$ because $k-i>0$ and $0<\theta<1$. Furthermore,

$$v_i - v_{i-1} \approx \left[\theta + \frac{\mu}{\sigma-1} + \frac{(1-\theta)(2(k-i)+1)}{2k+1}\right]\epsilon > 0 \qquad (1 \leq i \leq k-1). \tag{D.5}$$

Hence we have $v_k > v_{k-1} > \cdots > v_0$ $(\phi \to 1)$.

Using $\sum_{p=1}^{\delta} \phi^{\delta-2p} = \frac{\phi^{\delta}-\phi^{-\delta}}{\phi^2-1}$, we rewrite $\hat{v}_i$ in (D.3) with $1-\epsilon=\phi$ and evaluate $\ln \hat{v}_i$ as

$$\hat{v}_i \;=\; A\phi^{\delta} + B\left(i\phi^{-\delta} + \frac{\phi^{\delta}-\phi^{-\delta}}{\phi^2-1}\right); \quad A \equiv \theta k + k + 1 > 0, \quad B \equiv 1-\theta > 0. \tag{D.6}$$
$$\ln \hat{v}_i \;=\; -\delta \ln\phi + \ln\left[\left(A + \frac{B}{\phi^2-1}\right)\phi^{2\delta} + B\left(i + \frac{1}{1-\phi^2}\right)\right].$$

Using this equation, we can rewrite $v_i = v_{k-\delta}$ in (8) and evaluate its limit behavior as

$$v_i = \ln \frac{\theta}{1-\theta} + \delta\,\left(\frac{\mu}{\sigma-1} - 1\right)\ln\phi + \ln\left[\left(A + \frac{B}{\phi^2-1}\right)\phi^{2\delta} + B\left(i + \frac{1}{1-\phi^2}\right)\right], \tag{D.7}$$
$$\lim_{\phi\to+0} v_i = \ln\frac{\theta}{1-\theta} + \delta\,\left(\frac{\mu}{\sigma-1}-1\right)\left(\lim_{\phi\to+0}\ln\phi\right) + \ln\,(1-\theta)(i+1).$$

We consider the case $\mu < \sigma - 1$ and prove $v_k < v_{k-1} < \cdots < v_0$ $(\phi \to +0)$ by showing $v_k < v_i$ $(i \neq k)$ and $v_i < v_{i-1}$ $(1 \leq i \leq k-1)$. First, since v_k is constant and $\lim_{\phi\to+0} v_i = +\infty$ $(i \neq k)$, we have $v_k < v_i$ $(\phi \to +0, i \neq k)$. Next, using (D.7) with $\delta = k-i$ and $B = 1-\theta$, we have

$$v_{i-1} - v_i = (\rho-1)\ln\phi + V(\phi); \quad V(\phi) = \ln\left(\frac{\left(A + \frac{1-\theta}{\phi^2-1}\right)\phi^{2(k-i+1)} + (1-\theta)\left(i-1-\frac{1}{\phi^2-1}\right)}{\left(A + \frac{1-\theta}{\phi^2-1}\right)\phi^{2(k-i)} + (1-\theta)\left(i - \frac{1}{\phi^2-1}\right)}\right)$$

$(\rho = \frac{\mu}{\sigma-1} < 1)$. Since $\lim_{\phi\to+0} V(\phi) = \ln\,\left(\frac{i}{i+1}\right)$ and $\lim_{\phi\to+0}\,[(\rho-1)\ln\phi] = +\infty$, we have $\lim_{\phi\to+0}\,(v_{i-1} - v_i) = +\infty$. This shows $v_i < v_{i-1}$ $(\phi \to +0)$. □

Appendix D.4. Proof of Proposition 5

By Proposition 4, $\phi^{\mathrm{c}}_{\delta}$ exists for each δ and, accordingly, ϕ^{s} $(= \max_{\delta \in N^{\delta}} \phi^{\mathrm{c}}_{\delta})$ can be defined. Hence $v_{k-\delta} - v_k$ does not change its sign in $\phi \in (\phi^{\mathrm{s}}, 1)$. By Lemma 5, $v_{k-\delta} - v_k < 0$ as $\phi \to 1$ for any δ. Accordingly, $v_{k-\delta} - v_k < 0$ in $\phi \in (\phi^{\mathrm{s}}, 1)$ for any δ and the full agglomeration is sustainable for $\phi > \phi^{\mathrm{s}}$.

Appendix D.5. Proof of Proposition 6

We investigate the change of the value of a sustain point by the implicit function theorem. We first define $g(\phi, \sigma, \mu) \equiv v_i - v_k$ $(i = k - \delta)$. Substituting (7) and (8), we have

$$g(\phi, \sigma, \mu) = \frac{\delta\mu}{\sigma - 1}\ln\phi + \ln X - \ln(2k+1)$$

with $X = \left(\frac{\mu}{\sigma}k + k + 1\right)\phi^{\delta} + \left(1 - \frac{\mu}{\sigma}\right)\left[(k-\delta)\phi^{-\delta} + \sum_{p=1}^{\delta}\phi^{\delta-2p}\right] > 0$.

We assume that 1) $g(\phi^{\mathrm{s}}, \sigma^{\mathrm{s}}, \mu^{\mathrm{s}}) = 0$ holds and 2) $v_i = \max(v_0, \ldots, v_{k-1})$ is satisfied in the neighborhood of $(\phi^{\mathrm{s}}, \sigma^{\mathrm{s}}, \mu^{\mathrm{s}})$. ϕ^{s} is a sustain point by these assumption. Furthermore, by the implicit function theorem, there exist $\phi = \phi(\sigma, \mu)$ in the neighborhood of $(\phi^{\mathrm{s}}, \sigma^{\mathrm{s}}, \mu^{\mathrm{s}})$ and $\phi(\sigma, \mu)$ satisfies the followings:

$$g(\phi(\sigma, \mu), \sigma, \mu) = 0, \tag{D.8}$$

$$\frac{\partial\phi}{\partial\sigma} = -\frac{\partial g/\partial\sigma}{\partial g/\partial\phi}, \qquad \frac{\partial\phi}{\partial\mu} = -\frac{\partial g/\partial\mu}{\partial g/\partial\phi}. \tag{D.9}$$

The value of $\phi(\sigma, \mu)$ is that of a sustain point by assumption 2) and condition (D.8). Using (D.9), we investigate the change of the value.

Concrete forms of partial derivatives of $g(\phi, \sigma, \mu)$ to be used later are

$$\frac{\partial g}{\partial\sigma} = -\frac{\delta\mu}{(\sigma-1)^2}\ln\phi + \frac{1}{X}\frac{\partial X}{\partial\sigma}, \qquad \frac{\partial g}{\partial\mu} = \frac{\delta}{\sigma - 1}\ln\phi + \frac{1}{X}\frac{\partial X}{\partial\mu}; \tag{D.10}$$

$$\frac{\partial X}{\partial\sigma} = -\frac{1}{\sigma^2}E, \quad \frac{\partial X}{\partial\mu} = \frac{1}{\sigma}E; \qquad E = k\phi^{\delta} - (k-\delta)\phi^{-\delta} - \sum_{p=1}^{\delta}\phi^{\delta-2p}. \tag{D.11}$$

First, we investigate the sign of $\partial\phi/\partial\sigma$ (D.9). We generically have $\frac{\partial g}{\partial\phi}(\phi^{\mathrm{s}},\sigma,\mu) < 0$ because, by the definition of the sustain point, we have $g(\phi^{\mathrm{s}},\sigma,\mu) = 0$ and $g(\phi^{\mathrm{s}} + \mathrm{d}\phi^{\mathrm{s}},\sigma,\mu) < 0$ $(0 < \mathrm{d}\phi^{\mathrm{s}} \ll 1)$. We also have $\frac{\partial g}{\partial\sigma}(\phi^{\mathrm{s}},\sigma,\mu) > 0$ because, in (D.10), we have $\ln\phi < 0$, $X > 0$, and $\partial X/\partial\sigma|_{\phi=\phi^{\mathrm{s}}} > 0$ from (D.11) with $E|_{\phi=\phi^{\mathrm{s}}} < 0$:

$$E|_{\phi=\phi^{\mathrm{s}}} = k(\phi^{\mathrm{s}})^{\delta} - (k-\delta)(\phi^{\mathrm{s}})^{-\delta} - \sum_{p=1}^{\delta}(\phi^{\mathrm{s}})^{\delta-2p} < k(\phi^{\mathrm{s}})^{\delta} - (k-\delta)(\phi^{\mathrm{s}})^{\delta} - \sum_{p=1}^{\delta}(\phi^{\mathrm{s}})^{\delta} = 0.$$

Then the sign of $\partial\phi/\partial\sigma$ in (D.9) is positive.

On the other hand, the sign of (D.9) is negative as we already know $\frac{\partial g}{\partial\phi}(\phi^{\mathrm{s}},\sigma,\mu) < 0$ and have $\frac{\partial g}{\partial\mu}(\phi^{\mathrm{s}},\sigma,\mu) < 0$ in (D.10) ($\frac{\partial X}{\partial\mu}|_{\phi=\phi^{\mathrm{s}}} = \frac{1}{\sigma}\,E|_{\phi=\phi^{\mathrm{s}}} < 0$).

Note that the signs of $\partial\phi/\partial\sigma$ and $\partial\phi/\partial\mu$ do not depend on the value of δ. Therefore, the results do not change even if condition 2) changes to $v_j = \max(v_0,\ldots,v_{k-1})$ $(j \neq i)$.

Appendix E. Uniqueness of bifurcation points (Proof of Proposition 4(ii))

In preparation for the discussion regarding the uniqueness of bifurcation points in subsequent proofs, we have the following essential limits:

$$\lim_{\phi\to 1}(v_i - v_k) = 0, \qquad \lim_{\phi\to 0}(v_i - v_k) = +\infty, \tag{E.1}$$

$$\lim_{\phi\to 1}\frac{\partial(v_i - v_k)}{\partial\phi} = \delta\left[\frac{\delta(\sigma-\mu)}{2k\sigma+\sigma} + \mu\left(\frac{1}{\sigma} + \frac{1}{\sigma-1}\right)\right] > 0, \tag{E.2}$$

where (E.1) is apparent from (D.4) and Lemma 5 and (E.2) is given by a straightforward calculation. In the following discussion, the sign of $\frac{\partial^2(v_i-v_k)}{\partial\phi^2}$ plays an important role. We express $\frac{\partial^2(v_i-v_k)}{\partial\phi^2}$ such that its denominator is positive; accordingly, its sign is given by the sign of its numerator, being defined as $P_\delta(\phi)$ for a given δ. Then, we have:

Lemma 6. *If $\frac{\partial^2(v_i-v_k)}{\partial\phi^2}$ has at most one root for $\phi > 0$, there is a unique bifurcation point satisfying $v_i - v_k = 0$ for $\phi \in (0,1)$.*

Proof. We have: $P_\delta(0) = -\delta(k+1-\delta)^2(\mu-\sigma)^2(\mu-\sigma+1) > 0$. This means that $v_i - v_k$ is convex for $\phi = 0$. If $P_\delta(\phi)$ has at most one root for $\phi > 0$, then $v_i - v_k$ may become concave for some $\phi > 0$. This implies that $v_i - v_k$ may have either one zero or three zeros for $\phi \in (0, 1)$. However, the limits in (E.1) and (E.2) rule out the latter case and establish that there exists exactly one root of $v_i - v_k = 0$ for $\phi \in (0, 1)$. Therefore, there exists a unique bifurcation point satisfying $v_i - v_k = 0$ for $\phi \in (0, 1)$. □

We would like to show the following lemma for $\delta \in \{1, 2, \ldots, 6\}$. Then by Lemma 6 and Descartes' rule of sings, Proposition 4(ii) can be proven in a straightforward manner.

Lemma 7. *$P_\delta(\phi)$ takes a polynomial form of*

$$P_\delta(\phi) = a_1\phi^{4\delta} + a_2\phi^{4\delta-2} + \cdots + a_{2\delta}\phi^2 + a_{2\delta+1} \qquad (\delta = 1, 2, \ldots, 6)$$

and the sign of a series of coefficients a_1, a_2, $\ldots$, $a_{2\delta+1}$ changes once for $\mu < \sigma - 1$.

Proof. The sign of the series of coefficients changes once for each $\delta = 1, 2, \ldots, 6$ as expressed by the explicit forms of these coefficients listed below:

For $\delta = 1$ and for any $k \geq 1$, we have $P_1(\phi) = a_1\phi^4 + a_2\phi^2 + a_3$ with

$$a_1 = -(\mu+\sigma-1)\left[k(\mu+\sigma)+\sigma\right]^2 < 0,$$

$$a_2 = 2k(\mu-2\sigma+2)(\mu-\sigma)\left[k(\mu+\sigma)+\sigma\right] > 0, \quad a_3 = -k^2(\mu-\sigma)^2(\mu-\sigma+1) > 0.$$

For $\delta = 2$ and for any $k \geq 2$, we have $P_2(\phi) = a_1\phi^8 + \cdots + a_4\phi^2 + a_5$ with

$$a_1 = -(\mu+\sigma-1)\left[k(\mu+\sigma)+\sigma\right]^2 < 0, \quad a_2 = (\mu-\sigma)(2\mu-\sigma+1)\left[k(\mu+\sigma)+\sigma\right],$$

$$a_3 = (\mu-\sigma)\Big[2k\{k(\mu+\sigma)-\mu\}(\mu-4\sigma+4) - \mu^2 - \mu\sigma + 8(\sigma-1)\sigma\Big] > 0,$$

$$a_4 = -(k-1)(2\mu-3\sigma+3)(\mu-\sigma)^2 > 0, \;\; a_5 = -(k-1)^2(\mu-\sigma)^2(\mu-\sigma+1) > 0.$$

Note that the sign of the series of coefficients changes once, irrespective of the sign of a_2. Such is also the cases for $\delta = 5, 6$ below.

For $\delta = 3$ and for any $k \geq 3$, we have $P_3(\phi) = a_1\phi^{12} + \cdots + a_6\phi^2 + a_7$ with

$$a_1 = -3(\mu + \sigma - 1)(k(\mu + \sigma) + \sigma)^2 < 0, \quad a_2 = 6\mu(\mu - \sigma)(k(\mu + \sigma) + \sigma) < 0,$$

$$a_3 = (\mu - \sigma)\Big((6k - 3)\mu^2 + k\mu(14 - 8\sigma) - (14k + 13)(\sigma - 1)\sigma + 8\mu\sigma + \mu\Big) > 0,$$

$$a_4 = 2(\mu - \sigma)\Big[3((k - 2)k - 1)\mu^2 - \mu(3k(k(5\sigma - 6) - 11\sigma + 12) + \sigma + 2)$$
$$- 2(9(k - 1)k - 17)(\sigma - 1)\sigma\Big] > 0,$$

$$a_5 = -(\mu - \sigma)^2(6k(\mu - 3\sigma + 3) - 9\mu + 35(\sigma - 1)) > 0$$

$$a_6 = -2(k - 2)\phi^2(3\mu - 4\sigma + 4)(\mu - \sigma)^2 > 0, \quad a_7 = -3(k - 2)^2(\mu - \sigma)^2(\mu - \sigma + 1) > 0.$$

For $\delta = 4$ and for any $k \geq 4$, we have $P_4(\phi) = a_1\phi^{16} + \cdots + a_8\phi^2 + a_9$ with

$$a_1 = -2(\mu + \sigma - 1)\left[k(\mu + \sigma) + \sigma\right]^2 < 0, \quad a_2 = (\mu - \sigma)(4\mu + \sigma - 1)\left[k(\mu + \sigma) + \sigma\right] < 0,$$

$$a_3 = (\mu - \sigma)\Big[(4k - 2)\mu^2 - 2k\mu(\sigma - 3) - (6k + 5)(\sigma - 1)\sigma + 5\mu\sigma + \mu\Big] > 0,$$

$$a_4 = (\mu - \sigma)\Big\{4(k - 1)\mu^2 + \mu\left[k(17 - 13\sigma) + 9\sigma - 1\right] - (17k + 18)(\sigma - 1)\sigma\Big\} > 0,$$

$$a_5 = 2(\mu - \sigma)\Big(\left[2(k - 3)k - 3\right]\mu^2 + \mu\left\{2k\left[k(8 - 7\sigma) + 22\sigma - 24\right] + \sigma - 4\right\}$$
$$-4\left[4(k - 2)k - 11\right](\sigma - 1)\sigma\Big) > 0,$$

$$a_6 = -(\mu - \sigma)^2\left\{k\left[4\mu - 19\sigma + 19\right] - 8\mu + 54(\sigma - 1)\right\} > 0,$$

$$a_7 = -(\mu - \sigma)^2\left[2k(2\mu - 5\sigma + 5) - 10\mu + 29(\sigma - 1)\right] > 0,$$

$$a_8 = -(k - 3)(4\mu - 5\sigma + 5)(\mu - \sigma)^2 > 0, \quad a_9 = -2(k - 3)^2(\mu - \sigma)^2(\mu - \sigma + 1) > 0.$$

For $\delta = 5$ and for any $k \geq 5$, we have $P_5(\phi) = a_1\phi^{20} + \cdots + a_{10}\phi^2 + a_{11}$ with

$$a_1 = -5(\mu + \sigma - 1)\left[k(\mu + \sigma) + \sigma\right]^2 < 0, \quad a_2 = 2(\mu - \sigma)(5\mu + 2\sigma - 2)\left[k(\mu + \sigma) + \sigma\right] < 0,$$

$$a_3 = (\mu - \sigma)\Big[5(2k - 1)\mu^2 + \mu(10k + 12\sigma + 3) - (10k + 7)(\sigma - 1)\sigma\Big],$$

$$a_4 = 2(\mu - \sigma)\left[k(\mu + \sigma)(5\mu - 16\sigma + 16) - 5\mu^2 + 10\mu\sigma - 16(\sigma - 1)\sigma\right] > 0,$$

$$a_5 = (\mu - \sigma)\left\{5(2k - 3)\mu^2 + \mu\,[k(62 - 52\sigma) + 38\sigma - 13] - (62k + 75)(\sigma - 1)\sigma\right\} > 0,$$

$$a_6 = 10(\mu - \sigma)\Big([(k - 4)k - 2]\,\mu^2 + \mu\,\{k\,[k(10 - 9\sigma) + 37\sigma - 40] + 2(\sigma - 2)\}$$

$$-2(5(k - 3)k - 18)(\sigma - 1)\sigma\Big) > 0,$$

$$a_7 = (\mu - \sigma)^2\,[-2k(5\mu - 33\sigma + 33) + 5(5\mu - 49\sigma + 49)] > 0,$$

$$a_8 = -\,2(\mu - \sigma)^2\,[5k(\mu - 4\sigma + 4) - 15\mu + 76(\sigma - 1)] > 0,$$

$$a_9 = -\,(\mu - \sigma)^2\,[2k(5\mu - 11\sigma + 11) - 35\mu + 85(\sigma - 1)] > 0,$$

$$a_{10} = -\,2(k - 4)\phi^2(5\mu - 6\sigma + 6)(\mu - \sigma)^2 > 0, \quad a_{11} = -5(k - 4)^2(\mu - \sigma)^2(\mu - \sigma + 1) > 0.$$

For $\delta = 6$ and for any $k \geq 6$, we have $P_6(\phi) = a_1\phi^{24} + \cdots + a_{12}\phi^2 + a_{13}$ with

$$a_1 = -\,3(\mu + \sigma - 1)\,[k(\mu + \sigma) + \sigma]^2 < 0, \quad a_2 = 3(\mu - \sigma)(2\mu + \sigma - 1)\,[k(\mu + \sigma) + \sigma] < 0,$$

$$a_3 = (\mu - \sigma)\left\{(6k - 3)\mu^2 + \mu\,[2k(\sigma + 2) + 7\sigma + 2] - 2(2k + 1)(\sigma - 1)\sigma\right\},$$

$$a_4 = (\mu - \sigma)\left[6(k - 1)\mu^2 + k\mu(15 - 9\sigma) - (15k + 14)(\sigma - 1)\sigma + 11\mu\sigma + \mu\right] > 0,$$

$$a_5 = (\mu - \sigma)\left\{(6k - 9)\mu^2 + \mu\,[6k(5 - 4\sigma) + 20\sigma - 5] - 5(6k + 7)(\sigma - 1)\sigma\right\} > 0,$$

$$a_6 = (\mu - \sigma)\left\{6(k - 2)\mu^2 + \mu\,[k(49 - 43\sigma) + 36\sigma - 18] - (49k + 67)(\sigma - 1)\sigma\right\} > 0,$$

$$a_7 = (\mu - \sigma)\Big(3\,[2(k - 5)k - 5]\,\mu^2 + \mu\,\{6k\,[k(12 - 11\sigma) + 56\sigma - 60] + 5(5\sigma - 8)\}$$

$$-8(9(k - 4)k - 40)(\sigma - 1)\sigma\Big) > 0,$$

$$a_8 = -\,(\mu - \sigma)^2\,\{k\,[6\mu - 51\sigma + 51] - 18\mu + 233(\sigma - 1)\} > 0,$$

$$a_9 = -\,(\mu - \sigma)^2\,[k(6\mu - 34\sigma + 34) - 3(7\mu - 53\sigma + 53)] > 0,$$

$$a_{10} = -\,(\mu - \sigma)^2\,[3k(2\mu - 7\sigma + 7) - 4(6\mu - 25\sigma + 25)] > 0,$$

$$a_{11} = -\,(\mu - \sigma)^2\,[6k(\mu - 2\sigma + 2) - 27\mu + 58(\sigma - 1)] > 0,$$

$$a_{12} = -\,(k - 5)(6\mu - 7\sigma + 7)(\mu - \sigma)^2, \quad a_{13} = -3(k - 5)^2(\mu - \sigma)^2(\mu - \sigma + 1) > 0.$$

□

Appendix F. Equilibrium paths for $K = 9, 11$ cities and details of the analysis

Figures F.1 and F.2 show equilibrium paths for $K = 9$, 11cities.

Appendix F.1. Numerical analysis strategy

In this analysis of paths of equilibria, we employ the following innovative strategy that exploits the existence of invariant patterns and the bifurcation mechanism of the full agglomeration and twin cities (Section 2):

1. Stability analysis: Obtain the ranges of the trade freeness ϕ for stable and sustainable full agglomerations and twin cities, which are invariant patterns (Section 2.2).
2. Comparative static analysis: Obtain the equilibrium path connected to the almost uniform state at $\phi = 0$ and bifurcating equilibria from those invariant patterns to find a network of equilibrium paths.
3. Stability analysis: Find stable equilibrium paths on this network.

Appendix F.2. Bifurcation from the twin cities

The path for the twin cities resides on the horizontal line at $\lambda_k = 0$ in each figure; this horizontal line contains paths other than those of the twin cities (e.g., Point C for $K = 5$ in Fig 9). For $K = 5$ cities, for instance, there is a stable and sustainable Path DE for twin cities $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1^{\text{Twin}} = (0, 1, 0, 1, 0)$, enclosed by the two sustain points D and E. A sustain point E has the stable bifurcating equilibrium Path EFG, on which the central city regains population leading to an agglomeration to three cities, en route to the full agglomeration at the center as ϕ increases. From the sustain point D, there branches a stable bifurcating equilibrium Path DCBA,[24] on which the population of the two border

[24]The kink at the Point C is due to the vanishing of the population at the central city ($i = k = 2$).

cities ($i = 0$ and 4, i.e., $\delta = 2$) becomes non-zero,[25] en route to a nearly uniform state for small ϕ. This demonstrates a vital role in the progress of agglomeration played by sustain bifurcations on the twin cities, which connect the state of the twin cities to other agglomeration patterns.

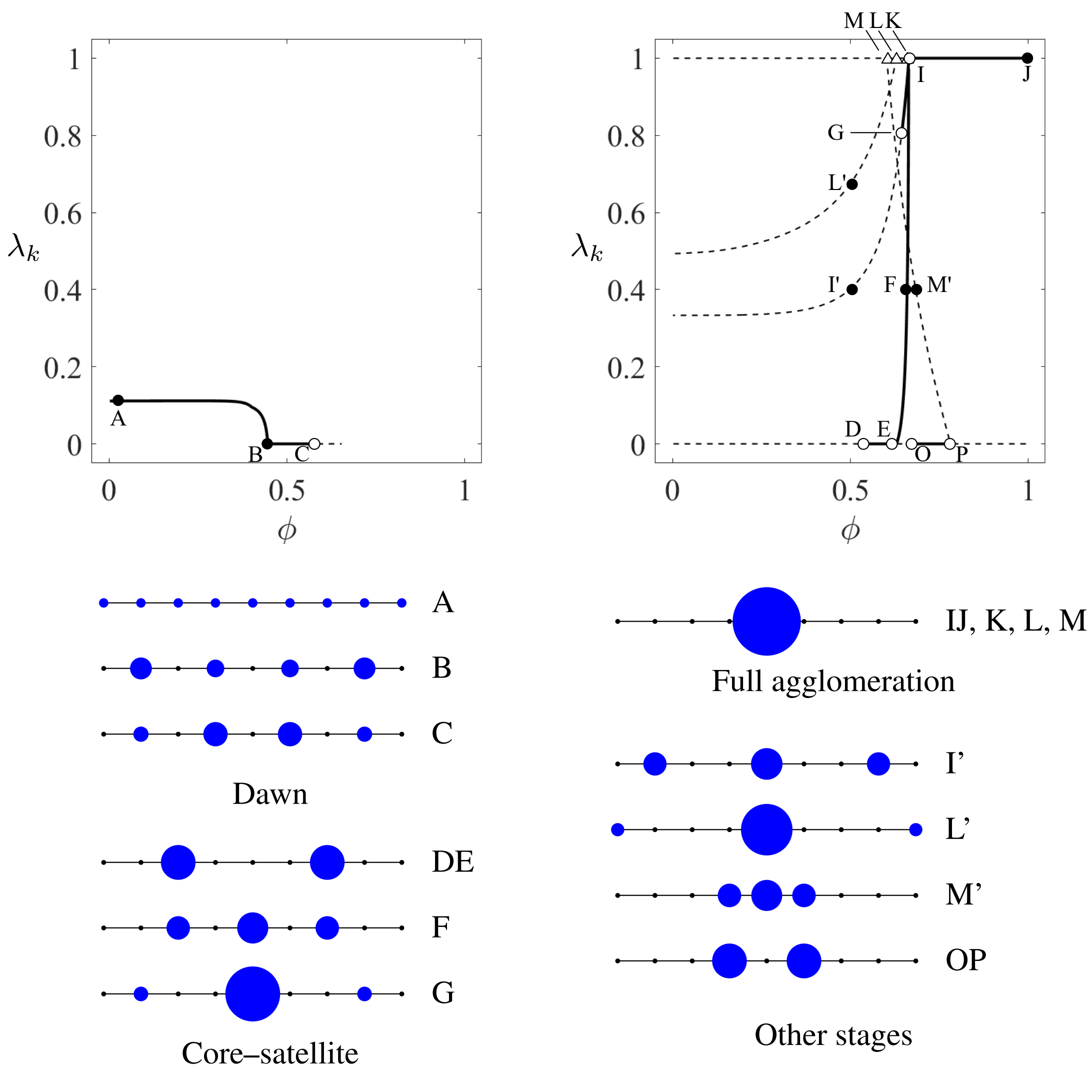


Figure F.1: Paths of equilibria for $K = 9$ cities for $(\sigma, \mu) = (6.0, 0.4)$ (the sustain point I with a branch IG is located closely to a bifurcation point K with a branch KF) (solid line: stable and sustainable; broken line: unstable and/or unsustainable; $\triangle$: bifurcation point; $\circ$: sustain point)

[25] Another bifurcating path for which the population of one satellite city at $i = 0$ or $i = 4$ becomes none-zero is unstable and is not included in Fig. 9.

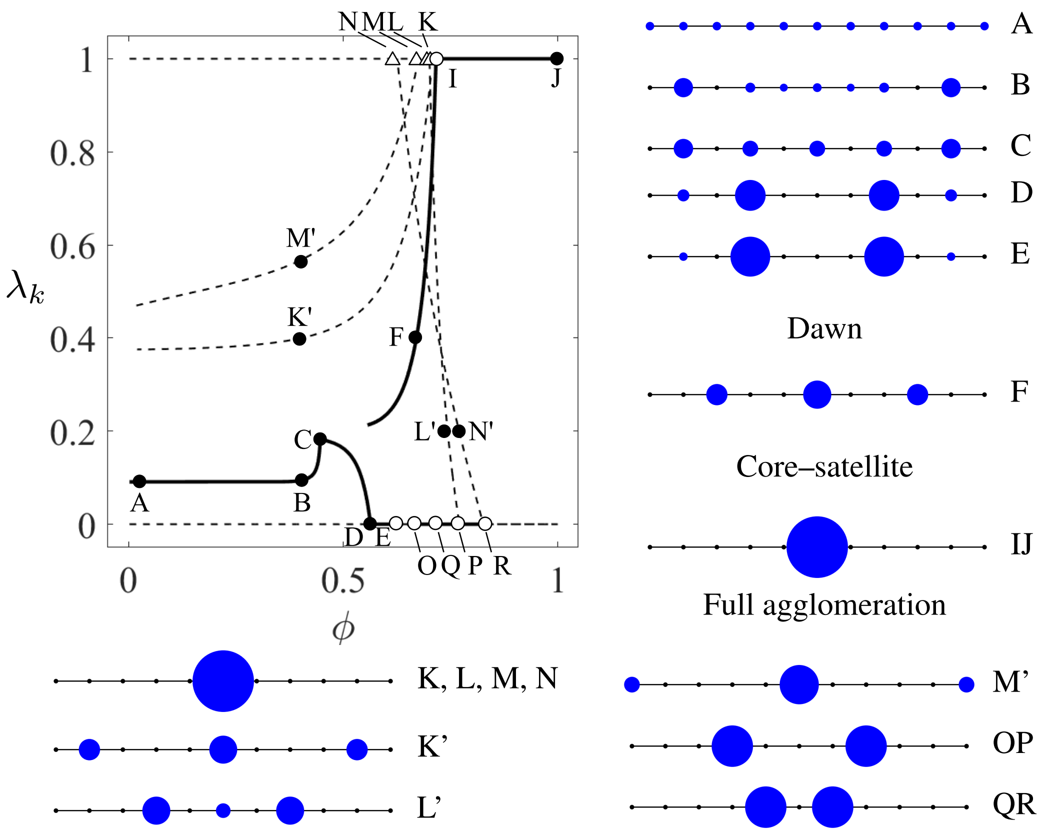


Figure F.2: Paths of equilibria for $K = 11$ cities for $(\sigma, \mu) = (6.0, 0.4)$ (solid line: stable and sustainable; broken line: unstable and/or unsustainable; $\triangle$: bifurcation point; $\circ$: sustain point)

References

[1] T. Ago, I. Isono, T. Tabuchi. Locational disadvantage of the hub, *Annals of Regional Science* 40(4) (2006) 819–848.

[2] H. Aizawa, K. Ikeda, M. Osawa, J. M. Gaspar. Breaking and sustaining bifurcations in S_N-Invariant equidistant economy, *International Journal of Bifurcation and Chaos* 30(16), 2050240.

[3] T. Akamatsu, T. Mori, M. Osawa, Y. Takayama. Endogenous agglomeration in a many-region world, arXiv preprint arXiv:1912.05113 (2020) 16-E-009, 2016.

[4] T. Akamatsu, Y. Takayama, K. Ikeda. Spatial discounting, Fourier, and racetrack economy: A recipe for the analysis of spatial agglomeration models, *Journal of Economic Dynamics and Control* 36(11) (2012) 1729–1759.

[5] T. Allen, K. Arkolakis. Trade and the Topography of the Spatial Economy, *The Quarterly Journal of Economics* 129(3) (2014) 1089–1140.

[6] Anderson, J. E., Van Wincoop, E. Trade costs, *Journal of Economic Literature* 42(3) (2004) 691–751.

[7] W. B. Arthur. 'Silicon Valley' locational clusters: When do increasing returns imply monopoly? *Mathematical Social Sciences* 19(3) (1990) 235–251.

[8] K. Behrens, Y. Murata. On quantitative spatial economic models, *Journal of Urban Economics* 123 (2021) 103348.

[9] S. B. S. D. Castro, J. Correia-da-Silva, J. M. Gaspar. Economic Geography meets

Hotelling: the home-sweet-home effect, *Economic Theory* forthcoming (2021) 1–27.

[10] W. Christaller. *Die zentralen Orte in Süddeutschland*, Gustav Fischer, Jena, 1933. English translation: *Central Places in Southern Germany*, Prentice Hall, Englewood Cliffs, 1966.

[11] Eurostat. Regional demographic statistics, Population on 1 January (2017).

[12] H. Flam, E. Helpman. Industrial policy under monopolistic competition, *Journal of International Economics* 22 (1987) 79–102.

[13] R. Forslid, G. I. P. Ottaviano. An analytically solvable core-periphery model, *Journal of Economic Geography* 3 (2003) 229–340.

[14] M. Fujita, P. Krugman, A. J. Venables. *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press, Cambridge, 1999.

[15] M. Fujita, T. Mori. Structural stability and the evolution of urban systems, *Regional Science and Urban Economics* 42 (1997) 399–442.

[16] M. Fujita, T. Mori. Frontiers of the New Economic Geography, *Papers in Regional Science* 84(3) (2005) 377–405.

[17] J. M. Gaspar. A prospective review on New Economic Geography, *The Annals of Regional Science* 61(2) (2018) 237–272.

[18] J. M. Gaspar. New Economic Geography: history and debate, *The European Journal of the History of Economic Thought* 28 (2021) 46–82.

[19] J. M. Gaspar, S. B. S. D. Castro, J. Correia-da-Silva. The Footloose Entrepreneur model with a finite number of equidistant regions, *International Journal of Economic Theory* 16(4) (2019) 420–446.

[20] J. M. Gaspar, K. Ikeda, M. Onda. Global bifurcation mechanism and local stability of identical and equidistant regions: Application to three regions and more, *Regional Science and Urban Economics* 86 (2021) 103597.

[21] M. Golubitsky, I. Stewart, D. G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, Vol.2, Springer-Verlag, New York, 1988.

[22] J. Gottmann. Megalopolis or the Urbanization of the Northeastern Seaboard, *Economic Geography* 33(3) (1957) 189–200.

[23] E. Helpman. The Size of Regions, In: Pines, D., Sadka. E. and Zilcha, I., Eds., *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge University Press, Cambridge (1998) 33–54.

[24] K. Ikeda, T. Akamatsu, T. Kono. Spatial period-doubling agglomeration of a core–periphery model with a system of cities, *Journal of Economic Dynamics and Control* 36 (2012) 754–778.

[25] K. Ikeda, K. Murota. *Imperfect Bifurcation in Structures and Materials*, 3rd ed., Springer-Verlag, New York, 2019.

[26] K. Ikeda, K. Murota, T. Akamatsu, Y. Takayama. Agglomeration patterns in a long narrow economy of a new economic geography model: Analogy to a racetrack economy, *International Journal of Economic Theory* 13 (1) (2017) 113-145.

[27] K. Ikeda, M. Onda, Y. Takayama. Spatial period doubling, invariant pattern, and break point in economic agglomeration in two dimensions, *Journal of Economic Dynamics and Control* 92 (2018) 129–152.

[28] K. Ikeda, Y. Takayama, M. Onda, D. Murakami. Group-theoretic spectrum analysis of population distribution in Southern Germany and Eastern USA, *International Journal of Bifurcation and Chaos* 28 (14) (2019) 1830045.

[29] Y. M. Ioannides, H. G. Overman. Spatial evolution of the US urban system, *Journal of Economic Geography* 4(2) (2004) 131–156.

[30] P. Krugman. Increasing returns and economic geography, *Journal of Political Economy* 99 (1991) 483–499.

[31] T. Mori. A modeling of megalopolis formation: the maturing of city systems, *Journal of Urban Economics* 42 (1997) 133–157.

[32] P. Mossay, P. M. Picard. On spatial equilibria in a social interaction model, *Journal of Economic Theory* 146(6) (2011) 2455–2477.

[33] Y. Murata, J.-F. Thisse. A simple model of economic geography á la Helpman–Tabuchi, *Journal of Urban Economics* 58(1) (2005) 137–155.

[34] G. Ottaviano, T. Tabuchi, J.-F. Thisse. Agglomeration and trade revisited, *International Economic Review* 43(2) (2002) 409–435.

[35] M. Pflüger. A simple, analytically solvable, Chamberlinian agglomeration model, *Regional Science and Urban Economics* 34(5) (2004) 565–573.

[36] D. Puga. The rise and fall of regional inequalities, *European Economic Review* 43(2) (1999) 303–334.

[37] S. J. Redding, E. Rossi-Hansberg. Quantitative spatial economics, *Annual Review of Economics* 9 (2017) 21–58.

[38] S. J. Redding, D. M. Sturm. The Costs of Remoteness: Evidence from German Division and Reunification, *American Economic Review* 99(5) (2008) 1766–1797.

[39] F. Robert-Nicoud. The structure of simple New Economic Geographymodels (or, On identical twins), *Journal of Economic Geography* 5(2) (2005) 201–234.

[40] W. H. Sandholm. *Population Games and Evolutionary Dynamics*, MIT Press, Cambridge, 2010.

[41] T. Tabuchi. Urban agglomeration and dispersion: a synthesis of Alonso and Krugman, *Journal of Urban Economics* 44(3) (1998) 333–351.

[42] T. Tabuchi, J.-F. Thisse. A new economic geography model of central places, *Journal of Urban Economics* 69 (2011) 240–252.

[43] T. Tabuchi, J.-F. Thisse, D.-Z. Zheng. On the number and size of cities, *Journal of Economic Geography* 5(4) (2005) 423–448.

[44] M. Turner, J.-F. Thisse., P. Ushchev. A unified theory of cities, Unpublished manuscript (2021) 1–45.

[45] United Nations, Department of Economic and Social Affairs, Population Division. World Urbanization Prospects: The 2018 Revision, Online Edition (2018).