Endogenous agglomeration in a many-region world

Akamatsu, Takashi and Mori, Tomoya and Osawa, Minoru and Takayama, Yuki

Tohoku University, Kyoto University, Kanazawa University

24 January 2017

Endogenous Agglomeration in a Many-region World*

Takashi Akamatsu
Tohoku University

Tomoya Mori
Kyoto University, RIETI

Minoru Osawa†
Tohoku University

Yuki Takayama
Kanazawa University

December 10, 2019

Abstract: We theoretically study a general family of economic geography models that features endogenous agglomeration. In many-region settings, the spatial scale—global or local—of the dispersion force(s) in a model plays a key role in determining the resulting endogenous spatial patterns and comparative statics. A global dispersion force accrues from competition between different locations and leads to the formation of multiple economic clusters, or cities. A local dispersion force is caused by crowding effects within each location and induces the flattening of each city. By distinguishing local and global dispersion forces, we can reduce a wide variety of extant models into only three prototypical classes that are qualitatively different in implications. Our framework adds consistent interpretations to the empirical literature and also provides general predictions on treatment effects in structural economic geography models.

Keywords: agglomeration; dispersion; economic geography; many regions; spatial scale.

JEL Classification: C62, R12, R13

*We are indebted to Michael Pflüger, Esteban Rossi-Hansberg, and Jacques-François Thisse for their constructive comments on the previous version of this paper, entitled "Spatial scale of agglomeration and dispersion: Theoretical foundations and empirical implications" (Akamatsu et al., 2017b). We are grateful for all seminar and conference participants at various institutions. This research has been supported by JSPS KAKENHI 17H00987, 18H01556, and 19K15108.

†Corresponding author. School of Engineering, Tohoku University, Aoba 6-6, Aoba, Sendai, Miyagi, 980-8579 Japan. Tel: +81 22 795 7419. Email: minoru.osawa.a5@tohoku.ac.jp / osawa.minoru@gmail.com

1 Introduction

Population and economic activities are astoundingly localized in space. For any spatial resolution—within countries, regions, or cities—disproportional concentrations of population, firms, or shops are present. For instance, just five cities (MSAs) of the United States, which make up about 5% of its cultivated land area, produces over 20% of the country's nominal GDP (as of 2017). The three major prefectures in Japan account for over 30% of the nominal GDP and 20% of the total population of the country, while taking up less than 5% of the total inhabitable area in the country (as of 2015).

Over the past four decades, the field of spatial economics has developed numerous theoretical and quantitative models to account for the uneven distribution of economic activities across cities and regions. The rich vein of theoretical modeling for endogenous agglomeration (e.g., Fujita et al., 1999a; Baldwin et al., 2003; Duranton and Puga, 2004; Fujita and Thisse, 2013) has been an important source of intuition-building devices in economics. In simplified geographical environments such as two-region models, the peaks and troughs in the space economy are explained as the endogenous outcomes of the various trade-offs between positive and negative incentives for spatial concentration. The accumulated knowledge for the general equilibrium modeling of spatial phenomena, together with the increased availability of fine spatial economic data, has allowed economists to construct quantitative models in a progressively detailed manner (see, e.g., Redding and Rossi-Hansberg, 2017 and Proost and Thisse, 2019, Section 5.2, for surveys). The exponential increase in the number of quantitative studies motivates us to ask the following question: is there any general and systematic means to classify and interpret the various spatial economic models that are proposed in different contexts?

This paper thus introduces a general classification that sets the basis for a unified taxonomy of theoretical or structural spatial economic models, irrespective of their micro-level assumptions. Our theory considers endogenous agglomeration based on ex-ante uniformity, in the spirit of Krugman (1991b). We study an important family of economic geography models that encompasses a wide range of extant models and covers all models—to the best of our knowledge—with a continuum of homogeneous agents with constant-elasticity-of-substitution preferences and a single type of iceberg interregional transportation costs (e.g., Krugman, 1991b; Helpman, 1998; Allen and Arkolakis,

(a) Many cities

(b) A single megacity

Figure 1: “Many-cities” and “single-megacity” patterns

2014).[1] To derive insights independent of the detailed microfoundations of the models, we fix a stylized geography as testbed. We assume a many-region *racetrack economy*, as in Krugman (1993), in which regions with the same local characteristics are symmetrically located over a circle (Figure 4). This simple geography serves as a dedicated prism through which the endogenous interactions in the model are decomposed according to their dependence on the underlying proximity structure between locations. We will show in Sections 5 and 6 that our results offer empirical implications, including regression approaches and structural modeling.

Our main result (i.e., **Proposition 1**) characterizes the spatial patterns of endogenous agglomeration that can emerge from ex-ante symmetry (i.e., the uniform distribution) in our many-region circular economy. In essence, it shows that the predictions of a model on the overall spatial pattern is governed by the *spatial scale* of the endogenous negative externalities (or *dispersion force*) in the model, but not on the model’s microfoundations. The spatial scale of the dispersion force is *local* when the force arises from the congestion effects inside each region (e.g., urban costs due to higher land rent in cities) and it is *global* when the force depends on the proximity to other regions due to, for example, competition between locations (e.g., interregional trade induces competition between firms in different regions that are geographically close).[2] If the dispersion force in a model is global, a “many-cities” pattern emerges from symmetry (Figure 1a). If it is local instead, a “single-megacity” pattern emerges (Figure 1b).[3] If a model includes both dispersion force types, then both possibilities arise, depending on the transportation cost level.

[1]See Definition 1 for the family of models we cover.

[2]See Definition 4 for formal definitions of local and global dispersion forces.

[3]This contrast in the “number” of cities is intrinsic and robustly generalizes to various geographical assumptions beyond our stylized circular economy. See Appendix D for a discussion.

		Global	
		Absent	Present
Local	Absent	–	**Class I**
	Present	**Class II**	**Class III**

Table 1: Spatial scale(s) of dispersion force(s) and model classes

The dichotomy between local and global dispersion forces allows us to infer the basic implications of the model and propose a simple taxonomy of economic geography models based on three prototypical classes as follows. A model is in Class I (II) if it has only a global (local) dispersion force and Class III if it has both (Table 1).[4]

Our numerical simulations supplement the theoretical predictions based on local stability analysis in the vicinity of a uniform distribution. The difference between model classes appears in their responses to interregional transportation costs. For Class I models, *many small cities* endogenously emerge when the transportation cost is high (cf. Figure 1a). A decrease in the transportation cost induces a decrease in the number of cities, an increase in the spacing between them, and an increase in the size of each city. By contrast, in Class II models, when the transportation cost is high, there is a *single dispersed city* (cf. Figure 1b). When the transportation cost decreases, it causes "suburbanization" by reducing the peak population density of the city. Class III, which is the most general, is a synthesis of Classes I and II. That is, when the transportation cost is high, a Class III model behaves as a Class I model and many small cities emerge. When the transportation cost is low, a single dispersed city exists, similar to a Class II model. At moderate levels of transportation cost, *multiple dispersed city* are generated (see Figure 13b). A decrease in transportation costs simultaneously causes a decrease in the number of cities (as in Class I models) and the flattening of each city (as in Class II models).

Notably, this behavior of Class III models provides a consistent interpretation of the evolution of the population distribution in Japan during 1970–2015. This period witnessed an almost from-scratch improvement in interregional accessibility in Japan, since the development of highways and high-speed railway networks was triggered by the Tokyo Olympics of 1964. Numerically, the total highway (high-speed railway) length increased from 879 km (515 km) to 14,146 km (5,350 km),

[4]Definition 5 formally defines the Class I, II, and III models.

which is more than a 16 (10) times increase. Suppose that a "city" is defined by the set of contiguous 1 km $\times$ 1 km cells with a population density of at least 1,000/km^2 and a total population of at least 10,000.[5] As such, 302 cities survived throughout the 45-year period, experiencing an average 21% increase in population size (controlling for national population growth). That is, there was a selective concentration towards a subset of cities, analogous to the implications of Class I and III models. The process was also associated with a flattening at the local scale: there was a 94% increase in area size and a 22% decrease in population density for an average individual city, analogous to the predictions of Class II and III models.

We also offer an additional result (**Proposition 2**) that reveals the effects of exogenous regional advantages (e.g., differences in amenities or productivity), which play a key role in counterfactual analyses based on calibrated quantitative economic geography models (see Redding and Rossi-Hansberg, 2017). Naturally, for a given transportation cost level, an exogenously advantageous region attracts more population than the average. We show that, when interregional access improves from the transportation cost level, the role of exogenous regional advantages is strengthened and weakened in Class I and II models, respectively. If exogenous heterogeneity causes one region to attract more population, then such asymmetry will be magnified and reduced in Class I and II models, respectively. This again indicates that the spatial scale of the dispersion force in a given model crucially governs the comparative static results of the model over other details.

In sum, our theoretical results reduce numerous economic geography models to a few model classes, according to the spatial scale of their dispersion forces. Therefore, our approach is philosophically related to those of Arkolakis et al. (2012) or Allen et al. (2019), who formulate general model classes that encompass a wide range of trade models in the literature as special cases, focusing on their macro-level restrictions rather than on their micro-level assumptions. Our approach is complementary to theirs, in that we focus on economic geography models that feature the multiplicity of equilibria and endogenous agglomeration. Recent evidence suggests that the multiplicity of equilibria and path dependence matter in the space economy in the long run (Bleakley and Lin, 2012; Michaels and Rauch, 2018). Consequently, the models that feature endogenous regional asymmetry can be useful for long-term counterfactual analyses. However, a well-known drawback

[5] See Appendix B for details.

of such models is that they may exhibit complex behaviors and cause technical and computational difficulties. In this regard, additional knowledge on the relationship between the spatial scale of dispersion forces and the resulting spatial patterns may be useful. For instance, our classification can be employed for choosing models to quantify the presence of the possible multiplicity of equilibria. Class III is the most general and may thus replicate the reality best among the three classes, as in the context of Japan discussed above. Class I would suffice if we are interested in the global patterns of economic agglomeration such as the number and population size of cities. If a major city region with a monopolar structure is the scope of the analysis, then Class II may be a reasonable choice.

The remainder of this paper is organized as follows. Section 2 introduces a general class of economic geography models, which we call *canonical models*. The simplest geographical setup, a two-region economy, is explored as a primer for our approach. The formal definitions of spatial scale of dispersion forces are also introduced in this section. Section 3 presents the main result, that is, **Proposition 1**. Section 4 illustrates the key implications of the main result with a minimal example. Section 5 provides a more extensive numerical examples and discusses the relationship with the empirical literature. Section 6 considers the effects of asymmetries in regional characteristics, leading to the additional result (**Proposition 2**). Section 7 concludes the paper.

2 Basic framework

We introduce a generic format for the many-region economic geography models and explore three specific models in the literature in a classical two-region setup. Definition 1 introduces the *canonical models*, the fundamental model class we focus on. Definition 4 introduces the *spatial scale* of a dispersion force, which is the main concept used in this paper.

2.1 A general format

We adhere to the simplest form of economic geography models, that is, static models with a single type of mobile agents. Consider an economy comprised of N regions, where a *region* is the discrete spatial unit. Let $\mathcal{N} \equiv \{1, 2, \ldots, N\}$ be the set of regions. There exists a unit-mass continuum of mobile agents. Each agent chooses a region to locate in. Let $x_i \geq 0$ be the mass of agents in region i, whereby $\boldsymbol{x} \equiv (x_i)_{i \in \mathcal{N}}$ is the spatial distribution of agents. The set of all possible spatial distributions is $\mathcal{X} \equiv \{\boldsymbol{x} \geq \mathbf{0} \mid \sum_{i \in \mathcal{N}} x_i = 1\}$. For each $\boldsymbol{x} \in \mathcal{X}$, a payoff function $v_i(\boldsymbol{x})$ gives the payoff for the agents in region i. We assume that $\boldsymbol{v}(\boldsymbol{x}) \equiv (v_i(\boldsymbol{x}))_{i \in \mathcal{N}}$ is one-time differentiable if $x_i > 0$ for all $i \in \mathcal{N}$.

Agents can freely relocate across N regions to improve their payoffs. Then, $\boldsymbol{x} \in \mathcal{X}$ is a *spatial equilibrium* if the following Nash equilibrium condition is met:

$$\begin{cases} v^* = v_i(\boldsymbol{x}) \text{ for all regions } i \in \mathcal{N} \text{ with } x_i > 0, \\ v^* \geq v_i(\boldsymbol{x}) \text{ for any region } i \in \mathcal{N} \text{ with } x_i = 0, \end{cases} \tag{1}$$

where v^* is the associated equilibrium payoff level.

An indispensable feature of an economic geography model is the presence of spatial frictions, or distance-decay effects, for the shipment of goods or for communication among agents. That is, $\boldsymbol{v}$ depends on a *proximity matrix* $\mathbf{D} = [\phi_{ij}]$ that summarizes the interregional transportation costs. Each entry $\phi_{ij} \in (0, 1]$ is the freeness of interactions between regions i and j. Such a structure of $\boldsymbol{v}$ is ubiquitous when we assume “iceberg” spatial frictions.

Payoff function $\boldsymbol{v}$ can include positive and negative externalities of spatial concentration, which may depend on interregional transportation costs. Owing to the positive externalities, economic geography models often face multiple spatial equilibria. As such, it is customary to introduce equilibrium refinement based on *local stability* under myopic dynamics. We follow this strategy. All the formal claims on the stability of equilibria in this paper hold true for the various standard dynamics employed in the literature. Remark C.4 in Appendix C provides concrete examples of the dynamics we cover.

Formal results in the remainder of this section are the corollaries of **Proposition 1** to be provided in Section 3. See the proof of **Proposition 1** in Appendix A.

2.2 A first view of endogenous agglomeration

The stability of a spatial equilibrium is parameter dependent. Particularly, changes in transportation costs can trigger a spontaneous emergence of regional asymmetry due to the instability of spatial uniformity (Papageorgiou and Smith, 1983).

For illustration purposes, we start with a classical two-region setup ($N = 2$). There are two regions that have identical characteristics, that is, there are no exogenous advantages. The proximity matrix for this setup is expressed as:

$$\mathbf{D} = \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix}, \tag{2}$$

where $\phi \in (0, 1)$ is the freeness of the interaction between the two regions. The interpretation of ϕ depends on context.

The uniform distribution of agents, $\bar{\boldsymbol{x}} \equiv (\bar{x}, \bar{x})$ with $\bar{x} = \frac{1}{2}$, is always a spatial equilibrium. How such a symmetric spatial distribution becomes unstable and an endogenous regional asymmetry of the form $\boldsymbol{x} = (x', x'')$ with $x' > x''$ is generated?

There is a general model-independent characterization: $\bar{\boldsymbol{x}}$ is stable (unstable) if the payoff gain of an agent relocating from one region to the other is negative (positive). The gain for a deviant can be evaluated by the following elasticity of the payoff difference:

$$\omega = \frac{\bar{x}}{\bar{v}} \frac{\partial(v_1(\bar{\boldsymbol{x}}) - v_2(\bar{\boldsymbol{x}}))}{\partial x_1} = \frac{\bar{x}}{\bar{v}} \left(\frac{\partial v_1(\bar{\boldsymbol{x}})}{\partial x_1} - \frac{\partial v_2(\bar{\boldsymbol{x}})}{\partial x_1} \right), \tag{3}$$

where $\bar{v}$ is the uniform payoff level at $\bar{\boldsymbol{x}}$, so that $\boldsymbol{v}(\bar{\boldsymbol{x}}) = (\bar{v}, \bar{v})$.

If $\omega < 0$, then $\bar{\boldsymbol{x}}$ is stable because there are no incentives for agents to migrate; $\omega < 0$ indicates that a marginal increase in the mass of agents in a region induces a *relative decrease* in the payoff therein. The instability of $\bar{\boldsymbol{x}}$ for $\omega > 0$ follows the same logic: if a small fraction of agents relocate from region 2 to 1, this induces a *relative increase* of the payoff in region 1, encouraging further migration from region 2. If we start from a state where $\bar{\boldsymbol{x}}$ is stable ($\omega < 0$), the endogenous regional asymmetry emerges when gains become positive ($\omega > 0$). The monotonic changes of freeness of interregional access ϕ can trigger such qualitative transitions, as demonstrated by Krugman (1991b).

Let $\mathbf{V} \equiv \frac{\bar{x}}{\bar{v}} \nabla \boldsymbol{v}(\bar{\boldsymbol{x}})$ be the matrix of the payoff elasticity, evaluated at $\bar{\boldsymbol{x}}$, where $\nabla \boldsymbol{v}(\bar{\boldsymbol{x}}) = [\frac{\partial v_i}{\partial x_j}(\bar{\boldsymbol{x}})]$ is the corresponding Jacobian matrix of $\boldsymbol{v}(\boldsymbol{x})$. Then, ω is an eigenvalue of $\mathbf{V}$ with eigenvector $\boldsymbol{z} \equiv (1, -1)$, because (3) implies that $\omega \boldsymbol{z} = \mathbf{V} \boldsymbol{z}$, which is the definition of an eigenvalue–

eigenvector pair.[6] Since $\boldsymbol{z}$ represents a population increase in one region and a decrease in the other, it is the *migration pattern* in the two-region economy. Obviously, $\boldsymbol{z}$ is *model independent*.

The concrete form of ω is model dependent. We focus on a specific family of models, which we call *canonical models*. Canonical models encompass a wide range of extant economic geography models. In particular, they include models that assume (i) a single type of homogeneous mobile agents with constant-elasticity-of-substitution preferences and (ii) a single sector that is subject to iceberg interregional transportation costs.[7]

Definition 1 (Canonical models)**.** Consider economic geography model $\boldsymbol{v}$ with proximity matrix $\mathbf{D} = [\phi_{ij}]$. Let $\bar{\mathbf{D}}$ be the row-normalized version of $\mathbf{D}$, whose (i, j)th element is given by $\frac{\phi_{ij}}{\sum_{k \in \mathcal{N}} \phi_{ik}}$. Let $\mathbf{V} = \frac{\bar{x}}{\bar{v}} \nabla \boldsymbol{v}(\bar{\boldsymbol{x}})$ be the payoff elasticity matrix at $\bar{\boldsymbol{x}}$. The model is *canonical* if there exists a rational function G that is continuous over $[0, 1]$ and satisfies

$$\mathbf{V} = G(\bar{\mathbf{D}}). \tag{4}$$

We call G the *gain function* of the model.

In Definition 1, for a rational function (i.e., the ratio of two polynomials) of form $G(t) = \frac{G^\sharp(t)}{G^\flat(t)}$ with polynomials $G^\sharp(t)$ and $G^\flat(t) \neq 0$, we define $G(\bar{\mathbf{D}}) = G^\flat(\bar{\mathbf{D}})^{-1} G^\sharp(\bar{\mathbf{D}})$, where, for a polynomial $G^\sharp(t) = c_0 + c_1 t + c_2 t^2 + \cdots$, we let

$$G^\sharp(\bar{\mathbf{D}}) = c_0 \mathbf{I} + c_1 \bar{\mathbf{D}} + c_2 \bar{\mathbf{D}}^2 + \cdots, \tag{5}$$

with $\mathbf{I}$ being the identity matrix.[8]

For a wide range of general equilibrium economic geography models that incorporate gravity-form interregional trade, there are two matrix polynomials, $G^\sharp(\bar{\mathbf{D}})$ and $G^\flat(\bar{\mathbf{D}})$ that satisfy $\mathbf{V} =$

[6] That $\omega \boldsymbol{z} = \mathbf{V}\boldsymbol{z}$ follows because the indices of the regions are interchangeable. As $\mathbf{V}$ is a 2×2 matrix, there exists another eigenvector, namely $\mathbf{1} = (1, 1)$. The only relevant eigenvector is $\boldsymbol{z}$, because $\mathbf{1}$ corresponds to population increases in *both* regions, obviously violating the assumption that the total mass of mobile agents is fixed.

[7] As noted by Allen et al. (2019); Arkolakis et al. (2012), this class of models includes various important models in the literature. However, we should also note that the iceberg cost is not an innocuous assumption for modeling a spatial economy (see, e.g., Hummels and Skiba, 2004; Irarrazabal et al., 2015; Proost and Thisse, 2019, Section 3.5.2), although it is widely employed in the literature for tractability.

[8] The assumption that G is rational is not restrictive because any continuous function defined on a closed interval can be approximated as closely as desired by a polynomial (the Weierstrass approximation theorem).

$G^{\flat}(\bar{\mathbf{D}})^{-1}G^{\sharp}(\bar{\mathbf{D}})$; thereby, there is a rational function G that satisfies the hypotheses in Definition 1. We will see two examples in Section 2.3.[9] Definition 1 covers, for example, models of endogenous city center formation (e.g., Beckmann, 1976), single-industry monopolistically competitive economic geography models (e.g., Krugman, 1991b; Helpman, 1998), and economic geography variants of the "universal gravity" framework (Allen et al., 2019), which in turn encompasses perfectly competitive Armington models with labor mobility (Allen and Arkolakis, 2014). Section 3 provides more examples.[10]

When we have $\mathbf{V} = G(\bar{\mathbf{D}})$ with a rational function G, it is standard in matrix analysis that ω, the eigenvalue of $\mathbf{V}$ with eigenvector $\boldsymbol{z} = (1, -1)$, is given by

$$\omega = G(\chi) \quad \text{and} \quad \chi = \frac{1-\phi}{1+\phi}, \tag{6}$$

where χ is the eigenvalue of $\bar{\mathbf{D}} = \frac{1}{1+\phi}\mathbf{D}$ associated with $\boldsymbol{z} = (1, -1)$.[11] We see that $\chi \in (0, 1)$ is a monotonically decreasing continuous function of $\phi \in (0, 1)$. If ϕ is small (large), χ is large (small). Since G is continuous, $\omega = G(\chi(\phi))$ smoothly varies with ϕ.

Gain function G of a model summarizes the endogenous effects under the model. For example, consider the seminal model of Beckmann (1976) on the formation of an urban center within a city.[12]

Example 2.1 (The Beckmann model)**.** Numerous variants of the model have been proposed since the original formulation of Beckmann (e.g., Mossay and Picard, 2011; Blanchet et al., 2016). Consider the following multiplicative specification:

$$v_i(\boldsymbol{x}) = x_i^{-\gamma} E_i(\boldsymbol{x}), \tag{7}$$

where $\gamma > 0$. The first component, $x_i^{-\gamma}$, reflects *negative externalities* due to congestion and the second, $E_i(\boldsymbol{x})$, represents *positive externalities* arising from agents' preference for proximity to

[9]See also Appendix F.1 in Appendix F for a general derivation.

[10]See Remark C.1 in Appendix C for examples of extant models we do *not* cover. Canonical models do not include models based on Ottaviano et al. (2002) which assume quadratic preference and urban models with multiple types of mobile agents such as in Fujita and Ogawa (1982); Lucas and Rossi-Hansberg (2002). For the Ottaviano et al. (2002) framework with a single type of mobile agents, the results are similar to canonical models (see Remark C.1).

[11]See Fact E.1 in Appendix E.

[12]In this respect, a "region" would best considered an "urban zone" in the model.

others. A typical specification for $E_i(\boldsymbol{x})$ is

$$E_i(\boldsymbol{x}) = \sum_{j \in \mathcal{N}} \mathrm{e}^{-\tau \ell_{ij}} x_j, \tag{8}$$

where $\tau > 0$ is the distance-decay parameter and $\ell_{ij} > 0$ is the distance between i and j. The proximity matrix is expressed as $\mathbf{D} = [\mathrm{e}^{-\tau \ell_{ij}}]$. If $N = 2$, $\phi = \mathrm{e}^{-\tau \ell_{12}} = \mathrm{e}^{-\tau \ell_{21}} \in (0, 1)$ represents the level of externalities that spill over from one location to the other.

We have $\mathbf{V} = -\gamma \mathbf{I} + \bar{\mathbf{D}}$ and $\omega = -\gamma + \chi$. The model is therefore a canonical model with gain function $G(\chi) = -\gamma + \chi$. Negative term $-\gamma$ in $G(\chi)$ corresponds to the congestion effect through $x_i^{-\gamma}$ and positive term χ corresponds to positive externalities $E_i(\boldsymbol{x})$. The former is the loss from congestion, whereas the latter represents the gains from the additional proximity to be induced by migration. Thus, ω is the *net* gain from migration. When ϕ is close to 1, so that the relative location in the economy becomes irrelevant, χ disappears, leaving only congestion effect $-\gamma$.

If $\gamma < 1$, then $\bar{\boldsymbol{x}}$ is stable for $\phi \in (\phi^*, 1)$ and unstable for $\phi \in (0, \phi^*)$, where $\phi^* \equiv \frac{1-\gamma}{1+\gamma}$. There is some endogenous asymmetry when $\phi \in (0, \phi^*)$. If $\gamma \geq 1$, then $\bar{\boldsymbol{x}}$ is stable for *all* $\phi \in (0, 1)$. That is, strong congestion effects suppress endogenous agglomeration. ■

As per the example, a positive (negative) term in $\omega = G(\chi)$ represents the agglomeration (dispersion) force. Therefore, ω is the *net* agglomeration force. We introduce the following formal definitions.

Definition 2. A *dispersion (agglomeration) force* in a canonical model is a negative (positive) term in its gain function G.

2.3 The reversed scenarios of Krugman and Helpman

Other examples of canonical models are the general equilibrium models of Krugman (1991b) and Helpman (1998). In the two-region case, proximity matrix $\mathbf{D}$ is given by (2), where $\phi \equiv \tau^{1-\sigma}$ is the *freeness of trade* defined with $\tau > 1$, the "iceberg" transportation cost parameter between the two regions, and $\sigma > 1$, the elasticity of substitution between horizontally differentiated varieties.

On the ϕ-axis, the models are known to exhibit a sharp contrast regarding *when* endogenous regional asymmetry emerges, that is, the "Krugman's scenario is reversed" (Fujita and Thisse, 2013, Chapter 8) in the Helpman model. In the Krugman (Helpman) model, uniform distribution $\bar{x}$ is stable when ϕ is low (high) and asymmetry exists when ϕ is high (low). The model predictions are thus "opposites" of each other.

We provide below brief definitions of the many-region extensions for the models.[13]

Example 2.2 (The Krugman model)**.** The payoff function (the indirect utility of mobile workers) for the Krugman model is given by

$$v_i(\boldsymbol{x}) = w_i(\boldsymbol{x}) P_i(\boldsymbol{x})^{-\mu}, \tag{9}$$

where $w_i(\boldsymbol{x})$ is the nominal wage of mobile workers for a given spatial distribution of mobile workers $\boldsymbol{x}$ and $P_i(\boldsymbol{x})$ is the Dixit–Stiglitz price index in region i:

$$P_i(\boldsymbol{x}) \equiv \left(\sum_{j \in \mathcal{N}} x_j \left(w_j(\boldsymbol{x}) \tau_{ji} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \tag{10}$$

where $\mu \in (0,1)$ is the expenditure share of manufactured goods and $\tau_{ij} \geq 1$ the iceberg transportation cost parameter. That is, τ_{ij} units should be shipped from origin i for one unit to arrive at destination j. Nominal wage $\boldsymbol{w}(\boldsymbol{x}) = (w_i(\boldsymbol{x}))_{i \in \mathcal{N}}$ is the unique solution for a system of nonlinear equations that summarizes the market equilibrium conditions under a fixed $\boldsymbol{x}$ (i.e., the gravity flows of interregional trade, goods and labor market clearing, and the zero-profit condition of firms):

$$w_i x_i = \sum_{j \in \mathcal{N}} \frac{x_i \left(w_i \tau_{ij} \right)^{1-\sigma}}{\sum_{k \in \mathcal{N}} x_k \left(w_k \tau_{kj} \right)^{1-\sigma}} e_j \qquad \forall i \in \mathcal{N}, \tag{11}$$

where $e_i \equiv \mu \left(w_i x_i + l_i \right)$ is region i's expenditure on differentiated goods and $l_i > 0$ the region-fixed immobile demand. The proximity matrix for the model is $\mathbf{D} = [\phi_{ij}] = [\tau_{ij}^{1-\sigma}]$. ■

Example 2.3 (The Helpman model)**.** Using the same notation as in the Krugman model, the payoff

[13] See Appendix F for details.

function of mobile agents in the Helpman model is given by:

$$v_i(\boldsymbol{x}) = \left(\frac{x_i}{a_i}\right)^{-\gamma} (w_i(\boldsymbol{x}) + \bar{r})^{\mu} P_i(\boldsymbol{x})^{-\mu}, \tag{12}$$

where a_i is the endowment of housing stock in region i, $\gamma \equiv 1 - \mu \in (0, 1)$ the expenditure share of housing goods, and $\bar{r}$ an equal dividend from the total rental revenue from housing in the economy. The market equilibrium conditions under a given $\boldsymbol{x}$ are summarized by (11) where $e_i = \mu(w_i + \bar{r})x_i$, with $\boldsymbol{w}(\boldsymbol{x})$ being the unique solution. The proximity matrix for the model is the same as in the Krugman model. ■

We now confirm the "reversed scenario" using our notation. Both the Krugman and Helpman models are canonical models. Appendix F shows that $\mathbf{V} = \frac{\bar{x}}{\bar{v}} \nabla \boldsymbol{v}(\bar{\boldsymbol{x}})$ is given by

$$\mathbf{V} = G^{\flat}(\bar{\mathbf{D}})^{-1} G^{\sharp}(\bar{\mathbf{D}}), \tag{13}$$

where $G^{\flat}(\chi) \equiv 1 - \frac{\mu}{\sigma}\chi - \frac{\sigma-1}{\sigma}\chi^2$ and

$$G^{\sharp}(\chi) = \qquad c_1\chi - \qquad c_2\chi^2 \qquad \text{(the Krugman model)}, \tag{14}$$

$$G^{\sharp}(\chi) = -\gamma + c_1\chi - (c_2 - \gamma)\chi^2 \qquad \text{(the Helpman model)}, \tag{15}$$

with $c_1 \equiv \mu\left(\frac{1}{\sigma-1} + \frac{1}{\sigma}\right)$ and $c_2 \equiv \frac{\mu^2}{\sigma-1} + \frac{1}{\sigma}$. The gain functions for the models are given by:

$$\omega = G(\chi) = \frac{G^{\sharp}(\chi)}{G^{\flat}(\chi)}, \tag{16}$$

where $\chi = \frac{1-\phi}{1+\phi}$ with $\phi = \tau^{1-\sigma}$, as in (6).[14] Figure 2 shows $G^{\sharp}(\chi)$ for the Krugman and Helpman models, which are both quadratic.[15]

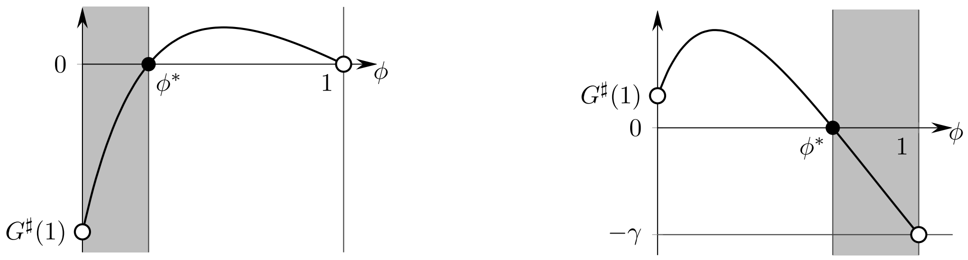
The stability of $\bar{\boldsymbol{x}}$ is dictated by the sign of numerator $G^{\sharp}(\chi)$ since $G^{\flat}(\chi) > 0$. That is, $\bar{\boldsymbol{x}}$ is stable if $G^{\sharp}(\chi) < 0$. In other words, $G^{\sharp}$ summarizes the *net* relative magnitudes of the agglomeration and

[14] Fujita et al. (1999a) calls χ "a sort of index of trade cost" (page 57), whereas Baldwin et al. (2003) calls it "a convenient measure of closed-ness" (page 46).

[15] Appendix F shows that, for many extant models, $\mathbf{V}$ is represented by up to the second-order term of the proximity matrix.

(a) The Krugman model (b) The Helpman model

Figure 2: $G^\sharp(\chi)$ for the Krugman and Helpman models



(a) The Krugman model (b) The Helpman model

Figure 3: $\omega^\sharp \equiv G^\sharp(\chi(\phi))$ for the Krugman and Helpman models

dispersion forces in each model.

The "reversed scenario" can be graphically verified using Figure 3. Composite function $\omega^\sharp \equiv G^\sharp(\chi(\phi))$ for each model is depicted in the figure. For each (14) and (15), there exists (at most) one root ϕ^* for $G^\sharp(\chi(\phi^*)) = 0$ in $(0,1)$.[16] We see

$$\bar{x} \text{ is stable when } \begin{cases} \phi \in (0, \phi^*) & \text{(the Krugman model),} \\ \phi \in (\phi^*, 1) & \text{(the Helpman model).} \end{cases}$$

As expected, $\bar{x}$ is stable for low (high) values of ϕ in the Krugman (Helpman) model and unstable otherwise. From (6), threshold ϕ^* is given by $\phi^* \equiv \frac{1-\chi^*}{1+\chi^*}$, where χ^* is the solution for $G^\sharp(\chi) = 0$ (see Figure 2).

[16]If no such ϕ^* exists, there is no switch in the stability of $\bar{x}$ for $\phi \in (0,1)$. If $\omega^\sharp = G^\sharp(\chi(\phi)) > 0$ for all $\phi \in (0,1)$, $\bar{x}$ is unstable for all ϕ, whereas $\omega^\sharp < 0$ for all $\phi \in (0,1)$ implies the contrary. We preclude these cases to focus on endogenous agglomeration due to the changes in ϕ.

2.4 Spatial scale of dispersion forces

According to Definition 2, the agglomeration force in the Krugman model is captured by the first term in (14) and the dispersion force by the second term. In the Helpman model, the second term in (15) reflects the agglomeration force, whereas the first and third terms reflect the dispersion forces.[17]

The two models have equivalent agglomeration forces. The common agglomeration force, $c_1\chi$ in $G^\sharp(\chi)$, arises from the price index of the differentiated varieties (10). Since a region with a larger set of suppliers in the market has a lower price index, mobile workers prefer such a region if the nominal wage is the same. This force is stronger when ϕ is low and declines as ϕ increases.

By elimination, the "reversed scenario" must stem from *differences in the dispersion forces*. The dispersion force in the Krugman model is the so-called *market-crowding effect* between firms (see Baldwin et al., 2003, Chapter 2). If a firm is geographically close to others, the firm can only pay a low nominal wage because of competition. Therefore, mobile workers are discouraged to enter a region in which firms face fierce market competition with other firms in that location as well as *nearby regions* thereof. The dispersion force thus depends on proximity structure $\mathbf{D}$ and appears as a negative second-order term, $-c_2\chi^2$. This force is stronger when χ is large, that is, when ϕ is small.

The main dispersion force in the Helpman model, on the other hand, represents a *local congestion effect*. The force stems from competition in the housing market of each region.[18] The local housing market does not depend on interregional trade cost structure $\mathbf{D}$ but only on the mass of agents *within* each region. The dispersion force thus appears in $G^\sharp(\chi)$ as negative constant term $-\gamma$. Since the agglomeration force ($c_1\chi$) declines as ϕ increases, the relative strength of the dispersion force rises with trade freeness ϕ.

The comparison between the Krugman and Helpman models highlights that the key difference is whether the dispersion force depends on the interregional transportation cost structure, $\mathbf{D}$. To denote this distinction, we introduce the formal notion of *spatial scale* of dispersion forces.

[17] The Helpman model exhibits endogenous asymmetry if $\mu > \frac{\sigma-1}{\sigma}$. This condition implies that $c_2 - \gamma > 0$ and, thus, the last term in (15) is negative.

[18] The market-crowding effect also exists in the Helpman model: $-(c_2-\gamma)\chi^2$ in $G^\sharp(\chi)$. However, in contrast to the Krugman model, it does not have a stabilizing power when ϕ is small, due to the absence of immobile factors in production. Technically, $G^\sharp(\chi(0)) > 0$ and $\bar{x}$ is unstable when ϕ is small.

We first define *net gain functions* to simplify the definition of the spatial scale of dispersion forces. In essence, we ignore the denominator of G, $G^\flat$, which is positive and thus irrelevant for the stability of $\bar{x}$.

Definition 3. A *net gain function* $G^\sharp$ for a canonical model with gain function G is a polynomial that satisfies $\mathrm{sgn}[\omega] = \mathrm{sgn}[G(\chi)] = \mathrm{sgn}[G^\sharp(\chi)]$ for all $\chi \in (0, 1)$.

The net gain functions for the Krugman and Helpman models are, respectively, given by (14) and (15) because $G^\flat(\chi) > 0$ for all $\chi \in (0, 1)$. For the Beckmann model, we see $G^\sharp(\chi) = G(\chi) = -\gamma + \chi$.

We can introduce the spatial scale of dispersion forces, which refines the definition of dispersion forces (Definition 2).

Definition 4 (Spatial scale of dispersion forces)**.** A negative constant term in net gain function $G^\sharp(\chi)$ is called a *local dispersion force*. A negative non-constant term in $G^\sharp(\chi)$ is called a *global dispersion force*.

The main dispersion force in the Krugman (Helpman) model is global (local). A global dispersion force is triggered when ϕ is low and a local dispersion force when ϕ is high due to an increase in its relative importance. The "reversed scenario" of Krugman model and Helpman model stems from the differences in the spatial scales of their dispersion forces.

3 Classification of canonical models

We show that there is a major watershed between "Krugman-like" and "Helpman-like" models in terms of endogenous spatial patterns in many-region economy, and that the spatial scale of the dispersion force plays the key role. By considering a *racetrack economy* à la Krugman (1993), this section presents the main result, **Proposition 1**. It provides a categorization of endogenous spatial distributions that can emerge from the spatially uniform distribution in canonical models.

Consider an N-region economy in which regions are symmetrically placed over a circumference and interactions are possible only through the circular network (Figure 4).

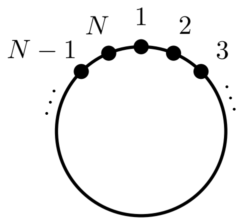


Figure 4: N-region racetrack economy.

Assumption RE. Proximity matrix $\mathbf{D} = [\phi_{ij}]$ is given by $\phi_{ij} = \phi^{\ell_{ij}}$, where $\phi \in (0,1)$ is the freeness of transportation between two consecutive regions and $\ell_{ij} \equiv \min\{|i-j|, N-|i-j|\}$ is the shortest-path distance over the circumference. N is a multiple of four.[19]

In line with Sections 2, we assume that payoff function v does not introduce any ex-ante asymmetries across regions. Technically, this can be formalized as:[20]

Assumption S. For all $x \in \mathcal{X}$, payoff function v satisfies $v(\mathbf{P}x) = \mathbf{P}v(x)$ for all permutation matrices $\mathbf{P}$ that satisfy $\mathbf{PD} = \mathbf{DP}$.

Example 3.1. Suppose $N = 4$. Then, Assumption RE is that

$$\mathbf{D} = \begin{bmatrix} 1 & \phi & \phi^2 & \phi \\ & 1 & \phi & \phi^2 \\ & & 1 & \phi \\ Sym. & & & 1 \end{bmatrix}. \tag{17}$$

The shape of the circular economy is invariant even if we swap the indices of regions 1 and 3. The following permutation matrix represents this re-indexing:

$$\mathbf{P} = \begin{bmatrix} & & 1 & \\ & 1 & & \\ 1 & & & \\ & & & 1 \end{bmatrix}, \tag{18}$$

and $\mathbf{P}$ satisfies the hypothesis of Assumption S that $\mathbf{PD} = \mathbf{DP}$. Condition $\mathbf{PD} = \mathbf{DP}$ ensures that the adjacency relationships between regions remain invariant under the permutation of the indices

[19] N is a multiple of four only for expositional simplicity. See Appendix A.1.

[20] Assumption S is called *equivariance*. See Golubitsky and Stewart (2003) for details.

represented by $\mathbf{P}$. The re-indexed spatial distribution is $\boldsymbol{x}' = \mathbf{P}\boldsymbol{x}$, where $\boldsymbol{x}$ is the original one. If $\boldsymbol{v}$ does not include any exogenous advantages, we must have $v_1(\boldsymbol{x}') = v_3(\boldsymbol{x})$, $v_2(\boldsymbol{x}') = v_2(\boldsymbol{x})$, $v_3(\boldsymbol{x}') = v_1(\boldsymbol{x})$, and $v_4(\boldsymbol{x}') = v_4(\boldsymbol{x})$, that is, $\boldsymbol{v}(\boldsymbol{x}') = \boldsymbol{v}(\mathbf{P}\boldsymbol{x}) = \mathbf{P}\boldsymbol{v}(\boldsymbol{x})$ as in Assumption S. ■

Uniform pattern $\bar{\boldsymbol{x}} \equiv (\bar{x}, \bar{x}, \ldots, \bar{x})$ ($\bar{x} \equiv \frac{1}{N}$) is a spatial equilibrium under Assumptions RE and S. The question is what are the spatial patterns that can emerge due to the destabilization of $\bar{\boldsymbol{x}}$ through purely endogenous mechanisms.

Consider an infinitesimally small migration of agents $\boldsymbol{z} = (z_i)_{i \in \mathcal{N}}$ from $\bar{\boldsymbol{x}}$ so that the new spatial distribution becomes $\boldsymbol{x}' \equiv \bar{\boldsymbol{x}} + \boldsymbol{z}$. We require $\sum_{i \in \mathcal{I}} z_i = 0$, so that the mass of agents does not change. Analogous to the two-region case, the marginal gain for agents due to such a deviation can be evaluated by

$$\bar{\omega} \equiv \frac{\bar{x}}{\bar{v}} \left(\sum_{i \in \mathcal{N}} v_i(\boldsymbol{x}') x_i' - \sum_{i \in \mathcal{N}} v_i(\bar{\boldsymbol{x}}) \bar{x} \right) = \boldsymbol{z}^\top \mathbf{V} \boldsymbol{z}, \tag{19}$$

where $\bar{v}$ is the uniform level of payoff at $\bar{\boldsymbol{x}}$ and $\mathbf{V} = \frac{\bar{x}}{\bar{v}} \nabla \boldsymbol{v}(\bar{\boldsymbol{x}})$ is the payoff elasticity matrix. If $\bar{\omega} < 0$ for *any migration pattern* $\boldsymbol{z}$, then $\bar{\boldsymbol{x}}$ is stable.

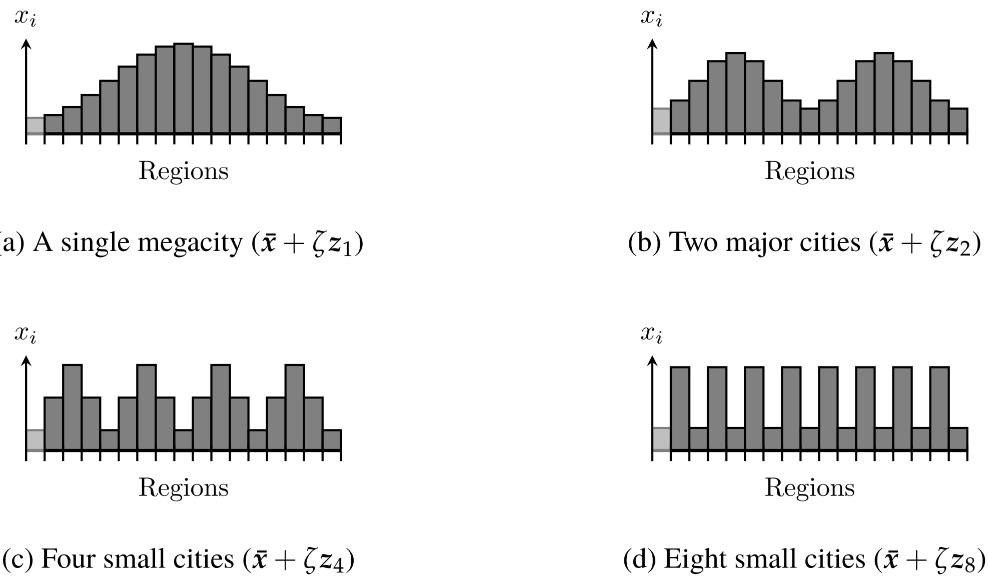
Under Assumptions RE and S, there is a *model-independent* way to conveniently represent all possible migration patterns:

$$\boldsymbol{z} = \sum_k \zeta_k \boldsymbol{z}_k, \tag{20}$$

where $\{\boldsymbol{z}_k\}$ are the eigenvectors of $\mathbf{V}$ and $\{\zeta_k\}$ are their coefficients. We normalize $\|\boldsymbol{z}_k\|^2 = \boldsymbol{z}_k^\top \boldsymbol{z}_k = 1$ for all k. Each $\boldsymbol{z}_k$ is in itself a migration pattern and is a cosine curve with k equally spaced peaks. We thus interpret $\boldsymbol{z}$ as the weighted sum of the "basic" migration patterns $\{\boldsymbol{z}_k\}$. Basic migration patterns are model-independent in the sense that they are the eigenvectors of $\mathbf{V}$ irrespective of the properties of payoff function $\boldsymbol{v}(\boldsymbol{x})$.

There are essentially $\frac{N}{2}$ basic migration patterns, since the concentration of agents in every other region achieves the maximum number of symmetric cities.[21] That is, in contrast to the two-region

[21]Concretely, $\{\boldsymbol{z}_k\}$ correspond to the real discrete Fourier modes for dimension N. They are of the form $\boldsymbol{z}_k \propto (\cos(\theta k i))$ where $\theta \equiv \frac{2\pi}{N}$ for $k = 1, 2, \ldots, \frac{N}{2}$, and $\boldsymbol{z}_k \propto (\sin(\theta(N-k)i))$ for $k = \frac{N}{2} + 1, \ldots, N$. Therefore, the number of cities (peaks) is the largest when $k = \frac{N}{2}$. See Appendix A.1.



(a) A single megacity ($\bar{x} + \zeta z_1$)

(b) Two major cities ($\bar{x} + \zeta z_2$)

(c) Four small cities ($\bar{x} + \zeta z_4$)

(d) Eight small cities ($\bar{x} + \zeta z_8$)

Figure 5: Schematic illustrations of migration patterns ($N = 16$).

Notes: We rotationally shift the spatial distribution and add the neighboring region of the leftmost region as a light gray bar, so that it is easier to grasp the overall shapes. For expositional simplicity, we present the cases when k is a power of 2.

economy where $z = (1, -1)$ is the only possible migration pattern, there are multiple possibilities in the many-region economy.

Example 3.2. Figure 5 shows spatial patterns $\bar{x} + \zeta z_k$ ($k = 1, 2, 4, 8$) for $N = 16$ with a small $\zeta > 0$. Basic migration patterns z_1, z_2, z_4, and z_8 express, respectively, the formation of a single megacity (Figure 5a), two major cities (Figure 5b), four small cities (Figure 5c), and eight small cities (Figure 5d). ■

Let ω_k be the eigenvalue of $\mathbf{V}$ associated with z_k (i.e., $\omega_k z_k = \mathbf{V} z_k$). Then, (20) yields

$$\bar{\omega} = \sum_k \zeta_k^2 z_k^\top \mathbf{V} z_k = \sum_k \zeta_k^2 \omega_k. \tag{21}$$

Thus, the migration pattern that maximizes $\bar{\omega}$ is the basic migration pattern that has the largest eigenvalue:

$$\omega_{\max} \equiv \max_{\|z\|^2=1} \bar{\omega} = \max_k \omega_k, \tag{22}$$

where we normalize $\|z\|^2 = z^\top z = 1$ without loss of generality.

Thus, $\bar{x}$ is stable when $\omega_{\max} = \max_k \omega_k < 0$. Further, $\bar{x}$ becomes unstable to form endogenous regional asymmetry when $\omega_{\max}$ become positive. To put it differently, each ω_k is the gain from migration towards z_k-direction, and $\bar{x}$ becomes unstable when migration towards some direction becomes profitable. This extends the discussion in the two-region case based on ω to our many-region setting.

We need the concrete formulae for eigenvalues $\{\omega_k\}$ of $\mathbf{V}$. Because we consider canonical models, we have $\mathbf{V} = G(\bar{\mathbf{D}})$ where $\bar{\mathbf{D}}$ is row-normalized proximity matrix and G some rational function. The two-region formula $\omega = G(\chi)$ can then be generalized as follows:

$$\omega_k = G(\chi_k) \qquad \forall k \in \mathcal{K}, \tag{23}$$

where χ_k is the eigenvalue of $\bar{\mathbf{D}}$ associated with z_k.

Each χ_k is an index of the average geographical proximity among agents when the k-city pattern $\bar{x} + \zeta z_k$ emerges. Further, χ_k decreases in number of cities k. This is because the average proximity from one agent to other agents is the largest in a single-city pattern (e.g., Figure 5a), while it decreases as the number of peaks in the spatial distribution increases. In particular,

$$\max_k\{\chi_k\} = \chi_1 = \frac{1-\phi}{1+\phi} \quad \text{and} \quad \min_k\{\chi_k\} = \chi_{\frac{N}{2}} = \left(\frac{1-\phi}{1+\phi}\right)^2 \tag{24}$$

for any given value of $\phi \in (0,1)$ (Akamatsu et al., 2012). Recall that the maximum possible number of symmetric cities is $\frac{N}{2}$ (cf. Example 3.2). Also, each χ_k takes value on $(0,1)$ and is a monotonically decreasing function of ϕ, reflecting that agents are less sensitive to the proximity to others when ϕ is larger (and vice versa).

Note that $\{z_k\}$ and $\{\chi_k\}$ are model independent. They encapsulate the properties of the underlying geography (Assumption RE) but not those of the payoff function. The model-dependent properties are instead represented by gain function G of a model.

That said, $\omega_{\max} = \max_k \omega_k = \max_k G(\chi_k)$ depends on the properties of $G(\chi)$. Section 2.3 demonstrated that the shape of gain function G of a model can crucially affect the resulting implications, where the most important distinction is in the spatial scale of the dispersion force in the model. We introduce a formal categorization of canonical models based on three prototypical shapes of G.

Definition 5. A canonical model with gain function G is said to be:

(a) *Class I*, if there is at most one $\chi^* \in (0,1)$ so that $G(\chi) > 0$ for $\chi \in (0,\chi^*)$, $G(\chi^*) = 0$, and $G(\chi) < 0$ for $\chi \in (\chi^*,1)$.

(b) *Class II*, if there is at most one $\chi^{**} \in (0,1)$ so that $G(\chi) < 0$ for $\chi \in (0,\chi^{**})$, $G(\chi^{**}) = 0$, and $G(\chi) > 0$ for $\chi \in (\chi^{**},1)$.

(c) *Class III*, if there are at most two $\chi \in (0,1)$ so that $G(\chi) = 0$, denoted by $\chi^{**} < \chi^*$, with $G(\chi) < 0$ for $\chi \in (0,\chi^{**}) \cup (\chi^*,1)$ and $G(\chi) > 0$ for $\chi \in (\chi^{**},\chi^*)$.

The Krugman and Helpman models are, respectively, Class I and II. The first two model classes in Definition 5 are, respectively, of "Krugman-type" and "Helpman-type." Class III features the combined characteristics of Classes I and II. We provide concrete examples of the three model classes in the following.

Example 3.3 (Class I)**.** Krugman (1991b), Puga (1999), Forslid and Ottaviano (2003), Pflüger (2004), and Harris and Wilson (1978). ■

Example 3.4 (Class II)**.** Helpman (1998), Murata and Thisse (2005), Redding and Sturm (2008), Allen and Arkolakis (2014), Redding and Rossi-Hansberg (2017) (§3), and Beckmann (1976). ■

Example 3.5 (Class III)**.** Tabuchi (1998), Pflüger and Südekum (2008), as well as Takayama and Akamatsu (2011). ■

As mentioned in the Introduction, Definition 5 classifies canonical models based on the spatial scale of the working dispersion force(s) (Table 1). Net gain functions $G^\sharp$ for all models in the above examples are at most *quadratic* (see Table F.1 in Appendix F). That is,

$$G^\sharp(\chi) = c_0 + c_1\chi + c_2\chi^2 \tag{25}$$

with model-dependent coefficients $\{c_0, c_1, c_2\}$, as in the Krugman or Helpman models. When $G^\sharp$ is quadratic and there exists an agglomeration force (a positive term), a model is Class I if and only if there is no local dispersion force ($c_0 \geq 0$) but only a global dispersion force that can stabilize $\bar{x}$. A model is Class II if and only if there exists local dispersion force ($c_0 < 0$) but no working global dispersion force is present (i.e., $G^\sharp(1) > 0$).[22] A model is Class III if and only if there exits both a local ($c_0 < 0$) and a global dispersion force ($c_2 < 0$), as well as an agglomeration force ($c_1 > 0$).

The following proposition characterizes the endogenous spatial patterns that Class I, II, or III models engender when $\bar{x}$ becomes unstable, which is essentially the characterizations based on $\omega_{\max}$ at the point $\bar{x}$ becomes unstable.

Proposition 1. *Suppose Assumptions RE and S. Consider a canonical model of either Class I, II, or III with gain function G. Assume $G(\chi) = 0$ has one root (two roots) in $\chi \in (0,1)$ if the model is Class I or II (Class III), so that endogenous agglomeration occurs in $\phi \in (0,1)$.*

(a) If the model is of Class I, there exists $\phi^* \in (0,1)$ so that $\bar{x}$ is stable for all $\phi \in (0,\phi^*)$ and unstable for all $\phi \in (\phi^*,1)$; the instability of $\bar{x}$ in ϕ^* leads to the formation of $\frac{N}{2}$ small cities.

(b) If the model is of Class II, there exists $\phi^{**} \in (0,1)$ so that $\bar{x}$ is stable for all $\phi \in (\phi^{**},1)$ and unstable for all $\phi \in (0,\phi^{**})$; the instability of $\bar{x}$ in ϕ^{**} leads to the formation of a single city.

(c) If the model is of Class III, there exist $\phi^*, \phi^{**} \in (0,1)$ with $\phi^* < \phi^{**}$ so that $\bar{x}$ is stable for all $\phi \in (0,\phi^*) \cup (\phi^{**},1)$; the instabilities of $\bar{x}$ at ϕ^* and ϕ^{**} lead to the formation of $\frac{N}{2}$ cities and a single city, respectively.

Proof. See Appendix A.1. □

Note that (a) and (b) generalizes the "reversed scenarios" of Krugman and Helpman in that $\bar{x}$ is stable for the low (high) values of ϕ in Class I (II) models. On the other hand, model classes differ in the *number of cities* they endogenously produce. Class I models engender $\frac{N}{2}$ small cities, whereas those of Class II entail a single megacity. Class III is a synthesis of Classes I and II.

Proposition 1 builds on the relationships in (24). This can most clearly be seen in the Beckmann model.

[22]Consider $G^\sharp$ for the Helpman model (15). If $\mu > \frac{\sigma-1}{\sigma}$, its second-order term can be negative but it cannot stabilize $\bar{x}$. If $\mu \leq \frac{\sigma-1}{\sigma}$, the term can may well be positive.

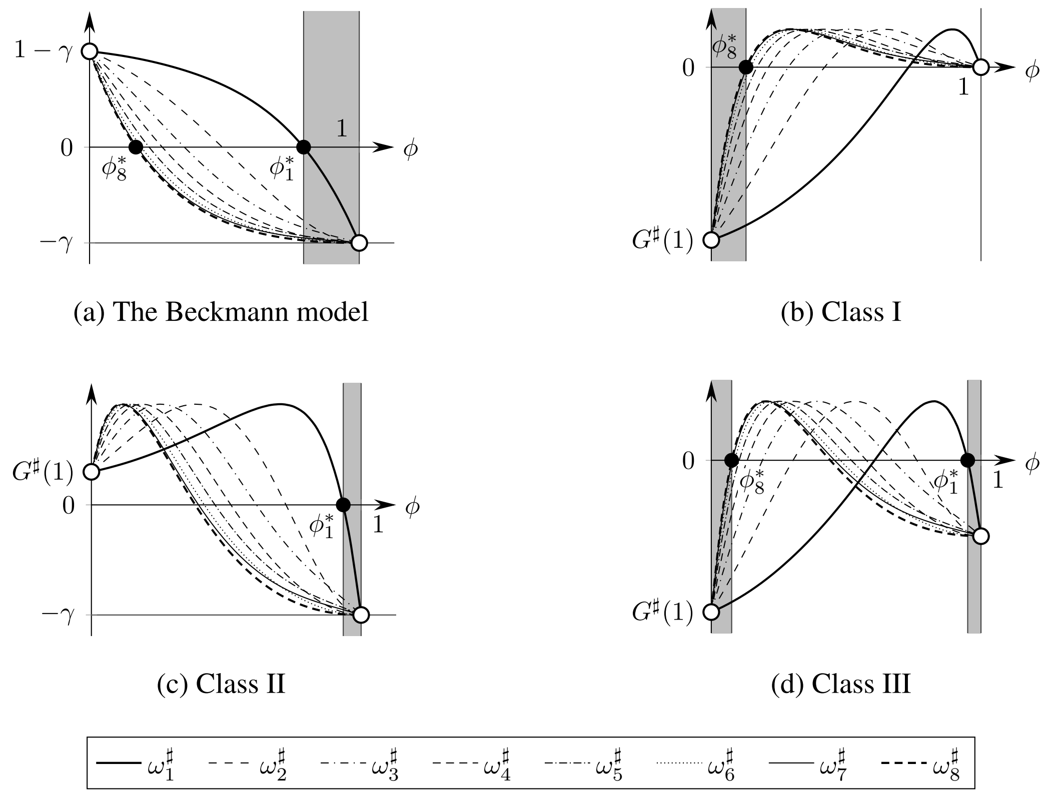


Figure 6: Examples of $\omega_k^\sharp \equiv G^\sharp(\chi_k(\phi))$ $(N = 16)$.

Example 3.6. The Beckmann model (Example 2.1) is Class II because $G(\chi) = G^\sharp(\chi) = -\gamma + \chi$ satisfies Definition 5 (b). Because $\max\{\chi_k\} = \chi_1$, we have $\max\{\omega_k\} = \max\{-\gamma + \chi_k\} = -\gamma + \chi_1$ for all ϕ. Figure 6a shows $\omega_k = \omega_k^\sharp \equiv G^\sharp(\chi_k(\phi))$ for $N = 16$. When all the curves stay below the horizontal axis, $\bar{x}$ is stable (the shaded area). The instability of $\bar{x}$ occurs at $\phi^{**} = \phi_1^*$, leading to the formation of a single megacity (Figure 5a). The maximality of ω_1 can be clearly interpreted. In the model, the formation of a single large city is the most beneficial outcome for every agent because agents prefer proximity to others, albeit agents must disperse around the city center to avoid local congestion effects. ■

Representative examples of general equilibrium models from all three classes are also shown in Figure 6.

Example 3.7. Figure 6b and Figure 6c respectively depict $\omega_k^\sharp$ for the Krugman and Helpman models, as the leading examples of Classes I and II. For all ϕ so that $\bar{x}$ is stable, $\max\{\omega_k^\sharp\} = \omega_{\frac{N}{2}}^\sharp$ in the Krugman model, whereas $\max\{\omega_k^\sharp\} = \omega_1^\sharp$ in the Helpman model. When $\bar{x}$ becomes unstable, $\frac{16}{2} = 8$ cities emerge for the former model, whereas a single city emerges for the latter. Figure 6d

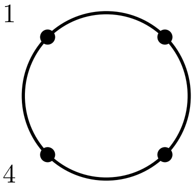


Figure 7: Four-region racetrack economy.

shows $\omega_k^\sharp$ for an instance of Class III, the Pflüger and Südekum (2008) model. Observe there are two ranges of ϕ under which $\bar{x}$ is stable. With both local and global dispersion forces, the model behaves as a Class I (II) model at a low (high) ϕ. ■

4 Illustration: Beyond the reversed scenarios

As a concrete illustration of **Proposition 1**, we reconsider the Krugman and Helpman models studied in Section 2.3. In a many-region world, the "reversed scenario" of the two models is no longer just a reversal of the binary process between symmetry and asymmetry. The Krugman and Helpman models provide minimal examples of the difference between Classes I and II. The difference in the spatial scale of dispersion forces induces an intrinsic contrast not only in *timing* but also in *endogenous spatial patterns*, as shown by **Proposition 1**.

Suppose Assumptions RE and S and let $N = 4$ (see Figure 7). This is the simplest setup in which different regions can have different neighbors. Example 3.1 provides the proximity matrix for this case.[23]

Uniform distribution $\bar{\boldsymbol{x}} = (\bar{x}, \bar{x}, \bar{x}, \bar{x})$ with $\bar{x} \equiv \frac{1}{4}$ is a spatial equilibrium. As discussed, $\bar{\boldsymbol{x}}$ is stable if all the eigenvalues of $\mathbf{V} = \frac{\bar{x}}{\bar{v}}\nabla v(\bar{\boldsymbol{x}})$ are negative. There are two $(= \frac{N}{2} = \frac{4}{2})$ eigenvalues of interest, which we denote by ω_1 and ω_2. Associated with them, there are two "basic" migration

[23] We may assume that the proximity between two regions on the antipodal points is $\phi' \in (0, 1)$ (with a natural restriction $\phi' < \phi$). It is inconsequential.

(a) A single megacity ($\bar{x} + \zeta z_1$)

(b) Two small cities ($\bar{x} + \zeta z_2$)

Figure 8: Schematic illustrations of spatial patterns

Notes: As region 1 is neighboring region 4 on the circle, we rotationally shift the spatial distributions for better understandability.

patterns:[24]

$$z_1 = \frac{1}{\sqrt{2}}(1, 0, -1, 0) \quad \text{and} \quad z_2 = \frac{1}{2}(1, -1, 1, -1). \tag{26}$$

Note that z_1 and z_2 do not include any parameters and thus are model independent.

Figure 8 shows the schematics of the two possible outcomes from $\bar{x}$. The two spatial configurations have distinct characteristics: one represents the formation of a *single megacity* that attracts all the population in the economy (Figure 8a), whereas the other represents the emergence of *two small cities* vying with each other (Figure 8b).

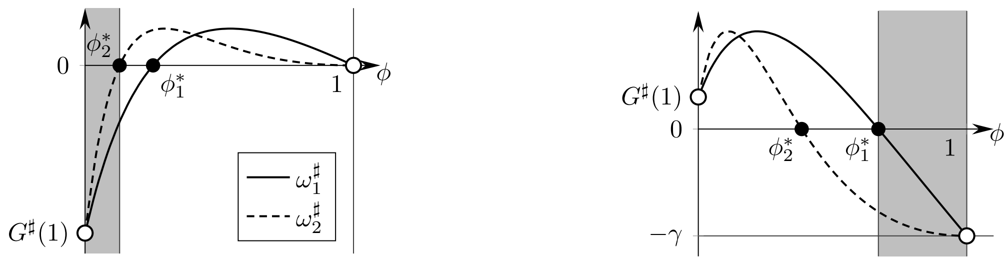
The question is *which* of the two patterns emerge endogenously in the Krugman and Helpman models—a single megacity or two small cities? **Proposition 1** (a) and (b) respectively show that the Krugman model produces two small cities and that the Helpman model produces a single city.

To show this, we ask whether $\omega_{\max} = \omega_1$ or $\omega_{\max} = \omega_2$ when $\bar{x}$ becomes unstable, as we recall each ω_k represents the net agglomerative force towards basic migration pattern z_k. In fact, by noting $\omega_k z_k = \mathbf{V} z_k$, we can show that

$$\omega_1 = G(\chi_1) = \frac{\bar{x}}{\bar{v}} \left(\frac{\partial v_1(\bar{x})}{\partial x_1} - \frac{\partial v_3(\bar{x})}{\partial x_1} \right), \tag{27}$$

$$\omega_2 = G(\chi_2) = \frac{\bar{x}}{\bar{v}} \left(\frac{\partial v_1(\bar{x})}{\partial x_1} - \frac{\partial v_2(\bar{x})}{\partial x_1} + \frac{\partial v_3(\bar{x})}{\partial x_3} - \frac{\partial v_4(\bar{x})}{\partial x_3} \right). \tag{28}$$

[24]Since $\mathbf{V}$ is 4×4, there exist two eigenvectors other than z_1 and z_2: one is uniform vector $\mathbf{1} = \frac{1}{2}(1, 1, 1, 1)$, in keeping with $N = 2$, and the other is $\frac{1}{\sqrt{2}}(0, 1, 0, -1)$, which has the same meaning as z_1 because of rotational symmetry. In fact, its associated eigenvalue is ω_1.



(a) The Krugman model

(b) The Helpman model

Figure 9: $\omega_1^\sharp$ and $\omega_2^\sharp$ for the Krugman and Helpman models.

Similar to (3), (27) indicates that if $\omega_1 > 0$, agents have an incentive to form a monocentric spatial pattern by migrating from region 3 to 1. Similarly, (28) indicates that agents may migrate to form a two-city pattern if $\omega_2 > 0$.

Figure 9 provides the answer to the question. For each model, it depicts $\omega_1^\sharp \equiv G^\sharp(\chi_1(\phi))$ and $\omega_2^\sharp \equiv G^\sharp(\chi_2(\phi))$ on the ϕ-axis, where $G^\sharp(\chi)$ is the same as in the two-region case for each model in (14) and (15) (Figure 2). As $\text{sgn}[\omega_k] = \text{sgn}[\omega_k^\sharp]$, $\bar{\boldsymbol{x}}$ is stable if the two curves stay below the horizontal axis (the shaded areas). That is,

$$\bar{\boldsymbol{x}} \text{ is stable when } \begin{cases} \phi \in (0, \phi_2^*) & \text{(the Krugman model)}, \\ \phi \in (\phi_1^*, 1) & \text{(the Helpman model)}. \end{cases}$$

There is an analogy with the "reversed scenario" regarding *when* $\bar{\boldsymbol{x}}$ is stable.

A sharp contrast is present in the *spatial patterns*. It is immediate that $\omega_{\max} = \omega_2 > \omega_1$ at ϕ_2^* for the Krugman model because $\omega_2(\phi_2^*) = 0$ and $\omega_1(\phi_2^*) < 0$. Similarly, $\omega_{\max} = \omega_1$ at ϕ_1^* for the Helpman model. Therefore, the spatial pattern that emerges from $\bar{\boldsymbol{x}}$ is

$$\begin{cases} \text{the two-city pattern (Figure 8b)} & \text{(the Krugman model)}, \\ \text{the single-megacity pattern (Figure 8a)} & \text{(the Helpman model)}, \end{cases}$$

as shown by **Proposition 1** (a) and (b).

The difference in the spatial scale of dispersion forces is the source of the contrast in the engendered spatial patterns. As discussed in Section 2.4, the dispersion force in the Helpman model is local and triggered when ϕ is high. Consider the process of a monotonic increase in trade freeness ϕ. When ϕ is at its lower extreme ($\phi \approx 0$), agents concentrate in a single region because the local dispersion force is less important than the benefits of agglomeration when interregional transportation is prohibitively costly.[25] The spatial pattern is close to a completely monopolar pattern, for example, $\boldsymbol{x} \approx (0, 1, 0, 0)$. As ϕ increases, the relative rise in the local dispersion force induces a crowding-out from the populated region. The spatial pattern become, for example $\boldsymbol{x} = (x', x, x', x'')$ with $x > x' > x''$, which can also be regarded as a monopolar pattern. As ϕ increases, the spatial pattern gradually flattens and, at threshold ϕ_1^*, it must connect to uniformity $\bar{\boldsymbol{x}}$. If we start from $\bar{\boldsymbol{x}}$ and gradually *decrease* ϕ to determine the dispersion process in a reverse-reproduced way, at ϕ_1^*, the spatial pattern must deviate in the direction of the "formation" of a single megacity (Figure 8a).

By contrast, the dispersion force in the Krugman model is global and triggered when ϕ is low. Recall that the dispersion force stems from firms' competition over consumers. When ϕ is low, there are few incentives for firms to concentrate on a small number of regions because the shipment of goods incurs large transportation costs. As ϕ increases, the size of the effective market area of a firm extends. For each firm, this brings more opportunities to access a wider range of consumers but also leads to tougher competition with other firms that are geographically close. At some point, firms are better off forming small cities so that each has its dominant market area but is relatively remote from other major concentrations of firms, as in the two-city pattern (Figure 8b).

To summarize, **Proposition 1** is the consequence of the difference in the spatial scale of dispersion forces. The global dispersion force represents the repulsive effects across different locations and supports the formation of multiple cities, whereas the local dispersion force represents the crowding effects that induce the flattening of each city. In a many-region economy, these forces lead to the formation of qualitatively different spatial patterns. This is most clearly seen by comparing the Krugman and Helpman models in the $N = 4$ case, which are the leading instances of Classes I and II, respectively. The contrast in spatial patterns is hidden in the two-region setup, as the only possible migration pattern is $\boldsymbol{z} = (1, -1)$.

[25]Note that mobile agents prefer concentrating towards a smaller number of regions because of the agglomeration force. In both models, agents should result in a "black-hole" concentration in a single region if there is no effective dispersion force.

There are several remarks on **Proposition 1**. First, it builds on *local analysis* around uniform distribution $\bar{x}$. It may be of interest whether we can formally *prove* that the local profitable deviation at the onset of instability is actually the ultimate spatial equilibrium the agglomeration force converges towards. The technically accurate answer is: "not always." To draw stronger conclusions beyond **Proposition 1**, we have to either introduce intricate classifications for the properties of the higher-order differentials of the payoff function v or focus on a specific model.[26] However, **Proposition 1** provides essential practical insights into the evolution of the spatial pattern in a circular economy.[27] To highlight this point, Section 5 will present a series of numerical examples for when $N = 8$.

Second, **Proposition 1** assumes a *complete geographical symmetry*. Assumptions RE and S abstract away regional heterogeneities and geographical advantages. It is thus of interest to what extent or in what sense the implications of **Proposition 1** generalize to asymmetric cases, given that the latest quantitative spatial models incorporate flexible structures regarding interregional transportation costs and differences in local characteristics. To address this issue, Section 6 provides formal analyses of the effects of heterogeneous local characteristics. We also include in Appendix D discussions on other geographical setups and provides numerical explorations for exogenous geographical advantages due to the existence of boundaries.

Appendix C also provides a brief discussion on the effects of idiosyncratic payoff shocks (Remark C.2) and on the forward-looking behaviors of agents (Remark C.3).

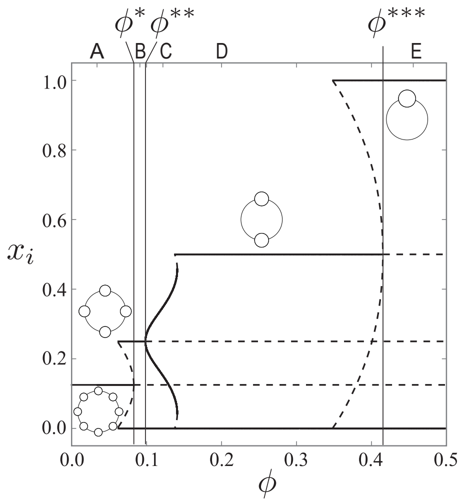
5 Evolution of spatial structure

We numerically explore an overall evolutionary path of the spatial structure for selected models from Classes I, II, and III in the $N = 8$ racetrack economy.[28] We will see that **Proposition 1** captures the intrinsic properties of the whole evolutionary process.

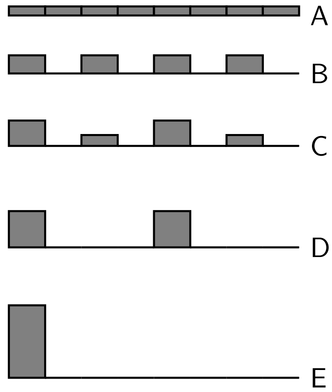
[26]Generally, this line of research converges to *bifurcation theory*. See Hirsch et al. (2012) and Kuznetsov (2004) for concise introductions. Additionally, under Assumptions RE and S, *equivariant bifurcation theory* allows one to draw various technical conclusions beyond **Proposition 1**. See Golubitsky and Stewart (2003) for an introduction.

[27]See, for instance, previous studies by (Akamatsu et al., 2012; Ikeda et al., 2012; Osawa et al., 2017; Ikeda et al., 2018).

[28]The formulations of the models and the parameter settings are shown in Appendix F.



(a) Bifurcation diagram



(b) Spatial patterns

Figure 10: Class I model (Krugman, 1991a)

Figure 10 reports the evolutionary path of stable equilibrium patterns in the course of increasing ϕ for the Krugman model, which is Class I. In Figure 10a, the black solid (dashed) curves depict the stable (unstable) equilibrium values of x_i at each ϕ. Figure 10b shows the schematic illustration of the stable spatial pattern on the path. The letters in Figure 10b correspond to those in Figure 10a.

Consider a gradual increase in ϕ from $\phi \approx 0$. Uniform distribution $\bar{x}$ is initially stable until ϕ reaches the so-called "break point" ϕ^* where a *bifurcation* from $\bar{x}$ occurs. At ϕ^*, the spatial pattern is pushed towards the formation of $\frac{8}{2} = 4$ cities. This confirms **Proposition 1** (a). The spatial pattern immediately converges towards a four-cities pattern after ϕ^* is passed. The number of populated regions halves from $8 \to 4$.

A further increase in ϕ triggers the second and third bifurcations at ϕ^{**} and ϕ^{***}, respectively. These bifurcations sequentially double the spacing between cities, each time halving their number, $4 \to 2 \to 1$, in a close analogy to the first bifurcation at ϕ^*.[29] At the higher extreme of ϕ, a complete monopolar pattern emerges. This behavior can be understood as a gradual increase in the effective market area of each city due to a decline in transportation costs. The spatial extent of each city is one regional unit at any level of ϕ because there exists no local dispersion force.

In the model, cities become larger when interregional access improves. However, such an effect is limited to the "selected" regions. The impact of an improvement in transportation on the size

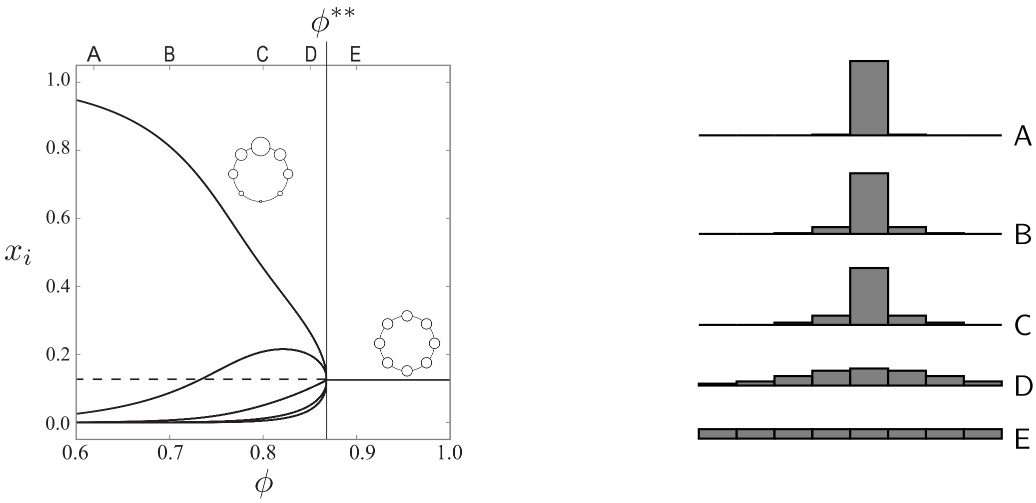
[29]This is the "spatial period-doubling cascade" behavior discussed by (Akamatsu et al., 2012; Osawa et al., 2017; Ikeda et al., 2018).

of *each* city can be either positive (for the selected ones) or negative (for the others). This point is already apparent in the two-region models that explicitly incorporate agglomeration economies combined with interregional transportation costs.

The many-region setup highlights another phenomenon. As ϕ increases, once selected regions can decline to form an *agglomeration shadow* (Arthur, 1994; Fujita and Krugman, 1995) of other regions. For example, consider the fifth region from the left in Figure 10b. This region is selected at the transitions at ϕ^* and ϕ^{**}, that is, the impact of an increase in ϕ is positive. However, after ϕ^{***} is encountered, it immediately loses its population. For the region, a monotonic increase in ϕ implies a winning situation followed by a losing one. The global dispersion force in Class I models is thus related to the *rise and fall of major cities*. Class I models do not provide robust predictions for each city, but they do for the *overall* spatial distribution of cities: the number of cities and spacing between them monotonically decreases and increases, respectively, with the monotonic reduction in interregional transportation costs.

Remark 5.1. The empirical evidence on regional agglomeration presented by Duranton and Turner (2012) and Faber (2014) is related to the theoretical predictions of Class I models. The former study focused on the growth of large metropolitan areas in the United States, while the latter analyzed the growth of peripheral counties in China. The former (latter) study revealed a positive (negative) correlation between the magnitude of growth and the interregional transportation infrastructure level of a given region. For Class I models, these opposite responses may simply reflect different sides of the same coin. That is, both results may indicate a tendency of selective concentration towards larger regions for an improvement in interregional transportation access (as discussed in the Introduction for Japan). ■

Next, Figure 11 shows the results for a Class II model, namely, the Allen and Arkolakis (2014). This model incorporates a local dispersion force, And, thus, $\bar{x}$ is stable for higher values of ϕ. As in Section 4, we see the evolutionary process in a reverse-reproduced way, that is, in the course of a monotonic *decrease* in ϕ. The bifurcation at ϕ^{**} leads to the "emergence" of a unimodal pattern. This is *the* bifurcation in the model: when ϕ decreases further, the spatial pattern monotonically and smoothly converges to a complete concentration in a single region. We interpret a region that



(a) Bifurcation diagram

(b) Spatial patterns

Figure 11: Class II model (Allen and Arkolakis, 2014)

locally maximizes population size (region i such that $x_i > x_{i-1}$ and $x_i > x_{i+1}$ where mod N for indices) as the location of a city. Then, this model endogenously produces at most one city. Class II models would be interpreted as expressing the evolution of the *spatial extent of a single city*, namely the flattening of a big city during improvement in interregional transportation access.

Remark 5.2. Class II models have an attractive property for quantitative applications. We can ensure the uniqueness of the spatial equilibrium *regardless of the level of interregional transportation costs* by imposing a strong local dispersion force (e.g., Redding and Sturm, 2008; Allen and Arkolakis, 2014). If the equilibrium is unique, then the calibrations and counterfactual analyses have determinate implications. Example 2.1 provides a prototypical situation in which a strong congestion force suppress the possibility of endogenous asymmetry (i.e., the $\gamma \geq 1$ case). The uniqueness of the equilibrium implies the stability of $\bar{x}$ for all ϕ in a racetrack economy, that is, no endogenous asymmetry can emerge, since $\bar{x}$ is always an equilibrium. Figure 12 provides our classification of possible spatial patterns and their stabilities for the Allen–Arkolakis model in a racetrack economy. The uniqueness condition for the Allen–Arkolakis model is $\beta \leq -\alpha$ (i.e., Range 3 in the figure), which is a sufficient condition for the stability of $\bar{x}$. See also Figure I of Allen and Arkolakis (2014) in comparison with Figure 12. ■

Finally, we consider a Class III model. Because both local and global dispersion forces exist, this class of models exhibits a rich and realistic interplay between the *number of cities and spacing*

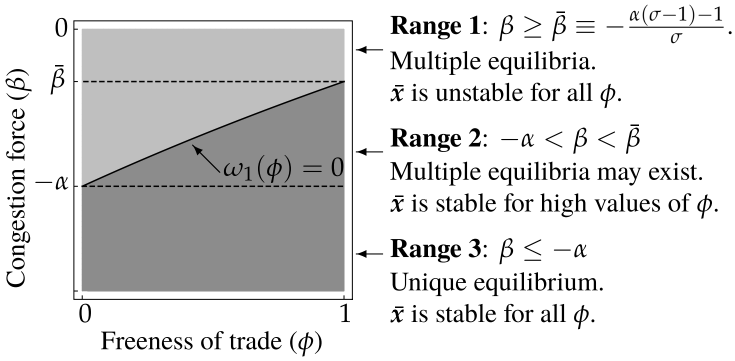
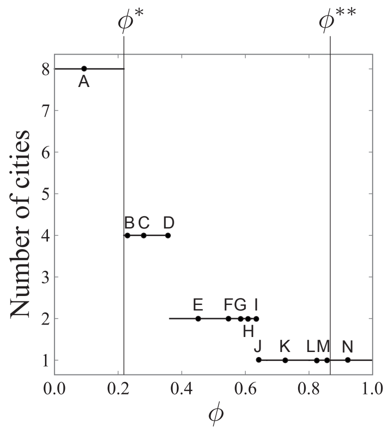


Figure 12: Uniqueness and stability of equilibria in the Allen–Arkolakis model

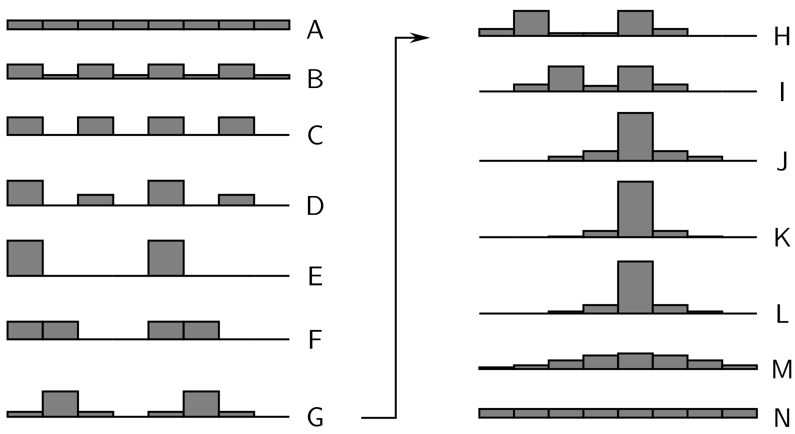
between them (as in Class I models) and the *spatial extent of each city* (as in Class II models).

Figure 13a shows the evolution of the number of cities in the course of increasing ϕ under the Pflüger and Südekum (2008)'s model in the $N = 8$ racetrack economy. The number of cities in a spatial distribution is defined by that of the local maxima therein. Figure 13a exhibits the mixed characteristics of Figures 10 and 11, as expected. Uniform distribution $\bar{x}$ is stable for $\phi < \phi^*$ or $\phi > \phi^{**}$. We interpret the number of cities in $\bar{x}$ as either 8 (for a low ϕ) or as 1 (for a high ϕ) to acknowledge that $\bar{x}$ at the low and high levels of ϕ are distinct. When ϕ gradually increases from $\phi \approx 0$, the number of cities reduces from $8 \to 4 \to 2 \to 1$ as in the Class I models (Figure 10), whereas it is always 1 in the latter stage as per the Class II models (Figure 11). The initial stage is governed by a decline in the global dispersion force, while the later stage is marked by a relative increase in the local dispersion force.

Figure 13b illustrates the spatial patterns associated with Figure 13a. Uniform pattern $\bar{x}$ is initially stable (Pattern A) and the first bifurcation at ϕ^* leads to a four-city pattern (B, C), whereas the second bifurcation to the formation of two cities (D, E). These transitions are in line with Figure 10 and are governed by the gradual decline in the global dispersion force. Then, the evolutionary behavior becomes more interesting: the decline in the global dispersion force increases the relative importance of the local one. As a result, the two cities in Pattern E gradually increase their spatial extents (F, G) because of the local dispersion effects. A further increase in ϕ means the local dispersion force succeeds and the two cities gradually merge (H, I) to form a megalopolis (J, K). As the relative importance of the local dispersion force further increases, a gradual flattening of the single megalopolis occurs (L, M), followed by complete dispersion (N) after ϕ^{**} is reached.



(a) Number of cities



(b) Spatial patterns

Figure 13: Class III model (Pflüger and Südekum, 2008)

Remark 5.3. Regarding the behavior of Class II and III models, there is ample empirical evidence for the flattening of once established economic clusters (i.e., cities) as a consequence of improved *interregional* access. Baum-Snow (2007) and Baum-Snow et al. (2017) presented evidence for US metro areas during 1950–1990 and Chinese prefectures during 1990–2010, respectively. These studies addressed the changes in the population or production size of the central area within the larger region, both reporting a significantly negative effect of improvements in interregional access. As discussed in these studies, the local flattening of cities can also be interpreted as suburbanization in response to the improved *intra-urban* transportation infrastructure in classical urban economic theory (e.g., Alonso, 1964). ■

6 Exogenous local characteristics

Proposition 1 builds on a complete geographical symmetry. However, exogenous asymmetries are inherent in real-world geography. As a relaxation of Assumption S, we study the sensitivity of spatial patterns to regional characteristics, for example, local amenities and productivity differences. This section shows that the spatial scale of dispersion force(s) in a model tends to determine whether the effects of any exogenous advantages are amplified (when transportation cost varies).[30] Throughout, we assume Assumption RE.

Let $\boldsymbol{a} = (a_i)_{i \in \mathcal{N}}$ with $a_i > 0$ indicate some regional characteristic, which may or may not affect the payoffs in other regions. For instance, a_i may be the level of amenities in region i exclusively enjoyed by the residents therein or the total factor productivity of the region. In the latter case, interregional trade flows and the resulting payoff levels in other regions can depend on a_i.

The regions are symmetric if $\boldsymbol{a} = \bar{\boldsymbol{a}} \equiv (\bar{a}, \bar{a}, \ldots, \bar{a})$, for some $\bar{a} > 0$. Therefore, pair $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{a}})$ is an equilibrium. Consider a variation in the local characteristic so that $\boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{\epsilon}$ with small $\boldsymbol{\epsilon} = (\epsilon_i)_{i \in \mathcal{N}}$. Then, there is a new equilibrium, say $\boldsymbol{x}(\boldsymbol{a})$, which is close to $\bar{\boldsymbol{x}}$. The "covariance" between region i's relative (dis)advantage $\epsilon_i = a_i - \bar{a}$ and the relative deviation of its population $x_i(\boldsymbol{a}) - \bar{x}$ is then evaluated by:

$$\rho \equiv (\boldsymbol{a} - \bar{\boldsymbol{a}})^\top (\boldsymbol{x}(\boldsymbol{a}) - \bar{\boldsymbol{x}}) = \sum_{i \in \mathcal{N}} (a_i - \bar{a}) (x_i(\boldsymbol{a}) - \bar{x}). \tag{29}$$

We here assume that $\bar{\boldsymbol{x}}$ is stable, since otherwise considering $\boldsymbol{x}(\boldsymbol{a})$ is nonsensical.

We expect $\rho > 0$ if regional characteristic $\boldsymbol{a}$ acts positively in the payoff for agents, that is, if $\boldsymbol{a}$ is in fact "advantageous." To formalize this intuition, we focus on a formulation class of local characteristics, $\boldsymbol{a}$, which encompass various standard specifications in the literature. Let $\mathbf{A} \equiv \frac{\bar{a}}{\bar{v}} [\frac{\partial v_i}{\partial a_j}]$ be the elasticity matrix of the payoff regarding $\boldsymbol{a}$, evaluated at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{a}})$. Analogous to our definition of canonical models in Definition 1, we suppose the following.

Assumption A. For the local characteristic $\boldsymbol{a}$ under consideration, there exists a rational function $G^\natural$ that is continuous over $(0, 1)$, positive whenever $\bar{\boldsymbol{x}}$ is stable, and satisfies $\mathbf{A} = G^\natural(\bar{\mathbf{D}})$.

[30] Apart from the sensitivity analysis, a general fact is that small perturbations in local factors do not qualitatively alter the predictions of **Proposition 1**. See Remark C.7 for a brief discussion.

For each model in Examples 3.3, 3.4, and 3.5, there exists in fact a function $G^\natural$ that satisfies the hypotheses of Assumption A for each natural choice of a local characteristic vector. The simplest example is that of *heterogeneous local amenity*, as considered by Allen and Arkolakis (2014).

Example 6.1. Assume that the payoff function takes the form $v_i(\boldsymbol{x}, \boldsymbol{a}) = a_i v_i(\boldsymbol{x})$, where $a_i > 0$ is the exogenous level of regional amenities and $\boldsymbol{v}(\boldsymbol{x}) = (v_i(\boldsymbol{x}))_{i \in \mathcal{N}}$ represents the symmetric (or homogeneous) component of the payoff function that satisfies Assumption S. Then, $\mathbf{A} = \frac{\bar{a}}{\bar{v}} \bar{v} \mathbf{I} = \bar{a} \mathbf{I}$ and, thus, $G^\natural(\chi) = \bar{a} > 0$. ■

Analogous to gain function G of a model, $G^\natural$ encodes the effect of the marginal changes in local characteristics $\boldsymbol{a}$ on the regional payoffs $\boldsymbol{v}$. Particularly, condition $G^\natural(\chi) > 0$ implies $\rho > 0$. As an example, consider a symmetric two-region economy as in Section 2.2. Then, we can show that

$$\rho = c\delta(\chi) \quad \text{where} \quad \delta(\chi) \equiv -\frac{G^\natural(\chi)}{G(\chi)} \tag{30}$$

with some $c > 0$ and $\chi = \frac{1-\phi}{1+\phi}$. Recall that $G(\chi) < 0$ if $\bar{\boldsymbol{x}}$ is stable. Therefore, $\rho > 0$ if $G^\natural(\chi) > 0$ for all χ so that $G(\chi) < 0$.

To understand formula (30), in line with ω in Section 2.2, we evaluate the payoff gain due to exogenous advantage in region 1 by the following elasticity:

$$\alpha \equiv \frac{\bar{a}}{\bar{v}} \left(\frac{\partial v_1(\bar{\boldsymbol{x}}, \bar{\boldsymbol{a}})}{\partial a_1} - \frac{\partial v_2(\bar{\boldsymbol{x}}, \bar{\boldsymbol{a}})}{\partial a_1} \right). \tag{31}$$

α is an eigenvalue of $\mathbf{A}$ associated with $\boldsymbol{z} = (1, -1)$. Particularly, when $\mathbf{A} = G^\natural(\bar{\mathbf{D}})$, we have

$$\alpha = G^\natural(\chi), \tag{32}$$

in close analogy with relationship $\omega = G(\chi)$. Note that $\alpha > 0$ by assumption.

Suppose that $\bar{\boldsymbol{x}} = (\bar{x}, \bar{x})$ is perturbed to $\boldsymbol{x} = (\bar{x} + \xi, \bar{x} - \xi)$ due to an exogenous regional asymmetry of the form $\boldsymbol{a} = (\bar{a} + \epsilon, \bar{a} - \epsilon)$ with some scalars ζ and ϵ. Then, $v_1(\boldsymbol{x}) = v_2(\boldsymbol{x})$ must hold true for $\boldsymbol{x}$ to be an equilibrium. Thus, the pair of deviations ξ and ϵ should cancel out two forces, namely, gain (or *loss*, since we assume $\bar{\boldsymbol{x}}$ is stable) $\omega < 0$ from endogenous migration and gain $\alpha > 0$ from exogenous asymmetry. That is,

$$\omega\xi + \alpha\epsilon = G(\chi)\xi + G^{\natural}(\chi)\epsilon = 0. \tag{33}$$

Then, because $\rho = \epsilon\xi + (-\epsilon)(-\xi) = 2\epsilon\xi$ by definition, formula (30) follows with $c = 2\epsilon^2$. The fraction $\delta(\chi) \equiv -\frac{G^{\natural}(\chi)}{G(\chi)} = \frac{G^{\natural}(\chi)}{|G(\chi)|} = \frac{\alpha}{|\omega|}$ thus compares the magnitudes of gain from marginal exogenous advantage and of loss from marginal endogenous migration, under the condition that the economy stays in equilibrium.

An important question from defining ρ is: *does ρ increase or decrease when ϕ increases?* In other words, does improved transportation access strengthen (weaken) the role of local characteristics and what are the responses of the spatial distribution of economic activities to an improvement in interregional access if $\boldsymbol{a}$ is fixed? These are the questions asked in counterfactual exercises employing calibrated quantitative spatial economic models (see, e.g., Redding and Rossi-Hansberg, 2017).

We have a general characterization for the response of ρ when ϕ varies.

Proposition 2. *Suppose Assumption RE. Consider a canonical model with gain function G. Consider local characteristic $\boldsymbol{a}$ that satisfies Assumption A with some $G^{\natural}$. Assume that $\bar{\boldsymbol{x}}$ is stable and define $\delta(\chi) = -\frac{G^{\natural}(\chi)}{G(\chi)}$. Then, the following hold true for ρ in (29):*

(a) $\rho'(\phi) > 0$, *if* $\delta'(\chi) < 0$ *for all* $\chi \in (0,1)$ *such that* $G(\chi) < 0$.

(b) $\rho'(\phi) < 0$, *if* $\delta'(\chi) > 0$ *for all* $\chi \in (0,1)$ *such that* $G(\chi) < 0$.

Proof. See Appendix A.2. □

Thus, the impacts of improved interregional access are inherently model dependent. An important observation is that *model class matters*—the response of ρ to a given model may be inferred by the spatial scale of the dispersion force in that model.

We provide examples employing common specifications of local characteristics. Below, all derivations are collected in Appendix F. Consider the simplest specification, that is, heterogeneous local amenity (Example 6.1). For this case, we observe a clear contrast between the Krugman and Helpman models, the leading instances of Classes I and II.

Example 6.2. We continue with Example 6.1, where $G^{\natural}(\chi) = \bar{a} > 0$. Then, $\operatorname{sgn}[\delta'(\chi)] = \operatorname{sgn}[\frac{\bar{a}G'(\chi)}{G(\chi)^2}] = \operatorname{sgn}[G'(\chi)]$. The Krugman model (Example 2.2) satisfies $G'(\chi) < 0$ and thus $\rho'(\phi) > 0$. On the other hand, the Helpman model (Example 2.3) satisfies $G'(\chi) > 0$ and thus $\rho'(\phi) < 0$ whenever the equilibrium is unique (as assumed in quantitative applications). ■

To determine why such contrast emerges, consider a two-region setup for which $\rho(\phi) = c\delta(\chi(\phi))$ from (30). Recall that $\delta(\chi) = \frac{G^{\natural}(\chi)}{|G(\chi)|}$ is the relative magnitude of the gains from an exogenous advantage against loss from migration at $\bar{\boldsymbol{x}}$. Assume that $\delta'(\chi) < 0$. Then, either gain $\alpha = G^{\natural}(\chi)$ from the exogenous regional asymmetry decreases in $\chi = \frac{1-\phi}{1+\phi}$ (i.e., increases in ϕ) or the magnitude of loss $|\omega| = |G(\chi)|$ from the endogenous migration increases in χ (i.e., decreases in ϕ). Recall that Class I models exhibit endogenous agglomeration due to the decline of the global dispersion force in the course of *increasing* ϕ (see Section 2.4). In other words, in Class I models, $|\omega|$ decreases in ϕ (i.e., increases in χ) as long as $\bar{\boldsymbol{x}}$ is stable. Therefore, if $G^{\natural}(\chi)$ is a constant, as in Example 6.2, we expect $\delta'(\chi) < 0$ in Class I models because δ is inversely proportional to $|\omega|$. A similar discussion applies to $\delta'(\chi) > 0$, and we expect that Class II models satisfy $\delta'(\chi) > 0$.

The contrast between Classes I and II generalizes to the regional characteristics that affect *interregional* trade flows, rather than purely local factors. For such cases, $G^{\natural}(\chi)$ become non-constant.

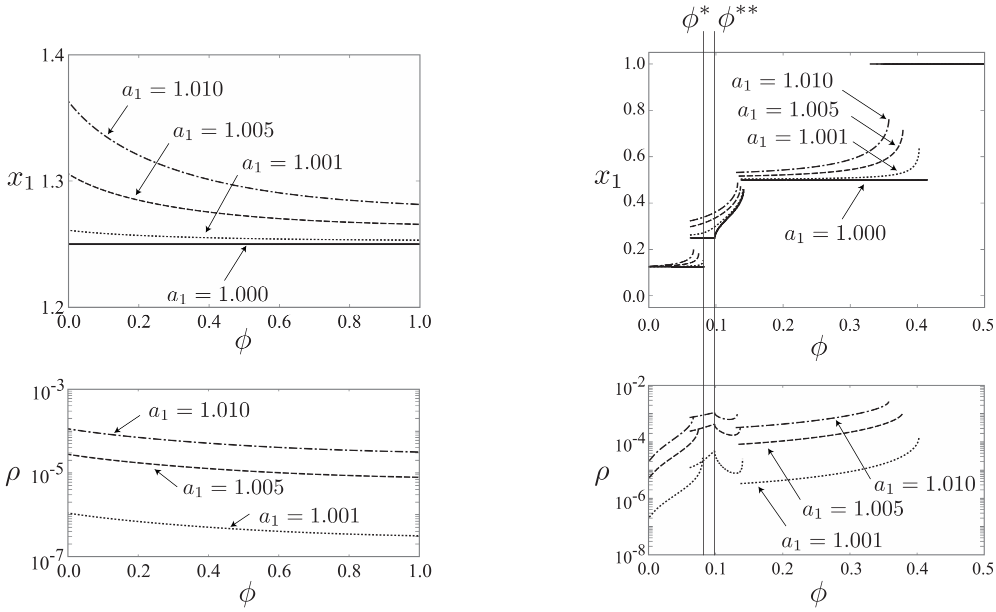
Example 6.3. Redding and Rossi-Hansberg (2017), §3, considered a Class II model, namely the Helpman model with a modified market equilibrium condition:

$$w_i x_i = \sum_{j \in \mathcal{N}} \frac{x_i a_i w_i^{1-\sigma} \phi_{ij}}{\sum_{k \in \mathcal{N}} x_k a_k w_k^{1-\sigma} \phi_{kj}} e_j \qquad \forall i \in \mathcal{N}, \tag{34}$$

where heterogeneities in $\boldsymbol{a}$ arises from *local productivity differences*. We can show that $G^{\natural}(\chi) > 0$ for all $\chi \in (0,1)$. When the equilibrium is unique, $\delta'(\chi) > 0$ and thus $\rho'(\phi) < 0$. ■

Example 6.4. The Krugman model (Example 2.2) is a Class I model. When one interprets immobile demand $\boldsymbol{l} = (l_i)_{i \in \mathcal{N}}$ in the Krugman model as regional characteristic, we have $G^{\natural}(\chi) > 0$ for all $\chi \in (0,1)$. Further, $\delta'(\chi) < 0$, so that $\rho'(\phi) > 0$ whenever $\bar{\boldsymbol{x}}$ is stable. ■

Examples 6.2, 6.3, and 6.4 demonstrate that the class a model belongs to can govern whether



(a) The Allen–Arkolakis model (b) The Krugman model

Figure 14: Population share of region 1 and covariance ρ

the endogenous causation of the model boosts the exogenous advantages when interregional transportation costs decrease. When interregional access improves, the endogenous mechanisms of a model strengthens (weakens) the effects of exogenous local advantages if the model has only a global (local) dispersion force. If exogenous heterogeneity causes one region to attract more population, such effects will be magnified (reduced) for Class I (II) models.

The qualitative differences between Classes I and II can be understood from the basic properties of the local and global dispersion forces in Section 2.4. For a Class I model, a larger ϕ means a relatively smaller global dispersion force, which tends to amplify (both the endogenous and exogenous) location-specific advantages towards the concentration of mobile agents. However, in a Class II model, a larger ϕ means a relatively larger local dispersion force, which reduces not only the benefit from concentration due to endogenous agglomeration externalities but also that due to location-specific exogenous advantages.

Figure 14 reports numerical examples for the $N = 8$ case. As in Example 6.1, we multiply positive term $a_1 \geq 1$ by the payoff in region 1, whereas we let $a_i = 1$ for all $i \neq 1$. The curves depicts region 1's population share, x_1, at stable equilibria against ϕ. We consider four incremental settings, $a_1 \in \{1.000, 1.001, 1.005, 1.010\}$, including the baseline case with no location-fixed advantage

($a_1 = 1.000$). Figure 14a reports the evolutionary paths of x_1 for the Allen–Arkolakis model (Class II) under the uniqueness of the equilibrium. We have $\delta'(\chi) > 0$ for all $\chi \in (0,1)$ and see that $x_1 - \bar{x} > 0$ when $a_1 > 1$ and $x_1 - \bar{x}$ increases as a_1 increases, which are intuitive. Additionally, $x_1 - \bar{x}$ decreases as ϕ increases. We confirm that $\rho(\phi) > 0$ and $\rho'(\phi) < 0$ for all ϕ. Even when the equilibrium can be nonunique in the Allen–Arkolakis model, we still have $\rho(\phi) > 0$ and $\rho'(\phi) < 0$ for all ϕ so that $\bar{x}$ is stable, as **Proposition** 2 predicts.[31] Figure 14b considers the Krugman model (Class I), where the model parameters are the same as in Figure 10.[32] For all $a_1 > 1$, $x_1 - \bar{x} > 0$. As **Proposition** 2 predicts, for all ϕ so that $\bar{x}$ is stable, that is, $\phi \in (0, \phi^*)$, we confirm $\rho(\phi) > 0$ and $\rho'(\phi) > 0$. Although **Proposition** 2 does not cover $\phi \in (\phi^*, 1)$, $\rho'(\phi) > 0$ holds true except for the transitional phase after ϕ^{**}.

7 Concluding remarks

In this paper, we introduce the dichotomy between "local" and "global" dispersion forces under a general framework that encompasses a wide range of extant economic geography models. We show that the spatial scale of the dispersion force in a model significantly affects the endogenous spatial patterns and comparative statics of that model. Three prototypical model classes are defined according to the spatial scale(s) of their dispersion forces (i.e., local, global, and local and global). Given the knowledge of the spatial scale of dispersion forces, we provide consistent interpretations to the empirical literature and provide qualitative characterizations of the comparative statics of structural economic geography models. We also hope our results and methods can be extended to achieve a unified understanding of the robust properties of a broader class of economic geography models.

There are two major directions for further research. First, the generalization of the theoretical results to asymmetric proximity structures is of importance. An efficient strategy would be to fix a few representative models—instead of geography—as test pilots and identify general insights

[31] **Proposition** 2 does not cover the case when $\bar{x}$ is unstable. Accordingly, $\rho'(\phi) < 0$ does not necessarily holds true when $\bar{x}$ is unstable. See Remark C.8 in Appendix C.

[32] Unlike in the Allen–Arkolakis model, the Krugman model admits multiple equilibria for some ϕ for *any* pair of the structural parameters (μ, σ).

when proximity matrix $\mathbf{D}$ varies systematically, as in Matsuyama (2017).[33] The basic implications of **Proposition 1** for the polarity of endogenous spatial patterns—a single megacity or multiple cities—may well be robust to the generalizations of an assumed geography by incorporating, for example, the presence of boundaries and/or two-dimensional location spaces (see Appendix D).

Second, apart from exploring exogenous asymmetries, the symmetric racetrack geography can be used as a standard testbed to investigate the implications of endogenous mechanisms for a given model. For instance, Dingel et al. (2018) employed a circular geography to theoretically characterize the welfare effects of exogenous productivity differences under a standard international trade model. Another important topic is the consideration of multiple types of mobile agents that are subject to different proximity matrices and/or different degrees of increasing returns.[34]

[33]This line of research has already been tackled by the authors, such as in Ikeda et al. (2017b) (line segment); Ikeda et al. (2014, 2017a, 2018) (two-dimensional regular lattices), confirming that Class I models generally feature stable many-cities patterns.

[34]Such a structure is ubiquitous in intra-city models with both firms and households (e.g., urban models of Fujita and Ogawa, 1982) or in multiple-sector models (Fujita et al., 1999b). A circular geography provides a canonical starting point for this type of models (Tabuchi and Thisse, 2011; Osawa and Akamatsu, 2019). See Remark C.9 in Appendix C for detailed discussions.

A Proofs

A.1 Proof of Proposition 1

We characterize stability of $\bar{x} = (\bar{x}, \bar{x}, \ldots, \bar{x})$ and the destabilization of, and *bifurcation* from, it. Appendix E collects the technical facts referenced in the following.

Part 1 (*Stability of $\bar{x}$*). To define stability of $\bar{x}$, some myopic dynamics must be assumed4. A myopic dynamic describes the rate of change in x. Denote the dynamic that adjusts x over $\mathcal{X}$ by $\dot{x} = f(x)$, where $\dot{x}$ represents the time derivative. For the majority of myopic dynamics in the literature, $f(x) \equiv \tilde{f}(x, v(x))$ where $\tilde{f}$ maps each pair $(x, v(x))$ of a state and its associated payoff to a motion vector $\dot{x}$. We will focus exclusively on such dynamics. Let *restricted equilibrium* be a state $x^* \in \mathcal{X}$ such that $v_j(x^*) = v_k(x^*)$ for all $j, k \in \{i \in \mathcal{N} \mid x_i^* > 0\}$, that is, a spatial distribution in which all populated regions earn the same payoff level. A spatial equilibrium is always a restricted equilibrium.

We assume that f and $\tilde{f}$ are differentiable and satisfy:

$$f(x) = \mathbf{0} \text{ if and only if } x \text{ is a restricted equilibrium,} \tag{RS}$$

$$\text{if } f(x) \neq \mathbf{0}, \text{ then } v(x)^\top f(x) > 0, \text{ and} \tag{PC}$$

$$\mathbf{P}\tilde{f}(x, v(x)) = \tilde{f}(\mathbf{P}x, \mathbf{P}v(x)) \text{ for all permutation matrices } \mathbf{P} \text{ such that } \mathbf{PD} = \mathbf{DP}. \tag{Sym}$$

We call dynamics that satisfy (RS), (PC), and (Sym) *admissible dynamics*. See Remark C.4 in Appendix C for examples of admissible dynamics. The conditions (RS) and (PC) ensure that f is consistent with the underlying economic geography model v. The condition (Sym) ensures that f is consistent with Assumption S, since it implies that $\mathbf{P}f(x) = f(\mathbf{P}x)$ for all permutation matrix $\mathbf{P}$ such that $\mathbf{PD} = \mathbf{DP}$.

A rest point x^* of f (i.e., $x^* \in \mathcal{X}$ such that $f(x^*) = \mathbf{0}$) is said to be *linearly stable* if all the eigenvalues $\{\eta_k\}$ of $\nabla f(x^*) = [\frac{\partial f_i}{\partial x_j}(x^*)]$, the Jacobian matrix of f at x^*, have negative real parts. A spatial equilibrium x^* is said to be *stable* (*unstable*) if and only if it is linearly stable (unstable) under admissible dynamics.

Consider $\bar{x}$. We assume that $\bar{x}$ is an isolated spatial equilibrium. Then, (PC) implies that there is a neighborhood $\mathcal{O} \subset \mathcal{X}$ of $\bar{x}$ such that $v(x)^\top f(x) > 0$ for all $x \in \mathcal{O} \setminus \{\bar{x}\}$. By expanding v and f about $\bar{x}$, we see

$$\left(v(\bar{x}) + \nabla v(\bar{x}) z\right)^\top \left(f(\bar{x}) + \nabla f(\bar{x}) z\right) > 0. \tag{A.1}$$

Note that $v(\bar{x}) = \bar{v}\mathbf{1}$, $\nabla v(\bar{x}) = \frac{\bar{v}}{\bar{x}}\mathbf{V}$, $f(\bar{x}) = \mathbf{0}$ by (RS), and $\mathbf{1}^\top \nabla f(\bar{x}) z = 0$. The last relationship $\mathbf{1}^\top \nabla f(\bar{x}) z = 0$ follows because $\dot{x} = f(x) \approx f(\bar{x}) + \nabla f(\bar{z}) z = \nabla f(\bar{x}) z$ and $\mathbf{1}^\top \dot{x} = \sum_{i \in \mathcal{N}} \dot{x}_i =$

0 must hold true for all $\boldsymbol{x}$, since the total mass of agents is a constant. From (A.1), we then see

$$\frac{\bar{v}}{\bar{x}} (\mathbf{V}\boldsymbol{z})^\top (\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})\boldsymbol{z}) > 0 \tag{A.2}$$

for any infinitesimal migration $\boldsymbol{z} = \boldsymbol{x} - \bar{\boldsymbol{x}}$ from the uniform distribution.

Because we consider canonical models, there is a rational function $G(t) = \frac{G^\sharp(t)}{G^\flat(t)}$ such that $\mathbf{V} = G(\bar{\mathbf{D}}) = G^\flat(\bar{\mathbf{D}})^{-1}G^\sharp(\bar{\mathbf{D}})$, where $G^\sharp(t)$ and $G^\flat(t)$ are some polynomials. We assume $G^\flat(t) > 0$, so that $G^\sharp(t)$ is a net gain function. $\bar{\mathbf{D}}$ is real, symmetric, and *circulant* matrix under Assumption RE (see Appendix E). Then, by Fact E.2, $\mathbf{V}$ is real, symmetric, and circulant matrix. Because of (Sym), $\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$ is also real, symmetric, and circulant matrix. Then, by Fact E.3, $\mathbf{D}$, $\mathbf{V}$, and $\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$ share the same set of eigenvectors $\{\boldsymbol{z}_k\}$.

For every eigenvector $\boldsymbol{z}_k$ (of $\mathbf{D}$, $\mathbf{V}$, or $\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$), (A.2) implies that

$$(\mathbf{V}\boldsymbol{z}_k)^\top (\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})\boldsymbol{z}_k) = \omega_k \eta_k > 0, \tag{A.3}$$

where ω_k and η_k are the (real) eigenvalues of $\mathbf{V}$ and $\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$ associated with $\boldsymbol{z}_k$. Thus, $\mathrm{sgn}[\eta_k] = \mathrm{sgn}[\omega_k] = \mathrm{sgn}[G^\sharp(\chi_k(\phi))]$. Therefore, $\bar{\boldsymbol{x}}$ is stable spatial equilibrium if and only if $\omega_k^\sharp \equiv G^\sharp(\chi_k(\phi)) < 0$ for all k. Note that η_k and ω_k are both real because $\nabla \boldsymbol{f}(\bar{\boldsymbol{x}})$ and $\mathbf{V}$ are both symmetric.

If the eigenpairs $\{(\chi_k, \boldsymbol{z}_k)\}$ of the row-normalized proximity matrix $\bar{\mathbf{D}}$ are available, then the eigenpairs of $\mathbf{V} = G(\bar{\mathbf{D}})$ are given by $\{(G(\chi_k), \boldsymbol{z}_k)\}$ (see Fact E.1). We have the following lemma:

Lemma A.1. Assume that N is an even and let $M \equiv \frac{N}{2}$. Then, $\bar{\mathbf{D}}$ satisfies the following properties:

(a) There are $M+1$ distinct eigenvalues. The eigenpairs $\{(\chi_k, \boldsymbol{z}_k)\}$ are

$$\chi_0 = 1, \qquad \boldsymbol{z}_0 \equiv \langle 1 \rangle_{i=0}^{N-1}, \tag{A.4}$$

$$\chi_k, \qquad \begin{cases} \boldsymbol{z}_k^+ \equiv \langle \cos(\theta k i) \rangle_{i=0}^{N-1}, \\ \boldsymbol{z}_k^- \equiv \langle \sin(\theta k i) \rangle_{i=0}^{N-1}, \end{cases} \qquad k = 1, 2, \ldots, M-1 \tag{A.5}$$

$$\chi_M, \qquad \boldsymbol{z}_M \equiv \langle (-1)^i \rangle_{i=0}^{N-1}. \tag{A.6}$$

where $\theta = \frac{2\pi}{N}$, and by $\langle z_i \rangle_{i=0}^{N-1} \equiv \frac{1}{\|z\|}(z_i)_{i=0}^{N-1}$ we denote a normalized vector.

(b) Every χ_k ($k \neq 0$) is a differentiable strictly decreasing function of ϕ with $\lim_{\phi \to 0} \chi_k = 1$ and $\lim_{\phi \to 1} \chi_k = 0$.

(c) For all ϕ, $\{\chi_k\}$ ($k = 0, 1, 2, \ldots, M$) are ordered as

$$\begin{cases} 1 = \chi_0 > \chi_2 > \cdots > \chi_{2k} > \cdots > \chi_M > 0, \\ 1 > \chi_1 > \chi_3 > \cdots > \chi_{2k+1} > \cdots > \chi_{M-1} > 0, \end{cases} \tag{A.7}$$

with $\chi_0 > \chi_1 > \chi_2$, so that $\max_{k \geq 1}\{\chi_k\} = \chi_1$.

(d) If N is a multiple of four, $\min_k\{\chi_k\} = \chi_M = \left(\frac{1-\phi}{1+\phi}\right)^2$ and $\chi_1 = \frac{1-\phi}{1+\phi}$.

Proof. See Fact E.3, Remark C.5, and Akamatsu et al. (2012) (Lemma 4.2). □

Thus, $\bar{x}$ is stable if $\omega_k^\sharp \equiv G^\sharp(\chi_k(\phi)) < 0$ for all $k \in \mathcal{K} \equiv \{1, 2, \ldots, M\}$. We can exclude $k = 0$ because z_0 represents an increase of total population.

Remark A.1. Assumption RE assumes that N is a multiple of four to ensure that $\min_{k \in \mathcal{K}}\{\chi_k\} = \chi_M$. The essential implication of **Proposition 1** on the *polarity* of spatial patterns does not alter because $\min_{k \in \mathcal{K}}\{\chi_k\} = \min\{\chi_{M-1}, \chi_M\}$ by **Lemma A.1** (c). ■

Refer to Figure 15, which shows $\{\omega_k^\sharp\}$, $G^\sharp(\chi)$, and $\{\chi_k\}$, to understand the following arguments.

Class I. By assumption, there is χ^* such that $G^\sharp(\chi) < 0$ for all $\chi \in (\chi^*, 1)$, that $G^\sharp(\chi^*) = 0$, and that $G^\sharp(\chi) > 0$ for all $\chi \in (0, \chi^*)$. By **Lemma C.1**, $\{\chi_k(\phi)\}$ are strictly decreasing from 1. Therefore, $\bar{x}$ is stable if and only if $\chi_k \in (\chi^*, 1)$, so that $\omega_k^\sharp \equiv G^\sharp(\chi_k) < 0$, for all $k \in \mathcal{K}$, i.e., if $\chi^* < \min_{k \in \mathcal{K}} \chi_k = \chi_M$. Thus, $\bar{x}$ is stable for all $(0, \phi_M^*)$ where $\phi_M^* = \frac{1-\sqrt{\chi^*}}{1+\sqrt{\chi^*}}$ is the unique solution for $\chi_M = \chi^*$. Because $G^\sharp(\chi) > 0$ for all $\chi \in (0, \chi^*)$ and χ_M is strictly decreasing, $\bar{x}$ is unstable for all $(\phi_M^*, 1)$ because $\omega_M^\sharp > 0$ for the range.

Class II. By assumption, there is χ^{**} such that $G^\sharp(\chi) < 0$ for all $\chi \in (0, \chi^{**})$, that $G^\sharp(\chi^{**}) = 0$, and that $G^\sharp(\chi^{**}) > 0$ for all $\chi \in (\chi^{**}, 1)$. Thus, $\bar{x}$ is stable if and only if $\chi_k \in (0, \chi^{**})$, so that $\omega_k^\sharp = G^\sharp(\chi_k) < 0$, for all $k \in \mathcal{K}$, i.e., if $\chi^{**} > \max_{k \in \mathcal{K}} \chi_k = \chi_1$. Thus, $\bar{x}$ is stable for all $(\phi_1^*, 1)$ where $\phi_1^* \equiv \frac{1-\chi^{**}}{1+\chi^{**}}$ is the unique solution for $\chi_1 = \chi^{**}$. Because $G^\sharp(\chi) > 0$ for all $\chi \in (\chi^{**}, 1)$ and χ_1 is strictly decreasing, $\bar{x}$ is unstable for all $(0, \phi_1^*)$.

Class III. Via a similar logic, we see $\bar{x}$ is stable if $\phi \in (0, \phi_M^*) \cup (\phi_1^*, 1)$.

Part 2 (*Bifurcation from* $\bar{x}$). Start from a state where $\bar{x}$ is stable. When one and only one ω_k ($k \in \mathcal{K}$) switches its sign from negative to positive at ϕ_k^*, then, from (A.3), η_k must switch its sign from negative to positive at ϕ_k^*. It is a standard fact in bifurcation theory that, at such point, $\bar{x}$ must deviate towards the direction of corresponding eigenvector z_k. See, e.g., Hirsch et al. (2012) and Kuznetsov (2004). That is, $M = \frac{N}{2}$ cities emerge at ϕ_M^*, whereas a single city emerges at ϕ_1^{**}.

Remark A.2. The bifurcation toward the single-city direction ($k = 1$) is a *double bifurcation* at which the relevant eigenvalue, ω_1, has multiplicity two, as seen from (A.5). For this case, possible migration patterns are linear combinations of the form $c^+ z_1^+ + c^- z_1^-$ with $c^+, c^- \in \mathbb{R}$. In fact, we have $(c^+, c^-) = (c, 0)$ or (c, c) for some $c \in \mathbb{R}$ under admissible dynamics along with Assumptions RE and S. Because any linear combination of z_1^+ and z_1^- is a one-peaked cosine curve, it is still interpreted as a single-city pattern. ■

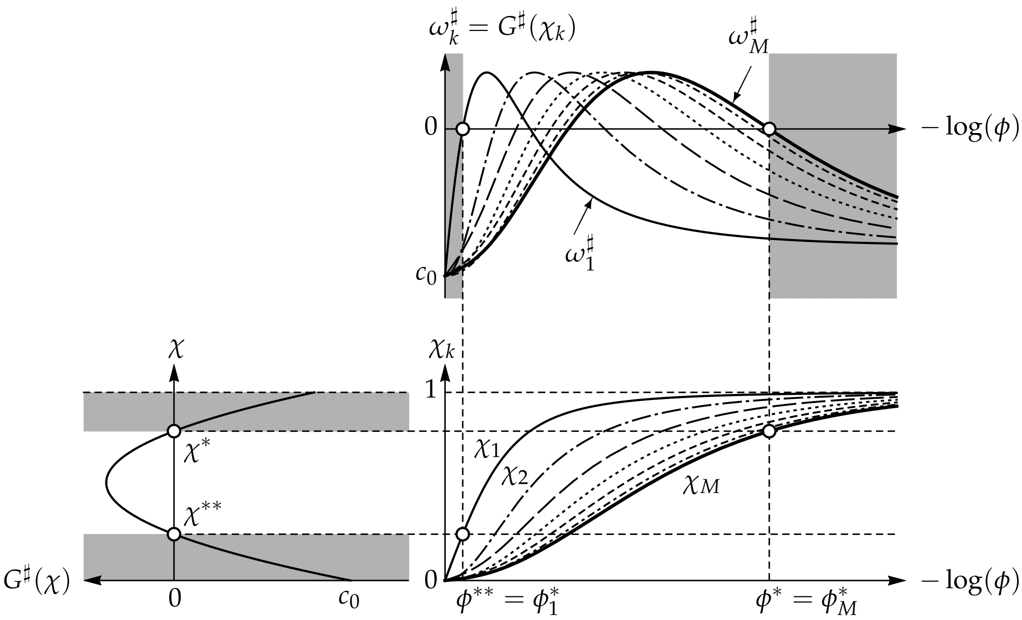


Figure 15: Net gain function $G^\sharp(\chi)$ and net agglomeration forces $\omega_k^\sharp$

Notes: Top: Graphs of $\omega_k^\sharp = G^\sharp(\chi_k)$. Bottom left: Net gain function $G^\sharp$ for Class III models with a quadratic net gain function of the form $G^\sharp(\chi) = c_0 + c_1\chi + c_2\chi^2$. Note that $G^\sharp(0) = c_0$. Bottom right: The eigenvalues $\{\chi_k(\phi)\}$ of $\bar{\mathbf{D}}$, which are model-independent. $\bar{\boldsymbol{x}}$ is stable in the shaded regions of ϕ or χ. For ϕ, log scale is used for better readability. Note that $\max\{\chi_k\} = \chi_1$ and $\min\{\chi_k\} = \chi_M$ at any given level of ϕ.

A.2 Proof of Proposition 2

The equilibrium condition when all regions are populated is given by

$$\boldsymbol{v}(\boldsymbol{x}, \boldsymbol{a}) - \bar{v}(\boldsymbol{x}, \boldsymbol{a})\mathbf{1} = \mathbf{0}, \tag{A.8}$$

where $\bar{v}(\boldsymbol{x}, \boldsymbol{a}) \equiv \sum_{i\in\mathcal{N}} v_i(\boldsymbol{x}, \boldsymbol{a})x_i$ is the average payoff and $\mathbf{1}$ is N-dimensional all-one vector. The pair $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{a}})$ is a solution to (A.8). When $\boldsymbol{a} = \bar{\boldsymbol{a}} + \boldsymbol{\epsilon}$ with small $\boldsymbol{\epsilon} = (\epsilon_i)_{i\in\mathcal{N}}$, there is a spatial equilibrium nearby $\bar{\boldsymbol{x}}$ because $\boldsymbol{v}$ is differentiable. Let $\boldsymbol{x}(\boldsymbol{a})$ denote the perturbed version of the uniform distribution, which is a function in $\boldsymbol{a}$. In the following, we consider some level of ϕ such that $\bar{\boldsymbol{x}}$ is stable, because otherwise studying a perturbed version of $\bar{\boldsymbol{x}}$ is meaningless.

The covariance ρ discussed in Section 6 is evaluated as follows:

$$\rho \equiv (\boldsymbol{a} - \bar{\boldsymbol{a}})^\top (\boldsymbol{x}(\boldsymbol{a}) - \bar{\boldsymbol{x}}) = (\mathbf{C}\boldsymbol{a})^\top \mathbf{C}\boldsymbol{x}(\boldsymbol{a}) = \boldsymbol{a}^\top \mathbf{C}\boldsymbol{x}(\boldsymbol{a}) \tag{A.9}$$

where $\mathbf{C} \equiv \mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\top$ is the N-dimensional centering matrix. Let $\mathbf{X} \equiv [\frac{\partial x_i}{\partial a_j}(\bar{\boldsymbol{a}})]$ is the Jacobian matrix of $\boldsymbol{x}$ with respect to $\boldsymbol{a}$ evaluated at $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{a}})$ Then, $\boldsymbol{x}(\boldsymbol{a}) \approx \bar{\boldsymbol{x}} + \mathbf{X}(\boldsymbol{a} - \bar{\boldsymbol{a}}) = \bar{\boldsymbol{x}} + \mathbf{X}\mathbf{C}\boldsymbol{a}$ and

$$\rho = \boldsymbol{a}^\top \mathbf{C}\mathbf{X}\mathbf{C}\boldsymbol{a} \tag{A.10}$$

since $\mathbf{C}\bar{x} = \mathbf{0}$. The implicit function theorem regarding (A.8) at $(\bar{x}, \bar{a})$ gives:

$$\mathbf{X} = -\left(\mathbf{V}_x - \mathbf{1}\bar{x}^\top \mathbf{V}_x - \mathbf{1}v(\bar{x})^\top\right)^{-1}\left(\mathbf{V}_a - \mathbf{1}\bar{x}^\top \mathbf{V}_a\right) \tag{A.11}$$

$$= \left(\frac{\bar{v}}{\bar{x}}\frac{1}{N}\mathbf{1}\mathbf{1}^\top - \left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\right)\mathbf{V}_x\right)^{-1}\left(\mathbf{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^\top\right)\mathbf{V}_a \tag{A.12}$$

$$= \frac{\bar{x}}{\bar{v}}\left((\mathbf{I}-\mathbf{C}) - \mathbf{C}\frac{\bar{x}}{\bar{v}}\mathbf{V}_x\right)^{-1}\mathbf{C}\frac{\bar{v}}{\bar{a}}\frac{\bar{a}}{\bar{v}}\mathbf{V}_a \tag{A.13}$$

$$= \frac{\bar{x}}{\bar{a}}\left((\mathbf{I}-\mathbf{C}) - \mathbf{C}\mathbf{V}\right)^{-1}\mathbf{C}\mathbf{A} \tag{A.14}$$

where $\bar{v}$ is the uniform level of payoff, $\mathbf{V}_x \equiv [\frac{\partial v_i}{\partial x_j}(\bar{x}, \bar{a})]$, $\mathbf{V}_a \equiv [\frac{\partial v_i}{\partial a_j}(\bar{x}, \bar{a})]$, $\mathbf{V} \equiv \frac{\bar{x}}{\bar{v}}\mathbf{V}_x$, and $\mathbf{A} \equiv \frac{\bar{a}}{\bar{v}}\mathbf{V}_a$, and we note $\bar{x} = \frac{1}{N}$. Under Assumptions RE, S, and A, the matrix $\mathbf{X}$ is real, symmetric, and circulant because all its components in (A.14) are.

The set of eigenvectors of $\mathbf{CXC}$ can be chosen as the same as in **Lemma C.1** (a) because it is a circulant matrix of the same size as $\bar{\mathbf{D}}$ (see Fact E.3). Let $\{\lambda_k\}_{k=0}^{M}$ be the distinct eigenvalues of $\mathbf{CXC}$. Because $\mathbf{CXC}$ is symmetric, $\mathbf{CXC}$ admits the following eigenvalue decomposition:

$$\mathbf{CXC} = \lambda_0 \mathbf{1}\mathbf{1}^\top + \sum_{k=1}^{M-1}\lambda_k\left(z_k^+ z_k^{+\top} + z_k^- z_k^{-\top}\right) + \lambda_M z_M z_M^\top, \tag{A.15}$$

which yields the following representation of ρ:

$$\rho = \boldsymbol{a}^\top \mathbf{CXC}\boldsymbol{a} = \sum_{k\neq 0}\tilde{a}_k^2\lambda_k. \tag{A.16}$$

where $\tilde{\boldsymbol{a}} \equiv (\tilde{a}_k)$ is the representation of $\boldsymbol{a}$ in the new coordinate system.[35] We can omit $k = 0$ since $\lambda_0 = 0$, which reflects that $z_0 = \mathbf{1}$ represents a uniform increase in $\boldsymbol{a}$ and thus is inconsequential. In concrete terms, as all the component matrices in (A.14) are circulant matrices and hence shares the same set of eigenvectors, we can translate the matrix relationship (A.14) to the following expression:

$$\lambda_k = \frac{\bar{x}}{\bar{a}}\left((1-\kappa_k) - \kappa_k\omega_k\right)^{-1}\kappa_k\alpha_k, \tag{A.17}$$

where κ_k, $\omega_k = G(\chi_k)$, and α_k are the kth eigenvalues of $\mathbf{C}$, $\mathbf{V}$, and $\mathbf{A}$, respectively. We have $\kappa_0 = 0$ and $\kappa_k = 1$ for all $k \neq 0$, thereby $\lambda_k = -\frac{\bar{x}}{\bar{a}}\frac{\alpha_k}{\omega_k}$ for all $k \neq 0$. Also, under the condition that $\mathbf{A} = G^\natural(\bar{\mathbf{D}})$, we have $\alpha_k = G^\natural(\chi_k)$. Summing up, we have

$$\lambda_k = -\frac{\bar{x}}{\bar{a}}\frac{G^\natural(\chi_k)}{G(\chi_k)} = \frac{\bar{x}}{\bar{a}}\delta(\chi_k) \qquad \forall k \in \mathcal{K}, \tag{A.18}$$

and $\lambda_0 = 0$ where $\delta(\chi) \equiv -\frac{G^\natural(\chi)}{G(\chi)}$, and $\{\chi_k\}_{k\in\mathcal{K}}$ are the eigenvalues of $\bar{\mathbf{D}}$.

Thus, $\rho > 0$ for all $\boldsymbol{a}$ if all $\{\lambda_k\}$ are positive except for $\lambda_0 = 0$. The denominator of (A.18),

[35] Observe that this representation implies upper and lower bounds for ρ when $\boldsymbol{a} \neq \bar{\boldsymbol{a}}$, namely, $\|\boldsymbol{a} - \bar{\boldsymbol{a}}\|\lambda^{\min} \leq \rho \leq \|\boldsymbol{a} - \bar{\boldsymbol{a}}\|\lambda^{\max}$ where $\lambda^{\min} \equiv \min_{k\geq 1}\{\lambda_k\}$ and $\lambda^{\max} \equiv \max_{k\geq 1}\{\lambda_k\}$.

$G(\chi_k)$, must be negative for all k because $\bar{x}$ is stable by assumption. Thus, we see that $\rho > 0$ if $G^{\natural}(\chi) > 0$ for all χ since $\chi_k \in (0,1)$ for all $k \in \mathcal{K}$.

Proposition 2 follows by noting that

$$\rho'(\phi) = \sum_{k \neq 0} \tilde{a}_k^2 \frac{\mathrm{d}\lambda_k}{\mathrm{d}\phi} = \frac{\bar{x}}{\bar{a}} \sum_{k \neq 0} \tilde{a}_k^2 \delta'(\chi_k) \frac{\mathrm{d}\chi_k}{\mathrm{d}\phi} = -\frac{\bar{x}}{\bar{a}} \sum_{k \neq 0} \tilde{a}_k^2 \delta'(\chi_k) \left| \frac{\mathrm{d}\chi_k}{\mathrm{d}\phi} \right|,$$

where we recall that $\{\chi_k\}_{k \in \mathcal{K}}$ are strictly decreasing in ϕ (**Lemma A.1**). If $\delta'(\chi) < 0$ ($\delta'(\chi) > 0$) for all relevant χ, then $\rho'(\phi) > 0$ ($\rho'(\phi) < 0$).

References

AHLFELDT, GABRIEL M., STEPHEN J. REDDING, DANIEL M. STURM, AND NIKOLAUS WOLF, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

AKAMATSU, TAKASHI, SHOTA FUJISHIMA, AND YUKI TAKAYAMA, "Discrete-space agglomeration model with social interactions: Multiplicity, stability, and continuous limit of equilibria," *Journal of Mathematical Economics*, 2017, *69*, 22–37.

_ , TOMOYA MORI, MINORU OSAWA, AND YUKI TAKAYAMA, "Spatial scale of agglomeration and dispersion: Theoretical foundations and empirical implications," 2017. Unpublished manuscript, MPRA Paper No. 84145.

_ , YUKI TAKAYAMA, AND KIYOHIRO IKEDA, "Spatial discounting, Fourier, and racetrack economy: A recipe for the analysis of spatial agglomeration models," *Journal of Economic Dynamics and Control*, 2012, *99* (11), 32–52.

ALLEN, TREB AND COSTAS ARKOLAKIS, "Trade and the topography of the spatial economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

_ AND DAVE DONALDSON, "The geography of path dependence," 2018. Unpublished manuscript.

_ , COSTAS ARKOLAKIS, AND YUTA TAKAHASHI, "Universal gravity," *Journal of Political Economy*, 2019, *Forthcoming.*

ALONSO, WILIAM, *Location and Land Use: Toward a General Theory of Land Rent*, Cambridge, MA: Harvard University Press, 1964.

ANAS, ALEX AND IKKI KIM, "General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration," *Journal of Urban Economics*, 1996, *40* (2), 232–256.

_ , RICHARD ARNOTT, AND KENNETH A SMALL, "Urban spatial structure," *Journal of Economic Literature*, 1998, *36* (3), 1426–1464.

ANDERSON, SIMON P., ANDRE DE PALMA, AND JACQUES FRANÇOIS THISSE, *Discrete Choice Theory of Product Differentiation*, MIT Press, 1992.

ARKOLAKIS, COSTAS, ARNAUD COSTINOT, AND ANDRÉS RODRÍGUEZ-CLARE, "New trade models, same old gains?," *American Economic Review*, 2012, *102* (1), 94–130.

ARMINGTON, PAUL S., "A theory of demand for product distinguished by place of production," *International Monetary Fund Staff Papers*, 1969, *16* (1), 159–178.

Arthur, W. Brian, *Increasing Returns and Path Dependence in the Economy*, Ann Arbor, MI: University of Michigan Press, 1994.

Baldwin, Richard E., Rikard Forslid, Philippe Martin, Gianmarco I. P. Ottaviano, and Frederic Robert-Nicoud, *Economic Geography and Public Policy*, Princeton University Press, 2003.

Baum-Snow, Nathaniel, "Did highways cause suburbanization?," *The Quarterly Journal of Economics*, May 2007, *122* (2), 775–805.

_ , Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang, "Roads, railroads and decentralization of Chinese cities," *Review of Economics and Statistics*, 2017, *99* (3), 435–448.

Beckmann, Martin J., "Spatial equilibrium in the dispersed city," in Yorgos Papageorgiou, ed., *Mathematical Land Use Theory*, Lexington Book, 1976.

Blanchet, Adrien, Pascal Mossay, and Filippo Santambrogio, "Existence and uniqueness of equilibrium for a spatial model of social interactions," *International Economic Review*, 2016, *57* (1), 36–60.

Bleakley, Hoyt and Jeffrey Lin, "Portage and path dependence," *The Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

Brown, George W. and John von Neumann, "Solutions of games by differential equations," in Harold W. Kuhn and Albert W. Tucker, eds., *Contributions to the Theory of Games I*, Princeton University Press, 1950.

Christaller, Walter, *Die Zentralen Orte in Süddeutschland*, Jena: Gustav Fischer, 1933.

Davis, Donald R. and Jonathan I. Dingel, "The comparative advantage of cities," April 2019. Unpublished manuscript, The University of Chicago Booth School of Business.

Desmet, Klaus and Esteban Rossi-Hansberg, "Spatial growth and industry age," *Journal of Economic Theory*, 2009, *144* (6), 2477–2502s.

_ and _ , "Spatial development," *American Economic Review*, 2014, *104* (4), 1211–1243.

_ and _ , "The spatial development of India," *Journal of Regional Science*, 2015, *55* (1), 10–30.

_ , Dávid Krisztián Nagy, and Esteban Rossi-Hansberg, "The geography of development," *Journal of Political Economy*, 2018, *126* (3), 903–983.

Dingel, Jonathan I., Kyle C. Meng, and Solomon M. Hsiang, "The spatial structure of productivity, trade, and inequality: Evidence from the global climate," 2018. Unpublished manuscript.

Dupuis, Paul and Anna Nagurney, "Dynamical systems and variational inequalities," *Annals of Operations Research*, 1993, *44* (1), 7–42.

Duranton, Gilles and Diego Puga, "Micro-foundations of urban agglomeration economies," in J. Vernon Henderson and Jacques-François Thisse, eds., *Handbook of Regional and Urban Economics*, Vol. 4, North-Holland, 2004, pp. 2063–2117.

_ and Matthew A. Turner, "Urban growth and transportation," *Review of Economic Studies*, 2012, *79* (4), 1407–1440.

Faber, Benjamin, "Trade integration, market size, and industrialization: Evidence from China's national trunk highway system," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Facchinei, Francisco and Jong-Shi Pang, *Finite-dimensional Variational Inequalities and Complementarity Problems*, Springer Science & Business Media, 2007.

Forslid, Rikard and Gianmarco I. P. Ottaviano, "An analytically solvable core-periphery model," *Journal of Economic Geography*, 2003, *33* (3), 229–240.

Fujita, Masahisa and Hideaki Ogawa, "Multiple equilibria and structural transformation of non-monocentric urban configurations," *Regional Science and Urban Economics*, 1982, *12*, 161–196.

_ and Jacques-François Thisse, *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth (Second Edition)*, Cambridge University Press, 2013.

_ and Paul R. Krugman, "When is the economy monocentric?: von Thuünen and Chamberlin unified," *Regional Science and Urban Economics*, 1995, *25* (4), 505–528.

_ , _ , and Anthony Venables, *The Spatial Economy: Cities, Regions, and International Trade*, Princeton University Press, 1999.

_ , _ , and Tomoya Mori, "On the evolution of hierarchical urban systems," *European Economic Review*, February 1999, *43* (2), 209–251.

Gilboa, Itzhak and Akihiko Matsui, "Social stability and equilibrium," *Econometrica*, 1991, *59* (3), 859–867.

Golubitsky, Martin and David G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, Vol. 1, Springer Science & Business Media, 2012.

_ and Ian Stewart, *The Symmetry Perspective: From Equilibrium to Chaos in Phase Space and Physical Space*, Vol. 200, Springer Science & Business Media, 2003.

_ , _ , and David G. Schaeffer, *Singularities and Groups in Bifurcation Theory*, Vol. 2, Springer Science & Business Media, 2012.

Harris, Britton and Alan G. Wilson, "Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models," *Environment and Planning A*, 1978, *10* (4), 371–388.

Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm, "The making of the modern metropolis: Evidence from London," 2018. Unpublished manuscript, Lodon School of Economics.

Helpman, Elhanan, "The size of regions," in David Pines, Efrainm Sadka, and Itzhak Zilcha, eds., *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge University Press, 1998, pp. 33–54.

Hirsch, Morris W., Stephen Smale, and Robert L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, Academic Press, 2012.

Horn, Roger A. and Charles R. Johnson, *Matrix Analysis (Second Edition)*, Cambridge University Press, 2012.

Hsu, Wen-Tai, "Central place theory and city size distribution," *The Economic Journal*, 2012, *122*, 903–932.

Hummels, David and Alexandre Skiba, "Shipping the good apples out? An empirical confirmation of the Alchian-Allen conjecture," *Journal of Political Economy*, 2004, *112* (6), 1384–1402.

Ikeda, Kiyohiro and Kazuo Murota, *Bifurcation Theory for Hexagonal Agglomeration in Economic Geography*, Springer, 2014.

_ , _ , and Yuki Takayama, "Stable economic agglomeration patterns in two dimensions: Beyond the scope of central place theory," *Journal of Regional Science*, 2017, *57* (1), 132–172.

_ , _ , Takashi Akamatsu, and Yuki Takayama, "Agglomeration patterns in a long narrow economy of a new economic geography model: Analogy to a racetrack economy," *International Journal of Economic Theory*, 2017, *13*, 113–145.

_ , _ , _ , Tatsuhito Kono, and Yuki Takayama, "Self-organization of hexagonal agglomeration patterns in new economic geography models," *Journal of Economic Behavior and Organization*, 2014, *99*, 32–52.

_ , Mikihisa Onda, and Yuki Takayama, "Bifurcation theory of a racetrack economy in a spatial economy model," *Networks and Spatial Economics*, 2018, pp. 1–26.

_ , Takashi Akamatsu, and Tatsuhito Kono, "Spatial period-doubling agglomeration of a core–periphery model with a system of cities," *Journal of Economic Dynamics and Control*, 2012, *36* (5), 754–778.

Irarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla, "The tip of the iceberg: A quantitative framework for estimating trade costs," *Review of Economics and Statistics*, 2015, *97* (4), 777–792.

Krugman, Paul R., "History versus expectations," *The Quarterly Journal of Economics*, 1991, *106* (2), 651–667.

_ , "Increasing returns and economic geography," *Journal of Political Economy*, 1991, *99* (3), 483–499.

_ , "On the number and location of cities," *European Economic Review*, 1993, *37* (2), 293–298.

_ and Anthony J. Venables, "Globalization and the Inequality of Nations," *The Quarterly Journal of Economics*, 1995, *110* (4), 857–880.

Kuznetsov, Yuri A., *Elements of Applied Bifurcation Theory (3rd Eds.)*, Springer-Verlag, 2004.

Lucas, Robert E. and Esteban Rossi-Hansberg, "On the internal structure of cities," *Econometrica*, 2002, *70* (4), 1445–1476.

Matsuyama, Kiminori, "Geographical advantage: Home market effect in a multi-region world," *Research in Economics*, 2017, *71*, 740–758.

Mertikopoulos, Panayotis and William H. Sandholm, "Riemannian game dynamics," *Journal of Economic Theory*, 2018, *177*, 315–364.

Michaels, Guy and Ferdinand Rauch, "Resetting the urban network: 117–2012," *The Economic Journal*, 2018, *128* (608), 378–412.

Mori, Tomoya, "A modeling of megalopolis formation: The maturing of city systems," *Journal of Urban Economics*, 1997, *42*, 133–157.

Mossay, Pascal and Pierre M. Picard, "On spatial equilibria in a social interaction model," *Journal of Economic Theory*, 2011, *146* (6), 2455–2477.

Murata, Yasusada and Jacques-François Thisse, "A simple model of economic geography à la Helpman–Tabucbi," *Journal of Urban Economics*, 2005, *58* (1), 137–155.

Nagy, Dávid Krisztián, "City location and economic growth," January 2017. Unpublished manuscript.

Nash, John, "Non-cooperative games," *Annals of Mathematics*, 1951, *54* (2), 286–295.

Osawa, Minoru and Takashi Akamatsu, "Equilibrium refinement for a model of non-monocentric internal structures of cities: A potential game approach," 2019. Unpublished manuscript.

— , — , and Yuki Takayama, "Harris and Wilson (1978) model revisited: The spatial period-doubling cascade in an urban retail model," *Journal of Regional Science*, 2017, *57* (3), 442–466.

Ota, Mitsuru and Masahisa Fujita, "Communication technologies and spatial organization of multi-unit firms in metropolitan areas," *Regional Science and Urban Economics*, 1993, *23* (6), 695–729.

Ottaviano, Gianmarco I. P., Takatoshi Tabuchi, and Jacques-François Thisse, "Agglomeration and trade revisited," *International Economic Review*, 2002, *43*, 409–436.

Owens, Raymond III, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte, "Rethinking Detroit," *American Economic Journal: Economic Policy*, forthcoming.

Oyama, Daisuke, "Agglomeration under forward-looking expectations: Potentials and global stability," *Regional Science and Urban Economics*, 2009, *39* (6), 696–713.

— , "History versus expectations in economic geography reconsidered," *Journal of Economic Dynamics and Control*, 2009, *33* (2), 394–408.

Papageorgiou, Yorgos Y. and Terrence R. Smith, "Agglomeration as local instability of spatially uniform steady-states," *Econometrica*, 1983, pp. 1109–1119.

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

— and Jens Südekum, "Integration, agglomeration and welfare," *Journal of Urban Economics*, March 2008, *63* (2), 544–566.

Picard, Pierre M. and Takatoshi Tabuchi, "On microfoundations of the city," *Journal of Economic Theory*, 2013, *148* (6), 2561–2582.

Proost, Stef and Jacques-François Thisse, "What can be learned from spatial economics?," *Journal of Economic Literature*, 2019, *57*, 575–643.

Puga, Diego, "The rise and fall of regional inequalities," *European Economic Review*, 1999, *43* (2), 303–334.

Redding, Stephen J. and Daniel Sturm, "The cost of remoteness: Evidence from German division and reunification," *American Economic Review*, 2008, *98* (5), 1766–1797.

— and Esteban Rossi-Hansberg, "Quantitative spatial economics," *Annual Review of Economics*, 2017, *9*, 21–58.

Sandholm, William H., *Population Games and Evolutionary Dynamics*, MIT Press, 2010.

Smith, Michael J., "The stability of a dynamic model of traffic assignment: An application of a method of Lyapunov," *Transportation Science*, 1984, *18* (3), 245–252.

Statistics Bureau, Ministry of Internal Affairs and Communications of Japan, "Population Census," 1970.

_ , "Population Census," 2015.

Tabuchi, Takatoshi, "Urban agglomeration and dispersion: A synthesis of Alonso and Krugman," *Journal of Urban Economics*, 1998, *44* (3), 333–351.

_ and Jacques-François Thisse, "A new economic geography model of central places," *Journal of Urban Economics*, 2011, *69* (2), 240–252.

Takayama, Yuki and Takashi Akamatsu, "Emergence of polycentric urban configurations from combination of communication externality and spatial competition," *Journal of JSCE Series D3: Infrastructure Planning and Management*, 2011, *67* (1), 001–020.

Taylor, Peter D. and Leo B. Jonker, "Evolutionarily stable strategies and game dynamics," *Mathematical Biosciences*, 1978, *40*, 145–156.

Supplementary Materials (Not for Publication)

December 10, 2019

By Takashi Akamatsu, Tomoya Mori, Minoru Osawa, and Yuki Takayama

We collect additional results, technical preliminaries, and various remarks that are referenced in the main text or proofs. Appendix B provides the summary of the Japan example in Section 1. Appendix C collects lengthy remarks. Appendix D provides discussions pertaining to the relaxation of the racetrack assumption (Assumption RE). Appendix E collects relevant facts from matrix analysis employed in the proofs (Appendix A). Appendix F collects supporting computations for the example models discussed in the main text.

B Cities in Japan: 1970–2015

Data. Population count data of Japan, obtained from Statistics Bureau, Ministry of Internal Affairs and Communications of Japan (1970, 2015).

Method. A city is represented by an *urban agglomeration* (*UA*), which is the set of contiguous 1 km-by-1 km cells with a population density of at least 1000/km^2 and total population of at least 10,000. The basic results below remain the same for alternative threshold densities and populations.

Below, UA i in year s is said to be *associated with* UA j in year t $(\neq s)$ if the intersection of the spatial coverage of i and that of j accounts for the largest population of i among all the UAs in year t. For years $s < t$, if i and j are associated with each other, they are considered to be *the same* UA. If i is associated with j but not vice versa, then i is considered to have been *absorbed* into j, while if j is associated with i but not vice versa, then j is considered to have *separated* from i. If i is not associated with any UA in year t, then i is considered to have *disappeared* by year t, while if j is not associated with any UA in year s, then j is considered to have newly *emerged* by year t.

For the part of Japan contiguous by roads to at least one of the four major islands (Hokkaido, Honshu, Shikoku, and Kyushu), 503 and 450 UAs are identified, as depicted in Panels (a) and (b) of Figure B.1 for 1970 and 2015, respectively, where the warmer color indicates a larger population. These together account for 64% and 78% of the total population in 1970 and 2015, respectively. Thus, there is a substantial 18% increase in the urban share over these 40 years. Of the 503 UAs that existed in 1970, 302 survived, while 201 either disappeared or integrated with other UAs by 2015. Of the 450 UAs that existed in 2015, 148 were newly formed after 1970 (including those split from existing UAs).

Panels (c) and (d) of Figure B.1 show the highway and high-speed railway networks in use in 1970 and 2015, respectively. The comparison of these panels indicates an obvious substantial expansion of these networks during these 45 years, as mentioned in the text.

Panels (a), (b), and (c) of Figure B.2 show the distributions of the growth rates of population share (in the national population), the areal size and population density of individual UAs for the set of the 302 UAs that survived throughout the 45-year period. A UA experienced an average growth rate of 21% (75%) of population share, 94% (105%) of areal size, and −22% (22%) of population density (per km^2), respectively, where the numbers in parentheses are the standard deviations.

As a larger population share was concentrated in a smaller number of UAs in 2015 than in 1970, the spatial size of an individual UA almost doubled on average. However, these spatial expansions are not simply due to the shortage of available land in UAs. Note that population density decreased by 22% on average. We take this as evidence of a decline in the number of major population concentrations combined with local flattening of each concentration in the course of the improvement in interregional transport access.

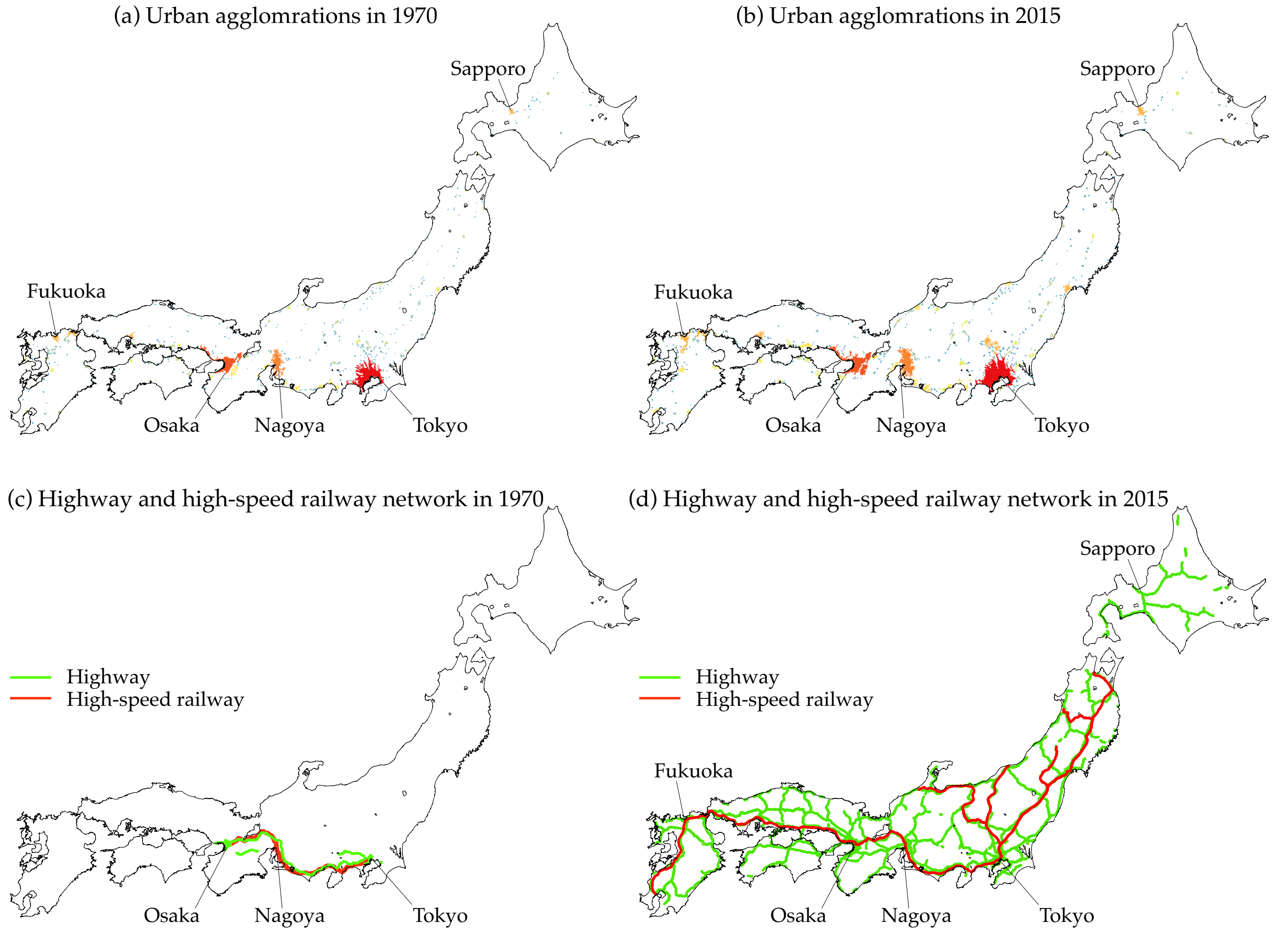


Figure B.1: UAs and transport network in Japan

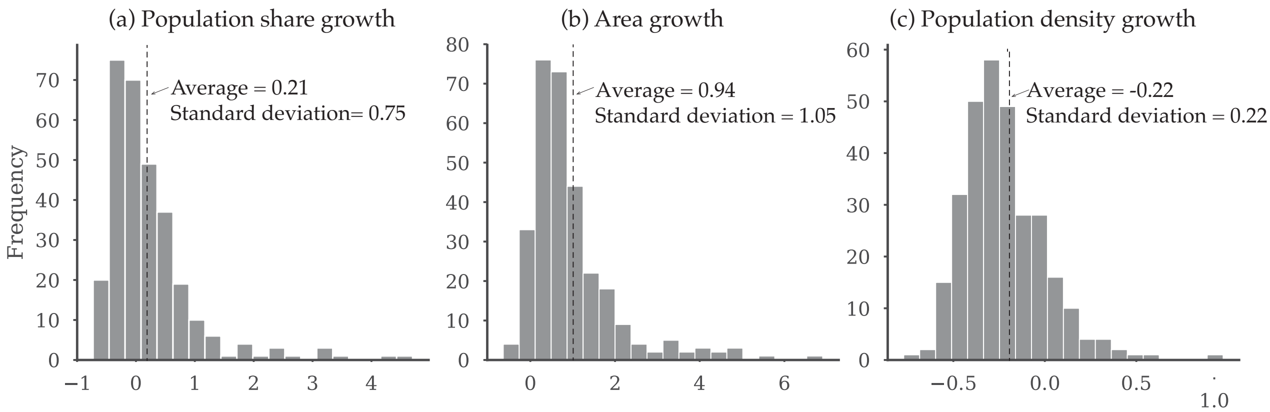


Figure B.2: Growth rates of the sizes of UAs in Japan.

C Additional remarks

This appendix collects lengthy remarks which may benefit interested readers.

Remark C.1. Section 2.1 and Definition 1 define canonical models. Canonical models do not include: (i) models with multiple types of mobile agents such as the urban models of Fujita and Ogawa (1982); Ota and Fujita (1993); Lucas and Rossi-Hansberg (2002) as well as Ahlfeldt et al. (2015); Owens et al. (forthcoming); (ii) models with sector-wise differentiated spatial frictions such as Fujita and Krugman (1995) and Mori (1997); (iii) models with multiple types of increasing returns such as Fujita et al. (1999b); Tabuchi and Thisse (2011), and Hsu (2012); and (iv) dynamic models such as Desmet and Rossi-Hansberg (2009, 2014, 2015); Desmet et al. (2018). Also, we do not cover models that build on Ottaviano et al. (2002) (OTT) framework, which assumes a quadratic preference and linear trade frictions. The OTT framework is not a canonical model as per Definition 1, because the model induces $\mathbf{V}$ of the form

$$\mathbf{V} = c_0\mathbf{I} + c_1\bar{\mathbf{D}} + c_2\bar{\mathbf{D}}^2 + \hat{c}_1\mathbf{D}^{[2]}, \tag{C.1}$$

where $\mathbf{D}^{[2]} \equiv [d_{ij}^2]$. That is, there is no rational function G such that $\mathbf{V} = G(\bar{\mathbf{D}})$ due to the existence of the special matrix, $\mathbf{D}^{[2]}$. We can still show that $c_2 < 0$, thereby the model has a global dispersion force. Also, under Assumptions RE and S, local stability analysis of $\bar{x}$ in the OTT framework can be done analytically. However, since the OTT framework assumes linear transport costs, the analysis can incorporate tiresome parametric classifications to handle possible corner solutions in market equilibrium (e.g., the cases where there are no interregional transport for some pairs of regions). ■

Remark C.2. On **Proposition 1**, the payoff function is assumed to be *homogeneous* across mobile agents. The effects of considering idiosyncratic payoff shocks are of interest, since it is a standard recipe in quantitative exercises (Redding and Rossi-Hansberg, 2017). We note that such idiosyncratic heterogeneity acts as a *local dispersion force*. It is a well-known fact that random utility models can be represented on the basis of deterministic utility (Anderson et al., 1992). Suppose that the idiosyncratic payoff function is defined by $\hat{v}_{ni}(\boldsymbol{x}) = \epsilon_{ni}\tilde{v}_i(\boldsymbol{x})$ where ϵ_{ni} is a random payoff shock for an individual agent n for choosing region i that are independent and identically distributed (i.i.d.) according to a Frèchet distribution, and $\tilde{v}_i(\boldsymbol{x})$ is the homogeneous component; a spatial equilibrium for this case is defined by $x_i = \Pr_i(\boldsymbol{x})$ where $\Pr_i(\boldsymbol{x}) \equiv \Pr\left(i = \arg\max_{j\in\mathcal{N}} \hat{v}_{nj}(\boldsymbol{x})\right) \in (0,1)$ denotes the probability for an agent to choose region i when the current spatial distribution is $\boldsymbol{x}$. Then, there is a *deterministic* (or homogeneous) payoff function $\boldsymbol{v}(\boldsymbol{x}) = (v_i(\boldsymbol{x}))_{i\in\mathcal{N}}$, associated with the *stochastic* (or heterogeneous) payoff function $(\hat{v}_{ni}(\boldsymbol{x}))$, such that $\boldsymbol{x}^*$ is a deterministic spatial equilibrium under $\boldsymbol{v}(\boldsymbol{x})$ if and only if it satisfies $x_i^* = \Pr_i(\boldsymbol{x}^*)$. In this sense, the two spatial equilibrium concepts are "isomorphic" in terms of equilibrium spatial distribution of agents. Local stability of equilibria under this kind of perturbed version of equilibrium condition can be investigated by the associated *perturbed best response dynamics*. See Sandholm (2010) for a unified discussion.

Let $\tilde{\mathbf{V}} \equiv \nabla \tilde{v}(\bar{x})$ be the Jacobian matrix of the homogeneous component of heterogeneous payoff. Then, $\tilde{\mathbf{V}}$ and the Jacobian matrix of deterministic version of payoff function, $\mathbf{V} = \nabla v(\bar{x})$, are connected via the relationship

$$\mathbf{V} = \tilde{\mathbf{V}} - \eta \mathbf{I}, \tag{C.2}$$

where η is a constant which is proportional to the dispersion parameter for the distribution ϵ_{ni} are drawn from. The formula implies that η appears as a *negative constant term* in the net gain function $G^\sharp$ for v. That is, idiosyncratic payoff shock acts as a local dispersion force. This is a natural consequence of assuming that ϵ_{ni} is i.i.d. over n and i. The idiosyncratic payoff shock acts as some kind of dispersion force, but it has no connection to the underlying geography. As such, introducing idiosyncratic payoff shocks to a Class I model can, in effect, change the model to Class III. ■

Remark C.3. On **Proposition 1**, We consider stability under *myopic dynamics*. An important venue of extension is to consider the forward-looking behavior of mobile agents in the spirit of Krugman (1991a). A rigorous theory on this issue can be found in Oyama (2009a,b). The papers explored an economic geography model subject to a deterministic *perfect foresight dynamic*, in which agents have a *complete* anticipation capability for the future; it is shown that the forward-looking behavior also drive the spatial distribution towards a state that is also locally stable under myopic dynamics when the future discount rate is high. That said, employing myopic dynamics for equilibrium refinement can also be interpreted as an approximation of forward-looking dynamics with a high discount rate.

Related to this, a recent literature on geography and development features an explicitly dynamic decision of mobile agents in the (discrete) time axis, with the anticipation capability of agents is supposed to be limited in favor of tractability (i.e., the discount rate for future utility is high) (Desmet and Rossi-Hansberg, 2009, 2014; Nagy, 2017; Desmet et al., 2018). One might be interested in how these dynamic models can be related to deterministic myopic dynamics. In this context, we note that "myopic" dynamics are interpreted as the average behavior of the behavioral assumptions imposed on agents' strategy switching protocol, or "revision protocol" (Sandholm, 2010). The question is, then, what is the average aggregate behavior induced by a dynamic economic geography model when we interpret the agent's dynamic choice in the time axis as the revision protocol of a hypothetical myopic dynamic. A similar discussion applies to overlapping generation models (e.g., Allen and Donaldson, 2018). It requires a model-by-model investigation and calls for another theory of independent interest. ■

Remark C.4. On admissible dynamics defined in Appendix A.1, the conditions (RS) and (PC) are, respectively, called *restricted stationality* and *positive correlation* (Sandholm, 2010), which are the most minimal assumptions we can impose on a dynamic f to be "consistent" with the underlying model v. Also, the symmetry assumption (Sym) ensures that the dynamic does not feature ex-ante preference over alternatives $\mathcal{N}$. We assume f is defined for all nonnegative orthant $\mathbb{R}^N_{\geq 0}$ to avoid unnecessary technical complication. Also, we suppose f is C^1 only because we employ linear stability as the definition of stability.

Admissible C^1 myopic dynamics specified by the conditions (RS), (PC), and (Sym) include, for instance, the *Brown–von Neumann–Nash dynamic* (Brown and von Neumann, 1950; Nash, 1951), the *Smith dynamic* (Smith, 1984), and *Riemannian game dynamics* (Mertikopoulos and Sandholm, 2018) that satisfy (Sym), e.g., the *Euclidian projection dynamic* (Dupuis and Nagurney, 1993) and the *replicator dynamic* (Taylor and Jonker, 1978). See Sandholm (2010) for more examples. Also, we note that **Proposition 1** can be extended to include the best response dynamic (Gilboa and Matsui, 1991), which is generally nondifferentiable. ■

Remark C.5. Define $\mathbf{D} = [\phi_{ij}]$ by $\phi_{ij} = \phi^{\ell_{ij}}$ with $\ell_{ij} \equiv \{|i-j|, N - |i-j|\}$. The eigenvalues $\{\chi_k\}$ of $\bar{\mathbf{D}}$ are given by the following lemma.

Lemma C.1. Assume that N is an even and let $M \equiv \frac{N}{2}$. Define

$$\Psi_k(\phi) \equiv \frac{1-\phi^2}{1-2\phi\cos[\theta k]+\phi^2} \quad \text{and} \quad \bar{\Psi}(\phi) \equiv \frac{1+\phi^M}{1-\phi^M} \tag{C.3}$$

with $\theta = \frac{2\pi}{N}$. Then,

$$\chi_k(\phi) = \begin{cases} \Psi_k(\phi)\Psi_M(\phi) & (k\text{: even}) \\ \Psi_k(\phi)\Psi_M(\phi)\bar{\Psi}(\phi) & (k\text{: odd}) \end{cases} \qquad k = 0, 1, 2, \ldots, M. \tag{C.4}$$

See Lemma 4.2 of Akamatsu et al. (2012). ■

Remark C.6. The following lemma is useful for characterizing the stability of $\bar{x}$. Notations are the same as Appendix A.1.

Lemma C.2. $\bar{x}$ is linearly stable under *all* admissible dynamics if and only if:

$$z^\top \mathbf{V} z < 0 \qquad \forall z \in T\mathcal{X} \setminus \{\mathbf{0}\} \tag{CND}$$

where $T\mathcal{X} \equiv \{z \in \mathbb{R}^N \mid \mathbf{1}^\top z = 0\}$.

Proof. From Fact E.3, $\mathbf{V}$ has $\frac{N}{2}+1$ distinct eigenvalues $\{\omega_k\}_{k=0}^{\frac{N}{2}}$. ω_0 is associated to $\mathbf{1}$, which is orthogonal to $T\mathcal{X}$. Thus, (CND) is equivalent to $\omega_k < 0$ for all $k \neq 0$. Then, from (A.3), $\eta_k < 0$ for all $k \neq 0$, which is the definition of linear stability of $\bar{x}$. ■

Sections 2.2, 2.3, and 4 builds on **Lemma C.2**. For instance, (CND) is equivalent to $\omega < 0$ in the two-region setup. ■

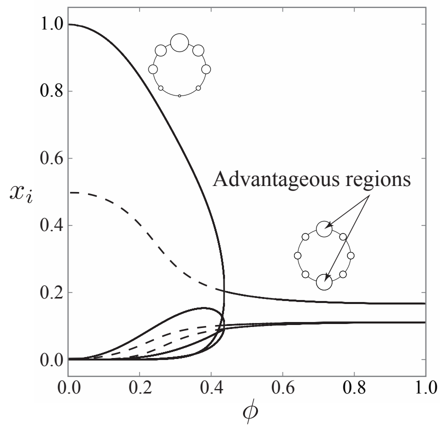


Figure C.1: Bifurcation in Class II model with exogenous heterogeneity

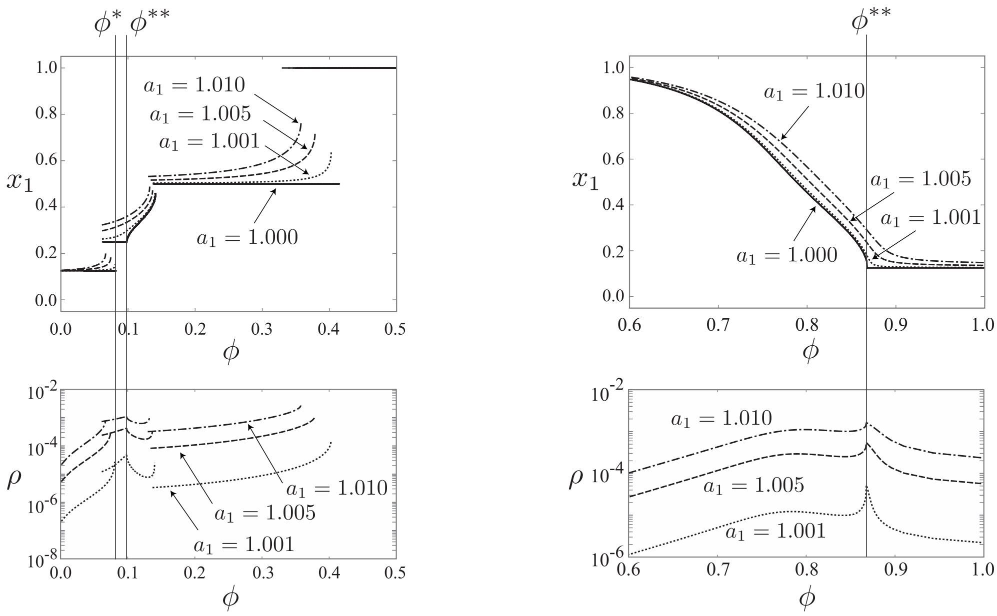
Notes: The black solid curves indicate stable equilibria, whereas the dashed curves unstable equilibria. For the range $\phi \in (\phi^{**}, 1)$ ($\phi^{**} \approx 0.43$), there are two large cities due to exogenous advantages. However, when endogenous agglomeration occurs, the spatial distribution become single-peaked, as **Proposition 1** (b) predicts.

Remark C.7. In Appendix A.2, the expression (A.18) provides an intuition for what pattern can emerge in the presence of small heterogeneities in local characteristics. Suppose that ϕ_k^* is a value of ϕ at which $\bar{x}$ changes its stability from stable to unstable under complete symmetry, resulting in a deviation towards k-cities direction ($k = 1$ or $k = \frac{N}{2}$). By definition, $G(\chi_k(\phi_k^*)) \approx 0$ near such ϕ_k^*. Combined with the condition $G^\natural(\chi_k) > 0$, it implies that λ_k become infinitely large as ϕ approaches to ϕ_k^*; if $\boldsymbol{a} = \bar{a}\mathbf{1} + \tilde{a}_k z_k$, then linearization $\boldsymbol{x}(\boldsymbol{a}) \sim \bar{\boldsymbol{x}} + \mathbf{X}(\boldsymbol{a} - \bar{\boldsymbol{a}}) = \bar{\boldsymbol{x}} + \tilde{a}_k \lambda_k z_k$ predicts that spatial pattern $\boldsymbol{x}(\boldsymbol{a})$ almost "kinks" towards k-cities pattern. This is another manifestation of **Proposition 1** in that the instability of $\bar{x}$ at ϕ_k^* implies the formation of k cities.

That said, small heterogeneities do not affect the predictions of **Proposition 1**. Figure C.1 illustrates that although local advantage can induce "multipolar" distribution in a Class II model, such spatial pattern vanish to form a single megacity when endogenous agglomeration force matters. The topic here is related to the so-called *universal unfolding* and there is an enormous body of general theory that justifies and generalizes the casual observation we have drawn here. See Golubitsky and Stewart (2003), Golubitsky and Schaeffer (2012), and Golubitsky et al. (2012). See also Ikeda and Murota (2014) for applications. ■

Remark C.8. Figure C.2 depicts the population share of region 1 under asymmetry as considered in Figure 14 in the main text. Basic model parameters are the same as Figure 10 and Figure 11 except that region 1 has exogenous advantage. Figure C.2a is a reproduction of Figure 14b, whereas Figure C.2b considers the AA model under multiplicity of equilibria. For both figures, we see that **Proposition 2** correctly predicts the sign of $\rho'(\phi)$ for the range of ϕ such that $\bar{x}$ is stable when $a_1 = 1$. We have $\rho'(\phi) > 0$ when $\phi \in (0, \phi^*)$ for the Km model, whereas $\rho'(\phi) < 0$ when $\phi \in (\phi^{**}, 1)$ for the AA model.

For Figure C.2a, we modify the definition of ρ for spatial patterns with unpopulated regions.



(a) The Krugman model

(b) The Allen–Arkolakis model

Figure C.2: Population share of region 1 and covariance ρ

For instance, for the range $\phi \in (\phi^*, \phi^{**})$, we define ρ with respect to the four-centric pattern $(2\bar{x}, 0, 2\bar{x}, 0, 2\bar{x}, 0, 2\bar{x}, 0)$:

$$\rho \equiv \sum_{i \in \mathcal{I}(\boldsymbol{x})} (x_i - 2\bar{x})(a_i - \bar{a}(\boldsymbol{x})), \tag{C.5}$$

where $\mathcal{I}(\bar{\boldsymbol{x}}) = \{i \in \mathcal{I} \mid x_i > 0\}$ is the set of populated regions, and $\bar{a}(\boldsymbol{x}) \equiv \frac{1}{|\mathcal{I}(\bar{\boldsymbol{x}})|} \sum_{i \in \mathcal{I}(\boldsymbol{x})} a_i$. We define ρ for two-centric pattern $(4\bar{x}, 0, 0, 0, 4\bar{x}, 0, 0, 0)$ in a similar way. For the transitional phase after ϕ^{**} we let

$$\rho \equiv \sum_{i \in \mathcal{I}(\boldsymbol{x})} (x_i - x_i^*)(a_i - \bar{a}(\boldsymbol{x})), \tag{C.6}$$

where x_i^* corresponds to the stable solution for $a_1 = 1$. Note that $\rho = (x_1 - 1)(a_1 - a_1) = 0$ for the complete monopolar pattern $(1, 0, 0, 0, 0, 0, 0, 0)$. It is natural that we have $\rho'(\phi) > 0$ for the four- and two-centric patterns because these patterns can be regarded as the uniform distribution on the four- and two-region cases, respectively.

For Figure C.2b, we employ (C.6) as the definition of ρ for the case $\phi \in (0, \phi^{**})$, i.e., we consider the deviation from the baseline equilibrium ($a_1 = 1$). We observe that $\rho'(\phi) < 0$ does not necessarily hold true for $\phi \in (0, \phi^{**})$. In particular, $\rho'(\phi) > 0$ when local dispersion force is relatively weak (when ϕ is small). ■

Remark C.9. Many economic geography models are subject to multiple proximity matrices and/or different degrees of increasing returns. We here discuss three major categories of economic geography models that can be studied by imposing racetrack assumptions.

One major strand of research in this line aims to explain the formation of the (possibly multiple) business districts together with residential land use and commuting patterns within a city. These intra-city models typically distinguish location behavior of firms and households (e.g., Fujita and Ogawa, 1982; Ota and Fujita, 1993; Lucas and Rossi-Hansberg, 2002; Ahlfeldt et al., 2015; Owens et al., forthcoming; Heblich et al., 2018; Osawa and Akamatsu, 2019).

Another possibility is to consider different transport cost structures by industry. For example, Fujita and Krugman (1995) introduced transport costs for land-intensive rural goods along with those of urban goods. In the presence of rural goods that are costly to transport, the delivered price for such goods is lower in regions farther away from the agglomerations, which generates a dispersion force. This is similar to the local dispersion force in that even a small deviation from an urban agglomeration will decrease the price of rural goods and increase the payoff of the deviant. However, the advantage of dispersion persists outside the agglomeration, i.e., it depends on the distance structure of the model. This type of dispersion force is known to result in the formation of an *industrial belt*, a continuum of agglomeration associated with multiple atoms of agglomeration as demonstrated by the simulations in Mori (1997) and Ikeda et al. (2017b). The formal characterization of industrial belts remains to be carried out.

The last relevant direction of research is the formalization of the classical central place theory of Christaller (1933) which investigates the diversity in sizes together with the spatial patterns of cities (e.g., Fujita et al., 1999b; Tabuchi and Thisse, 2011). It is an extension of Class I models by multiple industries subject to different degrees of increasing returns and/or transport costs.[36] As in Class I models, agglomerations of each industry are spaced apart from one another, but their spacing is larger for industries with greater increasing returns. The key in these central place models is that industries tend to co-agglomerate as they share demand externaltities through common consumers. As a consequence, there is a hierarchical structure in industrial agglomeration pattern: more localized industries (with greater increasing returns) tend to co-agglomerate with more ubiquitous ones (with smaller increasing returns). Since larger cities are formed at locations in which a larger number of industries co-agglomerate, the size diversity and spatial patterns of cities are determined by the spatial coordination of industries.

Since the endogenous mechanism is the key in both these types of models, a racetrack geography provides an ideal setup. Initial such explorations are found in Tabuchi and Thisse (2011) and Osawa and Akamatsu (2019). ■

[36] Hsu (2012) proposes an alternative formalization of central place theory in the context of spatial competition and firm entry. Davis and Dingel (2019) offer an alternative mechanism of spatial coordination among industries which in turn results in hierarchy principle and the diversity in city sizes in the context of a systems-of-cities model that abstracts from inter-city space.

D On relaxing the racetrack assumption

This section highlights the implications of **Proposition 1** qualitatively generalize to various geographical set-ups (line segment, square and hexagonal lattices with/without boundaries). In particular, the *polarity* of endogenous spatial patterns in each model class is unaffected; multiple cities can endogenously form in Class I, whereas a single city form in Class II, when we relax Assumption RE.

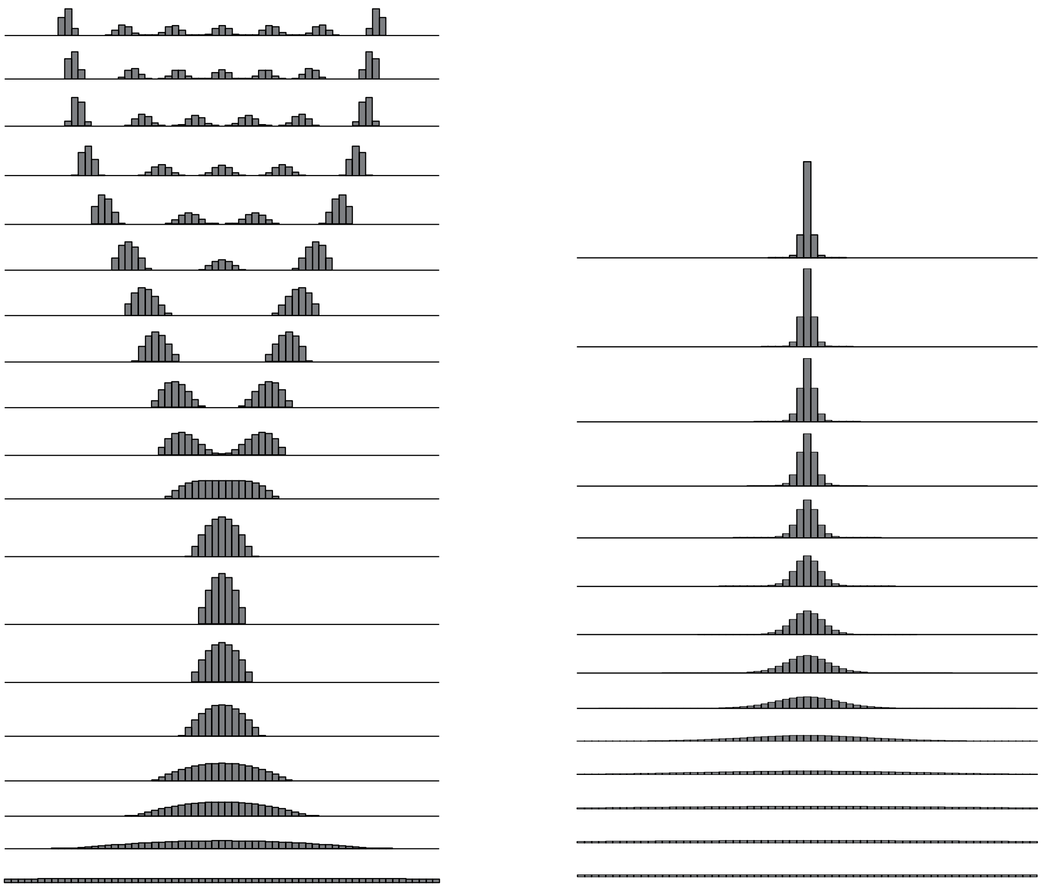
The simplest way to introduce geographical asymmetry into our one-dimensional set-up is to consider a bounded line segment, which is a standard spatial setting in urban economic theory. For instance, Ikeda et al. (2017b) considered a Class I model by Forslid and Ottaviano (2003) in a line segment and showed that multiple cities form in the set-up. It is also shown that the evolution of spatial structure on such geography follows a "period doubling" behavior, which is formally discussed for racetrack economy (Akamatsu et al., 2012; Osawa et al., 2017), as in Figure 10. For Class II and III, Figure D.1 shows endogenous agglomeration patterns in the Hm and PS models. Observe that for both the models, qualitative properties of the spatial patterns are consistent with those for circular geography (Section 5).

Also, Mossay and Picard (2011) considered a variant of Beckmann (1976)'s model (Class II) and showed that the only possible equilibrium is a unimodal distribution, as in Figure 11.[37] The numerical results of Anas and Kim (1996) and Anas et al. (1998) in line segments bear close resemblance to, respectively, agglomeration behaviors of Class I and II models.

The real-world geography is two-dimensional. The two-dimensional counterpart of the racetrack economy is a bounded lattices with periodic boundary conditions (i.e., flat torus). Ikeda et al. (2017a) and Ikeda et al. (2018) respectively considered a Class I model in two-dimensional hexagonal and square lattices. In both lattices, it is shown that multiple cities form and period-doubling behavior emerge from the model, as in the racetrack set-up. See, in particular, Ikeda et al. (2018) for detailed comparison of one-dimensional racetrack economy and two-dimensional square lattice economy. As concrete examples, Figure D.2 shows endogenous equilibrium spatial patterns over a bounded square economy with $9 \times 9 = 81$ regions in the course of increasing ϕ for the Krugman and Allen–Arkolakis models. The parameters are the same as Figure 10 and Figure 11. As **Proposition 1** and Section 5 predict, the Krugman model (Class I) engender multiple disjointed cities. When ϕ increases, the number of cities gradually decreases, while the spacing between them enlarges. For the AA model (Class II), in contrast, the spatial pattern is initially monopolar, i.e., there is a single big city. As ϕ increases, the city gradually flattens due to suburbanization. These behaviors are qualitatively consistent with **Proposition 1** and examples in Section 5.

Also, Blanchet et al. (2016) considered a Class II model over a two-dimensional space and showed that, for the Bm model, equilibrium spatial pattern is unique and given by a concave regular paraboloid, i.e., an "unimodal" pattern. Picard and Tabuchi (2013) also considered a Class II

[37] Mossay and Picard (2011) considered a continuous line segment, in contrast to this paper and Ikeda et al. (2017b). As shown by Akamatsu et al. (2017a), the model by Mossay and Picard (2011) can be considered as a continuous limit of an appropriate discrete-space model.

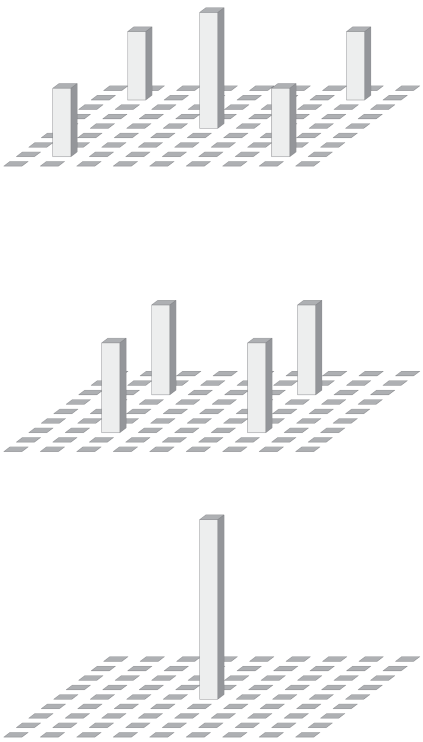


(a) The PS model (b) The Hm model

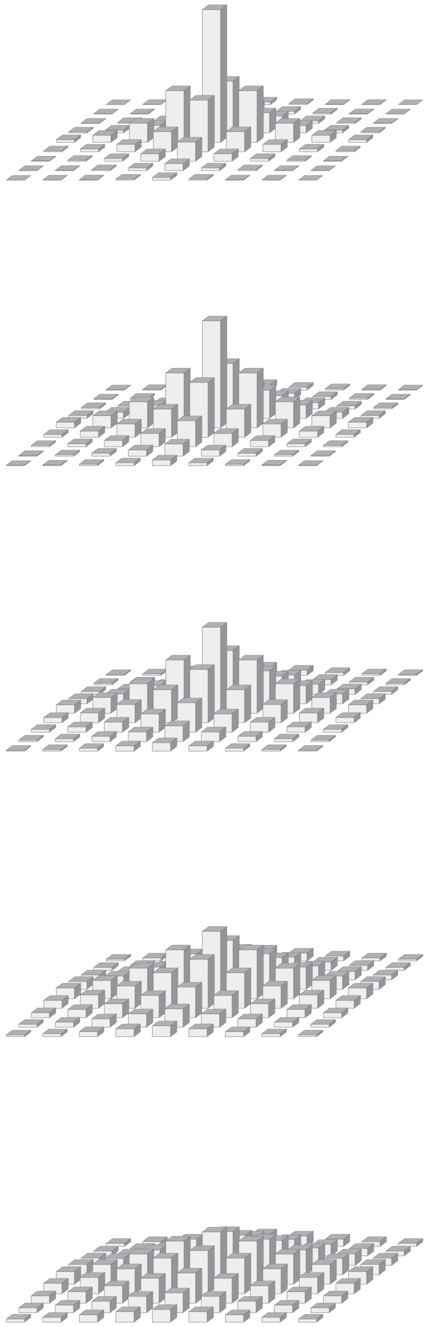
Figure D.1: Endogenous agglomeration patterns in a line segment with 65 regions

general equilibrium model in a two-dimensional space and showed that spatial distribution become unimodal.[38]

[38]Notably, the implications of **Proposition 1** seem to extend to non-iceberg transport costs that are not covered by the specification of **D** in Assumption RE. In fact, Mossay and Picard (2011); Picard and Tabuchi (2013); Blanchet et al. (2016) assume *linear* transport costs.



(a) The Krugman model



(b) The Allen–Arkolakis model

Figure D.2: Endogenous agglomeration patterns in a bounded square economy

E Preliminaries from matrix analysis

This section collects the relevant facts from matrix analysis for self-containedness. For a concise reference, see Horn and Johnson (2012) (henceforth HJ). First, we recall that when we know the eigenpairs (eigenvalue–eigenvector pairs) of a square matrix $\mathbf{A}$, we know those of matrix polynomials based on $\mathbf{A}$ (see HJ, Section 1.1).

Fact E.1. Let $\mathbf{A}$ be a square matrix with eigenpairs $\{(\lambda_k, \boldsymbol{z}_k)\}_{k\in\mathcal{K}}$. For a finite-degree polynomial $P(t) = \sum_{l=0}^{n} p_l t^l$, let $P(\mathbf{A})$ be defined by $P(\mathbf{A}) = \sum_{l=0}^{n} p_l \mathbf{A}^l$ with $\mathbf{A}^0 \equiv \mathbf{I}$. Then, the eigenpairs of $P(\mathbf{A})$ are given by $\{(P(\lambda_k), \boldsymbol{z}_k)\}_{k\in\mathcal{K}}$. Take another finite-degree polynomial $Q(t)$. If $Q(\mathbf{A})$ is nonsingular, then the eigenpairs of $Q(\mathbf{A})^{-1}$ are given by $\{(Q(\lambda_k)^{-1}, \boldsymbol{z}_k)\}_{k\in\mathcal{K}}$. Thus, the eigenpairs of the matrix $G(\mathbf{A}) \equiv Q(\mathbf{A})^{-1}P(\mathbf{A})$ is given by $\{(G(\lambda_k), \boldsymbol{z}_k)\}_{k\in\mathcal{K}}$ with $G(t) = \frac{P(t)}{Q(t)}$. ♢

Next, a *circulant matrix* $\mathbf{C}$ of size N generated by $\boldsymbol{c} = (c_i)_{i=0}^{N-1}$ is

$$\mathbf{C} = \mathrm{circ}[\boldsymbol{c}] \equiv \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{N-2} & c_{N-1} \\ c_{N-1} & c_0 & c_1 & c_2 & \cdots & c_{N-2} \\ c_{N-2} & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & c_2 \\ c_2 & \cdots & c_{N-2} & c_{N-1} & c_0 & c_1 \\ c_1 & c_2 & \cdots & c_{N-2} & c_{N-1} & c_0 \end{bmatrix}. \tag{E.1}$$

Each row of $\mathbf{C}$ are identical to the previous row moved one position to the right and wrapped around. Every row sum equals to $\boldsymbol{c}^\top \mathbf{1}$ by definition. Circulant matrices are known to satisfy the following properties (see HJ, Section 0.9.6 and Problem 2.2.P10):

Fact E.2. Circulant matrices of size N form a commutative algebra: linear combinations and products of circulants ar circulants; the inverse of a nonsingular circulant is a circulant; any two circulants of the same size commute. ♢

Fact E.3. Let $\mathbf{C} = \mathrm{circ}[\boldsymbol{c}]$ be a real and *symmetric* circulant matrix of size N. Then, $\mathbf{C}$ is diagonalized by an orthogonal matrix $\mathbf{Z}$: $\mathrm{diag}[\boldsymbol{\lambda}] = \mathbf{Z}^\top \mathbf{C}\mathbf{Z}$. The column vectors of the matrix $\mathbf{Z}$ are the eigenvectors of $\mathbf{C}$. Let $\theta \equiv \frac{2\pi}{N}$. Eigenpairs $(\lambda_k, \boldsymbol{z}_k)$ can be chosen to be

$$\lambda_0 = \boldsymbol{c}^\top \mathbf{1}, \qquad \boldsymbol{z}_0 \equiv \langle 1 \rangle_{i=0}^{N-1}, \tag{E.2}$$

$$\lambda_k, \qquad \begin{cases} \boldsymbol{z}_k^+ \equiv \langle \cos(\theta k i)\rangle_{i=0}^{N-1}, \\ \boldsymbol{z}_k^- \equiv \langle \sin(\theta k i)\rangle_{i=0}^{N-1}, \end{cases} \qquad k = 1, 2, \ldots, \lfloor \tfrac{N}{2} \rfloor - 1, \tag{E.3}$$

$$\lambda_{\frac{N}{2}}, \qquad \boldsymbol{z}_{\frac{N}{2}} \equiv \langle (-1)^i \rangle_{i=0}^{N-1}, \qquad \text{if } N \text{ is an even,} \tag{E.4}$$

where $\langle z_i \rangle_{i=0}^{N-1} \equiv \|\boldsymbol{z}\|^{-1}(z_i)_{i=0}^{N-1}$ denotes the normalized version of real vector $\boldsymbol{z}$. Thus, the distinct eigenvalues of $\mathbf{C}$ are given by $\boldsymbol{\lambda} = \{\lambda_k\}_{k\in\mathcal{K}}$ $(k \in \{0, 1, 2, \ldots, \lfloor \frac{N}{2} \rfloor\})$. Those eigenvalues with

$k = 1, 2, \ldots, \lfloor \frac{N}{2} \rfloor - 1$ are multiplicity two. $\boldsymbol{z}_0$ is a uniform vector and λ_0 is the row-sum of $\mathbf{C}$; the elements of other eigenvectors sum up to zero. If in addition $c_i > 0$ for all i and $\boldsymbol{c}^\top \mathbf{1} = 1$, then $\mathbf{C} = \text{circ}[\boldsymbol{c}]$ is positive and row-stochastic. We have $\mathbf{C1} = \mathbf{1}$ and thus $\lambda_0 = 1$; λ_0 is the maximal eigenvalue (or the spectral radius) of $\mathbf{C}$ and $\mathbf{1}$ is the only strictly positive eigenvector (the Perron–Frobenius theorem). Note that *all* real symmetric circulant matrices of size N share the same set of eigenvectors. For general (possibly asymmetric) circulant matrices, discrete Fourier transformation matix can be employed for diagonalization (see, e.g., Akamatsu et al., 2012, for an application). $\Diamond$

Remark E.1. Under Assumption RE, $\mathbf{D}$ is a circulant matrix because $\phi_{ij} = \phi^{\ell_{ij}} = \phi^{\ell_{i+1,j+1}} = \phi_{i+1,j+1}$ for all i, j (mod N for indices). ■

F Supporting computations

F.1 General observations

This section summarizes general relationships between various partial derivatives. Throughout this section, we let $\mathbf{F}_x = [\frac{\partial f_i}{\partial x_j}]$ denote the partial derivative of a vector-valued function $\boldsymbol{f}(\boldsymbol{x})$ with respect to the variable $\boldsymbol{x}$. For instance, $\mathbf{V}_x \equiv [\frac{\partial v_i}{\partial x_j}]$, $\tilde{\mathbf{V}}_x \equiv [\frac{\partial \tilde{v}_i}{\partial x_j}]$, $\tilde{\mathbf{V}}_w \equiv [\frac{\partial \tilde{v}_i}{\partial w_j}]$, $\mathbf{S}_x \equiv [\frac{\partial s_i}{\partial x_j}]$, $\mathbf{S}_w \equiv [\frac{\partial s_i}{\partial w_j}]$, and $\mathbf{W}_x \equiv [\frac{\partial w_i}{\partial x_j}]$.

The matrix $\mathbf{V} = \frac{\bar{x}}{\bar{v}}\mathbf{V}_x$. The payoff functions for most of the models we referenced in the main text reduce to the following form:

$$\boldsymbol{v}(\boldsymbol{x}) = \tilde{\boldsymbol{v}}(\boldsymbol{x}, \boldsymbol{w}), \tag{F.1}$$

$$\boldsymbol{s}(\boldsymbol{x}, \boldsymbol{w}) = \mathbf{0}. \tag{F.2}$$

The condition (F.2) represents the market equilibrium conditions for a given $\boldsymbol{x}$ that defines $\boldsymbol{w}$ as an implicit function of $\boldsymbol{x}$. For $\boldsymbol{v}(\boldsymbol{x})$ to be well-defined, (F.2) must admit a unique solution of $\boldsymbol{w}$ for all $\boldsymbol{x} \in \mathcal{X}$. We assume that (F.2) has a unique solution for all $\boldsymbol{x} \in \mathcal{X}^\circ$, where $\mathcal{X}^\circ \equiv \{\boldsymbol{x} \in \mathcal{X} \mid x_i > 0\ \forall i \in \mathcal{N}\}$ denote the interior of $\mathcal{X}$. In general, we have

$$\mathbf{V}_x = \tilde{\mathbf{V}}_x + \tilde{\mathbf{V}}_w \mathbf{W}_x, \tag{F.3}$$

$$\mathbf{W}_x = -\mathbf{S}_w^{-1}\mathbf{S}_x, \tag{F.4}$$

where $\mathbf{W}_x$ is obtained by the implicit function theorem regarding (F.2). The inverse $\mathbf{S}_w^{-1}$ exists for all $\boldsymbol{x} \in \mathcal{X}^\circ$ under our premise that $\boldsymbol{w}(\boldsymbol{x})$ exists.

If $\boldsymbol{x} = \bar{\boldsymbol{x}}$, $\mathbf{V}_x = \mathbf{S}_w^{-1}(\mathbf{S}_w\tilde{\mathbf{V}}_x - \tilde{\mathbf{V}}_w\mathbf{S}_x)$, since all matrices commute because they are real, symmetric, and circulant at $\bar{\boldsymbol{x}}$ (Fact E.2). In fact, $G^\flat(\bar{\mathbf{D}})$ in the Km and Helpman models arises from $\mathbf{S}_w$ and represents general equilibrium effects through (F.2). In this way, for any model whose payoff function reduces to the equations of the form (F.1) and (F.2), $\mathbf{V} = Q(\bar{\mathbf{D}})^{-1}P(\bar{\mathbf{D}})$ where polynomials $P(t)$ and $Q(t)$ are chosen such that $P(\bar{\mathbf{D}}) = \mathbf{S}_w\tilde{\mathbf{V}}_x - \tilde{\mathbf{V}}_w\mathbf{S}_x$ and $Q(\bar{\mathbf{D}}) = \mathbf{S}_w$.

Example F.1. In Examples 2.2 and 2.3, (F.2) is given by

$$s_i(\boldsymbol{x}, \boldsymbol{w}) = w_i x_i - \sum_{j\in\mathcal{N}} m_{ij} e_j = 0, \tag{F.5}$$

where $e_i = e(w_i, x_i)$ with some nonnegative function e and $\mathbf{M} = [m_{ij}]$ is defined by

$$m_{ij} = \frac{x_i w_i^{1-\sigma}\phi_{ij}}{\sum_{k\in\mathcal{N}} x_k w_k^{1-\sigma}\phi_{kj}}. \tag{F.6}$$

In matrix form, we may write (F.5) as $\boldsymbol{y} - \mathbf{M}\boldsymbol{e} = \mathbf{0}$. It gives

$$\mathbf{S}_x = \operatorname{diag}[\boldsymbol{w}] - \left(\operatorname{diag}[\mathbf{M}\boldsymbol{e}] - \mathbf{M}\operatorname{diag}[\boldsymbol{e}]\mathbf{M}^\top\right)\operatorname{diag}[\boldsymbol{x}]^{-1} - \mathbf{M}\mathbf{E}_x, \tag{F.7a}$$

$$\mathbf{S}_w = \operatorname{diag}[\boldsymbol{x}] + (\sigma - 1)\left(\operatorname{diag}[\mathbf{M}\boldsymbol{e}] - \mathbf{M}\operatorname{diag}[\boldsymbol{e}]\mathbf{M}^\top\right)\operatorname{diag}[\boldsymbol{w}]^{-1} - \mathbf{M}\mathbf{E}_w. \tag{F.7b}$$

Suppose Assumptions RE and S. Consider uniform distribution $\bar{\boldsymbol{x}}$ and let $\bar{w}$ be the uniform level of wage at $\bar{\boldsymbol{x}}$. Note that $\mathbf{M} = \bar{\mathbf{D}}$ when $\boldsymbol{x} = \bar{\boldsymbol{x}}$. Let $\mathbf{E}_x = \epsilon_x \bar{w}\mathbf{I}$ and $\mathbf{E}_w = \epsilon_w \bar{x}\mathbf{I}$ at $\bar{\boldsymbol{x}}$. Let $\bar{e} = e(\bar{w}, \bar{x})$ and $\zeta \equiv \frac{\bar{e}}{\bar{w}\bar{x}}$. We see that

$$\mathbf{S}_x = -\bar{w}\left((\zeta - 1)\mathbf{I} + \epsilon_x\bar{\mathbf{D}} - \zeta\bar{\mathbf{D}}^2\right), \tag{F.8a}$$

$$\mathbf{S}_w = \bar{x}\left((1 + \zeta(\sigma - 1))\,\mathbf{I} - \epsilon_w\bar{\mathbf{D}} - \zeta(\sigma - 1)\bar{\mathbf{D}}^2\right). \tag{F.8b}$$

For instance, if $e(w_i, x_i) = w_i x_i$, we see $\mathbf{W}_x = \frac{\bar{w}}{\bar{x}}(\sigma\mathbf{I} + (\sigma - 1)\bar{\mathbf{D}})^{-1}\bar{\mathbf{D}}$ at $\bar{\boldsymbol{x}}$. ■

The matrix $\mathbf{A} = \frac{\bar{x}}{\bar{a}}\mathbf{V}_a$. Note that $\mathbf{X} = [\partial x_i(\bar{\boldsymbol{a}})/\partial a_i] = \mathbf{X}_a$ in (A.14) acts essentially as $\hat{\mathbf{X}} \equiv -\mathbf{V}_x^{-1}\mathbf{V}_a$ for $\boldsymbol{z}$ such that $\boldsymbol{z}^\top\mathbf{1} = \mathbf{0}$. Thus, $\mathbf{V}_a$ is of interest.

For purely local regional characteristics (Example 6.1), since $v_i(\boldsymbol{x}, \boldsymbol{a}) = a_i v_i(\boldsymbol{x})$, it follows that $\mathbf{V}_a = \operatorname{diag}[\boldsymbol{v}(\boldsymbol{x})]$. At $\bar{\boldsymbol{x}}$, we have $\mathbf{V}_a = \bar{v}\mathbf{I}$. Thus, $\hat{\mathbf{X}} = -\bar{v}\mathbf{V}_x^{-1}$.

For regional characteristics that affect trade flows (Examples 6.3 and 6.4), the payoff function and the market equilibrium condition are respectively modified to $\boldsymbol{v}(\boldsymbol{x}, \boldsymbol{a}) = \tilde{\boldsymbol{v}}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{a})$ and $\boldsymbol{s}(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{a}) = \mathbf{0}$ with some $\boldsymbol{a}$. Then, by applying the implicit function theorem to the modified equation, we see $\mathbf{V}_a = \tilde{\mathbf{V}}_a + \tilde{\mathbf{V}}_w\mathbf{W}_a = \tilde{\mathbf{V}}_a - \tilde{\mathbf{V}}_w\mathbf{S}_w^{-1}\mathbf{S}_a$. Thus, it is equivalent to consider

$$\hat{\mathbf{X}} = -\left(\tilde{\mathbf{V}}_x - \tilde{\mathbf{V}}_w\mathbf{S}_w^{-1}\mathbf{S}_x\right)^{-1}\left(\tilde{\mathbf{V}}_a - \tilde{\mathbf{V}}_w\mathbf{S}_w^{-1}\mathbf{S}_a\right) \tag{F.9}$$

$$= \left(\mathbf{S}_w\tilde{\mathbf{V}}_x - \tilde{\mathbf{V}}_w\mathbf{S}_x\right)^{-1}\left(\tilde{\mathbf{V}}_w\mathbf{S}_a - \mathbf{S}_w\tilde{\mathbf{V}}_a\right), \tag{F.10}$$

where we also note that all matrices commute under Assumption RE.

Example F.2. In Example 6.3, we have

$$s_i(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{a}) = w_i x_i - \sum_{j\in\mathcal{N}} \frac{x_i a_i w_i^{1-\sigma}\phi_{ij}}{\sum_{k\in\mathcal{N}} x_k a_k w_k^{1-\sigma}\phi_{kj}} e_j = 0, \tag{F.11}$$

which induces $\mathbf{S}_a = -\left(\operatorname{diag}[\mathbf{M}\boldsymbol{y}] - \mathbf{M}\operatorname{diag}[\boldsymbol{y}]\mathbf{M}^\top\right)\operatorname{diag}[\boldsymbol{a}]^{-1}$ or, at the uniformity $\boldsymbol{x} = \bar{\boldsymbol{x}}$, $\mathbf{S}_a = -\frac{\bar{e}}{\bar{a}}\left(\mathbf{I} - \bar{\mathbf{D}}^2\right) = -\frac{\bar{e}}{\bar{a}}\left(\mathbf{I} - \bar{\mathbf{D}}\right)\left(\mathbf{I} + \bar{\mathbf{D}}\right)$. Section F.2.4 will derive $\delta(\chi)$ for this case. ■

Example F.3. In Example 6.4, we have

$$s_i(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{a}) = w_i x_i - \sum_{j\in\mathcal{N}} \frac{x_i w_i^{1-\sigma}\phi_{ij}}{\sum_{k\in\mathcal{N}} x_k w_k^{1-\sigma}\phi_{kj}} e(w_j, x_j, a_j) = 0 \tag{F.12}$$

where e maps the tuple (w_j, x_j, a_j) to the regional expenditure. Then, we have $\mathbf{S}_a = -\mathbf{M}\mathbf{E}_a$, or, at $\bar{x}$, $\mathbf{S}_a = -\epsilon_a \bar{\mathbf{D}}$ where $\epsilon_a = \frac{\partial e(\bar{x}, \bar{w}, \bar{a})}{\partial a_i}$. We will derive $\delta(\chi)$ for this case in Section F.2.1. ■

F.2 Model-by-model analyses

For self-containedness, this section provides omitted derivations of the net gain functions $G^\sharp(\chi)$ for the examples provided in the main text.[39] Table F.1 at the end of this appendix summarizes the exact mappings from each model to the coefficients of a model-dependent net gain function $G^\sharp(\chi) = c_0 + c_1\chi + c_2\chi^2$. Throughout, $\bar{v}$, $\bar{w}$, $\bar{e}$ and so on represent that they are the values of v_i, w_i, e_i evaluated at $\boldsymbol{x} = \bar{\boldsymbol{x}} = \bar{x}\mathbf{1}$, unless otherwise noted.

F.2.1 Krugman (1991b) model (Examples 2.2 and 6.4)

There are two types of workers, mobile and immobile, with the total endowments of them being 1 and L, respectively, and $\boldsymbol{x} \equiv (x_i)_{i\in\mathcal{N}}$ denotes the spatial distribution of mobile worker. Each worker inelastically supplies one unit of labor.

There are two industrial sectors: agriculture (abbreviated as A) and manufacturing (abbreviated as M). The A-sector is perfectly competitive and a unit input of immobile labor is required to produce one unit of goods. The M-sector is modeled by Dixit–Stiglitz monopolistic competition. M-sector goods are horizontally differentiated and produced under increasing returns to scale using mobile labor as the input. The goods of both sectors are transported. Transportation of A-sector goods is frictionless, while that of M-sector goods is of an iceberg form. For each unit of M-sector goods transported from region i to j, only the proportion $1/\tau_{ij}$ arrives, where $\tau_{ij} > 1$ for $i \neq j$ and $\tau_{ii} = 1$.

All workers share an identical preference for both M- and A-sector goods. The utility of a worker in region i is given by a two-tier form. The upper tier is Cobb–Douglas over the consumption of A-sector goods C_i^{A} and that of M-sector manufacturing constant-elasticity-of-substitution (CES) aggregate C_i^{M} with $\sigma > 1$

$$C_i^{\mathrm{M}} \equiv \left(\sum_{j\in\mathcal{N}} \int_0^{n_j} q_{ji}(\xi)^{\frac{\sigma-1}{\sigma}} \mathrm{d}\xi \right)^{\frac{\sigma}{\sigma-1}}, \tag{F.13}$$

that is, $u_i = (C_i^{\mathrm{A}})^{\mu}(C_i^{\mathrm{A}})^{1-\mu}$ where $\mu \in (0,1)$ is the constant expenditure of the latter. With free trade in the A-sector, the wage of the immobile worker is equalized, and we normalize it to unity by taking A-sector goods as the numéraire. Consequently, region i's expenditure on the M-sector goods is given by $e_i = \mu(w_i x_i + l_i)$ where l_i denotes the mass of immobile workers in region i.

In the M-sector, to produce q units of the differentiated product, a firm requires $\alpha + \beta q$ units of mobile labor. The profit maximization of firms yields the price of differentiated goods produced in region i and exported to j as $p_{ij} = \frac{\sigma\beta}{\sigma-1} w_i \tau_{ij}$, which in turn determines gravity trade flow from j to

[39] Routin derivations are omitted. For detailed derivations, see Akamatsu et al. (2017b).

i. That is, when X_{ij} denotes the price of M-sector goods produced in region i and sold in region j, $X_{ij} = m_{ij} e_j$ where the share $m_{ij} \in (0,1)$ is defined by (F.6) with $\phi_{ij} \equiv \tau_{ij}^{1-\sigma}$. The proximity matrix is thus $\mathbf{D} = [\phi_{ij}] = [\tau_{ij}^{1-\sigma}]$.

Given $\boldsymbol{x}$, we determine the market wage $\boldsymbol{w} \equiv (w_i)_{i\in\mathcal{N}}$ by the M-sector product market-clearing condition, zero-profit condition, and mobile labor market-clearing condition. These conditions are summarized by the trade balance $w_i x_i = \sum_{j\in\mathcal{N}} X_{ij}$, or (F.5) with $e(x_i, w_i) = \mu(w_i x_i + l_i)$. By adding up (F.5) for the Krugman model, we see $\sum_{i\in\mathcal{N}} w_i x_i = \frac{\mu}{1-\mu} L$, which constrains the total income of mobile workers at any configuration $\boldsymbol{x}$. The existence and uniqueness of the solution for (F.5) follow from standard arguments (e.g., Facchinei and Pang, 2007). Given the solution $\boldsymbol{w}(\boldsymbol{x})$ of (F.5), we have the indirect utility of mobile workes is given by $v_i = \Delta_i^{\frac{\mu}{\sigma-1}} w_i$, where $\Delta_i \equiv \sum_{k\in\mathcal{N}} x_k w_k^{1-\sigma} d_{ki}$.

To satisfy Assumption S, let $l_i = l \equiv \frac{L}{N}$ for all $i \in \mathcal{N}$. We have

$$\nabla \log \boldsymbol{v}(\bar{\boldsymbol{x}}) = \underbrace{\frac{\mu}{\sigma-1}\mathbf{M}^\top \operatorname{diag}[\boldsymbol{x}]^{-1}}_{\bar{v}^{-1}\tilde{\mathbf{V}}_x} + \underbrace{\left(\mathbf{I} - \mu\mathbf{M}^\top\right) \operatorname{diag}[\boldsymbol{w}]^{-1}}_{\bar{v}^{-1}\tilde{\mathbf{V}}_w} \mathbf{W}_x \tag{F.14}$$

$$= \frac{1}{\bar{x}}\frac{\mu}{\sigma-1}\bar{\mathbf{D}} + \frac{1}{\bar{w}}\left(\mathbf{I} - \mu\bar{\mathbf{D}}\right)\mathbf{W}_x, \tag{F.15}$$

where $\mathbf{W}_x$ is given by (F.4) and (F.8). Plugging $\delta = \frac{\mu(\bar{w}\bar{x}+l)}{\bar{w}\bar{x}} = 1$ (as $\bar{w}\bar{x} = \frac{\mu}{1-\mu} l$) and $\epsilon_x = \epsilon_w = \mu$ to (F.8),

$$\mathbf{W}_x = \frac{\bar{w}}{\bar{x}} \underbrace{\left(\sigma\mathbf{I} - \mu\bar{\mathbf{D}} - (\sigma-1)\bar{\mathbf{D}}^2\right)^{-1}}_{\bar{x}\mathbf{S}_w^{-1}} \underbrace{\bar{\mathbf{D}}\left(\mu\mathbf{I} - \bar{\mathbf{D}}\right)}_{\bar{w}^{-1}\mathbf{S}_x}. \tag{F.16}$$

Since circulant matrices commute (Fact E.2), (F.15) and (F.16) together imply $\mathbf{V} = \bar{x}\nabla \log \boldsymbol{v}(\bar{\boldsymbol{x}}) = G^\flat(\bar{\mathbf{D}})^{-1} G^\sharp(\bar{\mathbf{D}})$, where we define

$$G^\sharp(\chi) \equiv \mu\left(\frac{1}{\sigma-1} + \frac{1}{\sigma}\right)\chi - \left(\frac{\mu^2}{\sigma-1} + \frac{1}{\sigma}\right)\chi^2, \tag{F.17}$$

$$G^\flat(\chi) \equiv 1 - \frac{\mu}{\sigma}\chi - \frac{\sigma-1}{\sigma}\chi^2, \tag{F.18}$$

as presented in Section 2.2.

Remark F.1. In Figure 10, we set $\mu = 0.5$, $\sigma = 10$, and $L = 8$. ■

Remark F.2 (Derivation for Example 6.4)**.** To obtain $G^\natural$ with respect to $\boldsymbol{l} = (l_i)_{i\in\mathcal{N}}$, we need to evaluate $\mathbf{V}_l = -\tilde{\mathbf{V}}_w \mathbf{S}_w^{-1}\mathbf{S}_l$ since $\mathbf{A} = \frac{l}{\bar{v}}\mathbf{V}_l$. From Example F.3, we have $\mathbf{S}_l = -\mu\bar{\mathbf{D}}$. Also, $\tilde{\mathbf{V}}_w = \bar{v}\frac{\partial}{\partial w}\log \boldsymbol{v}(\bar{\boldsymbol{x}}) = \frac{\bar{v}}{\bar{w}}(\mathbf{I} - \mu\bar{\mathbf{D}})$ and $\tilde{\mathbf{V}}_l = \mathbf{0}$. We obtain

$$G^\natural(\chi) = c\frac{\chi(1-\mu\chi)}{G^\flat(\chi)} > 0 \tag{F.19}$$

where $c = \frac{l}{\bar{w}}\frac{\mu}{\sigma} = \frac{1-\mu}{\sigma}\bar{x} > 0$. It then follows that

$$\delta(\chi) = -\frac{\bar{x}}{\bar{a}}\frac{G^{\natural}(\chi)}{G(\chi)} = -\frac{c\bar{x}}{\bar{a}}\frac{\chi(1-\mu\chi)}{G^{\sharp}(\chi)}. \tag{F.20}$$

Straightforward algebra veryfies that $\delta'(\chi) < 0$ if $G^{\sharp}(\chi) > 0$. ■

F.2.2 Forslid and Ottaviano (2003) (FO) model (Example 3.3)

The FO model is a slightly simplified version the Krugman model. The model is sometimes called the *footloose-entrepreneur model*, since a unit of mobile (mobile) labor is required as the fixed input of a manufacturing firm, thereby x_i coincides with the mass of firms. The only difference is that the variable input of M-sector firms in the Krugman model is now replaced by immobile labor. Specifically, to produce q units of good, an M-sector firm now requires α units of mobile labor and βq units of immobile labor. Thus, the total cost of a firm in region i that produces q units of good is $\alpha w_i + \beta q$. It implies $p_{ij} = \frac{\sigma\beta}{\sigma-1}\tau_{ij}$ provided that A-sector goods are produced in every region. The market equilibrium conditions under a fixed $\boldsymbol{x} \in \mathcal{X}$ is

$$w_i = \frac{\mu}{\sigma}\sum_{j\in\mathcal{N}}\frac{\phi_{ij}}{\sum_{k\in\mathcal{N}}\phi_{kj}x_k}(w_j x_j + l_j). \tag{F.21}$$

This equation is analytically solvable. In vector–matrix form, we have

$$\boldsymbol{w} = \frac{\mu}{\sigma}\left(\mathbf{I} - \frac{\mu}{\sigma}\mathbf{M}\,\mathrm{diag}[\boldsymbol{x}]\right)^{-1}\mathbf{M}\boldsymbol{l}, \tag{F.22}$$

where $\boldsymbol{l} \equiv (l_i)$ and $\mathbf{M} \equiv [m_{ij}] = [\phi_{ij}/\Delta_j]$ with $\Delta_i = \sum_{j\in\mathcal{N}}\phi_{ji}x_j$. The indirect utility $\boldsymbol{v}(\boldsymbol{x})$ is expressed as $v_i = \Delta_i^{\frac{\mu}{\sigma-1}} w_i$. At $\bar{\boldsymbol{x}}$, we compute that $\mathbf{V} = G^{\flat}(\bar{\mathbf{D}})^{-1}G^{\sharp}(\bar{\mathbf{D}})$ with

$$G^{\sharp}(\chi) = \mu\left(\frac{1}{\sigma-1} + \frac{1}{\sigma}\right)\chi - \left(\frac{\mu}{\sigma-1}\frac{\mu}{\sigma} + 1\right)\chi^2, \tag{F.23}$$

$$G^{\flat}(\chi) = 1 - \frac{\mu}{\sigma}\chi. \tag{F.24}$$

F.2.3 Pflüger (2004) (Pf) model (Example 3.3)

The Pf model is a further simplified version of the FO model (and hence the Krugman model) in which we assume a quasi-linear form for the upper tier. It results in the following analytical expression for $\boldsymbol{w}$:

$$w_i = \frac{\mu}{\sigma}\sum_{j\in\mathcal{N}}\frac{\phi_{ij}}{\sum_{k\in\mathcal{N}}\phi_{kj}x_k}(x_j + l_j). \tag{F.25}$$

Indirect utility is then given by $v_i = \log \Delta_i^{\frac{\mu}{\sigma-1}} + w_i$, where $\Delta_i \equiv \sum_{j\in\mathcal{N}} \phi_{ji} x_j$. At the uniform distribution, we have

$$\mathbf{V} = \frac{1}{\bar{v}} \left(\mu \left(\frac{1}{\sigma - 1} + \frac{1}{\sigma} \right) \bar{\mathbf{D}} - \frac{\mu}{\sigma}(1 + L)\bar{\mathbf{D}}^2 \right) \tag{F.26}$$

so that we may let $G^\sharp(\chi) = \mu \left(\frac{1}{\sigma-1} + \frac{1}{\sigma} \right) \chi - \frac{\mu}{\sigma}(1+L)\chi^2$. Observe that the Pf model reveals the agglomeration force in the Km framework in its simplest form.

F.2.4 Helpman (1998) model (Examples 2.3 and 6.3)

Helpman (1998) removed the A-sector in the Krugman model and assumed that all workers are mobile. Instead of the A-sector, the model introduces the housing (abbreviated as H) sector. Each region i is endowed with a fixed stock a_i of housing. Workers' preference is Cobb–Douglas of M-sector CES aggregate $\mathrm{C}_i^{\mathrm{M}}$ and H-sector goods $\mathrm{C}_i^{\mathrm{H}}$, $u_i = (\mathrm{C}_i^{\mathrm{M}})^\mu (\mathrm{C}_i^{\mathrm{H}})^\gamma$, where $\mu \in (0, 1)$ is the expenditure share of the former and $\gamma = 1 - \mu \in (0, 1)$ is that for the latter. There are two variants for assumptions on how housing stocks are owned: *public landownership* (abbreviated as PL) and *local landownership* (LL). The original formulation by Helpman (1998) supposes housing stocks are equally owned by all workers (i.e., PL). The income of a worker in region i is the sum of the wage and an equal dividend $r > 0$ of rental revenue *over the economy*. On the other hand, Ottaviano et al. (2002), Murata and Thisse (2005), and Redding and Sturm (2008) assumed that housing stocks are locally owned (i.e., LL). The income of a worker in region i is the sum of the wage and an equal dividend of rental revenue *in each region*.

Regarding the market equilibrium conditions, the only difference from the Krugman model is regional expenditure e_i on M-sector goods in each region:

$$[\text{PL}] \quad e_i = \mu \left(w_i + r \right) x_i, \tag{F.27}$$

$$[\text{LL}] \quad e_i = w_i x_i, \tag{F.28}$$

and market wage is given as the solution for (F.5). For the LL case, $\boldsymbol{w}(\boldsymbol{x})$ is uniquely given up to normalization. The indirect utility function is

$$[\text{PL}] \quad v_i = \left(\frac{x_i}{a_i} \right)^{-\gamma} (w_i + r)^\mu \Delta_i^{\frac{\mu}{\sigma-1}}, \tag{F.29}$$

$$[\text{LL}] \quad v_i = \left(\frac{x_i}{a_i} \right)^{-\gamma} w_i^\mu \Delta_i^{\frac{\mu}{\sigma-1}}, \tag{F.30}$$

where $\Delta_i \equiv \sum_{j\in\mathcal{N}} x_j w_j^{1-\sigma} \phi_{ji}$ and $r > 0$.

Let $a_i = 1$ to satisfy Assumption S. We compute that

$$\mathbf{V} = \bar{x}\left(\frac{\mu}{\sigma-1}\mathbf{M}^\top \operatorname{diag}[\boldsymbol{x}]^{-1} + \hat{\mathbf{V}}_w \mathbf{W}_x - \gamma \operatorname{diag}[\boldsymbol{x}]^{-1}\right), \tag{F.31}$$

$$\text{where} \quad [\text{PL}] \quad \hat{\mathbf{V}}_w \equiv \mu\left(\operatorname{diag}[\boldsymbol{w} + r\mathbf{1}]^{-1} - \mathbf{M}^\top \operatorname{diag}[\boldsymbol{w}]^{-1}\right), \tag{F.32}$$

$$[\text{LL}] \quad \hat{\mathbf{V}}_w \equiv \mu\left(\mathbf{I} - \mathbf{M}^\top\right)\operatorname{diag}[\boldsymbol{w}]^{-1}, \tag{F.33}$$

and $\mathbf{M}$ is defined by (F.6). From (F.7), $\mathbf{V} = G^\flat(\bar{\mathbf{D}})^{-1}G^\sharp(\bar{\mathbf{D}})$ with

$$G^\sharp(\chi) \equiv -\gamma + \mu\left(\frac{1}{\sigma-1} + \frac{1}{\sigma}\right)\chi - \left(\left(\frac{\mu^2}{\sigma-1} + \frac{1}{\sigma}\right) - \gamma\right)\chi^2, \tag{F.34}$$

$$G^\flat(\chi) \equiv 1 - \frac{\mu}{\sigma}\chi - \frac{\sigma-1}{\sigma}\chi^2 \tag{F.35}$$

for the PL case, whereas for the LL case

$$G^\sharp(\chi) \equiv (1-\chi)\left(-\gamma + \left(\mu\left(\frac{1}{\sigma-1} + \frac{1}{\sigma}\right) - \gamma\frac{\sigma-1}{\sigma}\right)\chi\right), \tag{F.36}$$

$$G^\flat(\chi) \equiv (1-\chi)\left(1 + \frac{\sigma-1}{\sigma}\chi\right). \tag{F.37}$$

Remark F.3. The condition for the uniqueness of the equilibrium is $\gamma\sigma = (1-\mu)\sigma > 1$ (Redding and Sturm, 2008). For both PL and LL, it implies that $G^\sharp(\chi) < 0$ for all $\chi \in (0,1)$. ■

Remark F.4 (Derivation for Example 6.3)**.** The regional model formulated in §3 of Redding and Rossi-Hansberg (2017) is an enhanced version of the Helpman model with LL, in which the variable input of mobile labor is allowed to depend on region i (i.e., productivity differs across regions). That is, the cost function of firms in region i is given by $C_i(q) = w_i(\alpha + \beta_i q)$. The market equilibrium condition for this case is, with $a_i \equiv \beta_i^{1-\sigma} > 0$, given by

$$s_i(\boldsymbol{x}, \boldsymbol{w}) = w_i x_i - \sum_{j\in\mathcal{N}} \frac{x_i a_i w_i^{1-\sigma}\phi_{ij}}{\sum_{k\in\mathcal{N}} x_k a_k w_k^{1-\sigma}\phi_{kj}} w_j x_j = 0. \tag{F.38}$$

The payoff function is given by (F.33) with $\Delta_i = \sum_{k\in\mathcal{N}} x_k a_k w_k^{1-\sigma}\phi_{ki}$.

From Example F.2, $\mathbf{S}_a = -\frac{\bar{w}\bar{x}}{\bar{a}}(\mathbf{I} - \bar{\mathbf{D}})(\mathbf{I} + \bar{\mathbf{D}})$ as $\bar{e} = \bar{w}\bar{x}$. Also, we have $\tilde{\mathbf{V}}_w = \frac{\bar{v}}{\bar{w}}\mu(\mathbf{I} - \bar{\mathbf{D}})$, $\tilde{\mathbf{V}}_a = \frac{\bar{v}}{\bar{a}}\frac{\mu}{\sigma-1}\bar{\mathbf{D}}$, and $\mathbf{S}_w = \sigma\bar{x}G^\flat(\bar{\mathbf{D}})$. Since $\mathbf{V}_a = \tilde{\mathbf{V}}_a - \tilde{\mathbf{V}}_w\mathbf{S}_w^{-1}\mathbf{S}_a$ and $\mathbf{A} = \frac{\bar{a}}{\bar{v}}\mathbf{V}_a = G^\natural(\bar{\mathbf{D}})$, we compute

$$G^\natural(\chi) = c\frac{(\sigma-1) + \sigma\chi}{G^\flat(\chi)} > 0 \tag{F.39}$$

where $c \equiv \frac{\bar{v}}{\bar{a}}\frac{\mu}{\sigma} > 0$. This in turn implies

$$\delta(\chi) = -\frac{\bar{x}}{\bar{a}}\frac{G^{\natural}(\chi)}{G(\chi)} = -\frac{c\bar{x}}{\bar{a}}\frac{(\sigma-1)+\sigma\chi}{G^{\sharp}(\chi)} \tag{F.40}$$

where $G^{\sharp}(\chi)$ is that for the LL case (F.37). We can show that $\delta'(\chi) > 0$ for all χ whenever $(1-\mu)\sigma > 1$ so that equilibrium is unique (Remark F.3). ■

F.2.5 Puga (1999) model (Example 3.3)

Puga (1999) generalized the Krugman model in two directions, namely (i) the inter-sector mobility of workers between the A- and M-sector (without immobile workers but land) and (ii) intermediate inputs in the M-sector, both as in Krugman and Venables (1995).

There is only a unit mass of mobile workers. Let x_i^{M} and x_i^{A} the masses of workers engaged in the M- and A-sectors, respectively ($x_i = x_i^{\mathrm{M}} + x_i^{\mathrm{A}}$). The homogeneous preference of consumers is the same as in the Krugman model, with the expenditure share of the M-sector good μ and elasticity of substitution between M-sector varieties σ. Each region is endowed with a_i units of land owned by immobile landlords that have the same preference as the workers. We assume that if a worker relocates, then he or she first enters the M-sector of the destination region. The stability of the spatial pattern $\boldsymbol{x}$ is then reduced to the study of $\boldsymbol{x}^{\mathrm{M}} \equiv [x_i^{\mathrm{M}}]$.

The A-sector is perfectly competitive and produces a homogeneous output by using labor and land under constant returns to scale. A-sector goods are costless to trade and set as the numéraire. Let X_i^{A} be the gross regional product of the A-sector. In line with the original study, we specify a Cobb–Douglas production function with labor share $\bar{\mu}$; in concrete terms, we have $X_i^{\mathrm{A}} = (x_i^{\mathrm{A}})^{\bar{\mu}} a_i^{1-\bar{\mu}}$. This implies that the total labor costs of A-sector firms are given by $\bar{\mu} X_i^{\mathrm{A}} = w_i x_i^{\mathrm{A}}$, while their land costs (= the total rental revenue of landlords) are $(1-\bar{\mu})X_i^{\mathrm{A}} = \frac{1-\bar{\mu}}{\bar{\mu}} w_i x_i^{\mathrm{A}}$. In particular, labor demand in this sector is given by a function of the wage $x_i^{\mathrm{A}} = a_i (w_i/\bar{\mu})^{1/(\bar{\mu}-1)}$, because $w_i = \bar{\mu}(x_i^{\mathrm{A}}/a_i)^{\bar{\mu}-1}$. Let $x_i^{\mathrm{A}} = \epsilon_i x_i^{\mathrm{M}}$, meaning that $x_i = (1+\epsilon_i)x_i^{\mathrm{M}}$; we assume $x_i^{\mathrm{M}} \neq 0$, because we are interested in the stability of complete dispersion. We also have $\epsilon_i \equiv (a_i/x_i^{\mathrm{M}})(w_i/\bar{\mu})^{1/(\bar{\mu}-1)}$. The regional rental revenue from land, R_i, in terms of x_i^{M} is

$$R_i \equiv \frac{1-\bar{\mu}}{\bar{\mu}} \epsilon_i w_i x_i^{\mathrm{M}}. \tag{F.41}$$

By employing the above formulae, the elasticity ν_i of a region's labor supply to the M-sector with respect to wage is $\nu_i \equiv \frac{w_i}{x_i^{\mathrm{M}}}\frac{\partial x_i^{\mathrm{M}}}{\partial w_i} = \frac{x_i^{\mathrm{A}}}{x_i^{\mathrm{M}}}\frac{1}{1-\bar{\mu}} = \epsilon_i \frac{1}{1-\bar{\mu}}$.

By considering the simplest possible model of round-about intermediate inputs as in Krugman and Venables (1995), the minimized cost in the M-sector is $C_i(q) = P_i^{\hat{\mu}} w_i^{1-\hat{\mu}}(\alpha + \beta q)$ where P_i is the price index of M-sector goods in region i and $\hat{\mu}$ the share of intermediates in firms' costs. The (variety-independent) profit-maximizing price is given by

$$p_{ij} = \frac{\sigma\beta}{\sigma - 1} P_i^{\hat{\mu}} w_i^{1-\hat{\mu}} \tau_{ij}. \tag{F.42}$$

Resultingly, $\mathbf{P} = (P_i)_{i\in\mathcal{N}}$ should satisfy the following equation:

$$t_i(\boldsymbol{x}, \boldsymbol{w}, \mathbf{P}) = P_i^{1-\sigma} - \frac{1}{1-\hat{\mu}} \sum_{j\in\mathcal{N}} x_j^{\mathrm{M}} P_j^{-\hat{\mu}\sigma} w_j^{1-\sigma+\hat{\mu}\sigma} \phi_{ji}. \tag{F.43}$$

Land is locally owned by immobile landlords that share the same preference as mobile workers; their regional expenditure on M-sector goods is given by μR_i. In addition, the regional expenditure of firms on intermediates is given by $\hat{\mu} C_i n_i = \frac{\hat{\mu}}{1-\hat{\mu}} w_i x_i^{\mathrm{M}}$. Total expenditure in region i on M-sector goods is $e_i = \mu w_i x_i + \mu R_i + \hat{\mu} C_i n_i$. By using (F.41) as well as $x_i = (1+\epsilon_i) x_i^{\mathrm{M}}$, this is simplified to

$$e_i = \left(\mu \left(1 + \frac{\epsilon_i}{\bar{\mu}} \right) + \frac{\hat{\mu}}{1-\hat{\mu}} \right) w_i x_i^{\mathrm{M}}. \tag{F.44}$$

The market equilibrium condition for the model is given by

$$s_i(\boldsymbol{x}, \boldsymbol{w}, \mathbf{P}) = \frac{1}{1-\hat{\mu}} w_i x_i^{\mathrm{M}} - \sum_{j\in\mathcal{N}} m_{ij} e_j = 0 \tag{F.45}$$

where we define

$$m_{ij} = \frac{x_i^{\mathrm{M}} P_i^{-\hat{\mu}\sigma} w_i^{1-\sigma+\hat{\mu}\sigma} \phi_{ij}}{\sum_{k\in\mathcal{N}} x_k^{\mathrm{M}} P_k^{-\hat{\mu}\sigma} w_k^{1-\sigma+\hat{\mu}\sigma} \phi_{kj}}. \tag{F.46}$$

The market wage $\boldsymbol{w} = (w_i)$ and price index $\mathbf{P} = (P_i)$ are obtained as the solution for the system of non-linear equations (F.43) and (F.45). We require $\hat{\mu} < \frac{\sigma-1}{\sigma}$, meaning that $\mathbf{P}$ and $\boldsymbol{w}$ are uniquely determined for any transportation cost. Given $\mathbf{P}$ and $\boldsymbol{w}$, the indirect utility function is $v_i = \tilde{v}_i(\boldsymbol{x}, \boldsymbol{w}, \mathbf{P}) = \Delta_i^{\frac{\mu}{\sigma-1}} w_i$ with $\Delta_i = \sum_{j\in\mathcal{N}} x_j^{\mathrm{M}} P_j^{-\hat{\mu}\sigma} w_j^{1-\sigma+\hat{\mu}\sigma} \phi_{ji}$.

Let $a_i = a$ for all i to satisfy Assumption S. Note that $\boldsymbol{v}(\boldsymbol{x})$ is differentiated in $\boldsymbol{x}^{\mathrm{M}}$ as $\mathbf{V}_x = \tilde{\mathbf{V}}_x + \tilde{\mathbf{V}}_w \mathbf{W}_x + \tilde{\mathbf{V}}_P \mathbf{P}_x$, where $\mathbf{W}_x = [\partial w_i / \partial x_j^{\mathrm{M}}]$ and $\mathbf{P}_x = [\partial P_i / \partial x_j^{\mathrm{M}}]$ are evaluated by applying the implicit function theorem to (F.43) and (F.45). We must have $\mathbf{T}_x \mathrm{d}\boldsymbol{x} + \mathbf{T}_w \mathbf{W}_x \mathrm{d}\boldsymbol{w} + \mathbf{T}_P \mathbf{P}_x \mathrm{d}\mathbf{P} = \mathbf{0}$ and $\mathbf{S}_x \mathrm{d}\boldsymbol{x} + \mathbf{S}_w \mathbf{W}_x \mathrm{d}\boldsymbol{w} + \mathbf{S}_P \mathbf{P}_x \mathrm{d}\mathbf{P} = \mathbf{0}$ for any infinitesimal $(\mathrm{d}\boldsymbol{x}, \mathrm{d}\boldsymbol{w}, \mathrm{d}\mathbf{P})$, thereby

$$\mathbf{P}_x = -\left(\mathbf{T}_w \mathbf{S}_P - \mathbf{T}_P \mathbf{S}_w\right)^{-1} \left(\mathbf{T}_w \mathbf{S}_x - \mathbf{T}_x \mathbf{S}_w\right), \tag{F.47}$$

$$\mathbf{W}_x = \left(\mathbf{T}_w \mathbf{S}_P - \mathbf{T}_P \mathbf{S}_w\right)^{-1} \left(\mathbf{T}_P \mathbf{S}_x - \mathbf{T}_x \mathbf{S}_P\right). \tag{F.48}$$

A patient computation yields $\mathbf{V} = \frac{\bar{x}^{\mathrm{M}}}{\bar{v}} \mathbf{V}_x = G^{\flat}(\bar{\mathbf{D}})^{-1} G^{\sharp}(\bar{\mathbf{D}})$ with

$$G^{\sharp}(\bar{\mathbf{D}}) \equiv \left(\check{\mu} \left(\frac{1}{\sigma - 1} + \frac{1}{\sigma} \right) \bar{\mathbf{D}} - \left(\frac{\check{\mu}^2}{\sigma - 1} + \frac{1}{\sigma} + \eta \right) \bar{\mathbf{D}}^2 \right) \tag{F.49}$$

and $G^{\flat}(\bar{\mathbf{D}})$ is a positive definite matrix defined by $\bar{\mathbf{D}}$. We let $\check{\mu} \equiv \hat{\mu} + \mu(1 - \hat{\mu})$, which is loosely interpreted as the aggregate expenditure in the economy on M-sector goods, and $\eta \equiv \frac{\mu(1-\check{\mu})}{\sigma(\sigma-1)}(1 - \bar{v})$ is a constant that summarizes the effects of labor mobility between the A- and M-sectors at $\bar{\boldsymbol{x}}$, where $\bar{v} \equiv \frac{\bar{\mu}}{1-\bar{\mu}} \frac{1-\mu}{\mu}$ is the elasticity of labor supply from the A-sector to the M-sector with respect to wage at $\bar{\boldsymbol{x}}$.

F.2.6 Tabuchi (1998) model

The Tabuchi model introduces the internal structure of regions to the Krugman model. The main thrust of this model is that unlike the majority of regional models, the city boundary in each region is endogenously determined by the full-fledged monocentric city model of Alonso–Muth–Mills. This produces a rich structure of urban costs, because the trade-off between commuting costs and land rents is explicit.

There are three sectors, M, H, and A. The internal structure of each region is featureless, except that it is endowed with a single central business district (CBD) with negligible spatial extent. In each region, locations are indexed by the distance from the CBD, $\ell \geq 0$. At any point, the land endowment density is assumed to be unity. The total mass of mobile and immobile workers are given by 1 and L, respectively. The mass of mobile workers in region i is denoted by x_i, whereas the spatial distribution (density) in that region is, allowing notational abuse, denoted by $x_i(\ell)$. Thus, we have

$$\int_0^{\ell_i} x_i(\ell) \mathrm{d}\ell = x_i, \tag{F.50}$$

where $\ell_i \geq 0$ is the city boundary in region i that is endogenously determined. Immobile workers are employed by the A-sector and do not commute to the CBD, whereas mobile workers do. A mobile worker at distance ℓ from the CBD incurs the generalized cost of commuting $T(\ell)$, which is measured by the numéraire. For simplify, we assume that the internal structure of each region is one-dimensional and extends symmetrically around the CBD such that $[-\ell_i, \ell_i]$ á la Murata and Thisse (2005).

The utility of a representative worker living in region i and located at ℓ is given by $u_i = (C_i^{\mathrm{M}})^{\mu} (C_i^{\mathrm{H}})^{\gamma} (C_i^{\mathrm{A}})^{1-\mu-\gamma}$ where μ and γ with $\mu + \gamma < 1$ are the constant expenditure shares for M-sector goods and H-sector goods, respectively; C_i^{M} is the CES aggregate of M-sector goods, C_i^{H} the consumption of housing space (H-sector goods), C_i^{A} the consumption of agricultural products (A-sector goods) in region i. The M- and A-sectors are the same as in the Krugman model, whereas the H-sector is the same as in the Helpman model. By choosing A-sector goods as the numéraire, the budget constraint of a mobile worker at location ℓ in region i is

$$C_i^{\mathrm{A}} + r_i(\ell) C_i^{\mathrm{H}}(\ell) + \sum_{j \in \mathcal{N}} \int_0^{n_j} p_{ji}(\xi) q_{ji}(\xi) \mathrm{d}\xi = y_i(\ell) = y_i - T(\ell), \tag{F.51}$$

where $r_i(\ell)$ is the land rent prevailing at location ℓ in region i, $T(\ell)$ the generalized cost of commuting from location ℓ to the CBD, and y_i the income of the worker. We assume that $T(\ell)$ is differentiable and increasing in ℓ with $T(0) = 0$.

Following the tradition of urban economics, the model assumes absentee landowners who keep the rental revenue of housing, leading to $y_i = w_i$ for every mobile worker. Immobile workers live outside the city and do not commute to the CBD. Thus, they face the region-independent agricultural land rent $r^{\mathrm{A}} > 0$ and zero commuting cost and $y_i = 1$. Intracity transportation of M-sector goods is costless, so that workers in each region face the same M-sector product price.

As shown in the original paper, the population density in the region for the given ℓ_i and w_i is given by

$$x_i(\ell, w_i) = \frac{r^{\mathrm{A}}}{\gamma w_i} \left(1 - \frac{T(\ell_i)}{w_i}\right)^{-\frac{1}{\gamma}} \left(1 - \frac{T(\ell)}{w_i}\right)^{\frac{1}{\gamma}-1} \tag{F.52}$$

We define the total commuting costs in the region by

$$T_i(\ell_i, w_i) = \int_0^{\ell_i} T(\ell) x_i(\ell, w_i) \mathrm{d}\ell. \tag{F.53}$$

Note that (x_i, w_i) is uniquely mapped to ℓ_i due to (F.50) and (F.52), so that T_i is also a function of (x_i, w_i).

The market equilibrium condition is given by (F.5), where we let

$$e_i = \mu \left(\int_{-\ell_i}^{\ell_i} y_i(\ell) x_i(\ell) \mathrm{d}\ell + l_i \right) = \mu \left(w_i x_i - T_i + l_i \right). \tag{F.54}$$

Then, the indirect utility in region i may be given evaluating it at the city boundary $\ell = \ell_i$, since utility is equalized in each region: $v_i(\boldsymbol{x}) = \tilde{v}_i(\boldsymbol{x}, \boldsymbol{w}) = \Delta_i^{\frac{\mu}{\sigma-1}} y_i(\ell_i)$, where $\Delta_i = \sum_{j \in \mathcal{N}} x_j w_j^{1-\sigma} \phi_{ji}$ and $y_i(\ell_i) = w_i - T(\ell_i)$.

Let $l_i = l$ for all $i \in \mathcal{N}$ to impose Assumption S and consider $\bar{\boldsymbol{x}}$. Let $\bar{w}$ and $\bar{T}$ be the uniform level of the nominal wage rate and total commuting cost in each region, respectively. Note that $\bar{T}$ is a function of $\bar{w}$ and $\bar{x}$. For normalization, we require

$$\bar{w}\bar{x} - \bar{T}(\ell, \bar{w}) = \frac{\mu}{1-\mu} l. \tag{F.55}$$

Then, there must exist a unique positive solution for the location of city boundary and wage rate

$(\bar{\ell}, \varpi)$ for the system of non-linear equations defined by (F.50) and (F.55). By employing the solution $(\bar{\ell}, \varpi)$, total expenditure in a region is given by $\bar{Y} = \frac{l}{1-\mu}$. Define the ratios κ of the regional disposable income of mobile workers and $\hat{\kappa}$ of regional total expenditure to the total nominal wage: $\kappa \equiv \frac{\varpi\bar{x}-\bar{T}}{\varpi\bar{x}}$ and $\hat{\kappa} \equiv \frac{\bar{Y}}{\varpi\bar{x}}$. The latter implies that $\frac{\bar{Y}}{\varpi} = \hat{\kappa}\bar{x}$ and $\frac{\bar{Y}}{\bar{x}} = \hat{\kappa}\varpi$. Given $(\bar{\ell}, \varpi)$, we define positive constants ψ_0, ψ_1, ρ_0, and ρ_1 such that $\mathbf{T}_x = \psi_0\mathbf{I}$, $\mathbf{T}_y = \psi_1\mathbf{I}$, $\mathbf{E}_x = \rho_0\varpi\mathbf{I}$, and $\mathbf{E}_y = \rho_1\bar{x}\mathbf{I}$.

Following Tabuchi (1998), we consider the simplest case where the commuting cost function is linear with respect to distance: $T(\ell) = t\ell$. Then,

$$\bar{\ell} = \frac{1}{t}(1 - \epsilon^{\gamma})\varpi, \tag{F.56}$$

where $\epsilon \in (0, 1)$ is defined by $\epsilon \equiv (1 + \hat{t}\bar{x})^{-1}$. The parameter $\hat{t} \equiv \frac{t}{2r^{\mathrm{A}}}$ is interpreted as a measure of the relative magnitude of commuting costs to land rents. As expected, $\bar{\ell}$ is decreasing in the generalized commuting cost per distance t. By solving (F.55), we have

$$\varpi = \frac{1}{\kappa}\frac{\mu}{1-\mu}L \quad \text{and} \quad \kappa = \frac{1}{1+\gamma}\frac{1-\epsilon^{1+\gamma}}{1-\epsilon} \tag{F.57}$$

as well as $\bar{y} \equiv \varpi - T(\bar{\ell}) = \epsilon^{\gamma}\varpi$, $\bar{Y} = \frac{l}{1-\mu}$, and $\hat{\kappa} = \frac{\kappa}{\mu}$. Then, we have $\psi_0 = \frac{\bar{y}\gamma}{\bar{x}}(1-\epsilon)$, $\psi_1 = 1 - \epsilon^{\gamma}$, $\rho_0 = 1 - \gamma(1-\epsilon)\kappa$, and $\rho_1 = \kappa$. Summarizing computations up to here gives:

$$\mathbf{V}_x = \frac{1}{\bar{x}}\left(\frac{\mu}{\sigma-1}\bar{\mathbf{D}} + \frac{\bar{x}}{\varpi}\left(\mathbf{I} - \mu\bar{\mathbf{D}}\right)\mathbf{W}_x - \hat{\gamma}\mathbf{I}\right), \tag{F.58}$$

$$\mathbf{W}_x = \frac{\varpi}{\bar{x}}\left(\hat{c}_0\mathbf{I} + \hat{c}_1\bar{\mathbf{D}} + \hat{c}_2\bar{\mathbf{D}}^2\right)^{-1}\left(\bar{c}_0\mathbf{I} + \bar{c}_1\bar{\mathbf{D}} + \bar{c}_2\bar{\mathbf{D}}^2\right) \tag{F.59}$$

with the coefficients being

$$\begin{cases} \hat{c}_0 \equiv 1 + (\sigma-1)\kappa > 0, \\ \hat{c}_1 \equiv -\mu\kappa < 0, \\ \hat{c}_2 \equiv -(\sigma-1)\kappa < 0, \end{cases} \qquad \begin{cases} \bar{c}_0 \equiv -(1-\kappa) < 0, \\ \bar{c}_1 \equiv \mu(1-\hat{\gamma}\kappa) > 0, \\ \bar{c}_2 \equiv -\kappa < 0 \end{cases} \tag{F.60}$$

where $\hat{\gamma} \equiv \gamma(1-\epsilon)$. Note that κ and $\hat{\gamma}$ together summarize the net effects of the two types of urban costs; κ and $\hat{\gamma}$ represent those from commuting and non-tradable land, respectively. Algebra shows that a net gain function for the model is $G^{\sharp}(\chi) = c_0 + c_1\chi + c_2\chi^2$ with

$$c_0 = -\hat{\gamma}\left(\frac{1}{\sigma} + \frac{\sigma-1}{\sigma}\kappa\right) < 0, \tag{F.61}$$

$$c_1 = \mu\left(\frac{1}{\sigma-1} + \frac{1}{\sigma}\right) > 0, \tag{F.62}$$

$$c_2 = -\left(\frac{\mu^2}{\sigma-1}\theta + \frac{1}{\sigma}\hat{\theta}\right) \tag{F.63}$$

where we let $\theta \equiv \frac{\kappa}{\sigma} + \frac{\sigma-1}{\sigma}(1 - \hat{\gamma}\kappa)$ and $\hat{\theta} \equiv \kappa\,(1 - \hat{\gamma}\kappa(\sigma - 1))$. For non-extremal cases, $c_2 < 0$ and the Tabuchi model incorporates both local and global dispersion forces.

F.2.7 Pflüger and Südekum (2008) model (Example 3.5)

The Pflüger–Südekum model builds on Pflüger (2004), with the only difference being that it introduces the housing sector (again denoted by H), which produces a local dispersion force. The indirect utility of a mobile worker in region i is

$$v_i(\boldsymbol{x}) = \frac{\mu}{\sigma - 1} \ln[\Delta_i] - \gamma \ln \frac{x_i + l_i}{a_i} + w_i, \tag{F.64}$$

where $\Delta_i = \sum_{j \in \mathcal{N}} \phi_{ji} x_j$, and l_i and a_i denote the mass of immobile workers and amount of housing stock in region i, respectively. The nominal wage in region i is given by (F.25). Let $l_i = l$ and $a_i = a$ for all i to meet Assumption S. Then, we see that $\mathbf{V} = \frac{1}{\bar{v}} G^\sharp(\bar{\mathbf{D}})$ with

$$G^\sharp(\chi) = -\frac{\gamma}{1 + L} + \mu \left(\frac{1}{\sigma - 1} + \frac{1}{\sigma} \right) \chi - \frac{\mu}{\sigma}(1 + L)\chi^2. \tag{F.65}$$

Remark F.5. Figure 13a and Figure 13b assume the Pflüger–Südekum model. We set $\mu = 0.4$, $\sigma = 2.5$, $L = 4$, $\gamma = 0.5$, and $a = 1$. ■

F.2.8 Murata and Thisse (2005) model (Example 3.4)

Similar to the Tabuchi model, Murata and Thisse (2005) studied the interplay between commuting costs and interregional transport costs by employing a simplified yet reasonable specification. The internal structure of each region is assumed to be one-dimensional and featureless except that there is a given CBD; the city expands symmetrically around the origin. There are only mobile and mobile workers, who choose their own residential region i and location $\ell \geq 0$ in that region, where the CBD is located at $\ell = 0$.

Land endowment equals unity everywhere in a region and workers are assumed to inelastically consume one unit of land. The opportunity cost of land is normalized to zero in every region. Then, the city spreads in the interval $\mathcal{X}_i \equiv [-\ell_i, \ell_i]$, where $\ell_i \equiv \frac{x_i}{2}$ denotes the city boundary. Commuting costs take an iceberg form. Specifically, a worker located at ℓ supplies $s(\ell) = 1 - 4\theta|\ell|$ units of labor, where we require $\theta \in [0, \frac{1}{2})$ so that we have $s(\ell) \geq 0$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all region i at any configuration. Then, total effective labor supply in the CBD of region i is given by

$$S_i = \int_{\mathcal{X}_i} s(\ell) \mathrm{d}\ell = x_i(1 - \theta x_i). \tag{F.66}$$

Note that $S_i = x_i$ when commuting is costless so that $\theta = 0$. Land is locally owned as in the Helpman model with LL (Section F.2.4).

The homogeneous preference of mobile workers in region i is $u_i = \ln C_i^{\mathrm{M}}$ where C_i^{M} is the consumption of the CES aggregate of M-sector goods. Manufacturing firms are assumed to be the same as in the Krugman model. Specifically, to produce q units of a good, a firm requires $\alpha + \beta q$ units of mobile labor. The market equilibrium condition is

$$w_i S_i = \sum_{j \in \mathcal{N}} \frac{S_i w_i^{1-\sigma} \phi_{ij}}{\sum_{k \in \mathcal{N}} S_k w_k^{1-\sigma} \phi_{kj}} w_j S_j \tag{F.67}$$

To normalize $\boldsymbol{w}$, we assume $\sum_{i \in \mathcal{N}} w_i S_i = 1 > 0$. Given the solution $\boldsymbol{w}$ to the equation, the indirect utility of workers in region i is obtained as

$$v_i(\boldsymbol{x}) = \frac{1}{\sigma - 1} \ln[\Delta_i] + \ln[w_i] + \ln[1 - \theta x_i], \tag{F.68}$$

where $\Delta_i \equiv \sum_{k \in \mathcal{N}} S_k w_k^{1-\sigma} d_{ki}$.

We compute as follows:

$$\mathbf{V}_x = \frac{1}{\sigma - 1} \mathbf{M}^\top \operatorname{diag}[\mathbf{S}]^{-1} \mathbf{S}_x + (\mathbf{I} - \mathbf{M}^\top) \operatorname{diag}[\boldsymbol{w}]^{-1} \mathbf{W}_x - \theta \operatorname{diag}[\mathbf{1} - \theta \boldsymbol{x}]^{-1}, \tag{F.69}$$

where $\mathbf{S}_x = \operatorname{diag}[\mathbf{1} - 2\theta \boldsymbol{x}]$. At $\bar{\boldsymbol{x}}$, we have

$$\mathbf{W}_x = \frac{\bar{w}(1 - 2\theta \bar{x})}{\bar{x}(1 - \theta \bar{x})} (\sigma \mathbf{I} + (\sigma - 1) \bar{\mathbf{D}})^{-1} \bar{\mathbf{D}}, \tag{F.70}$$

implying that $\mathbf{V} = \frac{\bar{x}}{\bar{v}} \mathbf{V}_x = \frac{1 - 2\theta \bar{x}}{(1 - \theta \bar{x}) \bar{v}} G^\flat(\bar{\mathbf{D}})^{-1} G^\sharp(\bar{\mathbf{D}})$ with

$$G^\sharp(\chi) = -\hat{\theta} + \left((1 - \hat{\theta}) \left(\frac{1}{\sigma - 1} + \frac{1}{\sigma} \right) - \hat{\theta} \frac{\sigma - 1}{\sigma} \right) \chi, \tag{F.71}$$

$$G^\flat(\chi) = 1 + \frac{\sigma - 1}{\sigma} \chi, \tag{F.72}$$

where we define $\hat{\theta} \equiv \frac{\theta \bar{x}}{1 - 2\theta \bar{x}}$.

F.2.9 Harris and Wilson (1978) (Example 3.3)

The Harris–Wilson model is an archetypal economic geography model formulated in the field of geography well before economists started to emphasize the self-organization of the spatial allocation of economic activity. A detailed analysis of the model can be found in Osawa et al. (2017).

The city is discretized into N *zones* and associated centroids. There is a continuum of retailing firms in each zone that operate a shop. The mass of firms in zone i is denoted by $x_i \geq 0$; $\boldsymbol{x}$ denotes the spatial distribution of retailers. A fixed proportion of consumers resides in each zone. Consumers are assumed to inelastically buy retail goods from some shop located in the city. Total per capita consumer demand for a shopping activity in zone i is a constant O_i. Consumers' shopping behavior is captured by a set of origin-constrained gravity equations. For any given $\boldsymbol{x}$, consumer demand $S_{ij}(\boldsymbol{x})$ from zone i to j is given by

$$S_{ij} = \frac{x_j^\alpha \phi_{ij}}{\sum_{k \in \mathcal{N}} x_k^\alpha \phi_{ik}} O_i, \tag{F.73}$$

with $\alpha > 0$. The term x_i^α is the "attractiveness" of the retailers in zone i.

The payoff (profit) of a retailer in zone i is defined as follows:

$$\Pi_i(\boldsymbol{x}) = \frac{\sum_{j \in \mathcal{N}} S_{ji}}{x_i} - \kappa_i, \tag{F.74}$$

where κ_i is the fixed cost of entry.

Harris and Wilson (1978) assumed that the spatial pattern $\boldsymbol{x}$ gradually evolves in proportion to the profit $\boldsymbol{\Pi}(\boldsymbol{x})$ and the state $\boldsymbol{x}$. Specifically, we let $\dot{x}_i = F_i(\boldsymbol{x}) \equiv x_i \Pi_i(\boldsymbol{x}) = S_i - \kappa_i x_i$ where $S_i = \sum_{j \in \mathcal{N}} S_{ji}$.

To satisfy Assumption S, let $O_i = 1$ and $\kappa_i = \kappa$ for all $i \in \mathcal{N}$. The Harris–Wilson model is an open-city model. The total mass of retailers at an equilibrium is thus determined from the following equilibrium condition: $x_i \Pi_i(\boldsymbol{x}) = 0$, $x_i \geq 0$, $\Pi_i(\boldsymbol{x}) \leq 0$. At any equilibrium, we have $\sum_{i \in \mathcal{N}} \kappa_i x_i = \sum_{i \in \mathcal{N}} O_i$ and thus if $O_i = 1$ and $\kappa_i = \kappa$ then $\mathcal{X} \equiv \{\boldsymbol{x} \in \mathbb{R}^K \mid \sum_{i \in \mathcal{N}} x_i = \frac{N}{\kappa},\ x_i \geq 0\}$ is globally attracting under $\mathbf{F}$. It is immediately clear that

$$\nabla \mathbf{F}(\bar{\boldsymbol{x}}) = (\kappa\alpha) \left(\frac{\alpha - 1}{\alpha} \mathbf{I} - \bar{\mathbf{D}}^2 \right), \tag{F.75}$$

so that $G^\sharp(\chi) = \frac{\alpha-1}{\alpha} - \chi^2$ for the model.

F.2.10 Takayama and Akamatsu (2011) (Example 3.5)

Takayama and Akamatsu (2011) is a partial equilibrium model that introduces a spatial competition effect à la Harris and Wilson (1978) into the Beckmann model. Specifically, in essence, they introduced firms that sell goods at a fixed price to spatially immobile consumers.

In each area, l_i immobile consumers with $\sum_i l_i = L$ demand a single unit of goods produced by firms; immobile consumers are assumed to engage in jobs in other industries. Given the spatial distribution $\boldsymbol{n} = (n_i)_{i \in \mathcal{N}}$ of firms, demand from area j to i is given by the following origin-constrained gravity equation:

$$q_{ji} = \frac{\hat{\phi}_{ji}}{\sum_{k \in \mathcal{N}} \hat{\phi}_{jk} n_k} l_j \tag{F.76}$$

with $\hat{\phi}_{ij} \in (0, 1)$. A manufacturing firm produces a single unit of a manufactured good at a fixed price μ, using a single unit of the labor of mobile workers. Thus, we must have $n_i = x_i$. The profit function of the firm at i is given by

$$\Pi_i(\boldsymbol{x}) = \mu \sum_{j \in \mathcal{N}} \frac{\hat{\phi}_{ji}}{\sum_{k \in \mathcal{N}} \hat{\phi}_{jk} x_k} l_j - w_i. \tag{F.77}$$

Firms can freely enter and exit the city, thereby drawing zero profit. We abstract from commuting between different areas. Then, the wage of a mobile worker in area i equals

$$w_i(\boldsymbol{x}) = \mu \sum_{j \in \mathcal{N}} \frac{\hat{\phi}_{ji}}{\sum_{k \in \mathcal{N}} \hat{\phi}_{jk} x_k} l_j. \tag{F.78}$$

The indirect utility of the worker is set to be

$$v_i(\boldsymbol{x}) = w_i(\boldsymbol{x}) + \log[\Delta_i] - \gamma \log[x_i]. \tag{F.79}$$

where $\Delta_i \equiv \sum_{j \in \mathcal{N}} \phi_{ij} x_j$ denotes social utility as in the Beckmann model (Example 2.1).

Let $l_i = 1$ for all i. Also, assume that $\phi_{ij} = \hat{\phi}_{ij}$ for all i and j. Then, we see that $\mathbf{V} = \frac{1}{\bar{v}} G^{\sharp}(\bar{\mathbf{D}})$ with $G^{\sharp}(\chi) = -\gamma + \chi - \mu\chi^2$.

F.2.11 Allen and Arkolakis (2014) (AA) (Example 3.4)

The AA model is a perfectly competitive Armington (1969)-based framework with positive and negative local externalities. We introduce a discrete-space version of the AA model to fit our context. We also abstract away all exogenous differences across regions. In the model, productivity of a location is proportional to x_i^{α} with $\alpha > 0$, representing positive externalities. The market equilibrium condition is

$$s_i(\boldsymbol{x}, \boldsymbol{w}) = w_i x_i - \sum_{j \in \mathcal{N}} \frac{w_i^{1-\sigma} x_i^{\alpha(\sigma-1)} \phi_{ij}}{\sum_{k \in \mathcal{N}} w_k^{1-\sigma} x_k^{\alpha(\sigma-1)} \phi_{kj}} w_j x_j = 0. \tag{F.80}$$

With market wage $\boldsymbol{w}$, the payoff function is given by $v_i(\boldsymbol{x}) = x_i^{\beta} w_i \Delta_i^{\frac{1}{\sigma-1}}$ where $\Delta_i \equiv \sum_{k \in \mathcal{N}} w_k^{1-\sigma} x_k^{\alpha(\sigma-1)} \phi_{ki}$. The term x_i^{β} with $\beta < 0$ represents negative externalities from congestion, as in the Beckmann model (Example 2.1). Direct computation gives $\mathbf{V} = G^{\flat}(\bar{\mathbf{D}})^{-1} G^{\sharp}(\bar{\mathbf{D}})$ with

$$G^{\sharp}(\chi) = -(\alpha + \beta - \gamma_0) + (\alpha + \beta + \gamma_1)\chi, \tag{F.81}$$
$$G^{\flat}(\chi) = (\sigma + (\sigma - 1)\chi)(1 - \chi), \tag{F.82}$$

where $\gamma_0 \equiv \frac{1+\alpha}{\sigma}$ and $\gamma_1 \equiv \frac{1-\beta}{\sigma}$.

Remark F.6. In Figure 11, the parameters are set to $\alpha = 0.5$, $\beta = -0.3$, and $\sigma = 6$. For Figure 14a, we let $\beta = -0.6$ so that $\alpha + \beta < 0$. ■

Table F.1: Examples of quadratic net gain functions $G^{\sharp}(\chi) = c_0 + c_1\chi + c_2\chi^2$

Model class	Example	Local force	Global forces	
		c_0	c_1	c_2
Class I	Krugman (1991b)	0	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\left(\frac{\mu^2}{\sigma-1}+\frac{1}{\sigma}\right)$
	Puga (1999)	0	$\check{\mu}\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\left(\frac{\check{\mu}^2}{\sigma-1}+\frac{1}{\sigma}+\eta\right)$
	Forslid and Ottaviano (2003)	0	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\left(\frac{\mu^2}{\sigma(\sigma-1)}+1\right)$
	Pflüger (2004)	0	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\frac{\mu}{\sigma}(1+L)$
	Harris and Wilson (1978)	$1-\frac{1}{\alpha}$	0	-1
Class II	Helpman (1998)	$-\gamma$	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\left(\frac{\mu^2}{\sigma-1}+\frac{1}{\sigma}\right)+\gamma$
	Redding and Sturm (2008)	$-\gamma$	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)-\gamma\frac{\sigma-1}{\sigma}$	0
	Murata and Thisse (2005)	$-\hat{\theta}$	$(1-\hat{\theta})\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)-\hat{\theta}\frac{\sigma-1}{\sigma}$	0
	Allen and Arkolakis (2014)	$-(\alpha+\beta)+\frac{1+\alpha}{\sigma}$	$(\alpha+\beta)+\frac{1-\beta}{\sigma}$	0
	Beckmann (1976)	$-\gamma$	1	0
Class III	Tabuchi (1998)	$-\hat{\gamma}\left(\frac{1}{\sigma}+\frac{\sigma-1}{\sigma}\kappa\right)$	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\left(\frac{\mu^2}{\sigma-1}\theta+\frac{1}{\sigma}\hat{\theta}\right)$
	Pflüger and Südekum (2008)	$-\frac{\gamma}{1+L}$	$\mu\left(\frac{1}{\sigma-1}+\frac{1}{\sigma}\right)$	$-\frac{\mu}{\sigma}(1+L)$
	Takayama and Akamatsu (2011)	$-\gamma$	1	$-\mu$

Notes: The positive (negative) coefficients indicate agglomeration (dispersion) forces. Observe that Classes II and III incorporates negative constant term in χ. For definitions of parameters, see Appendix F. Although Helpman (1998) can have a global dispersion force, there is no such (μ, σ) that ensures $G(1) < 0$, thereby the model is Class II.