Equilibrium renement for a model of non-monocentric internal structures of cities: A potential game approach

Osawa, Minoru and Akamatsu, Takashi

Tohoku University

26 February 2019

Equilibrium Refinement for a Model of Non-monocentric Internal Structures of Cities: A Potential Game Approach[*]

Minoru Osawa[†] and Takashi Akamatsu[‡]

September 2, 2019

Abstract: Using the theory of potential games, this paper addresses the emergence of polycentric structures in cities, resulting from trade-offs between agglomeration economies and congestion effects. We consider a model that explains the formation of multiple business centers in cities as an equilibrium outcome under the presence of households' commuting costs and positive technological externalities between firms. We first show that the model is a large-population (non-atomic) potential game. To elucidate properties of stable spatial equilibria in the model, we consider local and global maximizations of the potential function of the model, which are known to correspond to various equilibrium refinement criteria. We find that (i) the formation of business centers (agglomeration of firms) is possible only when households' commuting costs are sufficiently low and that (ii) the size (number) of business centers increases (decreases) monotonically as communication between firms becomes easier.

Keywords: Agglomeration; multiple equilibria; equilibrium refinement; potential game.

JEL Classification: C62, C72, C73, R14

1   Introduction

The internal structure of a city is the outcome of the interplay between forces toward agglomeration and dispersion. The very high density of activities and population in cities induces various congestion effects, or dispersion forces, arising from higher land rents and longer commutes. At the same time, despite the disincentives, there are agglomeration forces that drive businesses or people to form clusters. In particular, agglomeration economies for businesses may arise due to nonmarket interactions, or knowledge spillovers, that enhance a firm's productivity when its employees can communicate closely with other workers in a city.[1]

In this paper, we address the emergence of polycentric city structures that arise from trade-offs between agglomeration economies and congestion effects. Toward this aim, we build on the seminal framework of Fujita and Ogawa (1982) (FO) to facilitate formal analyses. The FO framework incorporates two canonical agent types: *firms* and *households*. Firms are subject to agglomeration economies in the form of positive technological externalities, whereas households have to commute to their workplaces. Both agents compete for land. In this framework, firms benefit from proximity to other firms. If they agglomerate to form business districts, however, they must pay higher rents and also compensate their workers for longer commutes with higher wages. A spatial distribution of firms and households in equilibrium is determined by balancing these opposing forces. In urban economics, the framework is known as a canonical model that can describe the formation of polycentric urban spatial structures in equilibrium.[2]

Typical for models with positive externalities, a multiplicity of equilibria is inherent in the FO framework. Equilibrium refinement is crucial for drawing robust implications, since some equilibria may be unattainable under any dynamics. Toward this goal, we employ the theory of *potential games* (Monderer and Shapley, 1996; Sandholm, 2001). We show that the FO model is an instance of large-population (non-atomic) potential games; the model is reduced to a maximization problem of a scalar-valued function (a *potential function*) that encapsulates its incentive structure. In particular, the potential function is shown to capture the trade-off between the agglomeration and dispersion forces in the model.

In potential games, sensible approaches are available for the characterization and refinement of equilibria. First, the set of Nash equilibria of a potential game coincides with that of Karush–Kuhn–Tucker (KKT) points of the maximization of the associated potential function. Second, the local maximizers of the potential function is known to be *locally stable* under various standard dynamics, including the *best response dynamic* (Gilboa and Matsui, 1991), 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).[3] See Sandholm (2001) as well as Sandholm (2010), Section 8.2. Third, the set of the global potential maximizers are known to be *globally stable* under various premises. Building on Blume (1993, 1997), Sandholm (2010) (Chapter 12) showed that the set of global potential maximiz-

---

[1] Agglomeration economies can be categorized into sharing, matching, and learning (Duranton and Puga, 2004).

[2] See surveys by, e.g., Anas et al. (1998) or Duranton and Puga (2015).

[3] As shown in their paper, Riemannian game dynamics encompass the *replicator dynamic* (Taylor and Jonker, 1978) and the *projection dynamic* (Dupuis and Nagurney, 1993) as their special cases.

ers in a non-atomic potential game is *stochastically stable*.[4] Also, Oyama (2009a,b) considered equilibrium selection in a dynamic setting where the static payoff is given by a "new economic geography" model; it is shown that a perfect foresight dynamic selects the global potential maximizer if the instantaneous payoff admits a potential function. Thus, the properties of local and global potential maximizer(s) are expected to reveal the essence of a potential game.

Given this background, we employ local and global maximizations of the potential function as equilibrium refinement criteria. To concretely illustrate the utility of the potential game approach in the context of the FO framework, we study the properties of spatial equilibria in two stylized geographies, namely, a two-location city and a circular city. We find that the formation of business centers can be a global potential maximizer but only when the commuting costs of households are sufficiently low. It is also shown that the number of business centers monotonically decreases, while their size increases, when either the commuting costs of households goes lower monotonically or the level of agglomeration externalities between firms diminishes monotonically. These results are consistent with intuition and with conjectural predictions in the prior literature.

This paper is organized as follows. Section 2 relates this study to the literature. Section 3 introduces a model of urban spatial configurations and shows that a potential maximization problem characterizes equilibria of the model. Section 4 shows that the model can be decomposed and then reduced to an equivalent potential game in which only firms are the agents. Sections 5 and 6 are example sections that study Fujita and Ogawa (1982)'s specification using the potential game method. Section 5 studies the simplest geography in which a city is divided into two cells. Section 6 considers the circular geography to demonstrate that the formation of multiple business centers can be a global potential maximizer. Section 7 concludes the paper.

2 Related literature

As highlighted in Section 1, this paper adopts a seminal framework proposed by Masahisa Fujita and Hideaki Ogawa in two papers, Ogawa and Fujita (1980) and Fujita and Ogawa (1982). They initiated the study of urban spatial equilibrium models with multiple types of agents and economics of agglomeration.[5] This framework has since been extended in many directions, including two-dimensional space (Ogawa and Fujita, 1989), a monopolistic competition framework (Fujita, 1988), multi-unit firms (Ota and Fujita, 1993), and general equilibrium frameworks (Lucas, 2001; Lucas and Rossi-Hansberg, 2002; Berliant et al., 2002; Mossay et al., 2017; Malykhin and Ushchev, 2018). Most notably, the latest generation of quantitative spatial economic models (e.g., Ahlfeldt et al., 2015; Owens et al., 2019), implicitly or explicitly inherits many elements from these theories (see the survey by Redding and Rossi-Hansberg, 2017). Due to the lack of effective analytical methods, however, no previous study has addressed stability of equilibria in the FO framework when there are multiple equilibria. This paper is the first to theoretically investgate this issue.

---

[4]For surveys of stochastic stability concepts in game theory, see Sandholm (2010) (Chapters 11 and 12) as well as Wallace and Young (2015).

[5]Study of *socially optimal* urban spatial structure under similar assumptions was explored by other researchers, such as Imai (1982).

Section 6 employs circular geography. Circular geography is one of the canonical setups for theoretical investigations, since the setup abstracts away various effects of geographical asymmetries.[6] Related to this paper, Mossay and Picard (2011) and Blanchet et al. (2016) considered a spatial equilibrium model with a *single type* of mobile agents (Beckmann (1976)'s model) over a circle. This paper, in contrast, has *two types* of mobile agents. Mossay and Picard (2011) showed that the formation of equidistant, disjointed cities over the circumference is the only possible equilibrium outcome. Furthermore, the equilibria are Pareto-ranked according to the number of cities. In particular, the formation of a single city is the social optimum in their framework. In our context, the multiple-city equilibrium can be interpreted as the formation of multiple business centers in a city. In our setup, however, such symmetric spatial configurations cannot be Pareto-ranked according to the number of business centers (see Sections 6.2 and 6.3). Pareto ranking depends on parameters. The contrast is natural since our model considers two types of mobile agents.

To avoid technical complications, we consider a discrete-space model, where a city is divided into a finite number of cells. A standard approach in the urban economics, however, is to suppose a continuous space. Fujita and Ogawa (1982) considered a continuous real line as the underlying geography. Such models are related to non-atomic games with continuous strategy sets and a continuum of agents. A general theory of potential games for such class of games, along with associated dynamics, was recently proposed by Cheung and Lahkar (2018). Blanchet et al. (2016) characterized the spatial equilibria of a general class of continuous-space urban models with a single type of agent as a variational problem. Bragard and Mossay (2016) studied a relocation dynamics in the continuous-space framework of Mossay and Picard (2011). Akamatsu et al. (2017) formulated a discrete-space version of Beckmann (1976)'s model as a potential game and elucidated its properties. Unification of this line of researches can lead to a fruitful field of applications of the potential game method.

3   A model for urban spatial structure

Building on the FO framework, this section formulates a discrete-space urban model. The model generalizes the original FO model in that a location's production level (introduced immediately below) can take arbitral form as long as it satisfies certain assumptions.

Consider a city comprised of a set of discrete cells $\mathcal{I} \equiv \{1, 2, \ldots, I\}$ ($I \geq 2$). Each cell $i \in \mathcal{I}$ is endowed with $a_i > 0$ units of a fixed supply of land. For normalization, we let $\sum_{i \in \mathcal{I}} a_i = 1$. There are two continuums of agents: firms and households. Since the model focuses mainly on the spatial distribution of firms, households' incentives are simplified (see Section 3.2).

3.1   Firms

There is a continuum of business firms. The set of strategies for firms is $\mathcal{I}$. Each firm chooses a cell $i \in \mathcal{I}$ to locate in. The total mass of firms is a fixed constant M. Let $m_i \geq 0$ be the

---

[6]For instance, Papageorgiou and Smith (1983); Krugman (1993); Mossay and Picard (2011); Akamatsu et al. (2012); Ikeda et al. (2012); Blanchet et al. (2016); Osawa et al. (2017); Ikeda et al. (2019).

mass of firms in $i \in \mathcal{I}$, and let $\boldsymbol{m} \equiv (m_i)_{i \in \mathcal{I}}$ denote the spatial distribution, i.e., the strategy distribution of firms.

Every firm produces a single unit of goods exported to the outside world with a fixed price (which is normalized to unity) prevailing there. Each firm requires one unit of land and $\phi > 0$ units of labor to operate. The profit of a firm choosing cell i is given by

$$\pi_i = A_i(\boldsymbol{m}) - \phi w_i - r_i, \tag{3.1}$$

where $A_i : \mathbb{R}_+^I \to \mathbb{R}_+$ is the level of production in cell i, $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$ is the set of nonnegative real numbers, $w_i \geq 0$ and $r_i \geq 0$ are, respectively, wages and land rents prevailing in cell i, which are endogenously determined in the perfectly competitive labor and land markets. We let $\boldsymbol{w} \equiv (w_i)_{i \in \mathcal{I}}$ and $\boldsymbol{r} \equiv (r_i)_{i \in \mathcal{I}}$.

The set of all possible spatial configurations of firms is given by the following:

$$\mathcal{M} \equiv \left\{ \boldsymbol{m} \in \mathbb{R}_+^I \; \middle| \; \sum_{i \in \mathcal{I}} m_i = M, \, 0 \leq m_i \leq a_i \; \forall i \in \mathcal{I} \right\}, \tag{3.2}$$

because each firm uses a single unit of land in production.

External economies in production, or technological externalities, arise from nonmarket interactions. Firms are supposed to produce more goods when they are close to other firms. We impose the following assumption on $\mathbf{A}(\boldsymbol{m}) = (A_i(\boldsymbol{m}))_{i \in \mathcal{I}}$.

Assumption 1. *The function* $\mathbf{A} : \mathbb{R}_+^I \to \mathbb{R}_+^I$ *is* C^1 *and*

(i) $\nabla \mathbf{A}(\boldsymbol{m}) = \left[\frac{\partial A_i(\boldsymbol{m})}{\partial m_j} \right]$ *is positive definite with respect to* $T\mathcal{M} \equiv \left\{ \boldsymbol{\epsilon} = (\epsilon_i) \in \mathbb{R}^I \mid \sum_{i \in \mathcal{I}} \epsilon_i = 0 \right\}$
for all $\boldsymbol{m} \in \mathcal{M}$. *That is,* $\boldsymbol{\epsilon}^\top \nabla \mathbf{A}(\boldsymbol{m}) \boldsymbol{\epsilon} > 0$ *for all nonzero* $\boldsymbol{\epsilon} \in T\mathcal{M}$ *for all* $\boldsymbol{m} \in \mathcal{M}$.

(ii) $\nabla \mathbf{A}(\boldsymbol{m})$ *is symmetric for all* $\boldsymbol{m} \in \mathbb{R}_+^I$.

Assumption 1 (i) is a condition under which the production level $\mathbf{A}(\boldsymbol{m})$ features positive externalities. Consider a deviation in the spatial distribution of firms such that $\boldsymbol{m}' = \boldsymbol{m} + \boldsymbol{\epsilon} \in \mathcal{M}$ with an infinitesimal $\boldsymbol{\epsilon} \in T\mathcal{M}$. Then, Assumption 1 (i) implies that

$$\sum_{\epsilon_i \geq 0} |\epsilon_i| \left(A_i(\boldsymbol{m}') - A_i(\boldsymbol{m}) \right) - \sum_{\epsilon_i < 0} |\epsilon_i| \left(A_i(\boldsymbol{m}') - A_i(\boldsymbol{m}) \right) = \boldsymbol{\epsilon}^\top \nabla \mathbf{A}(\boldsymbol{m}) \boldsymbol{\epsilon} > 0. \tag{3.3}$$

The first (second) term on the left hand side represents the aggregate gain in the level of externalities in the cells with more (less) firms induced by the deviation. Thus, Assumption 1 (i) implies that the former always dominates the latter, meaning a *self-reinforcing externality*. Suppose that a hypothetical firm leaves a cell and joins another. Then, improvements in the level of positive externalities in the cell to which the firm switches dominate those found in one left.

Assumption 1 (ii) is *externality symmetry* (Sandholm, 2001), under which the marginal effect of adding a firm to cell i on $A_j(\boldsymbol{m})$ is the same as the marginal effect of adding a firm to cell j on $A_i(\boldsymbol{m})$. As we will see, Assumption 1 (ii) is a necessary and sufficient condition for $\mathbf{A}(\boldsymbol{m})$ to be integrable.

5

Example 1. Positive externalities *within* each cell is expressed by $A_i(\boldsymbol{m}) = m_i^\alpha$ with $\alpha > 0$. $\Diamond$

Example 2. Fujita and Ogawa (1982) assumed that $A_i(\boldsymbol{m}) = \sum_{j \in \mathcal{I}} \exp\left(-\tau \ell_{ij}\right) m_j$ where $\tau > 0$ and $\ell_{ij} \geq 0$ is the distance between cells i and j with $\ell_{ij} = \ell_{ji}$. $\Diamond$

3.2 Households

Households are another continuum of actors. The set of strategies for households is $\mathcal{I} \times \mathcal{I}$. Each household chooses in which cell $i \in \mathcal{I}$ to reside and to which cell $j \in \mathcal{I}$ to commute. The mass of households that commute from cell i to cell j is denoted by $n_{ij} \geq 0$. The commuting pattern as a whole (i.e., the strategy distribution of households) is denoted by $\boldsymbol{n} \equiv (n_{ij})_{ij \in \mathcal{I} \times \mathcal{I}}$. There also is an outside option. Households can freely migrate from or toward the outside world, where their reservation utility is normalized to be 0. The equilibrium mass of households in the city is therefore endogenously determined.[7]

Each household is endowed with a single unit of labor, supplied to firms, and compensated by a wage. Each household consumes one unit of land inelastically for residential purposes. The indirect utility of a household that chooses cell i for its residential location and cell j as its job location is supposed to be

$$v_{ij} = w_j - t_{ij} - r_i, \tag{3.4}$$

where the constant $t_{ij} \geq 0$ is the commuting cost from i to j in the monetary unit.

3.3 Spatial equilibrium

In equilibrium, every firm should maximize its own profit by choosing its location. Every household should maximize its own utility. These conditions are expressed as follows:

$$\text{(utility maximization)} \quad 0 \leq n_{ij} \perp v^* - v_{ij} \geq 0 \qquad \forall ij \in \mathcal{I} \times \mathcal{I}, \tag{3.5}$$

$$\text{(profit maximization)} \quad 0 \leq m_i \perp \pi^* - \pi_i \geq 0 \qquad \forall i \in \mathcal{I}, \tag{3.6}$$

where the notation $0 \leq q \perp p \geq 0$ represents the complementarity condition $q \geq 0$, $p \geq 0$, and $qp = 0$. The scalars $v^* = 0$ and $\pi^* \geq 0$ are, respectively, the reservation utility of households and an equilibrium profit level of firms.

Land and labor markets are perfectly competitive and markets clear when the following conditions are met, thereby $\boldsymbol{w}$ and $\boldsymbol{r}$ are endogenously determined:

$$\text{(land market clearing)} \quad 0 \leq a_i - m_i - \sum_{j \in \mathcal{I}} n_{ij} \perp r_i \geq 0 \qquad \forall i \in \mathcal{I}, \tag{3.7}$$

$$\text{(labor market clearing)} \quad 0 \leq \sum_{i \in \mathcal{I}} n_{ij} - \phi m_j \perp w_j \geq 0 \qquad \forall j \in \mathcal{I}. \tag{3.8}$$

Following conventions in urban economic theory, we assume that land is owned by absentee landlords who spend their rental revenues outside the city. We normalize the opportunity cost of land to 0, so that $\min_{i \in \mathcal{I}} r_i = 0$.

---

[7] In urban economics, this type of models is called *open-city models*. See Fujita (1989).

Lastly, since the total mass of firms in the city is fixed, we require

$$\text{(conservation)} \quad \sum_{i \in \mathcal{I}} m_i = M. \tag{3.9}$$

To summarize, an equilibrium in the model is defined as follows:

Definition 1. A *spatial equilibrium* is a pair of strategy distributions of households and firms $(\boldsymbol{n}, \boldsymbol{m})$ with its associated market prices and equilibrium profits level of firms $(\boldsymbol{r}, \boldsymbol{w}, \pi^*)$ that satisfy the conditions (3.5), (3.6), (3.7), (3.8), and (3.9).

Remark 1. Let N be the equilibrium mass of households. Then, $N = \phi M$ because every household in the city should supply its labor to some firm, and every firm requires ϕ units of labor to operate. Note also that, from (3.7), $M + N = (1 + \phi)M \leq \sum_{i \in \mathcal{I}} a_i = 1$, or $M \leq \frac{1}{1+\phi} < 1$, must hold true for a spatial equilibrium to exsit. The restriction on M ensures that the total land endowments meet the total land demand in equilibrium. ◇

3.4 Potential maximization formulation

Spatial equilibria are characterized by a maximization problem of a scalar-valued function over the set of all feasible strategy distributions $(\boldsymbol{m}, \boldsymbol{n})$ of firms and households. In this sense, the model can be interpreted as a non-atomic potential game.[8]

Let $h : \mathbb{R}_+^{I^2} \to \mathbb{R}$ denote total commuting costs of households:

$$h(\boldsymbol{n}) = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} t_{ij} n_{ij}. \tag{3.10}$$

Also, note that Assumption 1 (ii) is necessary and sufficient for the existence of a function $f : \mathbb{R}_+^I \to \mathbb{R}$ such that $\nabla f(\boldsymbol{m}) = \left[\frac{\partial f(\boldsymbol{m})}{\partial m_i}\right] = \mathbf{A}(\boldsymbol{m})$ for all $\boldsymbol{m} \in \mathbb{R}_+^I$. In the integral form,

$$f(\boldsymbol{m}) = \oint \mathbf{A}(\boldsymbol{\omega}) \cdot \mathrm{d}\boldsymbol{\omega}, \tag{3.11}$$

where the path integral is taken over any piecewise smooth path from $\mathbf{0}$ to $\boldsymbol{m}$.

In addition, we note that any pair of strategy distributions $(\boldsymbol{m}, \boldsymbol{n})$ should satisfy

$$\text{(land supply)} \quad m_i + \sum_{j \in \mathcal{I}} n_{ij} \leq a_i \qquad \forall i \in \mathcal{I}, \tag{3.12}$$

$$\text{(labor supply)} \quad \sum_{i \in \mathcal{I}} n_{ij} \geq \phi m_j \qquad \forall j \in \mathcal{I}, \tag{3.13}$$

as well as the conservation condition (3.9). We have the following result.[9]

Proposition 1. *Suppose Assumption 1 (ii). Then, any spatial equilibrium of the model is a KKT point of the following optimization problem:*

$$\max_{(\boldsymbol{m}, \boldsymbol{n}) \geq \mathbf{0}} f(\boldsymbol{m}) - h(\boldsymbol{n}) \quad \text{s.t.} \quad \text{(3.12), (3.13), and (3.9)} \tag{P0}$$

---

[8]See Appendix B for a discussion.
[9]All proofs are relegated to Appendix A.

where r, w, and π^* are determined as the Lagrange multipliers for the land supply constraint (3.12), the labor supply constraint (3.13), and the conservation constraint (3.9), respectively.

The objective function of (P0) encapsulates the structure of the model. The first term $f(m)$ is a measure of firms' benefit from agglomeration. The second term $h(n)$ is the total commuting cost that is spent in the city by the commuting pattern n. Thus, (P0) is a problem that pertains to maximizing the benefit of firms' agglomeration while also reducing the commuting costs of households. The former corresponds to the agglomeration force, whereas the latter corresponds to the dispersion force (see Section 4.4).

4 Reduction to firms' game

To simplify the analysis of the properties of spatial equilibria, we decompose the model into a two-stage game by postulating that the adjustment of households' decision is faster than that of firms. In other words, we interpret the model as a Stackelberg leader–follower game, where firms lead and households follow, assuming that households play only passive and static roles in the game.[10] Since an individual firm or household is infinitesimally small, such decomposition does not affect the properties of equilibria (see Remark 2).[11]

In the lower-stage game, the strategy distribution (spatial distribution) of firms $m \in \mathcal{M}$ is fixed, and only households choose their strategies $ij \in \mathcal{I} \times \mathcal{I}$ (residential location and workplace). In the upper-stage game, firms choose locations $i \in \mathcal{I}$ to maximize their profit in response to the households' aggregate behavior, as well as market prices, in the lower stage.

4.1 Lower-stage game

Let $\mathcal{N}(m)$ denote the closed and convex set of all feasible commuting patterns, or strategy distributions of households, under a fixed $m \in \mathcal{M}$:

$$\mathcal{N}(m) \equiv \left\{ n \in \mathbb{R}_+^{I^2} \;\middle|\; \text{s.t. (3.12) and (3.13)} \right\}. \tag{4.1}$$

Then, equilibria in the lower-stage game are defined as follows.

Definition 2. Fix $m \in \mathcal{M}$. An equilibrium in the lower-stage game is a commuting pattern $n \in \mathcal{N}(m)$ and its associated market prices (r, w) that satisfy conditions (3.5), (3.7), and (3.8).

We denote an equilibrium value of r in the lower-stage game by $\check{r} = (\check{r}_i)$. Note that $\check{r}$ do not necessarily coincide with the equilibrium land rent in the upper-stage game because firms' strategies (locations) are fixed in the lower stage. Since firms have no incentive to raise wage w in the upper stage, wage should coincide with the upper-stage value.

Solving a lower-stage game becomes an optimization problem.

---

[10]This type of decomposition is a standard approach in spatial economics. For example, economic geography models á la Krugman (1991) consider the so-called *short-run equilibrium*, where market prices are determined under the condition that the spatial distribution of mobile actors, m, is fixed. It defines the payoff of mobile agents as a function of m. Then, the so-called *long-run equilibrium* considers the spatial distribution of mobile agents in equilibrium, according to the payoff function.

[11]From an optimization theory perspective, this interpretation corresponds to a *Benders' Decomposition* (Benders, 1962; Geoffrion, 1972) of problem (P0). Other possible decompositions of (P0) are discussed in Appendix C.

Proposition 2. *Any lower-stage equilibrium under a fixed $m \in \mathcal{M}$ has the following properties:*

(a) The aggregate commuting pattern n is a KKT point for the following problem:

$$\min_{n \in \mathcal{N}(m)} h(n). \tag{S}$$

(b) The corresponding market prices $(\check{r}, w)$ are a KKT point for the following problem:

$$\max_{(\check{r}, w) \geq 0} \quad \phi \sum_{i \in \mathcal{I}} w_i m_i - \sum_{i \in \mathcal{I}} \check{r}_i (a_i - m_i) \quad s.t. \quad 0 \geq w_j - t_{ij} - \check{r}_i \quad \forall i, j \in \mathcal{I}, \tag{D}$$

where we normalize $\check{r}$ and w by letting $\min_{i \in \mathcal{I}} \{\check{r}_i\} = 0$.

The problem (S) minimizes total commuting costs across the city while satisfying land and labor market constraints. On the other hand, (D) is the dual problem for (S). The first term of its objective is total wages received by households, and the second term is total land rent paid. It determines the highest bid $\check{r}$ for land and the lowest possible wage w to compensate for commuting costs and land rents. Also note that a bounded solution exists for both (S) and (D), because (3.12), (3.13), and the nonnegativity constraint of n define a nonempty, closed, and convex set.

4.2 Upper-stage game

Knowing households' responses, firms choose their location $i \in \mathcal{I}$. The "lower-stage" profit $\check{\pi}(m) = (\check{\pi}_i(m))_{i \in \mathcal{I}}$ of firms is defined as

$$\check{\pi}_i(m) = A_i(m) - c_i(m), \tag{4.2}$$

where $c_i(m)$ is the firms' minimum possible cost required to operate in cell $i \in \mathcal{I}$. With lower-stage market prices $(w(m), \check{r}(m))$ (i.e., a solution for (D)), $c_i(m)$ is given by

$$c_i(m) \equiv \phi w_i(m) + \check{r}_i(m). \tag{4.3}$$

In the upper stage, firms may have to pay higher land rents $r_i > \check{r}_i(m)$ to win the competition for land. Let $\hat{r}_i \equiv r_i - \check{r}_i(m) \geq 0$. With $\check{\pi}(m)$ and $\hat{r}_i$, the profit-maximization condition for firms in (3.6) can be rewritten as follows:[12]

$$0 \leq m_i \perp \pi^* - (\check{\pi}_i(m) - \hat{r}_i) \geq 0 \qquad \forall i \in \mathcal{I}. \tag{4.4}$$

Regarding $\hat{r} = (\hat{r}_i)$, the following conditions should be met:

$$0 \leq a_i - m_i \perp \hat{r}_i \geq 0 \qquad \forall i \in \mathcal{I}. \tag{4.5}$$

That is, if households cannot afford to reside in cell i (i.e., if $r_i = \check{r}_i + \hat{r}_i > \check{r}_i$), the cell is completely occupied by firms. Conversely, if $m_i < a_i$ and $\sum_{j \in \mathcal{I}} n_{ij} > 0$, land rents paid by firms

---

[12]Firms' profit (3.1) can be rewritten as $\pi_i = A_i - \phi w_i - r_i = A_i - \phi w_i - (\check{r}_i + \hat{r}_i) = A_i - c_i - \hat{r}_i = \check{\pi}_i - \hat{r}_i$.

and households must be the same, and $\hat{r}_i = 0$.

Equilibrium conditions for the upper-stage game can be summarized as follows.

Definition 3. An equilibrium in the upper-stage game is a strategy distribution of firms $\boldsymbol{m} \in \mathcal{M}$ with its associated market prices and equilibrium profits of firms $(\hat{\boldsymbol{r}}, \pi^*)$ that satisfy (4.4), (4.5), and (3.9).

4.3 Potential maximization formulation for the upper-stage game

The upper-stage game reduces to an optimization problem. First, observe that the lower-stage game induces a function of $\boldsymbol{m}$. Let $\check{h} : \mathcal{M} \to \mathbb{R}$ denote the optimal value of (S) as a function of $\boldsymbol{m}$. With a solution $\boldsymbol{n}^*(\boldsymbol{m})$ to (S), we let[13]

$$\check{h}(\boldsymbol{m}) = h(\boldsymbol{n}^*(\boldsymbol{m})). \tag{4.6}$$

In the following, we assume an extension $\check{h} : \mathbb{R}_+^I \to \mathbb{R}_+$ with an abuse of notation.[14]

We observe that $\check{h}(\boldsymbol{m})$ is the integral of the firms' cost function $\boldsymbol{c}(\boldsymbol{m}) = (c_i(\boldsymbol{m}))_{i \in \mathcal{I}}$. To see this, note that $\check{h}(\boldsymbol{m})$ is also the optimal value of (D) because of the strong duality of linear programming. Thus, the envelope theorem regarding dual representation (D) implies that

$$\frac{\partial \check{h}(\boldsymbol{m})}{\partial m_i} = \phi w_i(\boldsymbol{m}) + \check{r}_i(\boldsymbol{m}) = c_i(\boldsymbol{m}) \tag{4.7}$$

for almost all $\boldsymbol{m} \in \mathcal{M}$.[15]

Define a function $g : \mathbb{R}_+^I \to \mathbb{R}$ by $g(\boldsymbol{m}) \equiv f(\boldsymbol{m}) - \check{h}(\boldsymbol{m})$. Then, g turns out to be the *potential function for* $\check{\boldsymbol{\pi}}(\boldsymbol{m})$. Recalling that $\nabla f(\boldsymbol{m}) = \mathbf{A}(\boldsymbol{m})$, we have

$$\frac{\partial g(\boldsymbol{m})}{\partial m_i} = \frac{\partial f(\boldsymbol{m})}{\partial m_i} - \frac{\partial \check{h}(\boldsymbol{m})}{\partial m_i} = A_i(\boldsymbol{m}) - c_i(\boldsymbol{m}) = \check{\pi}_i(\boldsymbol{m}) \tag{4.8}$$

for almost all $\boldsymbol{m} \in \mathcal{M}$. Thus, $g(\boldsymbol{m})$ is a potential function for $\check{\boldsymbol{\pi}}(\boldsymbol{m})$.

To summarize, under Assumption 1 (ii), the upper-stage game can be interpreted as a non-atomic potential game with strategy set $\mathcal{I}$, state space $\mathcal{M}$, payoff function $\check{\boldsymbol{\pi}} : \mathcal{M} \to \mathbb{R}^I$, and a potential function $g : \mathbb{R}_+^I \to \mathbb{R}$. In particular, the following result holds true.

Proposition 3. *Suppose Assumption 1 (ii). Then, any equilibrium spatial distribution $\boldsymbol{m}$ in the upper-stage game is a KKT point for the following maximization problem:*

$$\max_{\boldsymbol{m} \in \mathcal{M}} \; g(\boldsymbol{m}) = f(\boldsymbol{m}) - \check{h}(\boldsymbol{m}). \tag{P}$$

---

[13]Since (S) is a linear programming problem, in general, the equilibrium commuting pattern $\boldsymbol{n}^*(\boldsymbol{m})$ may not be uniquely determined. However, because the minimized commuting cost $h(\boldsymbol{n}^*(\boldsymbol{m}))$ is uniquely given for any $\boldsymbol{m} \in \mathcal{M}$, it can be regarded as a function of $\boldsymbol{m}$.

[14]Such an extension is simple because $\check{h}$ is a piecewise affine function in $\boldsymbol{m}$. We extend each affine function defined on a nonempty subset of $\mathcal{M}$ to that defined on $\mathbb{R}_+^I$ and then take the point-wise maximum.

[15]Note that, in the standard sense, $\check{h}$ can be nondifferentiable in a measure-zero subset of $\mathcal{M}$. See Remark B.5 in Appendix B.

Remark 2. Any equilibrium in the upper-stage game is a spatial equilibrium of the overall model (Section 3), and vice versa. This follows from a known fact that (P) is *equivalent* to (P0) under the condition that (S) admits solution for all $\boldsymbol{m} \in \mathcal{M}$, (Tuy, 1987, Proposition 2.1). In particular, it implies that a pair of strategy distributions $(\boldsymbol{m}, \boldsymbol{n})$ locally (globally) maximizes $f(\boldsymbol{m}) - h(\boldsymbol{n})$ if and only if $\boldsymbol{m}$ locally (globally) maximizes $g(\boldsymbol{m}) = f(\boldsymbol{m}) - \check{h}(\boldsymbol{m})$. $\Diamond$

4.4 General properties

The shape of the potential function g provides general insights into the properties of equilibria. In g, the two terms f and $\check{h}$ encapsulate the two basic forces in the model laid out in Section 3, one toward firms' dispersion, and the other toward firms' agglomeration. This is reflected in the convexity properties of f and $\check{h}$.

We observe that $\check{h}$ is convex.[16]

Observation 1. $\check{h} : \mathbb{R}_+^I \to \mathbb{R}$ *is piecewise affine and convex on* $\mathcal{M}$. $\Diamond$

This reflects that the second term corresponds to the *dispersion forces* for firms. Recall that $\check{h}$ represents households' minimized commuting costs (the primal representation (S)) and, equivalently, the firms' minimum possible operating costs (the dual representation (D)) at $\boldsymbol{m} \in \mathcal{M}$. Observation 1 means that concentration of firms in a smaller number of cells induces higher total commuting costs for households while also implying that firms' higher operational costs compensate for these greater commuting costs with higher wages. Also, $\check{h}(\boldsymbol{m}) = 0$ for all $\boldsymbol{m}$ such that $(1 + \phi)m_i \leq a_i$ for all $i \in \mathcal{I}$. That is, if all households can reside at their workplaces when firms are dispersed throughout a city, then no commuting costs are spent. In sum, the second term of the potential function g is larger *when firms are dispersed*.

On the other hand, f is strictly convex.[17]

Observation 2. *Under Assumption 1 (i)*, $f : \mathbb{R}_+^I \to \mathbb{R}$ *is strictly convex on* $\mathcal{M}$. $\Diamond$

The convexity of f reflects that it corresponds to the *agglomeration force* for firms. Firms' production levels are higher when they are located close to other firms. Representing the agglomerative effects, f is larger near the corners of $\mathcal{M}$, or *when firms are agglomerated* in a smaller number of cells.

Whether firms agglomerate or disperse is determined by the balance of two opposing forces expressed by the two terms f and $-\check{h}$. In particular, the potential function g is generally neither concave nor convex but *piecewise strictly convex*, indicating that there may be numerous possible equilibria.

Remark 3 (Efficiency). A welfare-maximization problem for the model may be formulated as follows:

$$\max_{\boldsymbol{m} \in \mathcal{M}} \; \tilde{g}(\boldsymbol{m}) \equiv \tilde{f}(\boldsymbol{m}) - \check{h}(\boldsymbol{m}), \tag{W}$$

---

[16]Note that $\check{h}(\boldsymbol{m})$ is a solution for (D) and therefore a point-wise maximum of affine functions of $\boldsymbol{m}$.

[17]Assumption 1 (i) is the necessary and sufficient condition for the strict convexity of f on the hyperplanes of the form $\mathcal{M}(\alpha) \equiv \left\{ \boldsymbol{m} \in \mathbb{R}^I \mid \sum_{i \in \mathcal{I}} m_i = \alpha \right\}$ $(\alpha \in \mathbb{R})$, implying the assertion since $\mathcal{M} = \mathcal{M}(M)$.

where $\tilde{f}(\boldsymbol{m}) \equiv \sum_{i \in \mathcal{I}} A_i(\boldsymbol{m}) m_i$, and $\tilde{g}$ represents firms' aggregate profits. In equilibrium, firms' agglomeration is inadequate when compared with a social optimum. To demonstrate this, we build on Sandholm (2001), Section 5. Suppose that $\mathbf{A}$ is homogeneous of degree $k > 0$, that is, $\mathbf{A}(\alpha \boldsymbol{m}) = \alpha^k \mathbf{A}(\boldsymbol{m})$ for all $\alpha \in \mathbb{R}$ and $\boldsymbol{m} \neq \boldsymbol{0}$. Then, we see that

$$f(\boldsymbol{m}) = \frac{1}{k+1} \tilde{f}(\boldsymbol{m}). \tag{4.9}$$

This relationship implies that (P) assigns a smaller weight to $\tilde{f}(\boldsymbol{m})$ than (W) does. Because $\tilde{f}(\boldsymbol{m}) = (k+1) f(\boldsymbol{m})$, $\tilde{f}(\boldsymbol{m})$ also preferres agglomeration. Clustering of firms in equilibria is therefore less than efficient.

Next, suppose that $\mathbf{A}$ is characterized by a parameter $\tau > 0$. Noting that $h(\boldsymbol{n})$ is a linear function of $\boldsymbol{t} = (t_{ij})$, suppose that $\check{h}(\boldsymbol{m})$ is linear in a parameter $t > 0$. Then, it follows that

$$g(\boldsymbol{m}) = f(\boldsymbol{m}) - \check{h}(\boldsymbol{m}) = \frac{1}{k+1} \left(\tilde{f}(\boldsymbol{m}) - (k+1)\check{h}(\boldsymbol{m}) \right) = \frac{1}{k+1} \left. \tilde{g}(\boldsymbol{m}) \right|_{t:=(k+1)t}.$$

The relationship implies that, if $\boldsymbol{m}^* \in \mathcal{M}$ locally (globally) maximizes the potential function g for a parameter pair (τ^*, t^*), then it locally (globally) maximizes the welfare function $\tilde{g}$ for the pair $(\tau^*, (k+1)t^*)$. Also, if a spatial pattern $\boldsymbol{m}$ cannot globally maximize the potential, it cannot be a global social optimum. $\diamond$

4.5 The Fujita–Ogawa specification

In the remainder of the paper, we focus on the simple specification of $\mathbf{A}(\boldsymbol{m}) = (A_i(\boldsymbol{m}))_{i \in \mathcal{I}}$ and $\boldsymbol{t} = (t_{ij})_{ij \in \mathcal{I} \times \mathcal{I}}$ to derive concrete insights. Under the specification, the model can be interpreted as a discrete-space version of Fujita and Ogawa (1982)'s seminal model.

Assumption 2. $A_i(\boldsymbol{m}) = \sum_{j \in \mathcal{I}} d_{ij} m_j$ where $d_{ij} = \exp\left(-\tau \ell_{ij}\right)$ with $\tau > 0$ and $t_{ij} = t \ell_{ij}$ with $t > 0$. $\ell_{ij} \geq 0$ denotes the distance between cell $i, j \in \mathcal{I}$, such that $\ell_{ii} = 0$, $\ell_{i \neq j} > 0$, and $\ell_{ij} = \ell_{ji}$.

The parameter τ determines how fast technological externalities decay over distance, where a larger τ indicates a higher decay rate. On the other hand, $t > 0$ is the commuting cost rate per unit of distance.[18] We note that Assumption 2 is sufficient for Assumption 1.[19] Also, Assumption 2 implies that

$$f(\boldsymbol{m}) = \frac{1}{2} \boldsymbol{m}^\top \mathbf{D} \boldsymbol{m} = \frac{1}{2} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} d_{ij} m_i m_j \tag{4.10}$$

where $\mathbf{D} = [d_{ij}]$. We confirm that $\nabla f(\boldsymbol{m}) = \frac{1}{2}(\mathbf{D} + \mathbf{D}^\top)\boldsymbol{m} = \mathbf{D}\boldsymbol{m} = \mathbf{A}(\boldsymbol{m})$ for all $\boldsymbol{m} \in \mathbb{R}_+^I$.

Remark 4. Under the FO specification, $\mathbf{A}$ is homogeneous of degree 1. The social welfare function is given by $\tilde{g}(\boldsymbol{m}) = 2f(\boldsymbol{m}) - \check{h}(\boldsymbol{m})$ (see Remark 3). If $\boldsymbol{m}^* \in \mathcal{M}$ locally (globally)

---

[18]Households' equilibrium commuting pattern $\boldsymbol{n}^*(\boldsymbol{m})$ in the lower-stage game may not be unique for the linear specification $t_{ij} = t \ell_{ij}$ (Berliant and Tabuchi, 2018). Minimized total commuting costs, however, are uniquely given and suffice for our analysis of the upper-stage game.

[19]First, $\nabla \mathbf{A}(\boldsymbol{m}) = \mathbf{D} = [d_{ij}]$ satisfies Assumption 1 (i). $T\mathcal{M}$ is spanned by $\{\boldsymbol{z}_j \equiv \boldsymbol{e}_i - \boldsymbol{e}_j \mid j \in \mathcal{I} \setminus \{i\}\}$ where $\{\boldsymbol{e}_j\}$ is the standard basis and $i \in \mathcal{I}$ may be arbitrarily chosen. We have $(c\boldsymbol{z}_j)^\top \mathbf{D}(c\boldsymbol{z}_j) = 2c^2(1 - d_{ij}) > 0$ for any $c \in \mathbb{R}$ because $d_{ii} = d_{jj} = 1$, $d_{ij} = d_{ji}$, and $d_{ij} \in (0, 1)$. Second, $\mathbf{D} = \mathbf{D}^\top$ since $\ell_{ij} = \ell_{ji}$, satisfying Assumption 1 (ii).

maximizes $g(\boldsymbol{m})$ for a parameter pair (τ^*, t^*), it locally (globally) maximizes $\tilde{g}(\boldsymbol{m})$ for the pair $(\tau^*, 2t^*)$. $\diamondsuit$

As we reviewed in Section 1, any local potential maximizer for problem (P) is a locally stable equilibrium under standard myopic dynamics. Also, any global potential maximizer is a globally stable equilibrium in the sense of stochastic stability, or alternatively selection under a perfect foresight dynamic. We define the sets $\mathcal{L}$ and $\mathcal{G}$ of firms' spatial patterns by

$$\mathcal{L} \equiv \{\text{The set of } local \text{ maximizers for problem (P)}\}, \tag{4.11}$$

$$\mathcal{G} \equiv \{\text{The set of } global \text{ maximizers for problem (P)}\}, \tag{4.12}$$

respectively. Obviously, $\mathcal{L}$ and $\mathcal{G}$ are parameter dependent, and $\mathcal{G} \subset \mathcal{L} \subset \mathcal{M}$. For the remainder of the paper, we study the properties of $\mathcal{L}$ and $\mathcal{G}$.

In the example sections below, we presume the following condition for simplicity.

Assumption 3 (No excess land supply). $(1 + \phi)M = 1$.

The left-hand side $(1 + \phi)M$ is the equilibrium mass of agents, and the right-hand side is the total supply of land. If $(1 + \phi)M < 1$, then extra land capacity exists in equilibrium. In certain situations, some firms (and all their employees) can freely relocate to another cell, which introduces a nonessential complication.[20]

5 The two-cell city

This section provides a concrete analysis of spatial equilibria in the model for the simplest possible two-cell setup ($I = 2$) to illustrate the basic structure of the model. In particular, this section addresses the roles of asymmetries in cells' capacities $\boldsymbol{a} = (a_1, a_2)$. Section 6 then focuses on the effects of τ and t by abstracting away all asymmetries across cells.

Let $\ell_{11} = \ell_{22} = 0$ and $\ell_{12} = \ell_{21} = 1$. Then, $d_{11} = d_{22} = 1$ and $d_{12} = d_{21} = \delta \equiv e^{-\tau}$, and $t_{11} = t_{22} = 0$ and $t_{12} = t_{21} = t$. We assume that $a_1 \geq a_2$, i.e., $a_1 \geq \frac{1}{2}$, without loss of generality.

For simplicity, we take m_1 as the variable. To satisfy $\boldsymbol{m} \in \mathcal{M}$, it must be that $m_1 \in [\check{m}, \hat{m}]$ with some $\check{m} \geq 0$ and $\hat{m} \leq M$.[21] Also, let $\bar{m} \equiv \frac{M}{2}$ and $\tilde{m} \equiv \frac{a_1}{1+\phi}$. Note that $\bar{m}, \check{m}, \tilde{m}$, and $\hat{m}$ are all independent of the transportation cost parameters t and τ.

If $m_1 = \tilde{m}$, then Assumption 3 implies that $\boldsymbol{m} = (\frac{a_1}{1+\phi}, \frac{a_2}{1+\phi})$ and that all firms and their employees are located in the same cell. Following Fujita and Ogawa (1982), we call it the *integrated pattern*. The two patterns $m_1 = \check{m}, \hat{m}$ are *segregated patterns*, since one of the cells becomes a devoted business center. For instance, $m_1 = \hat{m}$ corresponds to $\boldsymbol{m} = (M, 0)$ (if $M \leq a_1$).

To evaluate $\check{h}$, we first solve the lower-stage game. Upon inspection, the solution for the

---

[20]For an illustration of the effects of dropping Assumption 3, compare the shape of the potential function in Figure 1 with that of Figure D.4.

[21]Concretely, $\hat{m} \equiv \max_{\boldsymbol{m} \in \mathcal{M}} m_1 = \min\{M, a_1\}$ and $\check{m} \equiv \min_{\boldsymbol{m} \in \mathcal{M}} m_1 = \max\{0, M - (1 - a_1)\}$.
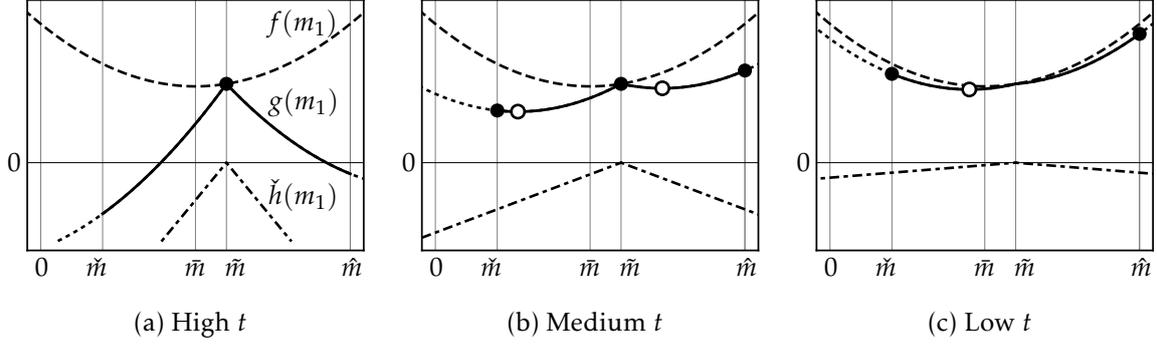
Figure 1: Graphs of $g(m_1)$ for high, medium, and low values of t. We let $t = \frac{3}{2}\hat{t}$, $t = \frac{2}{3}\tilde{t} + \frac{1}{3}\hat{t}$, and $t = \frac{1}{2}\tilde{t}$, respectively, where $\hat{t}$ and $\tilde{t}$ are to be defined in Proposition 4. $m_1 = \bar{m}$ is the minimizer of f, whereas $m_1 = \tilde{m}$ is the maximizer of $\check{h}$. ●: local maximizers; ○: locally minimizing KKT points. $M = \frac{1}{2}$, $a_1 = \frac{3}{5}$, $\delta = \frac{1}{10}$.

lower-stage equilibrium under Assumption 3 is given by:[22]

$$
\begin{cases}
n_{11} = \phi m_1, & n_{12} = (1+\phi)m_2 - a_2, & \check{r}_1 = w_1 = 0, \\
n_{22} = \phi m_2, & n_{21} = 0, & \check{r}_2 = w_2 = t,
\end{cases}
\quad \text{if } \check{m} \leq m_1 < \tilde{m}, \quad (5.1a)
$$

$$
\begin{cases}
n_{11} = \phi m_1, & n_{12} = 0, & \check{r}_1 = w_1 = 0, \\
n_{22} = \phi m_2, & n_{21} = 0, & \check{r}_2 = w_2 = 0,
\end{cases}
\quad \text{if } m_1 = \tilde{m}, \quad (5.1b)
$$

$$
\begin{cases}
n_{11} = \phi m_1, & n_{12} = 0, & \check{r}_1 = w_1 = t, \\
n_{22} = \phi m_2, & n_{21} = (1+\phi)m_1 - a_1, & \check{r}_2 = w_2 = 0,
\end{cases}
\quad \text{if } \tilde{m} < m_1 \leq \hat{m}. \quad (5.1c)
$$

If $m_1 = \tilde{m}$, then all households can reside in the same cells as their workplace, and there are no inter-cell commutes. In this case, we let $\check{r}_i = w_i = 0$ for both $i \in \mathcal{I}$. On the other hand, if $m_1 \neq \tilde{m}$, then $(1+\phi)m_i - a_i > 0$ for one of the cells, meaning that a shortage of labor exists in cell i and that $(1+\phi)m_i - a_i$ workers commute to i from the other cell. Here, $w_i = t$ is the minimal possible wage rate in cell i to compensate for commuting costs. The highest bid for the land rents in each cell i must coincide with its wage rate. Thus, we have $\check{r}_i = w_i = t$.

Employing the analytic solution (5.1), $\check{h}(\boldsymbol{m})$ is expressed in terms of m_1:

$$
\check{h}(m_1) = (1+\phi)t\,|m_1 - \tilde{m}|. \quad (5.2)
$$

Similarly, $f(\boldsymbol{m})$ is expressed as:

$$
f(m_1) = (1-\delta)(m_1 - \bar{m})^2 + (1+\delta)\bar{m}^2. \quad (5.3)
$$

where $\delta = e^{-\tau}$. As a function of m_1, $g(\boldsymbol{m})$ is computed as $g(m_1) = f(m_1) - \check{h}(m_1)$. Problem (P) is reduced to a maximization of $g(m_1)$ over the interval $[\check{m}, \hat{m}]$.

Figure 1 shows the graphs of g, f, and $\check{h}$ for three different values of t (high, medium, and low). The functions g, f, and $\check{h}$ are extended to be defined over all $\mathbb{R}$, whereas the state space is the closed interval $[\check{m}, \hat{m}]$. The potential function g is neither convex nor concave. Points in

---

[22]Appendix D explores the two-cell city in detail without Assumption 3.

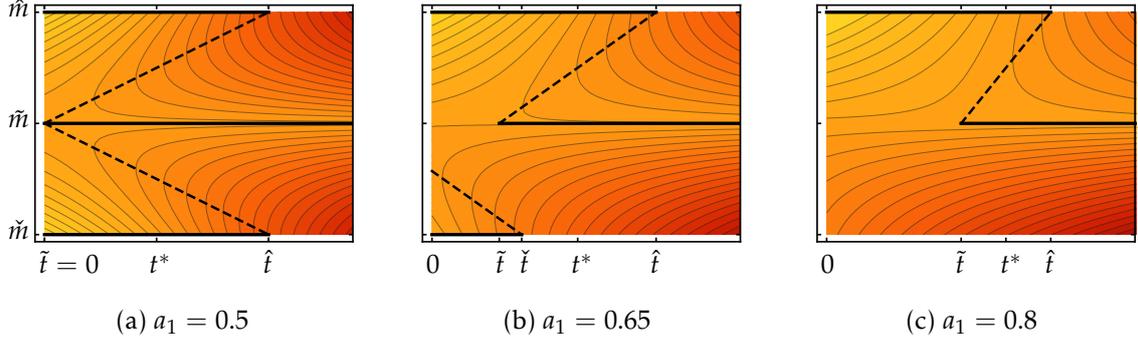(a) $a_1 = 0.5$ (b) $a_1 = 0.65$ (c) $a_1 = 0.8$

Figure 2: Local maximizer(s) of $g(m_1)$ in the t-axis. $M = 0.5$. The solid curves: local maximizers; The broken curves: locally minimizing KKT points; The background: the contours of $g(m_1)$ over (t, m_1)-space. $\mathcal{G} = \{\hat{m}\}$ if $t < t^*$ and $\mathcal{G} = \{\bar{m}\}$ if $t > t^*$.

$\mathcal{L}$, the local maximizers of g, are indicated by black marks, and locally minimizing KKT points are indicated by white circles. All these points are equilibria in the upper-stage game. For the high-commuting-cost case in Figure 1a, the integrated pattern $m_1 = \tilde{m}$ is the only local (and hence global) maximizer. For the medium-commuting-cost case in Figure 1b, the segregated patterns and the integrated pattern are the local maximizers. For the low-commuting-cost case in Figure 1c, only the segregated patterns are the local maximizers.

The following proposition characterizes $\mathcal{L}$ and $\mathcal{G}$, the local and global maximizer(s) of g.

Proposition 4. *Assume that $a_1 \geq \frac{1}{2}$. Then, only patterns in $\{\check{m}, \tilde{m}, \hat{m}\}$ ($\{\tilde{m}, \hat{m}\}$) can locally (globally) maximize $g(m_1)$ in $[\check{m}, \hat{m}]$. Let $\check{t} \equiv 2M(1-\delta)(\hat{m} - \bar{m})$, $\tilde{t} \equiv 2M(1-\delta)(\tilde{m} - \bar{m})$, $\hat{t} \equiv 2M(1-\delta)(\hat{m} - \bar{m})$, and $t^* \equiv \frac{1}{2}(\hat{t} + \tilde{t})$, with $\delta \equiv e^{-\tau}$. Then,*

$$
\begin{cases}
\check{m} \in \mathcal{L} & \forall t < \check{t}, \\
\tilde{m} \in \mathcal{L} & \forall t > \tilde{t}, \quad \text{and} \\
\hat{m} \in \mathcal{L} & \forall t < \hat{t},
\end{cases}
\qquad
\begin{cases}
\tilde{m} \in \mathcal{G} & \forall t \geq t^*, \\
\hat{m} \in \mathcal{G} & \forall t \leq t^*.
\end{cases}
\tag{5.4}
$$

Proposition 4 states that the integrated pattern ($m_1 = \tilde{m}$) is the only possible local maximizer for larger values of t, whereas the segregated patterns ($m_1 = \check{m}$ or $\hat{m}$) become local maximizer(s) for smaller values of t. The integrated pattern should be the only equilibrium if the commuting cost per distance t is sufficiently high because the costs of commuting overcome the benefit of agglomeration. When t is not too high, concentration of firms can be a local maximizer when τ is large. For a firm, proximity to other firms matters more than higher wage paid to its employees when τ is small and t is not too high. It is intuitive that $m_1 = \check{m}$, a concentration to a less advantageous cell 2, cannot be a global potential maximizer.

Figure 2 shows local maximizers in terms of m_1 in line with varying t. Figure 2a is the symmetric case ($a_1 = \frac{1}{2}$), for which we have $\tilde{m} = \bar{m} = \frac{M}{2}$ and $\check{t} = \hat{t}$ and $\tilde{t} = 0$. Figure 2c shows a case where cell 1 is so advantageous that a concentration to cell 2 ($m_1 = \check{m}$) can *never* be an equilibrium. Figure 2b is an intermediate case. Note that, while concentration to either cell 1 ($m_1 = \hat{m}$) or cell 2 ($m_1 = \check{m}$) can locally maximize $g(m_1)$, $\hat{m}$ is more likely to be a local maximizer (i.e., $\hat{t} > \check{t}$). Figure 1 shows snapshots of Figure 2b for three different values of t.
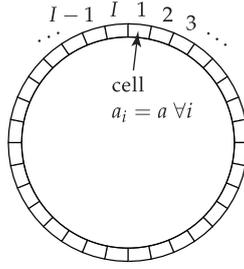
Figure 3: Circular city.

6 Circular city

This section considers a symmetric geography to elucidate the intrinsic role of transportation cost parameters t and τ. To abstract away the effects of asymmetries across cells, we assume that the underlying geography is a circle with circumferential length 1, on which cells of the same size are sequentially placed (Figure 3).

Assumption 4. $\ell_{ij} = \frac{1}{I}\min\{|i-j|, I-|i-j|\}$ *for all* $i, j \in \mathcal{I}$. *Also,* $a_i = a \equiv \frac{1}{I}$ *for all* $i \in \mathcal{I}$.

Assumption 4 abstracts away cell-specific advantages arising from the shape of the underlying transportation network or extra capacities, which allows us to isolate the endogenous forces that determine spatial patterns. For instance, in a line segment, the locations near city boundaries have fewer opportunities to access other cells. By contrast, in a circle, every cell has the same level of accessibility as the other cells. In particular, the uniform distribution of firms $\bar{\boldsymbol{m}} = (\bar{m}, \bar{m}, \ldots, \bar{m})$ $(\bar{m} \equiv \frac{M}{I})$ is always an equilibrium.

For simplicity, following Fujita and Thisse (2013), Section 6.5, we assume that $M = \frac{1}{2}$. In fact, the total mass of firms M does not affect the qualitative properties of equilibrium spatial *patterns* (i.e., integrated or segregated) under Assumption 4.[23]

Assumption 5. $M = \frac{1}{2}$.

For the remainder of the paper, we assume Assumptions 2, 3, 4, and 5. In addition, to simplify the analysis, *we consider global maximization of the potential function as the equilibrium refinement criterion* There is a pair of useful predictions on the properties $\mathcal{G}$.

Lemma 1. *Consider problem (P) under Assumptions 2, 3, 4, and 5. Then, the following hold true.*

(a) Any $\boldsymbol{m} \in \mathcal{L}$, *and thus any* $\boldsymbol{m} \in \mathcal{G}$, *should satisfy* $m_i \in \{0, \bar{m}, a\} = \{0, \bar{m}, 2\bar{m}\}$ *for all* $i \in \mathcal{I}$.

(b) For any $\tau > 0$, $\mathcal{G} = \{\bar{\boldsymbol{m}}\}$ *when* t *is sufficiently large.*

Condition (a) substantially simplifies the enumeration of relevant spatial configurations. It indicates that if $\boldsymbol{m} \in \mathcal{M}$ locally maximizes g, every cell has to be used either for purely residential ($m_i = 0$), integrated ($m_i = \bar{m}$), or purely business ($m_i = a = 2\bar{m}$), purposes. For (b), it would be natural that if t is very large, concentration of firms is inferior in potential value to the uniform distribution $\bar{\boldsymbol{m}}$ because, while $f(\boldsymbol{m})$ is bounded under Assumption 2, $\check{h}(\boldsymbol{m})$ can be arbitrarily large in t.

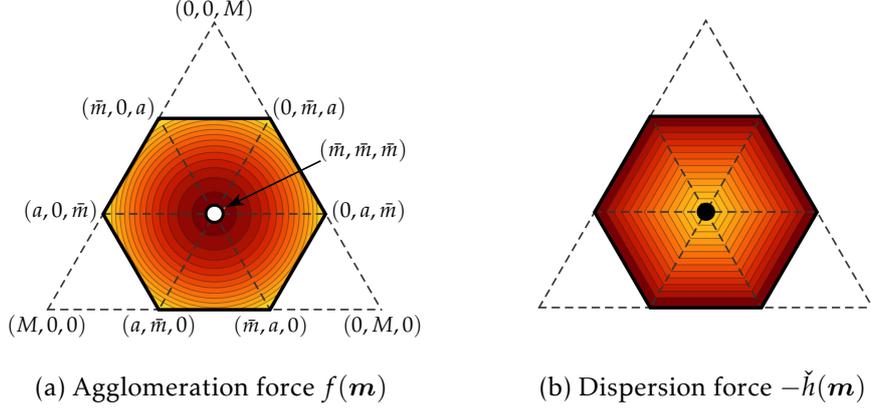[23] M acts as a scaling factor for t. See Remark D.2 in Appendix D for the $I = 2$ example.

(a) Agglomeration force $f(\boldsymbol{m})$ (b) Dispersion force $-\check{h}(\boldsymbol{m})$

Figure 4: Contours of $f(\boldsymbol{m})$ and $-\check{h}(\boldsymbol{m})$ for the three-cell city.

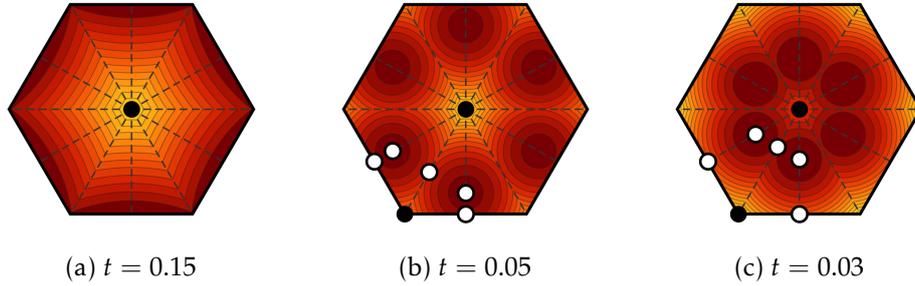(a) $t = 0.15$ (b) $t = 0.05$ (c) $t = 0.03$

Figure 5: Contours of $g(\boldsymbol{m})$ for the three-cell city ($e^{-\frac{\tau}{3}} = 0.5, M = \frac{1}{2}$). The coordinate system is the same as Figure 4. ●: local maximizer, ○: local minimizer or saddle point.

6.1 The three-cell city

Assuming $I = 3$, this section explores the landscape of the potential function $g(\boldsymbol{m})$. Because the basic results are the same with $I = 2$, we focus on graphical intuitions. Without loss of generality, we let $m_1 \geq m_2 \geq m_3$. Lemma 1 implies that only the uniform distribution $\bar{\boldsymbol{m}} = (\bar{m}, \bar{m}, \bar{m})$ and the monocentric pattern $\hat{\boldsymbol{m}} \equiv (a, \bar{m}, 0)$ can be global maximizers of the potential function.[24]

Figure 4 shows the contours of $f(\boldsymbol{m})$ and $-\check{h}(\boldsymbol{m})$ on $\mathcal{M}$ employing a simplex coordinate. The first term $f(\boldsymbol{m})$ is a parabola centered at $\bar{m}$. As discussed in Section 4.4, it is larger near the boundary of $\mathcal{M}$ because it corresponds to the agglomeration force in the model. The second term $-\check{h}(\boldsymbol{m})$ is a piecewise affine function attaining the maximum at $\bar{m}$.[25] It is larger when $\boldsymbol{m}$ is close to the integrated pattern $\bar{m}$, as total commuting costs is smaller. Being the sum of these two terms, working in opposite directions, the potential function g is nonconvex.

Figure 5 shows the contours of $g(\boldsymbol{m})$ for different values of t. The black markers indicate local maximizers, whereas the white markers indicate other KKT points. For simplicity, only those KKT points (i.e., equilibria) with $m_1 \geq m_2 \geq m_3$ are drawn. Qualitative properties are

---

[24] Note that all the patterns $(a, \bar{m}, 0)$, $(0, a, \bar{m})$, $(\bar{m}, 0, a)$, $(a, 0, \bar{m})$, $(0, \bar{m}, a)$, and $(\bar{m}, a, 0)$ are essentially the "same" spatial pattern and share the same potential value. There are five other local (global) maximizers of the same meaning when $\hat{\boldsymbol{m}} = (a, \bar{m}, 0)$ is a local (global) maximizer of the potential function.

[25] If we drop Assumption 4 so that a_i can vary across cells, $-\check{h}(\boldsymbol{m})$ attains maximum at $\tilde{\boldsymbol{m}} = \frac{1}{1+\phi}(a_1, a_2, a_3)$, analogous to the two-cell case with $a_1 > a_2$.
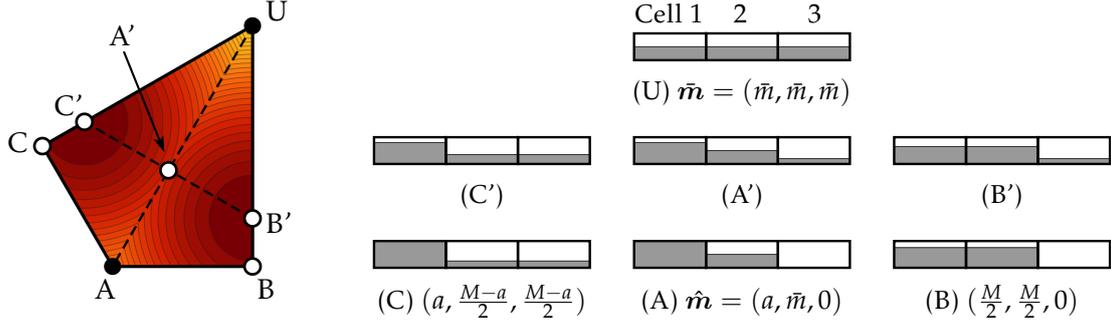
Figure 6: Equilibrium spatial distributions of firms ($I = 3$). The left-hand side: a subset of $\mathcal{M}$ that satisfy $m_1 \geq m_2 \geq m_3$. The right-hand side: spatial patterns, with letters corresponding to those in the contour plot on left-hand side.

consistent with the two-cell example. When t is large, $\bar{m}$ is the only maximizer. For lower values of t, there are numerous KKT points, all of which are equilibria in the model. Only $\bar{m}$ and $\hat{m}$ can be local maximizers, confirming Lemma 1. The figure also illustrates the landscape of the potential function behind the lemma, in that any interior equilibrium with $m_i \neq \bar{m}$ for some i is either a local minimizer or a saddle point.

Figure 6 schematically illustrates equilibrium patterns associated with the markers in Figure 5b, as well as a magnification of Figure 5b for the subset of $\mathcal{M}$ such that $m_1 \geq m_2 \geq m_3$. Only the uniform distribution $\bar{m}$ (pattern U) and the monocentric agglomeration $\hat{m}$ (pattern A) locally maximize the potential function. Patterns C and B are local minimizers on the boundaries. Patterns A', B', and C' are interior equilibria which may be interpreted as "in-between" patterns.

The (possible) number of equilibrium patterns increases considerably (from 5 to 25 without symmetry considerations), merely by moving from $I = 2$ to $I = 3$. The majority of equilibria (18 out of 25) are less important (i.e., local minimizers or saddles). Thus, imposing some equilibrium refinement is crucial for obtaining meaningful equilibria.

By focusing on patterns that can globally maximize $g(m)$, i.e., $\bar{m} = (\bar{m}, \bar{m}, \bar{m})$ and $\hat{m} = (a, \bar{m}, 0)$, we obtain the following characterization. In words, dispersion $\bar{m}$ is selected when commuting cost t is large, and agglomeration $\hat{m}$ is selected when t is small.

Proposition 5. *Suppose that $I = 3$. Let $t^* = \frac{1}{4}(1 - \delta)$ with $\delta \equiv e^{-\frac{\tau}{3}}$. Then, the uniform distribution $\bar{m}$ is a global potential maximizer for $t \geq t^*$ and the monocentric pattern $\hat{m}$ is a global potential maximizer for $t \leq t^*$.*

6.2 The four- and eight-cell cities

Fujita and Ogawa (1982)'s model describes the formation of multiple urban centers as an equilibrium outcome. The four-cell city ($I = 4$) is the minimal setup where the duo-centric pattern $m^2 \equiv (a, 0, a, 0)$ can emerge. If we focus on the set of global maximizers $\mathcal{G}$, however, only the complete dispersion $\bar{m} = (\bar{m}, \bar{m}, \bar{m}, \bar{m})$ and the monocentric pattern $m^1 \equiv (0, a, a, 0)$ can emerge.

18

Proposition 6. *Suppose that $I = 4$. Let $t^* \equiv \frac{1}{4}(1 - \delta^2)$ with $\delta \equiv \mathrm{e}^{-\frac{\tau}{4}}$. Then, only the uniform and monocentric patterns ($\bar{\boldsymbol{m}}$ and $\boldsymbol{m}^1$) can be global potential maximizers. The uniform distribution $\bar{\boldsymbol{m}}$ is a global potential maximizer for $t \geq t^*$ and the monocentric pattern $\boldsymbol{m}^1$ is a global potential maximizer for $t \leq t^*$.*

The duo-centric pattern cannot be a global potential maximizer because $g(\boldsymbol{m}^1) > g(\boldsymbol{m}^2)$ for all (τ, t). For households, there is no distinction between monocentric and duo-centric equilibria because in both cases every household commutes to a cell next to its residential cell, and the equilibrium commuting cost for a household is $\frac{t}{4}$, implying that $\check{h}(\boldsymbol{m}^1) = \check{h}(\boldsymbol{m}^2) = 2 \times a \times \frac{t}{4} = t\bar{m}$. On the other hand, since a monocentric concentration $\boldsymbol{m}^1$ offers a higher level of technological externalities for firms than the duo-centric pattern $\boldsymbol{m}^2$, we have $f(\boldsymbol{m}^1) > f(\boldsymbol{m}^2)$.

For $I = 8$, on the other hand, commuting length may vary for different spatial configurations. Possible configurations include the uniform pattern $\bar{\boldsymbol{m}} = (\bar{m}, \bar{m}, \ldots, \bar{m})$, the monocentric pattern $\boldsymbol{m}^1 \equiv (0, 0, a, a, a, a, 0, 0)$, the duo-centric pattern $\boldsymbol{m}^2 \equiv (0, a, a, 0, 0, a, a, 0)$, and the quad-centric pattern $\boldsymbol{m}^4 \equiv (a, 0, a, 0, a, 0, a, 0)$. Figure 7a illustrates $\bar{\boldsymbol{m}}$, $\boldsymbol{m}^1$, $\boldsymbol{m}^2$, and $\boldsymbol{m}^4$, with arrows indicating households' commuting patterns. As the commuting cost between two neighboring cells is $\frac{t}{I} = \frac{t}{8}$, a household's commuting cost in patterns $\bar{\boldsymbol{m}}$, $\boldsymbol{m}^1$, $\boldsymbol{m}^2$, and $\boldsymbol{m}^4$ are given by 0, $\frac{t}{8}$, $\frac{t}{8}$, and $\frac{t}{4}$, respectively.

The following proposition characterizes $\mathcal{G}$ for $I = 8$.[26]

Proposition 7. *Suppose $I = 8$. Then, only uniform, monocentric, and duo-centric patterns ($\bar{\boldsymbol{m}}$, $\boldsymbol{m}^1$, and $\boldsymbol{m}^2$) can be global potential maximizer(s). There exists τ^* such that:*

(a) If $\tau \leq \tau^*$, then there exists t^* such that (i) the uniform pattern $\bar{\boldsymbol{m}}$ is a global maximizer for $t \geq t^*$ and (ii) the monocentric pattern $\boldsymbol{m}^1$ is a global maximizer for $t \leq t^*$.

(b) If $\tau > \tau^*$, then there exist t^{**} and t^{***} with $t^{**} > t^{***}$ such that (i) the uniform pattern $\bar{\boldsymbol{m}}$ is a global maximizer for $t \geq t^{**}$, (ii) the duo-centric pattern $\boldsymbol{m}^2$ is a global maximizer for $t^{***} \leq t \leq t^{**}$, and (iii) the monocentric pattern $\boldsymbol{m}^1$ is a global maximizer for $t \leq t^{***}$.

It is far simpler to visualize Proposition 7 in the (τ, t)-space. Figure 7b shows a partitioning of the (τ, t)-space on the basis of $\mathcal{G}$. The letters U, M, and D in each region correspond, respectively, to the uniform, monocentric, and duo-centric patterns ($\bar{\boldsymbol{m}}$, $\boldsymbol{m}^1$, $\boldsymbol{m}^2$). When t is large, agglomeration cannot occur, and the uniform distribution is the global maximizer. Analogous to the two- and three-cell cases, a concentration of firms is possible if t is sufficiently small, where the benefit of proximity overcomes firms' costs. When the decay rate τ is small, the monocentric pattern is chosen. When τ is large, a new phenomenon emerges: duo-centric pattern becomes the global maximizer of the potential. When spillover from one cell to its neighbor is low (τ is high) and commuting cost t is intermediate, the benefit of agglomeration cannot compensate for longer households commutes.

A rough intuitive rationale behind the emergence of the duo-centric pattern when τ is large is as follows. Suppose that τ is large or that $\delta = \mathrm{e}^{-\frac{\tau}{8}}$ is small, so that either δ^2, δ^3, or δ^4 is negligibly small. Then, consider the third cell from the left in Figure 7a. In the monocentric

---

[26]The quad-centric pattern $\boldsymbol{m}^4$ cannot be in $\mathcal{G}$ because $g(\boldsymbol{m}^2) > g(\boldsymbol{m}^4)$ for all (τ, t), analogous to the $I = 4$ case.
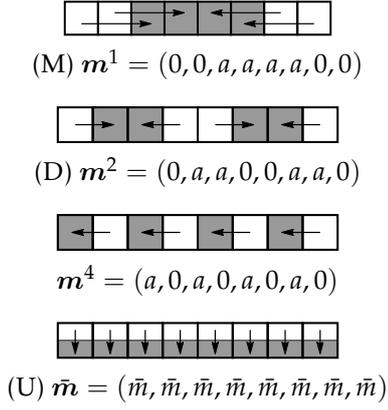
(a) Representative spatial patterns.

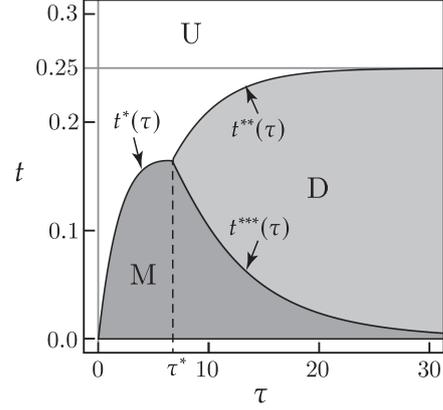(b) Partition based on $\mathcal{G}$.

Figure 7: Partition based on $\mathcal{G}$ and associated spatial patterns ($I = 8$). The letter on each region in Figure 7b corresponds to that in Figure 7a.

pattern $\boldsymbol{m}^1$, the cell is located at a boundary of the business center. The production level in the cell is given by:

$$\left(1 + \delta + \delta^2 + \delta^3\right) a \approx (1 + \delta)a,$$

whereas firms' operating costs in the cell are $(1 + \phi) \times \frac{2t}{8} = \frac{t}{2}$. On the other hand, every firm in the duo-centric pattern $\boldsymbol{m}^2$, including the third cell, enjoys the following production level:

$$\left(1 + \delta + \delta^3 + \delta^4\right) a \approx (1 + \delta)a,$$

whereas firms' operating costs are $(1 + \phi) \times \frac{t}{8} = \frac{t}{4}$. Thus, firms in the two boundary cells of the single business center in $\boldsymbol{m}^1$ have an incentive to deviate toward the duo-centric pattern $\boldsymbol{m}^2$ when δ is small (τ is large). The positive spillover is too low to compensate for higher operating costs when τ is large.

6.3 A city with a large number of cells

The above analyses for small Is have demonstrated that multiple-center equilibria can be local and global potential maximizers in the model. This exploratory section provides an additional numerical investigation into the case when the number of cells I is extremely large.

For simplicity, we restrict our attention to a collection of symmetric spatial distributions of firms: the uniform pattern (or the *integrated pattern*) $\bar{\boldsymbol{m}} = (\bar{m}, \bar{m}, \ldots, \bar{m})$ and *symmetric J-centric segregated patterns* $\boldsymbol{m}^J$ ($J = 1, 2, \ldots$). The latter are spatial patterns of firms such that

$$\boldsymbol{m}^J = \big(\underbrace{\check{\boldsymbol{m}}_J, \check{\boldsymbol{m}}_J \ldots, \check{\boldsymbol{m}}_J}_{J \text{ times}} \big) \quad \text{where} \quad \check{\boldsymbol{m}}_J \equiv \big(\underbrace{0, 0, \ldots, 0}_{\frac{I}{4J} \text{ times}}, \underbrace{a, a, \ldots, a}_{\frac{I}{2J} \text{ times}}, \underbrace{0, 0, \ldots, 0}_{\frac{I}{4J} \text{ times}} \big) \qquad (6.1)$$

where each $\check{\boldsymbol{m}}_J$ is interpreted as an urban subcenter with a self-contained job market.[27] For

---

[27] There are $\frac{I}{2J}$ purely business cells and $\frac{I}{2J}$ purely residential cells in each $\check{\boldsymbol{m}}_J$.

20

(a) Uniform $\bar{m}$ (b) Monocenter m^1 (c) Duo-center m^2 (d) Tri-center m^3 (e) Quad-center m^4

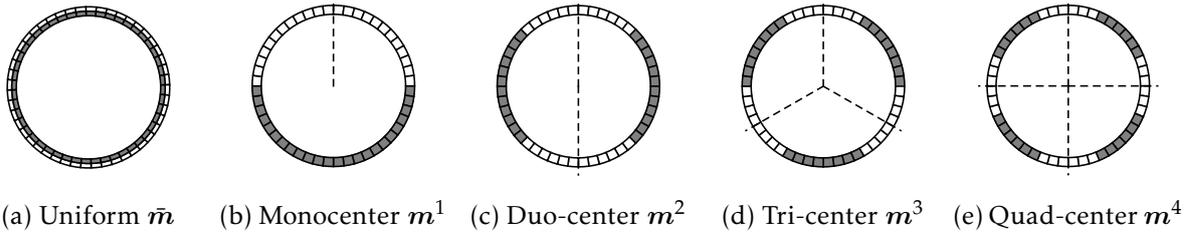**Figure 8:** Symmetric patterns in $\mathcal{M}^*$ ($I = 4\mathrm{LCM}(4) = 48$).

all m^J to exist up to K-center pattern ($K \geq 1$), $\frac{I}{4J}$ must be an integer for all $J \leq K$. That is, $I = 4\mathrm{LCM}(K)$ where $\mathrm{LCM}(K)$ denotes the least common multiple of $\{1, 2, 3, \ldots, K\}$.[28]

Let $\mathcal{M}^* \equiv \{\bar{m}, m^1, m^2, \ldots, m^K\}$ denote the set of symmetric spatial distributions. Figure 8 illustrates the patterns in $\mathcal{M}^*$ when $I = \mathrm{LCM}(4) = 48$. We consider global maximization of the potential function $g(m)$ over $\mathcal{M}^*$. The conjecture behind this analysis is formalized as follows.

Conjecture 1. *Suppose $I = 4\mathrm{LCM}(K)$ ($K \geq 1$). Let $\mathcal{G}^*$ be the set of global potential maximizers among $\mathcal{M}^*$. Then, $\mathcal{G} = \mathcal{G}^*$.*

The conjecture is, if $m \in \mathcal{M}^*$ has the maximum potential value among the patterns in $\mathcal{M}^*$, then it globally maximizes g over $\mathcal{M}$, and vice versa. Propositions 6 and 7 support the conjecture for $K = 1$ and $K = 2$ (i.e., $I = 4$ and $I = 8$), respectively, since $\mathcal{M}^* = \{\bar{m}, m^1\}$ for $K = 1$ and $\mathcal{M}^* = \{\bar{m}, m^1, m^2\}$ for $K = 2$. Also, a possible local maximizer other than the symmetric patterns $\mathcal{M}^*$ can be regarded as a transitional pattern connecting two symmetric patterns.[29] Thus, it is expected that $\mathcal{G}^*$ provides at least an approximate view of the overall properties of $\mathcal{G}$.

Figure 9 shows a partition of the (τ, t)-space on the basis of $\mathcal{G}^*$, i.e., the global potential maximizer among $\mathcal{M}^*$, when $K = 50$.[30] Segregated patterns $\{m^J\}_{J=1}^K$ dominate the integrated pattern $\bar{m}$ in the gray regions, whereas the uniform distribution is the global potential maximizer in the white region. Each gray region corresponds to one of the J-centric patterns, and the number on each region corresponds to the number of business centers J. On the τ-axis, gray regions are aligned in the increasing order of J.

Figure 9 has two basic implications. One is the effect of the commuting cost parameter t, and the other is the effect of the distance decay parameter τ.

Observation 3. *Fix $\tau > 0$, suppose that t is very large initially, and consider a monotonic decrease in t. Then, the global maximizer of g across $\mathcal{M}^*$ is given by the following sequence: $\bar{m} \to m^{J^*} \to m^{J^*-1} \to \cdots \to m^2 \to m^1$ for some $J^* \geq 1$.* ◇

Multiple business centers are formed from uniformity and then the number of business centers decreases as t decreases from a larger extreme. For instance, let $\tau = 25$ and consider a

---

[28]For instance, $I = 4, 8, 12, 48, 240, 240, 1,680$, and $3,360$ for each $1 \leq K \leq 8$.

[29]For example, $(0, a, \bar{m}, \bar{m})$ can be a local maximizer when $I = 4$. It connects two patterns $(0, a, 0, a)$ and $(0, a, a, 0)$. See Proof of Proposition 6 in Appendix A.

[30]In favor of efficiency of numerical computation, we employ a continuous version of the potential function g to draw Figure 9. See Appendix E for details.
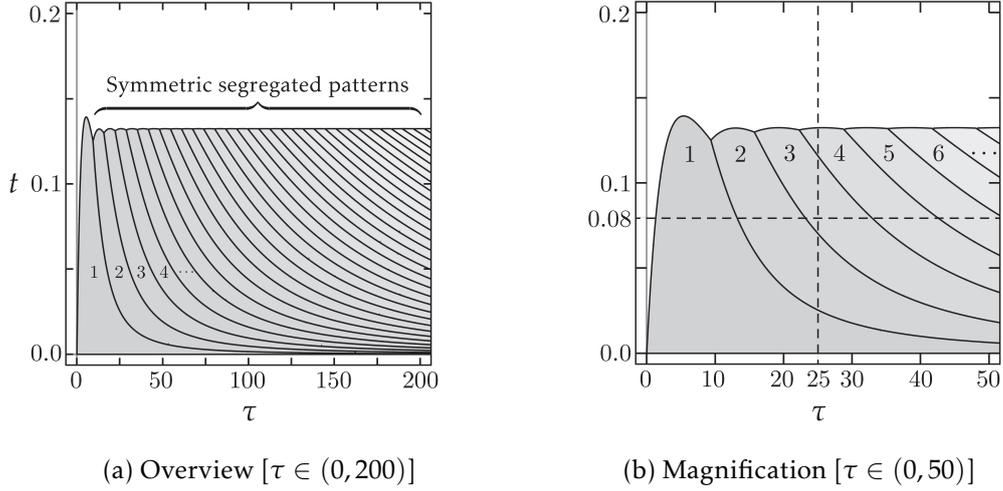
(a) Overview [$\tau \in (0, 200)$]

(b) Magnification [$\tau \in (0, 50)$]

Figure 9: Partition of the (τ, t)-space based on $\mathcal{G}^*$. The white region indicates that $\mathcal{G}^* = \{\bar{\boldsymbol{m}}\}$. The gray regions indicate that one of the symmetric segregated patterns $\{\boldsymbol{m}^J\}$ globally maximizes g among $\mathcal{M}^*$; the number in each region indicates the number of subcenters J.

decrease of t (the vertical broken line in Figure 9b). When t is sufficiently large, $\bar{\boldsymbol{m}}$ maximizes g (Lemma 1 (b)). But when t decreases and cuts a threshold, $\boldsymbol{m}^4$ emerges as the global potential maximizer in $\mathcal{M}^*$. As t decreases further, the potential maximizer sequentially switches, as $\boldsymbol{m}^4 \to \boldsymbol{m}^3 \to \boldsymbol{m}^2 \to \boldsymbol{m}^1$.[31]

Observation 4. *Fix a sufficiently small $t > 0$, suppose that τ is very small initially, and consider a monotonic increase in τ. Then, the global maximizer of g across $\mathcal{M}^*$ is given by the following sequence: $\bar{\boldsymbol{m}} \to \boldsymbol{m}^1 \to \boldsymbol{m}^2 \to \boldsymbol{m}^3 \to \cdots$.* ◇

A monotonic increase in the number of subcenters occurs when τ monotonically increases with a given t. For instance, let $t = 0.08$ and consider a monotonic increase in τ (the horizontal broken line in Figure 9b). For small τ, the uniform distribution $\bar{\boldsymbol{m}}$ has the maximum potential value on this line. A steady increase in τ induces a repetitive emergence of segregated patterns, with an increasing number of subcenters J as the global potential maximizer in $\mathcal{M}^*$.

If Conjecture 1 is true, Observations 3 and 4 generalize the implications of Proposition 7: when t is large, segregation cannot become a global maximizer; when t is not too large, then the number of business centers increases as decay rate τ rises.

7   Concluding remarks

Building on the seminal framework by Fujita and Ogawa (1982), this paper considered the emergence of multiple business centers in the cities. A potential game approach is shown to be efficacious for the analysis of the model. Through local and global maximizations of the potential function, we characterized stable spatial equilibria in the model. Our results provide an answer to the long-standing stability issue of polycentric equilibria of the FO model. We

---

[31] Figure E.7a and Figure E.7b in Appendix E show graphs of potential values for the symmetric patterns over the vertical and horizontal broken lines in Figure 9b ($\tau = 25$ and $t = 0.08$), respectively.

also demonstrated that the model can have *multiple local maximizers* of the potential function. The results presented in Sections 6.2 and 6.3 suggest that equilibrium refinement based on global maximization of the potential function makes the discussion cleaner, without affecting the basic implications of the model.

When t is large, the uniform distribution becomes the only global potential maximizer. When commuting costs are too large, firms cannot compensate its employees for their commuting costs with wage, because their production level is finite. We note that households' commuting costs has two interpretations. On the one hand, as emphasized in Sections 4, 5, and 6, they acts as *dispersion forces from a firm's perspective*. On the other hand, they can also be interpreted as an *inter-type agglomeration force* (or *co-location force*) that glues firms and households together (cf., Papageorgiou and Thisse, 1985). When t is large, the co-location force become strong, and the integration (the uniform pattern) emerges as the global potential maximizer.

When t is relatively small, segregation of firms and households can occur to form disjointed business districts (provided that the number of cells I are sufficiently large to express such patterns). Sections 6.2 and 6.3 illustrated that the number of business centers increases as τ increases. As τ grows large, the contributions of positive spillovers from firms in distant cells become negligible. At some point, firms can no longer pay the higher land rents and wages required to support the emergence of large business clusters.

There are several topics that worth further exploration. First, relaxing the simplifying assumptions in the basic framework (Section 3) would be important to provide policy implications. In particular, following Fujita and Ogawa (1982), we assume a fixed demand for land, which should be relaxed. To this end, the potential game approach is effective for any extensions that preserve the symmetry of externalities.[32] Second, for a given specification of $\mathbf{A}(\boldsymbol{m})$ or $\boldsymbol{t}$ (e.g., Assumption 2), the optimization representation (P) can be employed to study various assumptions about the transportation network of a city, i.e., the geographical distances $\{\ell_{ij}\}$ between cells, including a line segment as in the original FO paper, or a variety of stylized but interesting geographies, as in the paper by Matsuyama (2017).

---

[32] The Negishi's theorem (Negishi, 1960) implies that any competitive assumptions may be reduced to maximizing a specific kind of welfare function. The optimal value of a welfare function, in turn, acts as a potential function of the associated game when the mass of agents in each cell is taken as a variable.

23

A Proofs

Proof of Proposition 1. Let $\langle p, x\rangle \equiv p^\top x = \sum_{i\in\mathcal{I}} p_i x_i$. With multipliers $w = (w_i) \in \mathbb{R}_+^I$, $r = (r_i) \in \mathbb{R}_+^I$, v^*, and π^*, define the Lagrangian $\mathscr{L}$ for problem (P0) by

$$\mathscr{L} = -f(m) + h(n) + \left\langle r, -a + m + \left(\mathbf{I} \otimes \mathbf{1}_I^\top\right)n\right\rangle + \left\langle w, \phi m - \left(\mathbf{1}_I^\top \otimes \mathbf{I}\right)n\right\rangle$$
$$+ \left\langle v^*, \mathbf{1}_{I^2}^\top n - N\right\rangle + \left\langle \pi^*, \mathbf{1}_I^\top m - M\right\rangle,$$

where $\mathbf{I}$ is the I-dimensional identity matrix and $\mathbf{1}_K$ denotes the K-dimensional all-one vector. The first-order optimality condition with respect to the Lagrangian $\mathscr{L}$ coincides with the equilibrium conditions listed in Definition 1. $\qquad\square$

Proof for Proposition 2. Define the Lagrangian $\mathscr{L}$ for problem (S) by

$$\mathscr{L} = h(n) + \left\langle \check{r}, -a + m + \left(\mathbf{I} \otimes \mathbf{1}_I^\top\right)n\right\rangle + \left\langle w, \phi m - \left(\mathbf{1}_I^\top \otimes \mathbf{I}\right)n\right\rangle.$$

Then, the first-order optimality condition with respect to $\mathscr{L}$ coincides with the conditions in Definition 2. (D) is the Lagrangian dual problem for (S) and obtained by taking $\inf_{n\geq 0}\mathscr{L}$. $\quad\square$

Proof for Proposition 3. Define the Lagrangian $\mathscr{L}$ for problem (P) by

$$\mathscr{L} = -f(m) + \check{h}(m) + \langle \hat{r}, -a + m\rangle + \left\langle \pi^*, \mathbf{1}^\top m - M\right\rangle.$$

Then, the first-order optimality condition with respect to $\mathscr{L}$ is equivalent to the conditions in Definition 3 as well as the equilibrium conditions of the model (Definition 1). $\qquad\square$

Proof of Lemma 1. The lemma is a corollary of the following Lemma A.1.

Lemma A.1. *Suppose Assumptions 2, 3 and 4. Then, the following properties holds true for (P).*

(a) *The uniform distribution $\bar{m}$ is a local maximizer for all (τ, t).*

(b) *$\bar{m}$ is the unique global maximizer when t is sufficiently large.*

(c) *Any interior KKT point other than $\bar{m}$ (if it exists) is a local minimizer or a saddle point.*

(d) *No local maximizer involves more than one "unbalanced" cells, i.e., cells with $m_i > 0$, $\sum_{j\in\mathcal{I}} n_{ij} > 0$, and $\phi m_i \neq \sum_{j\in\mathcal{I}} n_{ij}$ so that there must be commuting from or toward the cell.*

(e) *If in addition Assumption 5 holds true, no local maximizer involves unbalanced mixed use cells.*

Proof. (a) $\bar{m}$ is a local minimizer of $\check{h}$ with $\check{h}(\bar{m}) = 0$. Evidently, we have $\check{h}(\bar{m} + \epsilon) > 0$ for any nonzero $\epsilon \in T\mathcal{M}$ under Assumption 3. Since $\check{h}(m)$ is a piecewise affine function, $\bar{m}$ is an isolated local minimizer of $\check{h}$ for any $t > 0$. Similarly, for $\epsilon \in T\mathcal{M}$, define $\tilde{f}(\epsilon) \equiv f(\bar{m} + \epsilon) = f(\epsilon) + f(\bar{m})$ (see (D.8) of Proof of Proposition D.1 in Appendix D). We have $\nabla\tilde{f}(\mathbf{0}) = \mathbf{0}$. It follows that $\bar{m}$ is a local maximizer of $g(m) = f(m) - \check{h}(m)$ for any finite $t > 0$.

(b) For anly spatial distribution $m \in \mathcal{M}$ we compute $\Delta g(m) \equiv g(m) - g(\bar{m}) = f(m - \bar{m}) - \check{h}(m)$ by (D.8) and $\check{h}(\bar{m}) = 0$. The first term is bounded above by $\frac{1}{2}\bar{m}^2\mathbf{1}^\top\mathbf{D}\mathbf{1}$, which is

finite for any $\tau > 0$. There exists t^* such that $\Delta g(m) < 0$ for any m whenever $t > t^*$ because $\check{h}(m)$ is a nonzero linear function of t when $m \neq \bar{m}$.

(c) Suppose that $m \neq \bar{m}$ is an interior equilibrium. Then, by definition there exists π and $c = (c_i)$ such that $c = Dm - \pi \mathbf{1} \geq 0$ and $g(m) = f(m) - c^\top(m - \bar{m})$, where $c > 0$ iff $m = \bar{m}$. Suppose that $m \neq \bar{m}$ and take $z \equiv \epsilon(m - \bar{m}) \in T\mathcal{M} \setminus \{0\}$ with some $\epsilon > 0$. Then,

$$
\begin{aligned}
g(m + z) &= \tfrac{1}{2}(m + z)^\top D(m + z) - c^\top(m - \bar{m} + z) \\
&= \tfrac{1}{2}m^\top Dm - c^\top(m - \bar{m}) + z^\top D(m + z) - c^\top z \\
&= g(m) + z^\top D(m + z) - z^\top(Dm - \pi \mathbf{1}) = g(m) + z^\top Dz > g(m),
\end{aligned}
$$

where we note that $z^\top \mathbf{1} = 0$ and thus Assumption 1 (ii) implies $z^\top Dz > 0$ because $z \neq 0$. It shows that any interior equilibrium m other than $\bar{m}$ is a local minimizer of g on the subspace spanned by $z = m - \bar{m}$. Therefore, m is either a local minimizer or a saddle point of g.

(d) Suppose that m is a boundary equilibrium with at least two *unbalanced* mixed-use cells j and k. Because m is an equilibrium, there exist $\{c_i\}$ and π such that $c_i = \sum_j d_{ij}m_j - \pi$ for all i with $m_i > 0$. Then, following the same procedure as in (a) with $z \equiv \epsilon(e_j - e_k) \in T\mathcal{M}$ with e_i being the ith standard basis, one shows that m is a local minimizer along the extreme line spanned by z, thereby showing the claim.

(e) Under Assumption 5, if a cell is unbalanced, then there must be another unbalanced cell. By (d), such KKT point (if exits) cannot be a local maximizer. □

Proof of Proposition 4. Noting that $g''(m_1) = 2(1 - \delta) > 0$, $g(m_1)$ is piecewise strictly convex. Since $\check{h}(m_1)$ has only one kink (at $m_1 = \tilde{m}$), there is at most three local maximizers, i.e., $\check{m}$, $\tilde{m}$, and $\hat{m}$. The left and right derivatives of $g(m_1)$ at $m_1 = \tilde{m}$ are, respectively, given by

$$
g'_-(\tilde{m}) = 2(1 - \delta)(\tilde{m} - \bar{m}) + (1 + \phi)t \quad \text{and} \quad g'_+(\tilde{m}) = 2(1 - \delta)(\tilde{m} - \bar{m}) - (1 + \phi)t. \quad \text{(A.1)}
$$

For $m_1 = \tilde{m}$ to be a local maximizer, we must have $g'_-(\tilde{m}) > 0$ and $g'_+(\tilde{m}) < 0$, which reduces to $t > \tilde{t} \equiv 2M(1 - \delta)(\tilde{m} - \bar{m})$ as we note $\frac{1}{1+\phi} = M$ under Assumption 3. At $m_1 = \hat{m}$, we compute

$$
g'(\hat{m}) = 2(1 - \delta)\min\{\bar{m}, a_1 - \bar{m}\} - (1 + \phi)t, \quad \text{(A.2)}
$$

so that $g'(\hat{m}) > 0$ implies that $t < 2M(1 - \delta)\min\{a_1 - \bar{m}, \bar{m}\}$. At $m_1 = \check{m}$, we compute

$$
g'(\check{m}) = 2(1 - \delta)\max\{-\bar{m}, \bar{m} - (1 - a_1)\} + (1 + \phi)t \quad \text{(A.3)}
$$

so that $g'(\check{m}) < 0$ implies $t < 2M(1 - \delta)\min\{\bar{m}, (1 - a_1) - \bar{m}\}$. Summarizing the above conditions, we shows that assertions concerning $\mathcal{L}$. The assertions on $\mathcal{G}$ are given by a simple comparison of the potential values for $\check{m}$, $\tilde{m}$, and $\hat{m}$. □

Proof of Proposition 5. It suffices to investigate $g(m)$ on the extreme line from $\bar{m}$ to $\hat{m}$:

$$
\bar{g}(\epsilon) \equiv g(\bar{m} + \epsilon(e_1 - e_3)) = f(\bar{m}) + (1 - \delta)\epsilon^2 - 2\bar{m}t\left((1 + \phi)(\bar{m} + \epsilon) - a\right) \quad \text{(A.4)}
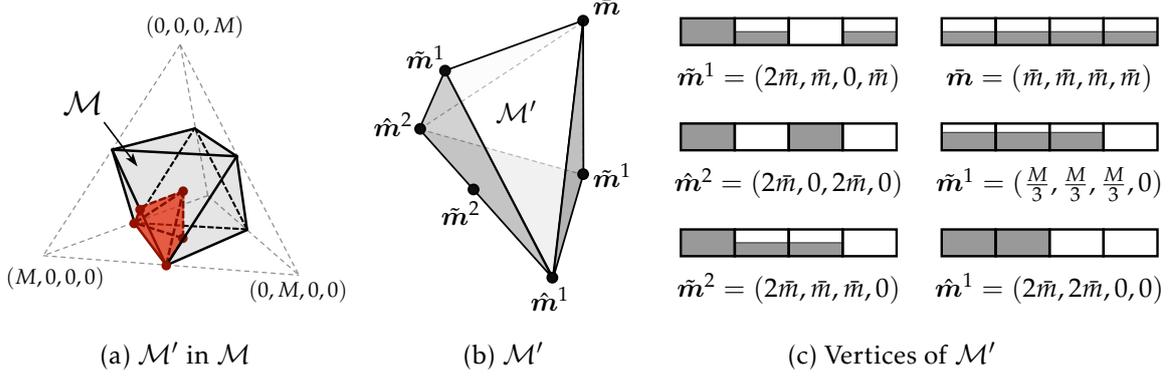$$

Figure A.1: $\mathcal{M}'$ and associated spatial patterns ($I = 4$).

where $\epsilon \in (0, \bar{m})$. Note that $\frac{t}{3} = 2\bar{m}t$. We note that $2(\bar{m} + \epsilon) - a$ is the shortage of labor in cell 1, i.e., the mass of commuters. Note that $\epsilon = 0$ ($\epsilon = \bar{m}$) corresponds to $\bar{m}$ ($\hat{m}$). We have $\bar{g}'(0) = -2\bar{m}(1 + \phi)t < 0$ and $\bar{g}'(\bar{m}) = 2\bar{m}(1 - \delta) - 2\bar{m}(1 + \phi)t$. $\hat{m}$ is a local maximizer iff $\bar{g}'(\bar{m}) > 0$, i.e., $t < \frac{1}{1+\phi}(1 - \delta) = \frac{1}{2}(1 - \delta)$. Also, $\bar{g}(0) \geq \bar{g}(\bar{m})$ is equivalent to $t \geq \frac{1}{4}(1 - \delta)$. $\qquad\square$

Proof of Proposition 6. For symmetry, we may focus on the following subset of $\mathcal{M}$:

$$\mathcal{M}' \equiv \{ \boldsymbol{m} \in \mathcal{M} \mid m_1 \geq \max\{m_2, m_3, m_4\}, \; m_2 \geq m_4, \; m_1 + m_3 \geq 2\bar{m} \}. \tag{A.5}$$

For instance, $(2\bar{m}, \bar{m}, \bar{m}, 0) \in \mathcal{M}^*$ but $(0, 2\bar{m}, \bar{m}, \bar{m}) \notin \mathcal{M}'$. Figure A.1 illustrates the convex polyhedron $\mathcal{M}'$ with associated spatial patterns.

From Lemma 1, $\mathcal{L}$ (and $\mathcal{G}$) restricted to $\mathcal{M}'$ should consist of the following patterns: $\bar{m} = (\bar{m}, \bar{m}, \bar{m}, \bar{m})$, $\tilde{m}^1 = (a, \bar{m}, 0, \bar{m})$, $\tilde{m}^2 = (a, \bar{m}, \bar{m}, 0)$, $\hat{m}^1 = (a, a, 0, 0)$, and $\hat{m}^2 = (a, 0, a, 0)$. Observe that all points other than $\bar{m}$ is located on the surface of $\mathcal{M}$.

Proposition 6 is a corollary of the following proposition. In the main text, we employ $(0, a, a, 0)$ instead of $\hat{m}^1 = (a, a, 0, 0)$ for the ease of interpretation at the first glance.

Proposition A.1. *Suppose $I = 4$ and let $t^* \equiv \frac{1}{2}(1 - \delta^2)$, $t^{**} \equiv \frac{1}{2}(1 - \delta)^2$, and $t^{***} \equiv \frac{1}{2}(1 - \delta)$ with $\delta \equiv e^{-\frac{\tau}{4}}$. Then,*

$$
\begin{cases}
\bar{m} \in \mathcal{L} & \forall t > 0 \\
\hat{m}^1 \in \mathcal{L} & \forall t < t^* \\
\hat{m}^2 \in \mathcal{L} & \forall t < t^{**} \\
\tilde{m}^1 \in \mathcal{L} & \forall t \in (\frac{1}{2}t^*, t^*) \\
\tilde{m}^2 \in \mathcal{L} & \forall t \in (t^{***}, \delta t^{***})
\end{cases}
\quad \text{and} \quad
\begin{cases}
\bar{m} \in \mathcal{G} & \forall t \geq \frac{1}{2}t^* \\
\hat{m}^1 \in \mathcal{G} & \forall t \leq \frac{1}{2}t^*
\end{cases}. \tag{A.6}
$$

Proof. Let $p'(\boldsymbol{x}; \boldsymbol{z})$ denote the one-sided directional derivative: $p'(\boldsymbol{x}; \boldsymbol{z}) \equiv \lim_{\epsilon \to +0}(p(\boldsymbol{x} + \epsilon \boldsymbol{z}) - p(\boldsymbol{x}))/\epsilon$. Note that at any state $\boldsymbol{m} \in \mathcal{M}$, the directional derivative of $f'(\boldsymbol{m}; \boldsymbol{z})$ toward $\boldsymbol{z} \in T\mathcal{M}$ is given by $f'(\boldsymbol{m}; \boldsymbol{z}) = \boldsymbol{m}^\top \mathbf{D} \boldsymbol{z} = \mathbf{A}(\boldsymbol{m})^\top \boldsymbol{z}$. For instance, when $\boldsymbol{z} = \boldsymbol{e}_j - \boldsymbol{e}_i$, $f'(\boldsymbol{m}; \boldsymbol{z}) = A_j(\boldsymbol{m}) - A_i(\boldsymbol{m})$. Note that $\frac{t}{4} = 2\bar{m}t$ and $a = 2\bar{m}$.

(i) The uniform pattern $\bar{m} = (\bar{m}, \bar{m}, \bar{m}, \bar{m})$ is a local maximizer for all δ and t (Lemma 1).

26

(ii) The monocentric pattern $\hat{m}^1 = (2\bar{m}, 2\bar{m}, 0, 0)$ is a local maximizer if and only if all of the directional derivatives $g'(\hat{m}^1; z)$ with $z \in \{z_{23}, z_{24}, z_{14} + z_{23}\}$ are negative, where $z_{ij} = e_j - e_i$ expresses a marginal replacement of firms from cell i to j.[33] The directions corresponds to the lines connecting $\hat{m}^1$ to other possible local maximizers $\{\tilde{m}^2, \tilde{m}^1, \bar{m}\}$. We compute that

$$g'(\hat{m}^1; z_{14} + z_{23}) = -4\bar{m}(1 - \delta^2) + 4\bar{m}(1 + \phi)t = 2g'(\hat{m}^1; z_{23}) = 2g'(\hat{m}^1; z_{24}).$$

Thus, $\hat{m}^1$ is a local maximizer of g when $t < \frac{1}{1+\phi}(1 - \delta^2) = \frac{1}{8}(1 - \delta^2)$.

(iii) The duocentric pattern $\hat{m}^2 = (2\bar{m}, 0, 2\bar{m}, 0)$ is a local maximizer if and only if all of the directional derivatives $g'(\hat{m}^2; z)$ with $z \in \{z_{12} + z_{34}, z_{32}, z_{32} + z_{34}\}$ are negative. We compute that

$$g'(\hat{m}^2; z_{12} + z_{34}) = g'(\hat{m}^2; z_{32} + z_{34}) = -4\bar{m}(1 - \delta)^2 + 4\bar{m}(1 + \phi)t = 2g'(\hat{m}^2; z_{32})$$

Thus, $\hat{m}^1$ is a local maximizer of g when $t < \frac{1}{1+\phi}(1 - \delta)^2$.

(iv) The pattern $\tilde{m}^1 \equiv (2\bar{m}, \bar{m}, 0, \bar{m})$ is a local maximizer if and only if all of the directional derivatives $g'(\tilde{m}^2; z)$ with $z \in \{z_{13}, z_{42}, z_{23} + z_{43}\}$ are negative. We compute that

$$g'(\tilde{m}^1; z_{13}) = -2\bar{m}(1 - \delta^2) + 2\bar{m}(1 + \phi)t, \tag{A.7}$$

$$g'(\tilde{m}^1; z_{42}) = -2\bar{m}(1 + \phi)t < 0, \tag{A.8}$$

$$g'(\tilde{m}^1; z_{23} + z_{43}) = -2\bar{m}(1 - \delta^2) - 4\bar{m}(1 + \phi)t. \tag{A.9}$$

It implies that $\tilde{m}^1$ is a local maximizer if and only if $\frac{1}{2(1+\phi)}(1 - \delta^2) < t < \frac{1}{1+\phi}(1 - \delta^2)$.

(v) The pattern $\tilde{m}^2 \equiv (2\bar{m}, \bar{m}, \bar{m}, 0)$ is a local maximizer if and only if all of the directional derivatives $g'(\tilde{m}^2; z)$ with $z \in \{z_{14}, z_{23}, z_{32}, z_{34}\}$ are negative. We compute that

$$g'(\tilde{m}^2; z_{14}) = -2\bar{m}(1 - \delta) + 2\bar{m}(1 + \phi)t, \tag{A.10}$$

$$g'(\tilde{m}^2; z_{23}) = 2\bar{m}\delta(1 - \delta) - 2\bar{m}(1 + \phi)t, \tag{A.11}$$

$$g'(\tilde{m}^2; z_{32}) = -2\bar{m}\delta(1 - \delta) - 2\bar{m}(1 + \phi)t < 0, \tag{A.12}$$

$$g'(\tilde{m}^2; z_{34}) = -\bar{m}(1 - \delta)^2 - 2\bar{m}(1 + \phi)t < 0. \tag{A.13}$$

Thus, $\tilde{m}^2$ is a local maximizer if and only if $\frac{1}{1+\phi}\delta(1 - \delta) < t < \frac{1}{1+\phi}(1 - \delta)$.

The claims concerning $\mathcal{G}$ follows by a comparison of the potential values for $\bar{m}, \hat{m}^1, \hat{m}^2, \tilde{m}^1$, and $\tilde{m}^2$. $\qquad \square$

Figure A.2 shows the contour of $g(m)$ on the polyhedron spanned by $\{\bar{m}, \hat{m}^1, \hat{m}^2\}$ for different values of t. Figure A.3 shows the partition of (τ, t)-space according to Proposition A.1. Observe that the partition based on $\mathcal{L}$ (Figure A.3b) become very complicated. The basic intuitions, however, does not change from that obtained by $\mathcal{G}$ (Figure A.3c). Under fixed δ, high (low) t encourages dispersion (agglomeration) of firms. $\qquad \square$

Proof of Proposition 7. Employing Lemma 1, one can enumerate all relevant spatial patterns

---

[33]If all of the directional derivatives are negative, it implies that the gradient of g points outward the polyhedron spanned $\mathcal{M}'$. For symmetry, it is the condition that the point is a local maximizer of g in $\mathcal{M}$.
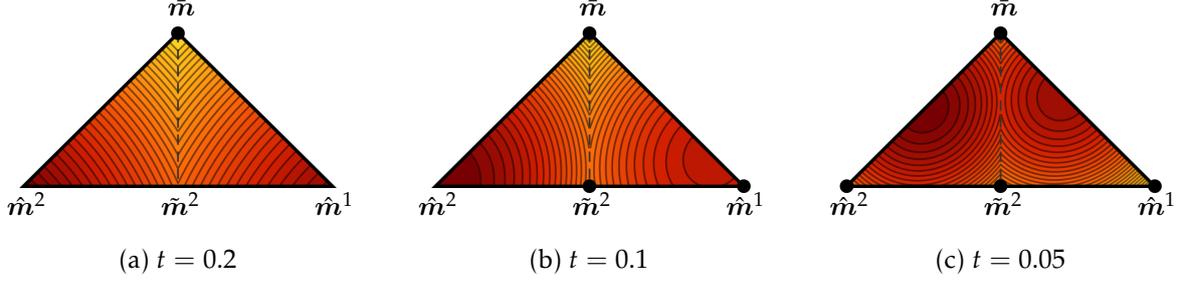
(a) $t = 0.2$ (b) $t = 0.1$ (c) $t = 0.05$

Figure A.2: Contours of $g(\boldsymbol{m})$ over the polygon spanned by $\{\hat{\boldsymbol{m}}^2, \hat{\boldsymbol{m}}^1, \bar{\boldsymbol{m}}\}$. $\delta = 0.1$. $\bullet$: points in $\mathcal{L}$.

(a) Partition based on $\mathcal{L}$ (b) Subset of Figure A.3a   (c) Partition based on $\mathcal{G}$
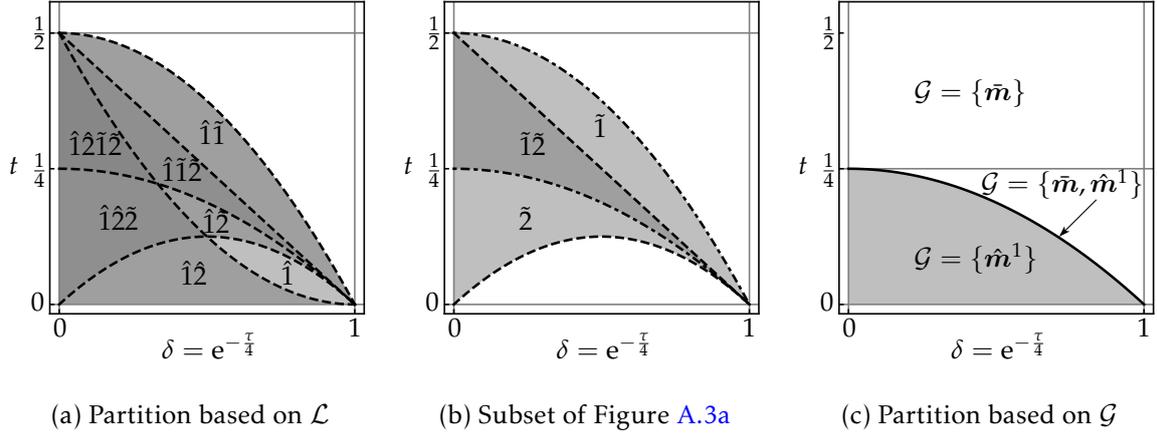
Figure A.3: Partitions of (δ, t)-space ($I = 4$). In (a) and (b), the numbers over a region indicate the patterns in $\mathcal{L}$ for the region. For instance, $\hat{1}\tilde{1}\tilde{2}$ means that $\hat{\boldsymbol{m}}^1, \tilde{\boldsymbol{m}}^2, \tilde{\boldsymbol{m}}^1 \in \mathcal{L}$. Also note that $\bar{\boldsymbol{m}} \in \mathcal{L}$ for all δ and t. (b) shows the partition for only $\tilde{\boldsymbol{m}}^1$ and $\tilde{\boldsymbol{m}}^2$.

such that $m_i \in \{0, \bar{m}, a\} = \{0, \bar{m}, 2\bar{m}\}$ for all $i \in \mathcal{I}$, which we refrain from explicitly listing up here. The proposition follows by just comparing the potential values for the spatial patterns. We note that, for any $\boldsymbol{m} \in \mathcal{M}$, $\check{h}(\boldsymbol{m})$ is computed by $\check{h}(\boldsymbol{m}) = \frac{t}{I} \sum_{i=1}^{I} |M_i - N_i - \alpha|$ where $M_i \equiv \sum_{k=1}^{i} m_k$, $N_i \equiv \sum_{k=1}^{i} (a - m_k)$, and α is the median of $\{M_i - N_i\}_{i \in \mathcal{I}}$ (see Rabin et al., 2011). The threshold values τ^*, t^*, t^{**}, and t^{***} are given by $\tau^* = -8 \log(\delta^*)$ with $\delta^* \equiv \frac{1}{6}(\sqrt{13} - 1)$, and $t^* = 2\bar{m}(1 - \delta)(1 + \delta)(1 + \delta + \delta^2)$, $t^{**} = 4\bar{m}(1 - \delta)^2(1 + \delta)^2$, and $t^{***} = 4\bar{m}\delta(1 - \delta)(1 + \delta)(1 + 2\delta)$ where $\delta = e^{-\frac{\tau}{8}}$. $\qquad\square$

B Non-atomic potential games

The precise definitions of *non-atomic game* and *non-atomic potential game* in this paper may be formalized as follows:

Definition B.1 (Nonatomic game). Let $\mathcal{Q} \equiv \{1, 2, \ldots, Q\}$ be a set of agent types, where Q is the number of types. Let $\mathcal{S}^q = \{1, 2, \ldots, S^q\}$ be the set of strategies for type $q \in \mathcal{Q}$ agents, where S^q is the number of strategies. Let $\mathcal{X} \subset \mathbb{R}_+^S$, where $S \equiv \sum_{q \in \mathcal{Q}} S^q$, denote the set of all possible strategy distributions, which is assumed to be closed and convex. A *non-atomic game* is a tuple $(\mathcal{S}, \mathbf{F})$ of the product of the strategy sets $\mathcal{S} \equiv \prod_{q \in \mathcal{Q}} \mathcal{S}^q$ and the payoff function $\mathbf{F}: \mathcal{X} \to \mathbb{R}^S$

that assigns a state $x \in \mathcal{X}$ the payoff for each strategy in $\mathcal{S}$.

Definition B.2 (Nonatomic potential game)**.** A non-atomic game $(\mathcal{S}, \mathcal{X}, \mathbf{F})$ is a *potential game* if there exists a scalar-valued *potential function* g on a neighborhood of $\mathcal{X}$ such that

$$\frac{\partial g(x)}{\partial x_i^q} = F_i^q(x) \qquad\qquad \forall i \in \mathcal{S}, q \in \mathcal{Q} \qquad\qquad \text{(B.1)}$$

for almost all $x \in \mathcal{X}$, where F_i^q is the payoff for an agent taking strategy $i \in \mathcal{S}^q$ and x_i^q is the mass of such agents in population $q \in \mathcal{Q}$.

Remark B.1. Definition B.1 is a simple generalization of the standard definition of non-atomic games with multiple types, or *population games* as synthesized by Sandholm (2010). The standard definition assumes that the state space $\mathcal{X}$ is the product of $(S^q - 1)$-simplexes $\mathcal{X} = \prod_{q \in \mathcal{Q}} \mathcal{X}^q$ where $\mathcal{X}^q \equiv \{x \in \mathbb{R}_+^{S^q} \mid \sum_{i \in \mathcal{S}^q} x_i^q = X^q\}$ and $X^q > 0$ is the total mass of agents in population $q \in \mathcal{Q}$. Definition B.2 is a non-atomic version of *generalized Nash equilibrium problems* introduced by Facchinei and Kanzow (2007) in the context of games with atomic players. A simple example is games with entry and exit (Sandholm, 2010, Exercise 3.1.8). $\diamond$

Remark B.2. Definition B.2 assumes that g is defined on a neighborhood of $\mathcal{X}$, not on $\mathcal{X}$, so the partial derivative (B.1) is well-defined. It is noted that, since we assume general convex state spaces, Definition B.2 is a non-atomic version of *generalized potential games* introduced by Facchinei et al. (2011) in the context of games with atomic players. $\diamond$

By employing Proposition 1, we can define an equivalent non-atomic potential game (Definition B.2) for the model laid out in Section 3.

Observation B.1. *The FO framework (Section 3) is a non-atomic potential game with two types of agents, i.e., $\mathcal{Q} = \{firms, households\}$. The strategy sets of firms and households are $\mathcal{I}$ and $\mathcal{I} \times \mathcal{I}$, respectively. The state space $\mathcal{X} \subset \mathbb{R}_+^{I+I^2}$ is given by the set of all feasible spatial distributions $\boldsymbol{m} = (m_i)_{i \in \mathcal{I}}$ and commuting patterns $\boldsymbol{n} = (n_{ij})_{ij \in \mathcal{I} \times \mathcal{I}}$:*

$$\mathcal{X} \equiv \left\{ (\boldsymbol{m}, \boldsymbol{n}) \in \mathbb{R}_+^I \times \mathbb{R}_+^{I^2} \,\middle|\, \text{s.t. (3.12), (3.13), and (3.9)} \right\}, \qquad\qquad \text{(B.2)}$$

which is closed and convex. The respective payoff functions of firms and households are:

$$\text{(firms)} \quad F_i(\boldsymbol{m}, \boldsymbol{n}) = A_i(\boldsymbol{m}) \qquad\qquad \forall i \in \mathcal{I}, \qquad\qquad \text{(B.3)}$$

$$\text{(households)} \quad F_{ij}(\boldsymbol{m}, \boldsymbol{n}) = -t_{ij} \qquad\qquad \forall ij \in \mathcal{I} \times \mathcal{I}, \qquad\qquad \text{(B.4)}$$

Define the function $g : \mathbb{R}_+^I \times \mathbb{R}_+^{I^2} \to \mathbb{R}$ by $g(\boldsymbol{m}, \boldsymbol{n}) = f(\boldsymbol{m}) - h(\boldsymbol{n})$ with $f(\boldsymbol{m})$ and $h(\boldsymbol{n})$ in (3.11) and (3.10). Then, g is a potential function for the above payoff functions because $\frac{\partial g(\boldsymbol{m}, \boldsymbol{n})}{\partial m_i} = A_i(\boldsymbol{m})$ and $\frac{\partial g(\boldsymbol{m}, \boldsymbol{n})}{\partial n_{ij}} = -t_{ij}$ for all $\boldsymbol{m} \in \mathbb{R}_+^I$ and $\boldsymbol{n} \in \mathbb{R}_+^{I^2}$. $\diamond$

As a non-atomic game, the model involves nonstandard constraints in the state space $\mathcal{X}$, the supply constraints (3.12) and (3.13) that determine the admissible pairs of strategy distributions $(\boldsymbol{m}, \boldsymbol{n})$. Such game can introduce subtleties from an evolutionary perspective. For

instance, when a cell is completely occupied by firms, households may not be able to choose the cell as their residential location. Thus, a dynamic for the game thus should express state-dependent interactions between the strategy switching behaviors of firms and households to respect supply constraints. Technically, one has to define appropriate dynamics whose orbits always stay in $\mathcal{X}$.

The reduction considered in Section 4 is meant to alleviate the subtlety discussed above. Interactions between firms and households are encapsulated by the lower-stage game and the resulting function $\check{h}(\boldsymbol{m})$. The state space in the upper stage $\mathcal{M}$, which consists only of the simplex constraint and the capacity constraints $m_i \leq a_i$, is independent of the strategy distribution of households $\boldsymbol{n}$.

Remark B.3. The projection dynamic (Dupuis and Nagurney, 1993) is well-defined for non-atomic games in Definition B.1, because the dynamic was originally developed for general convex state spaces. Other standard evolutionary dynamics in the literature, designed to be well-defined on the simplex (or products of them), may not remain in $\mathcal{X}$ in (B.2). Therefore, appropriate modifications, e.g., a projection step onto $\mathcal{X}$, are required. ◇

Remark B.4. In standard multiple-type non-atomic games, or population games, such inter-type interactions are summarized in the payoff function, and the strategy distributions only have simplex constraints (Remark B.1). In the model presented in Section 3, if (i) there is a fixed mass N of households and (ii) $\boldsymbol{r}$ and $\boldsymbol{w}$ can be obtained as simple functions of $(\boldsymbol{m}, \boldsymbol{n})$, one can define $F_i(\boldsymbol{m}, \boldsymbol{n}) = A_i(\boldsymbol{m}) - \phi w_i(\boldsymbol{m}, \boldsymbol{n}) - r_i(\boldsymbol{m}, \boldsymbol{n})$ and $F_{ij}(\boldsymbol{m}, \boldsymbol{n}) = w_i(\boldsymbol{m}, \boldsymbol{n}) - t_{ij} - r_i(\boldsymbol{m}, \boldsymbol{n})$ as the respective payoff functions for firms and households. The state space then becomes $\mathcal{X} = \{(\boldsymbol{m}, \boldsymbol{n}) \in \mathbb{R}_+^I \times \mathbb{R}_+^{I^2} \mid \sum_{i \in} m_i = M, \sum_{ij \in \mathcal{I} \times \mathcal{I}} n_{ij} = N\}$, reducing to a standard instance of multiple-type non-atomic games as laid out in Sandholm (2010). ◇

Observe that the lower- and upper-stage games are both an instance of non-atomic potential game, as in Definition B.2.

Observation B.2. *The lower-stage game is a non-atomic potential game with strategy set $\mathcal{I} \times \mathcal{I}$, state space $\mathcal{N}(\boldsymbol{m})$, payoff function $\mathbf{F} : \mathbb{R}_+^{I^2} \to \mathbb{R}_+^{I^2}$ in (B.4), and potential function $h : \mathbb{R}_+^{I^2} \to \mathbb{R}$ in (3.10).* ◇

Observation B.3. *Suppose Assumption 1 (ii). Then, the upper-stage game is a non-atomic potential game with strategy set $\mathcal{I}$, state space $\mathcal{M}$, payoff function $\check{\boldsymbol{\pi}} : \mathcal{M} \to \mathbb{R}^I$, and the potential function $g : \mathbb{R}_+^I \to \mathbb{R}$ employed in the main text.* ◇

Remark B.5. Regarding Definition B.2, it is also noted that we require differentiability of the potential function g for *almost all* $\boldsymbol{x} \in \mathbb{R}_+^S$, whereas the standard definition of non-atomic potential games in the literature by Sandholm (2001) requires differentiability for *all* admissible $\boldsymbol{x}$. This modification is required because the optimal value function of (S), $\check{h}$, can be nondifferentiable in the standard sense.

The function $\check{h}$ is instead *subdifferentiable*. That is, $\boldsymbol{c}(\boldsymbol{m})$ is a *subgradient*, or an element of *subdifferential*:

$$\partial \check{h}(\boldsymbol{m}) \equiv \left\{ \boldsymbol{z} \in \mathbb{R}^I \mid f(\boldsymbol{m}') \geq f(\boldsymbol{m}) + \boldsymbol{z}^\top (\boldsymbol{m}' - \boldsymbol{m}), \forall \boldsymbol{m}' \in \mathcal{M} \right\}. \tag{B.5}$$

This observation follows because (D) has a bounded solution for all $m \in \mathcal{M}$, $\check{h}$ takes a finite value and a bounded $c(m)$ exists for all $m \in \mathcal{M}$. See Rockafellar (1970) (Section 23). As illustrated by Figures 1, 4b, and D.4, $\check{h}$ is in fact differentiable (i.e., subdifferential is a singleton) almost everywhere in $\mathcal{M}$, and (4.7) holds true for almost all $m \in \mathcal{M}$. $\diamond$

C Alternative decompositions

In the main text, we decomposed the original model into a two-stage game, to reduce it to a non-atomic game where firms are the only agents. Technically, this corresponds to *partitioning* of the optimization problem (P0), or *Benders' Decomposition* after Benders (1962); Geoffrion (1972). In this Appendix, we briefly summarize two alternative decompositions, each corresponding to a different *interpretation* of the model. (The properties of equilibria are unaffected.) The reformulations can be in turn employed as a building blocks of extensions.

C.1 Households choose only job location in the lower stage

Suppose that households first choose job locations in the lower stage and then residential locations in the upper stage. Denote the residential location distribution of households by $\mathbf{N} \equiv (N_i)_{i \in \mathcal{I}}$, where $N_i \geq 0$ is the mass of households residing in cell $i \in \mathcal{I}$. The problem (P0) is equivalent to the following problem:

$$\max_{m \in \mathcal{M}, n \geq 0} \; f(m) - h(n) \tag{C.1a}$$

$$\text{s.t.} \quad m_i + N_i \leq a_i \qquad\qquad [r_i] \qquad\qquad \forall i \in \mathcal{I}, \tag{C.1b}$$

$$\sum_{i \in \mathcal{I}} n_{ij} \geq \phi m_j \qquad\qquad [w_j] \qquad\qquad \forall j \in \mathcal{I}, \tag{C.1c}$$

$$\sum_{j \in \mathcal{I}} n_{ij} = N_i \qquad\qquad [v_i] \qquad\qquad \forall i \in \mathcal{I}. \tag{C.1d}$$

Analogous to (S) and (D) in the main text, the following problem characterizes households' equilibrium behavior in the lower-stage game.

$$\min_{n \geq 0} \; \sum_{i \in \mathcal{I}} h(m), \quad \text{s.t.} \quad \text{(C.1c) and (C.1d)}. \tag{C.2}$$

$$\max_{(v,w) \geq 0} \; -\sum_{i \in \mathcal{I}} N_i v_i + \sum_{i \in \mathcal{I}} \phi m_i w_i, \quad \text{s.t.} \quad v_i \geq w_j - t_{ij} \qquad \forall i, j \in \mathcal{I}. \tag{C.3}$$

where v_i can be interpreted as the maximum possible payoff of households obtained by residing in $i \in \mathcal{I}$. Denote a solution for (C.3) by $\{w_i(m, \mathbf{N}), v_i(m, \mathbf{N})\}$.

The set of all feasible assignments of firms and households $(m, \mathbf{N})$ is

$$\mathcal{X}_1 \equiv \left\{ (m, \mathbf{N}) \in \mathbb{R}_+^I \times \mathbb{R}_+^I \; \middle| \; m_i + N_i \leq a_i \; \forall i \in \mathcal{I}, \; m \in \mathcal{M} \right\}, \tag{C.4}$$

which is a closed and convex set. Let $h_1(m, \mathbf{N})$ be the optimal value function of the problems

(C.2) and (C.3). Then, it follows that

$$\frac{\partial h_1(\boldsymbol{m}, \mathbf{N})}{\partial m_i} = \phi w_i(\boldsymbol{m}, \mathbf{N}) \quad \text{and} \quad \frac{\partial h_1(\boldsymbol{m}, \mathbf{N})}{\partial N_i} = -v_i(\boldsymbol{m}, \mathbf{N}) \tag{C.5}$$

almost everywhere. Let g_1 be defined by $g_1(\boldsymbol{m}, \mathbf{N}) \equiv f(\boldsymbol{m}) - h_1(\boldsymbol{m}, \mathbf{N})$. Then, we have

$$\frac{\partial g_1(\boldsymbol{m}, \mathbf{N})}{\partial m_i} = A_i(\boldsymbol{m}) - \phi w_i(\boldsymbol{m}, \mathbf{N}) \quad \text{and} \quad \frac{\partial g_1(\boldsymbol{m}, \mathbf{N})}{\partial N_i} = v_i(\boldsymbol{m}, \mathbf{N}) \tag{C.6}$$

almost everywhere. Also, (C.1) is equivalent to the problem $\max_{(\boldsymbol{m}, \boldsymbol{n}) \in \mathcal{X}_1} g_1(\boldsymbol{m}, \mathbf{N})$. If we interpret $A_i(\boldsymbol{m}) - \phi w_i(\boldsymbol{m}, \mathbf{N})$ and $v_i(\boldsymbol{m}, \mathbf{N})$ as the payoff functions of firms and households, respectively, the upper stage game may be interpreted as a non-atomic potential game.

C.2 Households choose only residential location in the lower stage

Next, suppose that households first choose their residential locations in the lower stage and then job locations in the upper stage. Denote the job distribution over households by $\mathbf{N} \equiv (N_j)_{j \in \mathcal{I}}$ in the lower-stage game, where, with notational abuse, $N_j \geq 0$ is the mass of households working in cell $j \in \mathcal{I}$. The problem (P0) is equivalent to the following problem:

$$\max_{\boldsymbol{m} \in \mathcal{M}, \boldsymbol{n} \geq \mathbf{0}} \quad f(\boldsymbol{m}) - h(\boldsymbol{n}) \tag{C.7a}$$

$$\text{s.t.} \quad m_i + \sum_{j \in \mathcal{I}} n_{ij} \leq a_i \qquad [r_i] \qquad \forall i \in \mathcal{I}, \tag{C.7b}$$

$$N_j \geq \phi m_j \qquad [w_j] \qquad \forall j \in \mathcal{I}, \tag{C.7c}$$

$$\sum_{i \in \mathcal{I}} n_{ij} = N_j \qquad [v_j] \qquad \forall i \in \mathcal{I}. \tag{C.7d}$$

The next problem characterizes households' equilibrium behavior in the lower stage.

$$\min_{\boldsymbol{n} \geq \mathbf{0}} \quad \sum_{i \in \mathcal{I}} h(\boldsymbol{m}), \quad \text{s.t.} \quad \text{(C.7b) and (C.7d).} \tag{C.8a}$$

$$\max_{(\boldsymbol{v}, \boldsymbol{w}) \geq \mathbf{0}} \quad -\sum_{i \in \mathcal{I}} N_j v_j - \sum_{i \in \mathcal{I}} r_i (a_i - m_i), \quad \text{s.t.} \quad v_j \geq -r_i - t_{ij} \qquad \forall i, j \in \mathcal{I} \tag{C.8b}$$

where v_j can be interpreted as households' maximum possible payoff from working in $j \in \mathcal{I}$. Denote a solution for (C.8b) by $\{w_i(\boldsymbol{m}, \mathbf{N}), v_j(\boldsymbol{m}, \mathbf{N})\}$.

The set of all feasible assignments of firms and households $(\boldsymbol{m}, \mathbf{N})$ is

$$\mathcal{X}_2 \equiv \left\{ (\boldsymbol{m}, \mathbf{N}) \in \mathbb{R}_+^I \times \mathbb{R}_+^I \,\middle|\, N_j \geq \phi m_j \,\forall j \in \mathcal{I}, \, M + \sum_{j \in \mathcal{N}} N_j \leq 1, \, \boldsymbol{m} \in \mathcal{M} \right\}, \tag{C.9}$$

which is a closed and convex set. Let $h_2(\boldsymbol{m}, \mathbf{N})$ be the optimal value function of the problems (C.2) and (C.8b). Then, it follows that

$$\frac{\partial h_2(\boldsymbol{m}, \mathbf{N})}{\partial m_i} = r_i(\boldsymbol{m}, \mathbf{N}) \quad \text{and} \quad \frac{\partial h_2(\boldsymbol{m}, \mathbf{N})}{\partial N_j} = -v_j(\boldsymbol{m}, \mathbf{N}) \tag{C.10}$$

if h is differentiable. Let g_2 be defined by $g_2(\boldsymbol{m}, \mathbf{N}) \equiv f(\boldsymbol{m}) - h_2(\boldsymbol{m}, \mathbf{N})$. Then, we have

$$\frac{\partial g_2(\boldsymbol{m}, \mathbf{N})}{\partial m_i} = A_i(\boldsymbol{m}) - r_i(\boldsymbol{m}, \mathbf{N}) \quad \text{and} \quad \frac{\partial g_2(\boldsymbol{m}, \mathbf{N})}{\partial N_j} = v_j(\boldsymbol{m}, \mathbf{N}) \tag{C.11}$$

almost everywhere. Also, (C.7) is equivalent to the problem $\max_{(\boldsymbol{m}, \boldsymbol{n}) \in \mathcal{X}_2} g_2(\boldsymbol{m}, \mathbf{N})$. If we interpret $A_i(\boldsymbol{m}) - r_i(\boldsymbol{m}, \mathbf{N})$ and $v_j(\boldsymbol{m}, \mathbf{N})$ as the payoff of firms and households, respectively, the upper stage game may be interpreted as a non-atomic potential game.

D The two-cell city in detail

This Appendix studies the two-cell case by straightforwardly investigating the equilibrium conditions without Assumption 3. Let $\mathcal{M}_0 \subset \mathcal{M}$ be defined by

$$\mathcal{M}_0 \equiv \{\boldsymbol{m} \in \mathcal{M} \mid (1 + \phi)m_i \leq a_i \ \forall i \in \mathcal{I}\}. \tag{D.1}$$

Upon inspection, the solution for the lower-stage equilibrium (Definition 2) is given by:

$$\begin{cases} n_{ii} = \phi m_i, & n_{ij} = 0, & \check{r}_i = w_i = 0, \\ n_{jj} = \phi m_j, & n_{ji} = 0, & \check{r}_j = w_j = 0, \end{cases} \quad \text{if } \boldsymbol{m} \in \mathcal{M}_0, \tag{D.2a}$$

$$\begin{cases} n_{ii} = \phi m_i, & n_{ij} = 0, & \check{r}_i = w_i = t, \\ n_{jj} = \phi m_j, & n_{ji} = (1 + \phi)m_i - a_i, & \check{r}_j = w_j = 0, \end{cases} \quad \text{if } (1 + \phi)m_i - a_i > 0, \ i \in \mathcal{I}, \tag{D.2b}$$

where we employ the convention that $i \neq j$.

If $\boldsymbol{m} \in \mathcal{M}_0$, or $(1 + \phi)m_i - a_i \leq 0$ for both cells 1 and 2, all households can reside at their workplace and there are no inter-cell commutes. In this case, we let $\check{r}_i = w_i = 0$ for both $i \in \mathcal{I}$. On the other hand, if $\boldsymbol{m} \notin \mathcal{M}_0$, then $(1 + \phi)m_i - a_i > 0$ for one of the cells, meaning that there is a shortage of labor in cell i. There must be $(1 + \phi)m_i - a_i$ workers who commute to cell i from the other cell. In this case, $w_i = t$ is the minimal possible wage rate in cell i to compensate for the commuting cost. The highest bid for the land rents in each cell i must coincide with the wage rates therein. Thus, we have $\check{r}_i = w_i = t$.

Take $m_1 = \hat{m}$ as a candidate of equilibrium in the upper-stage game. For simplicity, suppose $M > a_1$ and $a_1 > a_2$ so that $\hat{m} = a_1$. Since $(1 + \phi)\hat{m} - a_1 = \phi a_1 > 0$, there must be commuters from cell 2 to cell 1. Thus, $c_1 = (1 + \phi)t$ and $c_2 = 0$, which gives

$$\check{\pi}_1 = A_1 - c_1 = \hat{m} + \delta(M - \hat{m}) - (1 + \phi)t, \tag{D.3a}$$

$$\check{\pi}_2 = A_2 - c_2 = (M - \hat{m}) + \delta\hat{m} \tag{D.3b}$$

where $\delta = e^{-\tau}$. Since $0 < m_2 < a_2$, we have $\hat{r}_2 = 0$ from (4.5). For $\hat{r}_1$, it must be that

$$\hat{r}_1 = \check{\pi}_1 - \check{\pi}_2 = 2(1 - \delta)(\hat{m} - \bar{m}) - (1 + \phi)t \tag{D.4}$$

to satisfy equilibrium condition (4.4). The first term in $\hat{r}_1$ is $A_1 - A_2$, which is positive. Because of the asymmetry $a_1 > a_2$, firms in cell 1 enjoy higher level of technological externalities than
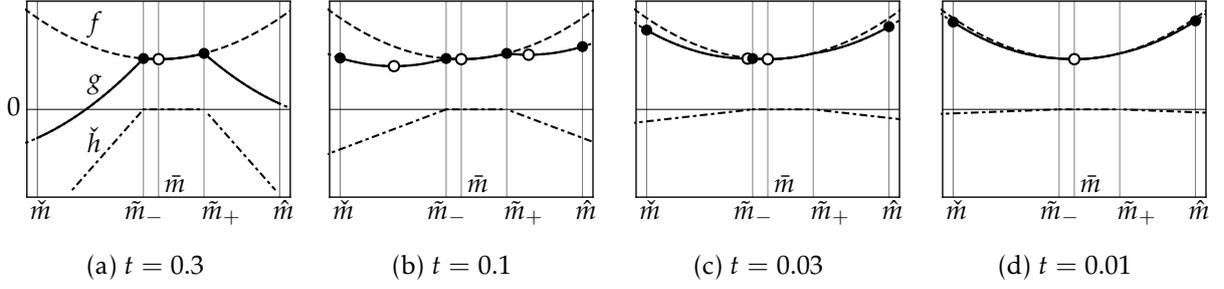
(a) $t = 0.3$ (b) $t = 0.1$ (c) $t = 0.03$ (d) $t = 0.01$

Figure D.4: Graphs of $g(m_1)$ for high, medium, and low values of t with excess land supply. $M = 0.4$, $a_1 = 0.55$, $\delta = 0.1$, $\phi = 1$. ●: local maximizers; ○: locally minimizing KKT points.

those in cell 2 at $m_1 = \hat{m}$. The second term in $\hat{r}_1$ is the minimized cost that firms have to bear. If $\hat{r}_1 \geq 0$, then $\hat{r}_1$ satisfies (4.5), and $m_1 = \hat{m}$ is an equilibrium for the upper-stage game.

Note that the condition $\hat{r}_1 \geq 0$ indicates either that δ is small (τ is large) or that t is small. If t is large, then firms cannot support the devoted business cell 1. After paying the commuting costs and land rents, firms' profits in cell 1 become lower than those in cell 2. Similarly, if δ is large, the difference between the production levels A_1 and A_2 becomes smaller, thereby reducing the benefit of agglomeration to cell 1.

We may continue the above procedure to enumerate all equilibria. With the aid of the potential function, the main text instead employs a more systematic approach.

The solution (D.2) gives the minimized total commuting costs $\check{h}(\boldsymbol{m})$ in the city:[34]

$$\check{h}(\boldsymbol{m}) = \max \left\{ t \left((1+\phi)m_2 - a_2 \right), 0, t \left((1+\phi)m_1 - a_1 \right) \right\} \tag{D.5}$$

$$= \begin{cases} t \left((1+\phi)(M - m_1) - a_2 \right) & \text{if } m_1 \leq \tilde{m}_-, \\ 0 & \text{if } \tilde{m}_- \leq m_1 \leq \tilde{m}_+, \\ t \left((1+\phi)m_1 - a_1 \right) & \text{if } m_1 \geq \tilde{m}_+. \end{cases} \tag{D.6}$$

where $\tilde{m}_- = M - \frac{a_2}{1+\phi}$ and $\tilde{m}_+ = \frac{a_1}{1+\phi}$. We confirm that $\check{h}(\boldsymbol{m})$ can be expressed as a piecewise maximum of affine functions and that $\nabla h(\boldsymbol{m}) = \boldsymbol{c}(\boldsymbol{m})$ almost everywhere.

The interval $[\tilde{m}_-, \tilde{m}_+]$ corresponds to $\mathcal{M}_0$. Observe that Assumption 3 implies $\tilde{m}_- = \tilde{m}_+$ and the plateau $(\tilde{m}_-, \tilde{m}_+)$ of $\check{h}(m_1)$ vanishes. Figure D.4 shows graphs of $g(m_1)$ without Assumption 3. As the figure illustrates, there is no essential difference in the properties of local maximizers with additional complications. Agglomeration of firms in either of the cells become local maximizers when t is low. Also note that $m_1 = \bar{m}$ is always a local minimizer. This generalizes to the circular case treated in Section 6 (see Proposition D.1 below).

Remark D.1. Suppose $m_1 \neq \tilde{m}_-, \tilde{m}_+$. Then, we have $\check{h}'(m_1) = c_1(\boldsymbol{m}) - c_2(\boldsymbol{m})$ and $f'(m_1) = 2(1-\delta)(m_1 - \bar{m}) = A_1(\boldsymbol{m}) - A_2(\boldsymbol{m})$, so that $g'(m_1) = \check{\pi}_1(\boldsymbol{m}) - \check{\pi}_2(\boldsymbol{m})$. For example, observe that $\hat{m} \in \mathcal{L}$ if and only if $g'(\hat{m}) = \check{\pi}_1 - \check{\pi}_2 = \hat{r}_1 > 0$. Thus, $\hat{m}$ is a local maximizer whenever it is an equilibrium, except for the borderline case $\hat{r}_1 = 0$.   ◇

---

[34]Observe that the function $\check{h}(\boldsymbol{m})$ can be interpreted interchangeably as the total commuting cost $\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} t_{ij} n_{ij}$ or as total household surplus $\phi \sum_{i \in \mathcal{I}} w_i m_i - \sum_{i \in \mathcal{I}} \check{r}_i (a_i - m_i)$ with the solution (D.2). This demonstrates the strong duality between (S) and (D).
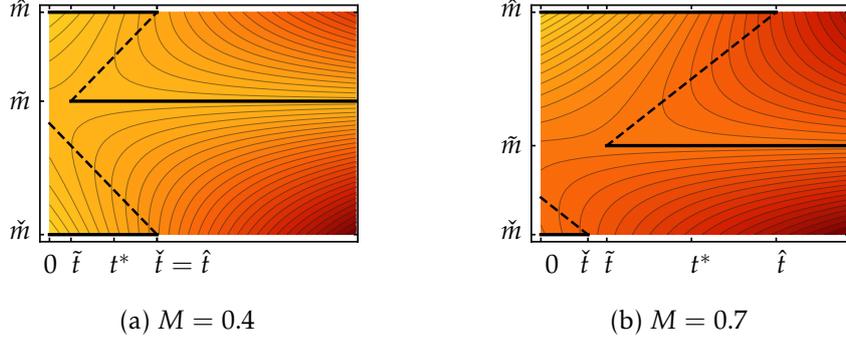
(a) $M = 0.4$ (b) $M = 0.7$

Figure D.5: Effects of varying M. $a_1 = 0.6$. When $M < a_1$, it is more likely that $m_1 = \check{m}$ become a local maximizer because cell 1 cannot contain all firms. When $M > a_1$, the converse holds true, and $m_1 = \check{m}$ is less likely to be a local maximizer. If $a_1 = a_2 = \frac{1}{2}$, the diagram become isomorphic to Figure 2a, and there are no qualitative differences when M varies.

Remark D.2. The mass of firms M acts as a scaling factor for both t and a_1. For t, observe that $|\check{h}'(m_1)| = (1+\phi)t = \frac{t}{M}$ whenever $\check{h}$ is differentiable under Assumption 3. For a_1, we note that when $M < a_1$, it is more likely that a concentration to cell 2 becomes a local maximizer, because cell 2 can contain the majority of firms (Figure D.5a); when $M > a_1$, the converse holds true (Figure D.5b). In Section 6, for simplicity we assume $M = \frac{1}{1+\phi} = \frac{1}{2}$ so that $M = a_1$ holds true under the symmetry of cells ($a_1 = a_2 = \frac{1}{2}$). $\diamond$

In addition, we note that the uniform distribution $\bar{m} = (\bar{m}, \bar{m}, \dots, \bar{m})$ can never locally maximize the potential function without Assumption 3 in the circular case in Section 6.

Proposition D.1. *Suppose Assumption 2. Suppose that $\ell_{ij} = \frac{1}{I} \min\{|i - j|, I - |i - j|\}$ and that $\bar{m} \in \mathcal{M}_0$. Then, $\bar{m}$ is a locally minimizing KKT point of (P) if and only if $(1+\phi)M < 1$.*

Proof. Since $\bar{m} \in \mathcal{M}_0$, $a_i - (1+\phi)\bar{m} \le 0$ for all $i \in \mathcal{I}$. Let

$$\mathcal{O} \equiv \left\{ \epsilon \in T\mathcal{M} \,\middle|\, \max_{i \in \mathcal{I}}\{\epsilon_i\} \le a_i - (1+\phi)\bar{m} \right\}. \tag{D.7}$$

Note that $\mathcal{O} = \{\mathbf{0}\}$ iff $(1+\phi)M = 1$, i.e., $\mathcal{O}$ admits nonzero elements iff $(1+\phi)M < 1$.

Under Assumption 2, we have

$$f(\boldsymbol{m}) = \frac{1}{2}\sum_{i \in \mathcal{I}}\sum_{j \in \mathcal{I}} d_{ij}(m_i - \bar{m})(m_j - \bar{m}) + \frac{1}{2}\bar{m}^2 \sum_{i \in \mathcal{I}}\sum_{j \in \mathcal{I}} d_{ij} = f(\boldsymbol{m} - \bar{m}) + f(\bar{m}). \tag{D.8}$$

For all nonzero $\epsilon \in \mathcal{O}$, we have $g(\boldsymbol{m}) = f(\boldsymbol{m})$, since by definition $\bar{m} + \epsilon \in \mathcal{M}_0$ and $h(\bar{m} + \epsilon) = 0$. For $\epsilon \in \mathcal{O}$ define $\bar{g}(\epsilon) \equiv g(\bar{m} + \epsilon) = f(\epsilon) + f(\bar{m})$ be the potential function defined on $\mathcal{O}$. It follows $\nabla \bar{g}(\mathbf{0}) = \mathbf{0}$ and $\nabla^2 \bar{g}(\mathbf{0}) = \mathbf{D}$. Note $\mathbf{D}$ is positive definite on $T\mathcal{M}$ and thus on $\mathcal{O}$. Thus, $\epsilon = \mathbf{0}$ is a local minimizer of $\bar{g}$, whence $\bar{m}$ is a locally minimizing KKT point of (P). $\square$

E   Computation of potential value when I is large

For $J \le K$, let $f_J = f(\boldsymbol{m}^J)$ and $h_J \equiv \check{h}(\boldsymbol{m}^J)$, so that $g(\boldsymbol{m}^J) = g_J \equiv f_J - h_J$. Let $f_\infty \equiv f(\bar{m})$, $h_\infty \equiv \check{h}(\bar{m}) = 0$, and $g_\infty \equiv f_\infty - h_\infty = f_\infty$. Figure 9 is obtained by numerically comparing the

potential values $\{g_\infty\} \cup \{g_J\}_{J=1}^K$, where $K = 50$, so that $I = 4\mathrm{LCM}(20) = 931{,}170{,}240$.

Let $R \equiv \frac{I}{4J}$ be the "radius" of a business center in $\boldsymbol{m}^J$; R is an integer for all $J \leq K$ because $I = 4\mathrm{LCM}(K)$. For the first block element $\check{m}_J$ in $\boldsymbol{m}^J$, the set of cell indices for the residential cells ($m_i = 0$) is $\mathcal{I}_J^- \cup \mathcal{I}_J^+$ with $\mathcal{I}_J^- \equiv \{1, 2, \ldots, R\}$ and $\mathcal{I}_J^+ \equiv \{3R + 1, \ldots, 4R\}$, whereas the business area ($m_i = a$) is $\mathcal{I}_J \equiv \{R + 1, R + 2, \ldots, 3R\}$.

In the lower stage, households in cell $i \in \mathcal{I}_J^-$ should commute to cell $i + R \in \mathcal{I}_J$, whereas those in cell $i \in \mathcal{I}_J^+$ should commute to cell $i - R \in \mathcal{I}_J$, thereby the commuting cost for every household is $\frac{tR}{I}$. The mass of such households is $(|\mathcal{I}_J^+| + |\mathcal{I}_J^-|)a = 2Ra$, and there are J symmetric small cities. The total commuting cost h_J is

$$h_J = J \times 2Ra \times \frac{tR}{I} = \frac{t}{8J}. \tag{E.1}$$

Observe that h_J is decreasing in J and it is independent of I.

Noting that I is extremely large, we employ continuous versions of f by taking limit $I \to \infty$ to simplify numerical computations. The continuous version of f is given as follows:

$$\frac{1}{2} \iint_{\mathcal{C} \times \mathcal{C}} \exp\left(-\tau \ell(x, y)\right) m(x) m(y) \mathrm{d}x \mathrm{d}y \tag{E.2}$$

where $\mathcal{C} = [0, 1)$ parameterizes the circle; $\ell(x, y) = \min\{|x - y|, 1 - |x - y|\}$ $(x, y \in \mathcal{C})$ is a continuous version of $\{\ell_{ij}\}$, and $m(x)$ is the density interpretation of the spatial pattern. For example, $\bar{\boldsymbol{m}}$ corresponds to $m(x) = \frac{1}{2}$ for all $x \in \mathcal{C}$, $\boldsymbol{m}^1$ may correspond to $m(x) = 1$ for $x \in [\frac{1}{4}, \frac{3}{4}]$ and 0 otherwise, $\boldsymbol{m}^2$ may correspond to $m(x) = 1$ for $x \in [\frac{1}{8}, \frac{3}{8}] \cup [\frac{5}{8}, \frac{7}{8}]$ and 0 otherwise, and so on. In concrete terms, we let the density interpretation of $\boldsymbol{m}^J$ be

$$m_J(x) = \begin{cases} 1 & \text{if } x \in \bigcup_{j=1}^J \mathcal{B}_j^J \\ 0 & \text{otherwise,} \end{cases} \tag{E.3}$$

where $\mathcal{B}_j^J \equiv [\frac{j}{2J} - \frac{1}{4J}, \frac{j}{2J} + \frac{1}{4J}]$ is the jth business center in $\boldsymbol{m}^J$. For $f(\bar{\boldsymbol{m}})$, letting $\delta = \mathrm{e}^{-\frac{\tau}{I}}$ gives

$$f(\bar{\boldsymbol{m}}) = \frac{\bar{m}^2}{2} \mathbf{1}^\top \mathbf{D} \mathbf{1} = \frac{1}{8I} \frac{1 + \delta}{1 - \delta} \left(1 - \delta^{\frac{I}{2}}\right) = \frac{1}{8I} \frac{1 + \mathrm{e}^{-\frac{\tau}{I}}}{1 - \mathrm{e}^{-\frac{\tau}{I}}} \left(1 - \mathrm{e}^{-\frac{\tau}{2}}\right) \longrightarrow \frac{1}{4\tau} \left(1 - \mathrm{e}^{-\frac{\tau}{2}}\right) \tag{E.4}$$

as $I \to \infty$. The last expression naturally coincides with a computation based on (E.2):

$$\frac{1}{2} \left(\frac{1}{2}\right)^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{e}^{-\tau |y|} \mathrm{d}y = \frac{1}{4\tau} \left(1 - \mathrm{e}^{-\frac{\tau}{2}}\right). \tag{E.5}$$

For $\boldsymbol{m}^J$, a computation based on (E.2) yields

$$f_J = (1 + \Psi_J) f_\infty \quad \text{where} \quad \Psi_J \equiv \left(1 + \frac{2C(\delta_J)}{\log(\delta_J)}\right) \bar{\Psi}_J \tag{E.6}$$

with $C(z) \equiv \frac{1-z}{1+z}$; $\delta_J \equiv \mathrm{e}^{-\frac{\tau}{2J}}$; $\bar{\Psi}_J \equiv 1$ for an even J and $\bar{\Psi}_J \equiv C(\mathrm{e}^{-\frac{\tau}{2}})^{-1}$ for an odd J. The expression can also be obtained by first computing an explicitly discrete potential value for
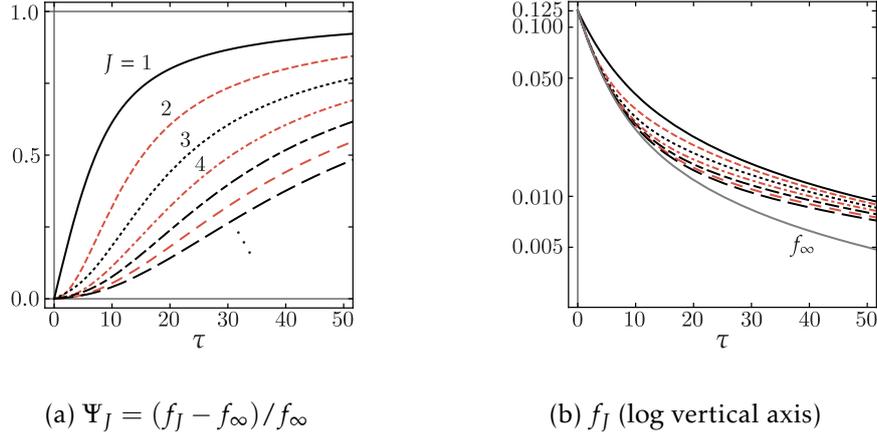
(a) $\Psi_J = (f_J - f_\infty)/f_\infty$

(b) f_J (log vertical axis)

Figure E.6: Graphs of Ψ_J and f_J for $J \in \{1, 2, \ldots, 7\}$.

(a) Δg_J in the t-axis ($\tau = 25$)
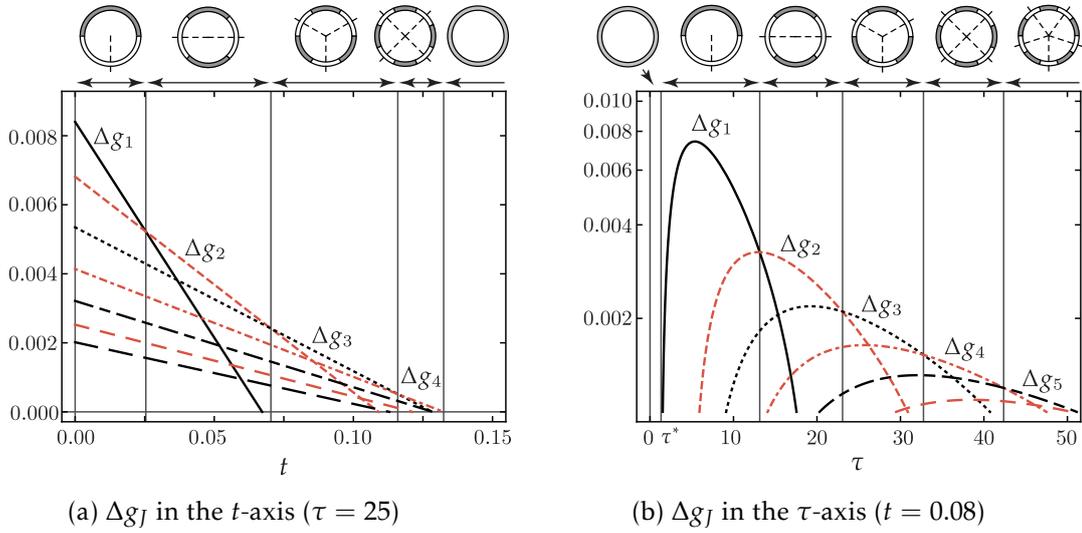
(b) Δg_J in the τ-axis ($t = 0.08$)

Figure E.7: Relative potential values $\Delta g_J \equiv g_J - g$ and potential maximizing patterns.

m^J and then taking the limit $I \to \infty$. We have $\Psi_J \in (0,1)$ (Figure E.6a). Additional calculation shows that Ψ_J is decreasing in J, implying that f_J is basically decreasing in J.[35] The benefit of agglomeration decreases with the number of business centers decreases (see Figure E.6b).

As an example of global potential maximization in $\mathcal{M}^*$, Figure E.7 depicts the differences $\Delta g_J \equiv g_J - g_\infty$ for $1 \leq J \leq 7$. When $\Delta g_J < 0$ for all J, the uniform distribution $\bar{m}$ globally maximizes g among $\mathcal{M}^*$. If $J^* = \arg\max_{1 \leq J \leq K} \Delta g_J$ and $\Delta g_{J^*} > 0$, then J^* globally maximizes g in $\mathcal{M}^*$. Thus, the patterns that correspond to the upper envelope of $\{0, \Delta g_1, \Delta g_2, \ldots, \Delta g_K\}$ also correspond to the global maximizers of the potential function in $\mathcal{M}^*$. Figure E.7a and Figure E.7b show $\{\Delta g_J\}_{J=1}^{7}$ on to the vertical and horizontal broken lines in Figure 9b, respectively. Observe that the sequence of potential maximizing symmetric patterns while decreasing t is given by $\bar{m} \to m^4 \to m^3 \to m^2 \to m^1$ for the former (i.e., $J^* = 4$), whereas the sequence for an increasing τ is given by $\bar{m} \to m^1 \to m^2 \to m^3 \to \cdots$ for the latter.

It is also noted that $g_J \geq g_\infty$ implies that m^J is an equilibrium when I grows large. The

---

[35]In concrete terms, $\{f_J\}$ can be shown to satisfy the following properties: f_J is monotonically decreasing in τ for all J, $f_J > f_\infty$ for all J, $f_J > f_{J+2}$ for all J, $f_1 > f_2$, $\lim_{J \to \infty} f_J = f_\infty$, and $\lim_{\tau \to 0} f_J = \lim_{\tau \to 0} f_\infty = \frac{1}{8}$.

market prices in the lower stage over $\mathcal{I}_J^- \cup \mathcal{I}_J \cup \mathcal{I}_J^+$ is then given by $\check{r}_i = w_i = \frac{t}{I}(i-1)$ and $\check{r}_{4R+(i-1)} = \check{r}_i$ for all $i \in \{1, 2, \ldots, 2R\}$. For $\boldsymbol{m}^J$ to be an equilibrium, it must be that

$$\hat{r}_i = A_i(\boldsymbol{m}^J) - \phi w_i - \check{r}_i - \pi^* = A_i(\boldsymbol{m}^J) - 2\check{r}_i - \pi^* \geq 0 \qquad \forall i \in \mathcal{I}_J \qquad \text{(E.7)}$$

where π^* is the equilibrium profit level. We first note that $\mathbf{A}(\boldsymbol{m})$ is symmetric and concave on $\mathcal{I}_J$. From symmetry of the spatial pattern, $A_k(\boldsymbol{m}^J) = A_{3R-k}(\boldsymbol{m}^J)$. With $A_k \equiv A_k(\boldsymbol{m}^J)$, we have

$$A_k'' \equiv \frac{(A_{k+1} - A_k) - (A_k - A_{k-1})}{2} = \frac{A_{k+1} - 2A_k + A_{k-1}}{2} < 0 \quad \forall k \in \mathcal{I}_J \setminus \{R+1, 3R\}, \quad \text{(E.8)}$$

thereby $\{A_k\}$ is mid-point concave on $\mathcal{I}_J$. In fact, for all $k \in \mathcal{I}_J \setminus \{R+1, 3R\}$ we compute

$$A_k'' = \left(\frac{\delta + \delta^{-1}}{2} - 1\right) \sum_{l=1}^{\frac{I}{2}-1} d_{k,k+l}(m_{k-l} + m_{k+l}) - (1-\delta)\left(m_k - \delta^{-1} d_{1,\frac{I}{2}} m_{k+\frac{I}{2}}\right) < 0 \qquad \text{(E.9)}$$

where $m_k \in \{0, a\}$ (mod I for indices); and $\delta = e^{-\frac{\tau}{I}}$, where the inequality follows by noting that $\frac{\delta + \delta^{-1}}{2} - 1 = \cosh(\frac{\tau}{I}) - 1 < 0$; $m_k - \delta^{-1} d_{1,\frac{I}{2}} m_{k+\frac{I}{2}} > m_k - m_{k+\frac{I}{2}} \geq a - a = 0$. In addition, one shows that $\frac{1}{2}(A_I + A_1) = \frac{1}{2}(A_{4R} + A_{4R+1}) = \bar{A}$ as well as that $A_1 > \bar{A}$ and $A_{4R} > \bar{A}$, where $\bar{A}$ is the uniform level of $\mathbf{A}(\boldsymbol{m})$ at $\bar{\boldsymbol{m}}$. Then, the equilibrium profit of firms may be evaluated by the average profit of firms in cell 1 and I as $\pi^* = \frac{1}{2}(\check{\pi}_1 + \check{\pi}_I) = \bar{A} - \frac{2t}{I}\left(R - \frac{1}{2}\right) = \bar{A} - \frac{t}{2J} + \frac{t}{I}$. Because $\{A_i\}$ is symmetric and mid-point concave on $\mathcal{I}_J$, (E.7) is satisfied when

$$\hat{r}_{2R} = \hat{r}_{2R+1} = A_{2R} - \frac{t}{2J} - \bar{A} - \frac{t}{I} \geq 0. \qquad \text{(E.10)}$$

On the other hand, $g_J - g_\infty$ is computed as

$$f_J - h_J - f_\infty = \frac{J}{2}\sum_{k \in \mathcal{I}_J} aA_k - J \times 2R \times \frac{t}{4J} - \frac{1}{2}I\bar{m}\bar{A} = \frac{J}{2I}\sum_{k \in \mathcal{I}_J}\left(A_k - \frac{t}{2J} - \bar{A}\right). \qquad \text{(E.11)}$$

If $g_J - g_\infty \geq 0$, then it must be that $A_{2R} - \frac{t}{2J} - \bar{A} > 0$ because A_k is (mid-point) concave and symmetric over $\mathcal{I}_J$. If instead $A_{2R} - \frac{t}{2J} - \bar{A} \leq 0$, then $A_k - \frac{t}{2J} - \bar{A} < 0$ for all $k \in \mathcal{I}_J \setminus \{2R, 2R+1\}$, implying $g_J - g_\infty < 0$. We thus observe that $g_J \geq g_\infty$ is a sufficient condition for (E.10) when I is extremely large so that the last term $\frac{t}{I}$ in (E.10) is negligible.

38

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 models with social interactions: Multiplicity, stability, and continuous limit of equilibria," *Journal of Mathematical Economics*, 2017, *69*, 22–37.

__ , **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, *36* (11), 1729–1759.

Anas, Alex, Richard Arnott, and Kenneth A. Small, "Urban spatial structure," *Journal of Economic Literature*, 1998, *36* (3), 1426–1464.

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

Benders, Jacques F., "Partitioning procedures for solving mixed-variables programming problems," *Numerische Mathematik*, 1962, *4* (1), 238–252.

Berliant, Marcus and Takatoshi Tabuchi, "Equilibrium commuting," *Economic Theory*, 2018, *65* (3), 609–627.

__ , **Shin-Kun Peng, and Ping Wang**, "Production externalities and urban configuration," *Journal of Economic Theory*, 2002, *104* (2), 275–303.

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.

Blume, Lawrence E., "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, 1993, *5* (3), 387–424.

__ , "Population games," *The Economy as an Evolving Complex System II*, 1997, *27*, 425–460.

Bragard, Jean and Pascal Mossay, "Stability of a spatial model of social interactions," *Chaos, Solitons & Fractals*, 2016, *83*, 140–146.

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.

Cheung, Man-Wah and Ratul Lahkar, "Nonatomic potential games: The continuous strategy case," *Games and Economic Behavior*, 2018, *108*, 341–362.

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** __ , "Urban land use," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5, Elsevier, 2015, pp. 467–560.

Facchinei, Francisco and Christian Kanzow, "Generalized Nash equilibrium problems," *4OR*, 2007, *5* (3), 173–210.

__ , **Veronica Piccialli, and Marco Sciandrone**, "Decomposition algorithms for generalized potential games," *Computational Optimization and Applications*, 2011, *50* (2), 237–262.

Fujita, Masahisa, "A monopolistic competition model of spatial agglomeration: Differentiated product approach," *Regional Science and Urban Economics*, 1988, *18* (1), 87–124.

— , *Urban Economic Theory: Land Use and City Size*, Cambridge University Press, 1989.

— **and Hideaki Ogawa**, "Multiple equilibria and structural transition 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 (2nd Edition)*, Cambridge University Press, 2013.

Geoffrion, Arthur M., "Generalized Benders decomposition," *Journal of Optimization Theory and Applications*, 1972, *10* (4), 237–260.

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

Ikeda, Kiyohiro, Mikihisa Onda, and Yuki Takayama, "Bifurcation theory of a racetrack economy in a spatial economy model," *Networks and Spatial Economics*, 2019, *19* (1), 57–82.

— **, 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.

Imai, Haruo, "CBD hypothesis and economies of agglomeration," *Journal of Economic Theory*, 1982, *28* (2), 275–299.

Krugman, Paul R., "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.

Lucas, Robert E., "Externalities and cities," *Review of Economic Dynamics*, 2001, *4* (2), 245–274.

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

Malykhin, Nikita and Philip Ushchev, "How market interactions shape the city structure," *Regional Science and Urban Economics*, 2018, *71*, 122–136.

Matsuyama, Kiminori, "Geographical advantage: Home market effect in a multi-region world," *Research in Economics*, 2017, *71* (4), 740–758.

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

Monderer, Dov and Lloyd S. Shapley, "Potential games," *Games and Economic Behavior*, 1996, *14* (1), 124–143.

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

— **, —, and Takatoshi Tabuchi**, "Urban structures with forward and backward linkages," Technical Report, Center for Research in Economic Analysis, University of Luxembourg 2017.

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

Negishi, Takashi, "Welfare economics and existence of an equilibrium for a competitive economy," *Metroeconomica*, 1960, *12* (2-3), 92–97.

Ogawa, Hideaki and Masahisa Fujita, "Equilibrium land use patterns in a nonmonocentric city," *Journal of Regional Science*, 1980, *20* (4), 455–475.

— **and** — , "Nonmonocentric urban configurations in a two-dimensional space," *Environment and Planning A*, 1989, *21* (3), 363–374.

Osawa, Minoru, Takashi Akamatsu, 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.

Owens, Raymond III, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte, "Rethinking Detroit," *American Economic Journal: Economic Policy*, 2019, (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 Jacques-François Thisse, "Agglomeration as spatial interdependence between firms and households," *Journal of Economic Theory*, 1985, *37* (1), 19–31.

— **and Terrence R. Smith**, "Agglomeration as local instability of spatially uniform steady-states," *Econometrica*, 1983, *51* (4), 1109–1119.

Rabin, Julien, Julie Delon, and Yann Gousseau, "Transportation distances on the circle," *Journal of Mathematical Imaging and Vision*, 2011, *41* (1-2), 147.

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

Rockafellar, R. Tyrrell, *Convex Analysis*, Princeton University Press, 1970.

Sandholm, William H., "Potential games with continuous player sets," *Journal of Economic Theory*, 2001, *97* (1), 81–108.

— , *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.

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

Tuy, Hoang, "Global minimization of a difference of two convex functions," in Bernard Cornet, V. Hien Nguyen, and Jean-Philippe Vial, eds., *Nonlinear Analysis and Optimization*, Vol. 30, Springer (Berlin, Heidelberg), 1987, pp. 150–182.

Wallace, Chris and H. Peyton Young, "Chapter 6 - Stochastic Evolutionary Game Dynamics," in H. Peyton Young and Shmuel Zamir, eds., *Handbook of Game Theory with Economic Applications*, Vol. 4, Elsevier, 2015, pp. 327 – 380.