Journal of Economic Dynamics & Control 42 (2014) 143–174

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Spatial externalities and agglomeration in a competitive industry ☆

William A. Brock [a,b,1], Anastasios Xepapadeas [c,*], Athanasios N. Yannacopoulos [d]

[a] Department of Economics, University of Wisconsin – Madison 1180 Observatory Drive, Madison, WI 53706, USA
[b] Department of Economics, University of Missouri – Columbia, 909 University Avenue, 118 Professional Building, Columbia, USA
[c] Department of International and European Economic Studies, Athens University of Economics and Business, Patission 76, Athens 10434, Greece
[d] Department of Statistics, Athens University of Economics and Business, Patission 76, Athens 10434, Greece

ARTICLE INFO

Article history:
Received 6 November 2013
Received in revised form
23 March 2014
Accepted 25 March 2014
Available online 4 April 2014

JEL classification:
D21
R3
C61

Keywords:
Competitive equilibrium
Social optimum
Spatial externality
Endogenous agglomerations

ABSTRACT

We introduce spatial spillovers as an externality in the production function of competitive firms operating within a finite spatial domain under adjustment costs. Spillovers may attenuate with distance and the overall externality could contain positive and negative components with the overall effect being positive. We show that when the spatial externality is not internalized by firms, spatial agglomerations may emerge endogenously in a competitive equilibrium. The result does not require increasing returns at the private or the social level, increasing marginal productivity of private capital with respect to the externality, or location advantages. In fact agglomerations may emerge with decreasing returns to scale, declining marginal productivity of private capital with respect to the externality, and no location advantage. The result depends on the interactions between the structures of production technology and spatial effects as shown in the paper. No agglomerations emerge at the social optimum when spillovers are internalized and diminishing returns both from the private and the social point of view prevail. Numerical experiments with Cobb–Douglas and CES technologies and an isoelastic demand confirm our theoretical predictions.

1. Introduction

A central result in the investment theory of the firm (Scheinkman, 1978) states that in a perfect foresight competitive equilibrium where firms take the price function as given and face convex adjustment cost in net investment, each firm's capital stock converges to a unique steady state which is independent of initial conditions. When firms are identical, all firms will converge in the long run to the same stock of capital.

In this paper we examine whether in a perfect foresight equilibrium for a competitive industry operating in a finite spatial domain with spatial interactions among firms, identical firms will end up with the same capital stock in the long run, or whether agglomeration emerges. Spatial interactions among firms are expressed as a spatial externality which in general

☆ This research has been co-financed by the European Union (European Social Fund ESF) and Greek National Funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) – Research Funding Program: “ARISTEIA Athens University of Economics and Business – Spatiotemporal Dynamics in Economics”.

* Corresponding author.
E-mail addresses: wbrock@ssc.wisc.edu (W.A. Brock), xepapad@aueb.gr (A. Xepapadeas), ayannaco@aueb.gr (A.N. Yannacopoulos).

[1] W.A. Brock is grateful to the Vilas Trust for financial support and the Department of Economics, University of Missouri, Columbia.

http://dx.doi.org/10.1016/j.jedc.2014.03.010
0165-1889/

attenuates with distance. One way of interpreting spatial interactions is to consider them as knowledge spillover effects from one firm to another. Knowledge spillovers are regarded as a positive intra-industry Marshalian externality which is bounded in space, the main idea being that innovation and new productive knowledge flows more easily among agents which are located within the same area (e.g. Krugman, 1991; Feldman, 1999; Breschi and Lissoni, 2001). Thus proximity is important in characterizing spatial spillovers (Baldwin and Martin, 2004; Breinlich et al., 2013). We incorporate knowledge spillovers by interpreting the capital stock of each firm in a broad sense to include knowledge along with physical capital (e.g. Romer, 1986). Following Quah (2002) we assume that the effect of capital on each firm's output, at any given point in time, does not depend just on the accumulated stock by the firm up to this time, but on capital accumulated in nearby locations by other firms. Thus the spatial externality takes the form of a Romer (1986) externality where, by keeping all other factors in fixed supply, output is determined by own capital stock and by an appropriately defined aggregate of capital stocks of firms across the spatial domain. The capital stock aggregate is determined by a distance-response function[2] that measures the strength of the spatial spillover on the output of a firm in a certain location associated with the capital stock accumulated by a firm in another location.

A positive distance-response function that attenuates with distance can be interpreted as reflecting knowledge spillovers. A distance-response which is negative indicates a negative externality such as generalized congestion effects. Thus, by combining a distance-response function, centripetal and centrifugal responses can be introduced. These forces are localized in the sense that their strength – positive or negative – diminishes with distance.[3]

Our purpose is to study whether optimal investment policy by forward-looking competitive firms combined with localized spatial spillovers generated from accumulated investment induces endogenous agglomerations and spatial clustering of firms.

It is known that spatial clusters may appear with localized knowledge spillovers when there are increasing returns. In this case the increasing returns activity concentrates to one location (e.g. Grossman and Helpman, 1991). Actually increasing returns underlie the generation of centripetal forces that favor cumulative causation and thus spatial clustering (e.g. Nocco, 2005). In our model the production function of each firm exhibits diminishing marginal productivity with respect to own capital for any fixed value of the spatial externality. To put it differently, private returns to capital are diminishing. The production function is strictly concave with respect to own capital and the spatial externality. That is, there are diminishing returns with respect to the spatial externality, for fixed levels of own capital. However, increasing social returns, in the sense of Romer (1986), are possible.

Our main result indicates that when diminishing returns from both the private and the social point of view prevail, then endogenous agglomeration may emerge at a perfect foresight rational expectations competitive equilibrium (PF-RECE). This agglomeration result does not require increasing returns at the private or the social level, increasing marginal productivity of private capital with respect to the externality, or location advantages.[4] In fact agglomerations may emerge with decreasing returns to scale, declining marginal productivity of private capital with respect to the externality, and no location advantage. The result depends on the interactions between the structures of production technology and spatial effects. The emergence of agglomeration may lead to a long-run steady state for the competitive industry where the distribution of capital stocks and outputs across space is not uniform On the other hand, at a social optimum (SO) where a planner fully endogenizes spatial spillovers, agglomerations do not emerge and all firms converge to the same stock of capital irrespective of location. The possibility of a potential agglomeration at a PF-RECE is related to the incomplete internalization of the spatial externality by optimizing firms and the structures of the production technology and spatial interactions, while the “no agglomerations” result at the SO stems from the full internalization of the spatial externality by a social planner and the strict concavity of the production function.

Our contribution is twofold. First, we provide a conceptual framework that explains dynamic endogenous emergence of spatial clustering in a competitive industry with optimizing forward-looking agents that do not require increasing returns to scale. Our model includes only the spatial externality and not other features of economic geography models such as transport costs, product differentiation or forward/backward linkages. We believe that this is a reasonable trade-off for being able to study agglomeration emergence in a fully dynamic optimizing model. Second, we show how convexity arguments and spectral theory can be used to study PF-RECE problems and SO problems in infinite horizon spatiotemporal economies, by properly decomposing the spatial and the temporal behavior. We thus provide valuable insights regarding the endogenous emergence (or not) of optimal agglomerations at a PF-RECE and the SO of a competitive industry.

2. Spatial externalities and adjustment costs

We consider an industry consisting of a large number of small firms with each firm located at point x of a one-dimensional bounded spatial domain $\mathcal{X}=[-L,L]$.[5] We further assume that $\mathcal{X}$ is discretized, i.e., it is divided into N cells or

[2] See Papageorgiou and Smith (1983) for an early use of distance-response functions.

[3] This is consistent with Prager and Thisse's second law of geography that states that what happens close to us is more important than what happens far from us (Prager and Thisse, 2012).

[4] We assume that the spatial domain is a circle to avoid the creation of agglomeration by the boundary conditions at the edge of the domain.

[5] Most of our results can be extended to general domains of characteristics $\mathcal{X}\subset\mathbb{R}^d$, $d\geq 1$.

intervals I_i, $i=1,\dots,N$, such that $\mathcal{X}=\cup_{i=1}^N I_i$. To save space we will denote by $\mathcal{N}:=\{1,2,\dots,N\}$ and use the compact notation $i\in\mathcal{N}$ in lieu of $i=1,\dots,N$.

At time $t\in\mathbb{R}_+$ and location $x\in\mathcal{X}$, each firm produces a single homogenous output $y(t,x)$. To simplify the model we assume that the output is uniform within each cell, i.e. $y(t,x)=y_i(t)$ for every $x\in I_i$, so that the state of the system at time t is given by a vector $y(t)=(y_1(t),\dots,y_N(t))\in\mathbb{R}^N$. Local output $y(t,x)$ is produced according to the production function $f:\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}_+$; $y(t,x)=f(k(t,x),K(t,x))$, which is strictly increasing in both arguments and sufficiently smooth with $\partial^2 f/\partial k^2, \partial^2 f/\partial K^2<0$. The arguments are: (i) broadly defined local capital stock $k(t,x)$ which includes knowledge that cannot be patented in full, and (ii) a spatial aggregate $K(t,x)$ of the broadly defined local capital stocks $K(t,x)$ which incorporates spatial externalities.[6] Strict monotonicity implies that the spatial externality acts as a productive input, i.e. it is a positive externality, while the assumption on the second derivatives implies that marginal productivities with respect to own capital and the spatial externality are diminishing. Thus for any fixed K the marginal productivity of capital from the private point of view is declining. Similarly for output, we assume that the inputs are uniform within each cell, so that $k(t,x)$ is replaced by a vector $k(t)=(k_1(t),\dots,k_N(t))\in\mathbb{R}^N$, and similarly $K(t,x)$ is replaced by a vector $K(t)=(K_1(t),\dots,K_N(t))\in\mathbb{R}^N$. Therefore, the production at time t and at cell i is given by $y_i(t)=f(k_i(t),K_i(t))$.

The spatial externality $K(t)$ plays the role of a productivity variable in a production function. The basic assumption is that the externality at time t and spatial point i is a weighted average of the broadly defined capital stocks at neighboring sites with weights declining with distance.[7] The weights determine the distance-response function. Thus local capital stock at each point j contributes to the total spatial spillover at site i according to a distance-response or weight function w_{ij}, and the total externality at location i is

$$K_i(t)=\sum_{j=1}^N w_{ij}k_j(t).$$

We also use the alternative compact notation $K=Wk$ where $W=(w_{ij})$, $i,j=1,\dots,N$, is an $\mathbb{R}^{N\times N}$ matrix. The rows of matrix W are called the kernel associated with location i, or simply the kernel. If $w_{ij}=0$ for a pair (i,j) that means that location j does not contribute at all to the total spillover at location i. Since the second law of geography suggests that distance is fundamental in the determination of spatial effects, we write $w_{ij}=\overline{w}(|i-j|)$ for some function $\overline{w}$, indicating that spatial impacts depend on distance and not on specific location. Note that the matrix W defines the connectivity of the "spatial network"[8] where the connectivity of sites 1 and N is related to the choice of boundary conditions. In order to eliminate the possibility of agglomeration creation by the edges of the one-dimensional spatial domain $[-L,L]$, periodic boundary conditions are imposed so that we consider the network as situated on a circle. Then site 1 interacts with site N that is now considered as its neighbor. We wish to emphasize that our analysis is valid for a general choice of networks, i.e., for a general choice of matrix W. See Fig. 1 for an illustration of the network modelling the spatial economy. However, the choice of a circle or a torus, for a two-dimensional spatial domain, eliminates the impact of boundary conditions on the formation of spatial patterns, which means that if agglomerations emerge they are emerging endogenously and not because of boundary conditions.

An important class of networks (equivalently connectivity matrices W) are those that satisfy the condition $\sum_j w_{ij}=\overline{w}$, independent of the choice of i. We call such a coupling, diffusive type coupling. Thus diffusive coupling means that if the capital stock is the same at all locations, say $\overline{k}$, then the spatial externality will be $\overline{wk}$ which is the same for all locations. Since our spatial domain is a circle, this assumption ensures that any agglomeration emergence is endogenous and not the result of boundary conditions or a location advantage for a specific site.

The spatial externality $K_i(t)$ will have different interpretations in different contexts. If $K_i(t)$ embodies a type of knowledge which is produced proportionately to capital usage, it is natural to assume that the distance-response function w_{ij}, considered as a function of $\zeta=i-j$ which expresses a positive externality, is single peaked and bell-shaped with a maximum at $\zeta=0$, and of possibly sufficiently fast decay to zero for sufficiently large $|\zeta|$. If the spatial externality $K_i(t)$ embodies damages to production at (t,i) from usage of capital at (t,j), then a composite externality can be created with $w_{ij}=w_{ij}^1+w_{ij}^2$.[9] If w_{ij}^1 is a bell-shaped positive externality and w_{ij}^2 is an inverted bell-shaped negative externality, which also decays to zero as $|\zeta|$ becomes large, then non-monotonic shapes of w_{ij} are possible with, for example, a single peak at $\zeta=0$ and two local minima located symmetrically around $\zeta=0$, with negative values indicating negative externality to production at i from usage of capital at j. Examples of such kernels are given in Section 6.1 and in Figs. 2 and 8. We assume throughout the paper that in the case of either single or composite externality, the overall effect is positive, that is, $\overline{w}>0$. A production function incorporating these externalities could be considered as a spatial version of a neoclassical production function with Romer/Lucas externalities modelled by geographical spillovers given by a Krugman (see e.g., Krugman, 1996) or Chincarini and Asherie type specification (see e.g. Chincarini and Asherie, 2008). Throughout the paper we also assume that $w_{ii}>0$, so that own effects and aggregate effects from the spatial externality are positive.

6 To simplify the exposition we assume that all other factors of production are in fixed supply.

7 See for example Lucas (2001) or Lucas and Rossi-Hansberg (2002) for a similar type of externality where the productivity variable is defined as the average of employment at neighboring sites.

8 If, for example, $w_{ij}=\delta_{j,i+1}+\delta_{j,i-1}-2\delta_{j,i}$, where $\delta_{i,j}$ is the Kronecker delta ($\delta_{ii}=1$, $\delta_{ij}=0$, for all $i\neq j$), we have a linear connectivity of the knowledge network, according to which site i interacts only with sites $i+1$ and $i-1$.

9 See also Papageorgiou and Smith (1983) for more details regarding composite distance-response functions.

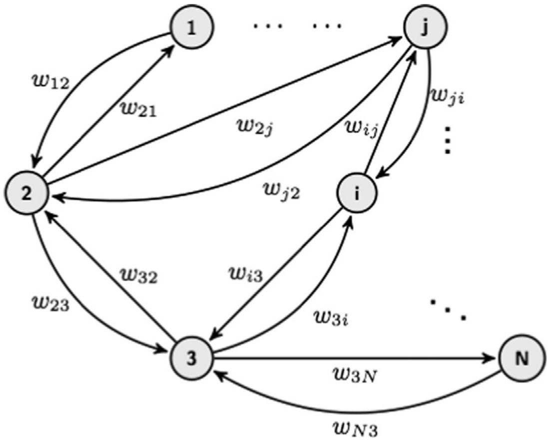

Fig. 1. An economy with spatial connections.

Net investment in each location i is given by the derivative with respect to time, k', of the vector valued function $k: \mathbb{R}_+ \to \mathbb{R}^{\cap N}$. The firm faces the cost of changing the capital stock, which is a function of net investment k'. This adjustment cost at time t and location i is expressed by a quadratic adjustment function $C_i(t) = (\alpha/2)(k_i'(t))^2$, $\alpha > 0$. Capital stock depreciates at the same rate η in all locations.

The output of the firms is sold at a market price determined by a demand function $D: \mathbb{R}_+ \to \mathbb{R}_+$

$$p(t) = D(Q) = D(Q(k,K)), \quad D > 0, D' \le 0 \tag{1}$$

$$Q{:=}Q(k,K) = \sum_{i=1}^{N} f(k_i(t), K_i(t)). \tag{2}$$

The k and K dependence is stated explicitly to emphasize that D can be understood as a functional $D: \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$. That is, given a vector k of capital stocks across locations and a kernel $W = (w_{ij})$, $i,j = 1, \dots, N$, we obtain $K{=}Wk$, and calculate the total output Q that determines p.

We are assuming a large number of identical small firms in the spatial domain $\mathcal{X}$ and identical agents at each cell I_i of $\mathcal{X}$. The nature of a positive spatial externality can be described in the following way. Firms produce a specific output along with knowledge related to the production processes which may increase productivity. Small firms and agents in each cell take the actions of the other firms at each location as given and beyond each firm's and agent's control within each location as well as across each location. Not all in-house knowledge is patented, so the public knowledge generated by the firms is combined together and creates an external knowledge aggregate that helps producers to increase their productivity. From the point of view of a certain location, the contribution of other locations to this knowledge aggregate attenuates with distance. The agents are however myopic and when they accumulate new knowledge they do not take into account their own contribution to this aggregate, but consider the aggregate as fixed and beyond their own control. This is a positive spatial externality.[10] A social planner who is not myopic realizes however that knowledge accumulation in each firm increases the knowledge aggregate, and benefits the productivity of each firm in the spatial domain.

Therefore, small agents in cell i optimize without taking into account own contribution as well as other agents' contributions within the cell and across locations on the aggregate externality K_i, taking thus the aggregate level of K_i affecting their cell as given. The assumption that each agent treats K_i as given could be rationalized in a model with a continuum of agents. Here we make the usual approximation of a large but finite number of small agents. Assuming furthermore perfect capital markets and that the unit price of capital is q, independent of time, the objective of a firm located at $i \in \mathcal{N}$ is to maximize the present value of profits by considering spatial spillovers as exogenous $K_i = K_i^e$. The firm's problem can be written as

$$\max_{k_i'} \int_0^\infty e^{-rt} \left[p(t) f\left(k_i, K_i^e\right) - \frac{\alpha}{2}(k_i')^2 - q\left(k_i' + \eta k_i\right) \right] dt \tag{3}$$

$$k_i(0) = k_{i0}, \quad k_i(t) \ge 0, \quad i \in \mathcal{N}. \tag{4}$$

In this set-up we define the industry equilibrium and derive conditions under which endogenous spatial clustering could emerge.

[10] A negative externality and a composite externality can be described in a similar way.

3. Industry equilibrium and social optimum

Following Lucas Jr. and Prescott (1971), Brock (1974), and Brock and Scheinkman (1977), we define a perfect foresight rational expectation competitive equilibrium (PF-RECE) as the price function $p(t)$ given by (1) where $k_i(t)$ solves (3) for all $i \in \mathcal{N}$ with optimality conditions evaluated at $K^e = Wk$. If the price path $p(t)$ is predicted by the competitive firms, this path will result in an aggregate output Q over the whole spatial domain such that the market is cleared at each t by $p(t)$.

The long-run properties of the industry equilibrium can be obtained by exploiting the concept of maximization of consumer surplus, that is the area under the demand curve (Lucas Jr. and Prescott, 1971; Brock, 1974; Brock and Scheinkman, 1977), which in the present model can be defined by

$$S(k,K) = \int_0^{Q(k,K)} D(s)\, ds. \tag{5}$$

Using the concept of consumer surplus, we consider two optimization problems leading to two different concepts of equilibrium:

(A) The problem of maximizing consumer surplus when firms regard knowledge spillovers as exogenous, that is when they do not internalize the spatial externality and they set $K_i(t) = K^e$. This problem is defined as

$$\max_{k'} \int_0^\infty e^{-rt} \left\{ S(k,K^e) - \sum_{i=1}^{N} \left[\frac{\alpha}{2}(k_i')^2 + q(k_i' + \eta k_i) \right] \right\} dt. \tag{6}$$

The solution to this problem determines the PF-RECE.

(B) The problem of maximizing consumer surplus when a social planner fully internalizes the spatial externality, which is defined as

$$\max_{k'} \int_0^\infty e^{-rt} \left\{ S(k,Wk) - \sum_{i=1}^{N} \left[\frac{\alpha}{2}(k_i')^2 + q(k_i' + \eta k_i) \right] \right\} dt. \tag{7}$$

The solution to this problem determines the SO.

The Euler equations for these two problems can be obtained in a straightforward manner, using the Pontryagin maximum principle. For problem (6), by setting $k_i'(t) = u_i(t)$, the current value Hamiltonian is

$$\mathcal{H}(k,u,\mu) = S(k,K^e) - \sum_{i=1}^{N} \left[\frac{\alpha}{2}(u_i)^2 + q(u_i + \eta k_i) \right] + \sum_{i=1}^{N} \mu_i u_i \tag{8}$$

with optimality conditions

$$u_i = \frac{\mu_i - q}{\alpha} = k_i' \tag{9}$$

$$\mu_i' = r\mu_i + q\eta - \frac{\partial}{\partial k_i} S(k,K^e) \tag{10}$$

and transversality conditions at infinity

$$\lim_{t\to\infty} e^{-rt} \sum_{i=1}^{N} \mu_i(t) k_i(t) = 0. \tag{11}$$

Using $k_i'' = \mu_i'/\alpha$ from (9) and substituting into (10) we obtain the Euler equations:

$$k_i'' - rk_i' + \frac{1}{\alpha}\left[\frac{\partial S(k,K^e)}{\partial k_i} - q(r+\eta) \right] = 0, \quad i \in \mathcal{N}. \tag{12}$$

Thus each firm treats the spatial externality K^e as parametric when deciding about its investment decisions. However the actions of all firms generate the "actual" value of the realized spatial externality which is Wk. Equilibrium requires that the spatial externality be consistent with the level that is assumed when firms make decisions about k. Thus in a PF-RECE, $K^e = Wk$ and the Euler equation that characterizes this equilibrium becomes

$$k_i'' - rk_i' + \frac{1}{\alpha}\left[\frac{\partial}{\partial k_i} S(k,K^e)|_{K^e = Wk} - q(r+\eta) \right] = 0, \quad i \in \mathcal{N}, \tag{13}$$

where the notation $(\partial/\partial k_i)S(k,K^e)|_{K^e = Wk}$ means that we first take the gradient of $S(k,K^e)$ with respect to k, treating K^e as fixed, and then substitute $K^e = Wk$ into the resulting function to determine the PF-RECE. Eq. (13) can be expressed in a more explicit form as

$$k_i'' - rk_i' + \frac{1}{\alpha}\left[D(Q(k,Wk)) f_k\left(k_i, \sum_r w_{ir} k_r \right) - q(r+\eta) \right] = 0, \quad i \in \mathcal{N}, \tag{14}$$

where $Q(k,Wk)=\sum_i f(k_i,\sum_r w_{ir}k_r)$. By f_k we denote the partial derivative of the production function f with respect to the first variable, and we employ the notation $(Wk)_i=\sum_r w_{ir}k_r$ for the i-th component of the vector Wk.

For the SO, problem (7), the corresponding Euler equation, is

$$k_i''-rk_i'+\frac{1}{\alpha}\left[\frac{\partial}{\partial k_i}S(k,Wk)-q(r+\eta)\right]=0,\quad i\in\mathcal{N}. \tag{15}$$

Eq. (15) can be expressed more explicitly as

$$k_i''-rk_i'+\frac{1}{\alpha}\left[D(Q(k,Wk))\left[f_k\left(k_i,\sum_r w_{ir}k_r\right)\quad+\sum_l w_{li}f_K\left(k_l,\sum_r w_{lr}k_r\right)\right]-q(r+\eta)\right]=0,\quad i\in\mathcal{N}. \tag{16}$$

By f_K we denote the partial derivative of the production function with respect to the second argument. This leads to the following definition:

Definition 1. We call the solution $k:\mathbb{R}_+\to\mathbb{R}^N$ of (13), with $K^e=Wk$, a PF-RECE and the solution of (15) an SO.

Note that in the SO, $(\partial/\partial k_i)S(k,Wk)$ are the components of the true gradient of the consumer surplus function S, treated as a function of k only, i.e., the true gradient of the function $S(k,Wk)$. This is in contrast to what happens for the PF-RECE where $(\partial/\partial k_i)S(k,K^e)|_{K^e=Wk}$ no longer corresponds to the components of a "true" gradient of a function. This remark will play a very important role in the qualitative long-term behavior of the two systems, and leads to important differences between them.

We close this section by noting that both the rational expectations and the SO Euler equations may be expressed in a single form, using the parameter σ, which takes the value $\sigma=0$ if we are studying the PF-RECE and the value $\sigma=1$ if we are studying the SO. The Euler equation thus takes the form

$$k_i''-rk_i'+\frac{1}{\alpha}\left[D(Q(k,Wk))\left[f_k\left(k_i,\sum_r w_{ir}k_r\right)\quad+\sigma\sum_l w_{li}f_K\left(k_l,\sum_r w_{lr}k_r\right)\right]-q(r+\eta)\right]=0,\quad i\in\mathcal{N}. \tag{17}$$

4. The steady state of the social optimum: a global result

The Euler equations characterizing the SO and the PF-RECE can be used to explore the emergence of agglomerations in the competitive industry. First we provide a global result about the possibility of agglomerations as a long-run outcome at the SO when the spatial externality is fully internalized.

Assumption 1. D is a strictly decreasing function and the production function $f(k,K)$ is a strictly concave function of (k,K).

Theorem 1. *Let Assumption1 hold. If the system of equations*

$$\frac{\partial}{\partial k_i}S(k,Wk)-q(r+\eta)=0,\quad i\in\mathcal{N} \tag{18}$$

admits the spatially uniform, or flat, $k_1=\cdots=k_N=\bar{k}$ solution, then there is no emergence of spatial agglomerations in the long-run equilibrium for the SO.

Proof. The function $S(x)=\int_0^x D(s)\,ds$ is strictly concave as the integral of a strictly decreasing function (see Lemma 1 in the Appendix, Section A.1) and, by the properties of the production function, the function $S(k,Wk)$ a strictly concave function of k. Therefore, function $\overline{S}(k){:=}S(k,Wk)-q(r+\eta)k$ is strictly concave. The Euler equation can be written as

$$k''-rk'=-\frac{1}{\alpha}\nabla\overline{S}, \tag{19}$$

and by the convexity of $-S$, the operator $-\nabla\overline{S}$ is a monotone operator on $\mathbb{R}^N$. By the results of Rouhani and Khatibzadeh (2009), any bounded solution of system (19) converges to a steady state which is a solution of (18). If the solution of (18) is unique then the long-run behavior of (19) is characterized by the unique solution of (18). However, note that the solution of (18) is recognized as the minimum of the function $-\overline{S}$, since at the steady state $\nabla\overline{S}=0$, which is unique by the strict convexity of $\overline{S}$. Therefore, if (18) admits a spatially uniform (or flat) solution, by uniqueness it may admit no other solution and the result follows. □

Thus, if there exists a steady state for the competitive industry at which all firms have the same capital stock, then no other steady state is possible. Moreover, this steady state is stable under the dynamics of the Euler equation and this eliminates the possibility of a steady state with spatially heterogeneous capital stock and hence the possibility of agglomerations at the SO. The result is an extension of Scheinkman's result (Scheinkman, 1978) in a spatial context and suggests that if the spatial externality is fully internalized, it cannot induce spatial clustering in a competitive industry with diminishing returns with respect to both own capital and the spatial externality. It should be noted that the result of Theorem 1 could have been shown by using Scheinkman's separable Hamiltonian approach (Scheinkman, 1978). The approach used here is more general in the sense that it does not require assumptions about the derivatives of the value function of the problem and provides insights for analyzing the PF-RECE, a case where the separable Hamiltonian approach cannot be applied.

It should also be noted that the special structure of the Euler equation for our particular formulation of the firm's problem, viz., $k''-rk'=-(1/\alpha)\nabla \overline{S}$, implies that at a steady state where $k''-rk'=0$ we must have a minimum of a convex function and this result is independent of r. Because of the particular way in which we model adjustment costs in the firm's problem, which produced a special structure to Eq. (19), the Euler equation admits this special structure. Thus convergence, in our model, to a unique steady state is independent of the discount rate. In general, for example, in the Intertemporal general equilibrium model of Bewley (1982), the discount rate needs to be sufficiently small to obtain convergence. In fact our model is a special case of Bewley's model, since we are dealing with only the firms' part of the general equilibrium problem, with the notable difference that our model includes an externality which is not present in Bewley's dynamic general equilibrium model. Bewley is able to use Pareto optimality to obtain his results. Pareto optimality holds in the SO problem but not in our PF-RECE problem. In our special case, the discount rate is relevant for determining the spatially homogeneous steady state towards which the system converges as the following proposition indicates.

To examine conditions under which such a spatially homogeneous or flat steady state at the SO exists, we make the following assumption:

Assumption 2. The coupling is of diffusive type, i.e. $\sum_j w_{ij}=\overline{w}$ for any $i \in \mathcal{N}$, and the production function is homogeneous of degree γ.

The first part of the assumption combined with the assumption that our spatial domain is a circle eliminates the possibility that agglomerations may emerge as a result of exogenous factors such as the shape of the spatial domain and the location advantage of one or more sites. The second part is a common assumption that simplifies the problem and allows us to determine solutions.

Theorem 1 refers to the convergence to a long run spatially uniform steady state. The following proposition provides a general result about the existence and the way to determine such a spatially homogeneous steady state within the structure of our model.

Proposition 1. *Let Assumptions 1 and 2 hold. If the scalar algebraic equation*

$$\gamma N^{(1-\gamma)/\gamma}\rho^{1/\gamma}D(s)s^{(\gamma-1)/\gamma}-q(r+\eta)=0, \quad \rho:=f(1,\overline{w}) \tag{20}$$

admits a solution $s^* \in \mathbb{R}_+$, then no agglomeration patterns will appear in the long-run equilibrium for the SO and the industry relaxes to a spatially homogeneous (flat) state $k_1=\cdots=k_N=k^*=(s^*/N\rho)^{1/\gamma}$.

Proof. The steady state is given by the solution of the system of equations

$$D(Q(k,Wk))\left(f_k\left(k_i,\sum_j w_{ij}k_j\right)+\sum_r w_{ri}f_K\left(k_i,\sum_j w_{rj}k_j\right)\right)-q(r+\eta)=0, \tag{21}$$

$i \in \mathcal{N}$, which for a spatially uniform solution $k_1=\cdots=k_N=k^*$ and using Assumption 2 reduces to a single algebraic equation, which is equivalent to (20), in terms of the variable $s=N\rho(k^*)^\gamma$. Then using Theorem 1 we obtain the stated result. □

For a standard Cobb–Douglas production function with $\gamma=\gamma_1+\gamma_2<1$, the spatially homogeneous steady state can easily be obtained by following the calculations of the proof to this proposition. Thus in a competitive industry with identical firms with Cobb–Douglas technology which is strictly concave in own capital and the spatial externality, and no location advantage or effects from the boundaries of the spatial domain, no agglomeration will occur if the spatial externality is fully internalized.

5. Agglomeration emergence

Our "no agglomeration" result holds for the SO ($\sigma=1$) under the strict concavity Assumption 1 of the production function. This means agglomerations do not emerge when the spatial externalities are fully internalized, the production function is strictly concave and the demand function is strictly decreasing. Therefore, agglomerations may emerge if any of the above assumptions is not satisfied.

When the spatial externality is not fully internalized in a PF-RECE, the no agglomeration result is no longer sustained,[11] and a result similar to Theorem 1 cannot be obtained for the PF-RECE. Thus, even if a flat steady state exists for the PF-RECE,

[11] To see this, note that in terms of the Euler equation characterizing the PF-RECE, the term $(\partial/\partial k_i)S(k,K^e)|_{K^e=Wk}$ is no longer a gradient, which means that a firm does not take into account the impact of its investment policy on the spatial externality.

we cannot exclude the existence of other spatially heterogeneous steady states. Spatial heterogeneity, however, means agglomerations.

In this section we provide explicit conditions under which agglomerations may occur either for the PF-RECE ($\sigma=0$) or for the SO ($\sigma=1$) if Assumption 1 does not hold.[12] We examine the potential emergence of spatial agglomerations by perturbing a spatially homogeneous, or flat, steady state in a fashion which is similar (but different in mechanism) to the celebrated Turing instability (Turing, 1952).[13] The instability concept presented in Theorem 2 is local saddle point instability of the steady state under scrutiny. Checking for this kind of local saddle point instability involves checking whether more than N eigenvalues have positive real parts of the linearization of the Euler equation (17) around a flat steady state.

To clarify the exposition we use the following definitions:

Definition 2. Define the real numbers

$$\rho := f(1,\overline{w}), \quad \rho_k := f_k(1,\overline{w}), \quad \rho_K := f_K(1,\overline{w}),$$
$$\rho_{kk} := f_{kk}(1,\overline{w}), \quad \rho_{kK} := f_{kK}(1,\overline{w}), \quad \rho_{KK} := f_{KK}(1,\overline{w}).$$

ρ denotes the output of a firm at a flat steady state where capital stock is normalized to one, ρ_k, ρ_K are the corresponding marginal products, ρ_{kk}, ρ_{KK} are the slopes of marginal products, while ρ_{kK} denotes the shift in the marginal product of capital from a small change in the spatial externality. All derivatives are evaluated at the flat normalized steady state.

Remark 1. For the Cobb–Douglas production function $f(k,K) = k^{\gamma_1}K^{\gamma_2}$, $\{\gamma_1,\gamma_2\} \in (0,1)$ these quantities become

$$\rho = \overline{w}^{\gamma_2}, \quad \rho_k = \gamma_1 \overline{w}^{\gamma_2}, \quad \rho_K = \gamma_2 \overline{w}^{\gamma_2 - 1},$$
$$\rho_{kk} = \gamma_1(\gamma_1 - 1)\overline{w}^{\gamma_2}, \quad \rho_{kK} : \gamma_1\gamma_2 \overline{w}^{\gamma_2 - 1}, \quad \rho_{KK} = \gamma_2(\gamma_2 - 1)\overline{w}^{\gamma_2 - 2},$$

while for a CES production function $f(k,K) = A[\beta k^{-\theta} + (1-\beta)K^{-\theta}]^{-\gamma/\theta}$, $A>0$, $\theta \geq -1, 0<\beta<1$, with degree of homogeneity γ, and elasticity of substitution $\hat{\sigma} = 1/(1+\theta)$ we have

$$\rho = A[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta}, \quad \rho_k = A\gamma\beta[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 1},$$

$$\rho_K = A\gamma(1-\beta)\overline{w}^{-\theta-1}[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 1},$$

$$\rho_{kk} = -A\gamma\beta(1+\theta)[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 1} + \left(1 + \frac{\nu}{\theta}\right)\theta A\gamma\beta^2[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 2},$$

$$\rho_{KK} = A\gamma(1-\beta)(1+\theta)\overline{w}^{-\theta-2}[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 1} + \left(1 + \frac{\nu}{\theta}\right)\theta A\gamma(1-\beta)^2\overline{w}^{-2\theta-2}[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 2},$$

$$\rho_{kK} = (\gamma + \theta)A(1-\beta)\beta\overline{w}^{-\theta-1}[\beta + (1-\beta)\overline{w}^{-\theta}]^{-\gamma/\theta - 2}.$$

Remark 2. It should be noticed that for the Cobb–Douglas production function $\rho_{kK} > 0$ or equivalently $f_{kK} > 0$, so that an increase in the externality will shift the marginal product of private capital upwards. The CES production function allows more flexibility with respect to the relationship between the marginal product of private capital and the externality as reflected by the cross partial derivative f_{kK}. In the CES case,

$$f_{kK} = (\gamma + \theta)(1-\beta)\beta\gamma A k^{-(1+\theta)}K^{-(1+\theta)}[\beta k^{-\theta} + (1-\beta)K^{-\theta}]^{-2-\gamma/\theta}, \tag{22}$$

with constant or increasing returns to scale and $\theta > -1$, $\gamma + \theta > 0$. Thus, as in the Cobb–Douglas case, an increase in the externality will shift the marginal product of private capital upwards. For a negative cross partial, meaning that an increase in the externality will shift the marginal product of private capital downwards, or $f_{kK} < 0$, decreasing returns to scale are required so that $\gamma + \theta < 0$. Since $\hat{\sigma} = 1/(1+\theta)$, the condition for a negative cross partial can be expressed in terms of the elasticity of substitution between k and K as $\hat{\sigma} > 1/(1-\gamma)$. Thus a negative cross partial requires decreasing returns to scale and $\hat{\sigma} > 1$. If $\gamma + \theta < 0$ then $\rho_{kK} < 0$ for the CES case. For a general production technology with two inputs $f(k,K)$ which is homogeneous of degree γ, it can be seen, by applying Euler's theorem to the fist order derivatives which are homogeneous of degree $\gamma - 1$, that

$$(\gamma - 1)\frac{f_k + f_K}{k + K} = f_{kK} + \frac{f_{kk}k + f_{KK}K}{k + K}. \tag{23}$$

Thus $f_{kK} < 0$ requires $\gamma < 1$.

[12] To ease the exposition, if $\sigma=0$ the stated results correspond to the PF-RECE, whereas if $\sigma=1$ the stated results correspond to the SO. When a quantity carries the subscript σ, this implies that it depends on the value of σ chosen, i.e., that it differs between the PF-RECE and the SO.

[13] The use of conditions under which a spatially homogenous steady state becomes unstable to spatially heterogeneous perturbation in order to establish the emergence of agglomerations has been used in spatial economics. See for example Papageorgiou and Smith (1983) or Krugman (1996). Our difference with this literature is that the perturbed steady state in our model is the outcome of actions of forward-looking optimizing agents.

Definition 3. Define the stability matrix

$$T_\sigma = C_1 I + C_2 \mathbf{1} + C_3 W + C_4 W^2, \quad \sigma = 0 \text{ or } 1 \tag{24}$$

where

$$C_1 := \frac{1}{\alpha} D\left(N\rho \overline{k}_\sigma^\gamma\right)\rho_{kk}\overline{k}_\sigma^{\gamma-2},$$
$$C_2 := \frac{1}{\alpha}(\rho_k + \overline{w}\rho_K)(\rho_k + \sigma\overline{w}\rho_K)\overline{k}_\sigma^{2(\gamma-1)} D'\left(\rho N \overline{k}_\sigma^\gamma\right),$$
$$C_3 := \frac{1}{\alpha}(1+\sigma)\rho_{kK} D\left(N\rho\overline{k}_\sigma^\gamma\right)\overline{k}_\sigma^{\gamma-2},$$
$$C_4 := \frac{1}{\alpha}\sigma\rho_{KK} D\left(N\rho\overline{k}_\sigma^\gamma\right)\overline{k}_\sigma^{\gamma-2},$$

and $\mathbf{1}$ is the $N \times N$ matrix whose every entry is equal to 1.

Definition 4. Let $\{\phi_\ell, \lambda_\ell\}$, $\ell = 1, \dots, N$, be the corresponding eigenvectors and eigenvalues of the stability matrix T_σ (see (24) in Definition 3), and define the sets

$$\mathcal{A} := \left\{ \ell \in \mathcal{N} : \frac{r^2}{4} < \lambda_\ell \right\},$$
$$\mathcal{B} := \left\{ \ell \in \mathcal{N} : 0 < \lambda_\ell < \frac{r^2}{4} \right\}.$$

Theorem 2. *Let Assumption 2 hold, let T_σ be the stability matrix defined in Definition 3 and $\mathcal{A}$, $\mathcal{B}$ the sets defined in Definition 4.*

(a) *If the scalar algebraic equation*

$$\left(\frac{1}{\rho N}\right)^{(\gamma-1)/\gamma}(\rho_k + \sigma\overline{w}\rho_K) D(s_\sigma) s_\sigma^{(\gamma-1)/\gamma} - r(q+\eta) = 0 \tag{25}$$

admits a unique solution $s_\sigma^* \in \mathbb{R}_+$, then a spatially homogeneous steady state $\overline{k}_\sigma = (s_\sigma^*/N\rho)^{1/\gamma}$ uniquely exists.

(b) *The following results hold concerning the linear stability of spatially homogeneous steady states*:

(i) *If $\mathcal{A} \neq \emptyset$, i.e., if T_σ has eigenvalues greater than $r^2/4$, then agglomerations may appear around the spatially homogeneous steady state $\overline{k}_\sigma$.*

(ii) *If $\mathcal{B} \neq \emptyset$, i.e., if T_σ has positive eigenvalues but less than $r^2/4$, then an agglomeration that fluctuates with respect to time may appear around the spatially homogeneous steady state $\overline{k}_\sigma$.*[14]

Eq. (25) is the steady-state equation for a flat steady state resulting from the Euler equation (17). If a spatially homogeneous steady state for PF-RECE ($\sigma = 0$) or the SO ($\sigma = 1$) exists, part (b) of the proposition presents the conditions under which it can be destabilized by spatial spillovers. Local instability of the spatially homogeneous steady state means that agglomerations may emerge.

Remark 3. While both regions $\mathcal{A}$ and $\mathcal{B}$ lead to linear instability of the flat steady state, we consider as a viable agglomeration pattern for the system only those patterns that correspond to region $\mathcal{B}$, for the following reason. Our system is a controlled system which is subject to a transversality condition at infinity. Only the patterns corresponding to region $\mathcal{B}$ satisfy the transversality condition, thus only these patterns are viable agglomeration patterns. We therefore consider as a condition for the occurrence of agglomeration the condition that at least one of the eigenvalues of the stability matrix T_σ lies in the interval $(0, r^2/4]$.

Remark 4. Since the instability is emerging as the optimal solution of the problem, we call this instability optimal spillover induced spatial instability. This type of instability is different from the celebrated Turing instability, or the instabilities identified in earlier models of economic geography (e.g. Papageorgiou and Smith, 1983, Krugman, 1996), because it is the result of forward-looking optimizing behavior.

Proof of Theorem 2. For the proof, we work with the compact formulation of the RE and SO model in a single equation (as in (17)) using the variable σ, where $\sigma = 0$ in the RE case and $\sigma = 1$ in the SO case. The equation of motion becomes

$$k_i'' - rk_i' + \frac{1}{\alpha}\left\{ D(Q(k, Wk))\left[f_1\left(k_i, \sum_r w_{ir}k_r\right) + \sigma\sum_\ell w_{\ell i} f_2\left(k_\ell, \sum_j w_{\ell j}k_j\right)\right] - q(r+\eta)\right\} = 0. \tag{26}$$

[14] This means that the solution will have the form $k_n(t), n \in \mathcal{N}$ such that, in general $k_n(t) \neq k_m(t)$ for $n \neq m$, but k_n are periodic functions of time (see also Eqs. (30) and (31) in the proof of the theorem).

We simplify the notation by using the definition

$$F_i := D(Q(k, Wk))\left[f_1\left(k_i, \sum_r w_{ir}k_r\right) + \sigma\sum_\ell w_{\ell i} f_2\left(k_\ell, \sum_j w_{\ell j}k_j\right)\right].$$

In the proof we employ the notation f_1 for f_k and f_2 for f_K.

We now perform the general calculation for the linearization of (26) around a homogeneous steady state $\overline{k}$. Note that $\overline{k}$ changes with σ, so we denote it as $\overline{k}_\sigma$. Consider then $\kappa = \overline{k}_\sigma + \epsilon k$ (meaning that $\kappa_i = \overline{k}_\sigma + \epsilon k_i$ for every i). We do the linearization of the three terms involved separately and in order to keep the flow of the text the detailed linearizations are presented in the Appendix, Section A.2

The linearization of F_i results in

$$F_i = A_0(B_0+C_0) + \epsilon(B_0+C_0)A_1\sum_\ell k_\ell + \epsilon A_0 B_{11} k_i + \epsilon A_0(B_{12}+C_{11})\sum_r w_{ir}k_r + \epsilon A_0 C_{12}\sum_\ell\sum_j w_{\ell i}w_{\ell j}k_j$$

where

$$A_0 = D(Nf(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)),$$
$$A_1 = D'(Nf(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma))(f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \overline{w}f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)),$$

$$B_0 = f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \quad C_0 = \sigma\overline{w}f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma),$$
$$B_{11} = f_{kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \quad C_{11} = \sigma f_{Kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma),$$
$$B_{12} = f_{kK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \quad C_{12} = \sigma f_{KK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma).$$

From the above calculations, we see that the homogeneous steady state $\overline{k}_\sigma$ will be a solution of the algebraic equation

$$A_0(B_0+C_0) - M = 0, \tag{27}$$

where $M = q(r+\eta)$. This yields

$$D(Nf(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma))(f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \sigma\overline{w}f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)) - M = 0.$$

Using the assumption that f is homogeneous with degree of homogeneity γ, and Definition 2 this is expressed as

$$(\rho_k + \sigma\overline{w}\rho_K)D(N\rho\overline{k}_\sigma^\gamma)\overline{k}_\sigma^{\gamma-1} - M = 0$$

which, when solved for $\overline{k}_\sigma$, provides the homogeneous steady state. Upon the change of variables $s_\sigma = N\rho\overline{k}_\sigma^\gamma$, the steady state equation becomes

$$\left(\frac{1}{\rho N}\right)^{(\gamma-1)/\gamma}(\rho_k + \sigma\overline{w}\rho_K))D(s_\sigma)s_\sigma^{(\gamma-1)/\gamma} - M = 0.$$

When $\sigma=0$ this yields the steady state for the RE case while when $\sigma=1$ this yields the steady state for the SO case.

Substituting into the equation we see that the linearized equation is

$$k'' - rk' + T_\sigma k = 0,$$

where

$$T_\sigma := \frac{1}{\alpha}\{A_0 B_{11} I + (B_0+C_0)A_1\mathbf{1} + A_0(B_{12}+C_{11})W + A_0 C_{12} W^2\}.$$

Using the homogeneity assumption for the production function, we may further simplify this matrix to the stability matrix

$$T_\sigma = C_1 I + C_2\mathbf{1} + C_3 W + C_4 W^2, \tag{28}$$

defined in 3 where $\overline{k}_\sigma$ is the solution of the steady-state equation (27).

Having obtained the linearized system, we study the evolution of a spatially nonhomogeneous perturbation of this homogeneous steady state. Consider a solution of (13), of the form $k_i = \overline{k}_\sigma + \epsilon p_i$, $i \in \mathcal{N}$, where ϵ is a small parameter. We substitute into (13) and linearize with respect to ϵ. The above results show that the vector $\mathfrak{p} = (p_1, \ldots, p_N)$ evolves according to the second order linear evolution equation

$$\mathfrak{p}'' - r\mathfrak{p}' + T_\sigma\mathfrak{p} = 0, \tag{29}$$

where T_σ is the matrix given in (28). Since W is a symmetric matrix, the same is true for the matrix T_σ, so by the spectral theorem there exists an orthonormal basis of $\mathbb{R}^N$ consisting of the eigenvectors of T_σ, each corresponding to real eigenvalues. Let $\{\phi_\ell, \lambda_\ell\}$, $\ell = 1, \ldots, N$, be the corresponding eigenvectors and eigenvalues. The general solution of (29) can be expressed as

$$\mathfrak{p}(t) = \sum_{\nu=1}^N q_\nu(t)\phi_\nu$$

so by substituting into (29), we obtain

$$\sum_{\nu=1}^{N} q_\nu''(t)\phi_\nu - r\sum_{\nu=1}^{N} q_\nu'(t)\phi_\nu + \sum_{\nu=1}^{N} q_\nu(t)\lambda_\nu\phi_\nu = 0,$$

and taking inner products with ϕ_ℓ, $\ell \in \mathcal{N}$ and using the orthogonality of the eigenvectors, $\langle \phi_\nu, \phi_\ell \rangle = \delta_{\nu,\ell}$ yields

$$q_\ell'' - rq_\ell' + \lambda_\ell q_\ell = 0, \quad \ell \in \mathcal{N}.$$

Now the system is decoupled. This implies that the general solution of (29) can be expressed as

$$\mathfrak{p}(t) = \sum_{\ell=1}^{N} (a_\ell \exp(s_\ell^+ t) + b_\ell \exp(s_\ell^- t))\phi_\ell, \tag{30}$$

where $a_\ell, b_\ell \in \mathbb{R}$ are constants related to the initial conditions $p(0), p'(0) \in \mathbb{R}^N$ and

$$s_\ell^\pm = \tfrac{1}{2}\left(r \pm \sqrt{r^2 - 4\lambda_\ell}\right), \quad \ell \in \mathcal{N}. \tag{31}$$

In the solution (30), each component p_i of the vector $\mathfrak{p}$ will determine the temporal evolution of the perturbation in each location near the spatially homogeneous steady state $\overline{k}_\sigma$. Note that the eigenvalues (31) are symmetric around $r/2$ and they could be either real and positive, or real one positive and one negative, or complex with positive real parts. Because the dynamical system (29) has been derived from the optimal control problem (6), satisfaction of the transversality conditions at infinity requires setting the constant corresponding to the eigenvalue that is larger than $r/2$ equal to zero. Therefore if all the eigenvalues $s_\ell < r/2$ are negative, $p_i(t)$ tends to zero for all i and the spatial perturbation will die out. In this case the flat steady state is stable and no agglomeration emerges. If however for some $\ell \in \mathcal{N}$ there are eigenvalues in the interval $(0, r/2)$, then the spatial perturbation will not die out as t increases while transversality conditions at infinity are satisfied. In this case the flat steady state is not locally stable and this is a sign of agglomeration emergence.

More precisely (31) implies three possibilities:

(A) $r^2/4 < \lambda_\ell$, so that $s_\ell^\pm = r/2 \pm i\sigma$, i.e., a pair of complex conjugate roots. This leads to oscillatory behavior compatible with the transversality condition (Hopf type behavior).
(B) $0 < \lambda_\ell < r^2/4$, so that $s_\ell^- < r/2 < s_\ell^+$, i.e., a pair of real roots, one larger and one smaller than $r/2$. The root which is larger than $r/2$ is incompatible with the transversality condition and the corresponding constant is set to zero, while the root s_ℓ^-, as long as $s_\ell^- > 0$, leads to an instability which is optimal and satisfies transversality conditions. Agglomeration may emerge at the PF-RECE.
(C) $\lambda_\ell < 0$, so that $s_\ell^- < 0 < r/2 < s_\ell^+$, i.e., a pair of real roots, one negative and one positive larger than $r/2$. The root s_ℓ^+ does not satisfy the transversality condition and the corresponding constant is set to zero. For the negative root s_ℓ^-, the perturbation is suppressed in the long run and the flat steady state is stable. No agglomeration emerges at the PF-RECE.

Thus case B may lead to agglomerations. □

Theorem 2 provides general conditions for agglomeration emergence. In the remainder of this section we try to identify the key economic parameters that may (or may not) induce agglomerations by further specifying our model.

5.1. Agglomeration in the perfect foresight rational expectations equilibrium

First we provide conditions for the existence of a flat steady state that can be destabilized by spatial perturbations according to Theorem 2.

Assumption 3. The elasticity of the demand is uniformly bounded and negative, i.e., if we define the quantities

$$\underline{E}_D := \inf_{s>0}\left(\frac{sD'(s)}{D(s)}\right), \quad \overline{E}_D := \sup_{s>0}\left(\frac{sD'(s)}{D(s)}\right),$$

then it holds that

$$-\infty < \underline{E}_D \le \overline{E}_D < 0.$$

For the isoelastic demand this assumption holds and $\underline{E}_D \le \overline{E}_D = -\delta$.

Proposition 2 (*Potential agglomeration at the PF-RECE*). *Define the matrix*

$$T_0 = C_1 I + C_2 \mathbf{1} + C_3 W$$

where C_1, C_2, C_3 are derived from Definition3 for $\sigma = 0$.

(i) *Let $\gamma < 1$ and assume that*

$$\lim_{s \to 0} D(s) s^{(\gamma-1)/\gamma} > \left(\frac{1}{\rho N}\right)^{-(\gamma-1)/\gamma} \frac{r(q+\eta)}{\rho_k}. \tag{32}$$

Then, a unique spatially homogeneous steady state $\overline{k}_0$ exists, which can be destabilized and may give rise to agglomerations if the matrix T_0, defined as in (24) in Definition3 calculated at $\overline{k}_\sigma = \overline{k}_0$, has eigenvalues in the interval $(0, r^2/4]$.

(ii) *Let $\gamma > 1$, D satisfy Assumption3 with $\overline{E}_D < -\gamma/(\gamma-1)$ and assume existence of an $\underline{s} > 0$ such that*

$$D(\underline{s}) \underline{s}^{(\gamma-1)/\gamma} > \left(\frac{1}{\rho N}\right)^{-(\gamma-1)/\gamma} \frac{r(q+\eta)}{\rho_k}.$$

Then, a unique spatially homogeneous steady state $\overline{k}_0$ exists, which can be destabilized and may give rise to agglomerations if the matrix T_0, defined as in (24) in Definition3 calculated at $\overline{k}_\sigma = \overline{k}_0$, has eigenvalues in the interval $(0, r^2/4]$.

Agglomeration emergence is related to γ, the degree of homogeneity of the production function. At a flat steady state, $\gamma > 1$ indicates increasing returns from a social point of view, while $\gamma < 1$ indicates decreasing returns to scale and diminishing returns from both the social and the private point of view. Proposition 3 suggests, therefore, that agglomeration is could be possible even with decreasing returns to scale. This proposition combined with Theorem 2 covers all possible cases in which agglomeration may be possible in the PF-RECE case, but identification of agglomerations requires the numerical calculation of the spectrum of the matrix T_0. This is straightforward even for the case of large dimensional systems (large N) but does not provide us with sufficient intuition regarding the forces and the parameters which are important in inducing agglomerations.

Since it not possible to provide closed form solutions for the eigenvalues of matrix T_0 we examine two special cases which provide more insights into the structure of industry fundamentals that could lead to agglomerations in the PF-RECE. It should be stressed however that these special cases provide only sufficient conditions for agglomerations. Actual detection of agglomerations requires numerical analysis, some examples of which are presented in Section 6. We start by examining in Proposition 3, the case where $w_{ij} > 0$ for all i,j, which means that the impact on site i from all sites j is beneficial. This assumption allows us to characterize the eigenvalues of T_0 in terms of stable and unstable Metzler matrices.[15]

Proposition 3 (*Potential agglomeration (or not) at the PF-RECE with $f_{kK} > 0$*). *Let $\sigma = 0$ and assume that a spatially homogeneous steady state $\overline{k}_0 > 0$ exists, that Assumption3 holds, that $w_{ii} > 0$, and that T_0 is a Metzler matrix, i.e., $w_{ij} \geq -(1/\rho N)(\rho_k/\rho_{kK})(\rho_k + \overline{w}\rho_K)\underline{E}_D \geq 0$ for all $i,j \in \mathcal{N}$.*

(i) *If the industry fundamentals are such that*

$$w_{ii} < -\frac{\rho_{kk}}{\rho_{kK}} - \frac{1}{\rho N} \frac{\rho_k}{\rho_{kK}} (\rho_k + \overline{w}\rho_K) \overline{E}_D, \quad i \in \mathcal{N},$$

then no agglomeration may emerge in the PF-RECE case.

(ii) *If the industry fundamentals are such that*

$$w_{ii} \geq -\frac{\rho_{kk}}{\rho_{kK}} - \frac{1}{\rho N} \frac{\rho_k}{\rho_{kK}} (\rho_k + \overline{w}\rho_K) \overline{E}_D, \quad i \in \mathcal{N},$$

then agglomerations may emerge. The top eigenvalue of the stability matrix is

$$\lambda^* = \frac{q(r+\eta)}{\alpha \overline{k}_0 \rho_k} \left(\rho_{kk} + \frac{\rho_k}{\rho} (\rho_k + \overline{w}\rho_K) \frac{sD'(s)}{D(s)} + \rho_{kK} \overline{w} \right).$$

The right-hand side of all inequalities are positive numbers which are defined in terms of the aggregate externality $\overline{w}$, the production function structure, and the elasticity of demand, while the terms w_{ii}, w_{ij} reflect "own" impact on the externality affecting site i, and impact of site j on the externality affecting site i respectively. Thus when w_{ij} satisfies the assumption of Proposition 3, no agglomeration requires small own impacts, while potential agglomeration requires large own impacts. Agglomeration may emerge when $\lambda^* > 0$. From its definition it is clear that this depends on how strong the

[15] Nonnegative and Metzler matrices have been used in economics, e.g., in input–output analysis (the input–output matrix) or in general equilibrium to model the gross substitution matrix.

complementarity between k and K is, and how large the aggregate externality $\overline{w}$ is, since these effects are combined to form the term $\rho_{kK}\overline{w}$ which is the only term that can make λ^* positive. It should be noted that Proposition 3 indicates that agglomeration can emerge both with increasing and decreasing returns to scale provided that $f_{kK}>0$ (which implies $\rho_{kK}>0$). From Remarks 1 and 2 we see that with a Cobb–Douglas technology, f_{kK} is always positive. With a CES technology f_{kK} is positive for constant and increasing returns to scale, while it is also positive for decreasing returns to scale ($\gamma<1$) if $(\gamma+\theta)>0$ or when the elasticity of substitution between k and K is $\hat{\sigma}<1/(1-\gamma)$. Thus Proposition 3 provides sufficient conditions for potential agglomerations under any type of returns to scale when technology is CES and $f_{kK}>0$. If $\hat{\sigma}>1/(1-\gamma)$ or $f_{kK}<0$, the proposition does not provide sufficient conditions for potential agglomeration emergence.

In order to examine conditions for agglomeration emergence with decreasing returns to scale and $f_{kK}<0$, we assume that $w_{ij}<0$ for some $i\neq j$. This is the case of a composite externality which is positive overall at a flat steady state, but negative local effects are present. Proposition 4 provides sufficient conditions for potential agglomeration with diminishing returns even when $f_{kK}<0$ (or equivalently $\rho_{kK}<0$).

Proposition 4 (*Potential agglomeration at the PF-RECE with $f_{kK}<0$*). *Consider the PF-RECE case ($\sigma=0$) and assume that $\rho_{kK}<0$. Let $(n,m)\in\arg\min_{(i,j),i\neq j}w_{ij}$, then a sufficient condition for agglomeration is that*

$$|w_{nm}|>\frac{1}{2}(w_{nn}+w_{mm})+\frac{\rho_{kk}}{\rho_{kK}}+\frac{(\rho_k+\overline{w}\rho_K)\rho_k}{\rho_{kK}}\frac{D'(\rho N\overline{k}_0^{\gamma})}{D(\rho N\overline{k}_0^{\gamma})}\overline{k}_0^{\gamma}. \tag{33}$$

Condition (33) clearly illustrates the interplay between local negative externalities and the cross partial ρ_{kK} in the formation of agglomeration patterns. A strong enough local negative externality allows the formation of agglomeration patterns. Since a negative cross partial f_{kK} (or ρ_{kK}) requires decreasing returns to scale, Proposition 4 provides sufficient conditions for potential agglomeration under decreasing returns to scale when an increase in the spatial externality shifts the marginal product of private capital downwards. It should be noted that this sufficient condition is purely local in terms of the spatial externality since it depends only on w_{nm}, w_{nn} and w_{mm}.

To summarize, Proposition 2 is more general in scope than Propositions 3 and 4, in the sense that Propositions 3 and 4 provide sufficient conditions and not the whole range of parameters for which agglomeration or no agglomeration is expected. However Propositions 3 and 4 provide explicit results regarding the structure of production technology, demand, and spatial externality as reflected in the spatial interaction matrix W, which could lead to agglomerations under any type of returns to scale and which allow for both positive and negative impacts of the externality on the marginal product of private capital.

For example, for agglomeration emergence for case (ii) of Proposition 3, we may obtain a simplified condition in terms of

$$\rho_{kK}>-\frac{\rho_k\rho_K}{\rho}\underline{E}_D,$$

$$\overline{w}>-\frac{\rho\rho_{kk}+\rho_k^2\underline{E}_D}{\rho_k\rho_K\underline{E}_D+\rho\rho_{kK}}$$

which again indicates more clearly the strength of complementarity and the size of externality as an agglomeration-inducing factor. The no-agglomeration-emergence condition of Proposition 3 with an isoelastic demand implies

$$w_{ii}-w_{ij}<-\frac{\rho_{kk}}{\rho_{kK}},\quad i,j\in\mathcal{N}.$$

Thus a small deviation between own impacts and other sites' impacts acts as an agglomeration-suppressing force. Since small deviations between w_{ii} and w_{ij} are suppressing potential agglomerations, an upper bound for w_{ij} would be expected such that, for fixed w_{ii}, if w_{ij} exceeds this upper bound potential agglomerations will be suppressed. The exact determination of this upper bound requires, however, the use of numerical methods in order to estimate the full spectrum of T_0. If we are dealing with a composite kernel where $w_{ij}<0$ for some j, Proposition 4 applies. If we assume a kernel where w_{ii} is the same for all i, then (33) implies that for agglomeration emergence $|w_{nm}|-w_{nn}$ should be sufficiently high when $\rho_{kK}<0$, since in this case the RHS of (33) is always positive. On the other hand if $\rho_{kK}<0$ and $|w_{nm}|>w_{nn}$, the agglomeration emergence condition will be satisfied since the RHS will be negative. Thus composite externalities with strong local negative effects can be regarded as an agglomeration-inducing factor.

Another way to provide additional clarifications regarding the impact of the cross partial f_{kK} and the spillover effects w_{ij} on the emergence of linear instability of the flat steady state is to consider a simplified case of two types of firms $i=1,2$. Writing the Hamiltonian system for each firm's problem in the state costate space, assuming a fixed price $\overline{p}$ to simplify, and N_i identical firms of type $i=1,2$ with varying q_i, a_i, the characteristic equation of the (4×4) linearization matrix is

$$Q_4(s_\ell)=s_\ell^4-2rs_\ell^3+\left(r^2-\frac{J_{13}}{a_1}\right)s_\ell^2+\left(\frac{J_{24}}{a_2}+r\frac{J_{23}}{a_1}\right)s_\ell+\frac{J_{14}J_{24}}{a_1a_2}=0 \tag{34}$$

$$J_{13}=-\overline{p}(f_{k_1k_1}+f_{k_1K_1}w_{11}N_1),\quad J_{14}=-\overline{p}(f_{k_1K_1}w_{12}N_2) \tag{35}$$

$$J_{23} = -\overline{p}(f_{k_2K_2}w_{21}N_1), \quad J_{24} = -\overline{p}(f_{k_2k_2} + f_{k_2K_2}w_{22}N_2) \tag{36}$$

where costate rows in the Hamiltonian system are labelled 1 and 2 and state rows are labelled 3 and 4. Instability emerges if (34) has three positive roots since the saddle point stability of the Hamiltonian system implies two positive and two negative roots.[16] If $w_{ij} > 0$ and $f_{k_iK_j} < 0$, $i,j = 1,2, i \neq j$ then all J's are positive and by Descartes' rules of signs the maximum number of positive and negative roots of (34) is two. To obtain a maximum of three positive roots, which means instability, we need some $w_{ij} < 0$ or some $f_{k_iK_j} > 0$. Basically this is another way of presenting the results of Propositions 3 and 4. If an increase in global externality shifts the marginal private product of capital downwards (i.e. $f_{k_iK_j} < 0$), then agglomeration emergence requires some local spillovers effects to be negative so that the desirability of a location changes according to the proximity of the location to the range of negative spillovers. This result holds for the PF-RECE under decreasing returns to scale.

Having studied the PF-RECE we turn now to providing further conditions for potential agglomeration emergence at the SO.

5.2. Agglomeration and the social optimum

There is *no* agglomeration emergence in the SO case if $\gamma < 1$, as the global result of Theorem 1 guarantees. One may obtain a generalization of Proposition 3 for the SO case, and provide conditions on the fundamentals of the economy under which no agglomerations will emerge even in the case $\gamma > 1$, which is not covered by Theorem 1.

Proposition 5 (*Potential agglomeration at the SO*). *Define the matrix*

$$T_1 = C_1 I + C_2 \mathbf{1} + C_3 W + C_4 W^2,$$

where C_1, C_2, C_3, C_4 are derived from Definition3 for $\sigma = 1$.

Let $\gamma > 1$, D satisfy Assumption3 with $\overline{E}_D < -\gamma/(\gamma-1)$ and assume existence of an $\underline{s} > 0$ such that

$$D(\underline{s})\underline{s}^{(\gamma-1)/\gamma} > \left(\frac{1}{\rho N}\right)^{-(\gamma-1)/\gamma} \frac{r(q+\eta)}{\rho_k + \overline{w}\rho_K}.$$

Then, a unique spatially homogeneous steady state $\overline{k}_1$ exists, which can be destabilized and may give rise to agglomerations if the matrix T_1, defined as in (24) in Definition 3 calculated at $\overline{k}_\sigma = \overline{k}_1$, has eigenvalues in the interval $(0, r^2/4]$.

Proposition 6 (*Potential agglomeration (or not) at the SO*). *Assume $\sigma = 1$, $\gamma > 1$ and let Assumption3 hold.*

(*i*) *If the industry fundamentals satisfy*

$$0 \geq \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\overline{E}_D + 2\rho_{kK}w_{ii} + \rho_{KK}\sum_r w_{ir}w_{ri},$$

$$0 \leq (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\underline{E}_D + 2\rho_{kK}w_{ij} + \rho_{KK}\sum_r w_{ir}w_{rj}, \quad i \neq j,$$

for all $i,j \in \mathcal{N}$, no agglomerations are possible for the SO case.

(*ii*) *If the industry fundamentals satisfy*

$$0 \geq \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\underline{E}_D + 2\rho_{kK}w_{ii} + \rho_{KK}\sum_r w_{ir}w_{ri},$$

$$0 \leq (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\underline{E}_D + 2\rho_{kK}w_{ij} + \rho_{KK}\sum_r w_{ir}w_{rj}, \quad i \neq j,$$

for all $i,j \in \mathcal{N}$, agglomerations could be possible for the SO case. The top eigenvalue of the stability matrix T_1 can be found explicitly in terms of the homogeneous steady state $\overline{k}_1$, as

$$\lambda^* = \frac{1}{\alpha(\rho_k + \overline{w}\rho_K)\overline{k}_1} \frac{q(r+\eta)}{} \left(\rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho}\frac{sD'(s)}{D(s)} + 2\rho_{kK}\overline{w} + \rho_{KK}\overline{w}^2\right),$$

where $s = \rho N \overline{k}_1^{\gamma}$.

For the proofs of Propositions 2, 3, 4, 5 and 6 see Appendix, Sections A.3–A.7.

Remark 5. Conditions on the demand function imply that: (a) $D(Q(k,K))$ is maintained above a critical level (related to $q(r+\eta)$) when $Q(k,K)$ falls below a critical level, and (b) the demand function decays fast enough for large enough values of

[16] From the well known Kurz's result (Kurz, 1968, p. 163) about the stability of optimal control in growth models, we know that the Hamiltonian system will have either saddle point stability (two positive and two negative roots), or will be unstable with more than two positive real roots, or with complex roots with positive real parts.

$Q(k,K)$. The second condition is quantified by $\overline{E}_D < -\gamma/(\gamma-1)$ (for $\gamma>1$) which in turn implies that $D(s)s^{(\gamma-1)/\gamma}$ is a decreasing function of s. The isoelastic demand function $D(s)=Bs^{-\delta}$ for large enough s satisfies this condition as long as $\delta>\gamma/(\gamma-1)$.

5.3. Agglomeration emergence: discussion

The results obtained above provide a number of points which are useful in understanding the potential emergence of agglomerations in a competitive industry.

In general $\overline{k}_1 \neq \overline{k}_0$, i.e. the steady state of the SO problem does not coincide with the steady state of the PF-RECE. Furthermore, for a strictly decreasing function $D(s)s^{(\gamma-1)/\gamma}$ and values of γ that are relevant for homogeneous production technologies, $\overline{k}_1 > \overline{k}_0$. Furthermore, $\overline{k}_\sigma$ $(\sigma=0,1)$ is an increasing function of $\overline{w}$.

The agglomeration emergence, related to part of the spectrum of T_σ being in region $\mathcal{B}$, is reminiscent of a Turing instability but with a major difference. It is related to an optimally controlled system, and this fact imposes major restrictions as to what will be an acceptable instability. As we see in the proof of Theorem 2 the instability condition needs to satisfy the transversality condition. Instabilities related to the part of the spectrum of T_σ being in region $\mathcal{A}$ are associated with a Hopf type bifurcation.

The conditions for agglomeration emergence in the linearized problem are related to the spectrum of the symmetric matrix T_σ. This is easily computed for concrete applications numerically (see e.g. Section 6.1). However, the concavity properties of the production function f as well as the monotonicity of the demand function provide important information on the signs of the constants C_1, C_2, C_3, C_4 and thus allow us to obtain general information concerning the position of the spectrum of the matrix T_σ. For example, consider first the PF-RECE case. We see that C_1 is always negative as long as $\rho_{kk}<0$ and since this term is responsible for a contribution $C_1 I$ to the stability matrix, this term will always contribute a negative eigenvalue, leading to stability. The diagonal part is perturbed by the term $C_2\mathbf{1}$ with C_2 being always negative, since D' is negative and $\rho_k, \rho_K>0$. Therefore, this term is not expected to lead to further destabilization for $\overline{w}>0$. The third term is a contribution $C_3 W$, where C_3 is positive when $\rho_{kK}>0$, which implies complementarity between own capital stock and the spatial externality. This term can induce spatial instability through the occurrence of positive spectrum if it is strong enough. If $\rho_{kK}<0$, which implies decreasing returns to scale with CES technology, then agglomeration is not possible if the externality is positive with respect to all locations, or $w_{ij}>0$ for all i,j. Agglomeration could be possible with decreasing returns to scale and $\rho_{kK}<0$ when the global externality is positive but there are sufficiently strong local negative effects to significantly reduce the desirability of locations which are close to the range of negative effects. The relative strength of C_3 with respect to C_1 and C_2 depends on the fundamentals of the problem, e.g. on N and $\overline{w}$, but the actual dependence is too complicated to be studied, unless explicit forms for $D$and f are assumed (see Section 6.1). In the SO the extra term $C_4 W^2$ is included in the stability matrix. If $\gamma<1$, this is a stabilizing term and this, in accordance with our global results in Section 4, eliminates agglomerations. If $\gamma>1$, this term may further contribute to instability and may lead to agglomeration formation.

It should be noted that for a Cobb–Douglas technology, if $(\gamma_1,\gamma_2)<(1,1)$ but $\gamma_1+\gamma_2=\gamma>1$, the flat steady state with spatial externalities is characterized by diminishing marginal productivity of capital from the private point of view, and by increasing marginal productivity from the social point of view. Increasing marginal productivity from the social point of view may induce agglomerations when the spatial externality is fully internalized. Thus the "no agglomeration" result, under full internalization, requires diminishing marginal productivity of capital from both the private and the social point of view.

With diminishing marginal productivity of capital from both the private and the social point of view, agglomeration may emerge as a long-run outcome of a PF-RECE but not at the SO. The local behavior described for the linearized system around the flat PF-RECE steady state $\overline{k}$, by Theorem 2 presents a potential scenario for the long-run spatial behavior of system (13). In particular, it is possible that some of the unstable modes leading to spatial patterns for the linearized system may persist, leading thus to stable agglomerations in the long run. The exact identification and characterization of the stability properties of such a long-run agglomeration require detailed analysis of the full nonlinear system, which is beyond the scope of the present paper, but represent a natural way of extending this research. It is interesting to note that this is in striking contrast to what happens for the SO equilibrium, where agglomerations and clustering in the long run are definitely ruled out by Theorem 1, which is based on the strict concavity of the production function in k and $K=Wk$. In terms of economics this means that diminishing returns from the social point of view, and full internalization of the spatial externality at the firm level, eradicate any spatial patterns. When, however, the spatial externality is not internalized in a competitive industry then spatial agglomerations may occur. It should be noted that this result does depend on increasing returns, geometry of the spatial domain and boundary conditions, or location advantages. In this case agglomeration-inducing forces, or centrifugal forces, include incomplete internalization of the spatial externality, strong complementarity between the stock of capital and the spatial externality when all local effects are positive, or a composite spatial externality which is positive overall but which includes positive and negative local spillovers, and relatively large deviations between own and other locations.

It should be noted that the spatial externality studied here is related to non-local or long-range spatial effects modelled by the spatial kernel. The effects of the spatial externality can also be analyzed by considering local effects modelled by local diffusion, as for example, in the growing literature related to spatial economic growth and the spatial Ramsey model (e.g., Camacho and Zou, 2004; Camacho et al., 2008; Boucekkine et al., 2009; Brito, 2011; Boucekkine et al., 2013a, 2013b), or the spatial diffusion of resources in resource management problems (e.g. Brock and Xepapadeas, 2008, 2010). Undoubtedly this is an important area for further research.

6. Illustrative examples

6.1. The Cobb–Douglas production function

In this section we provide an illustrative example using the Cobb–Douglas production function,

$$f(x,y)=Cx^{\gamma_1}y^{\gamma_2},\quad \gamma=\gamma_1+\gamma_2,\ \gamma_1<1,\ \gamma_2<1.$$

This function evaluated at $x=k_i$ and $y=(Wk)_i=\sum_j w_{ij}k_j$ gives the production at site i of the spatial economy as a homogeneous production function with degree of homogeneity γ. For a spatially homogenous steady state, $\gamma>1$ means increasing marginal productivity from the social point of view in the sense of Romer (1986), while $\gamma<1$ means diminishing marginal productivity from the social point of view. We assume, furthermore, that matrix W corresponds to a coupling of diffusive type, for which $\sum_j w_{ij}=\overline{w}>0$ for any $i\in\mathcal{N}$, in accordance with Assumption 2, and that the demand function is of the isoelastic form $D(s)=Bs^{-\delta}$, $\delta>0$. This demand function satisfies Assumption 3 with $\underline{E}_D=\overline{E}_D=-\delta$. Using this structure we calculate the flat steady states and the stability matrix for the PF-RECE ($\sigma=0$) and the SO ($\sigma=1$). Calculations are presented in the Appendix, Section A.8. Furthermore, conditions for non-emergence of agglomerations in the PF-RECE case simplify to

$$\frac{w_{ii}}{\overline{w}}<\frac{1-\gamma_1}{\gamma_2}+\frac{\delta\,\gamma}{N\gamma_2}, \tag{37}$$

$$\frac{w_{ij}}{\overline{w}}>\frac{\delta\,\gamma}{N\gamma_2},\quad i\neq j, \tag{38}$$

for all $i,j\in\mathcal{N}$, which may be reinterpreted as

$$\frac{w_{ii}-w_{ij}}{\overline{w}}<\frac{1-\gamma_1}{\gamma_2}. \tag{39}$$

Both these relations provide important insight into the mechanics of agglomeration in the PF-RECE case. The first interpretation of the stability conditions provides the interesting information that if w_{ij} are all positive, then agglomeration in the PF-RECE case is expected to take place for large enough values of γ. To see this we may reason as follows: Let all the $w_{ij}>0$ as is the case of a single positive spatial externality, and assume (37) holds. In order to have the possibility of agglomeration we need $w_{ij}/\overline{w}<(\delta/N)\gamma/\gamma_2$, and that can most easily be achieved if γ is large. We therefore expect occurrence of patterns for the PF-RECE case, for single positive spatial externalities for large enough values of γ. This is supported by numerical evidence as shown in Section 6.1.1. On the contrary, if we have composite kernels that combine positive and negative externalities, then while condition (37) may hold, condition (38) is never valid since for kernels of this type there exist $i,j\in\mathcal{N}$ such that condition (38) fails. Therefore, instability may occur more easily for composite kernels combining positive and negative externalities, and for smaller values of γ. This theoretical prediction is fully supported by the numerical results provided in Section 6.1.2. On the other hand, the alternative form of the stability condition (39) also provides some important qualitative information on the mechanics of agglomeration formation. If the difference between the diagonal and the off-diagonal terms of the interaction matrix W is small enough, no agglomeration is expected in the PF-RECE. The difference between w_{ii} and w_{ij}, and even more so the ratio $(w_{ii}-w_{ij})/\overline{w}$, can be interpreted as the importance of site i's contribution to the externality at site i, relative to the effect that site j has on the externality at site i. If this effect is small as quantified by (39), then no agglomerations are induced. If, on the contrary, it is larger than the critical value provided by (39), then agglomeration emergence may occur.

We provide numerical results concerning potential agglomeration emergence through the optimal spillover induced instability of a flat steady state. We choose a spatial economy consisting of $N=101$ sites. The parameters of the model are chosen as follows: $r=0.03$, $\eta=0.02$, $q=1$, $\alpha=0.025$, $\delta=1.25$, $B=100$, $C=1$ and these are kept fixed in all the numerical experiments that follow. We choose to vary the parameters of the production function γ_1, γ_2 as well as the type of spatial interaction kernel $\overline{w}$, which is used to generate the matrix W. We provide two sets of results corresponding to a single and a composite spatial externality.

6.1.1. Single spatial externality

In the first set of results, we model the spatial externality with an interaction kernel of the form

$$w(|i-j|)=A_1\exp(-\alpha_1|i-j|^2),$$

which is exponentially decaying with a single hump and corresponds to a single positive spatial externality. We choose $A_1=2$, $a_1=0.025$, and the form of the kernel is shown in Fig. 2. For the corresponding interaction matrix W, and using the fundamentals of the industry as described above, for each choice of parameters (γ_1,γ_2) we generate the corresponding stability matrix $T_\sigma=T(\gamma_1,\gamma_2)$, both for the PF-RECE case ($\sigma=0$) and for the SO case ($\sigma=1$) and study its spectrum as a function of the parameters (γ_1,γ_2). In Fig. 3, we present the region in the (γ_1,γ_2) plane, which corresponds to the top eigenvalue of the matrix $T(\gamma_1,\gamma_2)$ being in the interval $(0,r^2/4)$ (shaded region). For values of (γ_1,γ_2) within the shaded

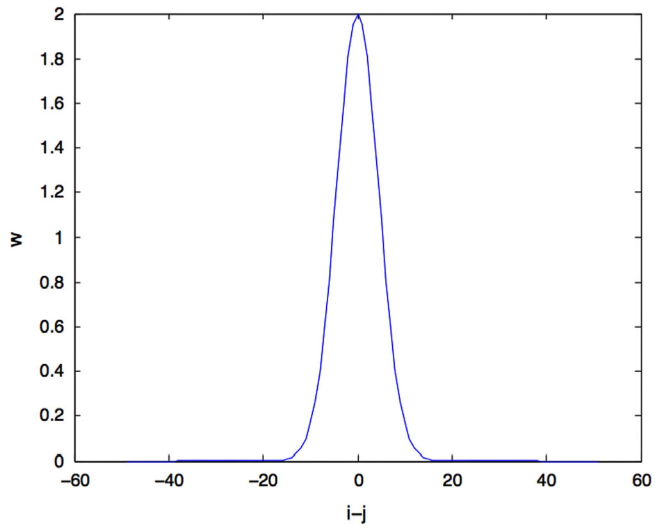

Fig. 2. The kernel of a single positive spatial externality.

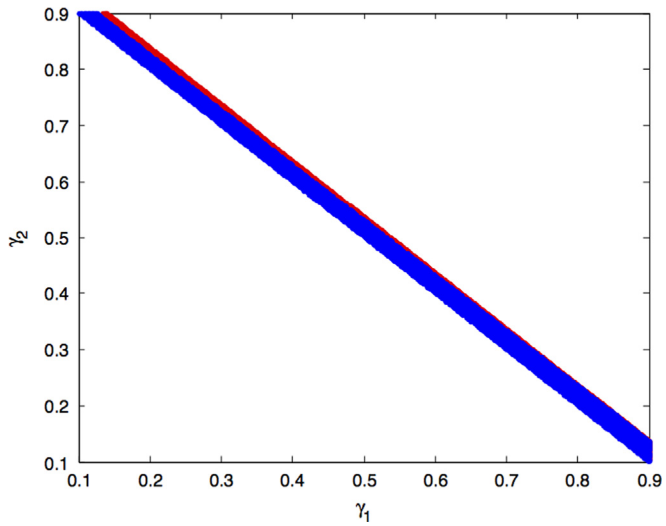

Fig. 3. Stability diagram with a single positive spatial externality. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

region, we therefore expect agglomeration to occur. Note that this region corresponds to values of (γ_1, γ_2) such that $\gamma_1+\gamma_2>1$, but $\gamma_1<1, \gamma_2<1$. The red band is the result for the matrix T_0 (the PF-RECE case) while the blue band is the result for the matrix T_1 (the SO case). It can be seen that both the PF-RECE and the SO equilibria may lead to agglomerations if $\gamma=\gamma_1+\gamma_2>1$, which implies increasing returns from the social point of view.

Keeping all parameters fixed for the same values as used in the two previous figures, we perturb the system from the spatially homogeneous PF-RECE steady state $\overline{k}_0$, by a small random spatially varying perturbation. In Fig. 4 we show the spatiotemporal evolution of the perturbed initial state, obtained by numerical integration of the resulting ODE, choosing the values for the parameters $\gamma_1=0.2, \gamma_2=0.7$. From the stability results shown in Fig. 3, we expect no agglomeration for these parameter values. As predicted by our theoretical results, the full numerical simulation indicates that the initial random spatial disturbance soon dies out and the system equilibrates once more to the spatially homogeneous steady state. In Fig. 5 we do the same as for Fig. 4, with the sole difference that we choose the values for the parameters $\gamma_1=0.16, \gamma_2=0.878$. From the stability results shown in Fig. 3, we now, contrary to the previous case, expect agglomeration for these parameter values. Indeed, as predicted by our theoretical results, the full numerical simulation indicates that the initial random spatial disturbance is strengthened and soon a spatial pattern is formed which, even though it started from a random initial perturbation, has a well defined shape, as the linear combination of the eigenvectors which correspond to the positive eigenvalues. Therefore, in this case we have the emergence of agglomeration (akin to Turing instability) which is generated from the destabilization by spatial interactions of a spatially homogeneous steady state. This pattern is compatible with the transversality condition, so we can call this pattern the optimal emerging potential agglomeration at a PF-RECE.

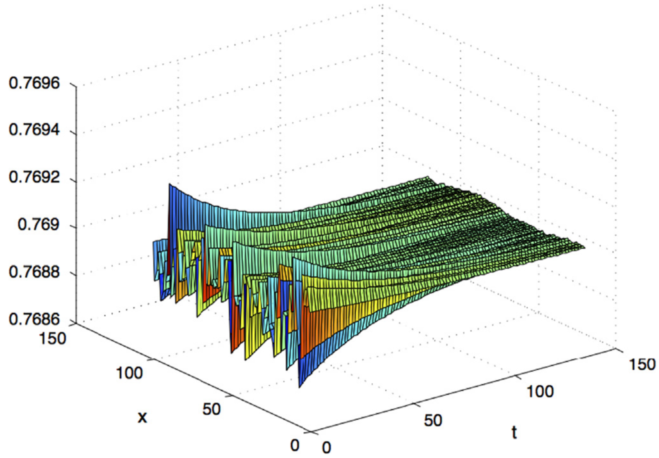

Fig. 4. No agglomeration at the PF-RECE for $\gamma < 1$.

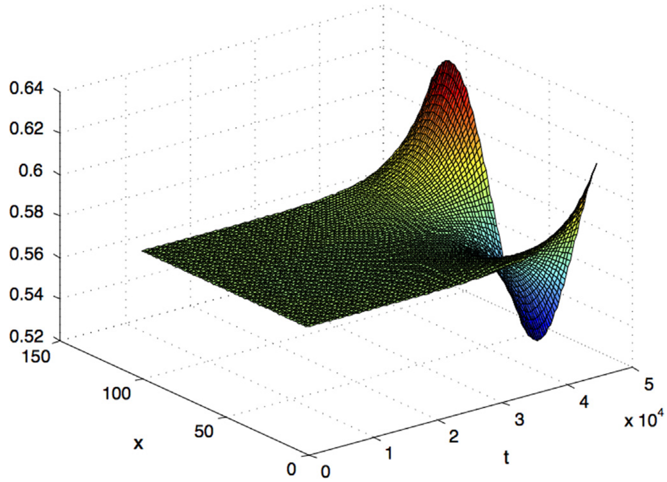

Fig. 5. Potential agglomeration emergence at the PF-RECE for $\gamma > 1$.

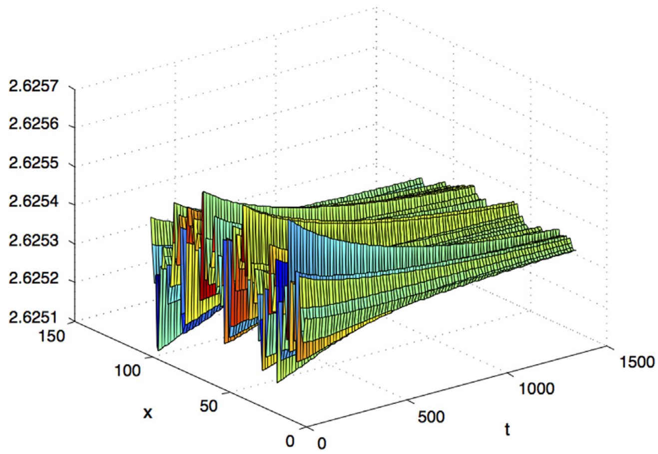

Fig. 6. No agglomeration at the social optimum for $\gamma < 1$.

Keeping all parameters fixed for the same values as used in the two previous figures, we perturb the system from the spatially homogeneous SO steady state $\overline{k}_1$, by a small random spatially varying perturbation. In Fig. 6 we show the spatiotemporal evolution of the perturbed initial state for the parameters $\gamma_1 = 0.2$, $\gamma_2 = 0.7$. As predicted by our theoretical results, the initial random spatial disturbance soon dies out and the system equilibrates once more to the SO spatially

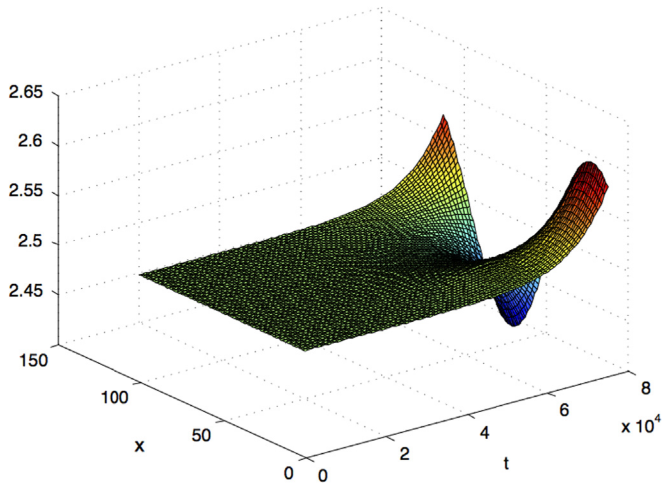

Fig. 7. Potential agglomeration emergence at the social optimum for $\gamma > 1$.

homogeneous steady state. In Fig. 7 we do the same as for Fig. 6, by allowing for increasing returns from the social point of view, i.e. $\gamma_1 = 0.122$, $\gamma_2 = 0.891$. As predicted by our theoretical results, the initial random spatial disturbance is strengthened and soon a spatial pattern is formed which, even though it started from a random initial perturbation, has a well defined shape, as the linear combination of the eigenvectors which correspond to the positive eigenvalues. We have again agglomeration (akin to Turing instability) which is generated from the destabilization by spatial interactions of a spatially homogeneous steady state. This pattern is compatible with the transversality condition, so we can call this pattern the emerging optimal agglomeration at the SO.

6.1.2. Composite spatial externality

In the second set of results we model a composite externality by an interaction kernel of the form

$$w(|i-j|) = A_1 \exp(-\alpha_1 |i-j|^2) + A_2 \exp(-\alpha_2 |i-j|^2). \tag{40}$$

By choosing $A_1 = 2$, $a_1 = 0.025$, $A_2 = -0.5$, $a_2 = 0.0025$, we now obtain a non-monotonic kernel with a single maximum and two local minima as shown in Fig. 8. In (40) the first term corresponds to the positive externality and the second to the negative externality. We generate as before the corresponding stability matrix $T = T(\gamma_1, \gamma_2)$ and study its spectrum as a function of the parameters (γ_1, γ_2). In Fig. 9, we present the region in the (γ_1, γ_2) plane, which corresponds to the top eigenvalue of the matrix $T(\gamma_1, \gamma_2)$ being in the interval $(0, r^2/4)$ (shaded region). For values of (γ_1, γ_2) within the shaded region we therefore expect agglomeration to occur. Note that this region now corresponds to values of (γ_1, γ_2) such that $\gamma_1 + \gamma_2 < 1$. Thus a composite externality induces agglomeration at a PF-RECE without increasing returns from the social point of view. Since the instability satisfies the transversality condition, we again have optimal agglomeration at a PF-RECE. In this numerical example we use a nonmonotonic kernel with negative local effects in order to be in line with the notion that spatial spillovers attenuate with distance. As our numerical simulations indicate, for both the Cobb–Douglas and the CES technologies, nonmonotonicity is not necessary for agglomeration emergence with decreasing returns to scale at the PF-RECE.

Keeping all parameters fixed for the same values as used in Figs. 8 and 9 we perturb the system from the spatially homogeneous PF-RECE steady state $\overline{k}_0$, by a small random spatially varying perturbation. In Fig. 10 we show the spatiotemporal evolution of the perturbed initial state for the parameters $\gamma_1 = 0.16$, $\gamma_2 = 0.256$. From the stability results shown in Fig. 9, we expect agglomeration for these parameter values. Indeed, as predicted by our theoretical results, the initial random spatial disturbance is strengthened and soon a spatial pattern is formed which, even though it started from a random initial perturbation, has a well defined shape, as the linear combination of the eigenvectors of T_0 which correspond to the positive eigenvalues. Therefore, in this case we have agglomeration (akin to Turing instability) at a PF-RECE with diminishing returns from both the private and the social point of view. Since the pattern satisfies the transversality condition, we again may have an emerging optimal agglomeration at the PF-RECE.

We now, keeping all parameters fixed for the same values as used in Figs. 8 and 9, perturb the system from the spatially homogeneous SO steady state $\overline{k}_1$ by a small random spatially varying perturbation. In Fig. 11 we show the spatiotemporal evolution of the perturbed initial state, obtained by numerical integration of the resulting ODE, choosing the values for the parameters $\gamma_1 = 0.2$, $\gamma_2 = 0.7$. From the stability results shown in Fig. 9, we expect no agglomeration for these parameter values. Indeed the initial random spatial disturbance soon dies out and the system equilibrates once more to the spatially homogeneous steady state. This is of course the result of strict concavity of the production function. In Fig. 12, we do the same as for Fig. 11, with the sole difference that we allow for increasing returns from the social point of view by choosing $\gamma_1 = 0.116$, $\gamma_2 = 0.9$. From the stability results shown in Fig. 3, we expect agglomeration for these parameter values. Indeed the full numerical simulation indicates that the initial random spatial disturbance is strengthened and soon a spatial pattern is formed which, even though it

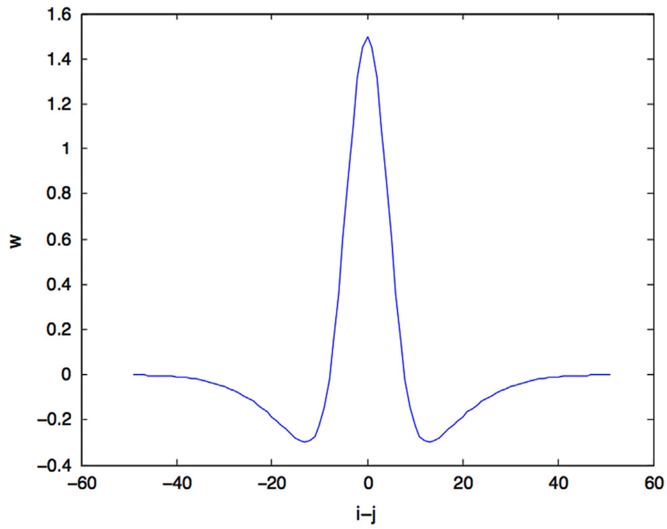

Fig. 8. The kernel of a composite – positive and negative – spatial externality.

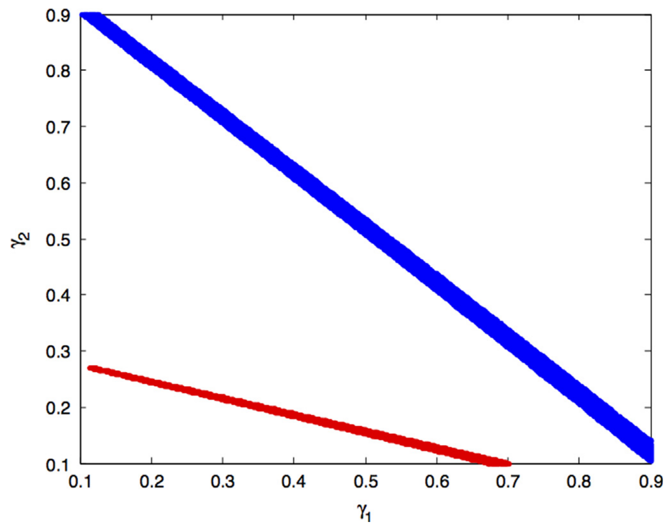

Fig. 9. Stability diagram with a composite spatial externality.

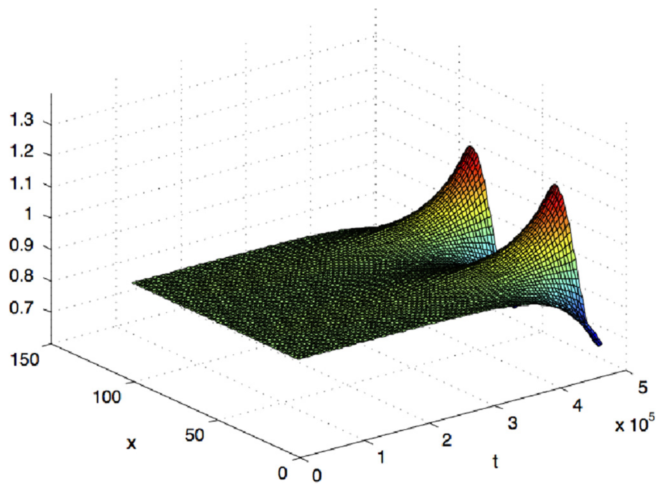

Fig. 10. Potential agglomeration emergence at the PF-RECE for $\gamma < 1$.

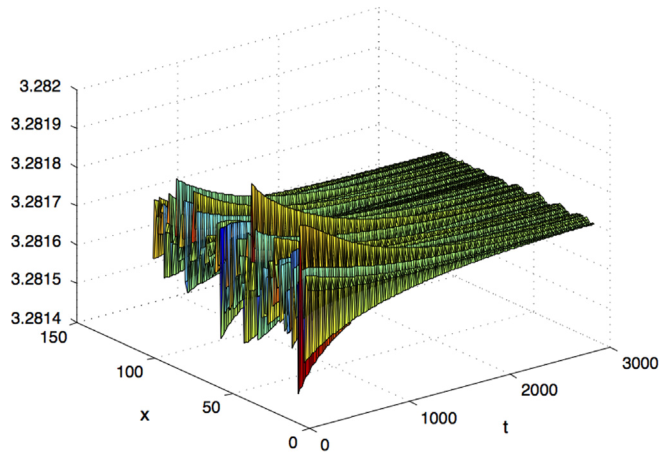

Fig. 11. No agglomeration at the social optimum for $\gamma < 1$.

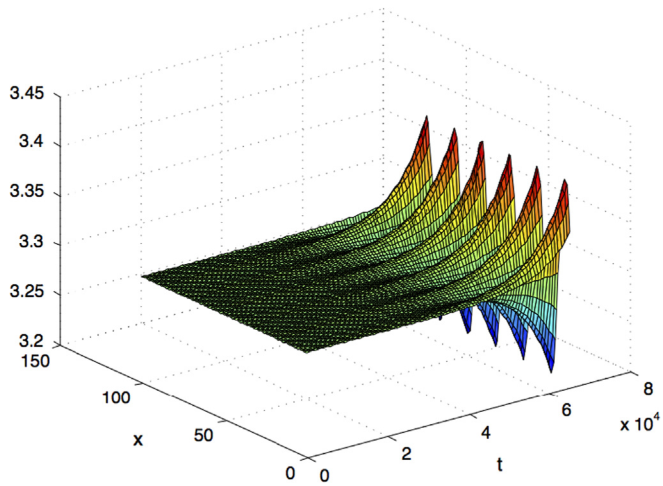

Fig. 12. Potential agglomeration emergence at the social optimum for $\gamma > 1$.

started from a random initial perturbation, has a well defined shape, as the linear combination of the eigenvectors which correspond to the positive eigenvalues. Therefore, in this case we have agglomeration (akin to Turing instability) at the SO. This pattern is compatible with the transversality condition and may induce optimal agglomeration.

6.2. The CES production function

As an illustration that increasing returns to scale, or $f_{kK} > 0$ are not required for potential agglomeration emergence, we now turn our attention to the study of a CES production function for which $f_{kK} < 0$. Consider the CES production function

$$f(x,y) = A(\beta x^{-\theta} + (1-\beta)y^{-\theta})^{-\gamma/\theta},$$

where $\theta \geq -1$, $0 < \beta < 1$ and γ is the degree of homogeneity. This function evaluated at $x = k_i$ and $y = (Wk)_i = \sum_j w_{ij} k_j$ gives the production at site i of the spatial economy. An elementary calculation yields

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = A\gamma\beta(1-\beta)(\gamma+\theta)(xy)^{-\theta-1}(\beta x^{-\theta} + (1-\beta)y^{-\theta})^{-\gamma/\theta-2},$$

so that at a flat steady state we have that

$$\rho_{kK} = A\gamma\beta(1-\beta)(\gamma+\theta)\overline{w}^{-\theta-1}(\beta + (1-\beta)\overline{w}^{-\theta})^{-\gamma/\theta-2},$$

which can be negative as long as $\gamma + \theta < 0$. This condition can hold for realistic values of the parameters (i.e. for $0 < \gamma < 1$), and yields decreasing returns to scale phenomena, in contrast with the Cobb–Douglas technology for which $\rho_{kK} > 0$ always.

As predicted by Proposition 4, in this case we may encounter instability of the flat steady state, leading to potential agglomeration formation, if we have local negative externalities, which are stronger than a critical threshold. An estimate for

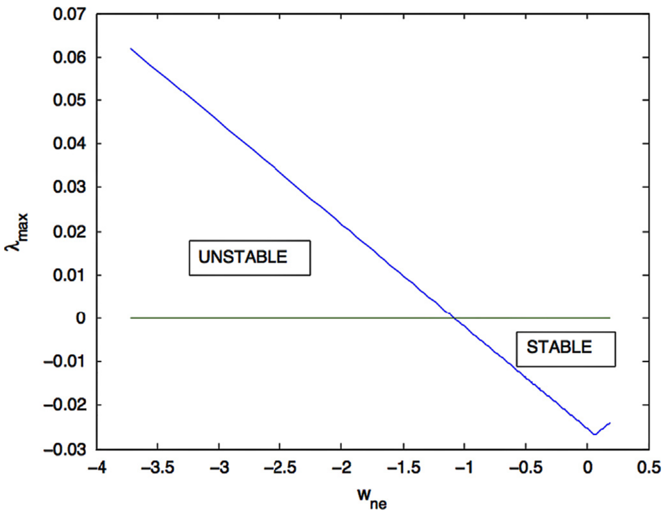

Fig. 13. Instability and the strength of the local negative externality.

this critical threshold is given in Proposition 4.[17] Note that the criterion of Proposition 4 is a sufficient criterion, meaning that if this criterion is met then we definitely have instability; however, we may still have instability in cases where this criterion is not met. Note also that the satisfaction of this criterion requires negative externalities *locally*, even on a single site. The rest of the matrix W may have positive entries, corresponding in the economic model to positive externalities, and most importantly, the average effect of the externalities as quantified by $\overline{w}$ can still be positive.

To illustrate the possibility of generation of agglomeration patterns in the case where $\rho_{kK}<0$ we choose the parameter values $A=1$, $\gamma=0.3$, $\beta=0.6$, $\theta=-0.8$ for the CES production function and consider a variety of kernels, scaled in such a way that $\overline{w}=1$. This yields $\rho=1$, $\rho_k=0.18$, $\rho_K=0.12$, $\rho_{kk}=-0.09$, $\rho_{KK}=-0.048$ and importantly $\rho_{kK}=-0.036<0$. The demand function is of the form $D(s)=Bs^{-\delta}$ for the choice $B=1$, $\delta=-0.5$, while $q=1$, $r=0.03$ and $\eta=0.02$. We consider a spatial economy of $N=15$ lattice sites. A straightforward application of Theorem 2 shows that in this case there exists a flat steady state for the PF-RECE at $\overline{k}_0=0.9176$. Since C_1, C_2, C_3 are negative, our analysis predicts that agglomeration is possible only when local negative externalities are present, i.e., when the matrix W has certain negative elements which are negative enough to drive part of the spectrum of the stability matrix T_0 to be positive. Note that to satisfy the transversality condition it will have to lie in the interval $[0,r^2/4]$. To make the effects of negative externalities on agglomeration evident, we choose a matrix W which is a perturbation of the kernel of Eq. (40) for the choice $A_1=2$, $a_1=0.0025$, $A_2=-1.2$ and $a_2=0.015$, for all lattice points apart from the case where $|i-j|=[N/2]+1$ (where $[\cdot]$ denotes the integer part). In such a case we perturb the matrix by choosing the relevant entry to be equal to w_-, which is allowed to take values which are positive or negative. The matrix is constructed so that it is symmetric, and normalized so that $\sum_j w_{ij}=\overline{w}>0$ and always equal to 1. Furthermore, the diagonal entries of the matrix are always positive.

Our theory predicts that we will not obtain the possibility of agglomeration unless there is strong negative externality, i.e., unless w_{ne} is less than a given (negative) threshold value. An estimate of the absolute value of this threshold is provided in Proposition 4 but it must be emphasized that this is an approximation of a sufficient criterion, therefore it is conceivable that patterns may occur even when the (absolute) value of the negative externality is smaller than the value predicted by this criterion. To show the mechanism of agglomeration as an effect of the interplay of positive and negative entries in the matrix W, we calculate numerically the top eigenvalue of the stability matrix T_0 as a function of the negative externality entry w_{ne} of the matrix W. In Fig. 13 we plot the value of the top eigenvalue of T_0 as a function of the strength of the possible negative externality w_{ne}. It is shown that if $w_{ne}>w_{cr}$ then the top eigenvalue of T_0 is negative, therefore the flat steady state is stable and no agglomeration phenomena are expected, whereas if $w_{ne}<w_{cr}$ then the top eigenvalue of T_0 becomes positive, thus leading to instability of the flat steady state, akin to Turing agglomeration instability. Fig. 13 can be treated as a bifurcation diagram for agglomeration in the system. In Fig. 14 we present a kernel profile that leads to potential agglomeration with decreasing returns to scale and $f_{kK}<0$.

7. Concluding remarks

We revisit the investment theory of a competitive firm in a spatial context where spatial externalities, which are regarded as a positive externality in the production function, are determined by spatial proximity of firms. We show that spatial

[17] There was no effort to make this estimate sharp; it is provided merely as an illustration that this phenomenon is indeed possible in the general case.

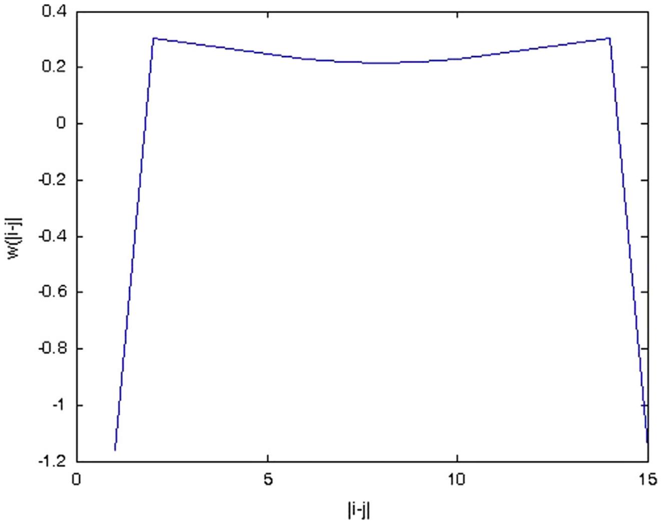

Fig. 14. A kernel profile leading to potential agglomeration.

agglomerations may emerge endogenously in a competitive industry where firms do not internalize spatial knowledge spillovers. The result does not require increasing returns either from the private or the social point of view, or location specific advantages at the location where the externality emerges, and does depend on boundary conditions since our spatial domain is a circle. In fact agglomeration at a PF-RECE is possible with decreasing returns to scale and positive or negative impact of the externality on the marginal product of private capital for an appropriate structure of the spatial interaction matrix. Potential agglomeration in a PF-RECE, when the externality is positively related to the marginal product of private capital, is driven by strong complementarity between the firms' stock of capital and the spatial externality, potential existence of positive and negative local spillovers but positive aggregate externality, and relatively large deviations between own and other-locations effects on the aggregate externality. These factors can be regarded as generalized centrifugal forces. Spatial agglomerations do not emerge as the SO when knowledge spillovers are internalized and the production function is characterized by strict concavity, i.e. we have diminishing returns both from the private and the social point of view. Agglomeration may emerge at the PF-RECE and the SO when the centrifugal forces are combined with diminishing returns from the private point of view but increasing returns from the social point of view.

Due to the well known complexity of spatial models, we try to obtain more insights through numerical experiments. Using a Cobb–Douglas production function and an isoelastic demand function, we show that agglomeration at the PF-RECE with diminishing returns from the private and social point of view may emerge when the global externality is positive but it consists of positive and negative components. Using a CES production function we show that agglomeration may emerge with decreasing returns to scale even when the externality has a negative impact on the marginal productivity of private capital. This result depends on the existence of strong negative local externality effects. Thus our numerical experiments confirm all our theoretical predictions about the potential emergence or not of agglomerations.

The deviation between the PF-RECE and the SO stemming from the fact that each firm neglects the impact of its own action on the aggregate externality suggests that, in the spirit of welfare analysis of models with externalities, a capital subsidy is required in order for the PF-RECE to reproduce the SO. This subsidy should be equal to $\psi_i(t) = \sum_{l=1}^{N} w_{li} f_K(k_i^*(t), \sum_{r=1}^{N} w_{ir} k_r^*(t))$ per unit of capital held by each firm at location i so that private and social marginal products are equal.

The results obtained in this paper suggest that agglomerations are possible as a long-run equilibrium outcome in a competitive industry with spatial spillovers and forward-looking agents. We think that the ability to study agglomeration emergence in the context of a full dynamic model with optimizing forward-looking agents is a reasonable trade-off for not taking into account some important features of new economic geography models, such as transport costs, product differentiation, mobile labor vs immobile "farmers" , or forward/backward linkages. Incorporating these aspects into our dynamic framework will bring our model closer to the economic geography models but will considerably increase its complexity. This is undoubtedly a direction for extension of our model.

Acknowledgments

The authors wish to thank two anonymous reviewers for valuable comments and suggestions on an earlier draft of the paper and Joan Stefan for her superb editorial work.

Appendix A. Proofs of stated results

A.1. A useful lemma

Lemma 1. *Assume that demand $D: \mathbb{R}_+ \to \mathbb{R}_+$ is a non-increasing function. Then the function $\mathbf{S}: \mathbb{R}^N \to \mathbb{R}$, defined by*

$$\mathbf{S}(k) := \int_0^{Q(k,Wk)} D(s)\, ds,$$

is a concave function of k as long as the production function f is a concave function.

Proof. We consider first the function $\overline{D}: \mathbb{R}_+ \to \mathbb{R}_+$, defined by $\overline{D}(x) := \int_0^x D(s)\, ds$. This is a concave function. Indeed, taking without loss of generality $x < (x+y)/2 < y$, we observe that

$$\frac{1}{2}(\overline{D}(x) + \overline{D}(y)) - \overline{D}\left(\frac{x+y}{2}\right) = \frac{1}{2}\left[\int_{(x+y)/2}^{y} D(s)\, ds - \int_x^{(x+y)/2} D(s)\, ds\right] \le 0,$$

since D is a non-increasing function. Therefore $\overline{D}$ is concave. If D is strictly non-increasing, then $\overline{D}$ is strictly concave.

Furthermore, by its definition, $\overline{D}$ is also an increasing function with respect to x. S^e is the composition of the concave and increasing function $\overline{D}$ with the function $Q^e: \mathbb{R}^N \to \mathbb{R}$, defined by $Q^e(k) = \sum_{i=1}^N f(k_i, K_i^e)$, which is clearly concave since the production function is assumed concave. Therefore S^e is a concave function of k as the composition of an increasing concave function with a concave function. Similarly, for $Q(k)$. This is the composition of the increasing and concave function $\overline{D}$, with $Q: \mathbb{R}^N \to \mathbb{R}$, where $Q(k) = f(k, Wk)$. Since Wk is a linear function and f is concave, Q is a concave function of k, therefore $\mathbf{S}$ is a concave function of k. □

A.2. Linearization for the Proof of Theorem 2

We linearize

$$F_i := D(Q(k, Wk))\left[f_k\left(k_i, \sum_r w_{ir} k_r\right) + \sigma \sum_\ell w_{\ell i} f_K\left(k_\ell, \sum_j w_{\ell j} k_j\right)\right] \tag{41}$$

around a homogeneous steady state $\overline{k}$. Note that $\overline{k}$ changes with σ, so we denote it as $\overline{k}_\sigma$ ($\sigma = 0$ or 1). Consider then $\kappa = \overline{k}_\sigma + \epsilon k$ (meaning that $\kappa_i = \overline{k}_\sigma + \epsilon k_i$ for every i). We do the linearization of the three terms involved separately:

(i) The term $D(Q)$: Since

$$Q(\kappa, W\kappa) = \sum_\ell f\left(\kappa_\ell, \sum_r w_{\ell r} \kappa_r\right) = \sum_\ell f\left(\overline{k}_\sigma + \epsilon k_\ell, \overline{w}\overline{k}_\sigma + \epsilon \sum_r w_{\ell r} k_r\right),$$

linearizing with respect to ϵ we get

$$\begin{aligned} Q(\kappa, W\kappa) &= \sum_\ell f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon \sum_\ell f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) k_\ell + \epsilon \sum_\ell f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_r w_{\ell r} k_r \\ &= N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_\ell k_\ell + \epsilon f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_r \sum_\ell w_{\ell r} k_r. \end{aligned}$$

For the last term we first perform the inner summation which yields that $\sum_\ell w_{\ell r} = \overline{w}$ for every r, so that

$$Q(\kappa, W\kappa) = N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_\ell k_\ell + \epsilon \overline{w} f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_r k_r,$$

which yields

$$Q(\kappa, W\kappa) = N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon (f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \overline{w} f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)) \sum_\ell k_\ell.$$

Therefore,

$$D(Q) = D(N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)) + \epsilon D'(N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma))(f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \overline{w} f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)) \sum_\ell k_\ell.$$

This is simplified, using the notation

$$D(Q) = A_0 + \epsilon A_1 \sum_\ell k_\ell$$

where

$$\begin{aligned} A_0 &= D(N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)), \\ A_1 &= D'(N f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma))(f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \overline{w} f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)). \end{aligned}$$

(ii) The term $f_k(k_i, \sum_r w_{ir} k_r)$: We have that

$$f_k\left(\kappa_i, \sum_r w_{ir}\kappa_r\right) = f_k\left(\overline{k}_\sigma + \epsilon k_i, \overline{w}\overline{k}_\sigma + \epsilon \sum_r w_{ir} k_r\right) = f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon f_{kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) k_i + \epsilon f_{kK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_r w_{ir} k_r.$$

This is expressed in the more compact notation

$$f_k\left(\kappa_i, \sum_r w_{ir}\kappa_r\right) = B_0 + \epsilon B_{11} k_i + \epsilon B_{12} \sum_r w_{ir} k_r$$

where

$$\begin{aligned} B_0 &= f_k(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \\ B_{11} &= f_{kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \\ B_{12} &= f_{kK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma). \end{aligned}$$

(iii) The term $\sigma \sum_\ell w_{\ell i} f_K(k_\ell, \sum_j w_{\ell j} k_j)$: We have that

$$\begin{aligned} \sigma \sum_\ell w_{\ell i} f_K\left(\kappa_\ell, \sum_j w_{\ell j}\kappa_j\right) &= \sigma \sum_\ell w_{\ell i} f_K\left(\overline{k}_\sigma + \epsilon k_\ell, \sum_j w_{\ell j}(\overline{k}_\sigma + \epsilon k_j)\right) \\ &= \sigma \sum_\ell w_{\ell i} f_K\left(\overline{k}_\sigma + \epsilon k_\ell, \overline{w}\overline{k}_\sigma + \epsilon \sum_j w_{\ell j} k_j\right) \\ &= \sigma \sum_\ell w_{\ell i} \left\{ f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon f_{Kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) k_\ell + \epsilon f_{KK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_j w_{\ell j} k_j \right\} \\ &= \sigma \overline{w} f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \epsilon\sigma f_{Kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_\ell w_{\ell i} k_\ell + \epsilon\sigma f_{KK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) \sum_\ell \sum_j w_{\ell i} w_{\ell j} k_j. \end{aligned}$$

This is expressed in the more compact notation

$$\sigma \sum_\ell w_{\ell i} f_K\left(\kappa_\ell, \sum_j w_{\ell j}\kappa_j\right) = C_0 + \epsilon C_{11} \sum_\ell w_{\ell i} k_\ell + \epsilon C_{12} \sum_\ell \sum_j w_{\ell i} w_{\ell j} k_j$$

where

$$\begin{aligned} C_0 &= \sigma \overline{w} f_K(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \\ C_{11} &= \sigma f_{Kk}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma), \\ C_{12} &= \sigma f_{KK}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma). \end{aligned}$$

We now calculate the linearization of F_i as

$$F_i = A_0(B_0 + C_0) + \epsilon(B_0 + C_0) A_1 \sum_\ell k_\ell + \epsilon A_0 B_{11} k_i + \epsilon A_0 (B_{12} + C_{11}) \sum_r w_{ir} k_r + \epsilon A_0 C_{12} \sum_\ell \sum_j w_{\ell i} w_{\ell j} k_j.$$

A.3. Proof of Proposition 2

We work in the setting of Theorem 2 setting $\sigma = 0$.

(i) If $\gamma < 1$ then the function $D(s)s^{(\gamma-1)/\gamma}$ is strictly decreasing, and if condition (32) holds then by standard continuity arguments the scalar algebraic equation defining the steady state admits a unique solution $\overline{k}_0$. The rest follows by routine application of Theorem 2 for $\sigma = 0$.

(ii) If $\gamma > 1$, the function $D(s)s^{(\gamma-1)/\gamma}$ is not necessarily strictly decreasing. Calculating the derivative of this function we note that for $s > 0$, this function is strictly decreasing if the condition on the elasticity of demand holds. The rest of the proof follows as in (i).

A.4. Proof of Proposition 3

We work in the setting of Theorem 2 setting $\sigma = 0$. We observe that the matrix T_0 consists of three contributions. The first one is diagonal $T_{0,1} = C_1 I$, and furthermore, $C_1 < 0$ always. The second contribution is $T_{0,2} = C_2 \mathbf{1}$ and $C_2 < 0$ always. Therefore, the matrix $T_{0,1} + T_{0,2}$ consists of negative elements. The first two contributions depend on the matrix W only through $\overline{w}$ (and we assume that overall externalities are positive in the sense that $\overline{w} > 0$). The third contribution to this matrix is fundamentally different: it is $T_{0,3} = C_3 W$ and this may contain positive or negative contributions depending on the particular elements on the matrix W, w_{ij}. For a composite kernel combining positive and negative spatial externalities, some elements w_{ij} may be positive and some may be negative. If

$$T_{0,ii} := C_1 + C_2 + C_3 w_{ii} \le 0, \quad \forall i \in \mathcal{N}$$

$$T_{0,ij} := C_2 + C_3 w_{ij} \geq 0, \quad \forall i,j \in \mathcal{N}, \ \ i \neq j,$$

then the matrix T_0 is a Metzler matrix and this provides detailed information concerning its stability properties. In particular, a Metzler matrix is asymptotically stable if and only if its diagonal elements are negative. Therefore, if W and the other fundamentals of the system are such that the two above inequalities hold, the first with strict inequality, then all the eigenvalues of T_0 have real parts which are negative, and by Theorem 2 we expect no agglomerations. What remains is to check the validity of the above conditions.

Using the definition of the terms C_1, C_2, C_3, the off-diagonal terms are expressed as

$$T_{0,ij} = \overline{k}_0^{\gamma-2} D\left(\rho N \overline{k}_0^{\gamma}\right) \left(\rho_k (\rho_k + \overline{w} \rho_K) \overline{k}_0^{\gamma} \frac{D'(\rho N \overline{k}_0^{\gamma})}{D(\rho N \overline{k}_0^{\gamma})} + \rho_{kK} w_{ij} \right)$$

and since $\overline{k}_0^{\gamma-2} D(\rho N \overline{k}_0^{\gamma}) > 0$, the off-diagonal terms will have the same sign as

$$I_{ij} := \rho_k (\rho_k + \overline{w} \rho_K) \overline{k}_0^{\gamma} \frac{D'(\rho N \overline{k}_0^{\gamma})}{D(\rho N \overline{k}_0^{\gamma})} + \rho_{kK} w_{ij}.$$

The first term of this sum is clearly negative, so I_{ij} can be positive if the second term is sufficiently large and positive. Since in principle we allow $\rho_{kK} > 0$, this implies that w_{ij} is sufficiently large. Using the notation $s = \rho N \overline{k}_0^{\gamma}$ and employing Assumption 3,

$$\frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \underline{E}_D + \rho_{kK} w_{ij} \leq I_{ij} \leq \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \overline{E}_D + \rho_{kK} w_{ij}.$$

A similar calculation allows us to express the diagonal terms as

$$T_{0,ii} = \overline{k}_0^{\gamma-2} D\left(\rho N \overline{k}_0^{\gamma}\right) \left(\rho_{kk} + \rho_k (\rho_k + \overline{w} \rho_K) \overline{k}_0^{\gamma} \frac{D'(\rho N \overline{k}_0^{\gamma})}{D(\rho N \overline{k}_0^{\gamma})} + \rho_{kK} w_{ii} \right),$$

and since $\overline{k}_0^{\gamma-2} D(\rho N \overline{k}_0^{\gamma}) > 0$, the diagonal terms will have the same sign as

$$I_{ii} := \rho_{kk} + \rho_k (\rho_k + \overline{w} \rho_K) \overline{k}_0^{\gamma} \frac{D'(\rho N \overline{k}_0^{\gamma})}{D(\rho N \overline{k}_0^{\gamma})} + \rho_{kK} w_{ii}.$$

Using the notation $s = \rho N \overline{k}_0^{\gamma}$ and employing Assumption 3,

$$\rho_{kk} + \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \underline{E}_D + \rho_{kK} w_{ii} \leq I_{ii} \leq \rho_{kk} + \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \overline{E}_D + \rho_{kK} w_{ii}.$$

The stability matrix is a stable Metzler matrix if $T_{ij} \geq 0$ and $T_{ii} < 0$. By the above estimates, this will happen if

$$0 \leq \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \underline{E}_D + \rho_{kK} w_{ij},$$
$$\rho_{kk} + \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \overline{E}_D + \rho_{kK} w_{ii} < 0.$$

On the other hand, if $T_{ij} \geq 0$ and $T_{ii} \geq 0$ then T_0 is a positive matrix . This will happen if the fundamentals of the economy are such that

$$0 \leq \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \underline{E}_D + \rho_{kK} w_{ij},$$
$$\rho_{kk} + \frac{\rho_k}{\rho N} (\rho_k + \overline{w} \rho_K) \underline{E}_D + \rho_{kK} w_{ii} \geq 0.$$

If the above conditions hold, then according to the Perron–Frobenious theorem, T_0 has a real maximal eigenvalue λ^*, which is positive. In particular we have an estimate for this eigenvalue in terms of $\min_i \sum_j T_{0,ij} \leq \lambda^* \leq \max_i \sum_j T_{0,ij}$. We can easily see that

$$\sum_j T_{0,ij} = \overline{k}_0^{\gamma-2} D(\rho N \overline{k}_0^{\gamma}) \left(I_{ii} + \sum_{j \neq i} I_{ij} \right) =: \overline{k}_0^{\gamma-2} D(\rho N \overline{k}_0^{\gamma}) J_i,$$

and

$$J_i := I_{ii} + \sum_{j \neq i} I_{ij} = \rho_{kk} + \frac{\rho_k}{\rho} (\rho_k + \overline{w} \rho_K) \frac{s D'(s)}{D(s)} + \rho_{kK} \sum_j w_{ij} = \rho_{kk} + \frac{\rho_k}{\rho} (\rho_k + \overline{w} \rho_K) \frac{s D'(s)}{D(s)} + \rho_{kK} \overline{w}.$$

This term is independent of i so in fact the maximal eigenvalue is equal to

$$\lambda^* = \overline{k}_0^{\gamma-2} D\left(\rho N \overline{k}_0^{\gamma}\right) \left(\rho_{kk} + \frac{\rho_k}{\rho} (\rho_k + \overline{w} \rho_K) \frac{s D'(s)}{D(s)} + \rho_{kK} \overline{w} \right)$$

where $s=\rho N\overline{k}_0^{\gamma}$, or recalling the definition of the spatially homogeneous steady state, $\overline{k}_0$,

$$\lambda^* = \frac{M}{\overline{k}_0\rho_k}\left(\rho_{kk}+\frac{\rho_k}{\rho}(\rho_k+\overline{w}\rho_K)\frac{sD'(s)}{D(s)}+\rho_{kK}\overline{w}\right).$$

The above estimate allows us to check the condition $\lambda^*\in(0,r^2/4]$, which by Theorem 2 is the condition for occurrence of agglomerations in the PF-RECE case. For example, one easily understands from this estimate that we expect instability if the kernel is such that

$$H := \rho_{kk}+\frac{\rho_k}{\rho}(\rho_k+\overline{w}\rho_K)\underline{E}_D+\rho_{kK}\overline{w}>0$$

which, taking into account the negativity of ρ_{kk} and $\underline{E}_D$ and the positivity of ρ_{kK}, is a condition on largeness of $\overline{w}$. To see this, rearrange the above term as

$$H=\rho_{kk}+\frac{\rho_k^2}{\rho}\underline{E}_D+\left(\frac{\rho_k\rho_K}{\rho}\underline{E}_D+\rho_{kK}\right)\overline{w}.$$

If the overall effect of the externalities is positive, then H can be positive if $(\rho_k\rho_K/\rho)\underline{E}_D+\rho_{kK}>0$ and if $\overline{w}$ is large enough. The conditions for instability in this case are

$$V := \frac{\rho_k\rho_K}{\rho}\underline{E}_D+\rho_{kK}>0,$$

$$\overline{w}>\frac{-\rho_{kk}-\frac{\rho_k^2}{\rho}\underline{E}_D}{V}=\frac{-\rho\rho_{kk}-\rho_k^2\underline{E}_D}{\rho_k\rho_K\underline{E}_D+\rho\rho_{kK}}.$$

A.5. Proof of Proposition 4

Recall the following result which holds for real symmetric matrices. Let A be a symmetric matrix. This has real eigenvalues which can be ordered from the minimum to the maximum one and let λ_{max} and λ_{min} be the maximum and the minimum eigenvalues. Consider the Rayleigh quotient

$$R_A(x)=\frac{(Ax,x)}{(x,x)},\quad \forall x\in\mathbb{R}^N.$$

Then we have the following theorem (Rayleigh–Ritz)

$$\lambda_{max}=\max_{x\neq 0}\frac{(Ax,x)}{(x,x)}=\max_{(x,x)=1}(Ax,x),$$

$$\lambda_{min}=\min_{x\neq 0}\frac{(Ax,x)}{(x,x)}=\min_{(x,x)=1}(Ax,x),$$

which implies that $\lambda_{max}\geq (Ay,y)/(y,y)$ for any $y\in\mathbb{R}^N$. This provides a variety of bounds for the top eigenvalue of varying sharpness, by using different choices of y. For example if $y=e_i$ is the unit vector in the i direction, then we see that $\lambda_{max}\geq a_{ii}$ and this is true for any $i=1,\ldots,N$. If $y=(1,1,\ldots,1)^{tr}$ then $(y,y)=N$ and the Rayleigh–Ritz characterization gives $\lambda_{max}\geq(1/N)\sum_i\sum_j a_{ij}$. A great variety of other choices in between can be used.

This approach can be used to provide lower bounds for the maximum eigenvalue of the stability matrix T_σ. If we manage to show that $\lambda_{max}>0$ then know that we have instability, so it is enough to show that by the Rayleigh–Ritz characterization we may obtain a strictly positive lower bound. Since $(y,y)\geq 0$ for any $y\in\mathbb{R}^N$ it is enough to consider the numerator $(T_\sigma y,y)=\sum_i\sum_j T_{\sigma,ij}y_iy_j$ for any $y=(y_1,\ldots,y_N)^{tr}$. Using the definition of the matrix T_σ we have that

$$\lambda_{max}\geq C_1\sum_i y_i^2+C_2\sum_i\sum_j y_iy_j+C_3\sum_i\sum_j w_{ij}y_iy_j+C_4\sum_i\sum_j w_{ij}^{(2)}y_iy_j$$

which is true for every $y=(y_1,\ldots,y_N)\in\mathbb{R}^N$ where $w_{ij}^{(2)}=\sum_k w_{ik}w_{kj}$. If we find a choice $y\in\mathbb{R}^N$ such that the term of the right-hand side is strictly positive, then we have that $\lambda_{max}>0$ and we are done! Note that we are allowed to choose vectors y with negative entries. Also note that if $\sigma=0$ then $C_4=0$.

So we focus on the quantity

$$\begin{aligned}I(y) &:= C_1\sum_i y_i^2+C_2\sum_i\sum_j y_iy_j+C_3\sum_i\sum_j w_{ij}y_iy_j+C_4\sum_i\sum_j w_{ij}^{(2)}y_iy_j\\ &= C_1\sum_i y_i^2+C_2\left(\sum_i y_i\right)^2+C_3\sum_i\sum_j w_{ij}y_iy_j+C_4\sum_i\sum_j w_{ij}^{(2)}y_iy_j.\end{aligned}$$

and our aim is to show that there exists $y\in\mathbb{R}^N$ such that $I(y)>0$.

To this end let w_{nm} be the most negative element of matrix W, i.e., $w_{nm} = \min(w_{ij}, i \neq j)$ and assume that $w_{nm} < 0$. Choose $y \in \mathbb{R}^N$ such that $y_i = 0$ if $i \neq n, m$ and $y_n = y_m = \frac{1}{2}$. We now calculate $I(y)$ for this choice of y. This yields

$$I = \tfrac{1}{2}C_1 + C_2 + \tfrac{1}{4}C_3(w_{nn} + 2w_{nm} + w_{mm}) + \tfrac{1}{4}C_4\left(w_{nn}^{(2)} + 2w_{nm}^{(2)} + w_{mm}^{(2)}\right).$$

Note that C_1 and C_2 are negative. Furthermore, since we concentrate on the case where $\rho_{kK} < 0$, we also have that $C_3 < 0$. Finally, in the PF-RECE case we have $\sigma = 0$ it follows that $C_4 = 0$. We rewrite I in such a way as to make clear the signs of the various contributions, i.e. using $C_1 = -|C_1|$, $C_2 = -|C_2|$, $C_3 = -|C_3|$, $w_{nm} = -|w_{nm}|$, $w_{nn} = |w_{nn}|$ and $w_{mm} = |w_{mm}|$. Doing that yields

$$I = \frac{|C_3|}{2}\left[|w_{nm}| - \frac{1}{2}(w_{nn} + w_{mm}) - \frac{|C_1|}{|C_3|} - 2\frac{|C_2|}{|C_3|}\right]$$

and $I > 0$ as long as

$$|w_{nm}| > \frac{1}{2}(w_{nn} + w_{mm}) + \frac{|C_1|}{|C_3|} + 2\frac{|C_2|}{|C_3|}.$$

This condition can be simplified even further to

$$|w_{nm}| > \frac{1}{2}(w_{nn} + w_{mm}) + \frac{\rho_{kk}}{\rho_{kK}} + \frac{(\rho_k + \overline{w}\rho_K)\rho_k}{\rho_{kK}} \frac{D'(\rho N \overline{k}_0^\gamma)}{D(\rho N \overline{k}_0^\gamma)} \overline{k}_0^\gamma.$$

A.6. Proof of Proposition 5

We work in the setting of Theorem 2 setting $\sigma = 1$, and apply arguments similar to the ones used in the proof of Proposition 2 (ii). The details are omitted.

A.7. Proof of Proposition 6

We work in the setting of Theorem 2 setting $\sigma = 1$. The off-diagonal elements of the matrix T_1 are

$$T_{1,ij} = \frac{1}{\alpha}\overline{k}_1^{\gamma-2} D(s)\left((\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\frac{sD'(s)}{D(s)} + 2\rho_{kK} w_{ij} + \rho_{KK}\sum_r w_{ir} w_{rj}\right),$$

whereas the diagonal terms are

$$T_{1,ii} = \frac{1}{\alpha}\overline{k}_1^{\gamma-2} D(s)\left(\rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\frac{sD'(s)}{D(s)} + 2\rho_{kK} w_{ii} + \rho_{KK}\sum_r w_{ir} w_{ri}\right)$$

where $s = \rho N \overline{k}_1^\gamma$. Since $\overline{k}_1^{\gamma-2} D(s) > 0$, for every $\overline{k}_1 > 0$, the signs of the terms $T_{1,ij}$, $T_{1,ii}$ coincide with the signs of the terms I_{ij}, I_{ii} respectively, where

$$I_{ij} := (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\frac{sD'(s)}{D(s)} + 2\rho_{kK} w_{ij} + \rho_{KK}\sum_r w_{ir} w_{rj}, \quad i \neq j,$$
$$I_{ii} := \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\frac{sD'(s)}{D(s)} + 2\rho_{kK} w_{ii} + \rho_{KK}\sum_r w_{ir} w_{ri}.$$

Similar arguments to those used in Proposition 6 provide us with lower and upper bounds for these terms, in particular,

$$(\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\underline{E}_D + 2\rho_{kK} w_{ij} + \rho_{KK}\sum_r w_{ir} w_{rj} \le I_{ij}$$
$$\le (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\overline{E}_D + 2\rho_{kK} w_{ij} + \rho_{KK}\sum_r w_{ir} w_{rj},$$

and

$$\rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\underline{E}_D + 2\rho_{kK} w_{ii} + \rho_{KK}\sum_r w_{ir} w_{ri} \le I_{ii}$$
$$\le \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\overline{E}_D + 2\rho_{kK} w_{ii} + \rho_{KK}\sum_r w_{ir} w_{ri}.$$

The matrix T_1 is a Hurwitz stable Metzler matrix if $T_{1,ij} \ge 0$ and $T_{1,ii} < 0$. Using the above bounds we see that this is the case if

$$0 \le (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\underline{E}_D + 2\rho_{kK} w_{ij} + \rho_{KK}\sum_r w_{ir} w_{rj},$$
$$\rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N}\overline{E}_D + 2\rho_{kK} w_{ii} + \rho_{KK}\sum_r w_{ir} w_{ri} \le 0.$$

If the above conditions are true, the spectrum of the matrix T_1 is negative, and using Theorem 2 no agglomeration patterns are expected to occur.

If $T_{1,ij} \geq 0$ and $T_{ii} \geq 0$, then T_1 is a positive matrix and using the Perron–Frobenius theorem the top eigenvalue is positive, therefore by Theorem 2 we expect the emergence of agglomeration patterns. Using the lower bounds obtained we see that this will happen if

$$0 \leq (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N} \underline{E}_D + 2\rho_{kK} w_{ij} + \rho_{KK} \sum_r w_{ir} w_{rj},$$
$$0 \geq \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho N} \underline{E}_D + 2\rho_{kK} w_{ii} + \rho_{KK} \sum_r w_{ir} w_{ri}.$$

We may furthermore estimate the top eigenvalue using the estimate $\min_i \sum_j T_{1,ij} \leq \lambda^* \leq \max_i \sum_j T_{1,ij}$. Some algebra yields that

$$\sum_j T_{1,ij} = \frac{1}{\alpha} \overline{k}_1^{\gamma-2} D\left(\rho N \overline{k}_1^{\gamma}\right) \sum_j I_{ij},$$

and

$$\sum_j I_{ij} = \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho} \frac{sD'(s)}{D(s)} + 2\rho_{kK}\overline{w} + \rho_{KK}\left(\sum_{j \neq i} \sum_r w_{ir} w_{rj} + \sum_r w_{ir} w_{ir}\right).$$

Note that

$$\left(\sum_{j \neq i} \sum_r w_{ir} w_{rj} + \sum_r w_{ir} w_{ir}\right) = \sum_j \sum_r w_{ir} w_{rj} \quad = \sum_j \left(\sum_r w_{ir} w_{jr}\right) = \sum_j w_{jr} \left(\sum_r w_{ir}\right) = \overline{w} \sum_j w_{jr} = \overline{w}^2,$$

by Assumption 2, so that

$$\sum_j I_{ij} = \rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho} \frac{sD'(s)}{D(s)} + 2\rho_{kK}\overline{w} + \rho_{KK}\overline{w}^2.$$

As this is independent of i, we obtain the maximal eigenvalue as

$$\lambda^* = \frac{1}{\alpha} \overline{k}_1^{\gamma-2} D\left(\rho N \overline{k}_1^{\gamma}\right) \left(\rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho} \frac{sD'(s)}{D(s)} + 2\rho_{kK}\overline{w} + \rho_{KK}\overline{w}^2\right),$$

which, keeping in mind the definition of $\overline{k}_1$, simplifies to

$$\lambda^* = \frac{1}{\alpha} \frac{M}{(\rho_k + \overline{w}\rho_K)\overline{k}_1} \left(\rho_{kk} + (\rho_k + \overline{w}\rho_K)^2 \frac{1}{\rho} \frac{sD'(s)}{D(s)} + 2\rho_{kK}\overline{w} + \rho_{KK}\overline{w}^2\right).$$

This expression allows us to check whether the top eigenvalue is less than $\frac{r^2}{4}$, as long as the steady state $\overline{k}_1$ is known, and this is usually obtained by a very straightforward calculation (even if we need to calculate it numerically, it only involves the solution of a single algebraic equation).

A.8. Details on the Cobb–Douglas example

For the Cobb–Douglas production function we have

$$f(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma},$$
$$f_1(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma_1 C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-1},$$
$$f_2(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma_2 C\overline{w}^{\gamma_2-1}\overline{k}_\sigma^{\gamma-1},$$
$$\overline{w}f_2(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma_2 C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-1},$$
$$f_1(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) + \overline{w}f_2(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-1},$$
$$f_{11}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma_1(\gamma_1-1)C\overline{k}_\sigma^{\gamma_1-2}\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma_2} = \gamma_1(\gamma_1-1)C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-2},$$
$$f_{12}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma_1\gamma_2 C\overline{k}_\sigma^{\gamma_1-1}\overline{w}^{\gamma_2-1}\overline{k}_\sigma^{\gamma_2-1} = \gamma_1\gamma_2 C\overline{w}^{\gamma_2-1}\overline{k}_\sigma^{\gamma-2},$$
$$f_{22}(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma) = \gamma_2(\gamma_2-1)C\overline{k}_\sigma^{\gamma_1}\overline{w}^{\gamma_2-2}\overline{k}_\sigma^{\gamma_2-2} = \gamma_2(\gamma_2-1)C\overline{w}^{\gamma_2-2}\overline{k}_\sigma^{\gamma-2}.$$

Using the isoelastic demand function D we obtain

$$D(Nf(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)) = BC^{-\delta}\overline{w}^{-\delta\gamma_2}\overline{k}_\sigma^{-\delta\gamma}N^{-\delta},$$
$$D'(Nf(\overline{k}_\sigma, \overline{w}\overline{k}_\sigma)) = -\delta BC^{-(1+\delta)}\overline{w}^{-(1+\delta)\gamma_2}\overline{k}_\sigma^{-(1+\delta)\gamma}N^{-(1+\delta)}.$$

Therefore

$$A_0 = BC^{-\delta}\overline{w}^{-\delta\gamma_2}\overline{k}_\sigma^{-\delta\gamma}N^{-\delta},$$

$$
\begin{aligned}
A_1 &= -\delta\gamma B C^{-\delta}\overline{w}^{-\delta\gamma_2}\overline{k}_\sigma^{-(1+\delta\gamma)}N^{-(1+\delta)},\\
B_0 &= \gamma_1 C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-1},\\
B_{11} &= \gamma_1(\gamma_1-1)C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-2},\\
B_{12} &= \gamma_1\gamma_2\overline{w}^{\gamma_2-1}C\overline{k}_\sigma^{\gamma-2},\\
C_0 &= \sigma\gamma_2 C\overline{w}^{\gamma_2}\overline{k}_\sigma^{\gamma-1},\\
C_{11} &= \sigma\gamma_1\gamma_2 C\overline{w}^{\gamma_2-1}\overline{k}_\sigma^{\gamma-2},\\
C_{12} &= \sigma\gamma_2(\gamma_2-1)C\overline{w}^{\gamma_2-2}\overline{k}_\sigma^{\gamma-2}
\end{aligned}
$$

and

$$
\begin{aligned}
A_0B_{11} &= \gamma_1(\gamma_1-1)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-\delta},\\
A_1(B_0+C_0) &= -\delta\gamma(\gamma_1+\sigma\gamma_2)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-(1+\delta)},\\
A_0(B_{12}+C_{11}) &= (1+\sigma)\gamma_1\gamma_2 BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)-1}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-\delta},\\
A_0C_{12} &= \sigma\gamma_2(\gamma_2-1)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)-2}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-\delta}.
\end{aligned}
$$

Therefore the stability matrix is of the form

$$T_\sigma = C_1 I + C_2\mathbf{1} + C_3 W + C_4 W^2,$$

where

$$
\begin{aligned}
C_1 &= \frac{1}{\alpha}\gamma_1(\gamma_1-1)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-\delta},\\
C_2 &= -\frac{1}{\alpha}\delta\gamma(\gamma_1+\sigma\gamma_2)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-(1+\delta)},\\
C_3 &= \frac{1}{\alpha}(1+\sigma)\gamma_1\gamma_2 BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)-1}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-\delta},\\
C_4 &= \frac{1}{\alpha}\sigma\gamma_2(\gamma_2-1)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)-2}\overline{k}_\sigma^{\gamma(1-\delta)-2}N^{-\delta}.
\end{aligned}
$$

This fully characterizes the stability matrix T_σ in both cases ($\sigma=0$ the RE case, $\sigma=1$ the SO case) in terms of the homogeneous steady state $\overline{k}_\sigma$. However, since $\overline{k}_\sigma$ depends on the fundamentals of the economy (the parameters of the system), it is best at this stage to explicitly calculate $\overline{k}_\sigma$ in terms of the fundamentals of the economy, substitute in the above expressions and thus obtain the matrix T_σ purely in terms of the parameters of the system.

To perform this calculation, recall that the steady state $\overline{k}_\sigma$ is given by the solution of the equation

$$A_0(B_0+C_0)-M=0$$

where $M=q(\rho+\eta)$. We see that

$$A_0(B_0+C_0) = (\gamma_1+\sigma\gamma_2)BC^{1-\delta}\overline{w}^{\gamma_2(1-\delta)}\overline{k}_\sigma^{-1+\gamma(1-\delta)},$$

so that the steady state is given by

$$\overline{k}_\sigma = \left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{1/(-1+\gamma(1-\delta))} B^{-1/(-1+\gamma(1-\delta))}C^{-(1-\delta)/(-1+\gamma(1-\delta))}\overline{w}^{-\gamma_2(1-\delta)/(-1+\gamma(1-\delta))}N^{\delta/(-1+\gamma(1-\delta))},$$

or in terms of $\rho_1=-1+\gamma(1-\delta)$,

$$\overline{k}_\sigma = \left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{1/\rho_1} B^{-1/\rho_1}C^{-(1-\delta)/\rho_1}\overline{w}^{-\gamma_2(1-\delta)/\rho_1}N^{\delta/\rho_1}. \tag{42}$$

We now substitute expression (42) for $\overline{k}_\sigma$ into the stability matrix, to obtain the final form in terms of the parameters of the system only. This gives (upon collecting all similar terms)

$$T_\sigma = C_1 I + C_2\mathbf{1} + C_3 W + C_4 W^2,$$

where

$$
\begin{aligned}
C_1 &= \frac{1}{\alpha}\gamma_1(\gamma_1-1)\left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{(\rho_1-1)/\rho_1} B^{1/\rho_1}C^{(1-\delta)/\rho_1}\overline{w}^{\gamma_2(1-\delta)/\rho_1}N^{-\delta/\rho_1},\\
C_2 &= -\frac{1}{\alpha}\delta\gamma(\gamma_1+\sigma\gamma_2)\left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{(\rho_1-1)/\rho_1} B^{1/\rho_1}C^{(1-\delta)/\rho_1}\overline{w}^{\gamma_2(1-\delta)/\rho_1}N^{(1-\delta)(1-\gamma)/\rho_1},
\end{aligned}
$$

$$C_3=\frac{1}{\alpha}(1+\sigma)\gamma_1\gamma_2\left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{(\rho_1-1)/\rho_1}B^{1/\rho_1}C^{(1-\delta)/\rho_1}\overline{w}^{\gamma_2(1-\delta)/\rho_1-1}N^{-\delta/\rho_1},$$
$$C_4=\frac{1}{\alpha}\sigma\gamma_2(\gamma_2-1)\left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{(\rho_1-1)/\rho_1}B^{1/\rho_1}C^{(1-\delta)/\rho_1}\overline{w}^{\gamma_2(1-\delta)/\rho_1-2}N^{-\delta/\rho_1},$$

and $\rho_1=-1+\gamma(1-\delta)$.

We finally note that the matrix T_σ can be expressed as

$$T_\sigma=\tau\{\gamma_1(\gamma_1-1)I-\delta\gamma(\gamma_1+\sigma\gamma_2)N^{-1}\mathbf{1}+(1+\sigma)\gamma_1\gamma_2\overline{w}^{-1}W+\sigma\gamma_2(\gamma_2-1)\overline{w}^{-2}W^2\},$$

where

$$\tau:=\frac{1}{\alpha}\left(\frac{M}{\gamma_1+\sigma\gamma_2}\right)^{(\rho_1-1)/\rho_1}B^{1/\rho_1}C^{(1-\delta)/\rho_1}\overline{w}^{\gamma_2(1-\delta)/\rho_1}N^{-\delta/\rho_1}.$$

We now have everything in terms of the primitives of the model, and by using numerical algebra techniques we can find the spectrum of the matrix T, and find parameter values for which agglomeration occurs. The patterns that occur can be obtained from the spatial form of the relevant eigenvectors.

In closing, we provide explicit forms for the stability criterion of Propositions 3 and 6. For the Cobb–Douglas utility function,

$$\rho=\overline{w}^{\gamma_2},\quad \rho_k=\gamma_1\overline{w}^{\gamma_2},\quad \rho_K=\gamma_2\overline{w}^{\gamma_2-1},\quad \rho_k+\overline{w}\rho_K=\gamma\overline{w}^{\gamma_2},$$
$$\rho_{kk}=\gamma_1(\gamma_1-1)\overline{w}^{\gamma_2},\quad \rho_{kK}=\gamma_1\gamma_2\overline{w}^{\gamma_2-1},\quad \rho_{KK}=\gamma_2(\gamma_2-1)\overline{w}^{\gamma_2-2}.$$

Therefore,

$$-\frac{\rho_{kk}}{\rho_{kK}}=\frac{1-\gamma_1}{\gamma_2}\overline{w},\quad \frac{1}{\rho N}\frac{\rho_k}{\rho_{kK}}(\rho_k+\overline{w}\rho_K)=\frac{\gamma}{N\gamma_2}\overline{w},$$

so that the stability criterion for PF-RECE becomes

$$w_{ii}<\frac{1-\gamma_1}{\gamma_2}\overline{w}+\frac{\delta\,\gamma}{N\gamma_2}\overline{w},$$
$$w_{ij}>\frac{\delta\,\gamma}{N\gamma_2}\overline{w}.$$

References

Baldwin, R.E., Martin, P., 2004. Agglomeration and regional growth. Handbook of Regional and Urban Economics, vol. 4, pp. 2671–2711.

Bewley, T., 1982. An integration of equilibrium theory and turnpike theory. J. Math. Econ. 10 (2), 233–267.

Boucekkine, R., Camacho, C., Fabbri, G., 2013a. On the optimal control of some parabolic partial differential equations arising in economics. AMSE, WP 2013-Nr 34.

Boucekkine, R., Camacho, C., Fabbri, G., 2013b. Spatial dynamics and convergence: the spatial ak model. CNRS, 2013.47.

Boucekkine, R., Camacho, C., Zou, B., 2009. Bridging the gap between growth theory and the new economic geography: the spatial Ramsey model. Macroecon. Dyn. 13 (1), 20.

Breinlich, H., Ottaviano, G.I., Temple, J.R., 2013. Regional growth and regional decline. CEPR, Discussion Paper No. DP9568.

Breschi, S., Lissoni, F., 2001. Localised knowledge spillovers vs. innovative milieux: knowledge "tacitness" reconsidered. Pap. Reg. Sci. 80 (3), 255–273.

Brito, P., 2011. Global endogenous growth and distributional dynamics. MPRA Paper No. 41653.

Brock, W., Xepapadeas, A., 2008. Diffusion-induced instability and pattern formation in infinite horizon recursive optimal control. J. Econ. Dyn. Control 32 (9), 2745–2787.

Brock, W., Xepapadeas, A., 2010. Pattern formation, spatial externalities and regulation in coupled economic–ecological systems. J. Environ. Econ. Manag. 59 (2), 149–164.

Brock, W.A., 1974. Money and growth: the case of long run perfect foresight. Int. Econ. Rev. 15 (3), 750–777.

Brock, W.A., Scheinkman, J.A., 1977. On the long-run behaviour of a competitive firm. In: Schwodiauer, G. (Ed.), Equilibrium and Disequilibrium in Economic Theory, D. Reidel Publishing Company, Dordrecht-Holland.

Camacho, C., Zou, B., 2004. The spatial Solow model. Econ. Bull. 18, 1–11.

Camacho, C., Zou, B., Briani, M., 2008. On the dynamics of capital accumulation across space. Eur. J. Oper. Res. 186 (2), 451–465.

Chincarini, L., Asherie, N., 2008. An analytical model for the formation of economic clusters. Reg. Sci. Urban Econ. 38 (3), 252–270.

Feldman, M.P., 1999. The new economics of innovation, spillovers and agglomeration: a review of empirical studies. Econ. Innov. New Technol. 8 (1–2), 5–25.

Grossman, G.M., Helpman, E., 1991. Trade, knowledge spillovers, and growth. Eur. Econ. Rev. 35 (2), 517–526.

Krugman, P., 1996. The Self-organizing Economy Mitsui, Lectures in Economics. Wiley-Blackwell, Cambridge, MA, USA.

Krugman, P.R., 1991. Geography and Trade. MIT Press, Cambridge, MA.

Kurz, M., 1968. The general instability of a class of competitive growth processes. Rev. Econ. Stud. 35 (2).

Lucas, R.E., 2001. Externalities and cities. Rev. Econ. Dyn. 4 (2), 245–274.

Lucas, R.E., Rossi-Hansberg, E., 2002. On the internal structure of cities. Econometrica 70 (4), 1445–1476.

Lucas Jr., R.E., Prescott, E.C., 1971. Investment under uncertainty. Econom.: J. Econom. Soc., 659–681.

Nocco, A., 2005. The rise and fall of regional inequalities with technological differences and knowledge spillovers. Reg. Sci. Urban Econ. 35 (5), 542–569.

Papageorgiou, Y.Y., Smith, T.R., 1983. Agglomeration as local instability of spatially uniform steady-states. Econom.: J. Econom. Soc., 1109–1119.

Prager, J.-C., Thisse, J.-F., 2012. Economic Geography and the Unequal Development of Regions, vol. 59. Routledge, New York.
Quah, D., 2002. Spatial agglomeration dynamics. Am. Econ. Rev. 92 (2), 247–252.
Romer, P.M., 1986. Increasing returns and long-run growth. J. Polit. Econ., 1002–1037.
Rouhani, B.D., Khatibzadeh, H., 2009. Asymptotic behavior of bounded solutions to a class of second order nonhomogeneous evolution equations. Nonlinear Anal.: Theory Methods Appl. 70 (12), 4369–4376.
Scheinkman, J.A., 1978. Stability of separable Hamiltonians and investment theory. Rev. Econ. Stud. 45 (3), 559–570.
Turing, A.M., 1952. The chemical basis of morphogenesis. Philos. Trans. R. Soc. London. Ser. B Biol. Sci. 237 (641), 37–72.