The Emergence of Optimal Agglomeration in Dynamic Economics

William Brock and Anastasios Xepapadeas

Department of Economics, University of Wisconsin, and Beijer Fellow
wbrock@ssc.wisc.edu
and
Department of International and European Economic Studies,
Athens University of Economics and Business, and Beijer Fellow
xepapad@aueb.gr

Abstract

We analyze endogenous pattern formation resulting from forward-looking optimizing behavior of economic agents in the presence of spatial spillovers modelled by continuous kernels. We use Fourier methods to identify necessary and sufficient conditions for the emergence of optimal agglomeration through an optimal spillover induced instability of a spatially homogeneous steady state. We apply our methods to study the emergence of optimal agglomeration for a rational expectations equilibrium and an optimal growth model. We believe that our analytical methods can be used to systematically study optimal agglomeration and clustering in dynamic economics.

KEY WORDS: Agglomeration, spatial spillovers, kernels, optimal control, spillover induced instability, optimal growth, rational expectations equilibrium.

October 16, 2009

1 Introduction

Modeling spatial interactions and studying spatial patterns which emerge endogenously as a result of interactions among agents has drawn considerable attention in scientific fields such as economics or biology. The main emphasis of this literature is on dynamical systems forces that cause agglomeration of economic activity (e.g. Krugman 1996) or general agglomeration phenomena in biology (e.g. Murray 2003).

A major approach to modelling spatial interactions is the use of an influence *kernel* which describes the effects of state variables located at different spatial sites on a state variable located at a given site.[1] Early writers such as Krugman (1996) and Fujita et al. (2001) used influence kernels to capture the tension between local centripetal forces and more distant centrifugal forces associated with the market potential of a location. Later writers such as Lucas (2001), Quah (2002), Lucas and Rossi-Hansberg (2002), Ioannides and Overman (2007) and Desmet and Rossi-Hansberg (2007) use kernels to incorporate spatial or geographical spillovers into economic models, to reflect, for example, the impact of employment at neighboring sites on productivity at a given site, or the impact of accumulated knowledge at neighboring sites on accumulated knowledge at a given site.

The purpose of this paper is to develop what we believe to be the first relatively general treatment of pattern formation and agglomeration in infinite horizon recursive forward-looking dynamical systems models, in which spatial effects are modelled by influence kernels. Our modelling includes kernel expressions in the law of motion and/or the payoff function, which makes it suitable for use in dynamic economics. The applicability of our method is demonstrated in section 4, through the study of the emergence of agglomeration in growth models with spatial effects.

Although earlier results in biology have used influence kernels to model

[1] The other major approach to modeling (mainly short-range) spatial interactions is the use of classical diffusion where a state variable moves from locations of high to low concentration. See for example Murray (2003, Vol. I, Ch. 11), or Brock and Xepapadeas (2008, 2009) for the analysis of infinite horizon forward-looking systems where the spatial interactions are of diffusion type.

long-range effects (e.g. Murray 2002), our contribution is that we explicitly solve for the optimization problem and derive conditions for the endogenous emergence of spatial patterns which result from the forward-looking optimization behavior of economic agents. The results obtained in biology do not incorporate optimizing behavior. Since in the majority of applications influence kernels are modelled by continuous functions and spatial spillovers by linear integral operators, the forward-looking optimization problems analyzed in this paper lead to infinite dimensional optimal control problems. We use Fourier type bases to decompose the infinite dimensional optimal control problems into a countable sequence of tractable finite ones. This approach allows us to fully characterize conditions for local stability/instability of a steady state to spatial spillovers and thus to study the emergence of economic agglomeration in terms of finite dimensional dynamical systems. We consider this to be an additional contribution as it provides tools for studying forward-looking dynamic optimization problems, which are at the core of dynamic economics, in infinite dimensional Hilbert space settings. We believe our methods could provide a useful basis for systematic analysis of agglomeration and clustering in dynamic economic models.

2 Spatial Spillover Dynamics and Optimization

This section presents results regarding necessary and sufficient conditions of optimal control under spatial spillovers modelled by a kernel. We consider a distributed control system where the state and the control are respectively represented by real functions $x(t,z)$ and $u(t,z)$ of time $t \in [0,\infty) := \mathcal{T}$, and the spatial variable $z \in [-\pi,\pi] := Z$.[2] Following Appel et al. (2000), the real function $x(t,z)$ of two variables (t,z) is identified with the abstract function $x = x(t)$ of one variable $t \in T$ which takes its values in a separable

[2]Circular spaces of the form $Z := [-|Z|, |Z|]$ can be handled by a change in units.

Hilbert space $\mathcal{X}$ of square integrable functions,[3] $\mathcal{X} = L^2(Z)$, defined as $x(t) = x(t)(z) = x(t, \cdot)$. Similarly the control $u(t, z)$ is identified with the abstract function $u = u(t) = u(t)(z) = u(t, \cdot)$, $u(t) \in \mathcal{U} = L^2(Z)$. We make the following assumptions:

A1: *For each* $[0, T] \in \mathcal{T}$, *the controls* $u(t, \cdot)$ *are measurable functions in* t *that lie in the subset* $B[0, T]$ *of* L^2 *such that*

$$\mathcal{B}[0, T] = \left\{ u(t, \cdot) \in L^2(Z) : \|u(t, \cdot)\|_{L^2(z)} \leq b(T) < \infty \right\} \tag{1}$$

where the bound $b(T)$ *is finite but may depend upon* T. *We call such controls* L^2- *bounded measurable controls.*[4]

A2: *The set of pairs* (x, u) *is admissible if for each* T, u is L^2- *bounded measurable control in* $\mathcal{B}[0, T]$ *and* x *is uniformly* L^2- *bounded on* $[0, T]$.

Long-range spatial effects describing the impact of the concentration of the state variable $x(t, z')$ in locations z' on $x(t, z)$ are modelled using the kernel formulation:

$$X(t, z) = \int_{z' \in Z} w(z - z')\, x(t, z')\, dz' := (\mathbf{K}x)(t, z). \tag{2}$$

A3: *The kernel function* $w(\cdot)$ *is continuous and symmetric.*

Assumption A3 implies that the operator $\mathbf{K}$ defined by $(\mathbf{K}x)(t, z)$ in (2) is a compact linear operator that maps the Hilbert space $L^2(Z; \mathbb{R})$ to itself. [5] The kernel function quantifies the impact of site z' on site z. Spatial impacts are assumed to be symmetric, or $w(z - z') = w(z' - z)$.

When spatial spillovers are combined with a temporal growth function

[3] A square integrable function $v(z)$ in the interval $a \leq z \leq b$ satisfies the condition $\int_a^b |v(z)|^2\, dz < \infty$.

[4] By Carleson's theorem each such $u(t, \cdot)$ has a Fourier series that converges pointwise for almost all z.

[5] A linear operator $\mathcal{A} : H \to \mathcal{H}$, where $H, \mathcal{H}$ are Hilbert spaces, is compact if the image of every bounded subset B of H under $\mathcal{A}$ is relatively compact in $\mathcal{H}$. For a continuous or square integrable kernel $w(\cdot)$, the operator $\mathcal{A}$ is compact. For a symmetric L^2-kernel the operator is self-adjoint. An operator $\mathcal{A}$ is linear if $\mathcal{A}(\beta_1 f_1 + \beta_2 f_2) = \beta_1 (\mathcal{A} f_1) + \beta_2 (\mathcal{A} f_2)$, for constants β_1, β_2 and square integrable functions f_1, f_2. Thus the operator we use is compact, linear and self-adjoint (for details, see e.g., Dieudonne, 1969, Chapter XI). To simplify notation, sometimes we write $\mathbf{K}v$ instead of $(\mathbf{K}v)(t, z)$ for some function $v(t, z)$.

$g(x(t,z), u(t,z), X(t,z))$, the rate of change of the state x at time t and location z depends also on the values of the state at locations $z' \in Z$ and can be written as:

$$\frac{\partial x(t,z)}{\partial t} = g(x(t,z), u(t,z), X(t,z)) \ , \ x(0,z) = x_0(z) \tag{3}$$

where $x_0(\cdot) \in L^2(Z)$. Equation (3) describes the effects of spatial spillovers on the evolution of the system's state (e.g. capital stock, knowledge, technology) both in the time and the space domain. Using the identification of $x(t,z), u(t,z)$ as $x = x(t)$, $u = u(t)$ which take values in separable Hilbert spaces, (3) can be written as the ordinary differential equation in $\mathcal{X} \times \mathcal{U}$:

$$\frac{dx}{dt} = g(x, u, X), \ x(0) = x_0. \tag{4}$$

A4: *The function* $g(x,u,X)$, $x = x(t), u = u(t)$, $X = X(t) = X(t,\cdot) = (\mathbf{K}x)(t,\cdot)$ *satisfies a Lipschitz condition; is* C^2; *and* $\sup \|\partial g/\partial v\|$, $v = x, u, X$ *is bounded. We also assume enough regularity in* (4) so that L^2-*bounded measurable control inputs yield* L^2- *bounded state outputs.*

From A4, (4) has an L^2- bounded solution $x(t)$ for a given u. The solution $x = x(t)$ defines a generalized solution $x(t,z) = x(t)(z)$ of the integro-differential equation (3).[6] Equation (3) can be used as a dynamic constraint in an optimal control problem where the objective is to choose admissible controls $u(t,z)$ which will maximize discounted benefits over the spatial domain Z associated with a payoff function. We associate with the control system (2)-(3) the payoff expression:

$$J(x_0(z), u(\cdot)) = \int_0^\infty e^{-\rho t} F(x(t), u(t), X(t))\, dt. \tag{5}$$

A5: $F(\cdot,\cdot,\cdot)$ *is a differentiably concave and upper-semicontinuous function on* $L^2 \times L^2 \times L^2$ *that satisfies the coercivity condition* $F(x(t), u(t), X(t)) \leq a - c\|x(t)\|_{L^2(Z)}$, $(a,c) > (0,0)$, (Leizarowitz 2008).

[6]A generalized solution $x = x(t,z)$ is measurable on $\mathcal{T} \times [-\pi, \pi]$, $x(\cdot, z)$ is absolutely continuous on $\mathcal{T}$ for each $z \in [-\pi, \pi]$ and satisfies (3) almost everywhere. For details see Appel et al. (2000, Chapter 1).

Payoff functional (5) corresponds to the optimization problem of a decentralized agent at site z who maximizes the discounted sum of net benefits at its own site taking the choices of agents at other sites as fixed and beyond its control. This implies that $X(t)$ is a fixed parameter. Of special interest is the payoff functional that corresponds to the maximization of discounted benefits over the entire spatial domain, when taking the choices of agents located in different sites explicitly into account, or

$$F(x(t), u(t), X(t)) = \int_{-\pi}^{\pi} f(x(t,z), u(t,z), X(t,z))\, dz. \tag{6}$$

A6: $f(\cdot,\cdot,\cdot)$ *is a differentiably concave and upper semi-continuous function* $f : \mathbb{R}^3 \to \mathbb{R}$.

Payoff functionals (5) and (6) can be associated with two important types of economic problems. When the payoff is maximized over the entire spatial domain, the problem can be stated as:

$$\max_{\{u(t,z)\}} J(x,u) = \int_{z \in Z} \int_0^{\infty} e^{-\rho t} f(x(t,z), u(t,z), X(t,z))\, dt dz \tag{7}$$

subject to (2) and (3).

Problem (7) provides a framework for the analysis of dynamic optimization problems in economics in which a social planner maximizes discounted benefits by internalizing spatial spillovers. For example, in the special case where f depends only on u, problem (7) can be regarded as describing the social planner's problem for the classic Ramsey growth model extended to include geographical spillovers. In this case f is a standard utility function, u is consumption and x is capital stock at time t and site z, and $X(t,z)$ reflects spatial spillovers on the production function g from capital located on sites around z. The introduction of x and X into the utility function could reflect stock effects and spatial effects on utility. Maximization of (5) subject to (2) and (3) with $X(t,z) = X^e$ can be regarded as the problem of a private agent located at z who does not internalize spatial spillovers. This problem can be associated with a rational expectations equilibrium. Both problems are analyzed below.

Problem (7) is an optimal control problem of an infinite dimensional system. Our assumptions in the context of the results developed by Papageorgiou (1990) suggest that (7) admits a solution.[7] Optimality conditions for (7) can be obtained using Pontryagin's maximum principle. Dropping (t, z) to ease notation, we introduce the current value Hamiltonian function:

$$H(x, u, X, \lambda) = f(x, u, X) + \lambda g(x, u, X) \tag{8}$$

and associate the costate variable $\lambda(t, z)$ with transition equation (3). Let $u^*(t, z)$ be a control function satisfying A1 and A2 and let $x^*(t, z)$ be a generalized solution of (3) corresponding to $u^*(t, z)$ and originating at $x_0(z)$. First-order necessary conditions state that, in order for $u^*(t, z)$ to be optimal for problem (7), a function $\lambda(t, z)$ must exist such that:[8],[9]

(i) *$x^*(t, z)$ and $\lambda(t, z)$ are a solution of the system*

$$\frac{\partial x}{\partial t} = g(x^*, u^*, X^*) = H_\lambda(x^*, u^*, X^*) \ , x(0) = x_0(z) \tag{10}$$

$$\frac{\partial \lambda}{\partial t} = \rho\lambda - (f_x^* + \lambda g_x^*) - (\mathbf{K} f_X^* + \mathbf{K}\lambda g_X^*) = \tag{11}$$

$$\rho\lambda - H_x(x^*, u^*, X^*) - \mathbf{K} H_X(x^*, u^*, X^*) \tag{12}$$

satisfying a temporal limiting transversality condition, and spatial transver-

[7] Actually Papageorgiou provides existence results for a more general problem then ours. Similar existence results can also be found in, e.g., da Prato and Ichikawa (1993).

[8] The conditions can be obtained using a variational argument and the linearity of the integral operator. For a heuristic discussion of the derivation, see Appendix 1.

[9] To ease notation we denote partial derivatives with subscripts. We write $H_\lambda(x, \lambda, X)$ in (11) to emphasize that H_λ is a function of three arguments (x, λ, X). We use the shorthand notation introduced in (4) to write (10) in a more compact way. In order to ease notational clutter we write the mappings

$$\mathbf{K} v(t, z) = \int_{z \in Z} w(z - z') \, v(x(t, z'), u(t, z'), X(t, z')) \, dz' dz' \tag{9}$$

where $v(x, u, X)$ stands for f_X, or λg_x, or $H_X(x, u, X)$, as $\mathbf{K} f_X^*, \mathbf{K}\lambda g_x^*, \mathbf{K} H_X(x^*, u^*, X^*)$. Superscript $(^*)$ indicates evaluation along the trajectory (x^*, u^*, X^*).

sality conditions for a finite spatial domain with circle boundary conditions[10]

$$\lim_{T\to\infty} e^{-\rho T}\int_{z\in Z}\lambda\left(T,z\right)x\left(T,z\right)dz=0 \tag{13}$$

$$\lambda\left(t,-\pi\right) = \lambda\left(t,\pi\right),\ x\left(t,-\pi\right)=x\left(t,\pi\right),\ \textit{for all } t. \tag{14}$$

(ii) *The Hamiltonian function* $H\left(x^{*},u,X^{*},\lambda\right)$ *has a (possibly local) maximum as a function of* u *at* $u^{*}\left(t,z\right)$ *for all* $t\geq 0$. *For an interior maximum,*

$$\frac{\partial H}{\partial u}=0,\ \textit{or}\ f_u+\lambda g_u=0\Rightarrow u^{*}=u^{*}\left(x,\lambda,X\right). \tag{15}$$

As shown in Appendix 2, under standard differentiable concavity assumptions the necessary conditions stated above are also sufficient.

If we interpret - as is common in such problems - the costate variable $\lambda\left(t,z\right)$ as the shadow price of the state variable at time t and location z, condition (11) suggests that geographical spillovers influence the evolution of shadow prices in both time and space. Assume for example that f does not depend on X and that $\partial x/\partial t=g\left(x,u\right)+\mathbf{K}x$. Then (11) becomes $\partial\lambda/\partial t=\rho\lambda-H_x-\left(\mathbf{K}\lambda\right)\left(t,z\right),\left(\mathbf{K}\lambda\right)\left(t,z\right)=\int_{z'\in Z}w\left(z-z'\right)\lambda\left(z',t\right)dz'$. That is, at the optimal solution, prices are affected in a similar way as the stocks (quantities) by geographical spillovers, but the effect on prices is in the opposite direction from the effect on quantities. Thus if the stock quantity in location z' has a positive effect on the stock accumulation in location z, the corresponding price in z' will have a negative effect on price changes in location z.

If we use (5) as the objective functional, the optimization can be interpreted as having a planner at each site z that maximizes discounted benefits on the site and considers the spatial spillover $X\left(t,z\right)$ affecting her/his site as an exogenous parameter X^{e}. This problem can be written as:

$$\max_{\{u(t)\}}\int_0^\infty e^{-\rho t}F\left(x\left(t\right),u\left(t\right),X^{e}\right)dt,\quad\text{subject to} \tag{16}$$

$$\frac{dx}{dt}=g\left(x\left(t\right),u\left(t\right),X^{e}\right),\ x\left(0,z\right)=x_0\left(z\right),\forall\ z\in Z. \tag{17}$$

[10]For a concave problem (7) the temporal limiting transversality condition is a necessary condition (e.g., Benveniste and Scheinkman 1982).

This is a standard optimal control problem with current value Hamiltonian function $h = F(x, u, X^e) + \lambda g(x, u, X^e)$. Setting $X(t, z) = X^e$ in the optimality conditions of problem (16)-(17), a rational expectations equilibrium is characterized by the Hamiltonian system:

$$\frac{\partial x(z,t)}{\partial t} = g(x, \hat{u}, \mathbf{K}x) = h_\lambda(x, \lambda, X) \tag{18}$$

$$\frac{\partial \lambda(z,t)}{\partial t} = \rho\lambda - (f_x + \lambda g_x) = \rho\lambda - h_x(x, \lambda, X) \tag{19}$$

where $\hat{u}$ maximizes the current value Hamiltonian h.

3 Optimal Spillover Induced Instability and Agglomerations

A question which arises in the study of problems described by (7) is whether solutions exhibit spatial homogeneity or spatial heterogeneity. Spatial homogeneity means that the state, costate and control variables which are solutions of (7) have a spatially uniform distribution along the optimal spatiotemporal path. Heterogeneity means that spatial distributions are not uniform and thus spatial patterns are formed. This implies that clusters or economic agglomeration emerge as a result of optimizing behavior and may become persistent at a spatially heterogeneous steady state.

To study the emergence of economic agglomeration we follow the general approach introduced by Turing (1952) which examines the stability of a stable spatially homogeneous, or *flat*, steady state of reaction-diffusion systems to spatially heterogenous perturbations.[11] We extend this approach to deal with the system of equations such as (10)-(11) which constitute the Hamiltonian system for problem (7). Define a flat optimal steady state (FOSS) (x^*, λ^*) as the steady state of system (10)-(11), which is obtained when $(\partial x/\partial t) = (\partial \lambda/\partial t) = 0$ where state, costate and control have the same value at all

[11] Turing's approach has been used in new economic geography (e.g. Krugman 1996, Fujita et al. 1999, Chincarini and Asherie 2008), in biology (e.g. Murray 2003) and in ecosystem management (Brock and Xepapadeas 2008, 2009).

spatial sites but are optimal given the same initial conditions as in the FOSS. To examine the stability of this FOSS we consider small perturbations off the FOSS, $(x(t,z)-x^*, \lambda(t,z)-\lambda^*)$. For sufficiently small perturbations the stability analysis can be obtained in terms of linearization of (10)-(11). This linearization is the Fréchet derivative of (10)-(11) evaluated at (x^*, λ^*). Since we integrate over finite limits when defining $(\mathbf{K}x)(t,z)$ and the kernel $w(\cdot)$ is continuous, the Fréchet derivative is a compact linear operator, which has an integral representation and countable numbers of eigenvalues besides the zero eigenvalue.[12] Using, by a slight abuse of notation, (x, λ) to denote deviations from (x^*, λ^*), and setting $X = \mathbf{K}x$ to simplify notation, the linearization is:

$$\frac{\partial x}{\partial t} = H^*_{\lambda x} x + H^*_{\lambda X} \mathbf{K} x + H^*_{\lambda\lambda} \lambda \tag{20}$$

$$\frac{\partial \lambda}{\partial t} = -H^*_{xx} x - 2H^*_{Xx} \mathbf{K} x - H^*_{XX} \mathbf{K}(\mathbf{K}x) + (\rho - H^*_{x\lambda})\lambda - H^*_{X\lambda} \mathbf{K}\lambda \tag{21}$$

where the superscript $(^*)$ indicates that the Fréchet derivatives are evaluated at (x^*, λ^*). Our approach in studying the stability of the FOSS to spatially heterogeneous perturbations off the FOSS, is to transform the infinite dimensional system (20)-(21) into a countable sequence of linear systems of ordinary differential equations so that we can use linear stability analysis. To do this we consider pairs of square integrable solutions $(x(t)(z), \lambda(t)(z)) = (x(t,z), \lambda(t,z))$ and we construct trial solutions using an orthogonal basis of $L^2(Z)$ created in terms of functions $\cos(kz)$, $\sin(kz)$, $z \in [-\pi, \pi]$, for mode $k = 0, 1, 2, ...$ which form a complete orthogonal basis over $[-\pi, \pi]$. Our assumptions about functions f and g suggest that the solution $(x(t,z), \lambda(t,z))$ of the optimal control problem will be smooth enough

[12]For a statement of these results in an applied context see Kot and Schaffer (1986).

to be expressed in terms of Fourier basis, or:[13]

$$x(t,z) = \sum_{k=0}^{\infty} \langle a_k(t), B_k(z) \rangle, \ x(0,z) = \sum_{k=0}^{\infty} (a_k(0) \cdot B_k(z)) \quad (22)$$

$$\lambda(t,z) = \sum_{k=0}^{\infty} \langle A_k(t), B_k(z) \rangle \quad (23)$$

where $B_k(z) = (\cos(kz), \sin(kz))$ is the sine/cosine basis, $a_k = (a_{1k}, a_{2k})$, $A_k = (A_{1k}, A_{2k})$. and $\langle \cdot, \cdot \rangle$ denotes inner product.

Proposition 1: *Assume that the state and costate variables are expressed by the Fourier basis (22)-(23), then the linearized infinite dimensional system (20)-(21) can be transformed to the following countable sequence of linear systems of ordinary differential equations for* $k = 0, 1, 2, \ldots$ *:*

$$\frac{dx_k}{dt} = [H^*_{\lambda x} + H^*_{\lambda X} W(k)] x_k + H^*_{\lambda\lambda} \lambda_k \ , \quad (24)$$

$$W(k) = \int_{\zeta} w(\zeta) \cos(k\zeta) \, d\zeta, \zeta = z - z' \quad (25)$$

$$\frac{d\lambda_k}{dt} = \left[-H^*_{xx} - 2H^*_{Xx} W(k) - H^*_{XX} W^2(k)\right] x_k + \quad (26)$$
$$[\rho - H^*_{x\lambda} - H^*_{X\lambda} W(k)] \lambda_k.$$

For Proof see Appendix 3.

This transformation suggests that the standard spatially homogeneous optimal control problem can be regarded as a special case of our more general spatially dependent problem (7) for $k = 0$. To analyze the stability of the FOSS for (7), it is sufficient to analyze the stability of the FOSS for each of the mode-k systems (24)-(26). If the FOSS becomes unstable at some mode k, this implies that spatially heterogenous perturbations destabilize the FOSS and economic agglomeration begins emerging. The stability of the FOSS for system (24)-(26) at mode k depends on the eigenvalues of a countable sequence of Jacobian matrices indexed by $k = 0, 1, 2, \ldots$. Let

[13] As shown in Priestley (1981, Section 4.2) any function in the class $L^2(-\pi, \pi)$ has a Fourier expansion which converges to the function in the mean square sense. Furthermore the Fourier coefficients $\alpha_k(t)$, $A_k(t)$ exist for each t and the Fourier series converges for each t. For more details see Appendix 3.

$J_k(x^*, \lambda^*) := J_k^*$ denote this sequence of Jacobian matrices with eigenvalues $(\sigma_{1k}, \sigma_{2k})$. The trace of J_k^* is $\rho > 0$ which implies at least one positive eigenvalue for all modes $k \geq 0$. Thus for $k = 0$, the spatially homogeneous case, the FOSS is either saddle point stable or completely unstable. As shown by Scheinkman (1976), the local solution manifold structure in this case is obtained by choosing, for a given initial state value $a_0(0)$ sufficiently close to the FOSS x^*, the initial costate value $A_0(0)$ such that $(a_0(t), A_0(t))$ lies on the one-dimensional manifold corresponding to the smallest eigenvalue of the "Poincare pair" $(\rho - \sigma_0, \sigma_0)$. This local manifold is tangent to the true nonlinear manifold at the FOSS. For a two-dimensional Hamiltonian system the tangent manifold is a line that passes through the FOSS and its slope is equal to that of the eigenvector that corresponds to the smallest eigenvalue.[14] This argument, extended to mode $k > 0$, suggests that the local solution manifold structure for mode k can be constructed by choosing, for each initial condition $a_k(0)$ which is sufficiently close in L^2 norm to x^*, the initial mode-k costate $A_k(0)$ such that $(a_k(t), A_k(t))$ lies on the one-dimensional manifold corresponding to the smallest eigenvalue of the "Poincare pair" $(\rho - \sigma_k, \sigma_k)$ with the initial value $A_k(0)$ determined by the eigenvector corresponding to the smallest eigenvalue and initial conditions $a_k(0)$. If the smallest eigenvalue is negative the local mode-k manifold is asymptotically stable and tangent to the true nonlinear mode-k manifold. In this case the FOSS retains the local saddle point property at mode k, and J_k^* has two real eigenvalues at this mode, one positive and one negative. If the smallest eigenvalue is positive the local mode-k manifold is unstable and optimal economic agglomeration due to spatial spillovers emerges around the FOSS at mode k. In this case, the FOSS is destabilized by spatial perturbations at mode k and J_k^* has two positive eigenvalues or two complex eigenvalues with positive real parts at this mode. Thus we derive a solution that satisfies the first-order necessary conditions for problem (7). Furthermore our sufficiency conditions under strict concavity imply that this solution is a solution to (7) and is thus

[14] If the smallest eigenvalue is negative then the local solution manifold converges to the FOSS. If the smallest eigenvalue is positive the solution manifold does not converge but it is optimal, provided it satisfies the optimality conditions and the temporal transversality condition at infinity.

unique.

Our results extend Turing's method to forward-looking dynamic optimization problems with spatial spillovers modelled by continuous symmetric kernels. Since our spatial instability is the result of optimizing behavior and spatial spillovers, we call it - by analogy to Turing's diffusion induced instability - *optimal spillover induced instability.* Thus from a saddle point FOSS, an optimal spillover induced instability emerges if the determinant of the Jacobian matrix J_k^* becomes positive for some mode $k > 0$. This determinant is called the *dispersion relationship* and is presented analytically in Appendix 3. Our stability/instability conditions for all modes $k = 0, 1, 2, ...$ are independent of the choice of basis for $L^2(Z)$ as shown below.

Proposition 2: *Let (x^*, λ^*) be a FOSS for the linearized system (20)-(21). The FOSS will be a local saddle point iff it is a local saddle point for all modes k. The FOSS will be locally unstable iff a mode k exists such that the FOSS is unstable for this mode. The local stability/instability result is independent of the basis chosen in $L^2(Z)$.*

For Proof see Appendix 4.

Proposition 2 means that if the FOSS is unstable to spatially heterogeneous perturbations for the sine/cosine basis at some mode k, it will be unstable at the same mode for any other complete basis in $L^2(Z)$. Conversely if the FOSS is stable for all modes $k \geq 0$ for the sine/cosine basis, it will be stable for any other complete basis in $L^2(Z)$. Thus the cyclic class of perturbations is sufficient to check for the emergence or not of spatial instabilities and economic agglomeration. If the dispersion relationship becomes positive for some mode k, then optimal spillover induced instability emerges. On the other hand, for optimal spillover induced instability *not to emerge*, the dispersion relationship must remain negative for all $k \geq 0$.

Our approach, like Turing's, focuses on the initial stages of a process where spatial spillovers cause deviations from the FOSS which do not die away, but grow over time to create agglomeration. Growth requires two positive eigenvalues or two complex eigenvalues with positive real parts for J_k^* at mode k, i.e. an unstable mode-k tangent manifold. The eigenvalues of J_k^* (see Appendix 3 for exact definition) depend on the fundamental parameters of our system, which include the discount rate; the second derivatives of the Hamiltonian function at the FOSS, which can be associated with benefits and costs of controlling the system to the FOSS; the spatial spillovers; and the mode itself through $W(k)$. Since we consider optimized systems, emergence of agglomeration at a specific mode may be interpreted as suggesting that at this mode the system can attain a higher value when not controlled towards the FOSS but letting an "optimal agglomeration" develop at mode k.[15] The optimal agglomeration will be realized as an emerging wave-like spatial pattern which grows over time on the unstable tangent manifold in the neighborhood of the FOSS. This agglomeration will persist if it is realized as a *spatially heterogeneous steady state*. This steady state, if it exists, will correspond to a time stationary solution $(x^*(z), \lambda^*(z))$ of the system of integral equations resulting from (10)-(11) for $(\partial x/\partial t) = (\partial \lambda/\partial \lambda t) = 0$. An example is provided in section 4.2.

We can also study spillover induced instability of the rational expectations equilibrium. Following the theory developed above, Fourier expansions imply the following sequence of linear systems of ordinary differential equations indexed by k:

$$\frac{dx_k}{dt} = \left(\bar{h}_{\lambda x} + \bar{h}_{\lambda X} W(k)\right) x_k + \bar{h}_{\lambda\lambda}\lambda_k \tag{27}$$

$$\frac{d\lambda_k}{dt} = \left(-\bar{h}_{xx} - \bar{h}_{xX} W(k)\right) x_k + \left(\rho - \bar{h}_{x\lambda}\right)\lambda_k \tag{28}$$

where all derivatives are evaluated at the flat steady state (FSS), $\left(\bar{x}, \bar{\lambda}\right)$.[16] Let $J_k\left(\bar{x}, \bar{\lambda}\right) := \bar{J}_k$ be the Jacobian matrix associated with (27)-(28). Destabilization of the FSS requires that $\text{trace}\bar{J}_k = \rho + \bar{h}_{\lambda X} W(k) > 0$ and $\det \bar{J}_k > 0$. By comparing (24)-(26) to (27)-(28) we see that the conditions for the desta-

[15]The relation between the system's fundamental parameters and the unstable mode is discussed in Krugman (1996) or Fujita et al. (2001, Chapter 6). We study this issue in a dynamic optimization framework with spatial spillovers, which as far as we know, has not been done before. Brock and Xepapadeas (2008, 2009) analyze a similar issue for short-range spatial interactions modelled by classical diffusion.

[16]FSS is the spatially homogeneous steady state defined by setting $(\partial x/\partial t) = (\partial \lambda/\partial t) = 0$ in (18)-(19). Thus we distinguish between the spatially homogeneous steady states corresponding to the social planner's problem (the FOSS) and the rational expectations equilibrium (the FSS). We denote FOSS with $(^*)$ and FSS with $(^-)$.

bilization of the FOSS and the FSS due to spatial spillovers are not the same. First note from (24), (26) that at each mode k the J_k^* of the linearization of the social planner's problem, i.e. , satisfies the property that trace$J_k^* = \rho$. Second, note that if σ is an eigenvalue of J_k^*, $\rho - \sigma$ is also. Hence as ρ approaches zero, the eigenvalues appear in opposite pairs, i.e. we have a saddle point. Since we have a saddle point for each mode k, we expect no pattern generation for small ρ for the social planner's problem under the usual differentiable concavity assumptions. This result is expected intuitively from the turnpike literature in infinite dimensional problems under the usual concavity assumptions. Turn now to a comparison with the linearization for the rational expectations system, i.e. equations (27)-(28). First we notice that trace$\bar{J}_k$ is not equal to ρ unless $\bar{h}_{xX} W(k) = 0$. But in this latter case there is no spatial externality at mode k. Thus it is intuitively clear that the saddle point property would be recovered without the spatial externality. Second, the "extra force" of concavity, i.e. the term, $\bar{h}_{XX} W^2(k)$, which is negative, is missing from $\bar{J}_k$ in contrast to J_k^*. Thus we would intuitively expect fewer patterns to be present under the usual concavity assumptions of economics, all other things equal, for the social optimization problem in contrast to the rational expectations equilibrium problem. We now use the above theoretical framework to study a classical problem of growth theory.

4 Geographical Spillovers, Pattern Formation and Optimal Growth

The classic Ramsey growth model, extended to include spatial spillover externalities in the production function, is a special case of problem (7). In this case, the problem of the social planner can be written as:

$$\max_{\{c(t,z)\}} \int_{-\pi}^{\pi} \int_{0}^{\infty} e^{-\rho t} U(c(t,z))\, dt dz \quad \text{subject to} \tag{29}$$

$$c(t,z) + \frac{\partial x(t,z)}{\partial t} = f(x(t,z), X(t,z), l) - \eta x(t,z), x(0,z) = x_0(z) \tag{30}$$

where $c(t,z)$ denotes consumption at site z at date t; $U(c(t,z))$ is a standard utility function which is C^2, strictly differentiably concave, strictly increasing, with $U'(0)=\infty$, $U'(\infty)=0$; $x(t,z)$ denotes capital stock at site z at date t which depreciates at the rate $\eta(t,z)$; l denotes labor; and $X(t,z)$ defined by (2) denotes an external effect on production at site z at date t. For the production function we assume $(f_x, f_X) > (0,0)$, $(f_{xx}, f_{XX}) < (0,0)$, $f_{xX} > 0$. As both $x(t,z), X(t,z)$ are treated as inputs, the quantity $X(t,z)$ will have different interpretations in different contexts. If $X(t,z)$ represents a type of knowledge which is produced proportionately to capital usage, it is natural to assume that the kernel $w(\zeta), \zeta = z - z'$ is single peaked with a maximum at $\zeta = 0$, like kernel $w_1(\zeta)$ in Appendix 5 (Figure 2). If $X(t,z)$ reflects aggregate benefits of knowledge produced at (t,z') for producers at (t,z) and damages to production at (t,z) from usage of capital at (t,z'), then non-monotonic shapes of $w(\zeta)$ in ζ, like kernel $w_2(\zeta)$ in Appendix 5 (Figure 4), are plausible. This production function could be considered a spatial version of a neoclassical production function with Romer (1986) and Lucas (1988) externalities modelled by geographical spillovers given by a Krugman (1996), Chincarini and Asherie (2008) specification. To concentrate on the impact of geographical spillovers, we assume zero exogenous technical change and fixed labor input in each site l normalized to unity, i.e. $l = 1$, for all sites z so $x(t,z)$ denotes total and per capita capital, and we write $f(x, X, l) := f(x, X)$. We assume that $f(x, X, l)$ exhibits constant returns to scale in (x, l) for each X. Notice that labor and capital cannot be moved across sites. Thus the social planner's problem (29)-(30) has an extreme assumption that capital and labour are completely immobile across locations. If capital, labor, and consumption goods are completely mobile, then it can be shown that it is easy to reduce problem (29)-(30) to one that is equivalent to a one-dimensional Ramsey type problem. Of course the cases of complete immobility of capital and labor and complete mobility of capital and labor are polar cases, but they can be used to provide insight into the more realistic case where there are frictional costs to the movement of capital and labor. A special case of the planner's problem (29)-(30) is the one where each economic agent considers the spatial externality $X(t,z)$ as given and maximizes discounted utility at each site z, subject to (30). This optimization problem can be associated with the concept of rational expectations equilibrium. We first analyze this problem and then move to the more general problem of the social planner.

4.1 Rational Expectations Equilibrium

Assume that within each site representative consumers maximize the discounted sum of utilities subject to the intertemporal budget constraint $\dot{\alpha}(t,z) = r(t,z)\,\alpha(t,z) + w(t,z) - c(t,z)$, where $\alpha(t,z)$ denotes net assets per capita at site z. Representative consumers at z rent out their capital at rate $r(t,z)$, receive wages $w(t,z)$, and take $r(t,z)$ and $w(t,z)$ as parametric. Representative firms take spatial spillovers $X^e(z,t)$ as parametric and, dropping (t,z) to simplify, hire capital and labor to maximize profits $\pi = f(x, X^e) - (r+\eta)\,x - w$ by facing rental rates on capital and wages parametrically. Constant returns to scale imply that after capital, labor are paid competitive rents, wages, the remaining net income for the firm is zero. A competitive equilibrium is produced in each site conditional on the commonly shared point expectations on $X^e(t,z)$. In equilibrium $\alpha(t,z) = x(t,z)$. Substituting profit maximization and the zero profit conditions, $r + \eta = f_x(x, X^e)$, $w = f(x, X^e) - f_x x$ respectively, into the consumer's budget constraint, the following representative consumer's problem generates a competitive equilibrium at each site z:

$$\max_{\{c(t,z)\}} \int_0^\infty e^{-\rho t} U(c(z,t))\,dt \quad \text{subject to} \tag{31}$$
$$c(t,z) + \frac{\partial x(t,z)}{\partial t} = f(x(t,z), X^e(t,z)) - \eta x(t,z),\ x(0,z) = x_0(z).$$

A rational expectations equilibrium is where $X^e(t,z)$ is actual $X(t,z)$ for each (t,z). Since we assume that capital is completely immobile, we interpret "capital" as a type of capital embodied in humans, knowledge or technology which does not move across "sites" z.[17] Using conditions (18)-(19) and the results of section 3, the Hamiltonian system for problem (31) can be trans-

[17] A richer model would allow mobility of capital by imposing some type of "haste makes waste" adjustment costs. This however is an area for future research.

formed, using the Fourier basis approach, into the following countable number of finite dimensional linear equilibrium problems, one for each mode, with derivatives evaluated at the FSS rational expectations equilibrium $(\bar{x}, \bar{p})$.

$$\frac{dx_k}{dt} = \left[\rho + \bar{f}_X W(k)\right] x_k - c'(\bar{p})\, p_k \ , c'(p) < 0. \tag{32}$$

$$\frac{dp_k}{dt} = -\left[\bar{f}_{xx} + \bar{f}_{xX} W(k)\right] \bar{p} x_k. \tag{33}$$

Mode $k = 0$ corresponds to a spatially homogeneous rational expectations equilibrium. For the spatial externality to generate economic agglomeration, the FSS should become unstable to spatially heterogeneous perturbations induced by the spatial spillovers at some mode k. Therefore, the stability analysis of section 3 suggests that clustering will emerge at the rational expectation equilibrium at mode k, if the mode-k Jacobian matrix of (32)-(33) has positive trace and determinant or:

$$\text{trace}\bar{J}(k) = \rho + \bar{f}_X W(k) > 0 \tag{34}$$

$$\det \bar{J}(k) = \varphi(k) = -c'(\bar{p})\, \bar{p} \left[\bar{f}_{xx} + \bar{f}_{xX} W(k)\right] > 0. \tag{35}$$

Thus $\varphi(k)$ is the dispersion relationship. With a Cobb-Douglas production function, $x^{\alpha_1} X^{\alpha_2}$, a FSS will solve the equation $\rho + \eta = \alpha_1 x^{\alpha_1 - 1} (W(0)\, x)^{\alpha_2}$, $W(0) = \int_{-\pi}^{\pi} w(\zeta)\, d\zeta = \bar{w}$. That is, only mode zero matters. Relationship (35) implies that this FSS becomes completely unstable and agglomeration emerges at mode k if $\alpha_1 + \alpha_2 (W(k)/W(0)) > 1$.[18] Thus the emergence of agglomeration requires increasing social returns to capital at some mode k. On the other hand, decreasing social returns to capital $(\alpha_1 + \alpha_2 < 1)$ at mode $k = 0$ imply that $\varphi(0) < 0$ and the rational expectations equilibrium FSS has the saddle point property, which is the expected result.

[18] The shape of the $W(k)$ for two different types of kernels is presented in Appendix 5 (figures 3 and 5).

4.2 The Social Planner's Optimum

The social planner, assuming that capital $x(t,z)$ is immobile in the sense described above and that consumption goods $c(t,z)$ are produced on site, solves problem (29)-(30). The Ramsey type optimality conditions for the social optimum can be derived by direct application of (10)-(15). Assume that a FOSS (x^*, p^*), as defined in section 3, exists and has the saddle point property. The Fourier basis approach transforms, as seen earlier, the planner's infinite dimensional linearized Hamiltonian system into a countable sequence of systems of linear ordinary differential equations indexed by mode k. Following section 3, the saddle point stable FOSS becomes completely unstable at a mode k, and agglomeration emerges if the determinant of the mode-k Jacobian matrix of the linear system of ordinary differential equations, i.e. the dispersion relationship, is positive at this mode k, or

$$\begin{aligned} \psi(k) &= (\rho + \eta - f_x^* - f_X^* W(k))(f_x^* + f_X^* W(k) - \eta) - \\ &\quad \bar{p}c'(\bar{p})\left[f_{xx}^* + 2f_{xX}^* W(k) + f_{XX}^* W^2(k)\right] > 0. \end{aligned} \tag{36}$$

It might be interesting to compare the rational expectations equilibrium steady state FSS and the social planner's FOSS, with respect to their size and likelihood of becoming unstable due to spatial spillovers. Let $(\bar{x}, \bar{p})$, (x^*, p^*) denote the FSS and the FOSS respectively, and assume that the production function is Cobb-Douglas with decreasing social returns at the spatially homogeneous case. Then it can be easily shown that, as expected, $\bar{x} < x^*$. To compare agglomeration forces we compare the dispersion relationships (35) and (36). Write (36) as $\psi(k) = T_1(k) - T_2(k)$ and assume that the function $g(x,k) := f(x, xW(k))$ is concave in x for each k. Define $T_3(k) = f_{xx}^* + 2f_{xX}^* W(k) + f_{XX}^* W^2(k) < 0$, then $T_2(k) = \bar{p}c'(\bar{p})\,T_3(k) > 0$. At a flat steady state $\rho + \eta = f_x^* - f_X^* \bar{W}(0)$, thus $T_1(k) = f_X^*(W(0) - W(k))\,(\,f_x^* + f_X^* W(k) - \eta)$. The emergence of clusters at the FOSS requires that $T_1(k) > 0$ and $T_1(k) > |T_2(k)|$. On the other hand, the emergence of clusters at the FSS requires that $\varphi(k) > 0$ or $\bar{f}_{xx} + \bar{f}_{xX} W(k) > 0$, since $-c'(\bar{p})\,\bar{p} > 0$. Numerical simulations, presented in Appendix 6, suggest that the FSS is more likely than the FOSS to become unstable under spatial spillovers, confirming the intuition discussed in section 3. Instability of the FSS means that an equilibrium steady state agglomeration may be realized in the long run. Such an agglomeration, obtained numerically, is presented in figure 1 (see Appendix 6 for details). The flat line corresponds to the FSS which is destabilized by spatial spillovers.

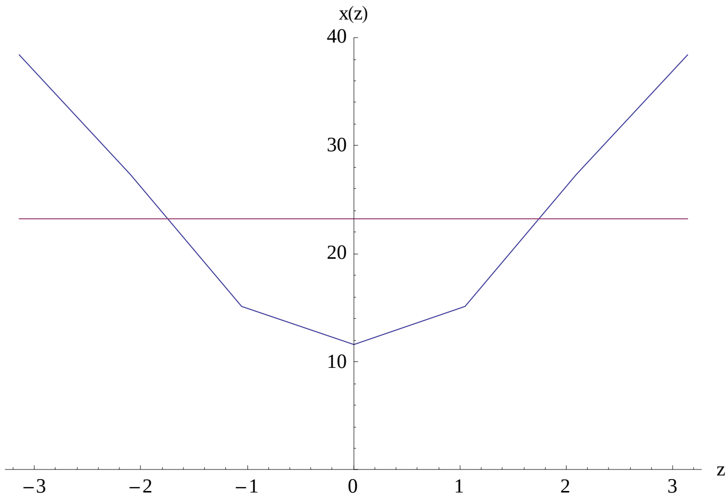

Figure 1: Equilibrium steady state agglomeration and FSS

5 Conclusions

This paper develops a fairly general approach to the study of infinite dimensional, infinite horizon, intertemporal recursive dynamic optimization models in continuous spatial settings, as well as analytical techniques for local stability analysis of spatially flat optimal steady states to spatial spillovers. Our work is related to the stability analysis of infinite dimensional, infinite horizon optimal control problems in Hilbert space settings (e.g. Carlson et al. 1991, Chapter 9; Leizarowitz 2008), but we formulate and analyze models with spillovers represented by kernels as in the new economic geography literature, technology spillover models, and elsewhere. We exploit Fourier basis techniques to organize the local stability analysis around an analytically tractable dispersion relation. Using the dispersion relation, which is a function of modes, we locate necessary and sufficient conditions for the local stability/instability) of a FOSS and FSS. Our stability analysis, which as we show is independent of the basis choice, allows us to study the emergence of optimal economic agglomeration in fairly general dynamic settings.

We apply our methods to the classic Ramsey model of growth theory extended to include spatial spillovers in the production function, and to a rational expectations competitive equilibrium under similar spatial spillovers. Our results suggest that there is a range of parameter values where the FSS associated with the rational expectations equilibrium is locally unstable to spatial spillovers, while the FOSS associated with the planner's problem is locally stable. This illustrates the economic point that in a world of low enough discounting, the social optimum would be stable due to the usual logic behind turnpike theorems (e.g. Scheinkman 1976), but the rational expectations competitive equilibrium can easily be unstable. In other words, it is socially optimal not to have agglomeration form, yet the competitive equilibrium produces agglomeration.

What about future research? We think the top priority is to extend the general forward-looking infinite dimensional, infinite horizon optimization approach developed here to new economic geography models, to structural change models, and to the general study of symmetry breaking in economics. We need to enrich the models studied here to include endogenous product variety at each site, increasing returns to production of each variety at each site, imperfect competition among varieties, backward/forward linkages, costly movement of resources, and other ingredients that expose the role of increasing returns, elasticity of substitution among varieties, costliness of moving resources, and so on. We view our paper as a contribution to the set of analytical techniques useful for analyzing models in this area.

References

APPEL, J., A. KALITVIN, AND P. ZABREJKO (2000): *Partial Integral Operators and Integro-differential Equations.* New York: Marcel Dekker.

BENVENISTE, L., AND J. SCHEINKMAN (1982): "Duality Theory for Dynamic Optimization Models of Economics: The Continuous Time Case," *Journal of Economic Theory*, 27, 1-19.

BROCK, W., AND A. XEPAPADEAS (2008): "Diffusion-Induced Instability and Pattern Formation in Infinite Horizon Recursive Optimal Control," *Journal of Economic Dynamics and Control*, 32, 2745–2787.

BROCK, W., AND A. XEPAPADEAS (2009): "Pattern Formation, Spatial Externalities and Regulation in Coupled Economic-Ecological Systems," Working paper, Available at SSRN: http://ssrn.com/abstract=1144071, forthcoming *Journal of Environmental Economics and Management*.

CARLSON, D., A. HAURIE, AND A. LEIZAROWITZ (1991): *Infinite Horizon Optimal Control*, Second Edition. Berlin: Springer-Verlag.

CHINCARINI, L., AND N. ASHERIE (2008): "An Analytical Model for the Formation of Economic Clusters," *Regional Science and Urban Economics*, 38, 252–270.

DA PRATO, G., AND A. ICHIKAWA (1993): "Optimal Control of Intergodifferential Equations of Parabolic Type," *SIAM Journal of Control and Optimization*, 31, 1167-1182.

DESMET, K., AND E. ROSSI-HANSBERG (2007): "Spatial Growth and Industry Age," CEPR Discussion Paper No. DP6421.

DIEUDONNE, J. (1969): *Foundations of Modern Analysis.* New York: Academic Press.

FUJITA, M., P. KRUGMAN, AND A. VENABLES (2001): *The Spatial Economy.* Cambridge, MA: The MIT Press.

IOANNIDES, Y. M., AND H. G. OVERMAN (2007): "Spatial Evolution and the U.S. Urban System," *Journal of Economic Geography*, 4, 131–156.

KOT, M., AND W. SCHAFFER (1986): "Discrete-Time Growth-Dispersal Models," *Mathematical Biosciences*, 80, 109-136.

KRUGMAN, P. (1996): *The Self-Organizing Economy.* Cambridge, MA: Blackwell Publishers.

LEIZAROWITZ, A. (2008): "Turnpike Properties of a Class of Aquifer Control Problems," *Automatica*, 44, 1460–1470.

LUCAS, R. E. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22 (1), 3–42.

LUCAS, R. E. (2001): "Externalities and Cities," *Review of Economic*

Dynamics, 4, 245–274.

LUCAS, R. E., AND E. ROSSI-HANSBERG (2002): "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.

MURRAY, J. (2003): *Mathematical Biology,* Volumes I and II, Third Edition. Berlin: Springer.

PAPAGEORGIOU, N. (1990): "Optimal Control of Nonlinear Evolution Inclusions," *Journal of Optimization theory and Applications*, 67, 321-354.

PRIESTLEY, M. (1981): *Spectral Analysis and Time Series*, Vol. 1. London: Academic Press.

QUAH, D. (2002): "Spatial Agglomeration Dynamics," *AEA Papers and Proceedings*, 92, 247–252.

ROMER, P. M. (1986): "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94 (5), 1002–10037.

SCHEINKMAN, J. (1976): "On Optimal Steady States of n-Sector Growth Models When Utility Is Discounted," *Journal of Economic Theory,* 12, 11–30.

TURING, A. (1952): "The Chemical Basis of Morphogenesis," *Philosophical Transactions of the Royal Society of London*, 237, 37–72.

Appendix 1

First-Order Necessary Conditions

In this Appendix we provide a heuristic discussion about first-order necessary conditions as a translation of the received work in the mathematical literature associated with optimal control in infinite dimensional spaces (e.g Ahmed 1985, Papageorgiou 1990) to our more special case, where the notation can be lightened and the first-order necessary conditions are more interpretable to economists.

We consider a distributed control system where the state and the control are respectively represented by real functions $x(t,z)$ and $u(t,z)$ of time $t \in [0,\infty) := \mathcal{T}$, and the spatial variable $z \in [-\pi,\pi] := Z$. [19] Following Appel et al. (2000), the real function $x(t,z)$ of two variables (t,z) is identified with the abstract function $x = x(t)$ of one variable $t \in T$ which takes its values in a separable Hilbert space $\mathcal{X}$ of square integrable functions,[20] $\mathcal{X} = L^2(Z)$, defined as $x(t) = x(t)(z) = x(t,\cdot)$. Similarly the control $u(t,z)$ is identified with the abstract function $u = u(t) = u(t)(z) = u(t,\cdot)$, $u(t) \in \mathcal{U} = L^2(Z)$. We make the following assumptions:

A1: *For each $[0,T] \in T$, the controls $u(t,\cdot)$ are measurable functions in t that lie in the subset $B[0,T]$ of L^2 such that*

$$\mathcal{B}[0,T] = \left\{ u(t,\cdot) \in L^2(Z) : \|u(t,\cdot)\|_{L^2(z)} \leq b(T) < \infty \right\} \tag{37}$$

where the bound $b(T)$ is finite but may depend upon T. We call such controls L^2- bounded measurable controls.[21]

A2: *The set of pairs (x,u) is admissible if for each T, u is L^2- bounded measurable control in $B[0,T]$ and x is uniformly L^2- bounded on $[0,T]$.*

Long-range spatial effects describing the effects of the concentration of the state variable $x(t,z')$ in locations z' on $x(t,z)$ are modelled using the

[19] Circular spaces of the form $Z := [-|Z|,|Z|]$ can be handled by a change in units.

[20] A square integrable function $v(z)$ in the interval $a \leq z \leq b$ satisfies the condition $\int_a^b |v(z)|^2 dz < \infty$.

[21] By Carleson's theorem each such $u(t,\cdot)$ has a Fourier series that converges pointwise for almost all z.

kernel formulation, as:

$$X(t,z) = \int_{z' \in Z} w(z-z')\, x(t,z')\, dz' := (\mathbf{K}x)(t,z). \tag{38}$$

A3: *The kernel function $w(\cdot)$ is continuous and symmetric.*

Assumption A3 implies that the operator $\mathbf{K}$ defined by $(\mathbf{K}x)(t,z)$ in (38) is a compact linear operator that maps the Hilbert space $L^2(Z;\mathbb{R})$ to itself.[22] The kernel function quantifies the impact of site z' on site z. Spatial impacts are assumed to be symmetric, or $w(z-z') = w(z'-z)$.

When spatial spillovers are combined with a temporal growth function $g(x(t,z), u(t,z), X(t,z))$, the rate of change of the state x at time t and location z depends also on the values of the state at locations $z' \in Z$ and can be written as:

$$\frac{\partial x(t,z)}{\partial t} = g(x(t,z), u(t,z), X(t,z)) \ , \ x(0,z) = x_0(z) \tag{39}$$

where $x_0(\cdot) \in L^2(Z)$. Equation (39) describes the effects of spatial spillovers on the evolution of the system's state (e.g. capital stock, knowledge, technology) both in the time and the space domain. Using the identification of $x(t,z), u(t,z)$ as $x = x(t)$, $u = u(t)$ which take values in separable Hilbert spaces, (39) can be written as the ordinary differential equation in $\mathcal{X} \times \mathcal{U}$:

$$\frac{dx}{dt} = g(x,u,X), \ x(0) = x_0. \tag{40}$$

A4: *The function* $g(x,u,X)$, $x = x(t), u = u(t)$, $X = X(t) = X(t,\cdot) = (\mathbf{K}x)(t,\cdot)$ *satisfies a Lipschitz condition; is* C^2; *and* $\sup \|\partial g / \partial v\|$, $v = x,u,X$ *is bounded. We also assume enough regularity in* (40) so that L^2*- bounded measurable control inputs yield* L^2*- bounded state outputs.*

[22] A linear operator $\mathcal{A} : H \to \mathcal{H}$, where $H, \mathcal{H}$ are Hilbert spaces, is compact if the image of every bounded subset B of H under $\mathcal{A}$ is relatively compact in $\mathcal{H}$. For a continuous or square integrable kernel $w(\cdot)$, the operator $\mathcal{A}$ is compact. For a symmetric L^2-kernel the operator is self-adjoint. An operator $\mathcal{A}$ is linear if $\mathcal{A}(\beta_1 f_1 + \beta_2 f_2) = \beta_1(\mathcal{A}f_1) + \beta_2(\mathcal{A}f_2)$, for constants β_1, β_2 and square integrable functions f_1, f_2. Thus the operator we use is compact, linear and self-adjoint (for details see e.g., Dieudonne 1960, Chapter XI). To simplify notation, sometimes we write $\mathbf{K}v$ instead of $(\mathbf{K}v)(t,z)$ for some function $v(t,z)$.

From A4 (40) has an L^2- bounded solution $x(t)$, for a given u. The solution $x = x(t)$ defines a generalized solution $x(t,z) = x(t)(z)$ of the integro-differential equation (39).[23] Equation (39) can be used as a dynamic constraint in an optimal control problem where the objective is to choose admissible controls $u(t,z)$ which will maximize discounted benefits over the spatial domain Z associated with a payoff function. We associate with the control system (38)-(39) the payoff expression

$$J(x_0(z), u(\cdot)) = \int_0^{\infty} e^{-\rho t} F(x(t), u(t), X(t))\, dt. \tag{41}$$

A5: *$F(\cdot,\cdot,\cdot)$ is a differentiably concave and upper-semicontinuous function on $L^2 \times L^2 \times L^2$ that satisfies the coercivity condition $F(x(t), u(t), X(t)) \leq a - c\,\|x(t)\|_{L^2(Z)}$, $(a,c) > (0,0)$,* (Leizarowitz 2008).

The payoff functional which corresponds to the maximization of discounted benefits over the entire spatial domain is defined by:

$$F(x(t), u(t), X(t)) = \int_{-\pi}^{\pi} f(x(t,z), u(t,z), X(t,z))\, dz. \tag{42}$$

A6: *$f(\cdot,\cdot,\cdot)$ is a differentiable concave and upper semi-continuous function $f : \mathbb{R}^3 \to \mathbb{R}$.*

When the payoff is maximized over the entire spatial domain, then the distributed parameter optimal control problem can be stated as:

$$\max_{\{u(t)\}} J(x,u,X) = \int_0^{\infty} e^{-\rho t} F(x(t), u(t), X(t))\, dt \tag{43}$$

subject to (39) and (38).

To develop a version of the maximum principle for this problem, we first

[23] A generalized solution $x = x(t,z)$ is measurable on $\mathcal{T} \times [-\pi,\pi]$, $x(\cdot,z)$ is absolutely continuous on $\mathcal{T}$ for each $z \in [-\pi,\pi]$ and satisfies (3) almost everywhere. For details see Appel et al. (2000, Chapter 1).

consider a fixed terminal time, free endpoint problem with discounting.

$$\max_{\{u(t,z)\}} J(x,u,X) = \int_0^{t_1} \int_{-\pi}^{\pi} e^{-\rho t} f(x(t,z), u(t,z), X(t,z))\, dzdt \tag{44}$$

subject to (39) and (38)

$$x(t_1, z) \text{ free for all } z \in Z. \tag{45}$$

Suppose that $u^*(t) = u^*(t,\cdot) = u^*(t,z)$ is an optimal control function for problem (44) and let $x^*(t) = x^*(t,\cdot) = x^*(t,z)$ represent the optimal path for the state of the system when $x(0,z) = x_0(z)$. We select an $\varepsilon > 0$ and we define the variation (see, for example, Athans and Falb 1966, Evans 2008)

$$u(t;\varepsilon) = u^*(t) + \varepsilon\beta(t) \ \ , 0 \leq t \leq t_1 \tag{46}$$

where $\beta(t) = \beta(t,\cdot)$ is a function which satisfies assumptions similar to A1 and which is selected such that $u(t;\varepsilon)$ satisfies A2 for all sufficiently small $\varepsilon > 0$. We call the function $\beta(t)$ an *acceptable variation* and we assume that such a function exists. Let $x^\varepsilon(t) = x(t;\varepsilon) = x(t,\cdot;\varepsilon)$ be a solution of (40) corresponding to $u^\varepsilon(t) = u(t;\varepsilon) = u^\varepsilon(t,\cdot;\varepsilon)$ which satisfies A2, and let $X^\varepsilon = X(t;\varepsilon) = (\mathbf{K}x^\varepsilon)(t,\cdot)$. Since a small variation in the control generates a small variation in the motion of the system,

$$x^\varepsilon(t) = x^*(t) + \varepsilon\psi(t) \tag{47}$$

where $\psi(t) = \psi(t,\cdot)$. Since x^ε is a solution of (40) we have, dropping t to ease notation,

$$\dot{x}^\varepsilon = g(x^\varepsilon, u^\varepsilon, X^\varepsilon) = g(x^* + \varepsilon\psi, u^* + \varepsilon\beta, X^* + \varepsilon\Psi) \tag{48}$$

where by the linearity of the integral operator, $X^\varepsilon = \mathbf{K}x^\varepsilon = \mathbf{K}(x^* + \varepsilon\psi) = \mathbf{K}x^* + \varepsilon\mathbf{K}\psi = X^* + \varepsilon\Psi$. Furthermore

$$\dot{x}^\varepsilon = \dot{x} + \varepsilon\dot{\psi}. \tag{49}$$

Expanding the right hand side of (48) around $\varepsilon = 0$, and denoting Fréchet derivatives with subscripts, we obtain

$$
\begin{aligned}
& g\left(x^*+\varepsilon\psi, u^*+\varepsilon\beta, X^*+\varepsilon\Psi\right)= \\
& g\left(x^*, u^*, X^*\right)+g_x\left(x^*, u^*, X^*\right)\varepsilon\psi+ \\
& g_u\left(x^*, u^*, X^*\right)\varepsilon\beta+g_X\left(x^*, u^*, X^*\right)\varepsilon\Psi+o(\varepsilon)
\end{aligned} \tag{50}
$$

where $o(\varepsilon)$ is defined in the L^2 norm sense, and $\lim_{\varepsilon\to 0}\frac{o(\varepsilon)}{\varepsilon}=0$. Using (47), (49) and (50),

$$
\varepsilon\dot{\psi} = g_x^*\varepsilon\psi+g_u^*\varepsilon\beta+g_X^*\varepsilon\Psi+o(\varepsilon) \quad \text{or} \tag{51}
$$

$$
\dot{\psi} = g_x^*\psi+g_u^*\beta+g_X^*\Psi+\frac{o(\varepsilon)}{\varepsilon} \tag{52}
$$

where $(^*)$ indicates that all derivatives are evaluated along the trajectory (x^*, u^*, X^*) and $\psi(0)=0$, since the trajectory $x(t,\cdot;\varepsilon)$ starts at $x(0;\varepsilon)=x_0(z)$.

Suppose that $y(t)=y(t,\cdot)=y(t)(z)$ is the solution of the linear ordinary differential equation

$$
\dot{y} = g_x\left(x^*, u^*, X^*\right)y+g_u\left(x^*, u^*, X^*\right)\beta+g_X\left(x^*, u^*, X^*\right)Y \tag{53}
$$

$$
y(0) = 0\ ,\ Y=(\mathbf{K}y)(t,\cdot)\,. \tag{54}
$$

Then it follows that

$$
x^\varepsilon(t)=x^*(t)+\varepsilon\psi(t)=x^*(t)+\varepsilon y(t)+o(\varepsilon) \tag{55}
$$

with $\dot{y}$ given by (53)-(54). In the following we will replace $x^\varepsilon(t)$ by $x^*(t)+\varepsilon y(t)$ instead of the more strictly accurate $x^*(t)+\varepsilon\psi(t)$, and $X^\varepsilon(t)$ with $X^*(t)+\varepsilon Y(t)$.

Since (x^*, u^*, X^*) is optimal,

$$
J\left(x^\varepsilon, u^\varepsilon, X^\varepsilon\right)=J(\varepsilon)-J\left(x^*, u^*, X^*\right)\le 0,\ \ (=0,\ \text{when}\ \varepsilon=0)\,. \tag{56}
$$

Thus $J(\varepsilon)$ assumes its maximum at $\varepsilon=0$, which implies $\left.\frac{dJ(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0}\le 0$, or

$$\left.\frac{dJ(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} \tag{57}$$

$$= \int_{-\pi}^{\pi}\left[\int_0^{t_1} e^{-\rho t}\frac{d}{d\varepsilon}\left(f\left(x^*+\varepsilon y, u^*+\varepsilon\beta, X^*+\varepsilon Y\right)\right)dt\right]dz =$$

$$\int_{-\pi}^{\pi}\left[\int_0^{t_1} e^{-\rho t}\left(f_x^* y + f_u^*\beta + f_X^* Y\right)dt\right]dz \leq 0 \tag{58}$$

where $(^*)$ indicates that all derivatives are evaluated along the trajectory (x^*, u^*, X^*). To define the adjoint dynamics we introduce an adjoint equation for a costate variable. The costate variable is a real function $\lambda(t, z)$, of two variables (t, z) and is identified with the abstract function $\lambda = \lambda(t)$ of one variable $t \in T$ which takes its values in a separable Hilbert space of square integrable functions and is defined by $\lambda(t) = \lambda(t)(z) = \lambda(t, \cdot)$. We require that the costate variable satisfy the ordinary differential equation

$$\dot{\lambda} = \rho\lambda - \left(f_x^* + \mathbf{K} f_X^*\right) - \left(\lambda g_x^* + \mathbf{K}\lambda g_X^*\right) \tag{59}$$

$$\text{with } e^{-\rho t_1}\lambda(t_1) = 0. \tag{60}$$

Taking the time derivative $\frac{d\left(e^{-\rho t}\lambda y\right)}{dt}$ we obtain

$$\frac{d\left(e^{-\rho t}\lambda y\right)}{dt} = e^{-\rho t}\left(-\rho\lambda y + \dot{\lambda} y + \lambda\dot{y}\right). \tag{61}$$

Integrating the right hand with respect to time, using integration by parts for the term $\lambda\dot{y}$, and noting that $y(0) = 0$ we obtain

$$\int_0^{t_1} e^{-\rho t}\left(-\rho\lambda y + \dot{\lambda} y + \lambda\dot{y}\right)dt \tag{62}$$

$$\int_0^{t_1} e^{-\rho t}\left(-\rho\lambda y + \dot{\lambda} y\right)dt + \left[e^{-\rho t_1}\lambda(t_1)\, y(t_1) - \lambda(0)\, y(0)\right] - \tag{63}$$

$$\int_0^{t_1} e^{-\rho t}\left(-\rho\lambda y + \dot{\lambda} y\right)dt = 0 \tag{64}$$

or using (53)

$$\int_0^{t_1} e^{-\rho t} \left[-\rho\lambda y + \dot{\lambda} y + \lambda g_x^* y + \lambda g_u^* \beta + \lambda g_X^* \mathbf{K} y \right] dt = 0. \tag{65}$$

Substituting this into (58) we obtain

$$\int_{-\pi}^{\pi} \left[\int_0^{t_1} e^{-\rho t} \left[\dot{\lambda} y - \rho\lambda y + f_x^* y + f_X^* \mathbf{K} y + \lambda g_x^* y + \lambda g_X^* \mathbf{K} y + (f_u^* + g_u^*) \beta \right] dt \right] dz \leq 0. \tag{66}$$

By changing the order of integration we note that two terms of the form

$$\int_{-\pi}^{\pi} f_X^* \int_{-\pi}^{\pi} w(z - z') y(t, z') dz' dz, \quad \int_{-\pi}^{\pi} \lambda g_X^* \int_{-\pi}^{\pi} w(z - z') y(t, z') dz' dz \tag{67}$$

appear above. Let ϕ denote f_X^*, or λg_X^*, then the two terms can be written as

$$\int_{-\pi}^{\pi} \phi \int_{-\pi}^{\pi} w(z - z') y(t, z') dz' dz. \tag{68}$$

By changing the order of integration we obtain

$$\int_{z' \in Z} \left[\int_{z \in Z} \phi w(z - z') dz \right] y(z') dz'.$$

Since the integration area is the same by re-labeling z as z' and z' as z, we obtain

$$\int_{z \in Z} \phi \left[\int_{z' \in Z} w(z - z') y(t, z') dz' \right] dz = \tag{69}$$

$$\int_{z \in Z} \left[\int_{z' \in Z} \phi w(z' - z) dz' \right] y(t, z') dz = \tag{70}$$

$$\int_{z \in Z} (\mathbf{K}\phi) y(t, z) dz. \tag{71}$$

Substituting into (66) we finally obtain

$$\int_{-\pi}^{\pi} \left[\int_0^{t_1} e^{-\rho t} \left[\left(\dot{\lambda} - \rho\lambda + f_x^* + \mathbf{K} f_X^* + \lambda g_x^* + \mathbf{K} \lambda g_X^* \right) y + (f_u^* + \lambda g_u^*) \beta \right] dt \right] dz \leq 0 \text{ or} \tag{72}$$

$$\int_{-\pi}^{\pi} \left[\int_{0}^{t_1} e^{-\rho t} \left(f_u^* + \lambda g_u^* \right) \beta dt \right] dz \leq 0 \tag{73}$$

because of (59). Inequality (73) must hold for every acceptable variation $\beta\left(t,\cdot\right)$. If we define the current value Hamiltonian function as

$$H\left(x,u,X,\lambda\right) = f\left(x,u,X\right) + \lambda g\left(x,u,X\right) \tag{74}$$

the coefficient of β in (73) is $\frac{\partial H(x^*,u^*,X^*,\lambda)}{\partial u}$. This implies that given $x^*\left(t,\cdot\right), X^*\left(t,\cdot\right)$ and $\lambda\left(t,\cdot\right)$ the optimal control $u^*\left(t,\cdot\right)$ should be selected to attain an extremum for the current value Hamiltonian function $H\left(x^*,u,X^*,\lambda\right)$ among admissible control functions.

If we let $t_1 \to \infty$, then (60) can be written as

$$\lim_{t \to \infty} e^{-\rho t} \lambda\left(t,\cdot\right) = 0 \tag{75}$$

which motivates the temporal transversality condition at infinity. Collecting our results we can state that:

In order for $u^*\left(t,z\right)$ to be optimal for problem (7), it is necessary that there exist a function $\lambda\left(t,z\right)$ such that:

(i) $x^*\left(t,z\right)$ and $\lambda\left(t,z\right)$ are a solution of the system

$$\begin{aligned} \frac{\partial x}{\partial t} &= g\left(x^*,u^*,X^*\right) = H_\lambda\left(x^*,u^*,X^*\right) \ , & (76) \\ x\left(0\right) &= x_0\left(z\right), \ X = \left(\mathbf{K}x\right)\left(t,z\right) \\ \frac{\partial \lambda}{\partial t} &= \rho\lambda - \left(f_x^* + \lambda g_x^*\right) - \left(\mathbf{K}f_X^* + \mathbf{K}\lambda g_X^*\right) = & (77) \\ & \rho\lambda - H_x\left(x^*,u^*,X^*\right) - \mathbf{K}H_X\left(x^*,u^*,X^*\right) \end{aligned}$$

satisfying a temporal limiting transversality condition,[24] and spatial transver-

[24] For a concave problem (43) the limiting transversality condition is a necessary condition (e.g. Benveniste and Scheinkman 1982). For general cases see Ekeland and Scheinkman (1986) for discrete time, and Kamihigashi (2001) for continuous time.

sality conditions for a finite spatial domain with circle boundary conditions,

$$\lim_{T\to\infty} e^{-\rho T} \int_{z\in Z} \lambda(T,z)\, x(T,z)\, dz = 0 \tag{78}$$

$$\lambda(t,-\pi) = \lambda(t,\pi),\ x(t,-\pi) = x(t,\pi),\ \text{for all } t. \tag{79}$$

With circle boundary conditions for the state variable $x(t,-\pi) = x(t,\pi) = \hat{x}(t)$, similar spatial transversality conditions $\lambda(t,-\pi) = \lambda(t,\pi)$ for all t, should be satisfied for the costate variable for the solution of the system of equations (76)-(77).

(ii) The Hamiltonian function $H(x^*, u, X^*, \lambda)$ has a (possibly local) maximum as a function of u at $u^*(t,z)$ for all $t \geq 0$. For an interior maximum,

$$\frac{\partial H}{\partial u} = 0,\ \mathit{or}\ f_u + \lambda g_u = 0 \Rightarrow u^* = u^*(x, \lambda, X)\,. \blacksquare \tag{80}$$

Appendix 2

Sufficient Conditions

Assume that $f(x,u,X)$ and $g(x,u,X)$ are differentiably concave functions. Suppose that $x^*(t,z)$, $u^*(t,z)$, $\lambda(t,z)$ satisfy conditions (76)-(80) and that $\lambda \geq 0$, and let functions $x(t,z)$, $u(t,z)$ satisfy (39) and initial and boundary conditions. Let f^*, g^* denote functions evaluated along $(x^*(t,z), u^*(t,z), X^*(t,z))$ and let f, g denote functions evaluated along the feasible path $(x(t,z), u(t,z), X(t,z))$. To prove sufficiency we need to show that

$$W \equiv \int_{z\in Z} \int_0^\infty e^{-\rho t} (f^* - f)\, dt dz \geq 0. \tag{81}$$

From the concavity of f it follows that

$$(f^* - f) \geq (x^* - x) f_x^* + (u^* - u) f_u^* + (X^* - X) f_X^*. \tag{82}$$

Setting $X = \mathbf{K}x$, and using, for functions $\phi_1(t,z), \phi_2(t,z)$, reasoning similar to (69)-(71) to write $\int_z \phi_1(\mathbf{K}\phi_2)\, dz = \int_z \phi_1(\mathbf{K}\phi_2)\, dz$, we obtain:

$$\int_{z\in Z} (f^* - f)\, dz \geq \int_{z\in Z} \left[(x^* - x)(f_x^* + \mathbf{K} f_X^*) + (u^* - u) f_u^*\right] dz.$$

Then,

$$W \geq \int_{z\in Z}\int_0^\infty e^{-\rho t}\left[(x^*-x)\left(f_x^*+\mathbf{K}f_X^*\right)+(u^*-u)\,f_u^*\right]dtdz = \tag{83}$$

$$\int_{z\in Z}\int_0^\infty e^{-\rho t}\left[(x^*-x)\left(\rho\lambda-\frac{\partial\lambda}{\partial t}-\lambda g_x^*-\mathbf{K}\lambda g_X^*\right)+(u^*-u)\left(-\lambda g_u^*\right)\right]dtdz = \tag{84}$$

$$\int_{z\in Z}\int_0^\infty e^{-\rho t}\lambda\left[(g^*-g)-(x^*-x)\,g_x^*-(X^*-X)\,g_X^*-(u^*-u)\,g_u^*\right]dtdz \geq 0. \tag{85}$$

Condition (84) follows from (83) by using conditions (77) and (80) to substitute for f_u^* and $f_x^*+\mathbf{K}f_X^*$. Condition (85) is derived from (84) in the following way.

(1) The term $\int_0^\infty e^{-\rho t}(x^*-x)\left(\rho\lambda-\frac{\partial\lambda}{\partial t}\right)dt$ is replaced, after integrating by parts $\int_0^\infty e^{-\rho t}\lambda\frac{\partial x}{\partial t}dt$ and rearranging terms, by:

$$\int_0^\infty e^{-\rho t}\lambda\left(\frac{\partial x^*}{\partial t}-\frac{\partial x}{\partial t}\right)dt \tag{86}$$

$$\frac{\partial x^*}{\partial t}=g^*,\ \frac{\partial x}{\partial t}=g.$$

(2)

$$\int_{z\in Z}(x^*-x)\left(\mathbf{K}\lambda g_X^*\right)dz=\int_{z\in Z}\lambda\left(\mathbf{K}\left(x^*-x\right)g_X^*\right)dz= \tag{87}$$

$$\int_{z\in Z}\lambda\left(X^*-X\right)g_X^*dz. \tag{88}$$

Substituting (86) and (88) into (84), the first term of (84) can be written as:

$$\int_{z\in Z}\int_0^\infty e^{-\rho t}\left[\lambda\left(g^*-g\right)-\lambda\left(x^*-x\right)g_x^*-\lambda\left(X^*-X\right)g_X^*\right]dtdz. \tag{89}$$

Finally by substituting (89) into (84) we obtain (85) which holds by the concavity assumption about g and the assumption that $\lambda \geq 0$. ■

Appendix 3

Proof of Proposition 1

The linearized Hamiltonian system of problem (43) at the FOSS can be written as:

$$\frac{\partial x}{\partial t} = H^*_{\lambda x}x + H^*_{\lambda X}\mathbf{K}x + H^*_{\lambda\lambda}\lambda \tag{90}$$

$$\frac{\partial \lambda}{\partial t} = \left(-H^*_{xx} - 2H^*_{Xx}\right)x - H^*_{XX}\mathbf{K}\left(\mathbf{K}x\right) + \left(\rho - H^*_{x\lambda}\right)\lambda - H^*_{X\lambda}\mathbf{K}\lambda \tag{91}$$

where the superscript $(^*)$ indicates that the Fréchet derivatives are evaluated at (x^*, λ^*). Our approach in studying the stability of the FOSS to spatially heterogeneous perturbations off the FOSS, is to transform the infinite dimensional system (90)-(91) into a countable sequence of linear systems of ordinary differential equations so that we can use linear stability analysis. To obtain this we consider pairs of square integrable solutions $(x(t)(z), \lambda(t)(z)) = (x(t,z), \lambda(t,z))$ and we construct trial solutions using an orthogonal basis of $L^2(Z)$ created in terms of functions $\cos(kz)$, $\sin(kz)$, $z \in [-\pi, \pi]$, for mode $k = 0, 1, 2, ...$ which form a complete orthogonal basis over $[-\pi, \pi]$. Our assumptions about functions f and g suggest that the solution $(x(t,z), \lambda(t,z))$ of the optimal control problem will be smooth enough for it to be expressed in terms of Fourier basis, as:

$$x(t,z) = e^{\sigma t}\sum_{k=0}^{\infty}\left[\alpha_{1k}\cos(kz) + \alpha_{2k}\sin(kz)\right], \; z \in [-\pi.\pi] \tag{92}$$

$$\lambda(t,z) = e^{\sigma t}\sum_{k=0}^{\infty}\left[A_{1k}\cos(kz) + A_{2k}\sin(kz)\right] \tag{93}$$

where $a_k(t) = \left(e^{\sigma t}a_{1k}, e^{\sigma t}a_{2k}\right)$, $A_k(t) = \left(e^{\sigma t}A_{1k}, e^{\sigma t}A_{2k}\right)$ are the Fourier coefficients. As shown in Priestley (1981, Section 4.2), any function in the class $L^2(-\pi, \pi)$ has a Fourier expansion which converges to the function in the mean square sense, and any continuous and bounded variation function of $(-\pi, \pi)$ has a Fourier series expansion which converges to values of the function in $(-\pi, \pi)$. To save on notation let $v(t,z)$ stand for $x(t,z)$ or $\lambda(t,z)$. Since $v(t,\cdot) \in L^2(-\pi, \pi)$ for each t, $v(t,\cdot)$ has a Fourier series. This is so because the Fourier coefficients $\alpha_k(t)$, $A_k(t)$ exist for each t and the Fourier series converges for each t (Priestley 1981, Section 4.2.1). As shown by Priestley the Fourier basis is a complete orthogonal basis for each t. For the Fourier coefficients to be Lipchitz in t on compact subsets $T_c \in [0, \infty)$, some more regularity is required. This regularity assures that solutions exist for the integral equation for each mode k. As one can see from Priestley (1981, equations (4.2.5), (4.2.6)), a sufficient condition for the Fourier coefficients of $v(t, z)$ to be uniformly Lipchitz in t on compact subsets $T_c \in [0, \infty)$ is the following: For each compact subset T_c of $[0, \infty)$, there is $L(T_c)$, $0 < L(T_c) < \infty$, such that for all $(t, t') \in T_c$, we have $|v(t, z) - v(t', z)| \leq L(T_c)|t - t'|$, for all $z \in Z$.

By the symmetry of the kernel $w(z - z') = w(z' - z)$, setting $\zeta = z' - z$ we obtain $\int_{z' \in Z} w(z' - z)\, v(t, z')\, dz' = \int_{\zeta \in Z} w(\zeta)\, v(t, \zeta + z)\, d\zeta$, $v = x, \lambda$. Substituting the trial solution under the integral we obtain, dropping t to simplify notation:

$$
\begin{aligned}
(\mathbf{K}v)(t, z) &= \int_{\zeta \in Z} w(\zeta)\, v(\zeta + z)\, d\zeta = \\
& e^{\sigma t} \int_{\zeta \in Z} w(\zeta) \sum_{k=0}^{\infty} \left[\alpha_k^v \cos(k(\zeta + z)) + \beta_k^v \sin(k(\zeta + z))\right] d\zeta.
\end{aligned}
$$

Using the formulas

$$
\begin{aligned}
\cos(A + B) &= \cos A \cos B - \sin A \sin B \\
\sin(A + B) &= \sin A \cos B + \sin B \cos A
\end{aligned}
$$

and noting that because of the symmetry of the kernel $w(\zeta) = w(-\zeta)$, it holds that $\int_{-|Z|}^{|Z|} w(\zeta) \sin(k\zeta)\, d\zeta = 0$ for any constants Z, k, we obtain for the terms $\mathbf{K}x$ and $\mathbf{K}\lambda$ in (90)-(91):

$$\begin{aligned}
(\mathbf{K}x)(t,z) &= e^{\sigma t}\sum_k \left[a_{1k}\cos(kz)+a_{2k}\sin(kz)\right]\int_{\zeta\in Z} w(\zeta)\cos(k\zeta)\,d\zeta = x(t,z)\,W(k)\\
(\mathbf{K}\lambda)(t,z) &= e^{\sigma t}\sum_k \left[A_{1k}\cos(kz)+A_{2k}\sin(kz)\right]\int_{\zeta\in Z} w(\zeta)\cos(k\zeta)\,d\zeta = \lambda(t,z)\,W(k)\\
W(k) &= \int_{\zeta\in Z} w(\zeta)\cos(k\zeta)\,d\zeta.
\end{aligned}$$

For the term $\mathbf{K}(\mathbf{K}x)$ in (90) we have:

Put $y(t,z)=(\mathbf{K}x)(t,z)$, then

$$\begin{aligned}
&\mathbf{K}(\mathbf{K}x)(t,z) = (\mathbf{K}y)(t,z) =\\
&\int w(\zeta)\left(\int w(\zeta)\,e^{\sigma t}\sum_{k=0}^{\infty}\left[\alpha_k^v\cos(k(\zeta+z))+\beta_k^v\sin(k(\zeta+z))\right]d\zeta\right)d\zeta =\\
&\int w(\zeta)\,x(t,\zeta+z)\,W(k)\,d\zeta = W(k)\int w(\zeta)\,x(t,\zeta+z)\,d\zeta =\\
&W(k)\int w(\zeta)\,e^{\sigma t}\sum_{k=0}^{\infty}\left[\alpha_k^v\cos(k(\zeta+z))+\beta_k^v\sin(k(\zeta+z))\right]d\zeta =\\
&W(k)\,x(t,z)\,W(k) = W^2(k)\,x(t,z).
\end{aligned}$$

Substituting the rest of the trial solutions into (90)-(91) and collecting terms we obtain

$$\frac{dx_k}{dt} = \left[H_{\lambda x}^* + H_{\lambda X}^* W(k)\right]x_k + H_{\lambda\lambda}^*\lambda_k \tag{94}$$

$$\begin{aligned}\frac{d\lambda_k}{dt} &= \left[-H_{xx}^* - 2H_{Xx}^* W(k) - H_{XX}^* W^2(k)\right]x_k + \\ &\left[\rho - H_{x\lambda}^* - H_{X\lambda}^* W(k)\right]\lambda_k.\end{aligned} \tag{95}$$

This is a sequence of linear systems of ordinary differential equations indexed by k which corresponds to mode k. Mode $k=0$ and $W(0)$ correspond to a spatially homogenous system.[25] From the Jacobian matrix J_k^*,

$$J_k^* = \begin{pmatrix} H_{\lambda x}^* + H_{\lambda X}^* W(k) & H_{\lambda\lambda}^* \\ -H_{xx}^* - 2H_{Xx}^* W(k) - H_{XX}^* W^2(k) & \rho - H_{x\lambda}^* - H_{X\lambda}^* W(k) \end{pmatrix} \quad (96)$$

of the sequence of the linear systems (94)-(95) it follows that mode k is saddle point stable if the pair of eigenvalues of (94)-(95) have opposite signs, and it is unstable if both eigenvalues are real and positive or complex with positive real parts. In (96) trace$J_k^* = \rho > 0$, while the determinant defines a quadratic expression in terms of $W(k)$. This is the dispersion relationship for the optimal control problem with spatial spillovers, which can be written as:

$$\begin{aligned} &\psi(k) = \left[H_{XX}^* H_{\lambda\lambda}^* - \left[H_{\lambda X}^*\right]^2\right] W^2(k) + \\ &\left[H_{\lambda X}^*\left(\rho - 2H_{\lambda x}^*\right) + 2H_{Xx}^* H_{\lambda\lambda}^*\right] W(k) + \left[\rho H_{\lambda x}^* - \left[H_{\lambda X}^*\right]^2 + H_{\lambda\lambda}^* H_{xx}^*\right] \end{aligned} \quad (97)$$

If there exists k such that $\psi(k) > 0$ for $k \in (k_1, k_2)$, then both eigenvalues (σ_1, σ_2) of (94)-(95) which characterize temporal growth are positive and the FOSS is not stable to spatially heterogeneous perturbations. The eigenvalues are obtained as the solution of the characteristic equation

$$\sigma^2 - \rho\sigma + \psi(k) = 0$$

with eigenvalues:

$$\sigma_{1,2}(k) = \frac{1}{2}\left(\rho \pm \sqrt{\rho^2 - 4\psi(k)}\right). \quad (98)$$

Spillovers induced spatial instability requires $\psi(k) > 0$ for $k \in (k_1, k_2)$. A linear approximation solution for (90)-(91) in the neighborhood of the FOSS can be obtained by setting the constant of the largest eigenvalue, which does

[25] A similar decomposition can be obtained by using as trial solutions $x_k(t,z) = c^x e^{\lambda t + ikz}$, $\lambda_k(t,z) = c^\lambda e^{\lambda t + ikz}$ $k = 0, 1, 2, ...$ for constants $\left(c^x, c^\lambda\right)$ and $Z = [0, 2\pi]$. In this case $W(k) = \int_\zeta w(\zeta) e^{ik\zeta} d\zeta$ is a scaled Fourier Transform and the sequence $\sqrt{\frac{1}{2\pi}} e^{ik\zeta}$ is a complete orthonormal basis in L^2.

not satisfy transversality conditions at infinity, equal to zero. ■

Appendix 4

Proof of Proposition 2

Let (x^*, λ^*) be a FOSS for problem (43), and let $B_k(z)$ be a complete orthonormal basis in $L^2(Z)$, such that the closed linear manifold generated by $B_k(z)$ is $L^2(Z)$, such that $(x(t,z) - x^*, \lambda(t,z) - \lambda^*) = \left(\sum_{k=0}^{\infty} a_k(t) B_k(z), \sum_{k=0}^{\infty} A_k(t) B_k(z)\right)$, where $(a_k(t), A_k(t))$ are Lipchitz in t on compact subsets $T_c \in [0, \infty)$. Such a basis exists since a separable Hilbert space has at least one complete orthonormal basis (e.g. Yosida 1980, Chapter III). This orthonormal basis can be constructed, for example, from the orthogonal sine/cosine basis using the Schmidt process.

Consider small perturbations $(x(t,z) - x^*, \lambda(t,z) - \lambda^*)$ off the FOSS and write

$$\xi^x(t,z) = x(t,z) - x^*, \xi^\lambda(t,z) = x(t,z) - x^*. \tag{99}$$

For the $B_k(z)$ basis we know that the *Fourier coefficients* for $\xi^x(t,z), \xi^\lambda(t,z)$ are

$$a_k(t) = \langle \xi^x(t,\cdot), B_k(\cdot) \rangle, A_k(t) = \left\langle \xi^\lambda(t,\cdot), B_k(\cdot) \right\rangle \tag{100}$$

where $\langle \cdot, \cdot \rangle$ denotes inner product. Then by Parseval's formula (Yosida 1980, Chapter 3)

$$\|\xi^x(t,\cdot)\|^2 = \sum_{k=0}^{\infty} |a_k(t)|^2, \left\|\xi^\lambda(t,\cdot)\right\|^2 = \sum_{k=0}^{\infty} |A_k(t)|^2. \tag{101}$$

Assume that another complete orthonormal basis $B'_k(z)$ exists in $L^2(Z)$ with $(A'_k(t), a'_k(t))$ Lipchitz in t on compact subsets $T_c \in [0, \infty)$. Then

$$\xi^x(t,z) = \sum_{k=0k}^{\infty\prime} a'_k(t) B'_k(z), \ \xi^\lambda(t,z) = \sum_{k=0}^{\infty} A'_k(t) B'_k(z). \tag{102}$$

Parseval's formula implies then that

$$\|\xi^x(t,\cdot)\|^2 = \sum_{k=0}^{\infty}|a_k(t)|^2 = \sum_{k=0}^{\infty}|a'_k(t)|^2 \tag{103}$$

$$\left\|\xi^\lambda(t,\cdot)\right\|^2 = \sum_{k=0}^{\infty}|A_k(t)|^2 = \sum_{k=0}^{\infty}|A'_k(t)|^2. \tag{104}$$

Then:

If there is a mode k such that $(a_k(t), A_k(t)) \to \infty$, as $t \to \infty$, the FOSS goes completely unstable for the basis $B_k(z)$ and any other complete orthonormal basis $B'_k(z)$ in $L^2(Z)$ for this mode since $\|\xi^v(t,\cdot)\|^2 \to \infty$, $v = x, \lambda$. If $\|\xi^v(t,\cdot)\|^2 \to \infty, v = x, \lambda$ then $(a_k(t), A_k(t)) \to \infty$ for at least one mode-k coefficient for the basis $B_k(z)$ and any other complete orthonormal basis $B'_k(z)$ in $L^2(Z)$.

If $(a_k(t), A_k(t)) \to 0$, as $t \to \infty$ for all $k \geq 0$, then the FOSS is stable for the basis $B_k(z)$ and any other complete orthonormal basis $B'_k(z)$ in $L^2(Z)$, since $\|\xi^v(t,\cdot)\|^2 \to 0, v = x, \lambda$. If $\|\xi^v(t,\cdot)\|^2 \to 0, v = x, \lambda$ then $(a_k(t), A_k(t)) \to 0$ for all mode-k coefficients for the basis $B_k(z)$ and any other complete orthonormal basis $B'_k(z)$ in $L^2(Z)$.

Considering as $B_k(z)$ the sine/cosine basis used in this paper, Parceval's formula, with appropriate orthonormalization, implies:

$$\|\xi^x(t,\cdot)\|^2 = \sum_{k=0}^{\infty}\left|e^{\sigma_{1k}t}a_1\right|^2 \tag{105}$$

$$\left\|\xi^\lambda(t,\cdot)\right\|^2 = \sum_{k=0}^{\infty}\left|e^{\sigma_{1k}t}A_1\right|^2 \tag{106}$$

where the mode-k coefficients of the sine/cosine basis are $(a_k(t), A_k(t)) = (e^{\sigma_{1k}t}a_1, e^{\sigma_{1k}t}A_1)$ and the square of the norm of the deviations are defined on the tangent mode-k manifold which corresponds to the smallest eigenvalue of J_k^*,[26] then:

[26]Note that the tangent manifold corresponding to the smallest eiegnvalue satisfies the temporal transversality condition at infinity. To control the system on this manifold the constants associated with the largest eigenvalue of J_k^* are set equal to zero.

If $\sigma_{1k} < 0$ for all $k \geq 0$, then all mode-k tangent manifolds are stable in the sense that the deviations from the FOSS tend to zero for all modes k as $t \to \infty$, and the FOSS is saddle point stable. The norms in (105) and (106) go to zero in this case and because of (103) and (104) the norms will go to zero as $t \to \infty$ for any other complete orthonormal basis too. If the FOSS is saddle point stable along all the mode-k tangent manifolds corresponding to the smallest eigenvalue, i.e. the norms in (105) and (106) go to zero, then all mode-k eigenvalues σ_{1k}, $k = 0, 1, 2, ...$ should be negative. Because of (103) and (104) all mode-k coefficients of any other complete orthonormal basis will go to zero.

If $\sigma_{1k} > 0$ for some mode $k > 0$, then the mode-k tangent manifold is unstable in the sense that the deviations from the FOSS tend to infinity for mode k as $t \to \infty$. The norms in (105) and (106) go to infinity in this case and because of (103) and (104) the norms will go to infinity as $t \to \infty$ for any other basis too. If the FOSS is unstable for a mode-k tangent manifold corresponding to the smallest eigenvalue, i.e. the norms in (105) and (106) go to infinity, then the mode-k eigenvalues σ_{1k} should be positive. Because of (103) and (104) at least one mode-k coefficient of any other complete orthonormal basis will go to infinity. ■

Appendix 5

Simple Exponential Kernels

We present the two simple exponential kernels with quadratic exponents which are considered in the optimal growth example.

Kernel	$w_1(\zeta) = b_1 \exp\left[-(\zeta/d_1)^2\right], b_1, d_1 > 0, \zeta = z - z'$
$z \in [-\pi, \pi]$	$W(k) = \frac{i\sqrt{\pi}}{2} b_1 d_1 \exp\left(-\frac{(d_1 k)^2}{4}\right) \times$ $\times \left[\operatorname{erf} \mathrm{i}\left(\frac{d_1 k}{2} - \frac{i\pi}{d_1}\right) + \operatorname{erf} \mathrm{i}\left(\frac{d_1 k}{2} + \frac{i\pi}{d_1}\right)\right]$
Kernel	$w_2(\zeta) = b_1 \exp\left[-(\zeta/d_1)^2\right] - b_2 \exp\left[-(\zeta/d_2)^2\right]$ $b_1 > b_2,\ d_1 < d_2$
$z \in [-\pi, \pi]$	$W(k) = \frac{i\sqrt{\pi}}{2}(A_1 - A_2),\ A_j = b_j d_j \exp\left(-\frac{(d_j k)^2}{4}\right) \times$ $\times \left[\operatorname{erf} i\left(\frac{d_j k}{2} - \frac{i\pi}{d_j}\right) + \operatorname{erf} i\left(\frac{d_j k}{2} + \frac{i\pi}{d_j}\right)\right], j = 1, 2$
	$\operatorname{erfi}(z) = \operatorname{erf}(iz/i)$: imaginary error function $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du$: the error function

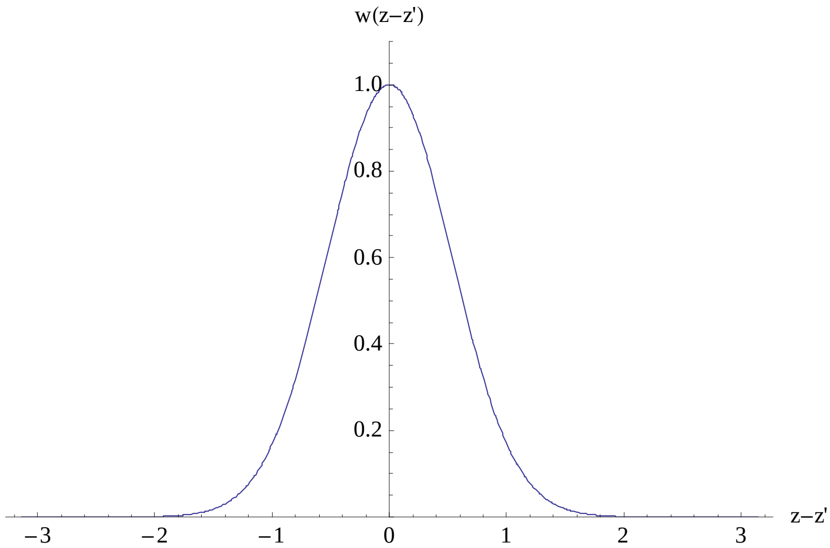

Figure 2: Kernel $w_1(z - z')$

Figures 2-5 present some typical shapes for $w(z - z')$ and the corresponding $W(k)$ in a discrete domain. Kernels of the type $w_1(z - z')$ imply that the influence of neighboring state variables on a local state variable is a weighted average of the state variable at neighboring locations, with weights decaying exponentially, and with this influence being always nonnegative. This is similar for example to Lucas's (2001) assumption for the case of labor productivity. Kernels of the type $w_2(z - z')$ imply similarly that the influence of neighboring state on local state is a weighted average of the state at neighboring locations, but that the influence from nearby locations is positive, while the influence is negative from relatively more distant locations. This is similar to Krugman's (1996) modelling of a market potential function.

Appendix 6

Optimal Spillover Induced Spatial Instability: FSS vs FOSS

We examine the strength of agglomeration forces acting on the FSS versus the FOSS with the help of a numerical example using a Cobb-Douglas production function. We assume $\alpha_1 = 0.4$, $\alpha_2 = 0.2$, $\rho = 0.03$, $\eta = 0.04$. We assume that the kernel is of the form $w_2(z - z')$ shown in Appendix 5 with $b_1 = 1, d_1 = 0.75, b_2 = 0.7, d_2 = 1$. The functions $w(\zeta)$ and $W(k)$ are shown

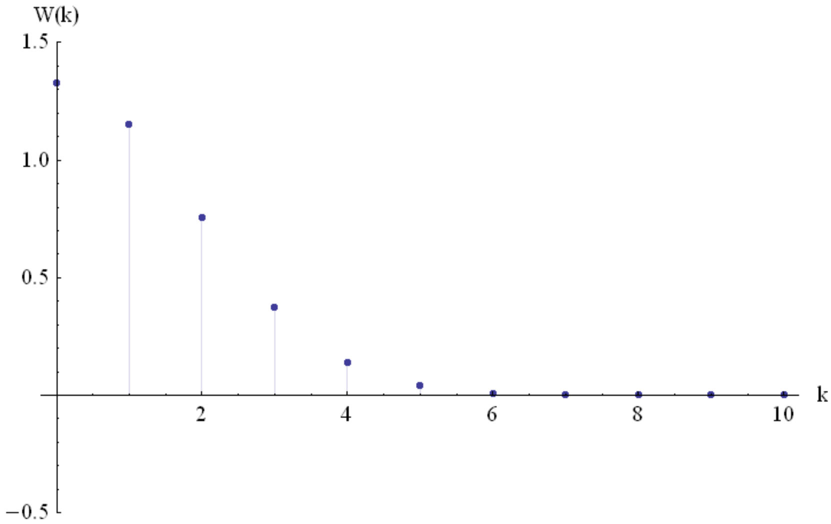

Figure 3: $W(k)$ for kernel $w_1(z - z')$

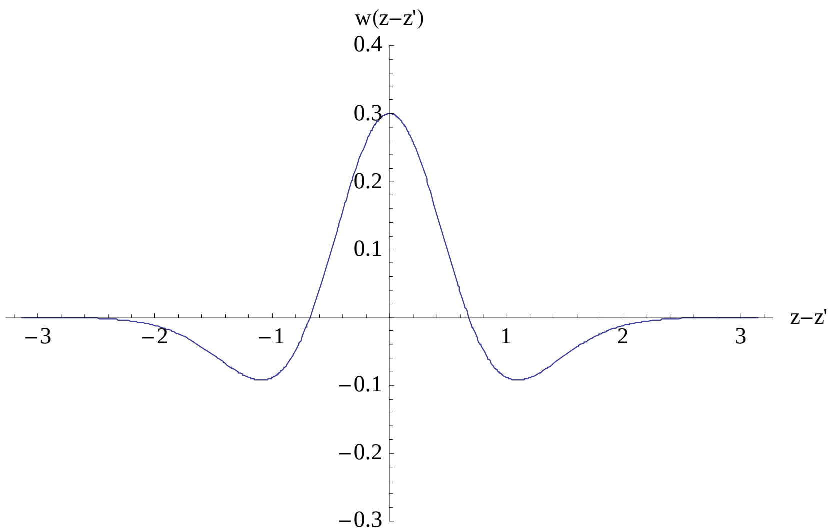

Figure 4: Kernel $w_2(z - z')$

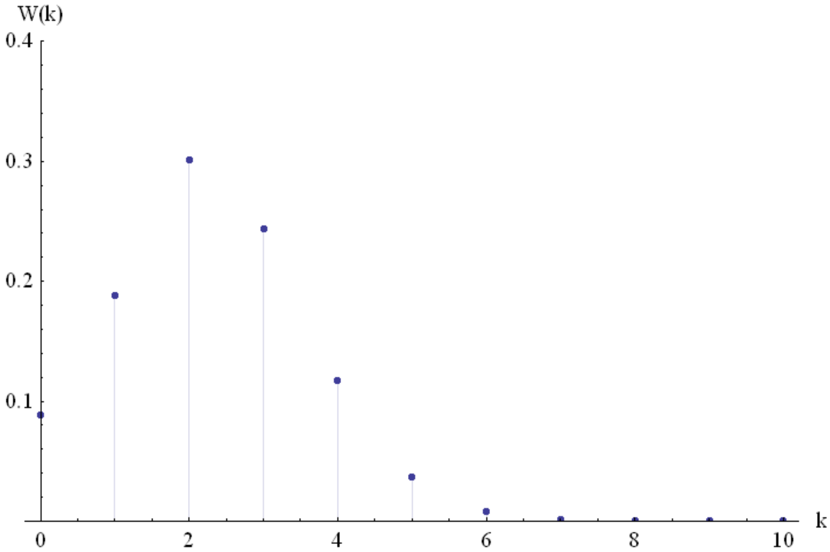

Figure 5: $W(k)$ for kernel $w_2(z - z')$

in Figures 4 and 5 respectively. The relationship $\bar{f}_{xx} + \bar{f}_{xX} W(k) = \varphi_1(k)$ is shown in Figure 6, while $T_1(k)$ is shown in Figure 7.

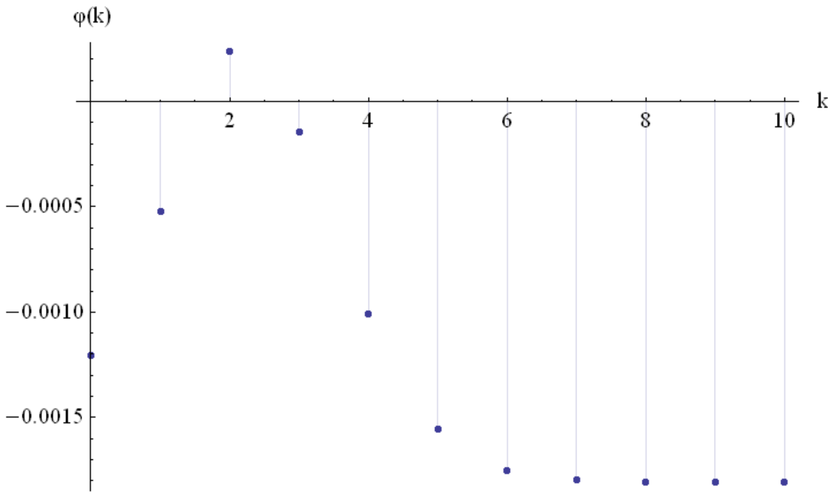

Figure 6: The dispersion relationship $\varphi(k)$

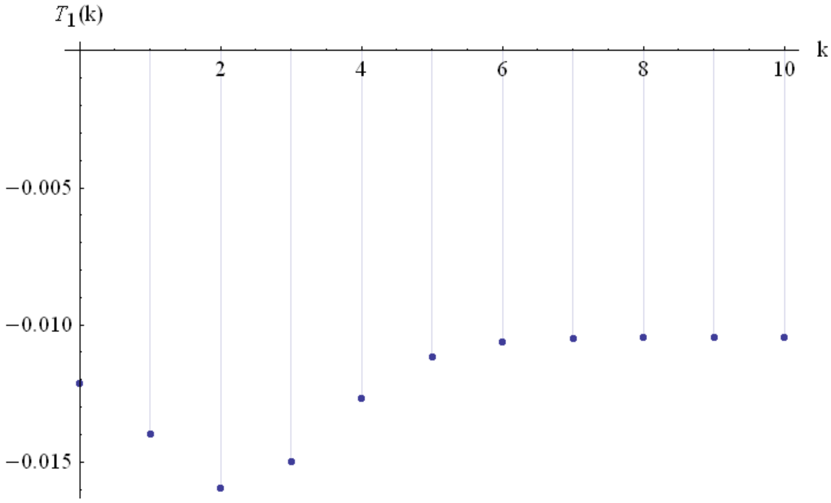

Figure 7: The relationship $T_1(k)$

For mode $k = 2$, $\varphi_1(2) > 0$, thus the FSS becomes unstable under the influence of spatial spillovers and economic agglomeration will start emerging. From Figure 5 we have that $a_1 + \frac{W(2)}{W(0)} a_2 = 1.07922$, indicating increasing social returns and instability at mode $k = 2$. The function $T_1(k)$ shown in Figure 7 is always negative, therefore that FOSS is *not* destabilized by spatial spillovers and the optimal steady state of the social planner's problem is spatially homogeneous. The same results hold for a neighborhood of parameters around the ones used above. $T_1(k)$ becomes positive, which is a necessary though not a sufficient condition for destabilization of the FOSS when the returns to externality are low and the discount rate high.

In the Cobb-Douglas example the FSS and FOSS solve respectively

$$\rho + \eta = a_1 x^{a_1 - 1} (W(0) x)^{a_2} \tag{107}$$

$$\rho + \eta = a_1 x^{a_1 - 1} (W(0) x)^{a_2} + a_2 x^{a_1} \left[(W(0) x)^{a_2 - 1} \right] W(0) \tag{108}$$

$$W(0) = \sqrt{\pi} \left[b_1 d_1 \operatorname{erf}\left(\frac{\pi}{d_1}\right) - b_2 d_2 \operatorname{erf}\left(\frac{\pi}{d_2}\right) \right]. \tag{109}$$

The values corresponding to the numerical example are $\bar{x} = 23.2383$, $x^* = 64.0372$. Since the FSS is destabilized by the spatial spillovers we seek a numerical approximation of the steady state optimal agglomeration. As can be seen by taking the optimality conditions for the rational expectations equilibrium, this agglomeration will be a function $\bar{x}(z)$ which will solve the steady state integral equation

$$
\begin{aligned}
0 &= (\rho+\eta) - a_1 (x(z))^{a_1-1} \left(\int_{-\pi}^{\pi} w(z-z')\, x(z')\, dz \right)^{a_2} && (110) \\
w(z-z') &= b_1 \exp\left[- \left(\frac{z-z'}{d_1} \right)^2 \right] - b_2 \exp\left[- \left(\frac{z-z'}{d_2} \right)^2 \right]. && (111)
\end{aligned}
$$

A search for a local numerical approximation can be conducted by choosing a set of n equal sub-intervals with length $\delta_n = 2\pi/n$ given by $-\pi = z_1 < z_2 < \ldots < z_r < \ldots < z_{n+1} = \pi$ with $z_r = -\pi + r\delta_n$. Approximating the Riemann integral in $\int_{-\pi}^{\pi} w(z-z')\, x(z')\, dz$ by a finite sum as

$$
\int_{-\pi}^{\pi} w(z-z')\, x(z')\, dz' \simeq \delta_n \sum_{m=1}^{n+1} w(z_r - z'_m)\, x(z'_m),
$$

the nonlinear integral equation (110) can be replaced by a system of nonlinear algebraic equations[27] which are written, taking logarithms, as:

$$
\begin{aligned}
&\ln\left(\frac{\rho+\eta}{a_1} \right) = (a_1 - 1) \ln x_r + a_2 \ln \left(\delta_n \sum_{m=1}^{n+1} w(z_r - z'_m)\, x(z'_m) \right) && (112) \\
&r = 1, \ldots, n+1.
\end{aligned}
$$

The system is solved in the neighborhood of the FSS for $n = 6$ and $\delta_n = \pi/3$. The results are shown in Figure 8. The flat line corresponds to the FSS which is destabilized by the spatial spillovers.

[27] This is based on the method introduced by Fredholm where the integral equation is treated as a limiting form of a finite system of linear algebraic equations.

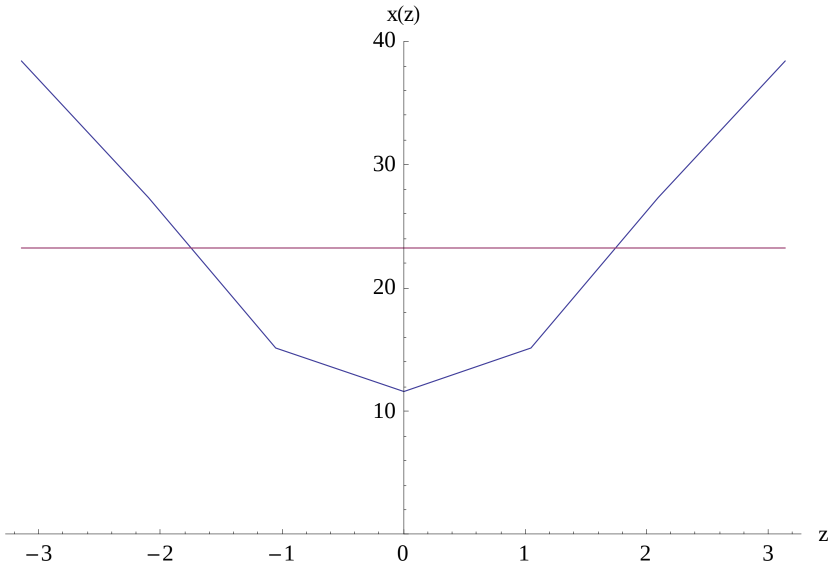

Figure 8: Equilibrium steady state agglomeration and FSS

References

AHMED, N. (1985): "On the Maximum Principle for Time-Optimal Controls for a Class of Distributed-Boundary Control Problems," *Journal of Optimization theory and Applications*, 45, 147-157.

APPEL, J., A. KALITVIN, AND P. ZABREJKO (2000): *Partial Integral Operators and Integro-differential Equations.* New York: Marcel Dekker.

ATHANS, M., AND P. FALB (1966): *Optimal Control.* New York: McGraw-Hill.

BENVENISTE, L. AND J. SCHEINKMAN (1982): "Duality Theory for Dynamic Optimization Models of Economics: The Continuous Time Case," *Journal of Economic Theory*, 27, 1-19.

DIEUDONNE, J. (1969): *Foundations of Modern Analysis.* New York: Academic Press.

EKELAND I., AND J. SCHEINKMAN (1986): "Transversality Conditions for Some Infinite Horizon Discrete Time Optimization Problems," *Mathematics of Operations Research*, 11, 216-229.

EVANS, L. (2008): "An Introduction to Mathematical Optimal Control Theory, Version 0.2," http://math.berkeley.edu/˜evans/

KAMIHIGASHI, T. (2001): "Necessity of Transversality Conditions for Infinite Horizon Problems," *Econometrica*, 69, 995-1012.

KRUGMAN, P. (1996): *The Self-Organizing Economy*. Cambridge, MA: Blackwell Publishers.

LUCAS, R. E. (2001): "Externalities and Cities," *Review of Economic Dynamics*, 4, 245–274.

PAPAGEORGIOU, N. (1990): "Optimal Control of Nonlinear Evolution Inclusions," *Journal of Optimization theory and Applications*, 67, 321-354.

PRIESTLEY, M. (1981): *Spectral Analysis and Time Series*, Vol. 1. London: Academic Press.

YOSIDA, K. (1980): *Functional Analysis,* 6th Edition. Berlin: Springer-Verlag.