DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

# Spatial Growth Theory: Optimality and Spatial Heterogeneity

Anastasios Xepapadeas

Athanasios Yannacopoulos

Working Paper Series

20-33

October 2020

# Spatial Growth Theory: Optimality and Spatial Heterogeneity

Anastasios Xepapadeas[1] and Athanasios. Yannacopoulos[2]
[1] Athens University of Economics and Business and University of Bologna
[2] Athens University of Economics and Business

**Abstract**

Spatiotemporal dynamics are introduced in a standard Ramsey model of optimal growth in which capital moves towards locations where the marginal productivity of capita is relatively higher. We extend Pontryagin's maximum principle to account for transition dynamics governed by a nonlinear partial differential equation emerging for spatial capital flows. The potential spatial heterogeneity of optimal growth as seen from the point of view of an optimizing social planner is examined. Our results suggest that for high utility discount rate the spatial capital flows induce the emergence of optimal spatial patterns while hor low utility discount a flat-earth steady state is socially optimal. Furthermore, when spatial heterogeneities exist due to total factor productivity differences across locations, we identify conditions under which the spatial capital flows could intensify or weaken spatial inequalities.

**Keywords:** Ramsey model, spatiotemporal dynamics, flat earth, pattern formation.

**JEL Classification:** O41, R11, C61, C62

## 1 Introduction

Optimal growth theory both in the context of traditional and new growth theory has been studied in a temporal domain (e.g. Aghion and Howitt 1998, Baro and Salai Martin 2005, Acemoglu 2008). The explanation of the temporal evolution of key variables such as output or capital per capita, the capital-output or the capital-labour ratio, or the evolution of positive externalities with temporal structure has been central to dynamic optimal growth models.

However, space and geography seems to be important when studying economic growth - Acemoglou ( 2008 in Chapter 1) points out the great inequality in income per capita and income per worker across countries, and that this inequality across nations increased between 1960 and 2000. Xepapadeas and Yannacopoulos (2016) provide evidence suggesting that geographical (or spatial) heterogeneity of per capita GDP increased between 1980 and 2011 across eleven world regions. Despite however the profound importance of the combined temporal and spatial dimension in the study of economic growth, little attention has been given in incorporating space in models of optimal growth.[1]

[1] Economic geography and economic growth has been discussed in the so-called second generation of new economic geography models, but not in a formal growth context (e.g. Martin and Ottaviano 2001, 2003; Baldwin et al. 2003; Baldwin and Krugman 2004; Fujita et al. 2001; Fujita and Mori 2005; Desmet and Rossi-Hansberg 2009, 2010; Breinlich et al. 2014.

Spatiotemporal models of economic growth appeared mainly in the 2000s.[2] Earlier research in which has provided the main mechanism for capital flows across space can be found in Issard (1979). A central reason for introducing spatial aspects in growth could be the question posed by Quah's (1996, p. 1053) "what one wants to know is what happens to the entire cross section of economies not whether a single economy is tending towards its own individual steady state". Answering such a question implies that the growth process should be defined in terms of the temporal evolution of the spatial distribution of capital or output when there are nontrivial interactions across locations.

This papers explores spatial growth by developing a spatial Ramsey optimal growth model in a two dimensional spatial domain. In this domain capital located in a certain spatial point has the tendency to move to locations where the marginal productivity of capital is higher relative to the marginal productivity in the location of origin. The assumption of the marginal-productivity-driven capital flows, was first introduced by Xepapadeas and Yannacopoulos (2016) and differs from the assumption underlying capital flows which is used in recent spatial growth model literature (e.g. Brito, Boucekkine et al., Fabri ). In this literature capital flows across locations are modeled through a trade balance approach with respect to a closed region. This approach leads to a model of classic local diffusion with a constant diffusion coefficient. Linear local diffusion implies that, with diminishing returns to capital, capital moves from locations of high capital concentration (i.e. rich countries) to locations of low capital concentration (i.e. poor countries). This property, however, seems not to be compatible with empirical findings since – as indicated by Lucas in the context of the Lucas paradox (see, Lucas 1990; 2003, Prasad et al. 2007) capital does not seem to flow from rich countries to poor countries.

The marginal-productivity-driven (MPD) assumption about capital flows adopted in this paper seems to be intuitive since even with diminishing returns, the marginal productivity of capital in a rich county could be high relative to a poor country because of total factor productivity effects or other positive externalities. Thus this approach can be regarded as imune to the Lucas paradox. The assumption of MPD capital flows leads, however, to a model of nonlinear local diffusion which introduces new challenges in the solution of the optimization problem which are required in order to study optimal growth in a spatiotemporal domain.

Thus, the present paper contributes to growth theory by extending the standard optimal growth Ramsey model with a traditional neoclassical production function exhibiting diminishing returns to capital, to a two dimensional spatial domain in which capital flows towards locations of high marginal productivity. In this context our contribution consists in trying to answer questions which emerge when the spatial dimension is combined with this type of capital flows across locations. More specifically we are seeking to answer two questions. First, suppose that the economies located within a bounded spatial domain with symmetric production functions converge in the long run to a "flat earth," – using Krugman's (1998), (see also Fujita et al, 2001) terminology – steady state, in which per capita output and capital is the same across all locations . Is it possible, when productivity-driven capital flows across locations take place, for a small perturbation of the flat-earth capital-labor ratio across locations to induce spatial heterogeneity to capital and output per capita which persists and eventually drives the economies to a "non flat earth " steady state, or the perturbation will die out with the

[2] See for example, Brito (2004), Camacho 2004, Bocekkine et al. (2009, 2013a , 2013b, 2016, 2018), Brock et al. (2014a, 2014b), Fabri (2016), Xepapadeas and Yannacopoulos (2016).

passage of time in which case capital flows will be a driver which homogenizes the economies across locations?

Second, suppose that in a flat earth economy there is a perturbation in total factor productivity (TFP) which, without productivity-driven capital flows, will eventually drive the economies to a spatially heterogenous steady state with respect to capital and output per capita. If along with the TFP perturbation capital starts flowing to locations with higher marginal productivity, will the economies be driven to a more or less spatially heterogenous steady state relative to the case of no productivity-driven capital flows?

The first questions explores whether optimal growth with capital seeking locations of high marginal productivity promotes or not spatial inequalities, while the second explores whether in world with TFP differences across locations optimal growth with capital seeking locations of high marginal productivity intensifies or not spatial inequalities.

In the process of answering these questions a second contribution of this paper is the extension of calculus of variations methods and Pontryagin's maximum principle to the solution of dynamic optimization problems, where capital accumulation is described by a partial differential equation with non linear diffusion.

Our main results, in relation to the questions posed above, is that when production functions are symmetric across locations optimal growth with MPD capital flows could lead to a growth process, in which output and capital per worker are different across locations if, the utility discount rate, the share of capital in a Cobb-Douglas production function and the elasticity of marginal utility are sufficiently high. For the conventional low discount rate, low capital share and low elasticity of marginal utility, MPD capital flows will act as homogenizing factor and reduce inequalities after a spatial perturbation of the capital labor ratio. On the other hand, under spatial TFP differences optimal growth under MPD capital flows could intensify or diminish spatial inequalities depending on the specific characteristics of capital flows.

The rest of the paper is organized as follows. Section 2 models capital flows under the assumption that the flows are driven by marginal productivity differentials across locations and states the spatial Ramsey model. Section 3 extends the Pontryagin's principle under nonlinear diffusion. Section 4 discusses the formation of spatial patterns in the Ramsey model and the relation between spatial heterogeneity and capital mobility, while section 5 concludes. All proofs are relegated to the Appendix.

## 2 Marginal Productivity Driven Spatial Capital Flows

Xepapadeas and Yannacopoulos (2016) extended the Solow model to include MPD spatial transport of capital. We briefly recall the derivation of the fundamental capital accumulation equation, before we present the related optimal control problem. Consider a spatial domain $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2$[3] and let

[3] Geographical distance is the most common metric of the distance between two spatial points . Conley and Ligon (2002) suggest, however, that a more appropriate metric for measuring distances associated with economic activities is that of the economic distance – the economic

metric – reflected in transportation costs. Thus in this paper space should be interpreted in terms of an economic metric. Since there is a one-to-one correspondence between the elements of the economic and the geographical space, any spatial distribution defined

$k(t,x)$ be capital (or the density of capital) at time $t$ and at the spatial location $x \in \mathcal{D}$.[4] By density we mean that the total quantity of capital in a subset $U \subset \mathcal{D}$ at time $t$ is $K(t) = \int_U k(t,x)dx$, where by $dx$ we denote the Lebesgue measure on $\mathcal{D}$.

Locally accumulated capital produces output at time $t$ and at the spatial location $x \in \mathcal{D}$ according to a standard neoclassical production function satisfying Inada conditions, $f(x, k(t,x))$, which exhibits spatial variability. Output is allocated to net capital formation, consumption $c(t,x)$, and locally capital depreciation at rate $\delta$, so that the density of depreciated capital at $(t,x)$ is $-\delta k(t,x)$. However, in contrast to nonspatial growth models capital can be transported in space, i.e. it may arrive to $(t,x)$ from other locations $(t,x')$ – where we assume for simplicity that capital transport is instantaneous in time – or it may depart from $(t,x)$ for other locations which are more advantageous in terms of marginal productivity. We will adopt a local in space model, and define the capital flux vector $J$ which is a vector field providing information on the direction and intensity of the capital motion. This vector field points to the direction that net capital transport takes place and its magnitude is related to the total quantity of transported capital.

To make the transport mechanism clear, consider for the moment no production, then the change of the total capital in any region $U \subset \mathcal{D}$ is given in terms of the surface integral $\int_{\partial U} J \cdot n dS$, where $\partial U$ is the boundary of $U$, $n$ is the outward unit vector at any point on $\partial U$, $J$ is the flux vector field and $dS$ is the surface volume element. This integral simply "adds" the quantity of capital that has left or entered $U$ through its boundary; at point $x \in \partial U$ the quantity of capital that enters or leaves – depending on the direction of the vector $J(x)$ – will be $J(x) \cdot n(x)dS(x)$ and the total quantity is the sum of all these quantities, which in the continuous limit is the surface integral of the scalar field $J \cdot n$.

The role of the flux vector in describing capital transport is clarified in terms of the Gauss divergence theorem according to which for any $U \subset \mathcal{D}$, with sufficiently smooth boundary, it holds that $\int_{\partial U} J \cdot n dS = \int_U \nabla \cdot J dx$, where the left hand side is a surface integral on $\partial U$ whereas the right hand side is a volume integral on $U$. The quantity $\nabla \cdot J$ is a scalar field which motivates the introduction of an operator which takes a vector field $(J)$ and maps it to a scalar field $(\nabla \cdot J)$. This is the divergence operator $\nabla \cdot$.[5] In the Cartesian system the divergence can be expressed as $\nabla \cdot = (\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2})$ the notation "$\cdot$" which means that considering any vector field $J = (J_1, J_2)$ we have that $\nabla \cdot J = \frac{\partial}{\partial x_1} J_1 + \frac{\partial}{\partial x_2} J_2$. [6]

A standard book keeping argument, complemented with the above discussion leads to the result

in economic space can be transformed into a corresponding distribution in the geographical space. Inequalities in the economic space can be immediately translated to inequalities in geographical space. The use of the economic space concept allows the meaningful use of local diffusion models.

[4]To simplify we assume that labour at each location is fixed and immobile across locations. So $k(t,x)$ can be interpreted as capital per worker or the capital-labor ratio at each location. Capital mobility combined with labor growth is undoubtedly an area for further research.

[5]The Gauss divergence theorem is coordinate free i.e., it can work for any coordinate system used to describe the domain $U$, such as Cartesian, polar etc., and for any geometry.

[6]We will keep the notation $\nabla \cdot$ in this paper, as it is allows us to present our results in a coordinate free fashion, allowing for a range of possible geometries, and not only the Cartesian case. However, any one unfamiliar with the $\nabla \cdot$ notation can safely think of $\nabla \cdot F$ for any vector field $F$, in the Cartesian representation.

that the net accumulation of capital at $(t,x)$ will be given by

$$\frac{\partial}{\partial t}k(t,x) = -\nabla\cdot\, J(t,x) + f(x,k(t,x) - \delta k(t,x) - c(t,x), \tag{1}$$

where the 3 last terms on the right hand side correspond to local output production, local capital depreciation and local consumption respectively, while the term $-\nabla\cdot\, J(t,x)$ provides information on the net capital transport from or to other locations than $x$.

In order to turn (1) into a useful tool that will allow us to monitor the spatiotemporal evolution of capital density we need to specify the vector field $J$. In Xepapadeas and Yannacopoulos (2016) it was assumed that capital tends to be relocated to regions of relatively higher marginal productivity of capital, which is defined as $m\,(t,x) = \frac{\partial}{\partial k} f(x,k\,(t,x))$. Clearly $m$ depends on $k$ and since $k$ is varying in space it also depends on the location of space that we consider. The marginal productivity of capital $m$ is thus a scalar field, and a quantity that reflects its spatial variability is its gradient $\nabla\, m$. The gradient is a vector field that locally points to the direction of greatest increase of $m$. As with the divergence this vector field can be defined in a coordinate free fashion. In Cartesian coordinates $(x_1, x_2)$ the gradient vector has a representation as $\nabla\, m = (\frac{\partial m}{\partial x_1}e_1, \frac{\partial m}{\partial x_2}e_2)$, $\;\nabla\cdot\, m = \frac{\partial m}{\partial x_1} + \frac{\partial m}{\partial x_2}$, where $e_1, e_2$ are the unit vectors in the direction $x_1$, $x_2$ respectively.[7]

Our basic modeling assumption is that the flux vector field $J$ is proportional to the gradient of the marginal productivity of capital $\nabla\, m$, with a proportionality factor which may depend on local conditions at point $x$ as well as the capital accumulation at this point. This assumption expresses the intuition that capital located at $x$ will relocate towards locations of higher marginal productivity of capital, but also that capital at different locations may have different propensity to relocate towards the higher marginal productivity locations, possibly on account of local regulations, taxes or tariffs). This propensity is likely to depend on the local concentration of capital, i.e. large capital concentrations may have different propensity to relocate towards higher returns than smaller capital concentrations. These considerations are expressed by defining the flux as

$$J(t,x) = \bar{D}_0\,(B(x)\psi(k(t,x)))\,k\nabla\, m(t,x) \text{ for any } (t,x)$$

where $\bar{D}_0$ is a constant capital transport parameter, the spatial term $B(x)$ could reflect the effect of regulations or geographical or commercial factors affecting the propensity of capital to relocate, and the term $\psi(k)$ models the fact that different capital densities could exhibit different propensity towards relocation. We take $J$ explicitly proportional to $k$ to stress the fact that $B\psi$ represents propensity, i.e. may be interpreted as probability to relocate. Assuming a production function of the

[7] Its representation differs in other coordinate systems, in which the unit vectors may also change direction from point to point, but the definition of the gradient as providing the direction where a scalar field presents the fastest increase remains. As with the divergence operator we will use the general notation $\nabla$ to allow for general geometries, but any one unfamiliar with this notation may use the definition of the gradient in Cartesian coordinates.

form $A(x)f(k(t,x))$, where $A(x)$ can be interpreted as local total factor productivity (TFP)[8], then

$$\nabla m(t,x) = A(x) f''(k(t,x)) \nabla k(t,x) + f'(k(t,x)) \nabla A(x)$$

and the flux $J$ can be expressed as

$$J(t,x) = \bar{D}_0 A(x)B(x)\psi(k(t,x))f''(k(t,x))k(t,x)\nabla k(t,x) + \bar{D}_0 B(x)\psi(k(t,x)f'(k(t,x))k(t,x)\nabla A(x).$$

From now onwards, to make notation easier we may omit the explicit dependence on $(t,x)$, keeping in mind that $A$ and $B$ could be functions of $x$, whereas $f$ is a function of $k$, while the composite function $f(k(t,x))$ will depend on $(t,x)$ through the dependence of $k$ on $(t,x)$. Finally by $'$ we denote the derivative of a function with respect to its argument, hence $f'(k(t,x))$ denote the derivative of the function $f:\mathbb{R}\to\mathbb{R}$ with respect to its argument, composed with the capital density.

Putting all the above together in (1)we end up with the following partial differential equation for the spatiotemporal evolution of the capital density

$$\frac{\partial}{\partial t}k = \nabla\cdot\Big( -\bar{D}_0 AB\psi(k)f''(k)k\nabla k - \bar{D}_0 B\psi(k)f'(k)k\nabla A\Big) + Af(k) - \delta k - c,$$

where upon defining

$$w_1(k) = -\psi(k)f''(k)k \text{ and } w_2 = \psi(k)f'(k)k, \tag{2}$$

which are both positive functions because of the properties of standard neoclassical production functions, $D_0(x) = \bar{D}_0 A(x)B(x)$. Noting that $\nabla \ln A = \frac{1}{A}\nabla A$, capital accumulation is defined as:

$$\frac{\partial}{\partial t}k = \nabla\cdot\Big(D_0 w_1(k)\nabla k - D_0 w_2(k)\nabla \ln A\Big) + Af(k) - \delta k - c. \tag{3}$$

This is a nonlinear diffusion equation in divergence form where capital transport is driven by two factors, the spatial variation of $k$, which induces the variation of $m$ and the spatial variation of the local TFP $A$.

The solution of the PDE (3) will provide the spatiotemporal evolution of capital density when the consumption density $c$ is given. This task requires an initial condition $k_0$ which corresponds to the capital density $k(0,x)$ and the determination of the behavior for the capital transport process on the boundary of the domain $\mathcal{D}$. A natural condition is to assume that capital may not be transported outside $\mathcal{D}$, a fact that implies that the capital flux $J$ vanishes perpendicular to the boundary i.e. that $J(t,x)\cdot n(x) = 0$ for $x\in\partial\mathcal{D}$ and for every $t\in[0,T]$ where $T$ is an appropriate time horizon (possibly infinite). Assuming that $\nabla \ln A = 0$ on $\partial\mathcal{D}$ this natural boundary condition is equivalent to a Neumann boundary condition of the form $\nabla k\cdot n = 0$ on $\partial\mathcal{D}$. [9]

[8] Note that while TFP varies locally it does not grow with time. The main reason for this simplification was to isolate and make clearer the impact of capital flows towards relatively higher $m(t,x)$ on the spatiotemporal evolution of capital density. Allowing for TFP growth i.e., $A(t,x)$ is relocated to further research.

[9] We will keep this equation in divergence form for two reasons (a) it is the natural form of this equation leading to conservation of the total capital if the terms $Af(k)-\delta k - c$ are omitted and (b) as written the equation is not dependent on the choice of geometry and is therefore valid for a general domain $\mathcal{D}$. A more familiar special form for this equation would be if $D := D_0 w_1(k)$ were a constant, and $A$ was spatially independent in which case the above equation reduces

It should be noted that in our model the diffusion coefficient $D = D_0 w_1(k)$ depends in a nonlinear fashion on the capital concentration. This dependency is induced by our assumptions concerning the nature of capital transport. This represents a difference of this model (3 realtive to the model for capital transport employed by Boucekkine et al. (2013b), in which the diffusion coefficient is assumed to be a constant, hereafter referred to as the linear diffusion model. Eventhough these two models come from different modelling assumptions concerning capital transport, our model reduces to the linear diffusion model in the special case where instead of using equation (2) for the definition of $w_1, w_2$ and $D_0$, we set $D_0 = \bar{D}_0$, a constant, $w_1 = 1$ and $w_2 = 0$ in (3).

### 2.1 The spatial Ramsey model

Having defined the capital accumulation equation we use it to study a Ramsey type optimal growth model, by considering a social planner who chooses optimal spatiotemporal consumption paths $c(t,x)$ to maximize the total discounted intertemporal utility over the whole domain $\mathcal{D}$ subject to spatiotemporal dynamics defined by (3).[10] The problem then becomes the optimal control problem

$$\max_c \int_0^\infty \int_{\mathcal{D}} e^{-rt} U(c(t,x)) dx dt, \text{ subject to } (3), \tag{4}$$

where $U$ is a standard utility function for consumption, satisfying Inada conditions and $r$ is the utility discount rate. Problem (4) has been studied in the case of linear transport ($D$ constant) using the Hamilton Jacobi equation framework by Boucekkine and coworkers and the Pontryagin maximum principle in the special case linear transport within the framework of the AK model (e.g. Boucekkine et al., 2013b). In this paper, we will study the optimal spatial consumption and capital allocation in the spatial Ramsey model with nonlinear transport mechanisms for capital and a concave neoclassical production function. While the Hamilton Jacobi equation can be employed, we prefer to use the Pontryagin maximum principle instead, mainly because this approach is very often used in the analysis of the temporal Ramsey model, and thus will provide easier access to our results to economists familiar with the classic temporal Ramsey model.

## 3 Pontryagin's principle under nonlinear diffusion

To solve problem (4) we extend Pontryagin's principle to the case in which the dynamic constraint is represented by a partial differential equation with nonlinear diffusion. In what follows we will assume the existence of a solution and an optimal path for problem (4) which admits sufficient regularity. We focus on the derivation of a necessary condition, in terms of a Pontryagin principle, for the optimal path and optimal control policy, with the specific interest of characterizing the long run (time independent)

to $\frac{\partial}{\partial t}k = D\nabla \cdot \nabla k + Af(k) - \delta k - c$, which using the definition of the Laplacian operator $\Delta k := \nabla \cdot \nabla k$ reduces to $\frac{\partial}{\partial t}k = D\Delta k + Af(k) - \delta k - c$. It is a matter of algebra to confirm that in Cartesian coordinates, using the definitions of $\nabla$, $\nabla \cdot$ given above we have that $\Delta k = \frac{\partial^2 k}{\partial x_1^2} + \frac{\partial^2 k}{\partial x_1^2}$ (or simply $\frac{\partial^2 k}{\partial x_1^2}$ for one dimensional domains) therefore reducing the model to the standard diffusion equation with a linear diffusion term.

[10] **?**, assuming that consumption was chosen as a fixed percentage of production, i.e., $c = (1-s)Af(k)$, used (3) studied a spatial Solow model with nonlinear diffusion.

optimal state, rather than studying in detail the dynamics of the optimal path.

**Proposition 3.1.** *Under the standing assumptions made earlier, a sufficiently regular optimal path* $(k^*, c^*)$ *for problem* (4) *such that* $c^*(t,x) > 0$ *a.e. can be characterized in terms of the optimality condition*

$$U'(c^*(t,x)) = p(t,x), \tag{5}$$

*where* $(k^*, p^*)$ *satisfy the following coupled forward-backward system of nonlinear PDE*

$$\frac{\partial k^*}{\partial t} = \nabla \cdot \Big( D_0 w_1(k^*) \nabla k^* - D_0 w_2(k^*) \nabla \ln A \Big) + Af(k^*) - \delta k^* - g(p), \tag{6}$$

$$\frac{\partial p}{\partial t} = -\nabla \cdot \Big( D_0 w_1(k^*) \nabla p \Big) + \underline{G}(k^*, x) \cdot \nabla p + (r + \delta - Af'(k^*))p, \tag{7}$$

*where* $g$ *is the inverse function of* $U'$ *and*

$$\underline{G}(k^*, x) = D_0(x) w_1'(k^*)) \nabla k^* - D_0(x) w_2'(k^*)) \nabla \ln A(x).$$

*The system (6)-(7) is supplemented with the boundary conditions*

$$\nabla k^*(t,x) \cdot n(x) = \nabla p(t,x) \cdot n(x) = 0, \;\; x \in \partial \mathcal{D},$$

*where for simplicity we have assumed that* $\nabla \ln A \cdot n = 0$ *on* $\partial\mathcal{D}$*, with the initial condition*

$$k^*(0,x) = k_0(x), \;\; x \in \mathcal{D},$$

*and the transversality condition*

$$\lim_{T \to \infty} e^{-rT} \int_{\mathcal{D}} k(T,x) p(T,x) dx = 0.$$

The proof of the proposition is given in Appendix A.

As it is traditional in optimal growth an alternative form for the optimality system (6)-(7), can be derived in terms of the optimal consumption path, by replacing the adjoint variable $p$, using the optimality condition (5). Since $\frac{\partial}{\partial t} p(t,x) = U''(c^*(t,x)) \frac{\partial}{\partial t} c^*(t,x)$, and $\nabla p(t,x) = U''(c^*(t,x)) \nabla c^*(t,x)$, given sufficient regularity of the utility function and the optimal path the following corollary is obtained.

**Corollary 3.2.** *If* $U$ *is twice continuously differentiable,and* $U''(c) \neq 0$*, the optimal path* $(k^*, c^*)$ *satisfies the system of PDE's (the explicit* $(t,x)$ *dependence dropped for ease of notation)*

$$\frac{\partial k^*}{\partial t} = \nabla \cdot \Big( D_0 w_1(k^*) \nabla k^* - D_0 w_2(k^*) \nabla \ln A \Big) + Af(k^*) - \delta k^* - c^*, \tag{8}$$

$$\frac{\partial c^*}{\partial t} = -\frac{1}{U''(c^*)} \nabla \cdot \Big( U''(c^*) D_0 w_1(k^*) \nabla c^* \Big) + \underline{G}(k^*, x) \cdot \nabla c^* + (r + \delta - Af'(k^*)) \frac{U'(c^*)}{U''(c^*)}, \tag{9}$$

*supplemented with Neumann boundary conditions, for both* $k^*, c^*$ *and with the transversality condition* *appropriately modified.*

Note that for traditionally used utility functions of the class $U(c) = \frac{1}{1-\lambda}c^{1-\lambda}$, $\lambda \geq 1$, the PDE for the optimal consumption simplifies to

$$\frac{\partial c^*}{\partial t} = -\frac{1}{U''(c^*)}\nabla \cdot \left(U''(c^*)D_0 w_1(k^*)\nabla c^*\right) + \underline{G}(k^*, x) \cdot \nabla c^* - \frac{1}{\lambda}(r + \delta - Af'(k^*))c^*.$$

Note also that when capital is immobile, that is $D_0 = 0$, and the production function does not explicitly depend on location $x$, the optimal paths satisfies the classic system of ordinary differential equations

$$\begin{aligned} \frac{dk^*}{dt} &= Af(k^*) - \delta k^* - c^* \\ \frac{dc^*}{dt} &= (r + \delta - Af'(k^*))\frac{U'(c^*)}{U''(c^*)} \end{aligned} \tag{10}$$

We conclude our study of the Pontryagin principle with an equivalent Hamiltonian formulation.

**Proposition 3.3.** *Define the Hamiltonian density*

$$\bar{H}(t, k, \bar{p}) = \max_{c \in \mathbb{R}_+} \left\{ \bar{p}(\nabla \cdot (D_0 w_1(k)\nabla k - D_0 w_2(k)\nabla \ln A) + Af(k) - \delta k - c) + e^{-rt}U(c) \right\}.$$

*and consequently the Hamiltonian functional*

$$\bar{\mathcal{H}}(t, k, \bar{p}) := \int_{\mathcal{D}} \bar{H}(t, k(t, x), \bar{p}(t, x))dx,$$

*The optimality condition can be brought into Hamiltonian form as*

$$\begin{aligned} k' &= D_{\bar{p}}\bar{\mathcal{H}}(t, k, \bar{p}), \\ \bar{p}' &= -D_k\bar{\mathcal{H}}(t, k, \bar{p}), \end{aligned}$$

*where $\bar{p}(t, x) = e^{-rt}p(t, x)$, and by $D_{\bar{p}}, D_k$, we denote the Gâteaux derivatives of the functional with respect to the $\bar{p}$ and $k$ respectively. The system is autonomous in terms of $(k, p)$.*

The proof is sketched in Appendix B.

The spatial structure of the steady state $(k^*, c^*)$ can be obtained either by using Proposition 3.1 directly and setting $\frac{\partial k^*}{\partial t} = \frac{\partial p^*}{\partial t} = 0$, along with the optimality condition $p(x) = U'(c(x))$, or by using the Corollary 3.2 and setting $\frac{\partial k^*}{\partial t} = \frac{\partial c^*}{\partial t} = 0$. Therefore,

**Proposition 3.4.** *Under sufficient smoothness on the optimal path and the utility function the optimal steady state $(k^*, c^*)$ solves the system of nonlinear elliptic equations*

$$\begin{aligned} 0 &= \nabla \cdot [D_0 w_1(k^*)\nabla k^* - D_0 w_2(k^*)\nabla \ln A)] + Af(k^*) - \delta k^* - c^*, \\ 0 &= -\nabla \cdot (U''(c^*)D_0 w_1(k^*)\nabla c^*) + U''(c^*)\underline{G}(k^*, x) \cdot \nabla c^* + (r + \delta - Af'(k^*))U'(c^*)), \end{aligned} \tag{11}$$

*supplemented with homogeneous Neumann boundary conditions.*

In Appendix C we show that the above system of elliptic equations may in fact be transformed into a more convenient equivalent form which is more useful for analysis and numerical treatment.

**Remark 3.5** (An alternative (equivalent) form for the Pontryagin formula)**.** For the convenience of the reader we provide an alternative equivalent form for the Pontryagin formula that will be used in Section 4.5. Using vector calculus identities and the definition of the functions $w_1, w_2$ we can express the state-costate equation Consider the state equation (3) expressed in the equivalent form,

$$\begin{aligned} \frac{\partial k^*}{\partial t} &= -\nabla \cdot [\bar{D}_0 B \psi(k^*) k^* \nabla (A f'(k^*)) - \delta k^* - c^*, \\ \frac{\partial p}{\partial t} &= \bar{D}_0 A f''(k^*) \nabla \cdot [B \psi(k^*) k^* \nabla p] - \bar{D}_0 B (\psi(k^*) k^*)' \nabla (A f'(k^*)) \cdot \nabla p + \\ &(\delta + r - A f'(k^*)) p. \end{aligned} \tag{12}$$

The system of equations (12) (supplemented with no flux boundary conditions and the transversality condition of Proposition 3.1) can be used to characterize the optimal path $(k^*, c^*)$, using $p = U'(c^*)$.

# 4 MPD Spatial Capital Flows and Spatial Inequalities

The tools developed in the previous section will help us answer to two main questions posed at the introduction of this paper. The first relates to whether MPD capital flux could destabilize a flat earth optimal steady state (FEOSS) of a Ramsey model and induce spatial patterns which in our case would imply spatial inequalities. It is well known that for the standard Ramsey optimal growth model without spatial interactions – system (10) – a steady state has the global saddle point property meaning that for any initial capital stock there is an initial level of consumption such that the system will converge to the steady state along the stable manifold The system (6)-(8) for $D_0 = 0$ could by analogy be regarded as the analogue of the saddle point in finite dimensional (temporal only) Ramsey model. In such a case the optimal policy which determines the optimal path for the state (capital) and the control (consumption) could be regraded as a "collection" of identical stable manifolds leading the system to the long run FEOSS.

Suppose now that at the state of the flat earth, $D_0 > 0$ so that MPD flux is introduced and a small perturbation of the flat earth capital landscape takes place. Can the optimal control of the perturbed system determine a "collection" of stable manifolds, similar to the $D_0 = 0$ case, along which the system will return of the FEOSS? If yes, the MPD capital mobility is a spatially homogenizing force and preserves spatial equality. If however the system does not returns to the flat earth state then spatial inequalities which are induced by optimal control emerge.

We know from the celebrated Turing (1952) paper that in a reaction-diffusion system, diffusion can be spatially homogenizing, but under certain conditions it could induce spatial heterogeneity and the emergence of spatial patterns and form - morphogenesis. Turing analysis has been used by Krugman (1996), (see also Fujita et al. 2001). to generate patterns from a flat earth-steady-state space with economies located on a circle.[11] Turing's and Krugman's analysis did not involve explicit dynamic

[11] This is the "50 Cadilac diagram" for a race track economy (see Krugman 1998).

optimization. Dynamic optimization and the emergence of patterns was first analyzed by Brock and Xepapadeas (2008, 2010) and subsequently by Brock, Xepapadeas and Yannacopoulos (2014c) in the context of optimal spatial resource management. These models were characterized by linear Fickian diffusion which is an appropriate assumption for biological resources, since they move from high to low concentrations. As we discussed however earlier, this assumption is not appropriate for capital movements in view of the Lucas paradox. So our objective in terms of our first question is to examine under what conditions nonlinear diffusion, induced by capital seeking higher marginal productivity, will act as a force of convergence for the economies in the spatial domain, or as force that generates spatial patterns and inequalities of per capita output and capital across space. In terms of our second question we need to know whether MPD capital mobility intensifies or weakens already existing, because of TFP differences, spatial inequality. These two cases are analyzed in the following sections.

## 4.1 Spatial homogeneity and optimal pattern formation

In the temporal Ramsey model saddle point stability means that the linearization matrix of system (10) at the steady state has a negative determinant and a small perturbation in the $(k, c)$ or the equivalent $(k, p)$ space along the stable manifold will die out and the system will return to the optimal steady state.

The linearized stability of a general steady state displaying arbitrary spatial dependence is not an easy task, and requires sophisticated techniques from the spectral theory of linear operators, and typically involves detailed numerical analysis. However, interesting detailed results can be obtained analytically by a perturbative approach, providing a clear view of the effects of capital mobility on optimal growth, in the case where a flat optimal steady state is perturbed by the effects of capital mobility, with the possible generation of optimal spatial patterns which emerge because the MPD capital mobility destabilizes the stable manifold of the zero MPD mobility system.

Consider the case where the TFP $A$ and capital mobility coefficient $B$ are independent of $x$. In trying to examine the potential emergence of pattern formation induced by MPD capital mobility we examine the stability of a flat-earth steady state of the optimizing system after spatiotemporal perturbations in capital and consumption. The analysis involves linearizing the optimality system (8) around a flat-earth steady state $(k_0^*, c_0^*)$, which is the solution of the algebraic system

$$A(f'(k_0^*)) = r + \delta, \qquad (13)$$
$$c_0^* = Af(k_0^*) - \delta k_0^*,$$

and then consider solutions for capital and consumption in terms of the corresponding linear PDE system. In doing so we can see that the perturbation $z(t, x) := (k(t, x), c(t, x))^T$ around $z^* := (k_0^*, c_0^*)^T$ evolves according to the constant coefficient PDE system

$$\frac{\partial z}{\partial t} = \mathbb{D}(z^*)z + L_F(z^*)z, \qquad (14)$$

where

$$L_F(z^*) = \begin{pmatrix} A(f_0'^*) - \delta & -1 \\ -\frac{U'(c_0^*)}{U''(c_0^*)} A f''(k_0^*) & 0 \end{pmatrix} = \begin{pmatrix} r & -1 \\ -\frac{U'(c_0^*)}{U''(c_0^*)} A f''(k_0^*) & 0 \end{pmatrix},$$

$\mathbb{D}(z^*)$ is the matrix differential operator

$$\mathbb{D}(z^*) = \begin{pmatrix} D(k_0^*)\Delta & 0 \\ 0 & -D(k_0^*)\Delta \end{pmatrix},$$

$\Delta$ is the Laplacian operator with homogeneous Neumann boundary conditions and

$$D(k_0^*) = D_0 w_1(k_0^*), \;\; D_0 = \bar{D}_0 AB,$$

is a constant diffusion coefficient, which depends on the level of the optimal $k_0^*$.

This is unlike the case of linear diffusion, where the perturbation $z$ will follow the linear evolution equation (14) with $\mathbb{D}(z^*)$ replaced by the operator

$$\mathbb{D}_0 = \begin{pmatrix} D_0\Delta & 0 \\ 0 & -D_0\Delta \end{pmatrix}.$$

Note also that in the limit of vanishing diffusion, this system reduces to the corresponding system for the Ramsey model in the absence of capital mobility, with the familiar saddle point property in the relevant phase space which is $\mathbb{R}^2$.

In the presence of capital mobility, spatial variability of the perturbation $(k(t,x), c(t,x))$ is possible, hence the relevant phase space is a function space, such as for example $L^2(\mathcal{D}) \times L^2(\mathcal{D})$, where by $L^2(\mathcal{D})$ we denote the space of square integrable functions on $\mathcal{D}$. Clearly such a dynamical system is not possible to visualize on the phase plane, eventhough one may still define a saddle point structure in this case as well, in terms of the spectral behavior of the linear operator $\mathbb{D}(z^*)$.

Since (14) is a constant coefficient problem in a bounded domain, we may characterize its solutions completely in terms of the eigenvalues of the operator $-\Delta$, on $\mathcal{D}$, with homogeneous Neumann boundary conditions. i.e., in terms of the set of functions $\{\phi_n : n \in \mathbb{N}\}$, which are solutions of the problem

$$\begin{aligned} -\Delta\phi_n &= \mu_n\phi_n, \\ \nabla\phi_n \cdot n &= 0. \end{aligned} \tag{15}$$

This discrete set is a complete orthonormal basis in $L^2(\mathcal{D})$, hence any function $u \in L^2(\mathcal{D})$ admits a Fourier expansion $u(x) = \sum_{n\in\mathbb{N}} u_n\phi_n(x)$, with almost everywhere convergence, and $u_n = \langle u, \phi_n\rangle$ with $\langle\cdot,\cdot\rangle$ denoting the inner product in $L^2(\mathcal{D})$ [12]. The spectrum has the property $\mu_n \geq 0$, with $\mu_n \to \infty$ and importantly for many cases of interest both $\mu_n$ and the eigenfunctions $\phi_n$ are known exactly in analytic form (see Appendix E). Since any solution of (14) can be expanded as $z(t,x) = \sum_{n\in\mathbb{N}} z_n(t)\phi_n(x)$, with $z_n(t) = (k_n(t), c_n(t))^T$ where $k_n(t) = \langle k(t,\cdot), \phi_n(\cdot)\rangle$, $c_n(t) = \langle c(t,\cdot), \phi_n(\cdot)\rangle$, by substituting this

[12] Typically, $u_n = \int_{\mathcal{D}} u(x)\phi_n(x)dx$, with the possible introduction of weights depending on the geometry of the domain.

expansion in (14) and using the orthogonality properties of the basis of eigenfunctions we see that the system of PDEs (14) reduces to an equivalent countable system of ODEs in the form

$$z_n' = L_n z_n, \quad n \in \mathbb{N},$$
$$z_n(0) = z_{n,0}.$$

where $z_n = (k_n, c_n)^T$, $z_{n,0} = (k_{n,0}, c_{n,0})^T$, with $k_{n,0} = \langle k_0, \phi_n \rangle$, $c_{n,0} = \langle c_0, \phi_n \rangle$, and the matrix

$$L_n = \begin{pmatrix} -D(k_0^*)\mu_n & 0 \\ 0 & D(k_0^*)\mu_n \end{pmatrix} + L_F. \tag{16}$$

The general solution of (14) can be expressed as the Fourier series

$$z(t,x) = \sum_{n \in \mathbb{N}} z_n(t)\phi_n(x) = \sum_{n \in \mathbb{N}} e^{tL_n} z_{n,0}\phi_n(x)$$

with the matrix exponential defined as $e^{tL_n} = I + \sum_{k=1}^{\infty} \frac{1}{k!} t^k L_n^k$. A detailed analysis of this solution is presented in Appendix E. This analysis reveals that for certain $n \in \mathbb{N}$, the vector $e^{tL_n} z_{n,0}\phi_n$ grows in time inducing a contribution to the full solution which has a spatial variability – as prescribed by the corresponding eigenfunctions $\phi_n$ – whereas for the remaining $n$ the vector $e^{tL_n} z_{n,0}\phi_n$ decays in time inducing a transient contribution which dies out in time converging asymptotically to the flat-earth steady state. The growing in time contribution can be considered as the emergence of a spatial pattern, compatible with the optimal growth structure, and a destabilization of the flat-earth steady state leading to possible spatial variability for the optimal capital concentration and consumption.

**Proposition 4.1.** *Let $(k_0^*, c_0^*)$ be the flat-earth steady state, attained for the case where $A$ is constant and capital is immobile, $\bar{D}_0 = 0$, and define $M = A\frac{U'(c_0^*)}{U''(c_0^*)} f''(k_0^*) > 0$. Then, the effect of MPD capital mobility $\bar{D}_0 \neq 0$ on the evolution of spatially dependent perturbations of $(k_0^*, c_0^*)$ and the potential emergence of spatial patterns from the flat-earth steady state can be summarized as follows:*

*(i) If $\frac{r^2}{4} < M$, no spatial pattern can develop.*

*(ii) If $\frac{r^2}{4} > M$, MPD capital mobility leads to the emergence of spatial patterns of the form*

$$z_p(t,x) = \sum_{n \in \mathcal{N}} e^{tL_n} z_{n,0}\phi_n(x),$$

*consisting of linear combinations of eigenmodes of the Laplacian $\{\phi_n : n \in \mathcal{N}\}$ where $\mathcal{N}$ is the set of $n \in \mathbb{N}$ such that*

$$D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M > 0, \quad \textit{where } D(k_0^*) = \bar{D}_0 ABw_1(\hat{k}_0^*).$$

*In particular, mode $n$ will become unstable if $D(k_0^*)$ satisfies*

$$\frac{1}{\mu_n}\left(\frac{r}{2} - \sqrt{\frac{r^2}{4} - M}\right) \leq D(k_0^*) \leq \frac{1}{\mu_n}\left(\frac{r}{2} + \sqrt{\frac{r^2}{4} - M}\right). \tag{17}$$

For the Proof see Appendix E.

It is interesting to note that MPD capital mobility will not destabilize the flat steady state if the utility discount rate is sufficiently low and in particular if $r^2 < 4M$. Alternatively, the same condition can be interpreted as a condition for the utility function. Assuming that$U(c) = c^{1-\lambda}/(1-\lambda)$ and the production function is of the Cobb-Douglas type of the form $f(k) = Ak^a$ , we easily obtain that

$$M = \frac{\delta + r}{\alpha} \frac{1-\alpha}{\lambda} \left[ (1-\alpha)\,\delta + r \right]$$

hence, for the destabilization condition $r^2/4 > M$ to hold we need

$$\lambda > \frac{4}{\alpha r^2} (1-\alpha)(\delta + r) \left[ (1-\alpha)\,\delta + r \right] \tag{18}$$

Figure 1 depicts the surface $\lambda(a, r)$ corresponding to (18). MPD capital mobility will destabilize the FEOSS for $(\alpha, r, \lambda)$ points above the $\lambda(\alpha, r)$ surface.

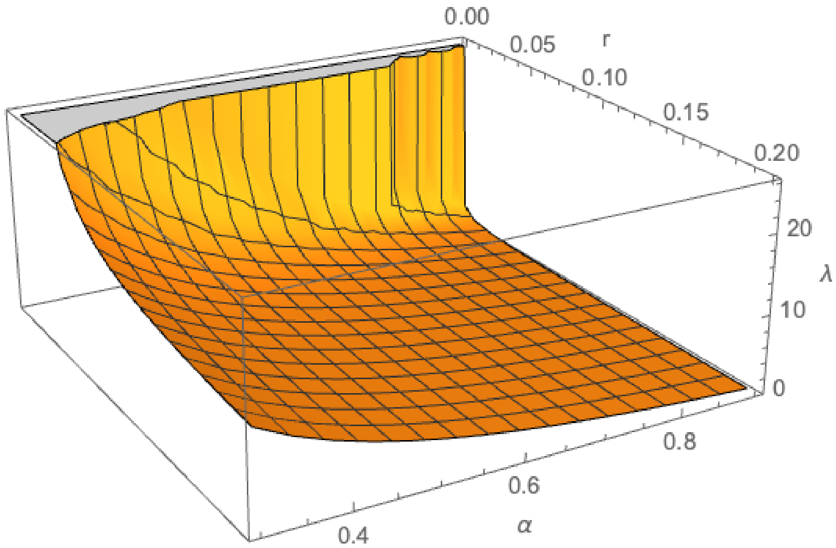



Figure 1: The $\lambda(\alpha, r)$ surface corresponding to (18) ($\delta = 0.03$)

The destabilization of the flat earth through MPD capital mobility is predominantly a high discount rate effect.[13]. For typical low utility discount rates optimal growth is not expected to generate spatial heterogeneities in output and capital per worker, and the optimal policy from the social planners point of view is to steer the spatial economy to a FEOSS when MPD driven capital mobility take place, provided of course that TFP is the same across locations. Destabilization requires a combination of high $(r, \lambda, \alpha)$

If destabilization occurs, the result of Proposition 4.1, may be seen as a perturbation result of the stable manifold of the optimality system for immobile capital, under the effect of capital mobility, so that the stable manifold acquires some spatial structure. This spatial structure is in principle supported by the optimal control procedure as long as the growth rate is compatible with the transversality

[13]The effect of the geometry of the spatial domain is introduced through the dependence of the above formulae on $\mu_n$.

condition, a fact which is guaranteed by the condition (17). Note that given the coefficients $A$, $B$, not any pattern is allowed to develop. The allowable patterns are those corresponding to linear combinations of eigenfunctions $\phi_n$, corresponding to these $n \in \mathbb{N}$ for which condition (17) holds. Since $\mu_n \to \infty$ as $n \to \infty$, one may easily see that for given $D(k^*)$ there is a critical $n_0$, so that only modes with $n < n_0$ may turn unstable. In this sense, the capital mobility mechanism selects the patterns that will appear. Given $D(k^*)$ one may predict exactly the modes that will turn unstable using condition (17) and hence the shape of the pattern. This can also be seen more explicitly in terms of the equivalent formulation of this criterion in terms of (18). In this respect, the effect of nonlinear diffusion may be to facilitate the emergence of certain patterns as long as the effect of nonlinearity is to enhance the capital mobility. To see this more clearly one simply has to observe that in the presence of linear diffusion $D_0$, condition (17) becomes

$$\frac{1}{\mu_n}\left(\frac{r}{2}-\sqrt{\frac{r^2}{4}-M}\right) \leq D_0 \leq \frac{1}{\mu_n}\left(\frac{r}{2}+\sqrt{\frac{r^2}{4}-M}\right)$$

so that if the effect of nonlinearity is to moderately enhance $D_0$ the pattern corresponding to $\phi_n$ is easier to develop. On the other hand either a dramatic decrease or dramatic increase of $D_0$ as an effect of nonlinearity will have an opposite effect, obstructing the appearance of the specific pattern.

The spectrum of the Laplace operator with Neumann boundary conditions as well as the eigenfunctions are known analytically for a number of interesting geometries, a fact that makes checking condition (17) or (18) feasible. The explicit knowledge of the eigenfunctions also allows for the determination of the spatial variability of the generated pattern.

The link between the emergence of spatial patterns by destabilization of the FEOSS through spatial capital flows and the utility discount rate can become clear using the destabilization condition (16). From this condition destabilization occurs at some mode $n$ if

$$\det L_n = \begin{vmatrix} r - D(k_0^*)\mu_n & -1 \\ -\frac{U'(c_0^*)}{U''(c_0^*)}Af''(k_0^*) & D(k_0^*)\mu_n \end{vmatrix} > 0, \tag{19}$$

where $\mu_n \geq 0$, with $\mu_n \to \infty$ as $n \to \infty$, while for $\mu_0 = 0$ and (19) is reduced to the Jacobian determinant of the standard Ramsey model with no spatial flows. Write $M = Af''(k_0^*)\frac{U'(c_0^*)}{U''(c_0^*)} > 0$ and $D(k_0^*)\mu_n = \alpha_n \geq 0$, with $\alpha_0 = 0$. Then

$$\det L_n = \psi_n(r) = \alpha_n(r - \alpha_n) - M \tag{20}$$

which is linear increasing in $r$ and shifts for different modes with $\psi_0 = -M$. Consider Figure 2 and the line $\psi_0 = -M$, and the three lines with intercepts $-a_i^2 - M$, $i = 1, 2, 3$. For the zero mode, which corresponds to the case in which capital is not mobile, $\det L_0 = \psi_0(r) = -M < 0$ and the FEOSS is stable in the saddle point sense. For $r < r_1$ the FEOSS is stable, since for all modes $\psi_n(r) < 0$, and no patterns emerge. For $r_1 < r < r_2$ only the first mode is destabilized, while for $r_2 < r < r_3$ both modes 1 and 2 are destabilized. In the last two cases spatial patterns emerge. A $\psi_n(r)$ line, like the lines shown in figure 2 defines a critical $r = r_n^+$ such that $\psi_n(r_n^+) = 0$. This critical $r_n^+$ can be interpreted as a mode-$n$ internal rate of return for which $\alpha_n\left(r_n^+ - \alpha_n\right) = M$, for $r < r_n^+$ the FEOSS is stable to spatial perturbations induced by capital flows, while for $r > r_n^+$ the FEOSS is destabilized at mode $n$ and spatial patterns emerge.

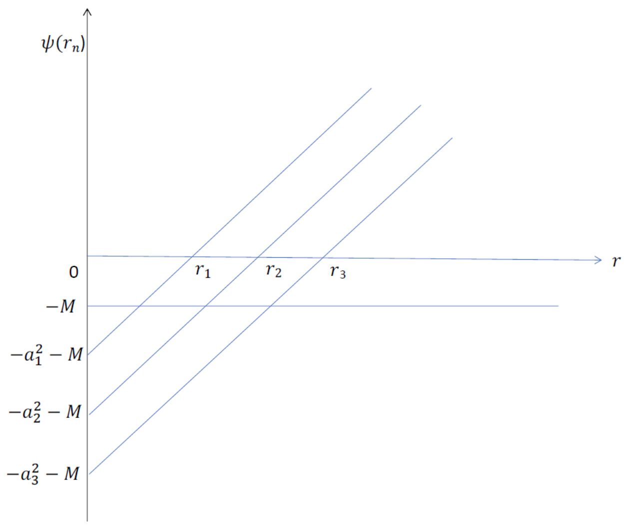



Figure 2: Critical $r$ and destabilization of a FEOSS.

To explore the economic intuition behind this result, note that each mode $n$ corresponds to a distinct optimal growth model. Following Magill (1977, p. 192), the term $M$ is a measure of benefits generated by the FEOSS, $k_0^*$; $M\left|k_n - k_0^*\right|$ is a measure of value loss induced by a deviation from $k_0^*$ for any mode $n$; and $-1$ reflects the cost of controlling the system using $c$ as the control. When $r = 0$ then $\det L_n = \psi_n(0) = -\alpha_n^2 - M < 0$ for all modes $n$. In this case controlling the system after the spatial perturbation to the FEOSS enhances that benefits of the FEOSS and thus no patterns emerge. However if $r$ is sufficiently high, so that $\alpha_n\left(r - \alpha_n\right) > M$, then the net benefits of controlling the system to the FEOSS at this mode become negative. From the social planner's point of view this can be interpreted as suggesting that it is preferable to let patterns emerge instead of controlling the system to the FEOSS.

**Example 4.2.** Consider the case where $\mathcal{D} = [0, L] \subset \mathbb{R}$. Then we have that $\mu_n = (\frac{n\pi}{L})^2$ and $\phi_n(x) = \cos\left(\frac{n\pi x}{L}\right)$, $n = 1, 2, \cdots$. In this case the spatial pattern generation condition selects these spatial patterns which correspond to linear combinations of $\phi_n(x) = \cos\left(\frac{n\pi x}{L}\right)$ with $n$ such that

$$D(k_0^*)\left(\frac{n\pi}{L}\right)^2\left(r - D(k_0^*)(\frac{n\pi}{L})^2\right) - M > 0.$$

It can be seen that only patterns corresponding to relatively small values of $n$ are expected to emerge. No patterns will emerge if

$$D(k_0^*)\left(\frac{\pi}{L}\right)^2\left(r - D(k_0^*)(\frac{\pi}{L})^2\right) - M < 0.$$

**Example 4.3.** Consider the case where $\mathcal{D} = [0, L_1] \times [0, L_2] \subset \mathbb{R}^2$. Then, setting $x = (x_1, x_2)$ and assuming an enumeration of $\mathbb{N} \times \mathbb{N}$ in terms of a multiindex $n = (n_1, n_2)$ we have for the spectrum

$$\mu_{n_1,n_2} = \left(\frac{n_1 \pi}{L_1}\right)^2 + \left(\frac{n_2 \pi}{L_2}\right)^2, \ n_1 = 1, 2, \cdots, \ n_2 = 1, 2, \cdots,$$

with corresponding eigenfunctions

$$\phi_{n_1,n_2}(x_1, x_2) = \cos\left(\frac{n_1 \pi}{L_1} x_1\right) \cos\left(\frac{n_2 \pi}{L_2} x_2\right).$$

To bring this example in the framework of Proposition 4.1 we may simply enumerate the pairs $(n_1, n_2)$ in terms of a single index $n \in \mathbb{N}$, chosen so that the eigenvalues $\mu_{n_1,n_2}$ are ordered in ascending order. For example $n = 1$ would correspond to the pair $(n_1, n_2) = (1, 1)$, $n = 2$ would correspond to the pair $(n_1, n_2) = (2, 1)$ etc. Defining $\nu = \frac{L_1}{L_2}$, the aspect ratio of the rectangle, the condition for the generation of patterns now can be expressed as follows: Consider the set of pairs of natural numbers

$$\mathcal{N} := \left\{ (n_1, n_2) \in \mathbb{N} \times \mathbb{N} \ : \ D(k_0^*) \left(\frac{\pi}{L_1}\right)^2 (n_1^2 + \nu n_2^2) \left( r - \left(\frac{\pi}{L_1}\right)^2 D(k_0^*)(n_1^2 + \nu n_2^2)^2 \right) - M > 0 \right\}.$$

Then emerging spatial patterns will be given by the double sum

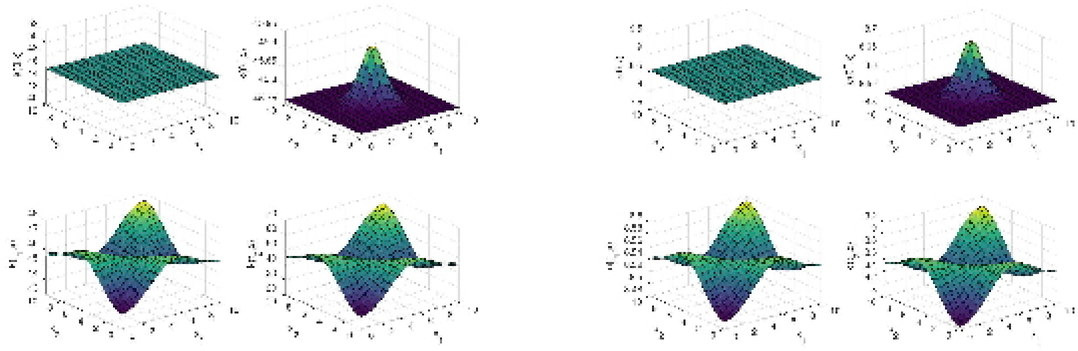
$$k(t, x_1, x_2) = \sum_{(n_1,n_2) \in \mathcal{N}} C_{1,n_1,n_2} \cos\left(\frac{n_1 \pi}{L_1} x_1\right) \cos\left(\frac{n_2 \pi}{L_2} x_2\right),$$
$$c(t, x_1, x_2) = \sum_{(n_1,n_2) \in \mathcal{N}} C_{2,n_1,n_2} \cos\left(\frac{n_1 \pi}{L_1} x_1\right) \cos\left(\frac{n_2 \pi}{L_2} x_2\right).$$

It can be seen again that only modes corresponding to relatively low values of $(n_1, n_2)$ can develop instabilities and no spatial structure will emerge if

$$D(k^*) \mu_0 \big(r - D(k^*) \mu_0\big) - M < 0,$$

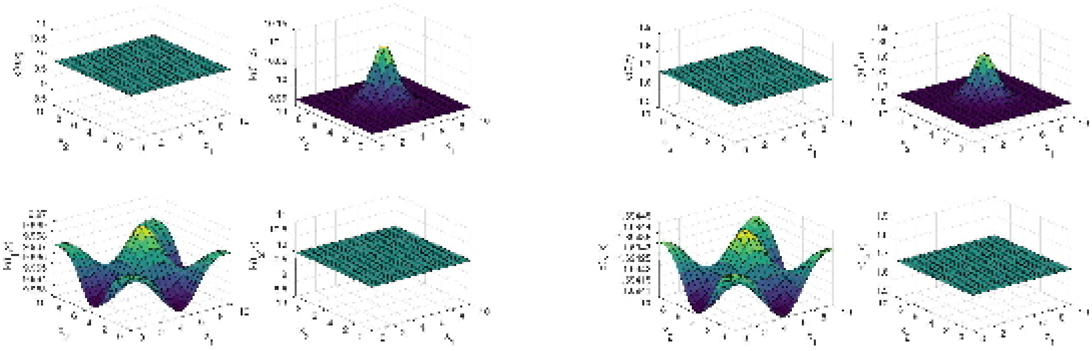
where $\mu_0 = \left(\frac{\pi}{L_1}\right)^2 + \left(\frac{\pi}{L_2}\right)^2$.

In the context of the two-dimensional spatial domain Figures 3 and 4 depict the cases of and emerging spatial pattern for high $r$,$\lambda$ and $\alpha$ after a spatial perturbation (Fig3) and a return to the FEOSS (decaying pattern Fig. 4) after the same spatial perturbation as in Figure 3. Each of the four graphs in each figure correspond to a different point in time starting at $t = 0$.



(a) (b)

Figure 3: Emergence of spatial patterns for capital (a) and consumption (b). $r = 0.1$, $\lambda = 6$, $\alpha = 0.6, \delta = 0.03$.



(a) (b)

Figure 4: Decaying spatial patterns for capital (a) and consumption (b). $r = 0.03$, $\lambda = 2$, $\alpha = 0.3, \delta = 0.03$.

**Example 4.4** (Pattern formation on the sphere)**.** A possible geometric model for the globe would be that of the surface of a sphere of radius $R$. Then, a convenient set of coordinates would be spherical coordinates $(\rho, \phi, \theta)$, where $r$ corresponds to the radial coordinate (assume $\rho = R$ constant), $\phi$ is the azimuthal angle that corresponds to the longitude and $\theta$ is the polar angle that corresponds to latitude. In particular, $\theta$ is the co-latitude ranging from 0 at the North pole to $\pi$ at the South pole, and $\phi$ is the longitude ranging from 0 to $2\pi$. The connection with Cartesian coordinates is in terms of the relations

$$x_1 = \rho \sin\theta \, \cos\phi, \;\; x_2 = \rho \sin\theta \, \sin\phi, \;\; x_3 = \rho \cos\theta.$$

The Laplace operator on the surface of the sphere (Laplace-Beltrami operator) becomes

$$\Delta u = \frac{1}{R^2 \sin\theta} \frac{\partial}{\partial\theta}\left( \sin\theta \frac{\partial u}{\partial\theta}\right) + \frac{1}{R^2 \sin^2\theta} \frac{\partial^2 u}{\partial\phi^2}.$$

The eigenvalue problem (15) for the Laplacian on the sphere (with periodic boundary conditions which are the natural boundary conditions for this geometry) has eigenvalues $\mu = \frac{\ell(\ell+1)}{R^2}$, $\ell = 0, 1, \cdots$ with corresponding eigenfunctions provided in terms of the spherical harmonic functions $\phi(\theta, \phi) = Y_\ell^m(\theta, \phi) = e^{im\phi} P_\ell^m(\cos\theta)$, $m = -\ell, \cdots, \ell$ where $P_\ell^m$ is an associated Legendre polynomial. For each $\ell$ the eigenfunction corresponding to $\mu_\ell = \frac{\ell(\ell+1)}{R^2}$ is a linear combination of the $Y_\ell^m$ for $m = -\ell, \cdots, \ell$. This leads to a complete set of real eigenfunctions as

$$\phi_\ell^{(1)} := cos(m\phi) P_\ell^m(\cos\theta), \;\; \phi_\ell^{(2)} := \sin(m\phi) P_\ell^m(\cos\theta), \;\; \ell = 0, 1, \cdots, \;\; m = 0, \cdots, \ell,$$

with corresponding eigenvalues $\mu_\ell = \frac{1}{R^2}\ell(\ell+1)$, which are orthogonal in terms of the inner product $\langle\cdot,\cdot\rangle$ defined in terms of the volume element on the surface of the sphere, i.e.,

$$\int_0^\pi \int_0^{2\pi} Y_\ell^m(\phi, \theta) Y_{\ell'}^{m'}(\phi, \theta) \sin\theta d\phi d\theta = 0,$$

unless $\ell = \ell'$ or $m = m'$ (they can be normalized but here for simplicity we do not provide the normalization constants). The associated Legendre polynomials can be provided explicitly, for example (omitting the explicit dependence on $\cos\theta$ from the polynomials for convenience and expressing them as functions of $\theta$ directly)

$$\begin{aligned}
&\ell = 0, \; m = 0, \; P_0^0(\theta) = 1, \\
&\ell = 1, \; m = 0, \; P_1^0(\theta) = \cos\theta, \; m = \pm 1 \; P_1^{\pm 1}(\theta) = \sin\theta, \\
&\ell = 2, \; m = 0, \; P_2^0(\theta) = \frac{1}{2}(3\cos^2\theta - 1), \; m = \pm 1, P_2^{\pm 1}(\theta) = 3\cos\theta\sin\theta, \; m = \pm 2, \; P_2^{\pm 2}(\theta) = 3\sin^2\theta.
\end{aligned}$$

In terms of the above, the eigenfunctions $\phi_\ell$ can be expressed as

$$\begin{aligned}
\ell = 0: \;\; & \phi_0(\phi, \theta) = c_{0,0}, \\
\ell = 1: \;\; & \phi_1(\phi, \theta) = c_{1,0} P_1^0(\theta) + c_{1,1}\cos\phi P_1^1(\theta) + s_{1,1}\sin\phi P_1^1(\theta) \\
\ell = 2: \;\; & \phi_2(\phi, \theta) = c_{2,0} P_2^0(\theta) + c_{2,1}\cos\phi P_2^1(\theta) + s_{2,1}\sin\phi P_2^1(\theta) \\
& + c_{2,2}\cos(2\phi) P_2^2(\theta) + s_{2,2}\sin(2\phi) P_2^2(\theta)
\end{aligned}$$

where $c_{\ell,m}$ and $s_{\ell,m}$ are appropriate (real valued) constants.

The eigenmode $\ell > 0$ will become unstable if $\mu_\ell = \frac{1}{R^2}\ell(\ell+1)$ satisfies

$$D(k_0^*)\mu_\ell(r - D(k_0^*)\mu_\ell) - M > 0.$$

Again, the spatial structure of the pattern will be given as a linear combination of the unstable eigenmodes, with the coefficients depending on the initial condition. High modes (large values of $\ell$) are unlikely to become unstable as condition (17) will be violated for large values of $\ell$. Of course for large $R$ the corresponding values of $\ell$ are expected to be larger, leading to more complex spatial structures.

Figure 5 presents the emergence of patterns on the sphere. It should be noted that the patterns presented have not association with actual geographical areas or counties on the earth-like globe, but are presented solely for illustrations. Associating the merging spatial patterns with actual location on earth is an important area for future research.

$(a)$ $(b)$

Figure 5: Pattern formation on a shere for capital (a) and consumption (b), $r = 0.1$, $\lambda = 6$, $\alpha = 0.6, \delta = 0.03$.

### 4.2 Spatial heterogeneity and MPD capital mobility

In the previous section we show that optimal growth under MPD capital mobility will not generate spatial heterogeneities for a typical Cobb-Douglas and utility parametrization and low discount rate. This result as derived under the assumption of equal TFP across locations or $A$ independent of location $x$. In this section we allow for for TFP differences across locations, that is $A = A(x)$ and we ask the question of weather MPD capital mobility combined with TFP heterogeneity intensifies or weakens spatial heterogeneity.

To study this problem we assume that capital is immobile but there is spatial variability on the TFP parameter $A(x)$. Then, the optimal capital allocation is given by the standard Ramsey model, parameterized by $A = A(x)$, i.e. the solution of the parametric steady state equation

$$A(x)f'(k) = r + \delta$$

which leads to an optimal allocation $k_0(x)$ with the $x$ dependence arising from the dependence of $A$ on $x$, and to a corresponding optimal consumption $c_0(x)$, obtained in terms of

$$c_0(x) = A(x)f(k_0) - \delta k_0.$$

An interesting question will be how would this non flat steady state $(k_0, c_0)$ evolve in the presence of weak capital mobility $\bar{D}_0$, if perturbed by a perturbation $(k_1, c_1)$. This requires the stability analysis of the Pontryagin system for such solutions, in the limit of small $\bar{D}_0$.

The following proposition provides some insight towards this.

**Proposition 4.5.** *Assume a spatially varying TFP, $A = A(x)$, and let $(k_0, c_0)$ be the non flat steady state corresponding to the optimal steady state in the absence of capital mobility. Assume small capital mobility $\bar{D}_0 \neq 0$.*

*(i) Then, an initial configuration $(k_0, c_0)$ will evolve under the Pontryagin optimality conditions to a spatially dependent steady state*

$$k(x) = k_0(x) + \bar{D}_0 \frac{1}{U'(c_0(x))} \nabla \cdot \Big( B\psi(k_0(x)) k_0(x) U''(c_0(x)) \nabla c_0(x) \Big) + O(\bar{D}_0^2),$$
$$c(x) = c_0(x) + \bar{D}_0 \frac{r}{U'(c_0(x))} \nabla \cdot \Big( B\psi(k_0(x)) k_0(x) U''(c_0(x)) \nabla c_0(x)) + O(\bar{D}_0^2) \Big).$$

*(ii) Assume the Cobb-Douglas case, where $f(k) = k^\alpha$, $\psi(k) = k^\rho$ and $U(c) = \frac{1}{1-\lambda} c^{1-\lambda}$ with $0 < \alpha < 1$, $\lambda > 1$, and set $B = 1$ for simplicity. Then,*

$$k_0(x) = M_0 A(x)^{\frac{1}{1-\alpha}}, \;\; c_0(x) = \frac{(1-\alpha)\delta + r}{\alpha} M_0 A^{\frac{1}{1-\alpha}}(x), \;\; M_0 = \Big( \frac{\delta + r}{\alpha} \Big)^{-\frac{1}{1-\alpha}},$$

*and*

$$k(x) = k_0(x) \Big( 1 - \bar{D}_0 \Psi(x) + O(\bar{D}_0^2) \Big),$$
$$c(x) = c_0(x) \Big( 1 - \bar{D}_0 \frac{\alpha r}{(1-\alpha)\delta + r} \Psi(x) + O(\bar{D}_0^2) \Big),$$
$$\Psi(x) = \frac{\lambda}{1-\alpha} M_0^\rho A^{\frac{\lambda - 1}{1-\alpha}}(x) \nabla \cdot \Big( A^{\frac{\rho - \lambda + \alpha}{1-\alpha}} \nabla A(x) \Big).$$

For the proof see Appendix F.

Proposition 4.5 allows us to approximate explicitly, analytically, the effect of small capital mobility on possibly sharp spatial gradients in the capital factor of productivity $A$.

**Example 4.6.** Consider the Cobb-Douglas case in Proposition 4.5(ii) and a variation in the TFP $A$ in the form of a sum of gaussians as

$$A(x) = C_0 + \sum_{i=1}^{N} C_i \exp\Big( - \frac{\|x - x_i\|^2}{2\sigma_i} \Big).$$

This general form can model local increases ($C_i > 0$) or decreases ($C_i < 0$) at locations $x_i$, of scale $\sigma_i > 0$ (small $\sigma_i$ corresponding to a localized spatial structure, large $\sigma_i$ corresponding to extended spatial structures).

In this case we can calculate $(k(x), c(x))$ in terms of the expression in Proposition 4.5(ii), using the facts that

$$\Psi(x) = \frac{\lambda}{1-\alpha} M_0^\rho \Big\{ A^{(\frac{\rho}{1-\alpha})-1} \Phi_1(x) + \frac{\rho - \lambda + \alpha}{1-\alpha} A^{(\frac{\rho}{1-\alpha})-2} \Phi_2(x) \Big\}$$
$$\Phi_1(x) = \sum_{i=1}^{N} C_i \Big( \frac{d}{\sigma_i} - \frac{1}{\sigma_i^2} \|x - x_i\|^2 \Big) \exp\Big( - \frac{|x - x_i|^2}{2\sigma_i} \Big)$$
$$\Phi_2(x) = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{C_i C_j}{\sigma_i \sigma_j} \exp\Big( - \frac{|x - x_i|^2}{2\sigma_i} \Big) \exp\Big( - \frac{|x - x_j|^2}{2\sigma_j} \Big) (x - x_i) \cdot (x - x_j),$$

where $d$ is the spatial dimension and $(x - x_i) \cdot (x - x_j)$ denotes the inner product between the position vectors $x - x_i$ and $x - x_j$.

We now consider the following question: Will the effect of capital mobility ($\bar{D}_0 \neq 0$) enhance (sharpen) the spatial gradients induced by the variability of $A$ or act as a mechanism for reducing them?

To answer this question we will consider the steady state solution of the Ramsey model for a spatially dependent $A$ and $\bar{D}_0 = 0$, denoted by $k_0$ and its steady state solution for the same $A$ but for $\bar{D}_0 \neq 0$, denoted by $k$, and compare the Sobolev norms $I_0 = \int_{\mathcal{D}} |\nabla k_0|^2 dx$, and $I = \int_{\mathcal{D}} |\nabla k|^2 dx$, which provide a measure for the overall spatial variability of the optimal capital allocation. If $I_0 \leq I$ then the effect of capital mobility will be an enhancement of the spatial variability while if $I_0 \geq I$ then the effect of capital mobility will be a smoothing of spatial variability in the optimal capital distribution.

While one may provide a priori estimates for the above quantities, this will require technicalities which are beyond the scope of the present paper. We prefer to adopt a more direct treatment to this question, which provides an easy to interpret analytic answer, in the limit of small values of $\bar{D}_0$ and in the case where $A$ assumes the form of a single Gaussian pulse. We further assume that $f(k) = Ak^\alpha$ and $\psi(k) = k^\rho$.

**Proposition 4.7.** *Under the additional assumptions that $f(k) = Ak^\alpha$, $\psi(k) = k^\rho$, $U(c) = \frac{1}{1-\lambda} c^{1-\lambda}$, and $A(x) = C \exp\left(-\frac{|x-x_0|^2}{2\sigma^2}\right)$, $\sigma > 0$, for sufficiently small $D_0$, it holds that*

$$I_0 := \int_{\mathcal{D}} |\nabla k_0|^2 dx, \quad I := \int_{\mathcal{D}} |\nabla k|^2 dx$$

*satisfy the relation*

$$I - I_0 = \bar{D}_0 S E(d) + O(\bar{D}_0^2),$$

*where $S > 0$ (explicitly given in the appendix) and $E(d)$ depends on the spatial dimension and is*

$$\begin{aligned} E(1) &= 3(\rho + 1) + (\rho - 1)\lambda, \quad \text{if } d = 1, \\ E(2) &= 8(\rho + 1) + 4\lambda\rho, \quad \text{if } d = 2. \end{aligned}$$

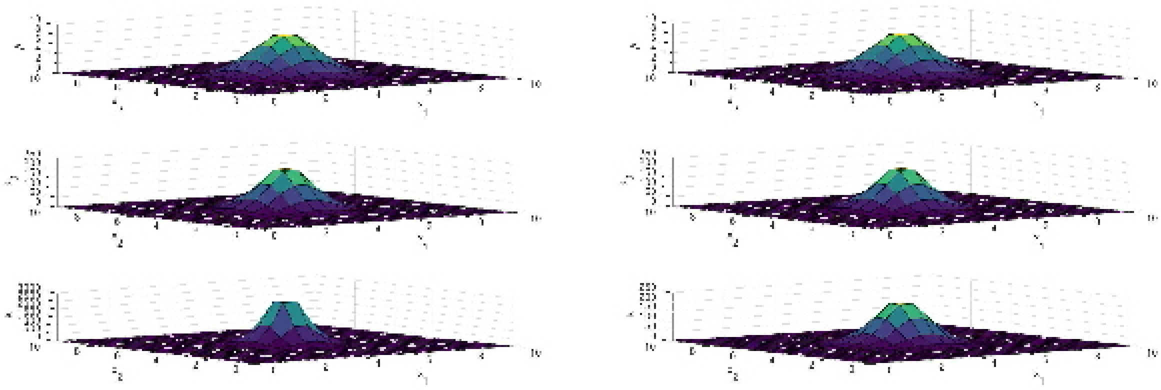
*For proof see Appendix F.*

The above calculations show that up to small order corrections in $\bar{D}_0$:

- The Sobolev norm of the solution will increase in an one-dimensional spatial domain ($1D$) as long as $\rho \geq 1$, and in a two-dimensional spatial domain ($2D$) for any $\rho > 0$.
- A decrease in the Sobolev norm in $1D$ is only feasible if $0 \geq \rho < 1$ and $\lambda > \frac{3(\rho+1)}{1-\rho}$ (i.e. for $\rho = 0$ if $\lambda > 3$), while in $2D$ a decrease in the Sobolev norm is only feasible for $\rho < 0$.

The change in the Sobolev norm depends on the value of $\rho$ which, given the specification $\psi(k) = k^\rho$, indicates that the propensity of capital to move to a high marginal productivity location is high if the capital stock at the location of origin is high $(\rho > 0)$, or that the propensity of capital to move to a high marginal productivity location is high if the capital stock at the location of origin is low $(\rho < 0)$. Since the Sobolev norm can be regarded as a summary measure of spatial heterogeneity within the spatial domain the above result suggests that MPD capital mobility could intensify or weaken spatial inequalities depending on the way that the propensity of capital to move depends on the size of the existing capital stock.

The case of increased or reduced spatial inequalities when MPD capital mobility occurs is depicted in Figure 6.



$(a)$ $(b)$

Figure 6: $(a)$ Increasing Sobolev norm, $\rho = 1, (b)$ Decreasing Sobolev norm $\rho = -1$

# 5 Concluding Remarks

Spatiotemporal dynamics are introduced in a standard Ramsey model of optimal growth by considering capital movement towards locations where the marginal productivity of capita is relatively higher. This induces nonlinear diffusion in the fundamental equation of capital accumulation. To accommodated this in the optimal control framework of Ramsey model we extend Pontryagin's maximum principle to the case in which transition dynamics are governed by a nonlinear partial differential equation.

In this context we examine questions related the potential spatial heterogeneity of optimal growth as seen from the point of view of a social planner which seeks to maximize discounted utility over a finite spatial domain by choosing optimal consumption paths for each location. Our results suggest that for high utility discount rate and appropriate parameters for the production and the utility function MPD capital flows, optimizing behavior by the social planner could induce the emergence of spatial patterns. For low utility discount rate the social planner will choose the optimal policy so that the spatial economy will return to a flat-earth steady state even after a perturbation caused by MPD capital flows. We also show that when spatial heterogeneities exists due to TFP differences across locations, MPD capital flows could intensify or weaken spatial inequalities. This depends on whether the tendency of capital from high capital accumulation locations to move once higher marginal productivity occurs in a different location is high or low.

Our results provide insights of the way in which an intuitive plausible mechanism of capital flows could generate spatial inequalities in the context of traditional Ramsey model of optimal growth. Further research should be directed towards studying market equilibrium under MPD capital flows. If market equilibrium outcomes in terms of spatial heterogeneity are not the same as the outcomes obtained in this paper for a social planner, this will provide the basis for exploring economic policies for attaining socially optimal spatial structures.

# Appendix

## A Derivation of the Pontryagin principle for Problem (4)

In this section we derive the Pontryagin maximum principle for problem (4). We define the functional $J$, by $c \mapsto \int_0^\infty \int_{\mathcal{D}} e^{-rt} U(c(t,x))dxdt$, where $c : \mathbb{R}_+ \times \mathcal{D} \to \mathbb{R}_+$, is the consumption function. Let us denote by $(k^*, c^*)$, the optimal capital path and the optimal consumption, and let us consider any perturbation of the optimal consumption $c^* + \epsilon c$, which will in turn lead to a perturbation of the optimal capital path $k^* + \epsilon k$. This perturbation will have to be consistent with the spatiotemporal dynamics of capital accumulation and transport, i.e. will have to satisfy the quasilinear PDE (3). Under differentiability assumptions, we may linearize (3) in $\epsilon$, and conclude that small deviations $(k, c)$ from the optimal path $(k^*, c^*)$ will satisfy the linear PDE

$$\begin{aligned}\frac{\partial k}{\partial t}(t,x) &= \nabla \cdot \, \{D_0(x) w_1(k^*(t,x)) \nabla \, k(t,x) + \underline{G}(k^*(t,x),x) k(t,x)\} \\ &\quad + f'^*(t,x)) k(t,x) - \delta k(t,x) - c(t,x),\end{aligned} \tag{21}$$

where

$$\underline{G}(k^*(t,x),x) := D_0(x) w_1'^*(t,x)) \nabla \, k^*(t,x) - D_0(x) w_2'^*(t,x)) \nabla \, \ln A(x),$$

with the perturbed boundary condition

$$\left\{ D_0(x)(k^*(t,x)) \nabla \, k(t,x) + k(t,x) \underline{G}(k^*(t,x),x) k(t,x) \right\} \cdot n(x) = 0, \quad x \in \partial\mathcal{D},$$

where $\partial\mathcal{D}$ is the boundary of $\mathcal{D}$ (assumed sufficiently smooth) and $n$ is the outward normal at $x \in \partial\mathcal{D}$. Note that $\underline{G}$ is a vector field.

Consider now an adjoint variable $p(t,x)$ at each point $(t,x) \in \mathbb{R}_+ \times \mathcal{D}$, the evolution law of which will be determined shortly, and define the temporal quantity

$$I(t) := \int_{\mathcal{D}} k(t,x) p(t,x) dx.$$

Assuming enough smoothness in order to be able to differentiate under the integral sign (invoking Lebesgue's dominated convergence theorem) we have that

$$\frac{d}{dt} I(t) = I_1 + I_2 := \int_{\mathcal{D}} \left( \frac{\partial k}{\partial t}(t,x) p(t,x) + \frac{\partial p}{\partial t}(t,x) k(t,x) \right) dx. \tag{22}$$

By (21) we have

$$\begin{aligned}I_1 &= \int_{\mathcal{D}} \left( \nabla \cdot \, \{D(k^*(t,x)) \nabla \, k(t,x) + D'^*(t,x) \nabla \, k^*(t,x) k(t,x)\} \right) p(t,x) dx \\ &\quad + \int_{\mathcal{D}} \left( f'(k^*(t,x)) k(t,x) - \delta k(t,x) - c(t,x) \right) p(t,x) dx.\end{aligned} \tag{23}$$

At this point we recall Green's theorem according to which for a vector field $\underline{F}$ and a scalar field $g$,

$$\int_{\mathcal{D}} (\underline{F}(x) \cdot \nabla g(x) + g \nabla \cdot \underline{F}(x)) dx = \int_{\partial D} g(x) \underline{F}(x) \cdot n(x) dS,$$

where by $n(x)$ we denote the outward normal vector at $x \in \partial \mathcal{D}$. Naturally sufficient smoothness on $\partial \mathcal{D}$ must be assumed for the outward normal to be well defined. The boundary term vanishes as long as either $\underline{F} \cdot n = 0$ or $g = 0$ on $\partial \mathcal{D}$, i.e. if either $\underline{F}$ satisfies no flux boundary conditions (e.g. Neumann or Robin boundary conditions) or $g$ satisfies Dirichlet boundary conditions.

We apply Green's theorem to the first integral of (23), setting $\underline{F} = D_0 w_1(k^*) \nabla k + \underline{G} k$ and $g = p$ (dropping the explicit $(t, x)$ dependence to facilitate the notation) and we have that

$$\begin{aligned}
\int_{\mathcal{D}} \nabla \cdot (D_0 w_1(k^*) \nabla k + \underline{G} k) p dx &= - \int_{\mathcal{D}} (D_0 w_1(k^*) \nabla k + \underline{G} k) \cdot \nabla p dx, \\
&= - \int_{\mathcal{D}} D_0 w_1(k^*) \nabla k \cdot \nabla p dx - \int_{\mathcal{D}} \underline{G} \cdot p \, k \, dx
\end{aligned}$$

where the boundary term vanishes because of the linearized boundary condition (**??**). We now apply once more Green's theorem to the first integral, choosing $\underline{F} = D_0 w_1(k^*) \nabla p$ and $g = k$, and setting $\nabla p \cdot n = 0$ on $\partial \mathcal{D}$ to eliminate the boundary term we obtain that

$$\int_{\mathcal{D}} \nabla \cdot (D_0 w_1(k^*) \nabla k + \underline{G} k) p dx = \int_{\mathcal{D}} \nabla \cdot (D_0 w_1(k^*) \nabla p) \, k \, dx - \int_{\mathcal{D}} (\underline{G}(k^*, x) \cdot \nabla p) \, k \, dx.$$

Substituting these results in (23) (and re-instated the explicit $(t, x)$ dependence) we obtain that

$$\begin{aligned}
I_1 = \int_{\mathcal{D}} \Big\{ & \nabla \cdot \{ D_0(x) w_1(k^*(t,x)) \nabla p(t,x) \} - \underline{G}(k^*(t,x), x) \cdot \nabla p(t,x) \\
& + f'(k^*(t,x)) p(t,x) - \delta p(t,x) \Big\} k(t,x) dx - \int_{\mathcal{D}} c(t,x) p(t,x) dx.
\end{aligned}$$

Combining that with (22) we see that

$$\begin{aligned}
\frac{dI}{dt} = \int_{\mathcal{D}} \Big\{ & \nabla \cdot \{ D_0(x) w_1(k^*(t,x)) \nabla p(t,x) \} - \underline{G}(k^*(t,x), x) \cdot \nabla p(t,x) \\
+ f'(k^*(t,x)) p(t,x) - \delta p(t,x) + & \frac{\partial p}{\partial t}(t,x) \Big\} k(t,x) dx - \int_{\mathcal{D}} c(t,x) p(t,x) dx.
\end{aligned} \tag{24}$$

and upon choosing

$$\begin{aligned}
& \nabla \cdot (D_0(x) w_1(k^*(t,x)) \nabla p(t,x)) - \underline{G}(k^*(t,x), x) \cdot \nabla p(t,x) \\
& + f'(k^*(t,x)) p(t,x) - \delta p(t,x) + \frac{\partial p}{\partial t}(t,x) = r p(t,x), \quad x \in \mathcal{D},
\end{aligned}$$

we express (24) as

$$\frac{dI}{dt}(t) = rI(t) - \int_{\mathcal{D}} c(t,x)p(t,x)dx.$$

Setting $\bar{I}(t) = e^{-rt}I(t)$, the above becomes

$$\frac{d\bar{I}}{dt}(t) = -e^{-rt} \int_{\mathcal{D}} c(t,x)p(t,x)dx$$

which integrated over the time interval $[0,T]$ yields

$$\bar{I}(T) - \bar{I}(0) = -\int_0^T e^{-rt} \int_{\mathcal{D}} c(t,x)p(t,x)dxdt,$$

or equivalently,

$$e^{-rT} \int_{\mathcal{D}} k(T,x)p(T,x)dx - \int_{\mathcal{D}} k(0,x)p(0,x)dx = -\int_0^T e^{-rt} \int_{\mathcal{D}} c(t,x)p(t,x)dxdt,$$

which upon choosing $k(0,x) = 0$ a.e $x \in \mathcal{D}$ and $\lim_{T\to\infty} e^{-rT} \int_{\mathcal{D}} k(T,x)p(T,x)dx = 0$ leads to the result that

$$0 = -\int_0^\infty e^{-rt} \int_{\mathcal{D}} c(t,x)p(t,x)dxdt. \tag{25}$$

We now consider the effect of the perturbation of the optimal policy on the control objective. Assuming once more differentiability of the utility function we see that for any $\epsilon > 0$,

$$\frac{1}{\epsilon}(J(c^* + \epsilon c) - J(c^*)) = \int_0^\infty \int_{\mathcal{D}} e^{-rt} U'(c^*(t,x))c(t,x)dxdt$$

and adding (25) to this we obtain

$$\frac{1}{\epsilon}\Big(J(c^* + \epsilon c) - J(c^*)\Big) = \int_0^\infty \int_{\mathcal{D}} e^{-rt}\Big(U'(c^*(t,x)) - p(t,x)\Big)c(t,x)dxdt.$$

Since a maximum is achieved for $J$ at $c^*$ we have that for $\epsilon > 0$ small enough,

$$\frac{1}{\epsilon}\Big(J(c^* + \epsilon c) - J(c^*)\Big) \leq 0,$$

for every $c$ such that $c^*(t,x) + \epsilon c(t,x) \geq 0$, a.e. $(t,x) \in \mathbb{R}_+ \times \mathcal{D}$, for sufficiently small $\epsilon > 0$, which when combined with (**??**) leads to the inequality

$$\int_0^\infty \int_{\mathcal{D}} e^{-rt}\Big(U'(c^*(t,x)) - p(t,x)\Big)c(t,x)dxdt \leq 0,$$

for every $c$ such that $c^*(t,x) + \epsilon c(t,x) \geq 0$, a.e. $(t,x) \in \mathbb{R}_+ \times \mathcal{D}$, for sufficiently small $\epsilon > 0$. The condition (**??**) is a variational inequality. If $c^*(t,x) > 0$, a.e. $(t,x) \in \mathbb{R}_+ \times \mathcal{D}$, then (under continuity assumptions on $c^*$) we see that if (**??**) holds for any $c$ then it will also hold for $-c$ as well, as long as $\epsilon > 0$ is chosen small enough, so that (**??**) yields

$$\int_0^\infty \int_{\mathcal{D}} e^{-rt}\Big(U'(c^*(t,x)) - p(t,x)\Big)c(t,x)dxdt = 0,$$

which means that

$$U'^*(t,x)) - p(t,x) = 0, \;\; a.e. \; (t,x) \in \mathbb{R}_+ \times \mathcal{D}.$$

Collecting all the above we see that if $(k^*, c^*)$ is an optimal path such that $c^*(t,x) > 0$, and $p$ solves the PDE

$$\frac{\partial p}{\partial t}(t,x) = -\nabla \cdot \left(D(k^*(t,x))\nabla p(t,x)\right) + \underline{G}'(k^*(t,x),x) \cdot \nabla p(t,x) \qquad (26)$$
$$-f'(k^*(t,x))p(t,x) + \delta p(t,x) + rp(t,x),,$$

with homogeneous Neumann boundary conditions for $p$,

$$\nabla p(t,x) \cdot n(x) = 0, \;\; x \in \partial\mathcal{D},$$

and the transversality condition

$$\lim_{T\to\infty} e^{-rT} \int_{\mathcal{D}} k(t,x)p(t,x)dx = 0,$$

then the optimal consumption is characterized by

$$U'^*(t,x)) - p(t,x) = 0, \;\; a.e. \; (t,x) \in \mathbb{R}_+ \times \mathcal{D}, \qquad (27)$$

and the nonlinear PDE

$$\frac{\partial k^*}{\partial t}(t,x) = \nabla \cdot \left(D(k^*(t,x))\nabla k^*(t,x)\right) + f(k^*(t,x)) - \delta k^*(t,x) - c^*(t,x),$$

with initial condition $k^*(0,x) = k_0(x)$ and homogeneous Neumann boundary conditions $\nabla k^*(t,x) \cdot n(x) = 0$ on $\partial\mathcal{D}$. Note that the transversallity condition is satisfied as long as $k(T,x)$ is bounded and $\lim_{T\to\infty} e^{-rT}p(T,x) = 0$ for every $x \in \mathcal{D}$ (note that the boundedness of $k(T,x)$ must be shown using (21) with initial condition $k(0,x) = 0$). Furthermore, by using the vector calculus identities and equivalent form for (26) is

$$\frac{\partial p}{\partial t}(t,x) = -D(k^*(t,x))\nabla \cdot \nabla p(t,x) - f'(k^*(t,x))p(t,x) + \delta p(t,x) + rp(t,x), \qquad (28)$$

where we will use the standard notation $\Delta p(t,x) := \nabla \cdot \nabla p(t,x)$, for the Laplacian.

Using (27) and defining by $G$ the inverse function of $U'$ we obtain the coupled forward-backward system of PDE

$$\frac{\partial k^*}{\partial t}(t,x) = \nabla \cdot \, (D(k^*(t,x))\nabla\, k^*(t,x)) + f(k^*(t,x)) - \delta k^*(t,x) - G(p(t,x)), \tag{29}$$

$$\frac{\partial p}{\partial t}(t,x) = -\nabla \cdot \, \{D(k^*(t,x))\nabla\, p(t,x)\} + D'(k^*(t,x))\nabla\, k^*(t,x)\cdot \nabla\, p(t,x) \tag{30}$$
$$+(r+\delta - f'(k^*(t,x)))p(t,x),$$

with

$$\nabla\, k^*(t,x)\cdot n(x) = \nabla\, p(t,x)\cdot n(x) = 0, \;\; x \in \partial\mathcal{D},$$
$$k^*(0,x) = k_0(x), \;\; x \in \mathcal{D},$$
$$\lim_{T\to\infty} e^{-rT} p(T,x) = 0, \;\; x \in \mathcal{D}.$$

The optimal state capital spatio-temporal allocation $k^*$ is characterized by the solution of this system whereas the optimal consumption $c^*$ is recovered from the adjoint variable upon setting $c^* = G(p)$ pointwise (or equivalently solving $U'(c^*(t,x)) = p(t,x)$ pointwise).

# B Hamiltonian structure

The Hamiltonian structure arises by first considering the Lagrangian form of the problem by eliminating consumption from the budget constraint (assumed to be binding on account of strict monotonicity of the utility function $u$). This yields the problem in the calculus of variations form

$$\max_k \int_0^\infty \int_{\mathcal{D}} e^{-rt} u(\nabla \cdot \, (D(k)\nabla\, k) + f(k) - \delta k - k')dxdt.$$

We will express that as a minimization problem of the form

$$\min_k \int_0^\infty \int_{\mathcal{D}} L(t, k(t,x), k'(t,x))dxdt,$$

where $L$ is the Lagrangian density

$$L(t,k,k' = e^{-rt} u(\nabla \cdot \, (D(k)\nabla\, k) + f(k) - \delta k - k').$$

Then, using standard techniques from convex analysis we can define the Hamiltonian density

$$\bar{H}(t,k,\bar{p}) := \sup_{k'} \Big\{ \bar{p}k' - L(t,k,k') \Big\},$$

which allows us to bring the Pontryagin system in Hamiltonian form.

## C The Kirkhoff transformation

The above system of elliptic equations may in fact be transformed into a more convenient equivalent form.

**Proposition C.1.** *Define the variables* $u = \Phi(k^*)$ *and* $v = \Psi(c^*)$ *where* $\Phi'(s) = w_1(s)$ *and* $\Psi = U'$, [14] *and assume that* $\Phi$ *and* $\Psi$ *are invertible and denote their inverse by* $\phi$, $\psi$ *respectively, so that* $k^* = \phi(u)$ *and* $c^* = \psi(v)$. *Then* $(u, v)$ *satisfies the system of semilinear elliptic equations*

$$0 = \Delta u + \nabla \ln D_0 \cdot \nabla u - w_2(\phi(u))\Delta \ln A - w_2(\phi(u))\nabla \ln D_0 \cdot \nabla \ln A - \tag{31}$$

$$(w_2(\phi(u))'\nabla u \cdot \nabla \ln A + \frac{1}{D_0}\Big(f(\phi(u)) - \delta\phi(u) - \psi(v)\Big), \tag{32}$$

$$0 = \Delta v + \nabla \ln D_0 \cdot \nabla v + \frac{w_2'(\phi(u))}{w_1(\phi(u))}\nabla \ln A \cdot \nabla v - \frac{(r + \delta - f'(\phi(u)))}{D_0 w_1(\phi(u))}v,$$

*which Neumann boundary conditions*

The proof is straightforward and uses the Kirkhoff transformation mentioned in the exposition and algebraic manipulation. The invertibility of the transformation is guaranteed by monotonicity arguments (e.g. it suffices that $D(s) > 0$ and that $U'$ is strictly monotone). Note however, that even in the absense of the above conditions the system (31) may be interpreted in terms of inclusions. Note that if the coefficients $A$ and $B$ (hence also $D_0 = AB$) are spatially independent this system simplifies considerably as all the gradient terms disappear, rendering it very convenient for analysis on account of its seilinear form. If at some point $w_1(\phi(u)) = 0$ then the second equation of (31) has to be interpreted as

$$0 = D_0 w_1(\phi(u))\Delta v - +\delta - f'(\phi(u)))v,$$

as a singular elliptic PDE, with a similar interpretation in the case of spatially dependent coefficients $A$ and $B$.

## D On the stability of steady states

The following proposition provides some information on the evolution of perturbations to a steady state.

**Proposition D.1.** *Consider a steady state solution* $(k^*, c^*)$ *of the nonlinear system* (**??**) *and assume a small time dependent perturbation of the optimal steady state of the form* $(k^* + \epsilon k, c^* + \epsilon c)$. *Then,*

[14] Note that $v = p$.

*under sufficient smoothness conditions the functions* $(k, c)$ *satisfy the linear parabolic system*

$$\frac{\partial}{\partial t}k = \nabla\cdot\Big(D_0 w_1(k^*)\nabla k + (D_0 w_1'^*)\nabla k^*)k - (D_0 w_2'^*)\nabla \ln A)k\Big) + (f'(k^*) - \delta)k - c,$$

$$U''(c^*)\frac{\partial}{\partial t}c = -\nabla\cdot\Big(D_0 w_1(k^*)\nabla(U''(c^*)c) + D_0 w_1'(k^*)(\nabla(U'(c^*))k\Big)$$

$$+\underline{G}(k^*, x)\cdot\nabla(U''(c^*)c) + \Big((D_0 w_1''(k^*)\nabla k^*)k + D_0 w_1'(k^*)\nabla k - (D_0 w_2''(k^*)\nabla \ln A)k\Big)\cdot\nabla U'(c^*)$$

$$+(r + \delta - f'(k^*))U''(c^*)c - f''(k^*)U'(c^*)k$$

*Consider the elliptic eigenvalue problem*

$$\rho k = \nabla\cdot\Big(D_0 w_1(k^*)\nabla k + (D_0 w_1'(k^*)\nabla k^*)k - (D_0 w_2'(k^*)\nabla \ln A)k\Big) + (f'(k^*) - \delta)k - c,$$

$$\rho U''(c^*)c = -\nabla\cdot\Big(D_0 w_1(k^*)\nabla(U''(c^*)c) + D_0 w_1'(k^*)(\nabla(U'(c^*))k\Big)$$

$$+\underline{G}(k^*, x)\cdot\nabla(U''(c^*)c) + \Big((D_0 w_1''(k^*)\nabla k^*)k + D_0 w_1'(k^*)\nabla k - (D_0 w_2''(k^*)\nabla \ln A)k\Big)\cdot\nabla U'(c^*)$$

$$+(r + \delta - f'(k^*))U''(c^*)c - f''(k^*)U'(c^*)k$$

*with Neumann boundary conditions. If the eigenvalues have positive real parts then* $(k^*, c^*)$ *compatible with the transversality condition this corresponds to a saddle point.*

The above proposition provides a generalization of the usual method of treating optimal control problems for temporal problems in the phase space, looking for a saddle point solution. The treatment of the linearized system and the corresponding elliptic eigenvalue problem is by no means an easy task as the relevant systems are now with spatially dependent coefficients. There are limited analytical tools to treat such systems, and one may either content to a priori estimates for the spectrum or to numerical analysis.

# E Proof of Proposition 4.1

Consider the eigenvalue problem

$$-\Delta\phi(x) = \mu\phi(x), \;\; x \in \mathcal{D},$$
$$\nabla\phi(x)\cdot n = 0, \;\; x \in \partial\mathcal{D}.$$

Setting this problem in the proper functional setting (e.g. considering that $\phi \in W^{1,2}(\mathcal{D})$ the set of $L^2(\mathcal{D})$ functions with first order weak derivatives in $L^2(\mathcal{D})$), it is well known that this problem admits a discrete set of eigenvalues $\{\mu_n \;\; n \in \mathbb{N}\}$ such that $\mu_n \to \infty$ as $n \to \infty$, corresponding to a set of eigenfunctions $\{\phi_n \;:\; n \in \mathbb{N}\}$ which importantly are orthogonal to each other with respect to the inner product in $L^2(\mathcal{D})$. Furthermore, this set of eigenfunctions constitutes a complete basis is $L^2(\mathcal{D})$, so that any $\psi \in L^2(\mathcal{D})$ can be expanded in a generalized Fourier series as $\psi = \sum_{n\in\mathbb{N}} \psi_n \phi$ with

$\psi_n = \langle \psi, \phi_n \rangle_{L^2(\mathcal{D})} = \int_{\mathcal{D}} \psi(x)\phi_n(x)dx$.

This observation allows us to expand any solution $(k(t,x), c(t,x))$ of (14) as

$$k(t,x) = \sum_{n\in\mathbb{N}} k_n(t)\phi_n(x),$$
$$c(t,x) = \sum_{n\in\mathbb{N}} c_n(t)\phi_n(x),$$

where $(k_n, c_n)$ are functions of time only to be determined and all the spatial variability (including the boundary conditions) is captured by the eigenfunction basis $\{\phi_n \ : \ n \in \mathbb{N}\}$. By the same token, we may expand the initial conditions $(k_0, c_0)$ in the same basis as

$$k_0(x) = \sum_{n\in\mathbb{N}} k_{0,n}\phi_n(x),$$
$$c_0(x) = \sum_{n\in\mathbb{N}} c_{0,n}\phi_n(x),$$

with $k_{0,n} = \int_{\mathcal{D}} k_0(x)\phi_n(x)dx$, and $c_{0,n} = \int_{\mathcal{D}} c_0(x)\phi_n(x)dx$.

Substituting the above expansions in (14) and taking into account the fact that $\phi_n$ are eigenfunctions of the Laplacian operator as well as the orthogonality of the eigenfunctions, upon projection on the eigenspaces spanned by these eigenvectors we conclude that the functions $z_n = (k_n, c_n)^T$ solve the countable system of ODEs

$$z_n' = L_n z_n, \qquad (33)$$
$$z_n(0) = z_{n,0},$$

where

$$L_n = \begin{pmatrix} -D(k^*)\mu_n + f'(k^*) - \delta & -1 \\ -\frac{U'^*)}{U''(c^*)} f''(k^*) & D(k^*)\mu_n \end{pmatrix} = \begin{pmatrix} -D(k^*)\mu_n + r & -1 \\ -\frac{U'(c^*)}{U''(c^*)} f''(k^*) & D(k^*)\mu_n \end{pmatrix},$$

and we took into account the fact that $\frac{w_2(k^*)}{w_1(k^*)} = -\frac{f'^*)}{f''(k^*)} > 0$.

The solution to system (33) can be expressed in terms of the exponential of the matrices $L_n$ defined by $\exp(tL_n) := I + \sum_{k=1}^{\infty} \frac{1}{k!} t^k L_n^k$, using the formula

$$z_n(t) = \exp(tL_n) z_{n,0}.$$

This expression shows that the behaviour of $z_n(t)$ is characterized by the behaviour of the matrix exponential $\exp(tL_n)$ and depends on the initial condition. If this matrix exponential grows as $t \to \infty$, then we will observe long lasting perturbations to the flat steady state, whereas is the matrix exponential decays as $t \to \infty$, then we will not observe long lasting perturbations to the flat steady state, and the system will asymptotically return to the flat steady state.

The matrix exponential $\exp(tL_n)$ can be calculated in terms of the eigenvalues and the eigenvectors of the matrix $L_n$, and it form depends on the spectrum of the matrix. In particular let $\rho_{1,n}, \rho_{2,n}$ be the eigenvalues of $L_n$. There are three possible cases

A. $\rho_{1,n}, \rho_{2,n}$ real and distinct

$$\exp(tL_n) = \frac{1}{\rho_{1,n} - \rho_{2,n}} \Big( e^{\rho_{1,n} t}(L_n - \rho_{2,n} I) - e^{\rho_{2,n} t}(L_n - \rho_{1,n} I) \Big),$$

B. $\rho_{1,n} = \rho_{2,n} = \rho_n$,

$$\exp(tL_n) = e^{\rho_n t} I + e^{\rho_n t} t (L_n - \rho_n I),$$

C. $\rho_{1,n} = a_n + ib_n$, $\rho_{1,n} = a_n - ib_n$, complex conjugate roots,

$$\exp(tL_n) = e^{a_n t} \Big( \cos(b_n t) I + \frac{1}{b_n} \sin(b_n t)(L_n - a_n I) \Big),$$

Case A corresponds to a saddle point solution to the linearized evolution equation (14). Clearly, the zero function $(0,0)$ is a solution of (14). For some spatio-temporal perturbations around this constant function we may expect a stable manifold like behaviour, which will allow for the controlled nature of the system. Working in terms of the Fourier expansions of these perturbations, $z_n(0) = (k_n(0), c_n(0))^T$ we see that as long as $k_n(0)$ and $c_n(0)$ are appropriately chosen then $z_n(t) = \exp(tL_n) z_n(0)$ will only contain the contribution of one of the two exponential factors. In the controlled system choosing $z_n(0)$ so that $(L_n - \rho_{1,n} I) z_n(0) = 0$, we can eliminate the contribution of the second exponential, and as long as $\rho_{1,n} < \frac{r}{2}$, this is an acceptable solution for the controlled system. The condition $(L_n - \rho_{1,n} I) z_n(0) = 0$ can be translated to conditions between $k_n(0)$ and $c_n(0)$ of the form $c_n(0) = Q_n k_n(0)$ for some suitable $Q_n \in \mathbb{R}$. Geometrically, this can be interpreted as the projection of the graph of the stable manifold for the saddle point on the subspace generated by the eigenfunctions of the Laplacian $\phi_n$. Functions of the form $(k(0,x), c(0,x))$ such that $k(0,x) = \sum_n k_n(0)\phi_n(x)$ and $c(0,x) = \sum_n Q_n k_n(0)\phi_n(x)$, with $Q_n$ defined as above, belong to the stable manifold with the above expressions providing a local parameterization of the (infinite) dimensional stable manifold (or rather its segments sufficiently close to the saddle point). The stable manifold clearly carries some spatial structure. Starting on an initial perturbation $(k(0,x), c(0,x))$ on the stable manifold, this initial condition will evolve as $z(t,x) = (k(t,x), c(t,x))^T$ with $z(t,x) = \sum_n \frac{1}{\rho_{1,n} - \rho_{2,n}} e^{\rho_{1,n} t}(L_n - \rho_{2,n} I) z_n(0)\phi_n(x)$ (since on the stable manifold we have that $(L_n - \rho_{2,n} I) z_n(0) = 0$ for every $n \in \mathbb{N}$). If $\rho_{1,n} < 0$ for all $n \in \mathbb{N}$ such spatio-temporal solutions starting on the linearized stable manifold, eventhough starting with a spatial structure asymptotically in time this spatial structure will be eliminated yielding a flatbehaviour. If, for some $n \in \mathbb{N}$, $\rho_{1,n} \in (0, \frac{r}{2})$ then such modes will not decay in time but yield a spatial structure. Clearly, linearized theory may not be provide the full picture, however, one may anticipate that such structures may be precursors to the generation of fully nonlinear patterns. This is beyond the scope of the present work.

One clearly sees by the above formulae that the asymptotic behaviour is dominated by the eigenvalues of $L_n$, leading to growth and sustainability of the perturbation if the eigenvalues have positive real part. The eigenvalues of the matrix $L_n$ can readily be calculated in terms of the roots of the quadratic equation

$$\rho^2 - r\rho + \left\{ D(k^*)\mu_n(r - D(k^*)\mu_n) - \frac{U'(c^*)}{U''(c^*)} f''(k^*) \right\} = 0,$$

which can easily be computed as

$$\rho_{j,n} = \frac{r}{2} + \frac{(-1)^j}{2}\sqrt{r^2 - 4\left[D(k^*)\mu_n(r - D(k^*)\mu_n) - \frac{U'(c^*)}{U''(c^*)} f''(k^*)\right]}, \;\; j = 1, 2.$$

To simplify the notation in this appendix we will denote $D(k^*)$ by $D$ and set $M = \frac{U'^*)}{U''(c^*)} f''(k^*) > 0$, by the properties of the utility function and the production function. Note also that since we would like the flat fixed point to be a saddle point it must hold that $M < \frac{r^2}{4}$.

Using the labeling above, we will be in Case A (i.e. have two real roots) as long as

$$D\mu_n(r - D\mu_n) < \frac{r^2}{4} + M.$$

To simplify notation further let us denote $D_n = D\mu_n$ and we express the condition for two real roots as $-D_n^2 + rD_n - (\frac{r^2}{4} + M) < 0$, or equivalently $-(D_n - \frac{r}{2})^2 - M < 0$ which is always true. Therefore, for any choice of $D_n$ we are always in Case A, where we have two real roots - and a saddle point. It is easily seen that $\rho_{1,n} < r/2 < \rho_{2,n}$, so that the part of the solution along the eigenvector corresponding to the second eigenvalue will be suppressed by the dynamics of the optimal control system as incompatible with the transversality condition. On the other hand, however, if $0 < \rho_{1,n}$ there will be an instability which will lead to the growth of a spatial pattern for the optimal system compatible with the tranversality condition. By algebraic manipulation we see that $0 < \rho_{1,n}$ as long as $D\mu_n(r - D\mu_n) - M > 0$. We therefore see that we will have two positive roots, with the acceptable root leading to instability if the parameters of the problem are such that

$$M < D\mu_n(r - D\mu_n).$$

Since $\mu_n \to \infty$ as $n \to \infty$ it can easily be seen that for large values of $n$ it will hold that $D\mu_n(r-D\mu_n) < 0$ hence condition (**??**) will not hold leading to eigenvalues $\rho_{1,n} < 0$, which correspond to stability of the flat steady state. Therefore, for given values of $D$ and $M$ only a finite number of modes can lead to instability and these modes will be in the set $\mathcal{N} := \{n \in \mathbb{N} \; : \; 0 < D\mu_n(r - D\mu_n)M\}$.

We now pose the following question: Suppose that we are interested in a particular unstable pattern corresponding to a spatial mode $\phi_n$ as presrcibed by the relevant eigenfunction of the Laplace operator. What will be the values of $D$ needed to support this mode? Clearly from (**??**) we see that for given $n$, hence given $\mu_n$, if $D$ is too large that (**??**) fails. Therefore, there must be a range of values of $D$ such that (**??**) holds. We express this condition as a quadratic polynomial in $D$, in the form

$$0 < -\mu_n^2 D^2 + r\mu_n DM.$$

This can never hold if the above quadratic polynomial admits no real roots, i.e., as long as $\frac{r^2}{4} < M$.

Therefore, a spatial pattern will not appear for small enough values of the discount factor $r$.

# F Proof of Propositions 4.5 and 4.7

We begin by expressing the state equation in the equivalent (but more convenient for our analysis) form

$$\frac{\partial k}{\partial t} = -\nabla \cdot [\bar{D}_0 B\psi(k) k \nabla(Af'(k))] + Af(k) - \delta k - c. \tag{34}$$

We restate the Pontryagin principle for this equivalent form of the equation using Remark 3.5 leading to an equation for the costate variable as

$$\begin{aligned}\frac{\partial p}{\partial t} = \bar{D}_0 Af''(k^*)\nabla \cdot [B\psi(k^*)k^*\nabla p] - \bar{D}_0 B\frac{\partial}{\partial k}(\psi(k^*)k^*)\nabla(Af'(k^*)) \cdot \nabla p + \\ (\delta + r - Af'(k^*))p.\end{aligned} \tag{35}$$

The solution of the Pontryagin system in the case where $\bar{D}_0 = 0$ (no capital mobility) will be denoted by $k_0$ and its is easily seen to correspond to $Af'(k_0) = \delta + r$, $c_0 = Af(k_0) - \delta k_0)$, $U'(c_0) = p_0$. We will consider an expansion of the solution of the Pontryagin system in the case of $\bar{D}_0 \neq 0$ (small) as $k = k_0 + \bar{D}_0 k_1 + \cdots$ and similarly for $p$ . By substitution in the Pontryagin system and keeping only first order terms in $D_0$ we obtain

$$\begin{aligned}0 &= -\nabla \cdot [B\psi(k_0)k_0\nabla(Af'(k_0))] + [Af'(k_0) - \delta]k_1 - c_1 \\ 0 &= Af''(k_0)\nabla \cdot (B\psi(k_0)k_0\nabla p_0) - B\frac{\partial}{\partial k}(\psi(k_0)k_0) \cdot \nabla p_0 + \\ &\qquad (\delta + r - Af'(k_0))p_1 - Af''(k_0)p_0 k_1.\end{aligned}$$

Since $Af'(k_0) = \delta + r$, if there is no spatial variation of $\delta$, then the above system simplifies to

$$\begin{aligned}0 &= rk_1 - c_1 \\ 0 &= Af''(k_0)\nabla \cdot (B\psi(k_0)k_0\nabla p_0) - Af''(k_0)p_0 k_1\end{aligned}$$

which provides an explicit form for $k_1$ and $c_1$ as

$$\begin{aligned}k_1 &= \frac{1}{p_0}\nabla \cdot (B\psi(k_0)k_0\nabla p_0), \\ c_1 &= rk_1.\end{aligned}$$

Up to the first order in $\bar{D}_0$ we have

$$\int_{\mathcal{D}} |\nabla k|^2 dx = \int_{\mathcal{D}} |\nabla k_0|^2 dx + 2\bar{D}_0 \int_{\mathcal{D}} \nabla k_0 \cdot \nabla k_1 dx,$$

hence

$$I - I_0 = 2\bar{D}_0 \int_{\mathcal{D}} \nabla k_0 \cdot \nabla k_1 dx,$$

and the sign of $I - I_0$ depends on the sign of $\int_{\mathcal{D}} \nabla k_0 \cdot \nabla k_1 dx$. If this quantity is positive then capital mobility enhances spatial inhomogeneity of capital whereas if it is negative it has the opposite effect.

We now compute this quantity in the case where $A(x) = C\exp(-\frac{|x-x_0|^2}{2\sigma^2})$, and for the case where the production function is of the Cobb-Douglas type $f(k) = Ak^{\alpha}$, $\psi(k) = k^{\rho}$, and $U(c) = \frac{1}{1-\lambda}c^{1-\lambda}$. A straight forward calculation yields,

$$k_0 = M_0 C^{1/(1-\alpha)} \exp(-\frac{|x-x_0|^2}{2(1-\alpha)\sigma^2}),$$
$$p_0 = N_0 C^{-\lambda/(1-\alpha)} \exp(\lambda\frac{|x-x_0|^2}{2(1-\alpha)\sigma^2}),$$

with

$$M_0 = \left(\frac{r+\delta}{\alpha}\right)^{-1/(1-\alpha)}, \quad N_0 = \left(\frac{(1-\alpha)\delta + r}{\alpha}\right)^{-\lambda} M_0^{-\lambda}.$$

After some straighforward calculations we obtain that

$$k_1 = \Lambda_0 \Big[ -\frac{\rho+1-\lambda}{(1-\alpha)\sigma^2}|x-x_0|^2 + d\Big] \exp(-(\rho+1)\frac{|x-x_0|^2}{2(1-\alpha)\sigma^2}),$$

where

$$\Lambda_0 = M_0^{\rho+1} C^{(\rho+1)/(1-\alpha)} \frac{\lambda}{1-\alpha}\sigma.$$

We now calculate $\nabla k_0$ and $\nabla k_1$ explicitly from the above expressions to get that

$$\nabla k_0 \cdot \nabla k_1 = \Xi_0(\rho+1-\lambda)\Big[|x-x_0|^4 + 2 - \frac{(1-\alpha)\sigma d}{\rho+1-\lambda}|x-x_0|^2\Big] \exp(-\rho\frac{|x-x_0|^2}{2(1-\alpha)\sigma^2})$$

We can now calculate the quantity $\int_{\mathcal{D}} \nabla k_0 \cdot \nabla k_1 dx$ in terms Gaussian integrals as long as $diam(\mathcal{D}) >> \sigma^2$. We can estimate

$$\int_{\mathcal{D}} |x-x_0|^2 \exp\Big(-\frac{1}{2\bar{\sigma}^2}|x-x_0|^2\Big)dx \simeq (2\pi\bar{\sigma}^2)^{d/2} d\bar{\sigma}^2,$$
$$\int_{\mathcal{D}} |x-x_0|^4 \exp\Big(-\frac{1}{2\bar{\sigma}^2}|x-x_0|^2\Big)dx \simeq (2\pi\bar{\sigma}^2)^{d/2} \nu(d)\bar{\sigma}^4,$$

for $\bar{\sigma}^2 = \frac{(1-\alpha)\sigma^2}{\rho}$, and $\nu(1) = 3$, $\nu(2) = 8$. Using the above formulae we obtain

$$\Delta I := \int_{\mathcal{D}} \nabla k_0 \cdot \nabla k_1 dx \simeq \Xi_0'(\rho+1-\lambda)\Big[\frac{(1-\alpha)\sigma^2}{\rho}\nu(d) + 2 - \frac{(1-\alpha)\sigma d}{\rho+1-\lambda}\Big]$$

where

$$\Xi_0' = \Xi_0 \frac{1}{(1-\alpha)\rho\sigma}(2\pi\bar{\sigma}^2)^{d/2} > 0.$$

Setting $z := \rho + 1 - \lambda$ we see that the sign of $\Delta I$ depends on the sign of

$$z\left(\frac{(1-\alpha)\sigma^2}{\rho}\nu(d) + 2 - \frac{(1-\alpha)\sigma d}{z}\right) = \left(\frac{(1-\alpha)\sigma^2}{\rho}\nu(d) + 2\right)z - (1-\alpha)\sigma d,$$

hence $\Delta I \leq 0$ as long as

$$z = \rho + 1 - \lambda \leq \frac{(1-\alpha)\sigma d}{\frac{(1-\alpha)\sigma^2}{\rho}\nu(d) + 2},$$

from which the stated result follows.

## References


Acemoglu, D. (2009). *Introduction to Modern Economic Growth.* Princeton University Press, Cambridge, MA.

Aghion, P. and P. Howitt (1998). *Engogenous Growth Theory.* The MIT Press, Cambridge, Massachussets.

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano, and F. Robert-Nicoud (2003). *Economic Geography and Public Policy.* Princeton University Press, Princeton, NJ.

Baldwin, R. E. and P. Krugman (2004). Agglomeration, integration and tax harmonisation. *European Economic Review* 48 (1), 1–23.

Baldwin, R. E., P. Martin, and G. I. Ottaviano (2001). Global income divergence, trade, and industrialization: The geography of growth take-offs. *Journal of Economic Growth* 6 (1), 5–37.

Boucekkine, R., C. Camacho, and G. Fabbri (2013a). On the optimal control of some parabolic partial differential equations arising in economics. *Serdica Mathematical Journal* 39, 331–354.

Boucekkine, R., C. Camacho, and G. Fabbri (2013b). Spatial dynamics and convergence: The spatial AK model. *Journal of Economic Theory* 148 (6), 2719–2736.

Boucekkine, R., C. Camacho, and B. Zou (2009). Bridging the gap between growth theory and the new economic geography: The spatial Ramsey model. *Macroeconomic Dynamics* 13 (1), 20–45.

Boucekkine, R., G. Fabbri, S. Federico, et al. (2016). Optimal economic growth through capital accumulation in a spatially heterogeneous environment. Technical report, Aix-Marseille School of Economics, Marseille, France.

Boucekkine, R., Fabbri,G., Federico,S. and Gozzi,F. (2018), Growth and agglomeration in the heterogeneous space: ageneralized AK approach", *Journal of Economic Geography*, (0),1-32.

Breinlich, H., G. I. Ottaviano, and J. R. Temple (2014). Regional growth and regional decline. In *Handbook of Economic Growth*, Volume 2, pp. 683–779. Elsevier, Amsterdam.

Brito, P. (2004). The dynamics of growth and distribution in a spatially heterogeneous world. Working Papers, Department of Economics 2004/14, ISEG – Lisbon School of Economics and Management, Department of Economics, Universidade de Lisboa.

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 (9), 2745–2787.

Brock, W. and A. Xepapadeas (2010). Pattern formation, spatial externalities and regulation in coupled economic–ecological systems. *Journal of Environmental Economics and Management* 59 (2), 149–164.

Brock, W. A., A. Xepapadeas, and A. N. Yannacopoulos (2014a). Optimal agglomerations in dynamic economics. *Journal of Mathematical Economics* 53, 1–15.

Brock, W. A., A. Xepapadeas, and A. N. Yannacopoulos (2014b). Spatial externalities and agglomeration in a competitive industry. J*ournal of Economic Dynamics and Control* 42, 143–174.

Brock, W. A., A. Xepapadeas, and A. N. Yannacopoulos (2014c). Optimal control in space and time and the management of Environmental Resources, *Annual Review of Resource Economics* 6, 33–68

Camacho, C. and B. Zou (2004). The spatial Solow model. *Economic Bulletin* 18 (2), 1–11.

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

Desmet, K. and E. Rossi-Hansberg (2010). On spatial dynamics. *Journal of Regional Science* 50 (1), 43–63.

Fabbri, G. (2016). Geographical structure and convergence: A note on geometry in spatial growth models. *Journal of Economic Theory* 162, 114–136.

Fujita, M., P. R. Krugman, and A. J. Venables (2001). *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Cambridge, MA.

Fujita, M. and T. Mori (2005). Frontiers of the new economic geography. *Papers in Regional Science* 84 (3), 377–405.

Isard, W. and P. Liossatos (1979). *Spatial Dynamics and Optimal Space-Time Development*. Elsevier, Amsterdam.

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

Krugman, P. (1998). Space: the final frontier. *Journal of Economic Perspectives* 12 (2), 161–174.

Lucas, R. E. (1990). Why doesn't capital flow from rich to poor countries? *American Economic Review* 80 (2), 92–96.

Lucas, R. E. (2003). Macroeconomic priorities. *American Economic Review* 93 (1), 1–14.

Mankiw N.G., D. Romer and D.N. Weil (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics* 107(2), 407–437.

Magill, M, (1977), Some new results on the local stability of the process of capital accumulation. *Journal of Economic Theory* 15, 174-210.

Martin, P. and G. I. Ottaviano (2001). Growth and agglomeration. *International Economic Review* 42 (4), 947–968.

Nonneman, W. and P. Vanhoudt (1996). A further augmentation of the Solow model and the empirics of economic growth for OECD countries. *The Quarterly Journal of Economics* 111(3), 943–953.

Quah D.T. (1996). Twin peaks: Growth and convergence in models of distribution dynamics. *Economic Journal* 106 (4370), 1045–1055.

Prasad, E., R. Rajan, and A. Subramanian (2007). The paradox of capital. *Finance and Development* 44 (1), 16–19.

Romer, P.M. (1986). Increasing returns and long-run growth. *Journal of Political Economy* 94(5) 1002–10037.

Turing, A. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London Series B* 237, 37–72.

Xepapadeas, A. and A. Yannacopoulos (2016). Spatial growth with exogenous saving rates. *Journal of Mathematical Economics* 67, 125–137.