Spatial growth theory: Optimality and spatial heterogeneity

*Published Version:*
Xepapadeas, A., Yannacopoulos, A.N. (2023). Spatial growth theory: Optimality and spatial heterogeneity. JOURNAL OF ECONOMIC DYNAMICS & CONTROL, 146, 1-35 [10.1016/j.jedc.2022.104584].

*Availability:*
This version is available at: https://hdl.handle.net/11585/950898 since: 2024-02-27

*Published:*

DOI: http://doi.org/10.1016/j.jedc.2022.104584

**Xepapadeas, A., & Yannacopoulos, A. N. (2023). Spatial growth theory: Optimality and spatial heterogeneity.** ***Journal of Economic Dynamics and Control*****,** ***146*****, 104584.**

https://doi.org/10.1016/j.jedc.2022.104584

# 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,
xepapad@aueb.gr
Corresponding author:
Athens University of Economics and Business,
76 Patission Str, 10434, Athens Greece.
[2] Athens University of Economics and Business, ayannaco@aueb.gr

**Abstract**

Spatiotemporal dynamics are introduced in a standard Ramsey model of optimal growth in which capital moves toward locations where the marginal productivity of capital is relatively higher and initiate a study of the effects of nonlinear capital transport in terms of a linearized analysis around steady state solution solutions.

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 a high utility discount rate, the spatial capital flows induce the emergence of optimal spatial patterns, while for a 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, in the context of both traditional and new growth theory, has been studied in a temporal domain (e.g., Aghion and Howitt, 1998; Barro and Sala-i-Martin, 2004; Acemoglu, 2009). The explanation of the temporal evolution of key variables – such as output or capital per capita, the capital-output or the capital-labor 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, chapter 1) points out the great inequality in income per capita and income per worker across countries, and the increase in this inequality across nations 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 11 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 to incorporating space into models of optimal growth.[1]

Spatiotemporal models of economic growth started appearing in the 2000s.[2] Earlier research which has provided the main mechanism for capital flows across space can be found in Isard and Liossatos (1979). A central reason for introducing spatial aspects in growth could be the question posed by Quah (1996, p. 1053): "What one wants to know here is, what happens to the entire cross sectional distribution 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 paper explores spatial growth by developing a spatial Ramsey optimal growth model in a two-dimensional spatial domain. In this domain, capital located at 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 (MPD) capital flows was first introduced by Xepapadeas and Yannacopoulos (2016) and differs from the assumption underlying capital flows which is used in much of the recent literature on spatial growth (e.g. Brito, 2004; Boucekkine et al., 2013b; Fabbri, 2016). 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 linear local diffusion with a constant diffusion coefficient which implies that capital moves from locations of high concentration to locations of low concentration. The MPD assumption about capital flows adopted in this paper seems to be plausible but it leads, however, to a model of nonlinear local diffusion. This introduces new challenges in the solution of the optimization problem, which is required in order to study optimal growth in a spatiotemporal domain.

In this context, the present paper contributes to growth theory by: (i) 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 toward locations of high marginal productivity; (ii) developing a maximum principle to the solution of dynamic optimization problems, where capital accumulation is described by a partial differential equation with nonlinear diffusion; and (iii) providing results on optimal spatiotemporal dynamics, especially regarding the endogenous emergence of spatial patterns both analytically and numerically. To the best of our knowledge the solution of the standard Ramsey problem with capital accumulation governed by nonlinear diffusion – induced by MPD flows – and the explicit identification of conditions for the endogenous emergence of optimal spatiotemporal patterns from a spatially homogenous state is new.[3]

By solving the spatiotemporal Ramsey problem we try to discuss issues which could emerge when the spatial dimension is combined with MPD capital flows across locations in the context of a social planner's model with spatial externalities. More specifically, we seek to explore two questions.

---

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

[2] See, for example, Brito (2004), Camacho and Zou (2004), Boucekkine et al. (2009, 2013a, 2013b, 2016, 2019), Brock et al. (2014a, 2014b), Fabbri (2016), Xepapadeas and Yannacopoulos (2016).

[3] This is relative to the existing literature cited in this paper, which considers various combinations of $AK$, production functions, linear diffusion of fixed savings without optimization.

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 the terminology of Krugman (1998) and Fujita et al. (2001) – steady state in which per capita output and capital is the same across all locations. Is it possible, when MPD 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 will the perturbation die out with the 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 MPD capital flows will eventually drive the economies to a spatially-heterogeneous 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-heterogeneous steady state relative to the case of no MPD capital flows?

The first questions explores whether or not optimal growth when capital is seeking locations of high marginal productivity promotes spatial inequalities. The second explores whether, in a world with TFP differences across locations, which can be connected with nature-first arguments, optimal growth with capital seeking locations of high marginal productivity intensifies or weakens spatial inequalities.

Our main results, in relation to the questions posed above, are 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 a homogenizing factor and reduce inequalities after a spatial perturbation of the capital per capita 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. These issues have been studied in the literature (e.g., Boucekkine et al. 2013b, 2016, 2019, Xepapadeas and Yannacopoulos 2016). However, we believe that by replacing the Fickian assumption about capital diffusion with the MPD assumption and by developing the appropriate maximum principle, well-founded insights regarding the endogenous emergence and the strengthening or weakening of spatial inequalities by MPD capital flows can be provided.

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 defines the spatial Ramsey model. Section 3 extends 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 MPD spatial capital flows

Xepapadeas and Yannacopoulos (2016) extended the Solow model to include MPD spatial transport of capital. Here we provide a brief review of the derivation of the fundamental capital accumulation equation, before presenting the related optimal control problem. Consider a spatial domain $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2,$[4] and let $k(t, x)$ be capital, or the density of capital, at time $t$ and at the spatial location $x \in \mathcal{D}$.[5] 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 $\mathcal{D}$.

Before proceeding with the model, we recall the following facts and notations: 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 $(\nabla \cdot J)$. The dot product notation following $\nabla$ in the notation for the divergence implies that the divergence is to be considered as the dot product of the gradient operator $\nabla$ with the vector field $J = (J_1, J_2)$. Another standard notation we will use in this paper is the notation $\Delta$ for the Laplace operator defined by its action on a scalar field $m$ as $\Delta := \nabla \cdot \nabla\, m = div(\nabla m)$. These operators can be defined in any coordinate system. In Cartesian coordinates in $\mathbb{R}$ with the standard basis $e_1, e_2$, and for a scalar field $\phi$ and a vector field $\Phi = \Phi_1 e_1 + \Phi_2 e_2$, the gradient, the divergence and the Laplacian admit the representations

$$\nabla\phi = \frac{\partial\phi}{\partial x_1}e_1 + \frac{\partial\phi}{\partial x_2}e_2, \;\; \nabla\cdot\Phi = \frac{\partial\Phi_1}{\partial x_1} + \frac{\partial\Phi_2}{\partial x_2}, \;\; \Delta\phi = \frac{\partial^2\phi}{\partial x_1^2} + \frac{\partial^2\phi}{\partial x_2^2},$$

respectively. This representation changes if a different coordinate system is adopted.

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 local 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 at $(t, x)$ 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 in the direction in which 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

---

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

[5] To simplify, we assume that labor at each location is fixed and immobile across locations. Thus $k(t, x)$ 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.

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 $U$.

A standard bookkeeping argument, in conjunction with the above discussion, leads to the result 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 three 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 locations other than $x$. The transport equation (1) must be complemented with a boundary condition on $\partial\mathcal{D}$, the boundary of the domain $\mathcal{D}$. A natural condition is the no flux boundary condition

$$n \cdot J(t,x) = 0, \quad \forall x \in \partial\mathcal{D}, \tag{2}$$

which models the situation where no capital can flow outside $\mathcal{D}$. In the absence of capital production, depreciation and consumption the dynamics of (1) supplemented with this boundary condition show the conservation of total capital in the region over time. While, other boundary conditions can also be chosen, as for example Dirichlet boundary conditions, Robin boundary conditions, periodic boundary conditions or mixed conditions, in this paper we will focus on the choice (2) which under our choice for the flux $J$ will be equivalent to setting Neumann boundary conditions.

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 $\nabla m$.

Our basic modelling 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 toward locations of higher marginal productivity of capital, but also that capital at different locations may have different propensity to relocate toward 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 toward higher returns than smaller capital concentrations. These considerations are expressed by defining the flux as:

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

where

- $\bar{D}_0$ is a constant capital transport parameter,
- $B(x)$ is a function reflecting the effect of regulations or geographical or commercial factors affecting the propensity of capital to relocate,
- $\psi(k)$ models the fact that different capital densities could exhibit different propensity toward relocation, and
- $\nabla m(t,x)$ is the gradient of capital productivity, defined in terms of the production function $F$ as $m = \frac{\partial}{\partial k}F(x,k)$, pointing towards the direction where $m$ increases

It is our convention to take $J$ explicitly proportional to $k$ to stress the fact that $B\psi$ represents propensity, i.e., it can be interpreted as probability to relocate, under this convention capital cannot relocate from regions where it has become extinct.

Assuming a production function of the form $F(x,k) = A(x)f(k)$, where $A\left(x\right)$ can be interpreted as local TFP,[6] then, denoting by $'$ the derivative of a function with respect to its argument,

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

and the flux $J$ can be expressed as

$$\begin{aligned} J(t,x) = {} & \bar{D}_0 A(x)B(x)\psi(k(t,x))f''(k(t,x))k(t,x)\nabla k(t,x) + \\ & \bar{D}_0 B(x)\psi(k(t,x)f'(k(t,x))k(t,x)\nabla A(x). \end{aligned} \tag{3}$$

This flux shows two factors affecting capital mobility, (a) spatial variability of capital (the first contribution) and (b) spatial variability of the TFP factor $A$, as both these terms may contribute to the spatial variability of the capital productivity $m$. Moreover, it is important to note that even in the presence of all these factors promoting capital mobility, no capital transport will occur if $B(x) = 0$, i.e., if there is no propensity for the capital to move.

---

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

Hereafter, to make notation easier we may omit the explicit dependence on $(t, x)$, keeping in mind that $A$ and $B$ could be functions of $x$, where $f$ is a function of $k$, while the composite function $f(k(t, x))$ will depend on $(t, x)$ through the dependence of $k$ on $(t, x)$. Recalling that by $'$ we denote the derivative of a function with respect to its argument, hence $f'(k(t, x))$ denotes the derivative of the function $f : \mathbb{R} \to \mathbb{R}$ with respect to its argument, composed with the capital density. Note that under this notation the capital flux assumes the form

$$J = \bar{D}_0 AB\psi(k)\Big(kf''(k)\nabla k + kf'(k)\nabla \ln A\Big). \tag{4}$$

Upon defining

$$D_0 = \bar{D}_0 AB, \;\; w_1(k) = -\psi(k)f''(k)k, \;\; w_2(k) = \psi(k)f'(k)k, \tag{5}$$

(note that by the properties of $f$ both functions $w_1$ and $w_2$ are positive) the flux can be expressed in the more compact form.

$$J = D_0(-w_1(k)\nabla k + w_2(k)\nabla \ln A) \tag{6}$$

Combining the transport equation (1) with the capital mobility law (6) we obtain the partial differential equation

$$\begin{aligned} \frac{\partial k}{\partial t} = \nabla \cdot \Big(D_0(w_1(k)\nabla k - w_2(k)\nabla \ln A)\Big) + Af(k) - \delta k - c \\ n \cdot J = 0, \;\; \text{on } \partial\mathcal{D}, \end{aligned} \tag{7}$$

supplemented with an appropriate initial condition $k_0$.

We make one further simplifying assumption,

**Assumption 2.1.** The TFP factor $A$ satisfies

$$n \cdot \nabla \ln A = 0, \;\; \text{on } \partial\mathcal{D}. \tag{8}$$

This assumption is natural for various reasons. Obviously it holds if the TFP factor $A$ does not depend on location. However, even if $A$ depends on $x$ in $\mathcal{D}$ we may naturally assume that we choose $\mathcal{D}$ large enough so that a subset $\mathcal{D}_0 \subset \mathcal{D}$, contains all the important spatial variability of $A$, so that $A$ can be assumed to be more or less constant in the complement $\mathcal{D} \setminus \mathcal{D}_0$. If the variability of $A$ takes place in a compact domain $\mathcal{D}_0$, we may choose $\mathcal{D}$ so that (8) is satisfied. If not, but $A$ has slow variability in space so that $\frac{1}{A}\nabla A \simeq 0$ then we may still assume that (8) is approximately true. For the above reasons, throughout this paper we will adopt assumption (8), which although allows for TFP spatial variability, it permits a simplification of the model in terms of the boundary condition. In particular, if (8) holds then (4) implies that

$$n \cdot J = \bar{D}_0 A\psi(k)(kf''(k)n \cdot \nabla k + kf'(k)n \cdot \nabla \ln A) = \bar{D}_0 A\psi(k)kf''(k)n \cdot \nabla k, \;\; \text{on } \partial\mathcal{D}, \tag{9}$$

so that, (as long as $A\psi(k)kf''(k)$ is finite on the boundary), if $n \cdot \nabla k = 0$ on the boundary then $n \cdot J = 0$ as well. For the above reason, we will replace the natural boundary condition (2) by the simpler homogeneous Neumann boundary condition.

We thus arrive at the following model for capital accumulation: Under the assumption that spatial TFP satisfies the boundary condition (8) the spatiotemporal evolution of the distribution of the capital stock across the spatial domain is the solution of the quasilinear parabolic equation

$$\begin{aligned} \frac{\partial k}{\partial t} = \nabla \cdot \Big( D_0 \big( w_1(k) \nabla k - w_2(k) \nabla \ln A \big) \Big) + A f(k) - \delta k - c \\ n \cdot \nabla k = 0, \quad \text{on } \partial \mathcal{D}, \end{aligned} \tag{10}$$

supplemented with an appropriate initial condition $k_0$.

This is a nonlinear diffusion equation in divergence form where capital transport is driven by two factors: the spatial variation of $k$ and the spatial variation of the local TFP, $A$, both of which induce the spatial variation of $m$ which is assumed to be the driving force behind capital transport. In fact if TFP does not display any spatial variation, equation (10) is in the form of a standard diffusion equation but with a diffusion coefficient $D(k) = D_0 w_1(k)$ which is a nonlinear function of capital, and displays varying diffusivity depending on the local capital concentration $k$. For example, there may be regions of space where because of the local capital concentration, $D(k)$ can be rather low leading to low mobility of capital hence accumulation or the opposite. The variability of the TFP factor $A$, leads to an additional term

$$-\nabla \cdot (D_0 w_2(k) \nabla \ln A) = -D_0 w_2'(k) \nabla \ln A \cdot \nabla k - \nabla \cdot (D_0 \nabla \ln A) w_2(k)$$

which is a contribution of two terms. The first one models streaming of capital with velocity vector driven by the TFP variability factor $\nabla \ln A$ , while the second acts as an adjustment to the production function.

The case where capital is immobile corresponds to setting $D_0 = 0$. Note that in this model, since $D_0 = \bar{D}_0 A B$, capital immobility may be realized by setting the capital propensity to migrate $\bar{D}_0 B = 0$, without having to assume that the TFP factor $A$ vanishes. In the case where $\bar{D}_0 B = 0$ model (10) reduces to the standard spatially homogeneous model

$$\frac{dk}{dt} = A f(k) - \delta k - c. \tag{11}$$

Note that $A$ can be assumed to be spatially dependent in this limit, while still keeping the capital to be immobile, if the propensity of capital to migrate as modelled by $\bar{D}_0 B = 0$ is zero. In fact an interesting limit that can be studied with the above model is the possible spatial patterns that may occur if $A$ is spatially varying, thus turning (11) to an ordinary differential equation in time whose solutions vary parametrically with $x$ on account of the spatial variability of the TFP factor $A$. An interesting question is whether allowing for small capital mobility $B$, will allow for transport of the localized structures generated by (11) (with $A = A(x)$) so as to generate new spatial structures and if so will transport contribute to smoothing of the original structures or have the opposite effect?

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. The nonlinear diffusion coefficient represents a difference of this model (10 relative 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. Even though 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 (5) 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 (10).

The nonlinear diffusion features in our model (10) make it a quasilinear parabolic PDE, of the porous medium type, a feature that introduces important mathematical difficulties and the need of introduction of advanced techniques from functional analysis and PDE theory for its detailed study. Concerning well posedness a general existence and uniqueness proof is outside the scope of this work however one may in principle adapt the arguments of Casas and Chrysafinos (2018) (see also references therein) which work in finite time and for Dirichlet boundary conditions to answer questions concerning global existence in time in an appropriate Sobolev space setting while the question of regularity can be treated by extending the methods in Lazyzhenskaia et al (1988). This would require assuming sufficient smoothness on the production function $f$ which would inherit sufficient smoothness to the functions defining the nonlinear diffusion coefficients $w_1$ and $w_2$. It will moreover require smoothness of the domain ($C^{1,1}$ domains would do) as well as smoothness of the TFP factor $A$. For well posedness in the Cobb-Douglas case see Xepapadeas and Yannacopoulos (2016) and references therein.

On the other hand, despite the technical difficulties, nonlinear diffusion introduces new interesting effects in the dynamics which are not present in the linear diffusion case, making it an interesting model to explore further.

## 3 The spatiotemporal planner problem

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 (10).[7] The problem then becomes the optimal control problem

$$\max_{c} \int_0^\infty \int_{\mathcal{D}} e^{-rt} U(c(t,x))dxdt, \text{ subject to } (10), \tag{12}$$

where $U$ is a standard utility function for consumption satisfying Inada conditions, and $r$ is the utility discount rate. Note that if $D_0 = 0$ problem (12) reduces to the standard temporal Ramsey model. Moreover, problem (12) has been studied in the case of linear transport ($D$ constant) by Boucekkine and coauthors (e.g., Boucekkine et al., 2013b) within the framework of the AK model.

**Remark 3.1.** We remark that, even though the constraints $k \geq 0$ and $c \geq 0$, would be natural to impose, here we do not do so, i.e., we do not treat problem (12) as a state constrained control

---

[7] Xepapadeas and Yannacopoulos (2016), assuming that consumption was chosen as a fixed percentage of output, i.e., $c = (1 - s)Af(k)$, used (10) to study a spatial Solow model with nonlinear diffusion.

problem. We rather take the alternative route, often followed in the literature (see e.g. Faggian et al (2021) and references therein) to solve problem (12) as an unconstrained problem and check ex post that the unconstrained optimal solution satisfies the constraints – either in a classical or weak form – so that it is also a solution of the constrained problem. The choice of the utility function (i.e. the Inada conditions) assist in this direction. Moreover, as the focus of the paper is to consider small perturbations of particular solutions which are known to satisfy such conditions, the proposed approach is sufficient for our current purpose. Even though the treatment of state constrained problems is feasible (see e.g. ) this analysis may be quite involved (see e.g. Calvia et al (2021) or Santambrogio et al (2020) so such an approach is not pursued here.

For reasons that will become clear shortly we also consider a related finite horizon optimal control problem of the form

$$\max_c \int_0^T \int_{\mathcal{D}} e^{-rt} U(c(t,x))dxdt + e^{-rT} \int_{\mathcal{D}} \phi(x, k(T,x))dx, \text{ subject to } (10), \tag{13}$$

for a suitable final bequest function $k$ and a suitably long horizon $T$. Model (13 may be more realistic, since economic planning takes place in finite horizon, even though it adds the complication of specifying the final bequest function $\phi$. The choice of $\phi$ can model for example sustainability issues, i.e. how much capital should be left for the generation at $T$ to continue consumption after $T$. The finite horizon spatial Ramsey model (for the case of linear diffusion) was studied by Camacho et al (2008).

The study of problem (12) for the nonlinear case is a complicated mathematical problem with the infinite time horizon adding more complications to an already hard task. One complication is the dynamics of the state equation (10) itself, since for some initial conditions $k_0$ equation (10) may not admit classical solutions and must therefore be considered in a weak sense in a Sobolev space framework. There is work on optimal control of quasilinear parabolic PDEs for finite horizons $T$, in the weak framework, while there is also some work in the case of state constraints which however complicate the situation even further. Such problems have been treated using generalizations of the Pontryagin maximum principle which characterize the optimal control process in terms of an adjoint process $p$, which is shown to satisfy a quasilinear parabolic equation as well, that now satisfies a final condition at $T$, related to a final bequest. To make the arguments more clear in the linear case ($w_1 = 1$, $w_2 = 0$, $D_0 = 1$) it has been shown (see Camacho et al (2008)) that the optimal control can be constructed in terms of the adjoint process $p$ satisfying

$$\begin{aligned} \frac{\partial p}{\partial t} + \Delta p + (Af'(k) - \delta - r) = 0 \\ p(T,x) = \phi'(k(T,x)), \end{aligned} \tag{14}$$

with suitable boundary conditions while the optimal consumption is given by $U'(c(t,x)) = p(t,x)$. Note the difference in the sign of the diffusion operator in (14) which however can be turned into a standard diffusion equation by reversing time. This transformation can be used to show solvability for the coupled system of the forward state equation with the backward equation (14) and complete the optimal control problem.

The infinite horizon problem (12) on the other hand presents another important issue. The adjoint equation in this case does not have a clear final condition, which must now be replaced by a limiting condition at infinity, called the transversality condition. Concerning the choice of the transversality condition there seem to be controversies in the literature, even for the case where there is no space dependence. In this case, the standard approach is to consider the system of ODEs

$$\begin{aligned} \frac{dk}{dt} &= Af(k) - \delta k - (U')^{-1}(p), \\ \frac{dp}{dt} &= (Af'(k) - \delta - r)p \end{aligned} \tag{15}$$

and look for a solution which remains bounded as $t \to \infty$. The obvious choice is to look for a fixed point of this system which (since the choice of the utility function does not allow $p = 0$) is given by the Ramsey golden rule

$$Af'(k) - \delta - r = 0,$$

which when solved specifies the level of $k$ at the fixed point whereas the state equation yields at the fixed point

$$Af(k) - \delta k - (U')^{-1}(p) = 0$$

which (having obtained the level of $k$ by the golden rule) can be solved to yield the level of $p$. Denoting this fixed point by $(k^*, p^*)$ a candidate solution for (14) such that both $k$ and $p$ remain bounded at infinity is a solution such that $(k(t), p(t)) \to (k^*, p^*)$ as $t \to \infty$. There is a wide consensus in economics that this is indeed the required solution, however a concrete mathematical proof requires careful considerations which can be offered using alternative techniques based on dynamic programming techniques and the Hamilton-Jacobi equation (see e.g. Ekeland (2011)). However, there is a long literature on the appropriate choice of asymptotic final condition for the adjoint variable with possible conditions involving $\lim_{t\to\infty} e^{-rt}p(t) = 0$ or $\lim_{t\to\infty} e^{-rt}k^*(t)p(t) \to 0$ or related exponential growth estimates (see e.g. Aseev and Kryazhimskiy (2004) and references therein) with counterexamples being available for some of the commonly used choices in the economic literature. Other authors attempt to deal with this problem choosing an appropriate function space setting for the treatment of the optimal control problem such as for example weighted Sobolev spaces with exponential weights related to the discount function (see e.g. Van, Boucekkine, and Saglam (2007), or Lykina, and Pickenhain (2017)), and the literature on the choice of the transversality conditions is still active with questions still being open when models in infinite dimensional spaces are being considered.

However, what we need to emphasize here is the intuitive approach to the solution of the Ramsey problem using the adjoint variable $p$ in the non spatial case, that characterizes the optimal path as the unique solution of the optimality system that drives the system (asymptotically) to the fixed point $(k^*, p^*)$, which can be shown to be a saddle point with a stable and an unstable manifold, with the optimal path in $(k, p)$ space corresponding to the stable manifold of the saddle point. In fact, the parametrization of the stable manifold $p = \Phi(k)$, in terms of an appropriate function $\Phi$ can be considered as a feedback control policy when combined with the optimality condition $p = U'(c)$, thus allowing for a formal closure of the system and the computation of the optimal capital path. We shall return to this geometric intuition shortly.

The situation becomes even more complicated when optimal control of spatial models is studied in the infinite horizon. While the Pontryagin maximum principle is well understood for finite horizon problems (such as (13) ) both for the linear diffusion case (including economic applications in the Ramsey model, see e.g., Camacho et al (2008)) or for quasilinear problems (see e.g. Casas and Chrysafinos (2018) and references therein) the infinite horizon version still presents open questions. These are related to the appropriate choice of transversality condition as well as the fact that the parabolic nature of the equation may turn its solvability backward in time an ill posed problem leading to problems with continuity with respect to the data of the problem. Such problems are amplified in the infinite horizon case, where the transformation involving time inversion proposed in Boucekkine et al (2009,2013) cannot be applied to facilitate the construction of a solution. In such cases a more abstract approach must be followed, using e.g. a fixed point approach in the proper function space setting accommodating the conditions for the adjoint variable at infinity (see e.g. Peng and Shi (1999) or for economics related problems Kartala et al (2020) or Yannacopoulos (2008) where a stochastic generalization of a saddle path is constructed, and references therein). An alternative is to treat such problems using techniques from dynamic programming and the associated Hamilton-Jacobi equation, an approach which bypasses the need for the introduction of the adjoint variable, however, in this case the Hamilton-Jacobi equation is far more complicated than the one used in the temporal case, as it is now a PDE on an infinite dimensional space whose solution provides the value function for the problem. Such equations in the general case present many mathematical intricacies among which the fact that they often do not admit solutions that are smooth enough so that the feedback control law is expressed as its gradient. However, for certain simple models explicit solutions can be constructed. This approach has been followed in Boucekkine et al (2013) who provided a complete solution of the Hamilton-Jacobi equation for the Ramsey problem for the AK model with linear diffusion in closed form and characterized both the value function for the problem as well as the optimal consumption policy. Their analysis highlights the importance of the choice of transversality condition if the problem is to be addressed in terms of the Pontryagin maximum principle. This is further documented in Ballestra (2016) who addressed the same problem solely in terms of the Pontryagin maximum principle and showed that if choosing the transversality condition as $\lim_{t\to\infty} p(t,x) = 0$ then the maximum principle cannot reproduce the solution obtained in terms of the dynamic programming approach, whereas choosing a different transversality condition does.

The above discussion shows that a full treatment for the spatial Ramsey model in infinite horizon is far from being complete. The intention of this paper is to add to this literature, by exploring possible behaviour of the forward backward PDE obtained by the Pontryagin maximum principle for the case of nonlinear diffusion in the capital evolution law as modelled by (10). While the exact form of the equation can be obtained beyond doubt, the choice of the tranversality condition for the adjoint variable is more tricky and involves delicate asymptotic estimates of quasilinear parabolic equations, which are well beyond the scope of this paper. As the exact solution of the relevant infinite dimensional

Hamilton-Jacobi equation coming from the dynamic programming approach is clearly an impossible task, comparisons such as those provided by Ballestra (2016) are clearly out of the question. As a result, we conjecture a type of transversality condition which is plausible by a formal argument (i.e. an argument where certain assumptions concerning the regularity of the solutions of problem (12) are made) and certainly compatible with the transversality conditions adopted for the AK model by Ballestra, or obtained as approximations of sufficiently long horizon problems of the form (13).

Based on arguments presented in Appendix A we make the following conjecture:

**Conjecture:** *A sufficiently regular optimal path* $(k^*, c^*)$ *for problem* (12) *such that* $c^*(t,x) > 0$ *a.e. can be characterized in terms of the optimality condition*

$$U'(c^*(t,x)) = p(t,x), \tag{16}$$

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

$$\frac{\partial k^*}{\partial t} = \nabla \cdot \Big(D_0 w_1(k^*)\nabla\, k^* - D_0 w_2(k^*)\nabla\, \ln A\Big) + Af(k^*) - \delta k^* - (U')^{-1}(p), \tag{17}$$

$$\frac{\partial p}{\partial t} = -\nabla \cdot \Big(D_0 w_1(k^*)\nabla\, p\Big) + \underline{G}(k^*) \cdot \nabla\, p + (r + \delta - Af'\,(k^*))p, \tag{18}$$

$$n \cdot \nabla k^* = n \cdot \nabla p = 0, \quad on\ \partial\mathcal{D}$$

*where*

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

*The system (17-18) is supplemented with the initial condition* $k(0) = k_0$ *and the transversality condition*

$$\lim_{T\to\infty} e^{-rT} \int_{\mathcal{D}} k^*(T,x)p(T,x)dx = 0. \tag{19}$$

*which requires that the present value of the capital stock is zero for* $t \to \infty$.

The grounds for choosing the above system (17-18) and the transversality condition (19) are presented in Appendix A. As they are based upon assumptions which we have not proved in full generality – which at any rate would not be necessary for our arguments – we prefer to call the above a conjecture based on which we analyse certain special cases of interest. Here we restrict in stating that (19) coincides with the transversality condition used by Ballestra (2016) to recover by the Pontryagin maximum principle the solution to the spatial AK problem with linear diffusion obtained by Boucikkine et al (2013) using the dynamic programming principle. Moreover, in the case where no spatial variation is considered it reduces to the popular transversality condition used in economic control problems motivated by Arrow and Kurz (1970).

In line with the optimality conditions of the traditional Ramsey model, equation (16) states that the marginal utility of consumption should be equal to the shadow value of the capital stock at each point of time and space. The inverse function of $U'$, $c = g\,(p)$ represents – in the terminology of Arrow and

Kurz (1970) – a short-term equilibrium of the system or a "derived demand equation" characterizing the optimal allocation of current output between consumption and savings at each point of time and space. The partial differential equations (17) and (18) describe the optimal spatiotemporal evolution of the stock of capital and its shadow value, when in the short-term output is allocated optimally between consumption and savings. Finally, the spatial boundary condition states that along the optimal path the flux of capital stock and its shadow value is zero at the boundary during the entire time horizon, while the temporal transversality condition states that at infinity the present value of the value of capital stock tends to zero for the entire spatial domain.

**Remark 3.2.** The following remarks are in order:

1. In the case where $A$ is constant and no capital mobility ($B = 0$) system (17-18) and the transversality condition (19) reduces to the standard Ramsey model optimality condition (Benveniste and Scheinkman (1982)). By requiring, from (19), that the present value of the capital stock is zero for $t \to \infty$ allows us to consider eigenvalues which are negative or positive but below r/2 in the stability analysis of the linearized system.

   On the other hand, in the case where $w_1 = 1$, $w_2 = 0$ and $D_0 = 1$ system (17-18) and the transversality condition (19) reduces to the optimality condition for the spatial AK model used by Ballestra (2016) to reproduce the results of Boucekkine et al (2013) who used the dynamic programming approach and the Hamilton-Jacobi equation. Hence, the proposed system is compatible in two special cases with the corresponding systems whose transversality conditions are known to reproduce in terms of the Pontryagin approach the results of the dynamic programming approach.

2. The solvability of system (17-18) in finite time horizon can be treated by inverting the time in the adjoint equation and extending the techniques in Casas and Chrysafinos (2018) or Boucekkine et al (2013). Concerning the general solvability of system (17-18) with the transversality condition (19), this is not an easy task and it is well beyond the scope of this paper, whose main focus is to initiate a discussion concerning the effects of nonlinear diffusion in economic growth, by exploring the effect of plausible transversality condition in certain special cases While the general solvability is under current investigation, we point out here that system (17-18) with the transversality condition (19) admits solutions in at least two special cases:

   (a) when $A$ is a constant it admits a spatially independent solution $(k(t,x), p(t,x)) = (k(t), p(t))$, which coincides with the solution of the standard Ramsey problem and

   (b) it admits a steady state solution $(k(t,x), p(t,x)) = (k(x), p(x))$, as long as the elliptic quasilinear equation

$$\begin{aligned} 0 &= \nabla \cdot \Big( D_0 w_1(k^*) \nabla k^* - D_0 w_2(k^*) \nabla \ln A \Big) + A f(k^*) - \delta k^* - (U')^{-1}(p), \\ 0 &= -\nabla \cdot \Big( D_0 w_1(k^*) \nabla p \Big) + \underline{G}(k^*) \cdot \nabla p + (r + \delta - A f'(k^*)) p, \qquad (20) \\ & \quad n \cdot \nabla k^* = n \cdot \nabla p = 0, \ \text{ on } \partial \mathcal{D}. \end{aligned}$$

has a solution. For $A$ constant and immobile capital ($B = 0$) the above system reduces to the system of algebraic equations which is used to find the saddle point in the standard Ramsey model. In the case of spatially varying $A$ and capital mobility equation (20) may reveal interesting features of the spatial distribution of capital in the steady state. In Appendix B we show that system (20) of elliptic equations may in fact be transformed into a more convenient equivalent form which is more useful for analysis and numerical treatment.

**Remark 3.3.** System (17-18) can be brought into the following alternative formulations:

1. The adjoint equation can also be expressed in terms of the optimal consumption. Upon the assumption that $U'' \neq 0$, using the optimality condition $p = U'(c)$ and (18) we obtain an evolution law for the optimal consumption in the form

$$\frac{\partial c^*}{\partial t} = -\tfrac{1}{U''(c^*)}\nabla\cdot\Big(U''(c^*)D_0 w_1(k^*)\nabla c^*\Big) + \underline{G}(k^*,x)\cdot\nabla c^* + (r+\delta - Af'(k^*))\tfrac{U'(c^*)}{U''(c^*)}, \quad (21)$$

   supplemented with homogeneous Neumann boundary conditions. Once again for $A$ constant and immobile capital ($B = 0$) this equation reduces to the relevant equation for the standard Ramsey model.

2. Upon defining the Hamiltonian functional

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

   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 $(k,p)$.

# 4 MPD spatial capital flows and spatial inequalities

In this section, based on the conjecture concerning the necessary optimality condition for problem (12), we explore various possibilities of the effect of capital transport on the formation of spatial capital structures compatible with the conjectured optimality condition. Based on the intuitive geometric approach obtained in the spatially independent case ($D_0 = 0$ and $A$ constant, hereafter referred to as the flat earth case) that characterizes the optimal path in terms of a saddle point and the corresponding stable manifold, we will try to see whether transferring this intuition to the spatially dependent case can lead to interesting new insights as to spatial structures of capital compatible with optimality. We should clarify here that since our results are based on perturbative and linearization arguments, there is no guarantee of global optimality of these patterns (a result that would require an exact nonlinear analysis) so that these structures should better be treated as local precursors of optimal spatial structures that may arise. In this spirit, and in relation with the two main questions posed in the introduction we will explore two cases:

(a) Assume that $A$ is constant, let $(k^*, p^*)$ be the fixed point of the temporal Ramsey model and $(k(t), p(t)$ be the saddle point solution of the temporal Ramsey model, such that $(k(t), p(t)) \to (k^*, p^*)$ as $t \to \infty$. Given that for $A$ constant, $(k(t, x), p(t, x)) = (k(t), p(t))$, is a solution of system (17-18) with the transversality condition (19), we will address the question of perturbations of this solution in terms of small spatial fluctuations in $k$ that will activate the action of the capital transport operator that may in turn stabilize or destabilize the original solution $(k(t), p(t))$ leading to spatial patterns.

(b) In the case where $A = A(x)$ we will consider the steady state solution of system (17-18) in the limit of weak capital mobility $B \to 0$, If $B = 0$, then the steady state solution of system (17-18) reduces to the solution of the parametric algebraic equation

$$A(x)f(k) - \delta k - (U')^{-1}(p) = 0,$$
$$A(x)f'(k) - \delta - r = 0,$$

which leads to a steady state $(k^*(x), p^*(x))$ which solves (17-18) for $B = 0$. The obtained steady state has a spatial dependence which is introduced by the spatial dependence of $A(x)$. If $\nabla A \cdot n = 0$ on $\partial \mathcal{D}$ (which is our standing assumption 2.1) one can check that upon differentiating the second of these equations the solution satisfies the boundary condition $\nabla k \cdot n = 0$ on $\partial \mathcal{D}$. Then, the question of what happens to this spatially dependent steady state arises in the presence of weak capital mobility $B \to 0$. This will lead to a perturbative construction of an approximate steady state which may add some light to the structure of possible optimal states.

The tools developed in this section will help us answer the two main questions posed in the introduction of this paper.

The first relates to whether MPD capital flux could destabilize a flat earth optimal steady state of a Ramsey model and induce spatial patterns, which in our case would imply spatial inequalities and is connected with item [(a)] above. It is well known that for the standard Ramsey optimal growth model without spatial interactions, for which the Pontryagin principle yields the system

$$\begin{aligned} \frac{dk}{dt} &= Af(k) - \delta k - (U')^{-1}(p), \\ \frac{dp}{dt} &= (Af'(x) - \delta - r)p, \\ \lim_{t \to \infty} e^{-rt} p(t) k(t) &= 0, \end{aligned} \tag{22}$$

the steady state, denoted hereafter by $(k_0^*, p_0^*)$ and obtained in terms of the algebraic equation

$$\begin{aligned} 0 &= Af(k) - \delta k - (U')^{-1}(p), \\ 0 &= (Af'(x) - \delta - r)p, \end{aligned} \tag{23}$$

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 (17)-(18) with the transversality condition (19) for $D_0 = 0$ could by analogy be regarded as the analogue of the saddle point in a 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 regarded as a "collection" of identical stable manifolds leading the system to the long-run flat earth optimal steady state.

The transversality condition (19) leaves us with a "sensible equilibrium" in which $k$ and $p$ converge to finite limits. This means that the stable manifold obtained by linear stability analysis is constructed by considering negative eigenvalues only. The transversality condition is also satisfied if the smallest eigenvalue of linear stability analysis is positive but less than $r/2$, so that the present value of capital stock is still zero as $t \to \infty$. These arguments will be used in the study of pattern formation induced by MPD.

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 to the flat earth optimal steady state? If yes, the MPD capital mobility is a spatially-homogenizing force and preserves spatial equality. If, however, the system does not return to the flat earth state, then spatial inequalities which are induced by optimal control emerge. We emphasize at this point that the obtained patterns are simply precursors to potential optimal patterns, and whether these precursors survive or not in the long term will depend on the action of the nonlinear terms and can in no way be determined solely in terms of the linear analysis presented here. However, as a minimal indication as to whether such patterns may be acceptable or not we impose the condition that the potential patterns at least satisfy the proposed transversality condition (19), in the sense that the present value of the value of the capital stock is 0 as $t \to \infty$, a fact that at least acts as an initial filter of possible spatial patterns.

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 – that is, morphogenesis. Turing's analysis has been used by Krugman (1996) and also Fujita et al. (2001) to generate patterns from a flat earth, steady-state space with economies located on a circle.[8] However, neither the analysis by Turing nor that by Krugman involve explicit forward-looking dynamic optimization. Dynamic optimization and the emergence of patterns was first analysed by Brock and Xepapadeas (2008, 2010) and subsequently by Brock et al. (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,

---

[8] This is the "50 Cadillac diagram" for a race track economy (see Krugman, 1998).

since they move from high to low concentrations. However, as we discussed earlier, this assumption is not appropriate for capital movements in the context of the Lucas paradox. Thus our objective in terms of our first question is to examine the conditions under which 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 a force that generates spatial patterns and inequalities of per capita output and capital across space.

The objective of our second question is to find out whether MPD capital mobility intensifies or weakens already existing – because of TFP differences – spatial inequality. These two cases are analysed in the following sections.

### 4.1 Spatial homogeneity and optimal pattern formation

In the temporal Ramsey model, saddle point stability means that the linearisation matrix of system (22) at the steady state has a negative determinant, and that 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.

Studying the linearised 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. While interesting results have recently been reported in this front for the spatial AK model in Calvia et al (2021) based on the detailed study of the Hamilton-Jacobi equation which is possible for this particular model, there are still open questions for more general models. 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. This perturbative approach could reveal the potential generation of optimal spatial patterns which emerge because the MPD capital mobility destabilizes the stable manifold of the zero MPD mobility system.

Throughout this section we make the following standing assumption:

**Assumption 4.1.** The TFP, $A$, and the capital mobility coefficient $B$ are independent of $x$.

We start by noting that if a solution to (17-18-19) exists then there must also be a $p_0 = p(0, x)$ such that if (17-18) is solved forward in time the solution is recovered. The same is true for the standard Ramsey model, in the absence of spatial effects, i.e. to the solution of system (22). As already mentioned, in the latter case, given a solution of (22) such that $k(0) = k_0$, then choosing $p_0$ on the stable manifold of the saddle point $(k_0^*, p_0^*)$ (i.e., the steady state of (22)) will be such an initial condition for $p$, so that starting $(k^{(0)}(0), p^{(0)}(0)) = (k_0, p_0)$ will generate a solution $(k^{(0)}(t), p^{(0)}(t))$ to (22) such that $(k^{(0)}(t), p^{(0)}(t)) \to (k_0^*, p_0^*)$ as $t \to \infty$.

In trying to examine the potential emergence of pattern formation induced by MPD capital mobility we propose the following experiment. Let us start (17-18-19) at an initial condition $k(0, x) = k_0^* + \epsilon k_1(x)$ and some initial control procedure corresponding to the adjoint process $p(0, x) = p_0^* + \epsilon p_1(x)$. For $\epsilon = 0$ this is clearly a solution of (17-18-19) for any $D$. What will be the solution of (17-18-19) for this initial condition for $\epsilon \neq 0$, and how will its behaviour change depending on the level of $D$?

Note that even though we treat (17-18) an initial value problem, we still impose the transversality condition (19). In fact, the spirit of this construction is to start with a given initial perturbation $k(0,x) = k_0^* + \epsilon k_1(x)$ and then treat $p(0,x)$ in the same fashion that we would treat a shooting method for a boundary value problem, i.e., leave it free and try to construct a solution to (17-18) with this initial condition that would satisfy (19). Given $k(0,x)$ (and after choosing $p(0,x)$ appropriately with the above shooting method) we construct the solution $(k(t,x), p(t,x))$ of (17-18) satisfying (19), and check whether this solution displays spatial dependence or not. This is performed by a linearised analysis of (17-18-19) near the solution $(k_0^*, p_0^*)$. The linearised analysis reveals that whether the approximate solution of (17-18-19) displays spatial variability or not depends on the value of the capital mobility $D$.

**Proposition 4.2.** *Let $(k_0^*, c_0^*)$ be the flat earth steady state*[9]*, define $M = A\frac{U'(c_0^*)}{U''(c_0^*)} f''(k_0^*) > 0$, and let $\{\mu_n, \phi_n \ : \ n \in \mathbb{N}\}$ be the eigenvalues and eigenfunctions of the operator $-\Delta$ with homogeneous Neumann boundary conditions on $\partial\mathcal{D}$.*

*For $D_0 = \bar{D}_0 B$ fixed, given a small initial perturbation $k(0,x)$ of $k_0^*$ of the form $k(0,x) = k_0^* + \epsilon k_1(x)$ there exists an approximate solution of (17-18) satisfying the tranversality condition (19) of the form*

$$Z(t,x) := \begin{pmatrix} k(t,x) \\ p(t,x) \end{pmatrix} = \begin{pmatrix} k_0^* \\ p_0^* \end{pmatrix} + \epsilon \sum_{n\in\mathbb{N}} e^{\rho_{1,n} t} \begin{pmatrix} u_n \\ v_n \end{pmatrix} \phi_n(x) + O(\epsilon^2) \tag{24}$$

*where $v_n, u_n$ are constants determined in terms of the expansion of the initial disturbance $k_1(x)$ in the basis $\{\phi_n \ : \ n \in \mathbb{N}$ (explicitly provided in the proof) and*

$$\rho_{1,n} = \frac{r}{2} - \frac{1}{2}\sqrt{r^2 - 4\left[D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M\right]} \tag{25}$$

*The solution provided by* (24) *has the following properties:*

*(a) $Z(t,x) \to Z^* := (k_0^*, p_0^*)^T$ if $D_0 = 0$, which indicates that no spatial patterns will emerge in the absence of capital mobility.*

*(b) $Z(t,x) \to Z^* := (k_0^*, p_0^*)^T$, for any $D_0$ if $r^2 < 4M$, which indicates that no spatial patterns will emerge for small discount rates.*

*(c) If $r^2 > 4M$ then,*

$$Z(t,x) \to \begin{pmatrix} k_0^* \\ p_0^* \end{pmatrix} + \epsilon \sum_{n\in\mathcal{N}} e^{\rho_{1,n} t} \begin{pmatrix} u_n \\ v_n \end{pmatrix} \phi_n(x) + O(\epsilon^2), \ \text{where}$$

$$\mathcal{N} := \{n \in \mathbb{N} \ : \ \rho_{1,n} > 0\} = \{n \in \mathbb{N} \ : \ \frac{1}{2}(r - \sqrt{r^2 - 4M}) \le D(k_0^*)\mu_n \le \frac{1}{2}(r + \sqrt{r^2 - 4M})\} \tag{26}$$

For the proof, and detailed statements of the instability conditions, see Appendix C.

---

[9] attained for the case where $A$ is constant and capital is immobile, $D_0 = \bar{D}_0 B = 0$ but also a solution of (17-18) satisfying the tranversality condition (19) for any value of $D_0 = \bar{D}_0 B$.

**Remark 4.3.** The following remarks are in order:

1. If $D_0 = 0$ then $\rho_{1,n} < 0$ for all $n \in \mathbb{N}$ (alternatively $\mathcal{N} = \emptyset$) hence spatial patterns will occur for a spatially dependent initial condition $k(0,x) = k_0^* + \epsilon k_1(x)$ for small $\epsilon$, only if $D_0 \neq 0$. This justifies the terminology diffusion induced instability.

2. If the discount rate $r$ is smaller than a critical value $M$ determined by the economic dynamics then it also holds that $\mathcal{N} = \emptyset$, so that no spatial patterns will occur.

3. If destabilization occurs, the result of Proposition 4.2 can 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 structure is in principle supported by the optimal control procedure, as long as the growth rate is compatible with the transversality condition, a fact which is guaranteed by the condition definining $\mathcal{N}$ in (26). Importantly, given $D(k_0^*)$ and a production function, not every possible pattern will develop but only those compatible with the transversality condition, i.e. those satisfying (26). This constitutes an important difference with the standard mechanism of the Turing instability.

4. A spatial pattern related to a mode $\phi_n$ will only emerge if the diffusion coefficient $D(k_0^*) = \bar{D}_0 B w_1(k_0^*)$ satisfies the condition in (26), i.e. requires $D(k_0^*)$ to be higher than a lower critical value, but certainly lower than a higher critical value (both determined by the discount factor, and the economic dynamics). It is interesting to note that just as well as low capital transport disables the pattern formation activity, high diffusion smooths our generated spatial gradients of capital density leading asymptotically to a homogeneous state.

5. On the other hand given a $D(k_0^*)$, then only a limited number of modes may appear in the asymptotic behaviour of the solution of (17-18) satisfying the tranversality condition (19), i.e., (24), and in fact the modes corresponding to $n \in \mathcal{N}$. Since $\mu_n \to \infty$ as $n \to \infty$, (26) implies that only a finite number of modes $n$ may be activated for given values of $D(k_0^*)$, $r$ and $M$ and in particular that there exists a $n^* \in \mathbb{N}$ such that only modes satisfying $n < n^*$ will appear in the asymptotic pattern.

6. 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 (26) feasible. The explicit knowledge of the eigenfunctions also allows for the determination of the spatial variability of the generated pattern.

7. The above remarks allow us to assess the effects of nonlinear diffusion in the formation of patterns. The linear diffusion case will yield a condition similar to the condition definining $\mathcal{N}$ in (26) with $D(k_0^*)$ replaced by $D_0 = \bar{D}_0 B$. Nonlinear effects, depending on the function $w_1$ (respectively on the production function $f$ and the capital mobility function $\psi$) may lead to $D(k_0^*) > D_0$ or $D(k_0) < D_0$, therefore having an important effect of the appearance or disappearance of new modes in the observed pattern.

It is interesting to expand a bit further on the economic intuition of the observations made in Remark 4.3.

(a) The condition $r^2 < 4M$ for the non existence of patterns (see Remark 4.3) can alternatively be interpreted as a condition on the utility function. For example if $U(c) = c^{1-\lambda}/(1-\lambda)$ and $f(k) = Ak^a$ then $M = \frac{\delta+r}{\alpha}\frac{1-\alpha}{\lambda}\left[(1-\alpha)\delta + r\right]$ so that pattern may occur only if

$$\lambda > \lambda(\alpha, r) := \frac{4}{\alpha r^2}(1-\alpha)(\delta+r)\left[(1-\alpha)\delta + r\right]. \tag{27}$$

Figure 1 depicts the surface $\lambda(a, r)$ corresponding to (27). MPD capital mobility will destabilize the flat earth optimal steady state for $(\alpha, r, \lambda)$ points above the $\lambda(\alpha, r)$ surface. The flat surface corresponds to $\lambda = 1$ of a logarithmic utility function.

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

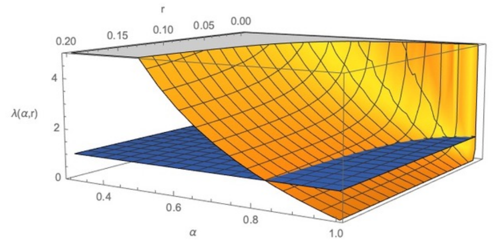


Figure 1: The $\lambda(\alpha, r)$ surface corresponding to (27) ($\delta = 0.03$).

(b) The link between the emergence of spatial patterns by destabilization of the flat earth optimal steady state through spatial capital flows and the utility discount rate can be made clear using the destabilization condition (57). From (57), destabilization occurs at some mode $n$ if

$$\det L_n = \begin{vmatrix} r - D(k_0^*)\mu_n & -1 \\ -\frac{U'(c_0^*)}{U''(c_0^*)}Af''(k_0^*) & D(k_0^*)\mu_n \end{vmatrix} > 0, \tag{28}$$

where $\mu_n > 0$ with $\mu_n \to \infty$ as $n \to \infty$, while for $\mu_0$, (28) 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,

$$\det L_n = \psi_n(r) = \alpha_n(r - \alpha_n) - M, \tag{29}$$

which is linearly 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 flat earth optimal steady state is stable in the saddle point sense. For $r < r_1$, the flat earth optimal steady state 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 $\alpha_n(r_n^+ - \alpha_n) = M$. For $r < r_n^+$, the flat earth optimal steady state is stable to spatial perturbations induced by capital flows, while for $r > r_n^+$, the flat earth optimal steady state is destabilized at mode $n$ and spatial patterns emerge.

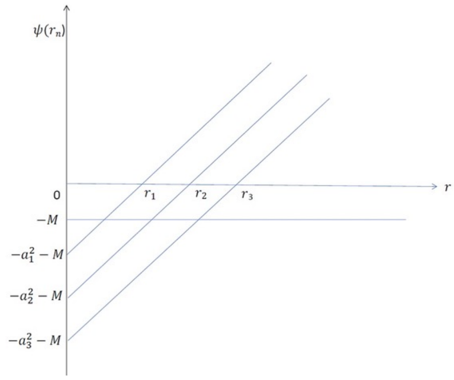


Figure 2: Critical $r$ and destabilization of a flat earth optimal steady state.

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 flat earth optimal steady state; $k_0^*$; $M|k_n - k_0^*|$ 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 flat earth optimal steady state enhances its benefits and thus no patterns emerge. However, if $r$ is sufficiently high, so that $\alpha_n(r - \alpha_n) > M$, then the net benefits of controlling the system to the flat earth optimal steady state 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 flat earth optimal steady state.

(c) Boucekkine et al. (2013b) provide a similar pattern-eliminating low discount rate result for an $AK$ model with linear diffusion in which $r^+ = A(1 - \lambda) + \lambda$. In this case optimal patterns depend on the productivity parameter and elasticity of marginal utility.[10] In our case the nonlinear diffusion induced by the MPD capital flows combined with the Cobb-Douglas production function provides more degrees of freedom for having emergence of optimal spatial patterns, as shown in Figure 1.

We close this section by providing some examples of the possible patterns that may emerge as a result of (24) for various geometries. In all examples in this section we use the Cobb-Douglas production function with exponent $\alpha$ and the isoelastic utility function parametrized as $U(c) = \frac{1}{1-\lambda}c^{1-\lambda}$.

**Example 4.4.** 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 the

---

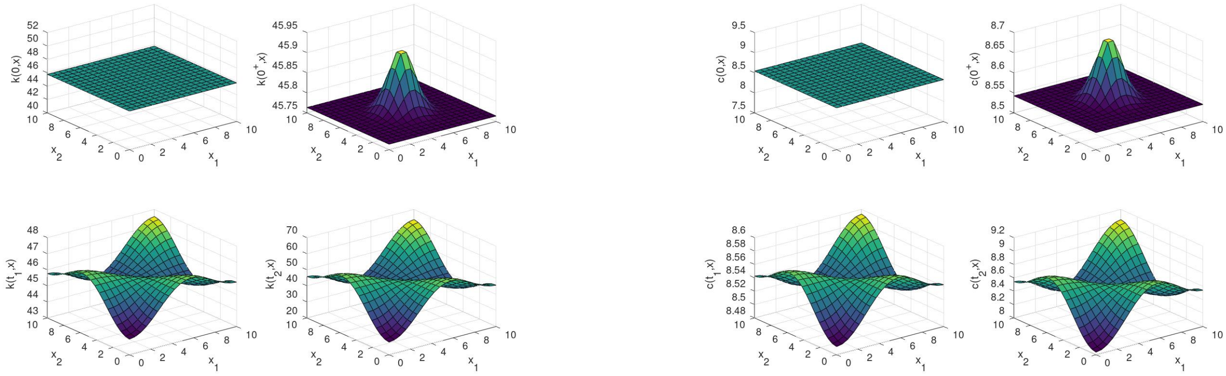
[10] For logarithmic utility, $\lambda = 1$ and patterns will never emerge for reasonable utility discount rates.

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

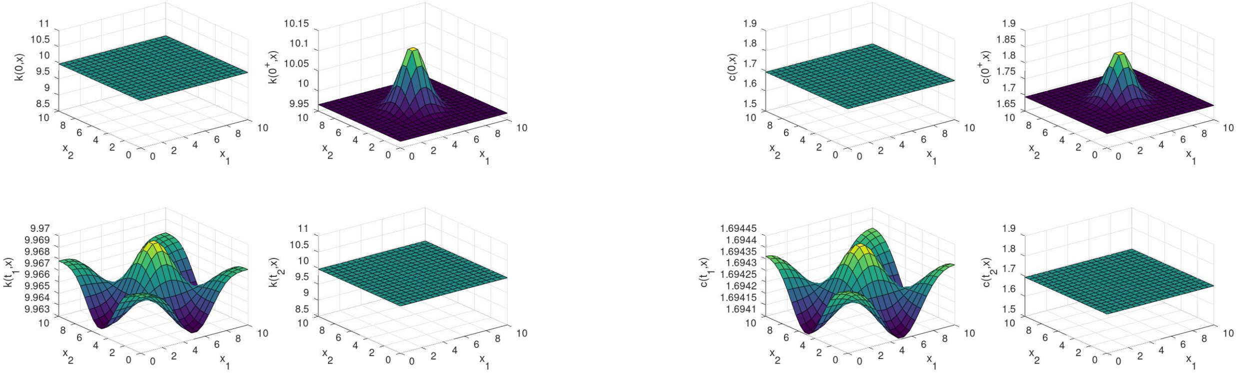
$$D(k_0^*)\left(\frac{\pi}{L}\right)^2\left(r - D(k_0^*)(\frac{\pi}{L})^2\right) - M < 0.$$



(a) Capital

(b) Consumption

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



(a) Capital

(b) Consumption

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

**Example 4.5.** 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 multi-index $n = (n_1, n_2)$, we have 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 into the framework of Proposition 4.2, we can 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 $(n_1, n_2) = (2, 1)$, and so on. Defining $\nu = \frac{L_1}{L_2}$, the aspect ratio of the rectangle, the condition for the generation of patterns can now be expressed as follows. Consider the set of pairs of natural numbers

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

Then, emerging spatial patterns will be given by the double sum

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

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

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

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

In the context of the two-dimensional spatial domain, Figure 3 depicts the case of an emerging spatial pattern for high $r$,$\lambda$ and $\alpha$ after a spatial perturbation and Figure 4 depicts the case of a return to the flat earth optimal steady state (decaying pattern) after the same spatial perturbation as in Figure 3. The four graphs in the panels of both Figures 3 and 4 each correspond to a different point in time, starting at $t = 0$. The solutions were generated using the method proposed in Appendix C.

**Example 4.6** (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 $\rho$ 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

(a) Capital

(b) Consumption

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

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 (55) for the Laplacian on the sphere 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. That is,

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

unless $\ell = \ell'$ or $m = m'$. 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, we obtain:

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

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

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

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 the condition in (26) 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 no association with actual geographical areas or countries on the earth-like globe, but are presented solely for illustration purposes. Associating the emerging spatial patterns with actual locations on earth is an important area for future research.

## 4.2 Spatial heterogeneity and MPD capital mobility

In the previous section we showed that optimal growth under MPD capital mobility will not generate spatial heterogeneities for a typical Cobb-Douglas utility parametrization and low utility discount rate. This result is derived under the assumption of equal TFP across locations ($A$ constant). In this section we allow for TFP differences across locations, that is $A = A(x)$, and we explore the question of whether MPD capital mobility combined with TFP heterogeneity intensifies or weakens spatial heterogeneity. The spatial distribution of the TFP parameter could be induced by nature-first factors or by population density (e.g., Allen and Arkolakis, 2014; Boucekkine et al., 2019)

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, parametrized by $A = A(x)$. That is, the solution of the parametric steady-state equation is given by

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

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

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

An interesting question is how this non-flat steady state $(k_0, c_0)$ would evolve in the presence of weak capital mobility $\bar{D}_0$, if perturbed by a perturbation $(k_1, c_1)$. This requires a study of the steady state of (17-18), in the limit of small $\bar{D}_0$.

The following proposition provides some insight into this.

**Proposition 4.7.** *Assume a spatially-varying TFP, $A = A(x)$, satisfying the standing Assumption 2.1, 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, for $\bar{D}_0$ small enough there is an approximate steady state of (17-18) given by*

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

*(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,*

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

*and*

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

For the proof, see Appendix D.

Proposition 4.7 allows us to approximate, explicitly and analytically, the effect of small capital mobility on possibly sharp spatial gradients in the TFP $A$.

**Example 4.8.** Consider the Cobb-Douglas case in Proposition 4.7(ii) and a variation in the TFP $A$

in the form of a sum of Gaussians as

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

where $\| \cdot \|$ denotes the Euclidean norm.

This general form can model local increases ($C_i > 0$) or decreases ($C_i < 0$) at locations $x_i$, of scale $\sigma_i > 0$, where small $\sigma_i$ corresponds to a localized spatial structure and large $\sigma_i$ corresponds to extended spatial structures. Moreover, by assuming that the domain $\mathcal{D}$ is large and that the centers of the Gaussians, $x_i$ are placed sufficiently far from $\partial\mathcal{D}$, an easy calculation yields that $A$ approximately satisfies the boundary condition $\nabla A \cdot n = 0$ on $\partial\mathcal{D}$.

In this case we can calculate $(k(x), c(x))$ in terms of the expression in Proposition 4.7(ii), using the facts that

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

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 $k$. We then 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 distribution. If $I_0 < I$, then the effect of capital mobility will be an enhancement of the spatial variability, while if $I_0 > I$, then the effect of capital mobility will be a smoothing of spatial variability in the optimal capital distribution.

Providing a priori estimates for the above quantities would require technicalities which are beyond the scope of the present paper. We prefer to adopt a more direct treatment of 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 4.9.** *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* $E(d)$, *which depends on the spatial dimension, is*

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

*For proof, see Appendix D.*

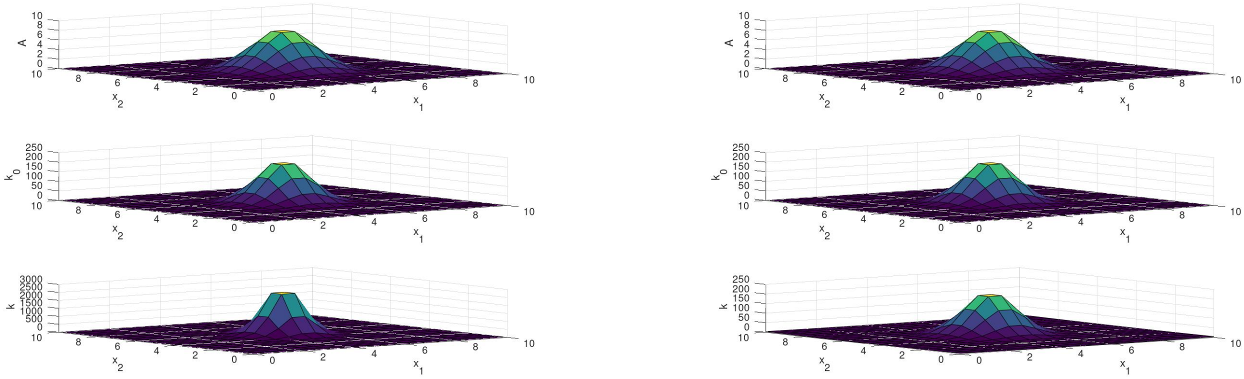
The above calculations show that up to small order corrections in $\bar{D}_0$:

- The Sobolev norm of the solution will increase in a 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 $0 \leq \rho < 1$ and $\lambda > \frac{3(\rho+1)}{1-\rho}$ (i.e., for $\rho = 0$ if $\lambda > 3$). 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 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. This depends on the relationship between the propensity of capital to move and 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.

This result suggests that MPD capital flows combined with the a high or low value of the elasticity of the propensity of capital to move to a high productivity location could intensify or weaken agglomerations which may already exist because of nature-first factors or distribution of population. Boucekkine et al., (2019) in the context of an $AK$ model with linear diffusion and exogenous location dependent TFP, also show the emergence of spatial agglomerations which might have different transitional dynamics for different initial conditions but they result in the same long-run capital distribution. Our results are comparable to theirs but in addition we show that initial exogenous agglomeration with respect to capital and consumption could be intensified or weakened by the MPD capital flows in the sense that the system could end with a sharp agglomeration or a nearly flat earth state. Furthermore we can observe intensification of agglomeration with diminishing returns to capital ($a < 1$) if the capital flows indicate high propensity to move to a high marginal productivity location.



(a)

(b)

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

## 5 Concluding remarks

In this work we initiate the study of a spatial Ramsey model of optimal growth displaying nonlinear capital transport effects by considering capital flows toward locations where the marginal productivity of capital is relatively higher. This induces nonlinear diffusion in the fundamental equation of capital accumulation.

In this context, based on perturbative solutions around the non spatial Ramsey model, we examine questions related to the potential spatial heterogeneity of optimal growth, as seen from the viewpoint of a social planner who seeks to maximize discounted utility over a finite spatial domain by choosing optimal consumption paths for each location. Our results suggest that for a high utility discount rate and appropriate parameters for the production and the utility function, MPD capital flows and forward-looking optimizing behaviour by the social planner could induce the emergence of spatial patterns. For a 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 exist due to TFP differences across locations, MPD capital flows could intensify or weaken spatial inequalities. This depends on whether there is a high or low tendency of capital from high capital accumulation locations to move once higher marginal productivity occurs in a different location.

Our results provide insights into the way in which an intuitive plausible mechanism of capital flows across locations could generate spatial inequalities in the context of a traditional Ramsey model of optimal growth. Further research should be directed toward studying market equilibrium under MPD capital flows. If spatial heterogeneity emerging from market equilibrium outcomes is different from the outcomes obtained in this paper for a social planner, this deviation could be the basis for exploring economic policies for attaining socially-optimal spatial structures.

**Acknowledgments**

The authors wish to thank the anonymous referees for their careful study of the manuscript and their

useful and insightful comments that helped improve considerably this paper.
ANY wishes to acknowledge the use of resources of the Stochastic Modelling and Applications Research Laboratory of the Department of Statistics, AUEB. Moreover, he wishes to thank Prof. E. V. Petracou for useful discussions and support.

# Appendix

## A Towards a derivation of the Pontryagin principle for problem (12)

In this section we provide some insight towards the derivation of the Pontryagin maximum principle for problem (12). While this is not a proof of the conjectured optimality condition it gives some indication as to its plausibility.

### A.0.1 Key assumptions

**Assumption A.1.** .

1. $U$ and the production function are sufficiently smooth with bounded derivatives.
2. $n \cdot \nabla A = 0$ on $\partial \mathcal{D}$.
3. An optimal consumption policy $c^* > 0$ exists for problem (12).
4. The evolution equation (10) admits a global in time solution of sufficient regularity for any admissible control of interest (in an appropriate function space setting).
5. The linearized evolution equation

$$\begin{aligned} \frac{\partial w}{\partial t} - \nabla \cdot (D_0 w_1(k^*) \nabla w + \underline{G}(k^*) w) - A f'(k^*) w + \delta w = v, \\ w(0) = w_0, \\ n \cdot \nabla w = 0, \ \text{ on } \partial \mathcal{D}, \end{aligned} \tag{30}$$

   admits a unique global solution in the same function space setting for any $(w_0, v)$ with continuous dependence on $v$.

The important assumption is 4, as it requires an appropriate function setting, which at least to start with must start in a Sobolev space setup. Related results for the finite horizon case can be found (see e.g. Casas and Chrysafinos (2018) for the case of homogeneous Dirichlet conditions). Assumption 5 is easier, since it requires solvability results for a linear parabolic equation, which under assumption 1, has continuous coefficients. Equations of this type are known to satisfy this property in finite horizons.

### A.0.2 Notation

We define the following:

- The functional

$$c \mapsto J(u) := \int_0^\infty \int_{\mathcal{D}} e^{-rt} U(c(t,x)) dx dt,$$

  where $c : \mathbb{R}_+ \times \mathcal{D} \to \mathbb{R}_+$, is the consumption function.

- An adjoint variable $p$ which is a function $p : \mathbb{R}_+ \times \mathcal{D}$, satisfying a Neumann boundary condition $\nabla p \cdot n = 0$ on $\partial\mathcal{D}$, whose evolution law will be determined shortly.
- The functions

$$t \mapsto I(t) := \int_{\mathcal{D}} k(t,x)p(t,x)dx,$$
$$t \mapsto \bar{I}(t) := e^{-rt} I(t).$$

- For any given suitable function $k : \mathbb{R}_+ \times \mathcal{D} \to \mathbb{R}$, the vector field

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

  Note that since the data of the problem are chosen so that $n \cdot \nabla \ln A = 0$ on $\partial\mathcal{D}$, it holds that if $k$ satisfies homogeneous Neumann boundary conditions then $\underline{G}(k) \cdot n = 0$ on $\partial\mathcal{D}$.

### A.0.3 Differentiability of the solution map

To show the differentiability of the solution map we may use the implicit function theorem in Banach spaces (see Theorem A.2).

Consider the map $F(c,k) : \mathcal{U} \times Y \to \mathcal{U} \times Y_0$ defined by

$$F(c,k) = (\frac{\partial k}{\partial t} - \nabla \cdot (D_0 w_1(k)\nabla k - D_0 w_2(k)\nabla A) - Af(k) + \delta k + c, k(0) - k_0),$$

where $Y$ and $Y_0$ are appropriate functional spaces, such that Assumption A.1.5 above holds (for this study these will be Sobolev type spaces see e.g. Casas and Chrysafinos 2018, however, they may also be chosen as spaces of continuously differentiable functions if classical solutions are to be considered) and $\mathcal{U}$ is an open subset of the set of allowable controls.

Let $c^* > 0$ be a solution to the optimal control problem and $k^*$ be the corresponding optimal path satisfying $k(0) = k_0$. Then $F(c^*, k^*) = (0,0)$. If $\frac{\partial F}{\partial k}(c^*, k^*) : Y \to \mathcal{U} \times Y_0$ is an isomorphism and $F$ is $C^k$ then the implicit function theorem guarantees the existence of an open neighbourhood $N(c^*) \subset \mathcal{U}$ and a function $u : N(c^*) \subset \mathcal{U} \to Y$ such that $F(c, u(c)) = F(c^*, k^*) = (0,0)$ which moreover if also $C^k$ (see Theorem A.2), This function $u$ is the solution map $S$, hence, such an argument guarantees the differentiability of the solution map.

It therefore remains to check the assumptions of the implicit function theorem. Using the differentiability assumption on $f$, we can show that

$$D_k F(c,k)w = \left(\frac{\partial w}{\partial t} - \nabla \cdot (D_0 w_1(k)\nabla w + \underline{G}(k)w) - Af'(k)w + \delta w, w_0\right), \tag{31}$$

supplemented with the homogeneous Neumann condition $n \cdot \nabla w = 0$ on $\partial\mathcal{D}$.

Indeed, consider $k + \epsilon w$ and consider $F(c, k + \epsilon w)$. Then,

$$\frac{1}{\epsilon}(F(c, k + \epsilon w) - F(c,k)) = (J_1(\epsilon), J_2(\epsilon)),$$

where

$$J_2(\epsilon) = \frac{1}{\epsilon}\Bigg(k(0) + \epsilon w(0) - k_0 - (k(0) - k_0)\Bigg) = w_0,$$

and setting $\bar{k} = k + \epsilon w$ to simplify the notation

$$\begin{aligned} J_1(\epsilon) = \frac{1}{\epsilon}\Bigg(&\frac{\partial \bar{k}}{\partial t} - \nabla \cdot (D_0 w_1(\bar{k})\nabla \bar{k} - D_0 w_2(\bar{k})\nabla A) - Af(\bar{k}) + \delta\bar{k} + c \\ &- \Bigg(\frac{\partial k}{\partial t} - \nabla \cdot (D_0 w_1(k)\nabla k - D_0 w_2(k)\nabla A) - Af(k) + \delta k + c\Bigg)\Bigg) \\ = \frac{\partial w}{\partial t} - \nabla \cdot (D_0 w_1(\bar{k})\nabla w &+ D_0\frac{1}{\epsilon}(w_1(\bar{k}) - w_1(k))\nabla k - D_0\frac{1}{\epsilon}(w_2(\bar{k}) - w_2(k))\nabla \ln A) \\ &-A\frac{1}{\epsilon}(f(\bar{k}) - f(k)) + \delta w \end{aligned} \tag{32}$$

By the differentiability properties of $w_1$ and $w_2$, which in turn follow from the differentiability properties of $f$, using the mean value we have (using $\varphi$ as a proxy for any of the functions $w_1$, $w_2$ or $f$) we have that there exists an $\epsilon_0 \in (0, \epsilon)$ such that

$$\varphi(\bar{k}) = \varphi(k) + \epsilon\varphi'(k + \epsilon_0 k)k,$$

hence using the continuity of $\varphi'$ we obtain that

$$\frac{1}{\epsilon}(\varphi(\bar{k}) - \varphi(k)) = \varphi'(k + \epsilon_0 k)w \to \varphi'(k)w, \text{ as } \epsilon \to 0.$$

Using the above observation in (32) and passing to the limit as $\epsilon \to 0$ we get that

$$J_1(\epsilon) \to \frac{\partial w}{\partial t} - \nabla \cdot (w_1(k)\nabla w + D_0 w_1'(k)\nabla k\, w - D_0 w_2'(k)\nabla \ln A\, w) - Af'(k)w + \delta w,$$

as $\epsilon \to 0$. Hence, (31) is obtained.

To show that $\frac{\partial F}{\partial k}(c^*, k^*) : Y \to \mathcal{U} \times Y_0$ is an isomorphism is equivalent to showing that the linear PDE

$$\begin{aligned} \frac{\partial w}{\partial t} - \nabla \cdot (D_0 w_1(k^*)\nabla w + \underline{G}(k^*)w) - Af'(k^*)w + \delta w = v, \\ w(0) = w_0, \\ n \cdot \nabla w = 0, \text{ on } \partial\mathcal{D}, \end{aligned} \tag{33}$$

admits a unique solution in $Y$ for every $v \in \mathcal{U}$, with this solution depending continuously on $v$.

In principle $k^*$ displays spatio-temporal dependence so its unique solvability must rely on the theory of non-homogeneous linear parabolic PDEs, with coefficients depending on both space and time. One difficulty here is that we want to guarantee solutions for all $t \in \mathbb{R}_+$, and this may require delicate estimates and the appropriate choice of $\delta > 0$.

Under the assumption that (33) assumes a unique solution in $Y$ for every $v \in \mathcal{U}$ (with continuous dependence), the above considerations show that the solution map $S : \mathcal{U} \to Y$ is differentiable with $z := G'(k)v$ satisfying a linear parabolic PDE. This follows by an application of the implicit function theorem. It is easy to see that

$$D_c F(c, k)v = (v, 0),$$

hence

$$DS(c) = -D_k(c, S(c))^{-1} D_c F(c, S(c))$$

implies that $z_v := DS(c)v$ is the solution of the linear parabolic PDE

$$\begin{aligned} \frac{\partial z_v}{\partial t} - \nabla \cdot (D_0 w_1(k) \nabla z_v + \underline{G}(k) z_v) - A f'(k) z_v + \delta z_v = -v, \\ z_v(0) = 0, \\ n \cdot \nabla z_v = 0, \text{ on } \partial \mathcal{D}, \end{aligned} \tag{34}$$

where $k = S(c)$. This essentially means that upon perturbing a consumption plan $c$ to $c+v$, if $k = S(c)$ is the capital distribution corresponding to $c$, then the capital distribution $\hat{k} = S(c+v)$ corresponding to the new perturbed consumption plan $c + v$ can be approximated as $k + z_v$ where $z_v$ is the solution of (34) with the error of approximation being controlled by an appropriate norm of $v$. In particular considering any $c \in \mathcal{U}$ and using a perturbed consumption plan of the form $c + \epsilon v$ then, the mean value theorem for the solution map $S$ yields that there exists an $\epsilon_0 \in (0, \epsilon)$ such that

$$S(c + \epsilon v) = S(c) + \epsilon S'(c + \epsilon_0 v)v,$$

which is expressed as

$$\hat{k} = k + \epsilon z_{\epsilon,v}$$

with $z_{\epsilon,v}$ solving (34) with $k$ replaced by $k_\epsilon = S(c + \epsilon v)$. Since $S$ is a continuous map, $k_\epsilon \to S(c) = k$ as $\epsilon \to 0$. The continuous dependence property of the solution to (34) then shows that $z_{\epsilon,v} \to z_v$ as $\epsilon \to 0$.

### A.0.4 Choice of the adjoint equation

We now show that if $z_v$ satisfies (34) while $p$ satisfies

$$\begin{aligned} \frac{\partial p}{\partial t} + \nabla \cdot (D_0 w_1(k) \nabla p) - \underline{G}(k) \cdot \nabla p + f'(k)p - \delta p = rp \\ n \cdot \nabla p = 0, \text{ on } \partial \mathcal{D}, \\ \lim_{t \to \infty} e^{-rt} \int_{\mathcal{D}} z_v p dx = 0 \end{aligned} \tag{35}$$

then

$$\int_0^\infty e^{-rt} \int_{\mathcal{D}} vpdxdt = \lim_{T\to\infty} \int_0^T e^{-rt} \int_{\mathcal{D}} vpdxdt = 0. \tag{36}$$

We prove these claims: For more clarity we reintroduce the explicit dependence on $t$ and $x$ only at the points where we think it is necessary.

Define the functions

$$t \mapsto I(t) := \int_{\mathcal{D}} z_v(t,x)p(t,x)dx,$$
$$t \mapsto \bar{I}(t) := e^{-rt} I(t).$$

and the vector field

$$(t, z_v) \mapsto \underline{J}_1(k, z_v) := D_0 w_1(k) \nabla z_v + \underline{G}(k) z_v,$$

that satisfies $n \cdot \underline{J}_1(k, z_v) = 0$ on $\partial\mathcal{D}$ since $n \cdot z_v = 0$ and $n \cdot \underline{G}(k) = 0$ on $\partial\mathcal{D}$.

For the remaining calculations we omit the subscript $v$ and denote $z_v$ as $z$ for simplicity. Assuming enough smoothness in order to be able to differentiate under the integral sign (invoking Lebesgue's dominated convergence theorem) we have that

$$\frac{d}{dt} I(t) = I_1 + I_2 := \underbrace{\int_{\mathcal{D}} \frac{\partial z}{\partial t}(t,x)p(t,x)dx}_{I_1} + \underbrace{\int_{\mathcal{D}} \frac{\partial p}{\partial t}(t,x)z(t,x)dx}_{I_2}. \tag{37}$$

We treat each of the above terms separately.

For $I_1$ we observe that by (34) we have

$$I_1 = \int_{\mathcal{D}} \nabla \cdot \underline{J}_1(k,z)pdx + \int_{\mathcal{D}} (f'(k)z - \delta z - v)pdx. \tag{38}$$

We apply Green's theorem (see Theorem A.3 to the first integral of (38), setting $\underline{F} = \underline{J}_1(k,z)$ and $g = p$, noting that $n \cdot \underline{J}_1(k,z) = 0$ on $\partial\mathcal{D}$, to obtain

$$\int_{\mathcal{D}} \nabla \cdot \underline{J}_1(k,z)p\,dx = -\int_{\mathcal{D}} \underline{J}_1(k,z) \cdot \nabla p\,dx. \tag{39}$$

Combining (38), (39) and (37) we conclude that

$$\frac{d}{dt} I = \int_{\mathcal{D}} \Big( - \underline{J}_1(k,z) \cdot \nabla p + (f'(k)z - \delta z - v)p \Big) dx + \int_{\mathcal{D}} \frac{\partial p}{\partial t} z\,dx \tag{40}$$

Using the definition of $\underline{J}_1(k,z)$ we rearrange (40) as

$$\begin{aligned}\frac{dI}{dt} &= \int_{\mathcal{D}} \Big( - D_0 w_1(k)\nabla z \cdot \nabla p - z\,\underline{G}(k)\cdot \nabla p + (f'(k)z - \delta z - v)p + \frac{\partial p}{\partial t} z \Big) dx \\ = \int_{\mathcal{D}} \Big( - \underline{G}(k)\cdot\nabla p + f'(k)p - \delta p + \frac{\partial p}{\partial t}\Big) z\, dx &+ \int_{\mathcal{D}} \Big( - D_0 w_1(k)\nabla z\cdot\nabla p\Big) dx - \int_{\mathcal{D}} vp\, dx\end{aligned} \tag{41}$$

We now consider the middle term in the above, using once more Green's theorem (see Theorem A.3, for the choice $\underline{F} = -D_0 w_1(k^*)\nabla p$ and $g = k$, where upon noting that $n \cdot \underline{F} = 0$ on $\partial\mathcal{D}$ since $p$ satisfies the homogeneous Neumann boundary condition we obtain

$$\int_{\mathcal{D}} \Big( - D_0 w_1(k)\nabla z\cdot\nabla p\Big) dx = \int_{\mathcal{D}} z\,\nabla\cdot\,(D_0 w_1(k)\nabla p) dx. \tag{42}$$

Combining (41) and (42) we obtain upon rearranging,

$$\frac{dI}{dt} = \int_{\mathcal{D}} \Big( \frac{\partial p}{\partial t} + \nabla\cdot\,(D_0 w_1(k)\nabla p) - \underline{G}(k)\cdot\nabla p + f'(k)p - \delta p\Big) z\, dx - \int_{\mathcal{D}} vp\, dx. \tag{43}$$

Hence,

$$\frac{d\bar{I}}{dt} = -re^{-rt} I + e^{-rt}\frac{dI}{dt} =$$
$$e^{-rt}\int_{\mathcal{D}} \Big( \frac{\partial p}{\partial t} + \nabla\cdot\,(D_0 w_1(k)\nabla p) - \underline{G}(k)\cdot\nabla p + f'(k)p - \delta p - rp\Big) z\, dx - e^{-rt}\int_{\mathcal{D}} vp\, dx,$$

so that choosing the adjoint variable $p$ as the solution of the evolution equation (35), the above becomes

$$\frac{d\bar{I}}{dt}(t) = -e^{-rt}\int_{\mathcal{D}} vp dx. \tag{44}$$

which integrated over the time interval $[0,T]$ yields (we reinstate now the explicit $(t,x)$ dependence for clarity)

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

or equivalently,

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

which upon choosing $z(0,x) = 0$ a.e $x \in \mathcal{D}$ and $\lim_{T\to\infty} e^{-rT}\int_{\mathcal{D}} z(T,x)p(T,x)dx = 0$ leads to (36).

### A.0.5 First order condition

Let $\mathcal{C} = \{c \geq 0\}$ be the set of admissible controls [11], which is a closed convex set. Choose any $c, v \in \mathcal{C}$ such that $w = v - c$ satisfies $w(0) = 0$. By the differentiability properties of $U$ we have that

$$\frac{1}{\epsilon}(J(c + \epsilon w) - J(c)) = \int_0^\infty \int_{\mathcal{D}} e^{-rt} U'(c) w dx dt + o(\epsilon). \tag{45}$$

Let $z_w$ be the solution of (34) with $v$ replaced by $w$. Then, by the steps in Section A.0.4 it holds that

$$0 = -\int_0^\infty e^{-rt} \int_{\mathcal{D}} wp dx dt. \tag{46}$$

Adding (45) and (46) we obtain

$$\frac{1}{\epsilon}(J(c + \epsilon w) - J(c)) = \int_0^\infty \int_{\mathcal{D}} e^{-rt} (U'(c) - p) w dx dt + o(\epsilon). \tag{47}$$

Assuming that $c \in \mathcal{C}$ is a local maximizer the above yields that

$$\int_0^\infty \int_{\mathcal{D}} e^{-rt} (U'(c) - p)(c - v) dx dt + o(\epsilon) \leq 0, \ \forall \epsilon > 0, \ v \in \mathcal{C},$$

and passing to the limit as $\epsilon \to 0^+$ we obtain

$$\int_0^\infty \int_{\mathcal{D}} e^{-rt} (U'(c) - p)(c - v) dx dt \leq 0, \ \forall \, v \in \mathcal{C}$$

If $c \in \mathcal{C}$ is a maximizer such that $c > 0$ then setting $v = c \pm \epsilon h$ for any function $h$ and $\epsilon > 0$ sufficiently small in the above condition yields $U'(c) = p$, which provides a characterization of the optimal consumption in this case, in terms of the adjoint process $p$.

### A.0.6 Some technical tools

**Theorem A.2** (Implicit function theorem)**.** *Let $X, Y, Z$ be three Banach spaces, $k \geq 1$, $\mathcal{A} \subset X \times Y$ be an open set and let $F : \mathcal{A} \to Z$ be a $C^k$ map. Let $(x_0, y_0) \in \mathcal{A}$ such that $F(x_0, y_0) = 0$.*

*If $D_y F(x_0, y_0) = D(F(x_0, \cdot))(y_0) : Y \times Z$ is a bounded invertible linear transformation (isomorphism) then there is an open neighbourhood $\mathcal{U}_0$ of $X$ such that for all connected open neighbourhoods $\mathcal{U}$ of $x_0$ contained in $\mathcal{U}_0$, there exists a unique continuous function $u : \mathcal{U} \to Y$ such that*

$$u(x_0) = y_0, \ \ (x, u(x)) \in \mathcal{A}, \ \ F(x, u(x)) = 0, \ \forall \, x \in \mathcal{U}.$$

*Moreover, $u$ is $C^k$ and*

$$Du(x) = -D_y F(x, u(x))^{-1} D_x F(x, u(x)), \ \ \forall \, x \in \mathcal{U}.$$

---

[11] If $c$ is considered as a function in a Lebesgue space, the inequality $c \geq 0$ is considered in an almost everywhere sense.

**Theorem A.3** (Green)**.** *Let $\mathcal{D}$ be a sufficiently smooth domain and consider a vector field $\underline{F}$ and a scalar field $g$. Then,*

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

*where by $n(x)$ we denote the outward normal vector at $x \in \partial\mathcal{D}$.*

Naturally sufficient smoothness on $\partial\mathcal{D}$ ( for the outward normal to be well defined ) and on $\underline{F}$ and $g$ (for the divergence and the gradient to be well defined) must be assumed. If

$$\begin{aligned} \text{either} \;\; \underline{F}(x) \cdot n(x) = 0 \;\; \text{on} \;\; \partial\mathcal{D}, \\ \text{or} \;\; g(x) = 0 \;\; \text{on} \;\; \partial\mathcal{D}, \end{aligned}$$

then the surface integral in Green's formula vanishes and the formula simplifies to

$$\int_{\mathcal{D}} \underline{F}(x) \cdot \nabla\, g(x) dx = -\int_{\mathcal{D}} g \nabla \cdot\, \underline{F}(x) dx$$

The classical version of the theorem requires all the above quantities to be $C^1$ with well defined values on the boundary of $\mathcal{D}$ (i.e. continuous up to the boundary). However, Green's theorem in the above form can still be valid in a Sobolev space setting with the derivatives of the vector and the scalar fields understood in the weak sense and the surface integral $\mathcal{D}$ understood in the sense of traces.

# B The Kirchhoff transformation

The system (20) of elliptic equations may in fact be transformed into a more convenient equivalent form.

**Proposition B.1.** *Define the variables $u = \Phi(k^*)$ and $v = \Psi(c^*)$ where $\Phi'(s) = w_1(s)$ and $\Psi = U'$,*[12] *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*

$$\begin{aligned} 0 &= \Delta u + \nabla\, \ln D_0 \cdot \nabla\, u - w_2(\phi(u))\Delta \ln A - w_2(\phi(u))\nabla\, \ln D_0 \cdot \nabla\, \ln A - \\ &\quad (w_2(\phi(u))'\nabla\, u \cdot \nabla\, \ln A + \frac{1}{D_0}\Big(f(\phi(u)) - \delta\phi(u) - \psi(v)\Big), \\ 0 &= \Delta v + \nabla\, \ln D_0 \cdot \nabla\, v + \frac{w_2'(\phi(u))}{w_1(\phi(u))}\nabla\, \ln A \cdot \nabla\, v - \frac{(r + \delta - f'(\phi(u)))}{D_0 w_1(\phi(u))} v, \end{aligned} \tag{48}$$

*with Neumann boundary conditions.*

The proof is straightforward and uses the Kirkhoff transformation (see e.g. Knight and Phillip (1974) ) 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

---

[12] Note that $v = p$.

monotone). Note, however, that even in the absence of the above conditions, system (48) can 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 semilinear form. If at some point $w_1(\phi(u)) = 0$, then the second equation of (48) 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 $B$.

## C Proof of Proposition 4.2

Let $k(0,x) = k_0^* + \epsilon k_1(x)$ and $p(0,x) = p_0^* + \epsilon p_1(x)$ where $(k_0^*, p_0^*)$ is the saddle point of the standard Ramsey model. Recall that for $\epsilon = 0$, this initial condition leads to the solution $(k(t,x), p(t,x)) = (k_0^*, p_0^*)$ for any value of the diffusion term $D$ (as diffusion has no effect in the absense of spatial gradients). We also emphasize that only $k(0,x)$ is given. The initial condition $p(0,x)$ is to be determined so that a solution of (17-18) compatible with (19) can be obtained. This solution will be approximately of the form $(k(t,x), p(t,x)) = (k_0^* + \epsilon k_1(t,x), p_0^* + \epsilon p(t,x))$ with the perturbation solving the linearized version of (17-18) subject to (19).

We now linearize (17-18) around $(k_0^*, p_0^*)$ and use the notation $g = (U')^{-1}$ to obtain the system

$$\begin{aligned} \frac{\partial k_1}{\partial t} &= \nabla \cdot (D_0 w_1(k_0^*)\nabla k_1) + (Af'(k_0^*) - \delta)k_1 - g'(p_0^*)p_1, \\ \frac{\partial p_1}{\partial t} &= \nabla \cdot (D_0 w_1(k_0^*)\nabla p_1) - Af''(k_0^*)p_0 k_1 + (r + \delta - Af'(k_0^*))p_1, \\ & n \cdot \nabla k_1 = n \cdot \nabla p_1 = 0. \ k_1(0,x) = k_1(x), \end{aligned} \tag{49}$$

Since $p = U'(c)$ by the implicit function theorem we obtain $g'(p_0^*) = \frac{1}{U''(c_0^*)}$ where $p_0^* = U'(c_0^*)$, while we note that at the steady state $r + \delta - Af'(k_0^*) = 0$. Hence, (49) simplifies to

$$\begin{aligned} \frac{\partial k_1}{\partial t} &= \nabla \cdot (D_0 w_1(k_0^*)\nabla k_1) + rk_1 - \frac{1}{U''(c_0^*)}p_1, \\ \frac{\partial p_1}{\partial t} &= \nabla \cdot (D_0 w_1(k_0^*)\nabla p_1) - Af''(k_0^*)p_0 k_1, \\ & n \cdot \nabla k_1 = n \cdot \nabla p_1 = 0. \ k_1(0,x) = k_1(x). \end{aligned} \tag{50}$$

The transversality condition (19) for this solution becomes

$$e^{-rt}\int_{\mathcal{D}} (k_0^* + \epsilon k_1(t,x))(p_0^* + \epsilon p_1(t,x))dx \to 0, \ \text{ as } \ t \to \infty$$

Since this is an asymptotic in time condition we do not linearize it. For growing perturbations the dominant part of this condition is

$$\lim_{t\to\infty} e^{-rt} \int_{\mathcal{D}} k_1(t,x) p_1(t,x) dx = 0.$$

We make one final transformation for the sake of convenience and economic interpretation. We define the new variable $c_1(t,x) = \frac{1}{U''(c_0^*)} p_1(t,x)$ so that recalling that $p_0^* = U'(c_0^*)$, and noting that $c_0^* = U^{-1}(p_0^*)$ is a finite constant, in terms of the new variables $(k_1, c_1)$ we bring (50) in the form

$$\begin{aligned} \frac{\partial k_1}{\partial t} &= \nabla \cdot (D_0 w_1(k_0^*) \nabla k_1) + r k_1 - c_1, \\ \frac{\partial c_1}{\partial t} &= \nabla \cdot (D_0 w_1(k_0^*) \nabla c_1) - A f''(k_0^*) \frac{U'(c_0^*)}{U''(c_0^*)} k_1, \\ n \cdot \nabla k_1 &= n \cdot \nabla c_1 = 0.\ k_1(0,x) = k_1(x). \end{aligned} \tag{51}$$

In terms of the same transformation (and since $U''(c_0^*)$ is a finite constant bounded away from zero) we also get the transversality condition in the form

$$\lim_{t\to\infty} e^{-rt} \int_{\mathcal{D}} k_1(t,x) c_1(t,x) dx = 0. \tag{52}$$

System (51) with the transversality condition (52) will be the starting point for our investigation.

To make its study easier we simplify the notation by using $k(t,x) = k_1(t,x)$ and $c(t,x) = c_1(t,x)$, then define the vectors

$$z(t,x) := (z^{(1)}(t,x), z^{(2)}(t,x))^T = (k(t,x), c(t,x))^T, \quad z^* := (z^{(1)*}, z^{(2)*})^T = (k_0^*, c_0^*)^T,$$

the constant

$$M := A f''(k_0^*) \frac{U'(c_0^*)}{U''(c_0^*)} \geq 0,$$

and

$$\mathbb{D}(z^*) = \begin{pmatrix} D(k_0^*)\Delta & 0 \\ 0 & -D(k_0^*)\Delta \end{pmatrix}, \quad L_F(z^*) = \begin{pmatrix} r & -1 \\ -M & 0 \end{pmatrix},$$

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

$$D(k_0^*) = D_0 w_1(k_0^*), \quad D_0 = \bar{D}_0 A B$$

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

In terms of the above express (51) as the constant coefficient PDE system

$$\frac{\partial z}{\partial t} = \mathbb{D}(z^*) z + L_F(z^*) z, \tag{53}$$

subject to the transversality condition

$$\lim_{t\to\infty} e^{-rt}\int_{\mathcal{D}} z^{(1)}(t,x)z^{(2)}(t,x)dx \to 0 \tag{54}$$

Note that in the case of linear diffusion, $D(k_0^*)$ must be replaced by $D_0$ in the above. 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 $\mathbb{R}^2$.

Since (53) is a constant coefficient problem in a bounded domain, we can 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

$$\begin{aligned} -\Delta\phi_n &= \mu_n\phi_n, \\ \nabla\phi_n \cdot n &= 0. \end{aligned} \tag{55}$$

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 $L^2(\mathcal{D})$.[13] The spectrum has the property $\mu_n \geq 0$, with $\mu_n \to \infty$ and, for many cases of interest, both $\mu_n$ and the eigenfunctions $\phi_n$ are known exactly in analytic form [14].

Any solution of (53) can be expanded as

$$z(t,x) = \sum_{n\in\mathbb{N}} z_n(t)\phi_n(x),$$

with

$$z_n(t) = (z_n^{(1)}(t), z_n^{(2)}(t))^T = (k_n(t), c_n(t))^T$$

where

$$k_n(t) = \langle k(t,\cdot), \phi_n(\cdot)\rangle, \;\; c_n(t) = \langle c(t,\cdot), \phi_n(\cdot)\rangle,$$

and a similar expansion for the initial condition.

By substituting this expansion in (53) and using the orthogonality properties of the basis of eigenfunctions, the system of PDEs (53) reduces to an equivalent countable system of ODEs of the

---

[13] 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.

[14] Note that the zero eigenvalue for the Laplace operator with homogeneous Neumann boundary conditions corresponds to a constant eigenfunction. Moroever, as we can see from the subsequent analysis this eigenvalue will not lead to instability, so it is not considered in our analysis.

form

$$\begin{aligned} z_n' &= L_n z_n, \;\; n \in \mathbb{N}, \\ z_n^{(1)}(0) &= z_{n,0}, \end{aligned} \tag{56}$$

where

$$L_n = \begin{pmatrix} -D(k_0^*)\mu_n & 0 \\ 0 & D(k_0^*)\mu_n \end{pmatrix} + L_F. \tag{57}$$

Moreover, using the orthogonality properties of the eigenfunction the transversality condition (54) becomes

$$\lim_{t\to\infty} e^{-rt} z_n^{(1)}(t) z_n^{(2)}(t) = \lim_{t\to\infty} e^{-rt} k_n(t) c_n(t) \to 0, \;\; \forall\, n \in \mathbb{N} \tag{58}$$

The solution to system (56) 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

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

hence leading to a solution of the form

$$z(t,x) = \sum_{n\in\mathbb{N}} z_n(t)\phi_n(x) = \sum_{n\in\mathbb{N}} e^{tL_n} z_{n,0}\phi_n(x), \tag{59}$$

Note that the solution is fully determined as long as the full vector of initial conditions $z_{n,0}$ is known for all $n$. However, here, for each $n \in \mathbb{N}$, we only have one component of the vector $z_{n,0}$ the one corresponding to $z_{n,0}^{(1)} = k_{n,0}$. We are going to choose the second one $z_{n,0}^{(2)} = p_{n,0}$ in such a way so that the solution to (56) satisfies the transversality condition (58). For this choice of $z_{n,0}^{(2)}$ the Fourier series (59) will be a solution of (53) satisfying the transversality condition (54) hence a feasible perturbation of the optimal solution $(k_0^*, p_0^*)$. If this $z(t,x)$ does not display any spatial dependence then the initial perturbation of the optimal solution will die out eventually and the system will equilibrate to the original optimal solution (as would happen for a perturbation along the stable manifold of the spatially independent Ramsey system). If on the other hand this $z(t,x)$ displays a genuine spatial dependence then this can be understood as a precursor instability that will generate spatially dependent patterns compatible with the transversality condition.

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 if 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 its 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. $\rho_{1,n} = a_n + i b_n$, $\rho_{1,n} = a_n - i b_n$, complex conjugate roots,

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

We will concentrate in case A and make the following observation. Let $z_0 = (k_n(0), c_n(0))^T$ with $k_n(0)$ known and $p_n(0)$ to be determined. Then

$$(z_n^{(1)}(t), z_n^{(2)}(t))^T = (p_n(t), z_n(t))^T = (a_{11}^{(n)} e^{\rho_{1,n} t} + a_{12}^{(n)} e^{\rho_{2,n} t}, a_{21}^{(n)} e^{\rho_{1,n} t} + a_{22}^{(n)} e^{\rho_{2,n} t})^T,$$

with

$$\begin{aligned}
a_{11}^{(n)} &= \frac{1}{\rho_{1,n} - \rho_{2,n}} (L_n - \rho_{2,n} I) z_n(0) \cdot e_1, \\
a_{12}^{(n)} &= \frac{1}{\rho_{1,n} - \rho_{2,n}} (L_n - \rho_{2,n} I) z_n(0) \cdot e_2, \\
a_{21}^{(n)} &= \frac{1}{\rho_{1,n} - \rho_{2,n}} (L_n - \rho_{1,n} I) z_n(0) \ \cdot e_1, \\
a_{22}^{(n)} &= \frac{1}{\rho_{1,n} - \rho_{2,n}} (L_n - \rho_{1,n} I) z_n(0) \ \cdot e_2.
\end{aligned} \tag{60}$$

Note that $a_{ij}^{(n)}$ depend linearly on the unknown initial condition for the adjoint variable $z_n^{(2)}(0) = c_n(0)$.

Then the tranversality condition becomes

$$\begin{aligned}
&\lim_{t\to\infty} e^{-rt} z_n^{(1)}(t), z_n^{(2)}(t) = \lim_{t\to\infty} e^{-rt} p_n(t) c_n(t) \qquad (61) \\
&= \lim_{t\to\infty} \Big( a_{11}^{(n)} a_{21}^{(n)} e^{(-r+2\rho_{1,n})t} + (a_{11}^{(n)} a_{22}^{(n)} + a_{12}^{(n)} a_{21}^{(n)}) e^{(-r+\rho_{1,n}+\rho_{2,n})t} + a_{12}^{(n)} a_{22}^{(n)} e^{(-r+2\rho_{n,2})t} \Big) = 0, \ \ \forall n \in \mathbb{N}
\end{aligned}$$

The eigenvalues of the matrix $L_n$ can readily be calculated in terms of the roots of the quadratic equation

$$\rho^2 - r\rho + \{D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M\} = 0,$$

which can easily be computed as

$$\rho_{j,n} = \frac{r}{2} + \frac{(-1)^j}{2} \sqrt{r^2 - 4\,[D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M]}, \ \ j = 1, 2.$$

and observe that, as long as these roots are real [15],

$$\rho_{1,n} < \frac{r}{2} < \rho_{2,n},$$
$$\rho_{1,n} + \rho_{2,n} = r$$

so that the only solution compatible with the tranversality condition (61) is the one such that

$$a_{11}^{(n)} a_{22}^{(n)} + a_{12}^{(n)} a_{21}^{(n)} = 0,$$
$$a_{12}^{(n)} a_{22}^{(n)} = 0,$$

which leads to choosing

$$a_{12}^{(n)} = a_{22}^{(n)} = 0.$$

Recalling the definition of $a_{ij}^{(n)}$ this implies that

$$\begin{aligned} \frac{1}{\rho_{1,n} - \rho_{2,n}} (L_n - \rho_{2,n} I) z_n(0) \cdot e_2 = 0, \\ \frac{1}{\rho_{1,n} - \rho_{2,n}} (L_n - \rho_{1,n} I) z_n(0) \cdot e_2 = 0, \end{aligned} \tag{62}$$

which since $z_n^{(1)}(0)$ is known provides a linear system for the determination of the unknown initial condition $z_n^{(2)}(0)$ such that the transversality condition holds. We do not proceed with this task, here, however this is an easy task (and was implemented in the numerical investigation presented in this paper).

Upon solving (62) for $z_n^{(2)}(0)$ we have now fully determined the initial condition

$$z_n(0) = (z_n^{(1)}(0), z_n^{(2)}(0))^T$$

and return to the definition of the remaining coefficients $a_{11}^{(n)}$ and $a_{21}^{(n)}$ using (60) with now $z_n(0)$ fully specified. We calculate these coefficients for this initial condition, and call them $\hat{a}_{11}^{(n)}$ and $\hat{a}_{21}^{(n)}$ respectively.

The solution of (56) compatible with the transversality condition (54) is therefore,

$$z(t, x) = \sum_{n \in \mathbb{N}} e^{\rho_{1,n} t} \begin{pmatrix} \hat{a}_{11}^{(n)} \\ \hat{a}_{21}^{(n)} \end{pmatrix} \phi_n(x). \tag{63}$$

This solution will be spatially dependent in the long run ($t \to \infty$) if at least one of the $\rho_{1,n} > 0$, otherwise $z(t, x) \to 0$ as $t \to \infty$ and the perturbation will relax to the optimal state $(k_0^*, p_0^*)$. If $\rho_{i,n} > 0$

---

[15] The condition for the roots being real is easily seen to be $[D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M] \leq \frac{r^2}{4}$. Setting $x_n = D(k_0^*)\mu_n$ this condition becomes $-x_n^2 + r x_n - (M + \frac{r^2}{4}) \leq 0$, which holds for every $x_n$ as long as $M > 0$, since in this case the quadratic has equation $-x_n^2 + r x_n - (M + \frac{r^2}{4}) \leq 0$ has no real root.

for some $n \in \mathcal{N} \subset \mathbb{N}$ then the asymptotic spatial behaviour of the perturbation will be given by

$$z(t,x) = \sum_{n \in \mathcal{N}} e^{\rho_{1,n} t} \begin{pmatrix} \hat{a}_{11}^{(n)} \\ \hat{a}_{21}^{(n)} \end{pmatrix} \phi_n(x), \qquad (64)$$
$$\mathcal{N} = \{n \in \mathbb{N} \ : \ \rho_{1,n} > 0\}$$

We now carefully examine the eigenvalues $\rho_{1,n}$ as a function of the diffusion coefficient $D(k_0^*)$ and the parameter $M$. Recall that

$$\rho_{1,n} = \frac{r}{2} - \frac{1}{2}\sqrt{r^2 - 4\left[D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M\right]} \qquad (65)$$

1. Since $D(k_0^*) = \bar{D}_0 B w_1(k_0^*)$, if there is no capital mobility ($B = 0$) then we have that

   $$\rho_{1,n} = \frac{r}{2} - \frac{1}{2}\sqrt{r^2 + 4M} < 0, \ \ \forall \, n \in \mathbb{N}.$$

   That means, in the absense of capital mobility effects the solution provided in (64) has the property $z(t,x) \to 0$ as $t \to \infty$.

2. Now let us consider the effect of increasing the mobility factor $B$ on $\rho_{1,n}$. Note that

   $$\frac{d}{dB}\rho_{1,n} = \frac{(r\bar{D}_0 w_1(k_0^*)\mu_n - 2\bar{D}_0^2 B w_1^2(k_0^*)\mu_n^2)}{2\sqrt{r^2 - 4\left[D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M\right]}}, \qquad (66)$$

   so that

   $$\left.\frac{d}{dB}\rho_{1,n}\right|_{B=0} \geq 0,$$

   implying that $\rho_{1,n}$ increases as $B$ increases for small values of $B$. This means that for sufficiently large values of the capital mobility certain of these $\rho_{1,n}$ may increase to values that are large enough so that $\rho_{1,n}$ and give rise to a spatial mode.

3. On the other hand, as (66) indicates, the eigenvalues $\rho_{1,n}$ are not monotone in $B$ for all values of $B$. So even though (66) clearly indicates that spatial structures can be initiated only where capital mobility is present, a more detailed analysis is required to answer the question: Given a fixed capital mobility $B$, i.e. $D(k_0^*) = \bar{D}_0 B w_1(k_0^*) \neq 0$ which spatial structures among the possible ones indicated in (64) will actually emerge.

   To answer this question given a fixed $n \in \mathbb{N}$ and a fixed $D(k_0^*)$ we try to find conditions such that the corresponding $0 < \rho_{1,n} < \frac{r}{2}$ (as required by the tranversality condition). Simple algebraic manipulations show that

   $$0 < \rho_{1,n} < \frac{r}{2} \quad \text{if} \quad 0 < D(k_0^*)\mu_n(r - D(k_0^*)\mu_n) - M,$$

   with the condition on $\mu_n$ guaranteeing that $\rho_{1,n} > 0$. Note that as mentioned in footnote 15 $\rho_{1,n} < \frac{r}{2}$ always holds when the roots are real, which in turn is always true if $M \geq 0$ (as is the case here). Defining $x_n = D(k_0^*)\mu_n$ to simplify the notation, we see that the condition for pattern formation reduces to

$$x_n^2 - r x_n + M < 0,$$

which is never true if the above quadratic polynomial has no real roots, i.e. where $r^2 - 4M < 0$, whereas if there are real roots $\lambda_1, \lambda_2$, it holds for $\lambda_1 < x_n < \lambda_2$.

Hence, the important observation is

1. If $r^2 < 4M$ then $\rho_{1,n} < 0$ for all $n \in \mathbb{N}$ no pattern emerges.
2. If $r^2 > 4M$ and for a given $n \in \mathbb{N}$ we may get $0 < \rho_{1,n} < \frac{r}{2}$ as long as

   $$\frac{1}{2}(r - \sqrt{r^2 - 4M}) \leq D(k_0^*)\mu_n \leq \frac{1}{2}(r + \sqrt{r^2 - 4M}), \tag{67}$$

   and the corresponding modes will form a spatial pattern compatible with the transversality condition.

Note that mode will not be activated if $D(k_0^*)$ is below a critical level, however, on the other hand too large values of $D(k_0^*)$ will kill the emergent pattern since the regularizing effects of diffusion in this case are too large and lead to dissipation of emergent spatial gradients.

# D Proof of Propositions 4.7 and 4.9

In this appendix we use the simplified notation $|\cdot|$ for the Euclidean norm $\|\cdot\|$.

The solution of (17)-(18-19) in the case where $\bar{D}_0 = 0$ (no capital mobility) will be denoted by $k_0$ and 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

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

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 straightforward calculation yields

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

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 straightforward calculations, we obtain that

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

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

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

We can now calculate the quantity $\int_{\mathcal{D}} \nabla k_0 \cdot \nabla k_1 dx$ in terms of 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

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

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

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

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, Princeton, NJ.

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

Allen, T. and C. Arkolakis (2014) Trade and the topography of the spatial economy. *Quarterly Journal of Economics*, 129, 1085-1139.

Arrow, K. and M. Kurz (1970), *Public investement the rate of return and optimal fiscal policy*, The Johns Hopkins University Press, Baltimore.

Aseev, S. M., and Kryazhimskiy, A. V. (2004). The Pontryagin maximum principle and transversality conditions for a class of optimal control problems with infinite time horizons. *SIAM Journal on Control and Optimization*, 43(3), 1094-1119.

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.

Ballestra, L. V. (2016). The spatial AK model and the Pontryagin maximum principle. *Journal of Mathematical Economics*, 67, 87-94.

Barro, R. J. and X. Sala-i-Martin (2004). *Economic Growth*, second edition. MIT Press, Cambridge, MA.

Benveniste, L.M. and Scheinkman, J.A., (1982). Duality theory for dynamic optimization models of economics: The continuous time case. *Journal of Economic Theory*, 27(1), 1–19

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, and S. Federico (2016). Optimal economic growth through capital accumulation in a spatially heterogeneous environment. AMSE Working Papers 1641, Aix-Marseille School of Economics, Marseille, France.

Boucekkine, R., G. Fabbri, S. Federico, and F. Gozzi (2019). Growth and agglomeration in the heterogeneous space: A generalized AK approach. *Journal of Economic Geography* 19 (6), 1287–1318.

Breinlich, H., G. I. Ottaviano, and J. R. Temple (2014). Regional growth and regional decline. In *Handbook of Economic Growth*, Volume 2, Aghion, P. and S. Durlauf (eds.), 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. *Journal 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.

Calvia, A., Federico, S., and Gozzi, F. (2021). State constrained control problems in Banach lattices and applications. SIAM Journal on Control and Optimization, 59(6), 4481-4510.

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

Camacho, C. B. Zhou and M. Briani (2008). On the dynamics of capital accumulation across space. *European Journal of Operations Research* 186, 451–465.

Casas, E. and Chrysafinos, K. (2018). Analysis and optimal control of some quasilinear parabolic equations. Mathematical Control & Related Fields, 8, 607.

Conley, T. G. and E. Ligon (2002). Economic distance and cross-country spillovers. *Journal of Economic Growth* 7, 157–187.

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.

Ekeland, I. (2010). From Frank Ramsey to René Thom: a classical problem in the calculus of variations leading to an implicit differential equation. *Discrete & Continuous Dynamical Systems*, 28(3), 1101.

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

Faggian, S., Gozzi, F. and Kort, P. M. (2021). Optimal investment with vintage capital: Equilibrium distributions. Journal of Mathematical Economics, 96, 102516.

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.

Kartala, X. I., Englezos, N., and Yannacopoulos, A. N. (2020). Future Expectations Modeling, Random Coefficient Forward?Backward Stochastic Differential Equations, and Stochastic Viscosity Solutions.*Mathematics of Operations Research*, 45(2), 403-433.

Knight, J. H. and Philip, J. R. (1974). Exact solutions in nonlinear diffusion.*Journal of Engineering*

*Mathematics*, 8(3), 219-227.

Kravvaritis, D. and A. N. Yannacopoulos (2020). *Variational Methods in Nonlinear Analysis: with applications in optimization and partial differential equations*, De Gruyter.

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

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

Ladyzhenskaia, O. A., Solonnikov, V. A., and Uralt́seva, N. N. (1988). Linear and quasi-linear equations of parabolic type (Vol. 23). American Mathematical Soc..

Lykina, V., and Pickenhain, S. (2017). Weighted functional spaces approach in infinite horizon optimal control problems: a systematic analysis of hidden opportunities and advantages. *Journal of Mathematical Analysis and Applications*, 454(1), 195-218.

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.

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.

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

Peng, S., and Shi, Y. (1999). Infinite horizon boundary value problems and applications.*Journal of Differential Equations*, 155(2), 405-422.

Santambrogio, F., Xepapadeas, A., and Yannacopoulos, A. N. (2020). Rational expectations equilibria in a Ramsey model of optimal growth with non-local spatial externalities. Journal de Mathématiques Pures et Appliquées, 140, 259-279.

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

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

Van, C.L., Boucekkine, R. and Saglam, C., 2007. Optimal control in infinite horizon problems: a Sobolev space approach. Economic Theory, 32(3), pp.497-509.

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

Yannacopoulos, A. N. (2008). Rational expectations models: An approach using forward–backward stochastic differential equations. Journal of Mathematical Economics, 44(3-4), 251-276.