DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Spatial Growth with Exogenous Saving Rates

Anastasios Xepapadeas

Athanasios Yannacopoulos

Working Paper Series

15-14

October 2015

Spatial Growth with Exogenous Saving Rates*

A. Xepapadeas and A. Yannacopoulos
Athens University of Economics and Business

September 22, 2015

Abstract

Economic growth has traditionally been analyzed in the temporal domain, while the spatial dimension is captured by cross-country income differences. Data suggest great inequality in income per capita across countries, and a slight but noticeable increase in inequality across nations (Acemoglu 2009). Seeking to explore the mechanism underlying the temporal evolution of the cross sectional distribution of economies, we develop a spatial growth model where saving rates are exogenous. Capital movements across locations are governed by a mechanism under which capital moves towards locations of relatively higher marginal productivity, with a velocity determined by the existing stock of capital. This mechanism leads to a capital accumulation equation augmented by a nonlinear diffusion term, which characterizes spatial movements. Our results suggest that under diminishing returns the growth process leads to a stable spatially non-homogenous distribution for per capita capital and income in the long run. Insufficient savings may lead to the emergence of persistent poverty cores where capital stock is depleted in some locations.

Keywords: Economic growth, space, capital flows, nonlinear diffusion, Solow model, steady state distributions, stability.

JEL Classification: O4, R1, C6

*We would like to thank William Brock for many challenging ideas and suggestions.This research has been co-financed by the European Social Fund – and Greek national funds through the Research Funding Program: Excellence(ARISTEIA) – AUEB "Spatiotemporal Dynamics in Economics."

1 Introduction

Economic growth, in formal growth models, has traditionally been analyzed in the temporal domain with the main focus of analysis being the development of models capable of explaining stylized facts, which are expressed in terms of the temporal evolution of key variables such as output or capital per capita or the capital labor ratio. A central issue, however, is cross-country income differences which exemplifies the spatial dimension of the problem. Acemoglu [1] (Chapter 1), using data on GDP per capita and per worker (in logs) since 1960, points out that there is great inequality in income per capita and income per worker across countries, and that there is a slight but noticable increase in inequality across nations. The geographical or spatial dimension is also taken into accounting the context of convergence. Data suggest (e.g. Acemoglu [1]) that there is no unconditional convergence during the post war perior. However, Barro and Sala-i-Martin [7] results suggest that conditional convergence takes place with poor countries growing faster in terms of per capita GDP than rich ones within a group that shares similar characteristics. Conditional convergence even at different steady states may not however adequately describe the evolution of the spatial distribution of per capita GDP across countries. As Quah [34, 36, 35] points out "Convergence concerns poor economies catching up with rich ones. 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."

Some insights into the characteristics of the spatial distribution of GDP per capita can be obtained by using the quantity

$$D_t = \sum_{i \neq j} \left(\frac{y_{it} - y_{jt}}{\bar{y}} \right)^2, \quad j = 1, \ldots N, \quad t = 1950, \ldots 2007$$

where y_{it}, y_{jt} denotes per capita GDP in countries i, j at time t for a sample of countries $i, j = 1, ..., N$, and $\bar{y}$ denotes the overall average (over all countries) per capita GDP. This quantity can be regarded as a measure of spatial inhomogeneity of GDP per capita, in the sense that an increasing D_t over time means that the spatial distribution of GDP becomes more spatially heterogenous or "less flat" relative to space.[1] Thus an increasing D_t over time indicates that the dispersal of per capita GDP across the countries of the sample was increased during the sample period. The inhomogeneity measure D_t, along with the corresponding linear trend is presented in the

[1]This measure can be related to a discretized version of a Sobolev norm.

figures below for eleven regions of the world[2] (Figure 1), and high income countries[3] (Figure 2), covering the period 1980–2011.

[Figure 1. Regional inhomogeneity measure]
[Figure 2. Inhomogeneity measure, high income countries]

The evolution of the inhomogeneity measure, and the associated linear trend, suggests that the overall dispersal is rather increasing both at the regional level, and within the group of high income countries. These observations although broad in nature, indicate that the spatial distribution of GDP per capita does not tend to become more uniform with the passage of time, or to put it differently does not seem to converge to a geographical homogenous state for countries grouped in the traditional way according to the level of their per capita GDP. Countries that start with lower per capital income in the region may growth faster than high income counties, which is consistent with β convergence arguments, but this growth does not seem to result in a spatially flatter distribution in the long run.

In this context the purpose of this paper is to develop a spatial model of economic growth and by doing so to explore mechanisms that could generate, through economic forces, persistent non uniform spatial distributions of per capita capital and GDP across locations, and determine the temporal evolution of these spatial distributions. In a sense we are exploring how traditional neoclassical growth theory can be extended to a spatial growth theory which would provide models capable of approximating persistent spatial heterogeneity across countries in terms of per capita GDP.

Economic geography and economic growth has been discussed in the so-called second generation of new economic geography models but not in a formal growth context (e.g.,Martin and Ottaviano [31], Baldwin et al. [6], [4], Baldwin and Martin [5], Fujita and Mori [24], Desmet and Rossi-Hansberg [22], [23]). Models of optimal development over space and time, which could

[2]For the GDP per capita (GDP per capita, PPP constant 2005 international $) the World Bank data base was used. The regions according to the World Bank classification are: Arab World, ARB; Caribbean small states, CSS; East Asia & Pacific (developing only), EAP; European Union, EUU; Europe & Central Asia (developing only), ECA; Latin America & Caribbean (developing only), LAC; Middle East & North Africa (developing only), MNA; North America, NAC; Pacific Island Small states, PSS; South Asia, SAS; Sub-Saharan Africa (developing only), SSA.

[3]The group of high income countries includes Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Estonia, Finland, France,Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Rep., Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States.

be regarded as a suitable vehicle for studying economic growth in a geographical context, were developed in the 1970s by Isard and Liossatos (e.g., [28], [27], [26], Carlson et al. [19]). Dynamic spatial economic modeling were developed in the context of economic growth and resource management mainly during the 2000s (e.g.Brito [12], Camacho and Zou [17], Boucekkine at al. [10], [8] [9] Brock and Xepapadeas [13], [14], Brock et al. [15], [16]). The main feature of current spatial growth models is that the spatial movements of the stock of capital across locations are modeled through a trade balance approach with respect to a closed region where capital flows are such that capital is received from the left of the region and flows away to the right of the region. This leads to a model of classic local diffusion with a constant diffusion coefficient. Modeling capital movements this way implies that capital stock moves from locations of high concentration to locations of low concentration. This property although consistent with diminishing returns to capital, since high concentration imply low marginal productivity and vise versa, seems not to be compatible with empirical findings. As indicated by Lucas in the context of the Lucas paradox ([29, 30]) although diminishing returns suggest that capital will flow from locations of high concentration to locations of low concentration, this is not happening in reality.

In the present paper we contribute to the ongoing research on spatiotemporal dynamics and spatial growth by developing a model where the basic mechanism underlying the movements of capital across space is the quest for locations where the marginal productivity of capital is relatively higher than the productivity at the location of origin, without imposing the constraint that capital moves from locations of high concentration to locations of low concentration. By assuming that capital flows towards locations of high returns, which is a plausible assumption underlying capital flows with endogenous velocity depending the existing stock of capital, our model implies that the spatiotemporal evolution of capital is governed by a nonlinear diffusion equation. In this case the "diffusion coefficient" is not constant but depends on the capital stock and the rate of change of marginal productivity of capital (the second derivative of the production function). This approach for modeling capital flows essentially differs from the classic diffusion models used in the existing literature which are based on the trade balance (e.g., Carlson et al. [19], Brito [12], Camacho and Zou [17], Boucekkine at al. [10], [8] [9]), and describe the spatiotemporal evolution of capital by a parabolic partial differential equation with constant diffusion coefficient.

Our contribution is that by using the plausible mechanism that capital moves towards locations of higher productivity, and not a mechanism where capital moves necessarily from higher to lower concentrations, we obtain using standard neoclassical growth assumptions spatial distributions for per capita capital and GDP which are characterized by large and persistent spatial inhomogeneities. These inhomogeneities could be regarded as compatible with existing observations. Furthermore, we are not confronted with Lucas paradox since our approach is based on the notion that capital moves to location of relatively higher productivity, but not necessarily from locations of high concentration to location of low concentration. The notion of capital we employ is a "mechanistic" kind that cannot move very fast, like financial capital, to areas of high marginal productivity because of adjustment costs and other potentially institutional barriers in this location.

By considering a distance metric concept based on economic distance we develop local models of capital diffusion and we develop an analytical framework that extend the standard Solow model in a geographical context. The spatial Solow model with a mechanism underlying capital flows which leads to nonlinear diffusion, generate solutions in which spatially nonhomogeneous distributions of per capita capital and income across locations persists over time. In certain cases spatial inhomogeneity may be amplified over time and locations may end up at a steady state in poverty cores with capital stock approaching zero. Our results about persistent spatial heterogeneity and non smoothing of spatial differences do not require increasing returns and are obtained under standard diminishing returns to capital.

2 Capital Flows and Distance Metrics

An issue that a spatial growth model should address is the topology of the space in which capital flows take place and the definition of an appropriate distance metric. The most common metric of the distance between two spatial points (say countries) where capital flows take place is geographical distance, as measured for example by the distance between capital cities. Conley and Ligon [20] suggest 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. They use United Parcel Service (UPS) distance as a proxy for transportation cost associated with physical capital, while airfare distance is used as a proxy for transportation cost associated with human capital. It turns out that the distance between countries might be very different depending on whether the geographic or the economic metric is used. For example while the geographical distance between Australia and Egypt is smaller than the distance Australia-UK and Australia-USA, the corresponding economic distance both in terms of UPS and airfare distance between Australia and Egypt is larger than the distance Australia-UK and Australia-USA.

The choice of the distance metric is important for modeling purposes since it provides a basis for choosing between a local model of capital diffusion, or non-local model of capital flows which will incorporate long range effects. If an economic metric is adopted, a local model might be regarded as adequate. This is because it is reasonable to assume that capital, given the restrictions imposed by technology and institutions, will flow among sites which are close in terms of the economic metric, since this would imply less frictions, with the flow directed towards sites where returns grow faster On the other hand if the geographic metric is used then a non-local model seems to be the most appropriate, since in this case the geographical distance might not be the good proxy for frictions associated with capital flows. In this case capital will flow again towards sites where returns grow faster, but these locations might not be close to each other in terms of the geographical metric, which means that a nonlocal model of spatial interactions is required.

To provide a picture of a potential shape of the distribution of GDP per capita in terms of an economic space we ordered in Figure 3 the high income countries so that the country with the highest average GDP per capita was in the middle locations of the space, while the rest of the countries where placed symmetrically on either sides of the middle location in a descending average GDP per capita order. This ordering could be interpreted as characterizing economic distance in term of GDP per capita differences. Small differences imply that countries are close in terms of the economic metric. Although this ordering is arbitrary it provides a bell shaped distribution, that does not become flatter with time as the quantity D_t and figure 2 indicate.

[Figure 3. Distribution of GDP per capita in economic space}

3 Modeling the Spatiotemporal Evolution of Capital

Following the previous discussion, we develop a local model that enable us to study the spatiotemporal evolution of capital in the context of an economic metric. Since each elements of the economic space can be mapped to one and only one element of the geographical space, any spatial distribution defined in economic space can be transformed to a corresponding distribution in the geographical space. This equivalence allows us to work with local models defined in the economic space. In these models the movement of capital to sites where returns are higher can be defined in a more tractable way through local transport operators, an approach which is not appropriate when capital flows are defined in the geographical space.

In what follows the "spatial" variable z can be considered as describing a point in a generalized notion of economic space. We will allow z to take values in a domain $U \subset \mathbb{R}^d$, of sufficiently smooth boundary $\Gamma = \partial U$ (for most applications a Lipschitz boundary is sufficient) where d is the dimension of space.

Restricting attention, momentarily, to a given location we assume that per capita aggregate production, y, depends on the per capita aggregate capital stock, k, at the location, though a neoclassical production function $f : \mathbb{R}_+ \to \mathbb{R}_+$, as $y = f(k, A)$, where A is a productivity factor characterizing the location. Our implicit assumption is that labour is immobile.[4] A location z, is characterized by an exogenous saving ratio $0 < s < 1$ and exponential capital depreciation at a given rate $\delta > 0$.

Since our main interest is the spatial allocation of per capita capital, we define a function $k : [0, T] \times U \to \mathbb{R}_+$ that describes the spatiotemporal distribution of per capita capital as well as a function $y : [0, T] \times U \to \mathbb{R}_+$ that describes the spatiotemporal distribution of per capita aggregate output. At any time t and spatial location z these are related through the production function f by

$$y(t, z) = A(z)f(k(t, z)). \tag{1}$$

The production function is assumed to be twice differentiable on $(0, \infty)$ and concave. The factor A takes into account spatial heterogeneities related to productivity that may reflect positive spatial externalities associated with location z.

The basic assumption of this paper is that aggregate capital stock is produced locally at location z through (1) but at the same time it could be moved out of z to locations z', or flow into z from locations z'' though a transport mechanism. The transport mechanism prescribes that capital moves towards locations where its marginal productivity, m, is higher than the location of origin.

The transport mechanism can be modelled by considering the local balance of the distribution of capital in any region V which is well included in U i.e. all the points in V are interior points of U and not boundary points. We consider the balance of capital in this region, in the sense that the temporal rate of change of the total capital accumulated in V will be equal to capital formation within V, plus capital inflow in V from neighbouring regions, minus capital outflow from V to neighbouring regions. This is simply a bookkeeping equation for the balance of capital stock in V.

[4]This assumption may be easily revisited by including labour mobility in the model, althought it seems justified taking into account the relative difficulty of labour mobility, as compared to capital mobility.

The inflow to or the outflow of capital from V is modeled through a vector field $\mathfrak{v} : [0,T] \times V \to \mathbb{R}^d$, such that $J := \mathfrak{v}k$ at $(t,z) \in [0,T] \times V$ provides the velocity (flux) of capital stock at location z at time t. In coordinate form $J(t,z) = (J_1(t,z), \cdots, J_d(t,z)$, which means that the velocity of capital motion is decomposed along the d coordinates that are needed to specify location, $z = (z_1, \cdots, z_d)$, and $J_i(t,z)$ corresponds to the component of capital velocity along this direction. The introduction of the auxiliary velocity field $\mathfrak{v} = (v_1, \cdots, v_d)$, reflects the fact, that only a fraction of the capital at location t will relocate, so expressing $J_i(t,z) = v_i(t,z)k(t,z)$, captures this fact. In a sense, $v_i(t,z)$ can be interpreted as the propensity of capital stock at location z at time t to move along direction i.

Assumption 1 *Consider a direction, represented by the direction vector $e = (e_1, \cdots, e_d)$, and fix a time t. The tendency of capital accumulated at location z to move along direction e, depends on the spatial rate of change of the marginal productivity of capital $m(t,z) = \frac{\partial y}{\partial k}(t,z)$ and the stock of capital $k(t,z)$ accumulated at z, as*

$$\mathfrak{v}(t,z) \cdot e = B(z)\psi(k(t,z))\nabla_z m(t,z) \cdot e = B(z)\psi(k(t,z)) \sum_{i=1}^{d} \frac{\partial}{\partial z_i} m(t,z) e_i,$$

where $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is a function reflecting the assumption that the tendency of capital to move from location z may depend on the existing stock of capital at z and $B : U \to \mathbb{R}_+$ is a function modelling specific location characteristics.

If ψ is an increasing function of k then an increase in the stock of capital at z increase the tendency of capital to move to another location, provided that it can attain a higher marginal productivity at the mew location. The opposite holds if ψ is a decreasing function.[5]

Proposition 1 *If capital movements follow Assumption 1, then the spatiotemporal evolution of the capital stock when saving rates $s(z)$ are exogenous across locations, is given in terms of the solution of the nonlinear diffusion equation*

$$\frac{\partial k}{\partial t} = -\mathrm{div}\left(B\psi(k)k \left(Af''(k)\nabla k + f'(k)\nabla A \right)\right) + sAf(k) - \delta k. \tag{2}$$

which is assumed to hold for any $(t,z) \in (0,T] \times \mathrm{int(U)}$, and $\mathrm{div} = \sum_{j=1}^{d} \frac{\partial}{\partial z_i}$.

[5] The existence and the monotonicity of ψ is an empirical issue. Our modelling frameworks is quite general and capable of incorporating alternative assumptions. If for example the stock of capital does not affect the flux than $\psi = 1$.

Proof. For the proof see Section A.1 in the Appendix. ■

On the boundary points, $z \in \partial U$, the behavior of k is prescribed by appropriate boundary conditions. Such boundary conditions can for example be homogeneous Dirichlet boundary conditions $k(t,z) = 0$ for $z \in \partial U$ which corresponds to annihilation of the capital stock at the boundary, or Neumann boundary conditions $J(t,z) \cdot n = 0$ for every $z \in \partial U$, where n is the outward normal at ∂U, corresponding to no flux of the capital stock from the boundary of U. Another set of appropriate boundary conditions would be to assume that U is an infinite domain and $\frac{\partial}{\partial z}k(t,z) \to 0$ as $|z| \to \infty$, or to assume that U is a bounded domain and use periodic boundary conditions. In this paper we will assume that U is bounded with sufficiently smooth boundary ∂U and we will adopt homogeneous Neumann boundary conditions on ∂U. We consider these boundary conditions as plausible from the economic point of view as they do not require the arbitrary specification of the capital stock on the boundary as the Dirichlet boundary conditions would require. The same would be true for periodic boundary conditions, which in fact lead to similar results as for the Neumann boundary conditions.

Equation (2) is a nonlinear diffusion equation in which the transport mechanism is nonlinear and depends on the underlying production function. This can contrasted with spatial models of capital movements based on the trade balance approach which assume a linear diffusion transport mechanism and usually admit the form of a semilinear equation as

$$\frac{\partial k}{\partial t} = D\mathrm{div}\nabla k + sAf(k) - \delta k. \tag{3}$$

The nonlinear form of the transport mechanism in our model (2) stems from the assumption that capital moves toward locations of higher marginal productivity which underlies the structure of capital flows across locations.

Some remarks are in order here.

1. Since the production function is concave (decreasing returns), $f'' < 0$ equation (2) is a nonlinear diffusion equation with positive nonlinear diffusion coefficient, which leads to a well posed parabolic problem.[6]

2. If A does not depend on space, or if the spatial dependence is slow, in the sense that the ratio $\frac{\nabla A}{A} = O(\epsilon)$ where ϵ is a small parameter,

[6] If we allowed for increasing returns, then the diffusion coefficient would be negative, leading to an ill posed parabolic problem. By ill posed we mean that even though solutions may exist, we may lose continuity with respect to the initial condition or have sharp concentrations of capital, blow up phenomena etc. The case of increasing returns is an interesting issue for future research.

then equation (2) reduces or can be approximated (respectively) by the simpler equation

$$\frac{\partial k}{\partial t} = -\text{div}\left(AB\psi(k)kf^{''}(k)\nabla k\right) + sAf(k) - \delta k. \tag{4}$$

3. An alternative way to derive (2) is to consider it as the continuous limit of a random walk on a discrete lattice, in which capital stock at a lattice site i may move to any of its neighbouring lattice sites j with probability proportional to the difference of marginal productivity between the site i and the site j. If $m(t,j) > m(t,i)$ then the capital stock at i will move to j but not otherwise. The continuous limit of this random walk will lead to a PDE similar to (2). Furthermore capital will only move from a site z to a site z' if $m(t,z^{'}) > m(t,z)$, independent of the relative concentration of capital between $z^{'}$ and z. This property seems to overcome issues related to Lucas paradox.

4 A Spatial Solow Model

We turn now to study in detail the implications of capital flows across locations modelled by (2) for the traditional Solow model with a Cobb-Douglas production function $f(k) = Ak^{\alpha}$, $\alpha \in (0,1)$.

Assumption 2 *The tendency of capital to move across locations follows the form in Assumption 1 with $\psi(k) = k^{\rho}$. If $\rho > 0$ then an increase in the capital stock will enforce the tendency of capital to move in search of higher marginal returns, whereas, while if $\rho < 0$ the opposite will take place.*

The spatial Solow model is defined in the following proposition.

Proposition 2 *Under Assumption 2, and assuming that A is constant or slowly varying with z in (4), the fundamental equation of economic growth describing the spatio-temporal evolution of the capital stock is given by the quasilinear degenerate partial differential equation*

$$\frac{\partial}{\partial t}k = \bar{D}\nabla_z \cdot (D_0(z)k^{\beta}\nabla_z k) + sA(z)k^{\alpha} - \delta k \tag{5}$$

or the equivalent form,

$$\frac{\partial}{\partial t}k = \frac{\bar{D}}{1+\beta}\nabla_z \cdot (D_0(z)\nabla_z k^{\beta+1}) + sA(z)k^{\alpha} - \delta k \tag{6}$$

where $\bar{D} = \alpha(1-\alpha)$, $\beta = \rho - (1-\alpha)$ and $D_0(z) = B(z)A(z)$.

The diffusion mechanism reduces to the linear diffusion mechanism in the special case where $\beta = 0$ or equivalently in the case where the parameters of the model are such that $\rho = (1-\alpha)$. The special case of an AK model where $\alpha = 1$, $\rho = 0$ leads to $\bar{D} = 0$.

The proof is straightforward and is omitted, however, the following remarks are important. Except for the special case where $\beta = 0$, our model is a nonlinear diffusion model with diffusion coefficient $D(z,k)$ depending on the state of the system as $D(z, k(t,z)) = \bar{D}_0(z)k(t,z)^{\beta}$, where $\bar{D}_0$ is a known function of space. Therefore, in our model the diffusion coefficient $D(z,k)$ which determines capital mobility across space, is determined endogenously under the assumption that capital flows to locations of relatively higher productivity. In models of linear diffusion the fixed diffusion coefficient D is determined exogenously. We feel that, although the degree of dependency of the diffusion coefficient on the stock of capital and the structure of capital velocity is an empirical issue, our approach by relating these factors to capital flows, provides a richer environment for studying the spatiotemporal evolution of capital stock.

Equation (5) is a generalization of the well studied porous medium equation in the sense that it is a porous medium equation with a reaction term.[7] It is interesting to note that this porous medium equation was not imposed as a modeling tool, but emerged from the assumption that capital flows seeks locations of high productivity and moves with the velocity which may depend on existing capital stock at the location of origin.

The parameter $\beta = \rho - (1-\alpha)$ plays a very important role in the porous medium equation. If $\beta > 0$ then this is traditionally called the *slow diffusion* case. On the other hand, if $\beta < 0$ this corresponds to the *fast diffusion* case.[8] In the slow diffusion case, the phenomenon of existence of compact support solutions is well known and very common. If the initial condition k_0 has support which is a compact subset of the domain, then, the solution presents the finite speed of propagation property. This means that for any $t > 0$, there will be regions of the domain for which the solution is identically equal to 0, i.e., the support has a free boundary, which separates the regions where $k > 0$ from the regions where $k = 0$. This phenomenon never holds for the linear

[7] The porous medium equation, in the absense of reaction term has been studied very actively, as a paradigm for nonlinear diffusion, and has served as a model for various physical or biological systems (see for example. Vasquez [39]).

[8] In the slow diffusion case we encounter a degenerate system as $D(k) \to 0$ as $k \to 0^+$. On the other hand, in the fast diffusion case we encounter a singular system as $D(k) \to \infty$ as $k \to 0^+$.

diffusion case $\beta = 0$, which presents infinite speed of propagation, meaning that even if the initial condition k_0 is of compact support, the solution for any $t > 0$, will not have this property Technical and abstract as it may sound at first, this qualitative behavior of the nonlinear (slow) diffusion may have interesting implications from the point of view of economic theory. This is because the compact support property may be interpreted in terms of the existence of regions where capital is depleted and remains depleted in the long run. This situation can be regarded as the limit of a poverty trap.

Furthermore, Proposition 2 elucidates the role of the production elasticity α in the capital concentration dynamics. To make the argument more transparent, consider the case $\beta = 0$ and let D_0 be independent of z. Equation (5) assumes the semilinear form

$$\frac{\partial}{\partial t}k = \bar{D}(1-\alpha)\Delta k + sA(z)k^{\alpha} - \delta\alpha, \quad \Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial z_i^2} \tag{7}$$

similar to the models employed so far in the literature on spatial growth, but with an important difference; the diffusion coefficient is proportional to $1-\alpha$. Therefore, if $\alpha < 1$ (diminishing returns to capital) then the diffusion coefficient is positive and this leads to a model similar to the one proposed by ([10], however within a totally different modeling framework. The positive diffusion coefficient corresponds to dynamics that tend to eliminate spatial gradients, thus leading to spatial convergence[9] phenomena. If $\alpha = 1$ the model is reduced to a growth model with an Ak production function which eliminates spatial heterogeneity. In the relevant literature based on trade balance (e.g. [9]) the diffusion coefficient $D = \bar{D}(1-\alpha)$ is set at the value of one, so that the relevant PDE is

$$\frac{\partial}{\partial t}k = \Delta k + sA(z)k^{\alpha} - \delta k \tag{8}$$

Models (5), (7) or (8) can thus be regarded as candidate specifications for a spacial growth equation with exogenous saving rates. The determination of the impact of spatial diffusion on capital accumulation, and therefore the choice of the appropriate model, is an empirical issues related to the estimation of coefficients in capital accumulation equations like (5), (7) or (8). Numerical simulation presented latter on suggest that the implication of these models for the long-run spatial distribution of capital, and convergence, in the context of the spatial Solow model, are not the same. In particular $D = 1$ combined with diminishing returns $(0 < \alpha < 1)$, or constant returns

[9] The term convergence used in its economic growth context.

($\alpha = 1$) tend to reduce an initial spatial heterogeneity and produce "flatter" spatial distributions. On the other hand models of linear or nonlinear diffusion tend to preserve and even amplify an initial spatial heterogeneity, implying that convergence is not attained in the long run.

5 Qualitative and Quantitative Aspects of the Spatial Solow Model

In this section we study the nonlinear spatial growth PDE characterizing the spatial Solow model defined Proposition2. By an appropriate rescaling of the spatial variable z the model can be written as:

$$\frac{\partial}{\partial t} k = D \Delta k^{\beta+1} + sAk^{\alpha} - \delta k \tag{9}$$

where $D > 0$ is a coefficient (independent of k but possibly spatially varying), s is the savings ratio, $A(z)$ is a productivity parameter, δ is the rate of capital depreciation and $\alpha \in (0, 1)$ is the production elasticity. Furthermore, without loss of generality, by a rescaling of the variable t, we may express the above equation in the form

$$\frac{\partial}{\partial t} k = \Delta k^{\beta+1} + c_1 k^{\alpha} - c_2 k \tag{10}$$

where

$$c_1 = \frac{sA}{D} \geq 0, \;\; c_2 = \frac{\delta}{D} > 0, \;\; \beta = \rho - (1 - \alpha).$$

In general c_1 and c_2 could depend on the spatial location z. We further allow for regions where $c_1 = 0$ (i.e. regions where no saving is possible). The possibility of allowing for a set $U_0 \subset U \subset \mathbb{R}^d$ with the property $c_1(z) = 0$ if $z \in U_0$, may provide insights regarding the existence of poverty traps.

The PDE (10) of the spatial Solow model will be complemented with an initial condition $k(0, z) = k_0(z)$, where $k_0 : U \subset \mathbb{R}^d \to \mathbb{R}_+$, is an initial capital stock distribution, and also with boundary conditions related to the prescribed behavior of the distribution of capital stock at certain parts of the domain U. We will consider the case where there is no flux of capital at the boundary, i.e., homogeneous Neumann boundary conditions $n \cdot \nabla k^{\beta+1} = 0$ on ∂U where n is the outward normal vector on ∂U.

Our final model will therefore be

$$\begin{aligned}
&\frac{\partial}{\partial t} k(t, z) = \Delta k^{\beta+1}(t, z) + c_1 k(t, z)^{\alpha} - c_2 k(t, z), \;\; (t, z) \in [0, T] \times U, \\
&n \cdot \nabla k(t, z)^{\beta+1}, \;\; (t, z) \in [0, T] \times \partial U, \\
&k(0, z) = k_0(z), \;\; z \in U.
\end{aligned} \tag{11}$$

On account of the following remark, the boundary condition may also be taken as $n \cdot \nabla k(t,z)$ for $(t,z) \in [0,T] \times \partial U$, without any significant change on the qualitative nature of the results.

Remark 1 *The no flux boundary condition is of the form $n \cdot J = 0$. For the model (10) the flux is $J = \nabla k^{\beta+1} = (\beta+1)k^{\beta}\nabla k$ so the full no flux boundary condition is actually $k^{\beta} n \cdot \nabla k = 0$ on ∂U. The linear diffusion model would correspond to $\beta = 0$ which gives the standard Neumann boundary condition $n \cdot \nabla k = 0$. Note that choosing $n \cdot \nabla k = 0$ on ∂U as the no flux boundary condition is an option which is also true for our case, and is further compatible with an endogeneous determination of the capital stock at suitable constant level rather than forcing it to be zero by the boundary condition as would happen for the Dirichlet case. We feel that our choice of boundary condition is more appropriate from the economic point of view than the choice of the Dirichlet boundary conditions.*

5.1 Steady state solutions

The starting point for our analysis will be the steady state solutions of model (11) i.e., solutions which are depending only on z and not on t. A steady state solution $k^* = k^*(z)$ can be regarded as a steady state distribution of the capital stock across space. For such solutions, the spatial Solow equation (11) simplifies to:

$$\begin{aligned} -\Delta k^{\beta+1} &= c_1 k^{\alpha} - c_2 k, \quad z \in U \\ n \cdot \nabla k^{\beta+1} &= 0, \quad z \in \partial U, \end{aligned} \tag{12}$$

where $\beta + 1 = \rho + \alpha$.

In the special case where c_1 and c_2 are constants, it is easy to see that (12) admits two possible solutions, $k = 0$ and $k = \left(\frac{c_1}{c_2}\right)^{1/(1-\alpha)}$, which will be called hereafter the flat solutions. These are the standard solutions of the non-spatial Solow model. However, the PDE (12) has a rich behavior which goes beyond these two standard flat solutions as the analysis of this section will show.

We use the Kirkhoff transformation $u = k^{1+\beta} = k^{\rho+\alpha}$ to express the system (12) in the form

$$\begin{aligned} -\Delta u &= c_1 u^{\gamma_1} - c_2 u^{\gamma_2}, \text{ in } U, \\ n \cdot \nabla u &= 0, \text{ on } \partial U. \end{aligned} \tag{13}$$

where

$$\gamma_1 = \frac{\alpha}{1+\beta} = \frac{\alpha}{\rho+\alpha}, \text{ and } \gamma_2 = \frac{1}{1+\beta} = \frac{1}{\rho+\alpha}.$$

Depending on the value of ρ we have 3 cases:

C.1 If $-\alpha < \rho < 0$ then $1 < \gamma_1 < \gamma_2$, and the corresponding elliptic problem (13) is superlinear. For this parameter range $0 < \rho + \alpha = \beta + 1 < \alpha$ hence $\beta < \alpha - 1 < 0$.

C.2 If $0 < \rho < 1 - \alpha$ then $\gamma_1 < 1 < \gamma_2$, so the corresponding elliptic problem (13) has a nonlinearity which consists of both a sublinear and a superlinear term. For this parameter range, $\alpha < \rho + \alpha = \beta + 1 < 1$ hence $\beta < 0$.

C.3 If $1 - \alpha < \rho$ then $\gamma_1 < \gamma_2 < 1$, and the corresponding elliptic problem (13) is sublinear. For this parameter range $1 < \rho + \alpha = \beta + 1$ hence $\beta > 0$. The limiting case where $1 - \alpha = \rho$ leads to $\gamma_1 < \gamma_2 = 1$ and $\beta = 0$. In this case, our model reduces to a linear (Fickian) diffusion model for the capital transport, with a nonlinear production term.

Cases C.2 and C.3 correspond to the case where $\rho > 0$ so that an increase in the capital stock will enforce the tendency of capital to move in search of higher marginal returns, while C.1 where $\rho < 0$ corresponds to the opposite case. The case that actually prevails regarding ρ is an empirical issue that goes beyond the scope of this paper, however, Case C3 is of special interest because it is compatible with the existence of compact support solutions, which lead to the emergence of the phenomenon of poverty traps. We will show that this phenomenon occur only when $\beta > 0$ and is never expected to appear in the case where $\beta = 0$.[10]

Remark 2 *In the case of homogeneous Neumann or periodic boundary conditions, if the coefficients c_1 and c_2 are constant, then the only steady state solutions of the spatial model are the flat (spatially independent) solutions. The argument is as follows. Let us express the elliptic Solow PDE, after the application of the Kikrhoff transformation in the form $-\Delta u = f(u)$ where $f(u) = c_1 u^{\gamma_1} - c_2 u^{\gamma_2}$. Note that $f(u) = 0$ has two solutions $u = 0$ and $u_* = \left(\frac{c_1}{c_2}\right)^{1/(\gamma_2 - \gamma_1)}$ which corresponds (upon inverting the Kirkhoff transformation) to the two steady states of the temporal Solow model. Note furthermore, that if $u \in [0, u_*]$ then $f(u) \geq 0$.*

[10] In mathematical terms this is related to the failure of the standard maximum principle which provides most of the nice properties of elliptic systems. It is this sublinearity that leads us to the need of more careful analysis for the elliptic system (12) or its equivalent form (13).

Assume that there exists a spatially dependent steady state u, with the property $0 \le u(z) \le u_*$ for every $z \in U$, which solves the PDE

$$-\Delta u(z) = f(u(z)), \;\; z \in U, \tag{14}$$

with homogeneous Neumann or periodic boundary conditions. Integrate (14) over U and using the boundary conditions we obtain that u must satisfy the consistency condition $\int_U f(u(z))dz = 0$. Since we have assumed that $0 \le u(z) \le u_*$ for every $z \in U$, by the nature of the nonlinearity we have that $f(u(z)) \ge 0$ for every $z \in U$, so that the consistency condition implies that $f(u(z)) = 0$ for every $z \in U$, therefore, $u(z) = 0$ for every $z \in U$ or $u(z) = u_*$ for every $z \in U$. The first option leads to the trivial solution. The second option leads to the flat steady state $u = u_*$.

In the case of Dirichlet boundary conditions (homogeneous or not) or non homogeneous Neumann or Robin boundary conditions the above argument does not hold and we have in principle some spatial dependence of the solution, mainly because the flat steady states do not satisfy the boundary conditions (unless these are selected very precisely, e.g. $u_*(z) = u_*$ for $z \in \partial U$. The spatial dependence is generated even in the absence of spatial variability of the coefficients. For example, a typical spatial distribution for u for homogeneous Dirichlet boundary conditions is that of an inverted parabola, having a maximum on some interior point of the domain U (in accordance to the maximum principle).

While the exact spatial distribution of capital may depend either on the boundary conditions or the specific spatial variability of the coefficients, there are certain qualitative features which are robust with respect to these two aspects and in our view offer interesting insight to the problem. These are collected in the following propositions.

The following proposition provides an existence and uniqueness result for the steady state PDE (12) in any spatial dimension allowing for spatial variability of the coefficients c_1, c_2. Clearly this PDE always has the solution $k(z) = 0$ for every $z \in U$, which will be hereafter called the trivial solution.

Proposition 3 *Let $\beta \ge 0$, $\alpha < 1$ and assume Hölder continuity properties*[11] *for the coefficients c_1, c_2. The steady state equation (12) has a classical (non*

[11] A function is $g : \mathbb{R}^d \to \mathbb{R}$ is called Hölder continuous when there exists $\alpha < 1$ and $C > 0$ such that $|g(x) - g(y)| \le C|x - y|^\alpha$ for every $x, y \in \mathbb{R}^d$. Hölder continuity is a form of uniform continuity which is weaker than Lipschitz continuity.

trivial) positive solution,[12] *satisfying the a priori bounds*

$$0 \leq k \leq \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{\frac{1}{1-\alpha}},$$

where

$$\bar{c}_1 = \sup_{z\in U} c_1(z), \ \ \underline{c}_2 = \inf_{z\in U} c_2(z).$$

The solution is unique if $\alpha - 1 < \beta$ (cases C2. and C.3). The results of the proposition remain true if we consider Dirichlet or periodic boundary conditions.

Proof. For the proof see Section A.3 in the Appendix. ■

It should be noted that if c_1 and c_2 are independent of z, then the upper bound for the steady state is the steady state of the standard Solow model without diffusion of capital. Thus the spatial Solow model with nonlinear diffusion is characterized by steady state which could be either spatially homogenous, i.e. flat, or exhibit spatial heterogeneity.

The spatial model with non linear diffusion allows us to explore cases in which the spatial domain contains locations and regions where savings do not take place. This might be a realistic situation for extremely impoverished locations. We will call these regions poverty cores and will define them as regions $V_0 \subset U$ with the property that $c_1(z) = 0$ if $z \in V_0$. The poverty core suggests the existence of regions where capital is identically zero at a steady state, implying that the steady state distribution of the capital stock contains regions with no capital and regions with positive capital. This is a result suggesting that convergence, in the sense used in growth theory, is not feasible and economies where savings are not possible could eventually be trapped in the poverty core where their capital stock is depleted. Since poverty cores will coexist with regions of positive capital, convergence is not possible at the steady state. The existence of poverty cores is verified by the existence of compact support solutions for the steady state equation (12) and is established in the following proposition.

Proposition 4 *Let $\beta > 0$, (Case C.3), assume that c_1, c_2 are Hölder continuous functions and that for some $z_0 \in U$ and $\rho > 0$, c_1 vanishes, on account of zero savings ratio s, inside a ball centered at z_0 with radius 2ϱ, situated in the interior of the spatial domain. Then,* ***any*** *non-trivial positive solution*

[12] The regularity of the solution depends on the regularity of the coefficients c_1, c_2. If $c_1, c_2 \in L^\infty(U)$ then the solution is a weak solution $k \in L^1(U)$, whereas if the coefficients c_1, c_2 enjoy Hölder continuity properties, the solution is classical.

of the steady state equation (12) will develop a poverty core, i.e. a region of total depletion of capital stock inside a ball centered at z_0 and of radius ϱ, as long as the parameters of the problem satisfy the condition

$$-\underline{c}_2 \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{-\frac{1-\gamma_2}{\gamma_2-\gamma_1}} + \nu\varrho^2((\nu-1)+(d-1)) \leq 0. \tag{15}$$

where $\nu = \frac{2}{1-\gamma_2}$ and d is the spatial dimension. Poverty traps cannot form if $\beta = 0$ (nor in cases C.1 and C.2). The proposition holds true also for the case of Dirichlet boundary conditions or periodic boundary conditions.

Proof. The proof is given in Section A.4 in the Appendix. ■

Poverty cores do not emerge in models of linear diffusion ($\beta = 0$) where capital moves from high to low abundance location even if c_1 vanishes in subsets of U. Thus non linear diffusion ($\beta > 0$) can help model the emergence of poverty cores where capital is depleted due to zero savings. This happens because although capital moves to locations with low capital stock, since these locations are characterized by high marginal productivity of capital as capital is depleted, no part of the inflow is used for capital accumulation since nothing is saved. If the parameters are such that condition (15) is satisfied then no accumulation will take place at this location and eventually the capital stock will be depleted. The vanishing of savings at a point is not however enough to guarantee the existence of a poverty core. Due to the spatial interactions, a poverty core emerges when relation (15) is satisfied. This relation links the maximum value of c_1 over a wider region (reflecting saving rates and productivity in nearby regions) with the minimum value of c_2 over a wider region, (reflecting depreciation rates), and depends on the characteristic of the velocity of capital flows (provided by β) within the region. As, for the parameter range for which the poverty trap may occur, $\frac{1-\gamma_2}{\gamma_2-\gamma_1} > 0$, we note that relation (15) implies that the smaller the ratio $\frac{\bar{c}_1}{\underline{c}_2}$ is the easier it is for the poverty trap to occur. Furthermore, the procedure followed in the proof, provides detailed information of the local behavior of capital stock near a point z_0 with zero savings. Finally note that capital stock can be identically zero inside a ball of center z_0 and radius ϱ, i.e. well inside the region where c_1 vanishes, but capital may start accumulating (still inside the region where $c_1 = 0$) on account of spatial effects and capital flow from nearby regions, since marginal productivity inside this ball is high. Thus a poverty core might exist "surrounded" by locations where savings vanish but due to capital flows a positive capital stock acummulated at the steady state.

To provide a possible picture of the steady state distribution of the capital stock, we consider the solution of (13) under the following parametrization for (9)

$$\alpha = 0.4, s = 0.15, \delta = 0.03, \rho = 1.3, \beta = 0.7, D = 0.01, \qquad (16)$$
$$A(z) = \exp\left(-z^2/4\right), z \in [-4, 4]$$

The productivity parameter $A(z)$ reflects the assumption that in an economic space like the one depicted in figure 3 more developed locations in terms of per capita GDP have a relatively higher productivity parameter. As shown in figure 4, the steady state spatial distribution of the capital stock is bell-shaped.

[Figure 4. Steady state distribution of capital stock]

To obtain some insight on the emergence of poverty cores we solve (13) with the saving ratio defined as $s(z) = 0.15\left(1 - \exp\left(-z^2/4\right)\right)$ which suggest that savings vanish at the center of the domain, and $A(z) = 1$ which suggest no spatial productivity differentiation. The emergence of the poverty core is depicted in figure 5.

[Figure 5. Poverty core at the steady state]

5.2 Time and space dependent solutions

Having studied steady state we now turn our attention to the analysis of the full spatiotemporal Solow model. This means finding the spatial distribution of capital stock $k(t, z)$ at each point of time t, that emerges if the fundamentals of the economies are determined by the basic assumptions of the Solow growth model and capital flows towards locations of relative higher marginal productivity with velocity determined endogenously by local capital stock and the size of marginal productivity. The corresponding mathematical problem reads as follows: given a function $k_0 : U \to \mathbb{R}$, find $k : [0, T) \times U \to \mathbb{R}$ such that the following initial boundary value problem is satisfied

$$\begin{aligned} &\frac{\partial k}{\partial t} = \Delta\Phi(k) + f(z, k), \;\; (t, z) \in (U)_T, \\ &\frac{\partial \Phi(k)}{\partial n}(t, z) = 0, \;\; (t, z) \in (\partial U)_T, , \\ &k(0, z) = k_0(z), \;\; z \in U, \end{aligned} \qquad (17)$$

where we use the notation $(U)_T = [0, T] \times U$, $(\partial U)_T = [0, T] \times \partial U$ and where $\Phi : [0, \infty) \to [0, \infty)$, $f : U \times [0, \infty) \to [0, \infty)$ are generic functions; for our purposes we have

$$\Phi(k) := k^{1+\beta} = k^{\rho+\alpha}, \;\; f(z, k) := c_1(z)k^{\alpha} - c_2(z)k. \qquad (18)$$

It is worth noting that for our choice of Φ, the no-flux (Neumann) boundary condition is in fact compatible with the standard no-flux boundary condition $\frac{\partial k}{\partial n}(t,z)=0$.

Problem (17) presents an interesting technical twist, which is directly related to economics. Since we are assuming decreasing returns, $\alpha < 1$, hence the function f is **not** Lipschitz continuous for k taking values in a neighbourghood of zero, but rather Hölder continuous. Eventhough this may sound as a boring technicality, it has important consequences for the uniqueness of solutions to (17), since the uniqueness theorem for the solution of PDEs needs the assumption of Lipschitz continuity of the nonlinearities. We should note that this non uniqueness problem is not a problem only of the model we propose here; it is also true for the standard temporal Solow model, as well as the PDE Solow version proposed by Boucekinne and coworkers (see e.g. [10]). The problem arises from the Cobb-Douglas production function and **not** from the transport term, whatever this may be. In the purely temporal case, this phenomenon is almost never discussed since we are usually interested in the region close to the non zero steady state $k^* \neq 0$, and in this region f is Lipschitz and no problem arises. For the standard diffusion case treated by Boucekinne, in which $\rho + \alpha = 1$, the maximum principle guarantees that $k > 0$ so again there is no need to pay too much attention to this problem. However, in the case we consider here, the emergence of poverty traps, dictates the need to consider seriously the pathological region $k = 0$, since as shown in Proposition 4 for a steady state we may have regions where $k > 0$ and regions where a poverty core emerges and $k = 0$.

We have the following existence result.

Proposition 5 *Let $\beta \geq 0$. If $k_0 \geq 0$, there exists a solution to problem (17) in the weak sense. In particular there exist two weak solutions of (17), $k_\diamond$ (the minimal solution) and $k^\diamond$ (the maximal solution) such that any weak solution k of (17) satisfies $0 \leq k_\diamond \leq k \leq k^\diamond \leq M_*$. If $k_0 > 0$, the solution to (17) is unique. The proposition remains true also for Dirichlet or periodic boundary conditions.*

Proof. For the proof see Section A.5 in the Appendix. ■

The asymptotic in time behavior of the solutions of system (17) is not a very easy problem on account of (a) the degeneracy of the problem resulting from the inclusion of the nonlinear diffusion term and (b) of the non Lipschitz property of the nonlinearity which is inherited by the use of the Cobb-Douglas production function. Its full treatment requires abstract techniques from nonlinear analysis, beyond the scope of the present article. However, here we wish to present a preliminary result in this direction, related to the problem of linearized stability of the steady state solution, which is important in its own right, and which highlights the various subtleties of problem (17).

Let $k^* = k^*(z)$ be a steady state solution of (17), that is k^* satisfies $-\Delta\Phi(k^*)+f(z,k) = 0$, and consider solutions of the time dependent problem (17) with initial condition $k(0,z) = k^*(z) + \epsilon u(0,z)$, where $\epsilon u(0,z)$ is a small initial perturbation around the steady state. This initial condition evolves according to the evolution equation (17). It is intuitively clear that if $u(t,z) \to 0$ for all $z \in U$ as $t \to \infty$ then the solution $k(t,z)$ of the full system (17) will have the property that $k(t,z) \to k^*(z)$ for all $z \in U$ as $t \to \infty$, hence all initial perturbation around the steady state solution k^* will die out and be eliminated as an effect of the dynamics and the steady state is asymptotically stable. If on the other hand $|u(t,z)| \to \infty$ for all $z \in U$ as $t \to \infty$, then the dynamics of the system will has as effect the attenuation of the initial small disturbance around the steady state, the state of the system for any time will deviate from k^* and the steady state is unstable.

The problem of stability or instability of the steady state is usually approximated by using the linearized version of the evolution equation (17) around the steady state k^* and looking at the spectrum of an appropriate elliptic eigenvalue problem which depends on the particular form of k^*. It the principal eigenvalue is positive the we have instability whereas if it is negative we have stability. However, this argument relies on the fact that we may use the Taylor expansion to approximate $\Phi(k^*(z) + \epsilon u(t,z))$ and $f(k^*(z) + \epsilon u(t,z))$ for small ϵ by a linear form in u. This is clearly inappropriate for (17) since f is non Lipschitz (let alone continuously differentiable) in the neighbourghood of zero. Therefore, if k^* approaches zero, or even worse if k^* develops a poverty trap as shown in Proposition 4 the linearization argument does not apply and the issue of treating the stability of the steady state k^* for small initial perturbations becomes more involved and requires special attention.

The following proposition provides a stability result for small perturbations of a steady state k^* which is valid also in the case where a poverty trap may occur.

Proposition 6 *Let $\beta \geq 0$ and k^* be a steady state.*

(i) If $k^* > 0$ the steady state is linearly asymptotically stable.

(ii) If $k^* \geq 0$, i.e., when the steady state develops a poverty trap, the core of the trap is persistent.

The proposition remains true also for Dirichlet or periodic boundary conditions.

Proof. For the proof see Section A.6 in the Appendix. ■

5.3 Numerical simulations

Having established existence of solutions, steady states,and stability properties for the steady states, we turn now to some simulation results to determine the shape of the spatiotemporal distribution of capital emerging for the spatial Solow model under plausible parameter choice. Our simulations solve numerically[13] model(9). Using the same parameter choice as in section 5.1 i.e.

$$\begin{aligned} &\alpha = 0.4, s = 0.15, \delta = 0.03, \rho = 1.3, \beta = 0.7, D = 0.01, \\ &A(z) = \exp\left(-z^2/4\right), z \in [-4, 4] \end{aligned} \tag{19}$$

figure 6 depicts the spatiotemporal evolution of the stock of capital with initial condition $k(0, z) = \mathsf{e}^{-z^2/4} + 0.01 \sin[50\pi z] - 0.0183156,\ z \in [-4, 4]$. This is a bell-shaped distribution chosen with the purpose of approximating, through the initial conditions, a distribution which could potentially emerge if we consider economic - not geographical - space with the distance defined in terms of GDP per capita differences. The boundary conditions are of zero flux type or $\partial k(t, -4)/\partial z = \partial k(t, 4)/\partial z = 0$.

[Figure 6 Spatiotemporal distribution of capital]

Figure 6 depicts the evolution of the Sobolev norm defined as

$$Sb(t) = \int_{-4}^{4} \left(\frac{\partial \hat{k}(t, z)}{\partial z} \right)^2 dt$$

where $\hat{k}(t, z)$ is the solution of (9) as depicted in figure 6.

[Figure 7. The time path of the Sobolev norm]

The convergence of the Sobolev norm to a fixed number means that the spatial gradients remain constant after a certain point in time, implying that the system converges to spatially nonhomogeneous distribution of the stock of capital. Furthermore the peak of the distribution in Figure 6 converges for $t > 200$ to a fixed positive number. Combining this with the convergence of the Sobolev norm suggests that the growth model converges in a spatiotemporal sense to a nonhomogeneous bell-shaped capital stock distribution. This result is consistent with our theory about the stability of spatially nonhomogeneous steady states and the steady state of figure 4.. Since per capita

[13] Wolfram Mathematica was used for the numerical simulations. The PDEs were solved for $t \in [0, 1000]$.

output is given by $\hat{y}(t,z) = \left(\hat{k}(t,z)\right)^{\alpha}$, per capita output also converges to a spatially non homogenous distribution.

The bell-shaped pattern remains when we assume that $\rho < 0$ i.e., an increase in the stock of a capital at a given location will reduce the tendency of the capital stock to seek for locations of higher productivity, and $\beta < 0$. If there is no spatial variability of the productivity parameter i.e., $A(z) = 1$, the spatial distribution becomes flat with or without spatially differentiated initial conditions. This result is consistent with the steady state result obtained in section 5.1, and is shown in figure 8.

[Figure 8. A flat spatiotemporal distribution of capital]

Our numerical results suggest therefore that under the plausible zero flux boundary conditions, spatial variability of productivity is important in generating persistent spatially nonhomogeneous distribution for the stock of capital in the spatial Solow model. Ak models when $\alpha = 1$ result also in flat spatial distribution. This was anticipated in view of (7).

The spatiotemporal evolution of a poverty trap, emerging from spatially flat initial conditions, can also be shown if we assume as in section 5.1 that $s(z) = 0.15\left(1 - \exp\left(-z^2/4\right)\right)$ and $A(z) = 1$. This is shown in figure 9. The Sobolev norm converges suggesting that the poverty core is persistent.

[Figure 9. Spatiotemporal evolution of a poverty trap]

Spatially nonhomogeneous bell-shaped pattern emerge and persists in time with Dirichlet boundary conditions $k(-4,t) = k(4,t) = k_0 \geq 0$, circle boundary conditions $k(-4,t) = k(4,t)$, and with time dependent boundary conditions $k(-4,t) = k(4,t) = \gamma t$ or $k(-4,t) = k(4,t) = \mathrm{e}^{\gamma t}$ which may reflect the assumption that location with low capital stock at the beginning may grow fast. The bell-shaped patterns with Dirichlet boundary conditions emerge in even with no spatial variability of the productivity parameter. they also emerge in the Ak model. This is also anticipated since in general Dirichlet boundary conditions especially of the hostile boundary $k(-4,t) = k(4,t) \approx 0$ tend to "force" the formation of non spatially homogeneous patterns. This result is also consistent with the steady state analysis of section 5.1.

Finally when we parametrize for the trade balance model, which corresponds to $\beta = 1$ and $D = 1$, with zero flux boundary conditions, the result shown in figure 10 is that initial spatial differentiation $k(0,z)$ becomes flatter but it does not disappear completely. The Sobolev norm for this model converges to 0.11. The comparison of this value with the value of 17.9 for the model corresponding to figure 6, suggests that the strong spatial gradients and spatial heterogeneity are more persistent in model with nonlinear diffusion than models with linear diffusion. This observation could provide some insights into the mechanisms driving the spatiotemporal evolution of capital stock.

[Figure 10. The spatiotemporally evolution for the trade balance model]

The numerical simulations seem to support the theory developed in the context of a spatial Solow model regarding the spatiotemporal evolution of the capital stock and the existence of steady states for a plausible set of parameter values regarding, savings rates depreciation and production elasticity. Furthermore they seem to suggest that capital flows characterized by capital seeking locations of high returns and an endogenous flow velocity, result in a persistent spatially nonhomogeneous distribution of capital and per capita output across locations if, as it is plausible, there are productivity differentials across locations and boundary conditions are zero flux. . This result holds under various types of boundary conditions.

6 Conclusions and Possible Extensions

Seeking to explore mechanisms underlying the temporal evolution of the cross sectional distribution of per capita capital and output across space we develop a spatial growth model where saving rates are exogenous. Capital movements across locations are governed by a mechanism where capital moves towards locations of relatively higher marginal productivity, with a velocity determined by the existing stock of capital. Considering that the spatial domain corresponds to economic space we developed a local model where the fundamental growth equation of the Solow model is augmented by a nonlinear diffusion term, which characterizes spatial movements.

We show that the augmented Solow equation has a solution and that steady states exist. Furthermore under diminishing returns the growth process could lead under plausible assumptions, to a stable spatially non-homogenous distribution for per capita capital and income in the long run. Insufficient savings may lead to the emergence of poverty cores where capital stock is depleted in some locations. Stability analysis indicates that a steady state with poverty cores is stable. This suggests that economies can persistently remain in the poverty core while economies in other locations will have a positive capital stock. In the spatial Solow model zero capital stock in some locations is consistent with the long run stability of the entire spatial distribution of the stock of capital. Numerical simulations confirm our theoretical results.

Our approach, by endogenizing the velocity of the capital flow provides a rich environment for studying growth processes in a spatiotemporal context. Furthermore by linking capital flows with differences in the marginal productivity of capital across locations and not with differences in the stock of capital across locations our approach seems not to suffer from the critiques associated with Lucas paradox. The emergence of spatial distributions where persistent poverty cores coexist with locations where the stock of capital is high - that is when the solution of the growth equation results in distribution with compact support - is a potentially interesting result suggesting that the nonlinear diffusion approach could support outcomes, which could be in line with observed situations.

A Proofs

A.1 Proof of Proposition 1

Proof. Consider any $V \subset U$ which consists only of interior points of U. Assuming only the integrability of J we may express the net inflow and outflow of capital from V in terms of a surface integral,

$$\text{Inflow} - \text{Outflow} = \int_{\partial V} J \cdot n dS = \int_{\partial V} (\mathfrak{v} k) \cdot n dS,$$

where dS is the surface element and n is the outward normal on the boundary of V. The models the natural observation that anything that moves in or out of V must definitely pass through its boundary ∂V. An application of Gauss' divergence theorem yields that

$$\int_{\partial V} J \cdot n dS = -\int_V \mathrm{div} J(t,z) dz := -\int_V \sum_{i=1}^{d} \frac{\partial}{\partial z_i} J_i(z,t) dz$$

The book-keeping equation for V assumes the form

$$\frac{\partial}{\partial t} \int_V k(t,z) dz = -\int_V \mathrm{div} J(t,z) dz + \int_V \left(s(z) A(z) f(k(t,z)) - \delta k(t,z) \right) dz.$$

Dividing by the volume of V, and then since V is arbitrary, by shrinking V to z and passing to the limit as the volume tends to 0, we obtain the differential equation

$$\frac{\partial}{\partial t} k(t,z) = -\mathrm{div} J(t,z) + s(z) A(z) f(k(t,z)) - \delta k(t,z).$$

We now assume that $J = \mathfrak{v}k$ with $\mathfrak{v}$ given as in Assumption 1. Note that

$$m(t,z) = \frac{\partial}{\partial k} y(t,z) = A(z) f'(k(t,z)),$$

so that

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

Therefore,

$$J(t,z) = B\psi(k)k \left(A f''(k) \nabla k + f'(k) \nabla A \right),$$

where on the right hand side we drop the explicit dependence on (t,x) of k and A for notational simplicity. This leads to the required form for the PDE. ∎

A.2 Concepts of solutions

The vanishing of Φ' at zero makes the problem degenerate and thus the concept of classical solution for (17) is not appropriate. For this reason, and by following the usual procedure, we introduce a weak notion of solution. We let $U_T := (0,T) \times U$, $(\partial U)_T := (0,T) \times \partial U$ and define the space of test functions as follows:

$$\mathcal{J} := \left\{ \psi \in C(U_T) \quad : \quad \psi \geq 0, \ \left. \frac{\partial \psi}{\partial n} \right|_{(\partial U)_T} = 0 \ \text{ and } \ \frac{\partial \psi}{\partial t}, \ \Delta \psi \in L^2(U_T) \right\}.$$

Definition 1 *Let $k_0 \in L^\infty(U)$.*[14] *The function $k \in L^\infty(U_T)$, $k \geq 0$, is called a (weak) solution of* (17) *if*

$$\int_U k(t,z)\psi(t,z)\,dz = \int_U k_0(z)\psi(0,z)\,dz$$
$$+ \int_0^t \int_U \left[k(s,z) \frac{\partial \psi}{\partial t}(s,z) + \Phi(k(s,z)) \Delta \psi(s,z) + f(k(s,z))\psi(s,z) \right] dz\,ds,$$

for $0 \leq t < T$ and every $\psi \in \mathcal{J}$.

Definition 2 *If we substitute the equality above with $\geq$, resp. $\leq$, then we obtain the concept of supersolution, resp. subsolution of* (17)*. A function which is at the same time a supersolution and a subsolution is a solution.*

[14] Meaning that the initial condition is essentially bounded on U.

The concept of weak solutions can also be extended to the steady state (elliptic problem) by choosing test functions $\psi \geq 0$ which are depending only on the spatial variable and not on time. By the density of test functions in the Sobolev space $W^{1,2}(U)$ we may consider the test functions as belonging to $W^{1,2}(U)$. For the sake of convenience of the reader we provide the definition of weak solution for an elliptic equation of the form

$$\begin{aligned} -\Delta u + f(z,u) = 0, \ \text{ in } U, \\ n \cdot \nabla u = 0, \ \text{ on } \partial U. \end{aligned} \tag{20}$$

Definition 3 *A function $u \in W^{1,2}(U)$ is called a weak solution of (20) if*

$$J(u) := \int_U \nabla u \cdot \nabla \psi dz + \int_U f(z,u)\psi dz = 0, \ \ \forall \ \psi \in W^{1,2}(U).$$

If we restrict to test functions $\psi \geq 0$, and the above (a) holds for a function $\underline{u} \in W^{1,2}(U)$ as the inequality $J(\underline{u}) \leq 0$ we say that $\underline{u}$ is a sub-solution, (b) holds for a function $\bar{u} \in W^{1,2}(U)$ as the inequality $J(\bar{u}) \geq 0$ we say that $\bar{u}$ is a super-solution of (20) respectively.

The concepts of super and sub-solutions are very important in the construction of solutions to both elliptic and parabolic equations, and in providing a priori bounds and estimates for the solutions of partial differential equations.

A.3 Proof of Proposition 3

Proof. Using the Kirkhoff transformation we bring the system to the equivalent form

$$-\Delta u = c_1 u^{\gamma_1} - c_2 u^{\gamma_2}, \tag{21}$$

with homogeneous Neumann boundary conditions.

For existence we can use the method of sub and supersolutions (see Definition 3 in Section A.2). For that it is convenient to express equation (21) as $J(u) := -\Delta u - c_1 u^{\gamma_1} + c_2 u^{\gamma_2} = 0$ in U with $\frac{\partial}{\partial n} u = 0$ on ∂U, and recall that a (weak) subsolution $\underline{u}$ is a $W^{1,2}(U)$ function such that $J(\underline{u}) \leq 0$ whereas a (weak) supersolution $\bar{u}$ is a $W^{1,2}(U)$ function such that $J(\bar{u}) \geq 0$ where the above inequalities are considered in a weak sense. For the generalization of the concept of sub and supersolutions in the weak sense for Neumann boundary conditions see e.g. [38] and references therein. If a pair of sup and supersolutions $\underline{u}$ and $\bar{u}$ exist such that $\underline{u} \leq \bar{u}$ with $|c_1 u^{\gamma_1}(z) - c_2 u^{\gamma_2}(z)| \leq \phi(z)$

for every $z \in U$ and $u(z) \in [\underline{u}(z), \bar{u}(z)]$ with $\phi \in L^2(U)$, then, there exists a solution of (21) u such that $\underline{u} \leq u \leq \bar{u}$ (see e.g. Theorem 2.3 in [38] for the case of weak solutions).

A standard candidate for a supersolution is a constant function, $\bar{u} = M$. It is easily seen that $J(M) = -c_1 M^{\gamma_1} + c_2 M^{\gamma_2} = M^{\gamma_1}\left(-c_1 + c_2 M^{\gamma_2-\gamma_1}\right) \geq M^{\gamma_1}\left(-\bar{c}_1 + \underline{c}_2 M^{\gamma_2-\gamma_1}\right)$, so that choosing $\bar{u} = M = \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{1/(\gamma_2-\gamma_1)}$ we guarantee that $\bar{u}$ is a supersolution. On the other hand, a standard candidate for a subsolution is a proper multiple of ϕ_1 where ϕ_1 is the eigenfunction related to the dominant eigenvalue λ_1 of the problem $-\Delta\phi = \lambda\phi$, with homogeneous Neumann boundary conditions. It is a well known fact that $\lambda_1 > 0$ and $\phi_1 > 0$ in U. We will look for subsolutions of the form $\underline{u} = \epsilon\phi_1$ for a proper choice of ϵ. We observe that $J(\epsilon\phi_1) = \epsilon\lambda_1\phi_1 - \epsilon^{\gamma_1} c_1 \phi_1^{\gamma_1} + \epsilon^{\gamma_2} c_2 \phi_1^{\gamma_2} = \epsilon\phi_1\left(\lambda_1 - \epsilon^{\gamma_1-1} c_1 \phi_1^{\gamma_1-1} + \epsilon^{\gamma_2-1} c_2 \phi_1^{\gamma_2-1}\right)$. If $J(\epsilon\phi_1) \leq 0$ then $\underline{u} = \epsilon\phi_1$ is a subsolution. The only negative term in $J(\epsilon\phi_1)$ is the middle term (if $c_1 > 0$). Then, in cases C.2 and C.3, for $\epsilon > 0$ small enough $J(\epsilon\phi_1) \leq 0$ and $\underline{u} = \epsilon\phi_1$ is a subsolution. Therefore, there exists a weak solution $u \in W^{1,2}(U)$ such that $\epsilon\phi_1 \leq u \leq M$. The proof then concludes using a standard bootstrapping argument (see e.g. [3] Chapter 7). For the benefit of the reader, we sketch that here: Since $u \in W^{1,2}(U)$ by the Sobolev embedding theorem $u \in L^{r_0}(U)$ (where $r_0 = \infty$ if $d = 2$ and $r_0 = \frac{2d}{d-2}$ if $d > 2$). Then, we rewrite (21) as $-\Delta u = f(x)$ with $f(x) = c_1(x)u(x)^{\gamma_1} - c_2(x)u(x)^{\gamma_2}$ and by the properties of u we see that $f \in L^{r_0/\gamma_2}(U)$. Then the analogue of the Agmon-Douglis-Nirenberg estimates for the Neumann problem $-\Delta u = f$ (see e.g. [2] or Lemma 5.2 in [40]) guarantees that $u \in W^{2,p}(U)$ and a further application of the Sobolev embedding theorem implies that $u \in C^{\mu}(\bar{U})$ for some $\mu \in (0,1)$. Since c_1, c_2 are Holder functions, we have by the previous estimates that f is Holder and considering once more the Neumann problem $-\Delta u = f$ with a Holder right hand side using the extension of the Schauder theory for such boundary conditions (see e.g. Theorem 6.26 in [25]) we conclude the higher regularity of u.

For uniqueness we need the extension of the classic results of [11] which are valid for the Dirichlet case, to the Neumann case. The uniqueness is guaranteed if the function $u \mapsto \varphi(x,u) = \frac{f(x,u)}{u}$ is strictly decreasing for every $z \in U$ for $u \in [0,\infty)$ (see e.g. Theorem 2 in [32]). Since $\varphi(x,u) = c_1 u^{\gamma_1-1} - c_2 u^{\gamma_2-1}$ we can calculate $\varphi'(x,u) = (\gamma_1-1)c_1 u^{\gamma_1-2} - (\gamma_2-1)c_2 u^{\gamma_2-2} = (\gamma_1-1)u^{\gamma_1-2}\left(c_1 - \frac{\gamma_2-1}{\gamma_1-1} c_2 u^{\gamma_2-\gamma_1}\right)$, from which the claims follow. ■

A.4 Proof of Proposition 4

Proof. We apply the Kirkhoff transformation $u = k^{\rho+\alpha}$ and work with the transformed steady state equation

$$-\Delta u = c_1 u^{\gamma_1} - c_2 u^{\gamma_2}, \text{ in } U, \tag{22}$$

with Neumann boundary conditions on ∂U. We will show that if $\gamma_1 < \gamma_2 < 1$, and c_1 vanishes on a subset of U (of positive measure) then u (hence k) develops a poverty trap, i.e., u is a non-trivial solution of (22) that vanishes on a region of positive measure. Our argument relies heavily on [21] who studied a very similar system with Dirichlet boundary conditions. In fact it turns out that since the argument relies on local considerations, only minor modifications are required, however, it is reproduced here as it allows us to obtain concrete conditions on the parameters of the system for the poverty trap to exist which are of interest from the economic point of view.

Without loss of generality assume that c_1 vanishes inside a ball of radius 2ϱ centered at $x_0 = 0$. We will use the notation $B_2 = B(0, 2\varrho)$ for this ball, and $B_1 = B(0, \varrho)$ for the ball with the same center but half the radius. Clearly $B_1 \subset B_2 \subset U$ and $c_1(z) = 0$ for every $z \in B_2$.

Consider the function

$$\Psi(z) = \begin{cases} 0 & z \in B_1 & \text{Region I} \\ \psi(z) & z \in B_2 \setminus B_1, & \text{Region II} \\ 1 & z \in U \setminus B_2. & \text{Region III} \end{cases}$$

where ψ is a function, the exact form of which will be specified soon. For reasons that will become apparent shortly, we will require ψ such that $\frac{\partial \psi}{\partial n} = 0$ on ∂B_1. This function vanishes in B_1. If we show that **any** positive solution of (22) satisfies the property $u \leq M\,\Psi$ for some constant M large enough, then clearly any positive solution will develop a poverty core, which will be located within the region where c_1 vanishes.

For that, it is enough to show that for appropriate choice of M, $W := u - M\,\Psi \leq 0$, or equivalently, $W^+ := max(u - M\,\Psi, 0) = 0$. If $M \geq M_* = \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{1/(\gamma_2-\gamma_1)}$, then for $z \in U \setminus B_2$ (Region III), it clearly holds that $u(z) \leq M\,\Psi(z)$, so that $W^+(z) = 0$ for such z. If we show that W^+ does not vary with z, i.e., that $\nabla W^+(z) = 0$ a.e., then clearly $W^+(z) = 0$ for every $z \in U$ and our claim is thus valid. Note that the gradient of W^+ is considered in the weak sense. One can easily see that it is enough to show that

$$I := \int_U |\nabla W^+|^2 dz = \int_U \nabla W \cdot \nabla W^+ dz \leq 0.$$

By the choice of M and Ψ it is straightforward to note that

$$I = \int_U \nabla W \cdot \nabla W^+ dz = \int_{B_2} \nabla u \cdot \nabla W^+ dz - M \int_{B_2} \nabla \Psi \cdot \nabla W^+ dz$$
$$= \int_{B_2} \nabla u \cdot \nabla W^+ dz - M \int_{B_2 \setminus B_1} \nabla \Psi \cdot \nabla W^+ dz,$$

since $\nabla \Psi = 0$ on B_1. Applying Green's theorem on the last integral we conclude that

$$I = \int_{B_2} \nabla u \cdot \nabla W^+ dz + M \int_{B_2 \setminus B_1} \Delta \Psi \, W^+ dz$$
$$-M \int_{\partial B_1} \frac{\partial \psi}{\partial n} W^+ ds - M \int_{\partial B_2} \frac{\partial \psi}{\partial n} W^+ ds.$$

On ∂B_2 we have that $W^+ = 0$, so the last integral vanishes. We may eliminate the penultimate contribution by choosing ψ so that $\frac{\partial \psi}{\partial n} = 0$ on ∂B_1. With this choice,

$$I = \int_{B_2} \nabla u \cdot \nabla W^+ dz + M \int_{B_2 \setminus B_1} \Delta \Psi \, W^+ dz.$$

We now consider the first integral. Since $W^+ \in W^{1,2}(U)$, using W^+ as a test function in the weak form of (22) and noting that W^+ is concentrated on B_2 we find that

$$I = \int_{B_2} (c_1 u^{\gamma_1} - c_2 u^{\gamma_2}) W^+ dz + M \int_{B_2 \setminus B_1} \Delta \Psi \, W^+ dz$$
$$= - \int_{B_2} c_2 u^{\gamma_2} dz + M \int_{B_2 \setminus B_1} \Delta \Psi \, W^+ dz,$$

since $c_1 = 0$ in B_2. Express $B_2 = B_1 \cup (B_2 \setminus B_1)$ to obtain the estimate

$$I = - \int_{B_2} c_2 u^{\gamma_2} dz - \int_{B_2 \setminus B_1} c_2 u^{\gamma_2} dz + M \int_{B_2 \setminus B_1} \Delta \Psi \, W^+ dz$$
$$\leq \int_{B_2 \setminus B_1} (-c_2 u^{\gamma_2} + M \Delta \Psi) \, W^+ dz =: J,$$

by the positivity of c_2 and u. If we manage to show that $J \leq 0$ we are done.

Note that by the definition of W^+, we have that

$$J = \int_{B_2 \setminus B_1} (-c_2 u^{\gamma_2} + M \Delta \Psi)(u - M\Psi) \mathbf{1}_{\{u - M\Psi \geq 0\}} dz$$
$$= \int_{B_2 \setminus B_1} (-c_2 u^{\gamma_2} + M \Delta \psi)(u - M\psi) \mathbf{1}_{\{u - M\psi \geq 0\}} dz,$$

recalling that $\Psi = \psi$ on $B_2 \setminus B_1$, and since $(u - M\psi)\mathbf{1}_{\{u-M\psi\geq 0\}} \geq 0$ we simply need to show that $(-c_2 u^{\gamma_2} + M\Delta\psi)\mathbf{1}_{\{u-M\psi\geq 0\}} \leq 0$. Noting that since $c_2 \geq 0$,

$$\begin{aligned}(-c_2 u^{\gamma_2} + M\Delta\psi)\mathbf{1}_{\{u-M\psi\geq 0\}} \leq 0 &\leq -c_2 M^{\gamma_2}\psi^{\gamma_2} + M\Delta\psi \\ &\leq -\underline{c}_2 M^{\gamma_2}\psi^{\gamma_2} + M\Delta\psi\end{aligned}$$

we see that it is enough to choose ψ such that it satisfies the inequality

$$\begin{aligned}-\underline{c}_2 M^{\gamma_2}\psi^{\gamma_2} + M\Delta\psi \leq 0, &\text{ in } B_2 \setminus B_1 \\ \frac{\partial\psi}{\partial n} = 0, &\text{ on } \partial B_1, \\ \psi = 1, &\text{ on } \partial B_2.\end{aligned}$$

Since we have some liberty on the choice of ψ we assume that it is sperically symmetric, so that using the expression for the Laplacian in sperical coordinates and setting $r = |z|$ reduces the above inequality to an ODE inequality of the simpler form

$$\begin{aligned}-\underline{c}_2 M^{\gamma_2}\psi^{\gamma_2} + M\frac{d^2\psi}{dr^2} + M\frac{d-1}{r}\frac{d\psi}{dr} &\leq 0, \\ \frac{d\psi}{dr}(\varrho) = 0, \ \ \psi(2\varrho) &= 1.\end{aligned}$$

We look for solutions of this inequality of the form $\psi(r) = \varrho^{-\nu}(r-\varrho)^{\nu}$, for $\nu > 1$ which will be specified shortly. Note that for $\nu > 1$, ψ satisfies both boundary conditions. Substituting this ansatz into the differential inequality we obtain the equivalent condition,

$$\begin{aligned}\varrho^{-\gamma_2\nu}(r-\varrho)^{\gamma_2\nu}\{-\underline{c}_2 M^{\gamma_2-1} &+ \nu(\nu-1)\varrho^{-\nu+\gamma_2\nu}(r-\varrho)^{\nu-2-\gamma_2\nu} \\ &+(d-1)\nu\varrho^{-\nu+\gamma_2\nu}r^{-1}(r-\varrho)^{\nu-1-\gamma_2\nu}\} \leq 0,\end{aligned}$$

which leads to

$$\begin{aligned}S := -\underline{c}_2 M^{\gamma_2-1} &+ \nu(\nu-1)\varrho^{-\nu+\gamma_2\nu}(r-\varrho)^{\nu-2-\gamma_2\nu} \\ &+(d-1)\nu\varrho^{-\nu+\gamma_2\nu}r^{-1}(r-\varrho)^{\nu-1-\gamma_2\nu} \leq 0.\end{aligned}$$

Only the first term in S is negative, so the condition will hold only if the first term dominates the other two terms. Since in $B_2 \setminus B_1$ we have that $0 \leq r - \varrho \leq \varrho$, and we are interested in the limit where ϱ is small, it is clear that ν must be chosen so that the second and the third term in the expression for S do not blow up to $+\infty$. The worst term in this respect is the second one, since the second exponent on the term $(r-\varrho)$ is the smallest. We may choose then ν so as to eliminate this exponent, i.e., choose $\nu = \frac{2}{1-\gamma_2}$. Note that since $\gamma_2 < 1$ we have that $\nu > 1$ as required. For this choice

$$S = -\underline{c}_2 M^{\gamma_2 - 1} + \nu \varrho^2((\nu-1) + (d-1)r^{-1}(r-\varrho))$$

Since in $B_2 \backslash B_1$ we have that $0 \leq r - \varrho \leq \varrho$, we easily see that $0 \leq r^{-1}(r-\varrho) \leq 1$ so that

$$\begin{aligned} S &\leq -\underline{c}_2 M^{\gamma_2-1} + \nu\varrho^2((\nu-1)+(d-1)) \\ &\leq -\underline{c}_2 \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{\frac{\gamma_2-1}{\gamma_2-\gamma_1}} + \nu\varrho^2((\nu-1)+(d-1)), \end{aligned}$$

since M is initially chosen so that $M \geq M_* = \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{1/(\gamma_2-\gamma_1)}$. We therefore see that it is enough to choose the parameters such that

$$-\underline{c}_2 \left(\frac{\bar{c}_1}{\underline{c}_2}\right)^{\frac{\gamma_2-1}{\gamma_2-\gamma_1}} + \nu\varrho^2((\nu-1)+(d-1)) \leq 0.$$

We close the proof by proving that a poverty trap is impossible to develop if $\beta = 0$ or in cases C.1 and C.2. In these cases we may express system (22) as

$$-\Delta u + (-c_1 u^{\gamma_1 - 1} + c_2 u^{\gamma_2-1})u = 0.$$

Since we know that any solution u satisfies $0 \leq u \leq M_*$ and since the function $u \mapsto g(z,u) := -c_1 u^{\gamma_1-1} + c_2 u^{\gamma_2-1}$ is strictly increasing for cases I and II it is bounded above by a positive constant K for every $u \in [0, M^*]$ we conclude that

$$0 = \Delta u + (-c_1 u^{\gamma_1-1} + c_2 u^{\gamma_2-1})u \leq -\Delta u + Ku,$$

and an application of the strong maximum principle for $-\Delta u + Ku \geq 0$ leads to the result that u cannot vanish anywhere in the interior of U, therefore a poverty core may not develop in cases I and II. ■

A.5 Proof of Proposition 5

Proof. The proof uses a regularization argument, according to which we approximate the non-Lipschitz (with respect to the variable k) function f by a sequence of Lipschitz functions $\{f_\epsilon\}$ which converges in a monotone fashion to f as $\epsilon \to 0$, and in particular $f_\epsilon \uparrow f$. This is possible for any continuous function, for the present case one such possible regularization scheme is the sequence of functions

$$f_\epsilon(z,k) = \begin{cases} c_1(z)\epsilon^{\alpha-1}k - c_2(z)k, & 0 \le k \le \epsilon, \\ c_1(z)k^\alpha - c_2(z)k & k \ge \epsilon, \end{cases}$$

which is easily seen to satisfy the required properties. We now consider the approximate problem

$$\begin{aligned} &\frac{\partial w}{\partial t} = \Delta\Phi(w) + f_\epsilon(w), \text{ in } (U)_T, \\ &\frac{\partial \Phi(w)}{\partial n} = 0, \text{ on } (\partial U)_T, \\ &w(0,z) = k_0(z) + \epsilon, \ z \in U, \end{aligned} \tag{23}$$

which on account of the Lipschitz property of f_ϵ has a unique solution for every $\epsilon > 0$, the we will denote by w_ϵ. Furthermore, the comparison principle holds for (23), meaning that if $\underline{w}$ and $\bar{w}$ are a sub and super solution of (23) with $\underline{w}(0,z) \le \bar{w}(0,z)$, then $\underline{w}(t,z) \le \bar{w}(t,z)$ for every $t \in [0,T]$ and the same result holds for solutions (see e.g. [37]). Using the comparison principle for (23) repeatedly, we may conclude that $\{w_\epsilon\}$ is a non increasing sequence in ϵ which is bounded below by 0, hence the limit $k^\diamond := \lim_{\epsilon\to 0} w_\epsilon \ge 0$ is well defined and is also a weak solution of the original problem (17). We can now show that **any** solution k of (17) must satisfy the inequality $k \le k^\diamond$, hence $k^\diamond$ is the maximal solution. To show that consider the function $k_\epsilon := max(k,\epsilon)$ which can be shown to be a subsolution of (23) so that $k_\epsilon \le w_\epsilon$ by the comparison principle for (23). Furthermore since by definition $k \le k_\epsilon$ we see that $k \le k_\epsilon \le w_\epsilon$ for every $\epsilon > 0$ and passing to the limit as $\epsilon \to 0$ we obtain the required result $k \le k^\diamond$. Again by a comparison principle it is straightforward to see that $k^\diamond \le M_*$.

The minimal solution is constructed in terms of the solution of the regularized problem (23) with the sole difference that the initial condition is now replaced by k_0 (instead of $k_0 + \epsilon$). This is again a well posed problem with a unique solution which for any $\epsilon > 0$ will be denoted by v_ϵ. Repeated applications of the comparison principle for this problem l ead to the conclusion that $\{v_\epsilon\}$ is a non decreasing family in $\epsilon > 0$ bounded above hence the limit $k_\diamond := \lim_{\epsilon\to 0} v_\epsilon \ge 0$ is well defined. It can be shown further that for any weak solution k of (17) it holds that $k_\diamond \le k$, hence $k_\diamond$ is the minimal solution.

Suppose now that $k_0 \ge \epsilon_0 > 0$. Choose $\epsilon < \epsilon_0$ and consider the solution of the regularized problem (23) for this choice of ϵ and initial condition k_0. This problem has a unique solution and by comparison results as well as by the construction of the approximate family $\{f_\epsilon\}$ it is clear that $f_\epsilon(k) = f(k)$

so that the solution of the original problem coincides with the solution of the regularized problem and uniqueness thus follows. ■

A.6 Proof of Proposition 6

(i) The proof of (i) uses arguments very similar to the arguments used in proof of (ii) in the region $U \setminus B_2$ and is omitted.

(ii) Consider case C.3 ($\gamma_1 < \gamma_2 < 1$) and assume that the parameters of the problem are such that a poverty trap occurs. Using the notation of the proof of Proposition 4, we assume that c_1 vanishes in B_2 and that the steady state solution u_0 vanishes in B_1, where $B_1 \subset B_2 \subset U$. Consider also a solution of the time dependent problem u which is assumed to be close to u_0. The standard way of treating this problem would be to expres$u$ in terms of the expansion $u(t,z) = u_0(z) + \epsilon v(t,z)$ substitute that into the equation, linearize in terms of ϵ and obtain an equation for the perturbation v. Doing that, and taking into account that u_0 solves the steady state equation leads to the evolution equation

$$\gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} = \Delta v + \gamma_1 u_0^{\gamma_1 - 1} c_1 v - \gamma_2 u_0^{\gamma_2 - 1} c_2 v,$$

for the perturbation. The validity of this equation relies on a Taylor expansion of the nonlinearity, which in turn relies the property that the nonlinearity is C^1. This assumption clearly fails if u_0 is zero which is in fact what happens at the poverty trap case. This can be seen in the above linearized equation, and is displayed in the fact that the potentials $u_0^{\gamma_2 - 1}$ and $u_0^{\gamma_1 - 1}$ blow up to infinity in the core of the poverty trap, hence the linearized equation is not well posed.

The above comment shows that in the presence of a poverty trap a linearization argument is not possible and careful considerations are needed in order to study the stability of the steady state u_0. Consider the weak solutions of the problem

$$\frac{\partial}{\partial t} u^{\gamma_2} = \Delta u + c_1 u^{\gamma_1} - c_2 u^{\gamma_2},$$

and assume $u = u_0 + \epsilon v$ where $\epsilon > 0$ is a small parameter modelling the fact that we are looking for small "deviations" around the steady state solution u_0. We take the weak form of the equation using a test function $\varphi \in W^{1.2}(U)$, where $W^{1.2}(U)$ is the Sobolev space containing functions whose first order weak derivatives are in $L^2(U)$,

$$\int_{U_T} \frac{\partial}{\partial t} u^{\gamma_2} \varphi + \int_{U_T} \nabla u \cdot \nabla \varphi = \int_{U_T} (c_1 u^{\gamma_1} - c_2 u^{\gamma_2}) \varphi,$$

and to ease notation $U_T = [0,T]\times U$ and $\int_{U_T} g := \int_0^T \int_U g(t,z)dzdt$. We break U_T as $U_T = [0,T]\times(B_1\cup(B_2\setminus B_1)\cup(D\setminus B_2)) =: (B_1)_T\cup(B_2\setminus B_1)_T\cup(U\setminus B_2)_T$, and consider each of the integrals in the weak solution separately. We have that

$$\begin{aligned} I_1 := \int_{U_T} \frac{\partial}{\partial t} u^{\gamma_2}\varphi &= \int_{(B_1)_T} \epsilon^{\gamma_2}\frac{\partial}{\partial t} v^{\gamma_2}\varphi + \int_{(D\setminus B_1)_T} \frac{\partial}{\partial t}(u_0+\epsilon v)^{\gamma_2}\varphi \\ &= \int_{(B_1)_T} \epsilon^{\gamma_2}\frac{\partial}{\partial t} v^{\gamma_2}\varphi + \int_{(D\setminus B_1)_T} \epsilon\gamma_2 u_0^{\gamma_2-1}\frac{\partial v}{\partial t}\varphi + O(\epsilon) \end{aligned}$$

where we used the fact that $u_0 = 0$ for $z\in B_1$ and also used the fact that $(u_0+\epsilon v)^{\gamma_2} = u_0^{\gamma_2} + \gamma_2 u_0^{\gamma_2-1}v + O(\epsilon)$ for $z\in U\setminus B_1$, since $u_0\neq 0$ in this case.

For the second term

$$I_2 := \int_{U_T} \nabla u\cdot\nabla\varphi = \int_{U_T} \nabla u_0\cdot\nabla\varphi + \epsilon\int_{U_T} \nabla u_0\cdot\nabla\varphi,$$

on account of linearity.

For the third term

$$\begin{aligned} I_3 : &= \int_{U_T} c_1 u^{\gamma_1}\varphi = \int_{(U\setminus B_2)_T} c_1(u_0+\epsilon v)^{\gamma_1} \\ &= \int_{(U\setminus B_2)_T} c_1 u_0^{\gamma_1}\varphi + \int_{(U\setminus B_2)_T} \epsilon\gamma_1 c_1 u_0^{\gamma_1-1} v\varphi + O(\epsilon) \\ &= \int_{(U)_T} c_1 u_0^{\gamma_1}\varphi + \int_{(U\setminus B_2)_T} \epsilon\gamma_1 c_1 u_0^{\gamma_1-1} v\varphi + O(\epsilon), \end{aligned}$$

where we used the fact that $c_1 = 0$ on B_2 (so that $c_1 u_0^{\gamma_1}$ on B_2) and that $(u_0+\epsilon v)^{\gamma_1} = u_0^{\gamma_1} + \epsilon\gamma_1 u_0^{\gamma_1-1}v + O(\epsilon)$, in $(U\setminus B_2)_T$ (since $u_0\neq 0$ there and the function is differentiable).

Finally for the fourth term

$$\begin{aligned} I_4 := -\int_{(U)_T} c_1 u^{\gamma_1}\varphi &= -\int_{(B_1)_T} \epsilon^{\gamma_2} c_2 v^{\gamma_2}\varphi - \int_{(U\setminus B_1)_T} c_2 u_0^{\gamma_2}\varphi \\ &\quad - \int_{(U\setminus B_1)_T} \epsilon c_2 c_2 u_0^{\gamma_2-1} v\varphi + O(\epsilon) \\ = -\int_{(B_1)_T} \epsilon^{\gamma_2} c_2 v^{\gamma_2}\varphi &- \int_{(U)_T} c_2 u_0^{\gamma_2}\varphi - \int_{(U\setminus B_1)_T} \epsilon\gamma_2 c_2 u_0^{\gamma_2-1} v\varphi + O(\epsilon), \end{aligned}$$

since $u_0 = 0$ in $(B_1)_T$ and $(u_0+\epsilon v)^{\gamma_2} = u_0^{\gamma_2} + \epsilon\gamma_2 u_0^{\gamma_2-1}v + O(\epsilon)$ in $(U\setminus B_1)_T$ (since $u_0\neq 0$ there and the function is differentiable).

Collecting all the above in the weak form of the parabolic equation we obtain that

$$\int_{U_T} (\nabla u_0 \cdot \nabla - c_1 u_0^{\gamma_1} + c_2 u_0^{\gamma_2})\varphi + \epsilon^{\gamma_2} \int_{(B_1)_T} \left(\frac{\partial}{\partial t} v^{\gamma_2} + c_2 v^{\gamma_2} \right) \varphi$$
$$+\epsilon \left\{ \int_{(D \setminus B_1)_T} \gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} \varphi + \int_{U_T} \nabla u_0 \cdot \nabla \varphi - \int_{(U \setminus B_2)_T} \gamma_1 c_1 u_0^{\gamma_1 - 1} v \varphi \right.$$
$$\left. + \int_{(U \setminus B_1)_T} \gamma_2 c_2 u_0^{\gamma_2 - 1} v \varphi \right\} + O(\epsilon) = 0.$$

Note the different orders of magnitude in the expansion, which now overcomes the technical difficulties of linearizing around the poverty core. The zeroth order term in ϵ is identified as the weak form of the steady state PDE and since u_0 is the steady state solution this term vanishes. When we are within the poverty core, i.e., in $(B_1)_T$, the next significant order of magnitude which is ϵ^{γ_2} (recall that $\gamma_2 < 1$) is activated yielding that

$$\int_{(B_1)_T} \left(\frac{\partial}{\partial t} v^{\gamma_2} + c_2 v^{\gamma_2} \right) \varphi = 0,$$

which is the weak form for

$$\frac{\partial}{\partial t} v^{\gamma_2} + c_2 v^{\gamma_2} = 0, \;\text{ in } (B_1)_T. \tag{24}$$

Finally, outside the poverty core, i.e., in $(U \setminus B_1)_T$, the next order of magnitude ϵ^1 is activated, yielding,

$$\int_{(D \setminus B_1)_T} \gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} \varphi + \int_{U_T} \nabla u_0 \cdot \nabla \varphi - \int_{(U \setminus B_2)_T} \gamma_1 c_1 u_0^{\gamma_1 - 1} v \varphi$$
$$+ \int_{(U \setminus B_1)_T} \gamma_2 c_2 u_0^{\gamma_2 - 1} v \varphi = 0.$$

As this is true for any test function φ, if we first consider a test function concentrated on $(B_2 \setminus B_1)_T$, and since $c_1 = 0$ on this set, we get that

$$\int_{(B_2 \setminus B_1)_T} \gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} \varphi + \int_{(B_2 \setminus B_1)_T} \nabla u_0 \cdot \nabla \varphi + \int_{(B_2 \setminus B_1)_T} \gamma_2 c_2 u_0^{\gamma_2 - 1} v \varphi = 0,$$

which is the weak form for

$$\gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} = \Delta v - \gamma_2 c_2 u_0^{\gamma_2 - 1} v, \;\; z \in B_2 \setminus B_1. \tag{25}$$

If we consider now a test function concentrated on $(U \setminus B_2)_T$, we get that

$$\int_{(U\setminus B_2)_T} \left(\gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} \varphi + \nabla u_0 \cdot \nabla + \gamma_1 c_1 u_0^{\gamma_1 - 1} v - \gamma_2 c_2 u_0^{\gamma_2 - 1} v \right) \varphi = 0,$$

which is the weak form for

$$\gamma_2 u_0^{\gamma_2 - 1} \frac{\partial v}{\partial t} = \Delta v + \gamma_1 c_1 u_0^{\gamma_1 - 1} v - \gamma_2 c_2 u_0^{\gamma_2 - 1} v, \;\; z \in U \setminus B_2. \tag{26}$$

The evolution of a small perturbation of the poverty trap up to $O(\epsilon)$ is governed by the evolution equations (24), (25) and (26). Within the poverty core, in B_1, and since $c_2 > 0$, equation (24) can be explicitly estimated to give that

$$\frac{\partial}{\partial t} v^{\gamma_2} = -c_2 v^{\gamma_2} \leq -\underline{c}_2 v^{\gamma_2}$$

which gives

$$v(t,z)^{\gamma_2} \leq v(0,z)^{\gamma_2} e^{-\underline{c}_2 t}, \;\; t > 0, \;\; z \in B_1,$$

hence

$$v(t,z) \leq v(0,z) e^{-\frac{\underline{c}_2}{\gamma_2} t}, \;\; t > 0, \;\; z \in B_1, \tag{27}$$

which leads to the conclusion that $\lim_{t\to\infty} v(t,z) = 0$ for any $z \in B_1$, a fact that guarantees the exponential stability of the poverty core. The above argument can be modified if $v(0,x)$ is not continuous slightly, by taking the spatial average of (24) over B_1, and obtaining the exponential estimate $\|v(t,\cdot)\|_{L^{\gamma_2}(B_1)} \leq \|v(0,\cdot)\|_{L^{\gamma_2}(B_1)} e^{-\frac{\underline{c}_2}{\gamma_2} t}$.

The situation is similar in $B_2 \setminus B_1$. In this region the evolution of the perturbation v is given by the linear PDE (25) which may also be expressed in the equivalent form

$$\frac{\partial v}{\partial t} = \frac{1}{\gamma_2} u_0^{1-\gamma_2} \Delta v - c_2 v, \;\; z \in B_2 \setminus B_1. \tag{28}$$

This equation is still singular, but only on the boundary ∂B_1. This happens since by continuity, $u_0(z) = 0$ for $z \in \partial B_1$, so that the diffusion coefficient vanishes for $z \in \partial B_1$, however it is nonzero elsewhere. Furthermore, by the estimates in the proof of Proposition 4 we know that $u_0 \leq C\Psi$ and in the relevant region $(B_2 \setminus B_1)$ we have that $\Psi(z) = \varrho^{-\nu}(|z| - \varrho)^{\nu}$, with $\nu = \frac{2}{1-\gamma_2}$ so that we have the estimate $0 \leq u_0^{1-\gamma_2} \leq C||z| - \varrho|^2$ for an appropriate constant

C. Equation (28) is a linear equation. Looking for separable solutions of the form $v(t,z) = e^{\lambda t} w(z)$, upon substituting into the equation yields

$$-\frac{1}{\gamma_2} u_0^{1-\gamma_2} \Delta w = -(\lambda + c_2) w,$$

or equivalently

$$-\Delta w + \gamma_2 u_0^{\gamma_2 - 1} (\lambda + c_2) w = 0. \tag{29}$$

If $\lambda < 0$ then $v(t,z) = e^{\lambda t} w(z)$ will tend to 0 as $t \to \infty$, hence the poverty trap will be stable. This can be shown by the folowing simple argument. Multiply (29) by w and integrate over $B_2 \setminus B_1$, to obtain

$$\int_{B_2 \setminus B_1} |\nabla w|^2 + \gamma_2 \int_{B_2 \setminus B_1} u_0^{\gamma_2 - 1} c_2 w^2 + \lambda \gamma_2 \int_{B_2 \setminus B_1} u_0^{\gamma_2 - 1} w^2 = 0,$$

which can be rearranged as

$$\lambda \gamma_2 \int_{B_2 \setminus B_1} u_0^{\gamma_2 - 1} w^2 = - \int_{B_2 \setminus B_1} |\nabla w|^2 - \gamma_2 \int_{B_2 \setminus B_1} u_0^{\gamma_2 - 1} c_2 w^2,$$

which, since $c_2 > 0$, leads to the result that $\lambda < 0$.

Finally, for $z \in U \setminus B_2$, the evolution of the perturbation is governed by (26) which can be rearranged as

$$\frac{\partial v}{\partial t} = \frac{1}{\gamma_2} u_0^{1-\gamma_2} \Delta v + \frac{\gamma_1}{\gamma_2} c_1 u_0^{\gamma_1 - \gamma_2} v - c_2 v, \quad z \in U \setminus B_2,$$

the difference now being that for $z \in U \setminus B_2$, we are way out of the core of the poverty trap and therefore there exists $K > 0$ such that $u_0(z) > K$ for all such z. This is now a standard linear diffusion equation, with non vanishing diffusion coefficient whose asymptotic behaviour is determined by a weighted eigenvalue problem. As before, assume that we look for separable solutions of (26) of the form $v(t,z) = e^{-\Lambda t} w(z)$, (here for reasons that will become obvious soon we use the reparametrization $\lambda = -\Lambda$). Upon substitution of this ansatz into (26) we see that Λ and w must be solutions of the weighted eigenvalue problem

$$(-\Delta + \mathcal{A}(z)) w = \Lambda \mathcal{W}(z) w, \tag{30}$$

where $\mathcal{A}$ and $\mathcal{W}$ is a potential and weight function given respectively by

$$\mathcal{A} = \gamma_2 c_2 u_0^{\gamma_2 - 1} - \gamma_1 c_1 u_0^{\gamma_1 - 1},$$
$$\mathcal{W} = \gamma_2 u_0^{\gamma_2 - 1}.$$

If the principle eigenvalue of the weighted eigenvalue problem (30) is positive then $\lambda < 0$ (recall that $\lambda = -\Lambda$) hence we have linearized stability. The existence of a principle eigenvalue follows from the results of [18] (see Theorem 1.1, op cit). To check the positivity of Λ we consider the following simple argument, based on the method of sub and super solutions. According to [33] (Chapter 2, Section 8) [15] if $\mathcal{W} > 0$ and for some $\sigma \in \mathbb{R}$ we may find a function $\bar{w}$ such that

$$J(\bar{w}) := (-\Delta + \mathcal{A}(z) - \sigma\mathcal{W})\bar{w} \geq 0,$$

then $Re(\Lambda) > \sigma$. If we manage to find a supersolution $\bar{w}$ of the above problem for some $\sigma > 0$ then our claim is valid. Let us choose $\bar{w} = u_0$. Substituting into J we get that

$$\begin{aligned} J(u_0) &= -\Delta u_0 + \gamma_2 c_2 u_0^{\gamma_2} - \gamma_1 c_1 u_0^{\gamma_1} - \sigma\gamma_2 u_0^{\gamma_2} \\ &= (1-\gamma_1)c_1 u_0^{\gamma_1} - \{(1-\gamma_2)c_2 + \sigma\gamma_2\} u_0^{\gamma_2} \\ &= u^{\gamma_1}[(1-\gamma_1)c_1 - \{(1-\gamma_2)c_2 + \sigma\gamma_2\} u_0^{\gamma_2-\gamma_1}] \\ &\geq u^{\gamma_1}[(1-\gamma_1)\underline{c}_1 - \{(1-\gamma_2)\bar{c}_2 + \sigma\gamma_2\} u_0^{\gamma_2-\gamma_1}] \end{aligned}$$

If $\underline{c}_1 \neq 0$, we see that the first term in the bracket is positive, and since $\gamma_2 < 1$, the second term will be negative. We will therefore, expect $J(u_0)$ to be positive as long as u_0 is small enough and in particular as long as

$$u_0 \leq \left(\frac{(1-\gamma_1)\underline{c}_1}{(1-\gamma_2)\bar{c}_2 + \sigma\gamma_2}\right)^{\frac{1}{\gamma_2-\gamma_1}} \tag{31}$$

Recall that $u_0 \leq \left(\frac{\underline{c}_1}{\bar{c}_2}\right)^{\frac{1}{\gamma_2-\gamma_1}}$. This implies that condition (31) will definitely hold as long as we may find a $\sigma > 0$ such that

$$\frac{\underline{c}_1}{\bar{c}_2} \leq \frac{(1-\gamma_1)\underline{c}_1}{(1-\gamma_2)\bar{c}_2 + \sigma\gamma_2}.$$

Simple algebra shows that this will always hold as long as

$$0 < \sigma < \frac{\gamma_2 - \gamma_1}{\gamma_2}\bar{c}_2,$$

so that a possible choice for σ is e.g. $\sigma^* = \frac{\gamma_2-\gamma_1}{2\gamma_2}\bar{c}_2$ (or in fact $\sigma^* = \frac{\gamma_2-\gamma_1}{2\gamma_2}\bar{c}_2 - \varepsilon$, for every $\varepsilon > 0$). If $\underline{c}_1 = 0$ as would happen for example on ∂B_2 then a different choice for supersolution w can be used, for example $w(z) = 1 - e^{-\zeta|z|}$ for $\zeta > 0$ and large enough.

[15] Note that in this reference the convention is to work if the operator Δ rather than Δ which explains the change in the signs.

References

[1] Daron Acemoglu. *Introduction to modern economic growth.* MIT press, 2009.

[2] Shmuel Agmon, Avron Douglis, and Louis Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. i. *Communications on Pure and Applied Mathematics*, 12(4):623–727, 1959.

[3] Antonio Ambrosetti and David Arcoya Álvarez. *An introduction to nonlinear functional analysis and elliptic problems*, volume 82. Birkhauser, 2011.

[4] Richard Baldwin, Rikard Forslid, Philippe Martin, Gianmarco Ottaviano, and Frédéric Robert-Nicoud. Public policies and economic geography. *PUP, Princeton*, 2003.

[5] Richard E Baldwin and Philippe Martin. Agglomeration and regional growth. *Handbook of regional and urban economics*, 4:2671–2711, 2004.

[6] Richard E Baldwin, Philippe Martin, and Gianmarco IP Ottaviano. Global income divergence, trade, and industrialization: the geography of growth take-offs. *Journal of Economic Growth*, 6(1):5–37, 2001.

[7] Robert J Barro and Xavier Sala-i Martin. Economic growth: Mit press. *Cambridge, Massachusettes*, 2004.

[8] Raouf Boucekkine, Carmen Camacho, and Giorgio Fabbri. On the optimal control of some parabolic partial differential equations arising in economics. *AMSE, WP 2013-Nr 34*, 2013.

[9] Raouf Boucekkine, Carmen Camacho, and Giorgio Fabbri. Spatial dynamics and convergence: The spatial ak model. *Journal of Economic Theory*, 148(6):2719–2736, 2013.

[10] Raouf Boucekkine, Carmen Camacho, and Benteng Zou. Bridging the gap between growth theory and the new economic geography: The spatial ramsey model. *Macroeconomic Dynamics*, 13(1):20, 2009.

[11] Haïm Brezis and Luc Oswald. Remarks on sublinear elliptic equations. *Nonlinear Analysis: Theory, Methods & Applications*, 10(1):55–64, 1986.

[12] Paulo Brito. The dynamics of growth and distribution in a spatially heterogeneous world. 2004.

[13] William Brock and Anastasios Xepapadeas. Diffusion-induced instability and pattern formation in infinite horizon recursive optimal control. *Journal of Economic Dynamics and Control*, 32(9):2745–2787, 2008.

[14] William Brock and Anastasios Xepapadeas. Pattern formation, spatial externalities and regulation in coupled economic–ecological systems. *Journal of Environmental Economics and Management*, 59(2):149–164, 2010.

[15] William A Brock, Anastasios Xepapadeas, and Athanasios N Yannacopoulos. Optimal agglomerations in dynamic economics. *Journal of Mathematical Economics*, 53:1–15, 2014.

[16] William A Brock, Anastasios Xepapadeas, and Athanasios N Yannacopoulos. Spatial externalities and agglomeration in a competitive industry. *Journal of Economic Dynamics and Control*, 42:143–174, 2014.

[17] Carmen Camacho and Benteng Zou. The spatial solow model. *Econ. Bull*, 18:1–11, 2004.

[18] Santiago Cano-Casanova and Julián López-Gómez. Properties of the principal eigenvalues of a general class of non-classical mixed boundary value problems. *Journal of Differential Equations*, 178(1):123–211, 2002.

[19] Dean A Carlson, Alain Haurie, and Arie Leizarowitz. *Infinite horizon optimal control: deterministic and stochastic systems*. Springer-verlag, 1991.

[20] Timothy G Conley and Ethan Ligon. Economic distance and cross-country spillovers. *Journal of Economic Growth*, 7(2):157–187, 2002.

[21] Manuel Delgado and Antonio Suárez. On the existence of dead cores for degenerate lotkaï£¡volterra models. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 130(04):743–766, 2000.

[22] Klaus Desmet and Esteban Rossi-Hansberg. Spatial growth and industry age. *Journal of Economic Theory*, 144(6):2477–2502, 2009.

[23] Klaus Desmet and Esteban Rossi-Hansberg. On spatial dynamics*. *Journal of Regional Science*, 50(1):43–63, 2010.

[24] Masahisa Fujita and Tomoya Mori. Frontiers of the new economic geography*. *Papers in Regional Science*, 84(3):377–405, 2005.

[25] David Gilbarg and Neil S Trudinger. *Elliptic partial differential equations of second order.* springer, 2015.

[26] W. Isard and P. Liossatos. *Spatial Dynamics and Optimal Space-Time Development.* North-Holland, Amsterdam, 1978.

[27] Walter Isard and Panagis Liossatos. On optimal development over space and time. *Journal of Regional Analysis and Policy*, 3(1), 1973.

[28] Walter Isard and Panagis Liossatos. Transport investment and optimal space-time development. *Papers in Regional Science*, 31(1):31–50, 1973.

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

[30] Robert E Lucas. Macroeconomic priorities. *American Economic Review*, 93(1):1–14, 2003.

[31] Philippe Martin and Gianmarco IP Ottaviano. Growth and agglomeration. *International Economic Review*, 42(4):947–968, 2001.

[32] C Morales-Rodrigo and Antonio Suárez. Uniqueness of solution for elliptic problems with nonlinear boundary conditions. *Comm. Appl. Nonlinear Anal*, 13(3):69–78, 2006.

[33] Murray H Protter and Hans F Weinberger. *Maximum principles in differential equations.* Springer Science & Business Media, 2012.

[34] Danny T Quah. Empirics for economic growth and convergence. *European economic review*, 40(6):1353–1375, 1996.

[35] Danny T Quah. Regional convergence clusters across europe. *European economic review*, 40(3):951–958, 1996.

[36] Danny T Quah. Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of economic growth*, 2(1):27–59, 1997.

[37] WH Ruan. Monotone iterative method for degenerate nonlinear parabolic equations. *Nonlinear Analysis: Theory, Methods & Applications*, 34(1):37–63, 1998.

[38] Klaus Schmitt. Revisiting the method of sub-and supersolutions for nonlinear elliptic problems. *Electronic Journal of Differential Equations*, page 377, 2007.

[39] Juan Luis Vázquez. *The Porous Medium Equation: Mathematical Theory: Mathematical Theory.* Oxford University Press, 2006.

[40] Xu-Jia Wang. Neumann problems of semilinear elliptic equations involving critical sobolev exponents. *Journal of Differential equations*, 93(2):283–310, 1991.

Figures

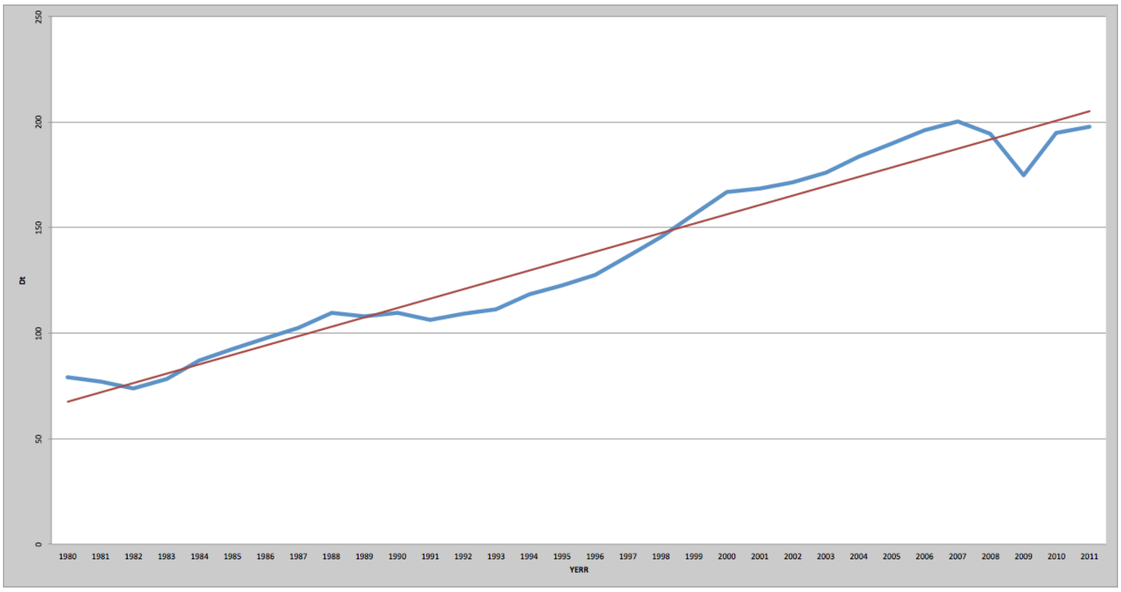


Figure 1. Regional inhomogeneity measure

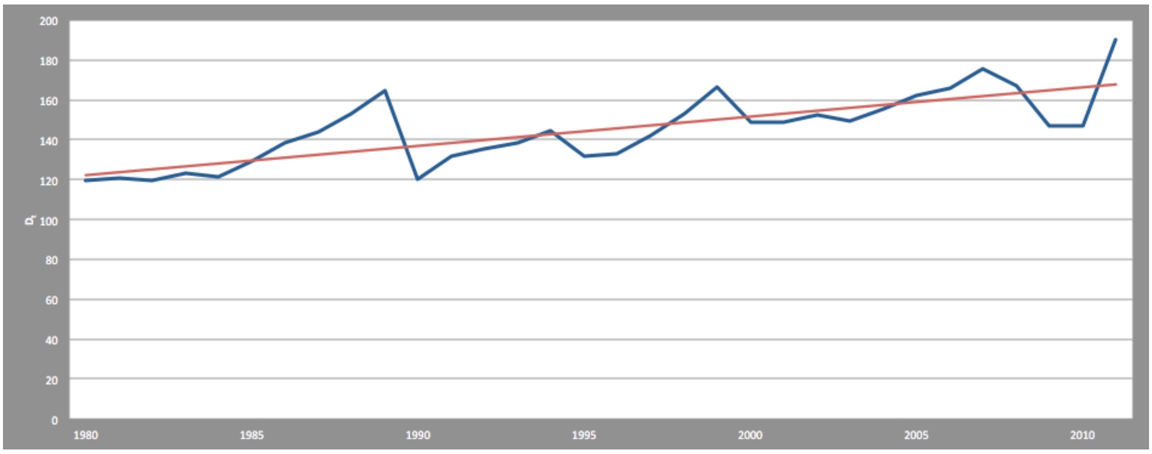


Figure 2. Inhomogeneity measure, high income countries

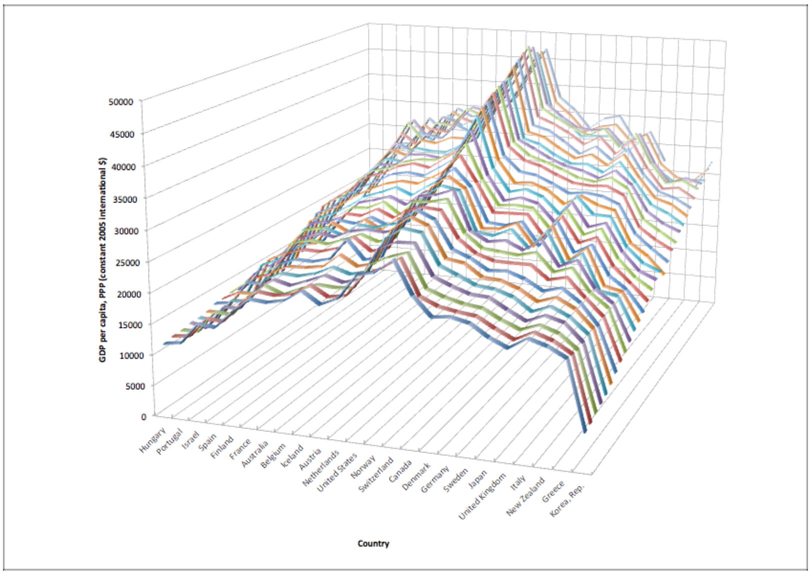


Figure 3. Distribution of GDP per capita in economic space

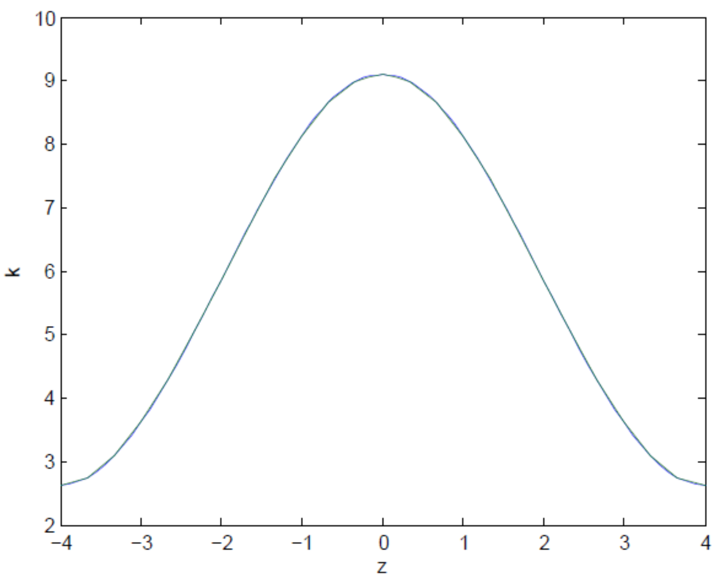
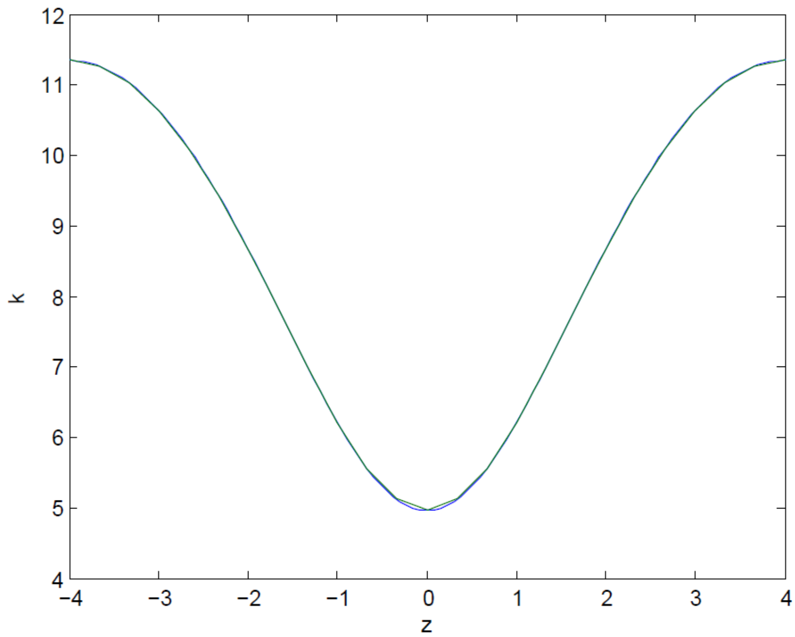
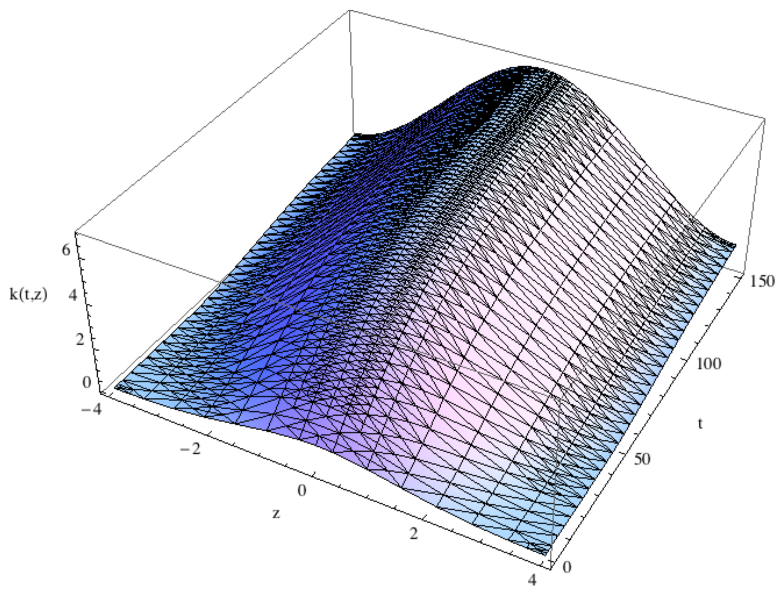


Figure 4. Steady state distribution of capital stock



[Figure 5. Poverty core at the steady state]



[Figure 6 Spatiotemporal distribution of capital]

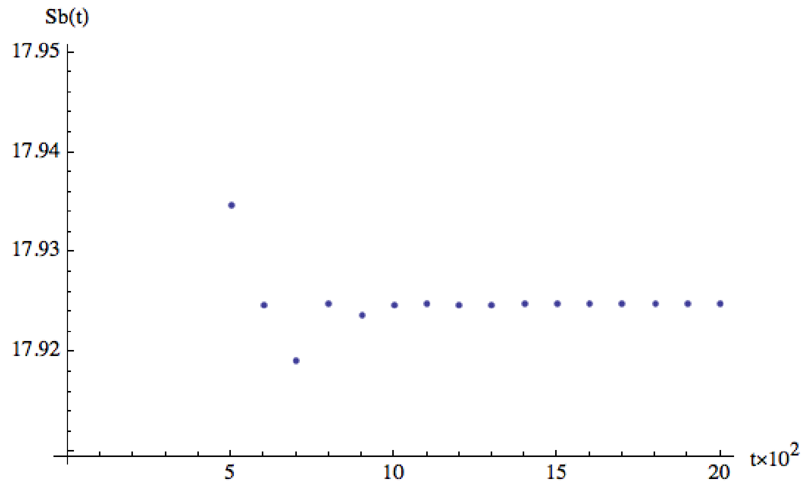
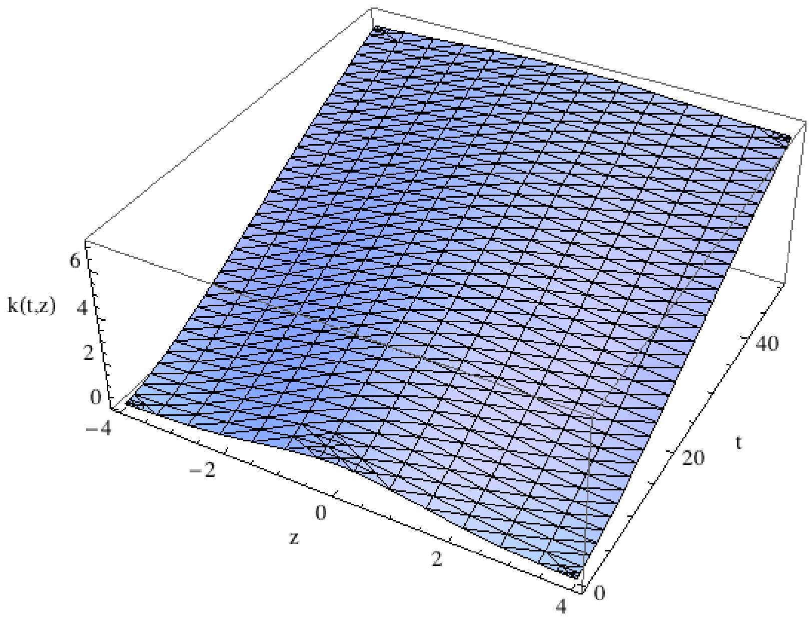
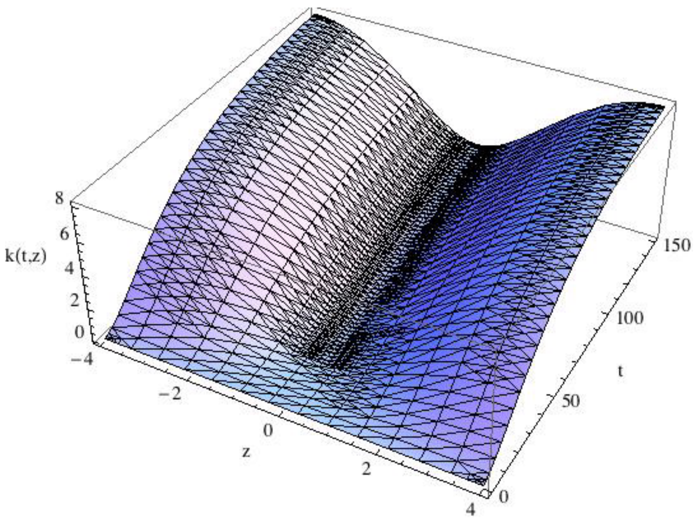


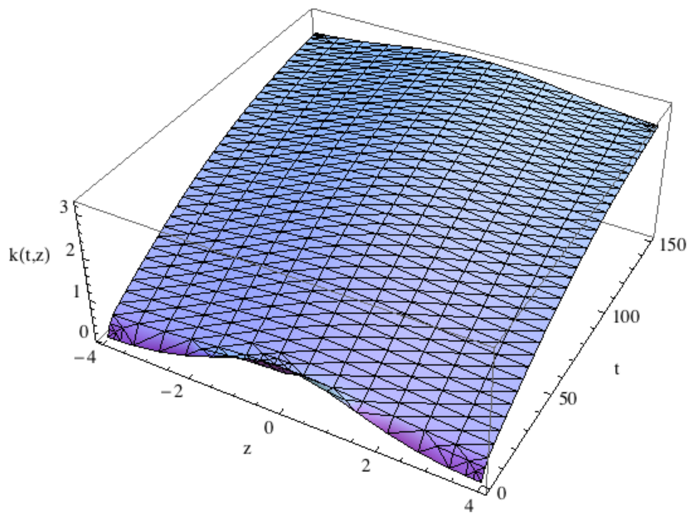
Figure 7. The time path of the Sobolev norm



[Figure 8. A flat spatiotemporal distribution of capital]



[Figure 9. Spatiotemporal evolution of a poverty trap]



[Figure 10. The spatiotemporal evolution for the trade balance model]