DOCUMENT DE RECHERCHE

EPEE

CENTRE D'ETUDES DES POLITIQUES ECONOMIQUES DE L'UNIVERSITE D'EVRY

Ecological Barriers and Convergence: A Note on Geometry in Spatial Growth Models

G. Fabbri

14-05

ECOLOGICAL BARRIERS AND CONVERGENCE: A NOTE ON GEOMETRY IN SPATIAL GROWTH MODELS

GIORGIO FABBRI

ABSTRACT. We introduce an AK spatial growth model with a general geographical structure. The dynamics of the economy is described by a partial differential equation on a Riemannian manifold. The morphology interacts with the spatial dynamics of the capital and is one determinant of the qualitative behavior of the economy. We characterize on the geographical structure the conditions that guarantee, in the long run, the convergence of the detrended capital across locations and those inducing spatial capital agglomeration.

Key words: Dynamical spatial model; growth; agglomeration; convergence; infinite dimensional optimal control problems; Riemannian manifolds.

Journal of Economic Literature Classification: R1; O4; C61.

1. INTRODUCTION

> *[...] The Americas were [more] fragmented by areas unsuitable for food production or for dense human populations. These ecological barriers included the rain forests of the Panamanian isthmus separating Mesoamerican societies from Andean and Amazonian societies; the deserts of northern Mexico separating Mesoamerica from U.S. southwestern and southeastern societies; dry areas of Texas [...]. As a result, there was no diffusion of domestic animals, writing, or political entities, and limited or slow diffusion of crops and technology.*

> Jared Diamond, "Guns, germs and steel", 1997.

Even if the theoretical importance of geography in development processes was already clear more than two centuries ago (see *e.g.* Smith, 1776, Book 1, Chapter 3), the effort to merge the continuous spatial dimension with benchmark growth theory models is rather recent.

In a seminal paper, Brito (2004) first[1] introduced spatial capital accumulation and capital mobility in the Ramsey growth framework. In Brito's model the population lives on a straight line. Production and capital accumulation are distributed in space and capital differentials drive the spatial capital dynamics. Boucekkine et al. (2009) further improved and studied this model in the linear utility case. More recent contributions in the same stream are the study of the endogenous growth case by Brito (2012) and the characterization of the optimal dynamics of the AK model on the a circle *à la Salop* described by Boucekkine et al. (2013)[2],

[1]Quah (2002), with his study, can be considered a precursor of this stream of literature but there the unique production factor ("knowledge") is exogenously generated and no spatial factor mobility is considered.

[2]Brock and Xepapadeas (2008) present a methodological contribution for a class of models including the spatial growth models, focusing again on the circle case.

see also the generalization proposed by Aldashev et al. (2014). A different approach considers spatial spillover and excludes capital mobility. This allows a technical simplification of the problem since the diffusion term disappears from the state equation; this is the method chosen for example by Brock and Xepapadeas (2009) and Mossay (2013). For an interesting survey on the subject, see the paper by Desmet and Rossi-Hansberg (2010). A longer list of references is available in the introduction of the paper by Brock et al. (2014).

The contribution of the present note is taking into account the specific role of the geography structure in the growth process and in the *agglomeration vs convergence* long run behavior of the system. In fact, the mentioned literature refers to several models of space: the straight line, the segment, the circle. Still, apart from a reflexion on the role of the "right" boundary conditions in the state equation, the study of the specific role of the chosen geographic structure was never taken into account. As we argue in this paper, there is probably a technical reason that could explain why this aspect of the problem has not been considered yet.

In this note we present a spatial growth model with the same AK production function as considered by Boucekkine et al. (2013), with the same law of motion of capital[3] but with a generic geographic structure. The morphology turns out to interact with the spatial dynamics of the capital and to be one determinant of the qualitative optimal behavior of the system. Keeping all the other parameters fixed and changing only the geographic structure may lead to a completely different qualitative behavior of the economy. Above all, the convergence result found by Boucekkine et al. (2013) is proved to be a particular case of a more complex picture that includes, when varying preferences parameters and geographical environment, the possibility of long run convergence on one side and clustering and agglomeration on the other.

As already mentioned, in order to see the whole picture, there is a price to pay in terms of mathematical complexity. Indeed, when the general geographic structure is taken into account, the planner optimization problem leads to an optimal control problem driven by a partial differential equation on a Riemannian manifold. So, in addition to the difficulties led by the infinite dimensional structure of the problems that appear in the previously mentioned spatial growth models, the role played by the metric structure of the manifold remains a specific challenge.

After having redrafted the model in the form of an optimal control problem (Section 2), we find (Section 3.1) its explicit solution in closed form, describing the optimal dynamics of the spatial distribution of the capital as the solution of a parabolic equation on the geography M, connected, compact and without boundary. Such a spatio-temporal equation describes the evolution of the economy in the whole transition towards convergence or agglomeration. The proofs, contained in Appendix A, make use of dynamic programming in a (infinite dimensional) Hilbert space. This result has a specific methodological interest in itself since it is, in our knowledge, the first optimal control problem driven by a diffusion equation on an abstract manifold (solved and) used in the literature.

The main contribution of this paper is contained in Section 3.2 where the role of the geography M fully appears. In Theorem 3.5 a sharp condition involving the total factor productivity, the discount rate, the elasticity of intertemporal substitution and an index that sums up the geometric characteristics of the geography, distinguishes situations where the spatial distribution of the (detrended) capital tends to an homogeneous distribution in the long run from situations of long-run capital agglomeration and cluster formation.

—————————

[3]It is in fact the law of motion initially introduced by Brito (2004).

Besides making the model tractable from an analytical point of view, the use of the linear production function with a total factor productivity constant in time and in space allows us to emphasize that, in the model, the returns of scale have no role in convergence and agglomeration phenomena: the spatio-temporal dynamics, the preferences of the planner and the geographic structure do the job.

This note is organized as follows. Section 2 introduces the model setup and presents its main features. Section 3 displays the analytical results and explains their economic implications. Section 4 concludes. Appendix A contains all the proofs.

2. The model

We consider an economy developing on a geography M, modeled as an n-dimensional compact, connected, oriented Riemannian manifold (with metric g) without boundary. Examples of geography with this structure are for example the Salop circle or the earth surface. This kind of structure allows to avoid the problems related to the way of assigning spatial-boundary conditions (at each time) in the evolution of the economy, see Brito (2012) or Boucekkine et al. (2009): one could indeed argue that convergence versus agglomeration questions naturally arise at a global level.

Through the paper we denote by $x \in M$ a generic (spatial) point in the geography and by $k(t,x)$ the (spatial density of) capital at space point x and time t. The population is assumed to be constant in time and uniformly distributed so $k(t,x)$ is the per-capita distribution of capital. We denote by $k_0(\cdot)$ the initial capital distribution: $k_0 \colon M \to \mathbb{R}$ is then a function representing for each point x the initial spatial density of capital at space point x.

We consider an AK production structure for a (spatially and temporally homogeneous) level of technology A. If $\tau(t,x)$ the trade balance at point x and time t, the evolution of the capital is given by

$$(1) \qquad \frac{\partial k(t,x)}{\partial t} = Ak(t,x) - c(t,x) - \tau(t,x)$$

where $c(t,x) \geq 0$ denotes the consumption at space point x and time t. Using the same argument as [4] Brito (2004) and several others for the one-dimensional case, given a region B (here a connected open subset of M with regular boundary) the trade balance over B is given by the sum (*i.e.* the integral) of what enters each point of the boundary ∂B, in formulas:

$$(2) \qquad \int_B \tau(t,x)\,\mathrm{d}x = -\int_{\partial B} \frac{\partial k(t,x)}{\partial \hat{n}}\,\mathrm{d}x.$$

We apply the divergence theorem to the last expression and we obtain

$$\int_{\partial B} \frac{\partial k(t,x)}{\partial \hat{n}}\,\mathrm{d}x = \int_B \nabla \cdot \nabla k(t,x)\,\mathrm{d}x = \int_B \Delta_x k(t,x)\,\mathrm{d}x$$

so that, for almost every $x \in M$,

$$(3) \qquad -\tau(t,x) = \Delta_x k(t,x)$$

where Δ_x is the Laplace-Beltrami operator (from now simply *Laplacian*) on the geography M. It reduces to the spatial second derivative when we use a one-dimensional model for the space; it is the case in the following papers: Brito (2004, 2012); Brock and Xepapadeas (2008); Boucekkine et al. (2009, 2013); Aldashev et al. (2014).

---

[4]In fact Brito (2004) is the first to adapt an idea coming from classical spatial economics (see *e.g.* Chapter 8 of Beckmann and Puu, 1985 or Isard et al., 1979) to the benchmark growth models context. The same idea was later used by several other authors, among them Camacho et al. (2008); Brock and Xepapadeas (2008); Boucekkine et al. (2009); Brito (2012); Boucekkine et al. (2013); Aldashev et al. (2014).

Using relation (3) in (1), provided an initial distribution of capital $k_0(\cdot)$ on M and given a consumption profile $c(\cdot, \cdot)$, we finally have a partial differential equation on M that describes the evolution of the capital density $k(t, x)$:

$$(4) \qquad \begin{cases} \dfrac{\partial k(t, x)}{\partial t} = \Delta_x k(t, x) + A k(t, x) - c(t, x) \\[2mm] k(0, x) = k_0(x). \end{cases}$$

The policy maker chooses the consumption $c(\cdot, \cdot)$ to maximize the following CRRA-Benthamite utility functional

$$(5) \qquad J(c(\cdot, \cdot)) = \int_0^{+\infty} e^{-\rho t} \int_M \frac{(c(t, x))^{1-\sigma}}{1 - \sigma} \, \mathrm{d}x \, \mathrm{d}t.$$

3. SPATIO-TEMPORAL DYNAMICS AND CONVERGENCE: THE SOLUTION OF THE MODEL

3.1. The explicit solution of the model.
In this subsection we present the solution of the described model. The proofs of the results are in Appendix A.

First, we describe the behavior of the aggregate capital $K(t) := \int_M k(t, x) \, \mathrm{d}x$ and of the aggregate consumption $C(t) := \int_M c(t, x) \, \mathrm{d}x$. Their evolution can be sketched by a simple one-dimensional differential equation, as shown by the following proposition.

Proposition 3.1. *The dynamics of $K(t)$ is described by*

$$(6) \qquad \dot{K}(t) = A K(t) - C(t), \qquad K(0) = \int_M k_0(x) \, \mathrm{d}x.$$

In other words, the proposition tells us that at the aggregate level the model is equivalent to the standard one-dimensional AK model: we are looking at the internal spatial dynamics of a benchmark AK model without altering the global structure.

Since we solve the problem using dynamic programming in the Hilbert space $L^2(M)$ of the square integrable functions[5] from M to $\mathbb{R}$ (see Appendix A for details) we will make use of the *value function* of the problem. For a given initial capital distribution $k_0(\cdot)$ we define the value function of our problem starting from $k_0(\cdot)$ as

$$(8) \qquad V(k_0) := \sup_{c(\cdot, \cdot)} J(k_0, c(\cdot, \cdot))$$

where the supremum is calculated by varying the positive spatio-temporal consumption distributions that ensure the aggregate capital to remain non-negative at any time. $V(k_0)$ corresponds to the maximal (utilitarian) aggregate welfare that can be guaranteed by the planner for a given initial capital distribution $k_0(\cdot)$. Its form can be described explicitly by the following proposition.

Proposition 3.2. *Suppose that*

$$(9) \qquad \rho > A(1 - \sigma)$$

and consider an initial positive capital distribution $k_0 \in L^2(M)$. Then the explicit expression of the value function of the problem at point k_0 is

$$V(k_0) = \frac{1}{1 - \sigma} \left(\frac{\rho - A(1 - \sigma)}{\sigma \, vol(M)} \right)^{-\sigma} \left(\int_M k_0(x) \, \mathrm{d}x \right)^{1-\sigma},$$

where $vol(M) := \int_M 1 \, \mathrm{d}x$ is the volume of the geography M.

---

[5]More formally $L^2(M)$ is the set:

$$(7) \qquad L^2(M) := \left\{ f \colon M \to \mathbb{R} \; : \; \int_M |f(x)|^2 \, \mathrm{d}x < \infty \right\}.$$

It can be endowed with a Hilbert space structure as described in Appendix A.

Observe that, as in the standard one-dimensional AK model, the condition (9) is needed to ensure the finiteness of the functional and of the value function. Since we chose the dynamic programming approach, we use the characterization of the value function in order to find the optimal dynamics and the optimal control of the problem. The former is described in the following theorem.

Theorem 3.3. *Under the hypotheses of Proposition 3.2, the optimal evolution of the capital distribution starting from k_0 is the solution of the following partial differential equation:*

$$(10) \qquad \begin{cases} \dfrac{\partial k(t,x)}{\partial t} = \Delta_x k(t,x) + Ak(t,x) - \left(\dfrac{\rho - A(1-\sigma)}{\sigma\, vol(M)} \right) \displaystyle\int_M k(t,x)\, \mathrm{d}x \\ k(0,x) = k_0(x). \end{cases}$$

Equation (10) is a parabolic equation on the geography M and it describes the optimal evolution of the system from time 0 to $+\infty$. The corresponding optimal spatio-temporal consumption can be expressed explicitly as shown in the next proposition.

Proposition 3.4. *Assume that the hypotheses of Proposition 3.2 are satisfied. Then the optimal consumption is constant in space and exponential in time: $c(t,x) = c_0 e^{\beta t}$ where $\beta := \left[\frac{A-\rho}{\sigma} \right]$ and $c_0 = \left(\frac{\rho - A(1-\sigma)}{\sigma\, vol(M)} \right) K(0)$, where $K(0) = \int_M k_0(x)\, \mathrm{d}x > 0$ is the initial level of aggregate capital.*

Furthermore the aggregate variables do not have a transitional dynamics along the optimal path and they are given by: $K(t) = K(0)e^{\beta t}$ and $C(t) = \left(\frac{\rho - A(1-\sigma)}{\sigma} \right) K(0)e^{\beta t}$.

The described optimal dynamics of the consumption $c(t,x)$ is elementary: the planner, on one hand, maximizes the utility if, at each time, all individuals in the economy can access the same level of consumption and the per-capita consumption grows exponentially in time. On the other hand, the capital distribution $k(t,x)$ has a much more elaborated behavior: it is described by the parabolic equation (10) that contains a second order term $\Delta_x k(t,x)$ and manifests a complex transitional dynamics. This dual behavior is common to other infinite dimensional AK models with CRRA utility[6]. Unlike in other cases, the system here can persist in a spatial-unequal capital distribution state. This is the argument of the next section.

We have already remarked that, at the aggregate level, the model is equivalent to the standard one-dimensional AK model. Indeed, the previous proposition establishes that the optimal aggregate capital and consumption growth rate are the same as in the one-dimensional case: $\frac{A-\rho}{\sigma}$. Moreover the same proportion of the aggregate production is consumed at each time. This global behavior is not reproduced at each spatial point where the dynamics of the capital is the non-trivial solution of (10).

3.2. **Geography and convergence.** In this subsection we study the role of the geographical structure in shaping the long run behavior of the system.

To state the results we need to recall some facts about the Laplacian operator Δ_x on the geography M. Some more details, useful for the proofs, are given in the Appendix A. A (non identically zero) regular function $\phi\colon M \to \mathbb{R}$ is called *eigenfunction* of Δ_x if there exists a real number (eigenvalue) λ such that[7] $\Delta_x \phi = -\lambda \phi$. It can be proved (see *e.g.* Chow et al.

---

[6]See for example the works by Boucekkine et al. (2005); Fabbri and Gozzi (2008); Boucekkine et al. (2010, 2013). In the first three of these works the capital accumulation takes the form of a delay differential equation but, even if the mathematical structure is completely different, the same dual behavior in the evolution of capital and consumption is reproduced.

[7]We define here an eigenvalue as a number λ such that $\Delta_x \phi = -\lambda \phi$. This is indeed the standard in the differential geometry literature where people often study the operator $(-\Delta_x)$ instead of Δ_x.

(2006) page 468) that the set of possible eigenvalues is discrete, that they form a sequence

(11) $$0 = \lambda_0 < \lambda_1 < \lambda_2 < ... < \lambda_k < ...$$

with $\lambda_k \xrightarrow{k \to \infty} +\infty$ and that the constant functions are the unique eigenfunctions associated to the eigenvalue $\lambda_0 = 0$.

We see now how this values, and in particular the first non-zero eigenvalue of the Laplacian, are important to determine the long run behavior of the spatial growth model. In order to look at the spatial distribution of the capital in the long run, we discount it by the growth rate of the aggregate variables thus obtaining the *detrended* spatial distribution of the capital at time t $k_D(t,x)$:

$$k_D(t,x) := e^{-\beta t} k(t,x).$$

If we divide k_D by $K(0)$, we obtain exactly the (density of the) portion of aggregate capital localized at the point x at time t. The following theorem makes explicit how the technological and preferences parameters and the geographic characteristics interact to determine the spatial convergence or capital cluster formation.

Theorem 3.5. *Assume that the hypotheses of Proposition 3.2 hold. Then[8] the detrended capital distribution $k_D(t,x)$ tends to a spatially equally distributed state if and only if*

(12) $$\rho < A(1 - \sigma) + \sigma \lambda_1,$$

otherwise we obtain long-run spatial capital agglomeration.

The previous theorem is the main result of this paper. In Theorem A.10 and in Remark A.11 in the Appendix A we present a more detailed result, showing the long run level of the detrended capital in case of convergence and characterizing, in case of agglomeration across locations, the limit of the detrended capital in terms of (non-constant) eigenfunctions of the Laplacian. However, in (12) we already have all the important ingredients to see how the various elements of the models interact to determine the convergence or the agglomeration in the long run behavior of the system.

A greater consumption impatience, measured by the discount rate ρ, tends to prevent the convergence outcome: increasing ρ pushes the consumption level c_0 characterized in Proposition 3.4. It is the same at each spatial point but it still is relatively higher in the depressed areas, more affected by the fall in investment levels.

The role of A depends on its impact on the consumption level of c_0 and it changes depending on the value of σ, the inverse of the elasticity of intertemporal substitution. This is not a surprising result: in the one-dimensional AK model already, the effect A on the level of the consumption in terms of physical capital $\left(\frac{\rho - A(1-\sigma)}{\sigma} \right)$ varies if σ is greater or lower than 1. Realistic values of the elasticity of intertemporal substitution are well below the unity, the corresponding values of σ are greater than 1. Consequently the term $A(1 - \sigma)$ is negative and an increase of the total factor productivity level diminishes the possibility of convergence. This means that the prevailing effect is the differential push: the impact of a gain in the TFP on the production at a spatial point x is proportional to the capital at that point and then stronger in richer areas.

λ_1 is the first non-zero eigenvalue of the Laplacian on the geography M and it summarizes its geometric properties. It can be shown that λ_1 is smaller if the geography M presents narrower bottlenecks, more precisely, under suitable curvature conditions (see Buser, 1982, Theorem 1.2) λ_1 goes to zero if the minimal area of a hypersurface disconnecting M does

---

[8]Apart from a subspace of initial data of co-dimension 1 in $L^2(M)$ *i.e.* apart from a "small" set of initial data, including in particular initial data in which the capital k_0 is constant in space.

so[9]. The smaller λ_1, the bigger the ecological barriers and geographical obstacles to capital diffusion. Consistently the smaller λ_1, the more difficult it is to verify (12).

In the case of the n-sphere $\lambda_1 = n$ (see *e.g.* Theorem 22.1 page 169 of the book by Shubin, 2013), when we consider the Salop circle (*i.e.* the 1-sphere) as a special example of geography, (12) becomes

$$\rho < A(1 - \sigma) + \sigma.$$

Furthermore, if we restrict our attention to the values of ρ, A and σ satisfying such an inequality we reproduce exactly the convergence result found by Boucekkine et al. (2013). If the geography is the (2-dimensional) ball surface, *i.e.* the 2-sphere, (12) reads as

$$\rho < A(1 - \sigma) + 2\sigma$$

and under a strictly larger set of parameters ρ, A and σ the optimal evolution of the detrended capital tends to a spatial uniform distribution in the long run.

Moreover, by changing the geography, the value of λ_1 can become as close to zero as we wish (see Randol, 1974). So, since the condition (9) always needs to be satisfied to ensure the finiteness of the functional, the set of possible parameters satisfying the long-run convergence condition (12) can be very thin.

Increasing the intertemporal elasticity of substitution $1/\sigma$, in case of a positive growth rate (*i.e.* $A > \rho$), increases also the part of the production used for investment for each point in space and time. The consequent effect depends on which of the two effects is stronger: the divergent effect, measured by A, due to different gains in production given different capital densities or the homogenizing effect of capital spread quantified by λ_1.

4. Conclusions

The specific contribution of the present work is to investigate the role of geography in the evolution of a spatial growth model. To this extent, we consider an AK spatial model with capital mobility with a generic geographic structure.

The main finding of this paper is that changing the geography changes the qualitative behavior of the system: by keeping the same parameters for preferences, discount rate, total factor productivity we observe convergence of the detrended capital across the locations or agglomeration depending on the geography structure. We have precisely characterized the analytical conditions that lead to different qualitative behaviors.

References

G. Aldashev, S. Aldashev, and T. Carletti. On convergence in the spatial ak growth models. *arXiv preprint arXiv:1401.4887*, 2014.

M. Beckmann and T. Puu. *Spatial economics: density, potential, and flow.* North-Holland, 1985.

A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and control of infinite dimensional systems.* Birkhäuser Boston Inc., Boston, MA, 2007.

R. Boucekkine, O. Licandro, L. A. Puch, and F. del Rio. Vintage capital and the dynamics of the AK model. *Journal of Economic Theory*, 120(1):39–72, 2005.

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

R. Boucekkine, G. Fabbri, and F. Gozzi. Maintenance and investment: Complements or substitutes? a reappraisal. *Journal of Economic Dynamics and Control*, 34(12):2420–2439, 2010.

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

P. Brito. The Dynamics of Growth and Distribution in a Spatially Heterogeneous World. Working Papers Department of Economics 2004/14, ISEG, University of Lisbon, 2004.

P. Brito. Global endogenous growth and distributional dynamics. Munich Personal RePEc Archive, Paper No. 41653, October 2012.

---

[9]The reversed fact is also true: if λ_1 goes to zero then the minimal area of a hypersurface disconnecting M does so see *e.g.* Cheeger (1970).

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

W. Brock and A. Xepapadeas. General pattern formation in recursive dynamical systems models in economics. Working Papers 2009.49, Fondazione Eni Enrico Mattei, Jun 2009.

W. Brock, A. Xepapadeas, and A. Yannacopoulos. Optimal Control in Space and Time and the Management of Environmental Resources. DEOS Working Papers 1402, Athens University of Economics and Business, Feb 2014.

P. Buser. A note on the isoperimetric constant. In *Annales scientifiques de l'École Normale Supérieure*, volume 15, pages 213–230. Société mathématique de France, 1982.

C. Camacho, B. Zou, and M. Briani. On the dynamics of capital accumulation across space. *European Journal of Operational Research*, 186(2):451–465, 2008.

J. Cheeger. A lower bound for the smallest eigenvalue of the laplacian. In R. C. Gunning, editor, *Problems in analysis*, pages 195–199. Princeton Univ. Press, 1970.

B. Chow, P. Lu, and L. Ni. *Hamilton's Ricci flow*, volume 77 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006.

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

J. M. Diamond. *Guns, Germs, and Steel: a short history of everybody for the last 13,000 years*. W. W. Norton, New York, 1997.

G. Fabbri and F. Gozzi. Solving optimal growth models with vintage capital: The dynamic programming approach. *Journal of Economic Theory*, 143(1):331–373, 2008.

A. Grigor'yan. *Heat kernel and analysis on manifolds*. American Mathematical Soc., 2012.

W. Isard, P. Liossatos, Y. Kanemoto, and P. C. Kaniss. *Spatial dynamics and optimal space-time development*. North-Holland, 1979.

P. Mossay. A theory of rational spatial agglomerations. *Regional Science and Urban Economics*, 43(2): 385–394, 2013.

D. Quah. Spatial agglomeration dynamics. *The American Economic Review*, 92(2):247–252, 2002.

B. Randol. Small eigenvalues of the laplace operator on compact riemann surfaces. *Bulletin of the American Mathematical Society*, 80(5):996–1000, 1974.

M. A. Shubin. *Pseudodifferential operators and spectral theory*. Springer, 2013.

A. Smith. *An inquiry into the nature and causes of the wealth of nations*. First edition, 1776.

Appendix A. Proofs

A.1. **Notations and preliminary results.** For the reader convenience, we first provide some definitions and preliminary results necessary to the proofs.

We start by recalling how to rewrite the optimal control problem stated in Section 2 as an optimal control problem in the space of square integrable functions from M to $\mathbb{R}$. The space $L^2(M)$, defined in (7), is a Hilbert space, we denote by $\langle f, g \rangle := \int_M f(x)g(x)\,\mathrm{d}x$ its scalar product. We define the operator G on $L^2(M)$ as follows[10]

$$\begin{cases} D(G) := H^2(M) \\ G(f) = \Delta_x f. \end{cases}$$

G is the (self-adjoint) generator of the heat semigroup on $L^2(M)$. It is a C_0 semigroup on $L^2(M)$ (see Section 4.3 of Grigor'yan, 2012 and Bensoussan et al., 2007 for the general theory of C^0-semigroups). The state equation (4) can be rewritten as an evolution equation in the Hilbert space $L^2(M)$ as follows:

$$(13) \qquad \begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - c(t) \\ k(0) = k_0 \end{cases}$$

where $k(t)$ and $c(t)$ are interpreted as the functions of variable x defined by $k(t)(x) \equiv k(t, x)$ and $c(t)(x) \equiv c(t, x)$. The mild solution of (13), see Defintion 3.1, page 129 of Bensoussan et al. (2007), is given by

$$(14) \qquad k_{k_0,c}(t) = e^{Gt}k_0 + \int_0^t e^{(t-s)G}\left(Ak_{k_0,c}(s) - c(s)\right)\mathrm{d}s$$

or, called $\tilde{G} := G + A$ (G plus A times the identity operator),

$$(15) \qquad k_{k_0,c}(t) = e^{\tilde{G}t}k_0 - \int_0^t e^{(t-s)\tilde{G}}c(s)\,\mathrm{d}s.$$

---

[10]The Sobolev space $H^2(M)$ is the completion, w.r.t. the norm $|f|_{H^2} := \left(\int_M \sum_{i=0}^2 \int_M \left|\nabla^i f(x)\right|^2 \mathrm{d}x\right)^{1/2}$ of the space of the C^∞ functions (see Grigor'yan, 2012, Section 4 for details). We write $D(G)$ to denote the *domain* of the operator G.

Observe that, chosen a control c, the aggregate capital is given by $K(t) = \langle k_{k_0,c}(t), \mathbb{1} \rangle$ where $\mathbb{1} \colon M \to \mathbb{R}$ is the function that is identically equal to 1.We use the notation $L^2_{loc}(0, +\infty; L^2(M))$ for the following functions space

$$L^2_{loc}(0, +\infty; L^2(M)) := \left\{ f \colon [0, +\infty) \times M \to \mathbb{R} \ : \ \int_0^T \int_M |f(t,x)|^2 \, \mathrm{d}x \, \mathrm{d}t < \infty \ \forall T > 0 \right\}.$$

The set of admissible controls of the optimal control problem (4)-(5) can be then written as

(16) $$\mathcal{U}_{k_0} := \left\{ c(\cdot, \cdot) \in L^2_{loc}(0, +\infty; L^2(M)) \ : \ c(\cdot, \cdot) \geq 0 \text{ and } K(t) > 0 \text{ for } t \geq 0 \right\}.$$

So the optimal control problem described in Section 2 is equivalent to the problem of maximizing the functional

$$J(k_0, c) := \int_0^{+\infty} e^{-\rho t} \langle \mathbb{1}, U(c(t)) \rangle \, \mathrm{d}t,$$

where $U(\eta)(x) = \frac{(\eta(x))^{1-\sigma}}{1-\sigma}$, among the controls of (16), subject to (13).

As already mentioned in Section 3, we call eigenfunction of Δ_x a (non identically zero) regular function $\phi \colon M \to \mathbb{R}$ such that $\Delta_x \phi = -\lambda \phi$ for a real number (eigenvalue) λ and it can be proved (see *e.g.* Chow et al. (2006) page 468) that the set of the possible eigenvalues is discrete and they form a sequence $0 = \lambda_0 < \lambda_1 < \lambda_2 < ... < \lambda_k < ....$ We call eigenspace associated with the eigenvalue λ_n the vector space of the eigenfunctions associated with the eigenvalue λ_n and we denote it by S_n. It can be proved that the dimension of S_n is finite (see *e.g.* Chow et al., 2006, page 469), we denote it by θ_n. It is possible to choose an orthonormal basis of $L^2(M)$ of (normalized) eigenfunctions of the Laplacian ϕ_n^j, for $n \in \mathbb{N}$ and $j \in \{1, .., \theta_n\}$, where, for any $n \geq 0$, $\phi_n^1, .., \phi_n^{\theta_n}$ are eigenfunctions associated to the eigenvalue λ_n. Any $f \in L^2(M)$ can be written as the $L^2(M)$-limit of the series $f = \sum_{n=0}^\infty \sum_{j=1}^{\theta_n} \langle f, \phi_n^j \rangle \phi_n^j$ and $|f|_{L^2(M)} = \sum_{n=0}^\infty \sum_{j=1}^{\theta_n} \langle f, \phi_n^j \rangle^2$.

It can also be shown that the dimension of S_0 is exactly 1 and it contains only constant functions *i.e.* the functions of the form $\alpha \mathbb{1}$ for some $\alpha \in \mathbb{R}$. In particular $e^{Gt} \mathbb{1} = \mathbb{1}$ and $e^{\tilde{G}t} \mathbb{1} = e^{-At} \mathbb{1}$. The unique normalized function (w.r.t. L^2-norm) of S_0 is $\phi_0 = \frac{1}{\sqrt{\mathrm{vol}(M)}} \mathbb{1}$.

Proof of Proposition 3.1. Given an initial datum k_0 and chosen an admissible control $c(\cdot, \cdot)$, using (15), we have

$$K(t) = \langle k_{k_0,c}(t), \mathbb{1} \rangle = \left\langle e^{\tilde{G}t} k_0, \mathbb{1} \right\rangle - \int_0^t \left\langle e^{(t-s)\tilde{G}} c(s), \mathbb{1} \right\rangle \, \mathrm{d}s.$$

Since G and then $e^{\tilde{G}t}$ are self-adjoint the expression above equals

$$\left\langle k_0, e^{\tilde{G}t} \mathbb{1} \right\rangle - \int_0^t \left\langle c(s), e^{(t-s)\tilde{G}} \mathbb{1} \right\rangle \, \mathrm{d}s$$

and, since $e^{\tilde{G}t} \mathbb{1} = e^{-tA} \mathbb{1}$ and the aggregate consumption is given by $C(t) = \langle c(t), \mathbb{1} \rangle$, we obtain

$$K(t) = e^{-tA} K(0) - \int_0^t e^{A(t-s)} C(s) \, \mathrm{d}s$$

so $K(t)$ is exactly the solution of (6). This proves the claim. $\qquad\square$

A.2. **Proofs of results of Section 3.1.** In this subsection we solve the optimal control problem using dynamic programming in infinite dimensions and we prove Proposition 3.2, Theorem 3.3 and Proposition 3.4.

First, we observe in the following proposition that the condition (9) demanded in Proposition 3.2 (and then in all the subsequent results) is sufficient to ensure the finiteness of the value function. In fact what we prove here is just that $V < +\infty$. The other bound will be a corollary of the following results since the utility along the optimal trajectory will be bigger than $-\infty$, and so will the supremum of the utility varying the control.

Proposition A.1. *If*

(17) $$\rho > A(1 - \sigma)$$

then all the trajectories give a bounded utility from above, more precisely, for all positive $k_0 \in L^2(M)$, $V(k_0) := \sup_{c \in \mathcal{U}_{k_0}} J(c(\cdot, \cdot)) < +\infty$.

Proof of Proposition A.1. Since the claim is obvious if $\sigma > 1$ we prove it only for $\sigma \in (0, 1)$. Observe first that, from (6), since $C(\cdot) \geq 0$, we have

(18) $$K(t) \leq K(0) e^{At}.$$

Using Jensen inequality on the space $[0, +\infty) \times M$ with the measure $\frac{1}{\rho \operatorname{vol}(M)} e^{-\rho t} \, \mathrm{d}t \otimes \mathrm{d}x$, (6) and then integration by part, we have

$$(19) \quad \frac{1}{\rho \operatorname{vol}(M)} \int_0^{+\infty} \int_M e^{-\rho t} c(t,x)^{1-\sigma} \, \mathrm{d}x \, \mathrm{d}t \leq \left(\frac{1}{\rho \operatorname{vol}(M)} \int_0^{+\infty} e^{-\frac{\rho}{1-\sigma} t} C(t) \, \mathrm{d}t \right)^{1-\sigma}$$

$$= \left(\frac{1}{\rho \operatorname{vol}(M)} \int_0^{+\infty} e^{-\frac{\rho}{1-\sigma} t} (AK(t) - \dot{K}(t)) \, \mathrm{d}t \right)^{1-\sigma}$$

$$\leq \left(\frac{1}{\rho \operatorname{vol}(M)} \int_0^{+\infty} e^{-\frac{\rho}{1-\sigma} t} AK(t) \, \mathrm{d}t - \left| K(t) e^{-\frac{\rho}{1-\sigma} t} \right|_{t=0}^{t=+\infty} - \int_0^{+\infty} \frac{\rho}{1-\sigma} e^{-\frac{\rho}{1-\sigma} t} K(t) \, \mathrm{d}t \right)^{1-\sigma}.$$

Applying (17) and (18) the last expression can be easily seen to be lower than a constant independent of the control c. So we have the claim. $\qquad \square$

We now study the optimal control problem using the dynamic programming approach in the space $L^2(M)$. So first we write, in (20), the Hamilton-Jacobi-Bellman (HJB) equation of the problem then we find an explicit solution (Proposition A.3) and we use such a solution to derive, in (22), a feedback. Eventually we demonstrate (Proposition A.8) that such a feedback is optimal and that the solution of the HJB equation we found is in fact the value function of the problem. These two last results prove Proposition 3.2 and Theorem 3.3. The proof of Proposition 3.4 follows as a corollary.

The HJB equation associated to our problem is[11]:

$$(20) \qquad \rho v(k) = \langle k, G\mathcal{D}v(k) \rangle + A \langle k, \mathcal{D}v(k) \rangle + \sup_{c \geq 0} \left\{ -\langle c, \mathcal{D}v(k) \rangle + \langle \mathbb{1}, U(c) \rangle \right\}$$

where $\mathcal{D}v$ represents the Fréchet differential of the function $v \colon L^2(M) \to \mathbb{R}$.

Definition A.2. *Let $O \subseteq L^2(M)$ an open set. $v \colon O \to \mathbb{R}$ is a solution of (20) on O if $v \in C^1(O)$, $\mathcal{D}v \in C(O, D(G))$ and v solves pointwise (20) on O.*

Proposition A.3. *The function*

$$(21) \qquad \qquad v(k) = \alpha \langle k, \mathbb{1} \rangle^{1-\sigma},$$

with $\alpha = \frac{1}{1-\sigma} \left(\frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right)^{-\sigma}$, is a solution of (20) on the halfspace $\Omega := \{ f \in L^2(M) \ : \ \langle \mathbb{1}, f \rangle > 0 \}$.

Proof. We verify the statement directly. We observe that $\mathcal{D}v(k) = \alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \mathbb{1}$ so the candidate-solution is a solution of (20) if and only if

$$\rho \alpha \langle k, \mathbb{1} \rangle^{1-\sigma} = \alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle k, G\mathbb{1} \rangle + A\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle k, \mathbb{1} \rangle + \sup_{c \geq 0} \left\{ -\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle c, \mathbb{1} \rangle + \langle \mathbb{1}, U(c) \rangle \right\}.$$

Observing that $G\mathbb{1} = 0$ and that the supremum is attained when $c = (\alpha(1-\sigma))^{-1/\sigma} \langle k, \mathbb{1} \rangle \mathbb{1}$ the expression above becomes:

$$\rho \alpha \langle k, \mathbb{1} \rangle^{1-\sigma} = A\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{1-\sigma} - \operatorname{vol}(M)\alpha(1-\sigma) (\alpha(1-\sigma))^{-1/\sigma} \langle k, \mathbb{1} \rangle^{1-\sigma} + \operatorname{vol}(M) \frac{\left[(\alpha(1-\sigma))^{-1/\sigma} \langle k, \mathbb{1} \rangle \right]^{1-\sigma}}{1-\sigma}$$

simplifying the non-zero factor $\alpha \langle k, \mathbb{1} \rangle^{1-\sigma}$ the previous expression is equivalent to

$$\rho = A(1-\sigma) - \operatorname{vol}(M)(1-\sigma) (\alpha(1-\sigma))^{-1/\sigma} + \operatorname{vol}(M) (\alpha(1-\sigma))^{-1/\sigma}.$$

Using the explicit expression of α given in the statement, we can easily see that the previous equation is verified. This proves the claim. $\qquad \square$

Definition A.4. *Given O an open subset of $L^2(M)$, a function $\Psi \colon O \to L^2(M)$ is said to be a feedback in O if, for any $k_0 \in O$, the equation*

$$\begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - \Psi(k(t)) \\ k(0) = k_0 \end{cases}$$

has a unique solution $k_{\Psi,k_0}(\cdot)$ and $k_{\Psi,k_0}(t) \in O$ for all $t \geq 0$.

Definition A.5. *Given O an open subset of $L^2(M)$, a function $\Psi \colon O \to L^2(M)$ is said to be an admissible feedback in O if it is a feedback in O and, for any $k_0 \in O$, $\Psi(k_{\Psi,k_0}(\cdot)) \in \mathcal{U}_{k_0}$.*

Definition A.6. *Given O an open subset of $L^2(M)$ a function $\Psi \colon O \to L^2(M)$ is said to be an optimal feedback in O if it is an admissible feedback in O and, for any $k_0 \in O$, $\Psi(k_{\Psi,k_0}(\cdot))$ is an optimal control.*

---

[11]Observe that we read the term $\langle Gk, \mathcal{D}v(k) \rangle$ appearing in the usual expression of the HJB equation as $\langle k, G^* \mathcal{D}v(k) \rangle$. Since G is self-adjoint, it equals $\langle k, G\mathcal{D}v(k) \rangle$ and we find the form given in (20).

The feedback associated to the solution of the HJB equation found in Proposition A.3 is given by

$$(22) \quad \begin{cases} \Phi \colon \Omega \to L^2(M) \\ \Phi \colon k \mapsto \sup_{c \in L^2(M), \, c \geq 0} \left\{ -\alpha(1-\sigma) \langle k, \mathbb{1} \rangle^{-\sigma} \langle c, \mathbb{1} \rangle + \langle \mathbb{1}, U(c) \rangle \right\} = \left(\frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right) \langle k, \mathbb{1} \rangle \mathbb{1}. \end{cases}$$

Proposition A.7. Φ, *defined in (22), is a in admissible feedback in* Ω *(defined in Proposition A.3). More precisely, for any* $k_0 \in \Omega$, *if we define* $K_{\Phi, k_0}(t) := \langle \mathbb{1}, k_{\Phi, k_0}(t) \rangle$ *we have*

$$(23) \qquad\qquad\qquad K_{\Phi, k_0}(t) = K(0) e^{\beta t}$$

where $K(0) = \langle \mathbb{1}, k_0 \rangle$ *and*

$$\beta := \left[\frac{A - \rho}{\sigma} \right].$$

Proof. We choose $k_0 \in \Omega$ and to lighten the notation, we write k_Φ instead of k_{Φ, k_0} and K_Φ instead of K_{Φ, k_0}. Using (14) and replacing $c(t)$ by the feedback $\Phi(k_\Phi(t))$ we have

$$k_\Phi(t) = e^{Gt} k_0 + \int_0^t e^{(t-s)G} \left(A k_\Phi(s) - c(s) \right) \, \mathrm{d}s = e^{Gt} k_0 + \int_0^t e^{(t-s)G} \left(A k_\Phi(s) - \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \langle k_\Phi(s), \mathbb{1} \rangle \mathbb{1} \right) \, \mathrm{d}s,$$

so

$$K_\Phi(t) = \left\langle k_0, e^{Gt} \mathbb{1} \right\rangle + \int_0^t \left\langle \left(A k_\Phi(s) - \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \left\langle k_\Phi(s), e^{(t-s)G} \mathbb{1} \right\rangle \mathbb{1} \right), e^{(t-s)G} \mathbb{1} \right\rangle \, \mathrm{d}s.$$

Given that $e^{Gt} \mathbb{1} = \mathbb{1}$, the expression above becomes:

$$(24) \quad K_\Phi(t) = \langle k_0, \mathbb{1} \rangle + \int_0^t \left\langle \left(A k_\Phi(s) - \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \langle k_\Phi(s), \mathbb{1} \rangle \mathbb{1} \right), \mathbb{1} \right\rangle \, \mathrm{d}s$$

$$= K_\Phi(0) + \int_0^t K_\Phi(s) \left[A - \operatorname{vol}(M) \frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right] \, \mathrm{d}s = K_\Phi(0) + \int_0^t K_\Phi(s) \left[\frac{A - \rho}{\sigma} \right] \, \mathrm{d}s$$

and the claim is proved. $\qquad\square$

Proposition A.8. *Assume that (17) is satisfied. Then* Φ *defined in (22) is an optimal feedback in* Ω *and the value function of the problem computed at* k_0 *is*

$$V(k_0) = \frac{1}{1-\sigma} \left(\frac{\rho - A(1-\sigma)}{\sigma \operatorname{vol}(M)} \right)^{-\sigma} \langle k_0, \mathbb{1} \rangle^{1-\sigma}$$

Proof. Call $c^*(t) := \Phi(k_\Phi(t))$. To prove that $c^*(\cdot)$ is an optimal control, we have to prove that for any other admissible control $\tilde{c}(\cdot)$ ($\tilde{k}(\cdot)$ being the related trajectory), $J(k_0, c^*) \geq J(k_0, \tilde{c})$. Denote by $w(t, k) \colon \mathbb{R} \times L^2(M) \to \mathbb{R}$ the function $w(t, k) := e^{-\rho t} v(k)$. If we fix $T > 0$, we have:

$$v(k_0) - w(T, \tilde{k}(T)) = w(t, \tilde{k}(0)) - w(T, \tilde{k}(T)) = -\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} w(t, \tilde{k}(t)) \, \mathrm{d}t$$

$$= \int_0^T e^{-\rho t} \left[\rho v(\tilde{k}(t)) - \left\langle G\tilde{k}(t) + A\tilde{k}(t) - \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle \right] \, \mathrm{d}t.$$

The last expression makes sense thanks to the regularizing properties of the heat semigroup: for any $t > 0$, $\tilde{k}(t) \in D(G)$. Using (18), the explicit form of v given in (21) and the hypothesis (12) we can easily see that $w(T, \tilde{k}(T)) \to 0$ when $T \to \infty$ so we can pass to the limit in the previous equation and find that

$$v(k_0) = \int_0^{+\infty} e^{-\rho t} \left[\rho v(\tilde{k}(t)) - \left\langle A\tilde{k}(t) - \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle - \left\langle \tilde{k}(t), G\mathcal{D}v(\tilde{k}(t)) \right\rangle \right] \, \mathrm{d}t$$

and then

$$v(k_0) - J(k_0, \tilde{c}) = \int_0^{+\infty} e^{-\rho t} \left[\left(\rho v(\tilde{k}(t)) - \left\langle A\tilde{k}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle - \left\langle \tilde{k}(t), G\mathcal{D}v(\tilde{k}(t)) \right\rangle \right) \right.$$

$$\left. + \left(\left\langle \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle - \langle \mathbb{1}, U(\tilde{c}(t)) \rangle \right) \right] \, \mathrm{d}t$$

$$= \int_0^{+\infty} e^{-\rho t} \left[\left(\sup_{c \in L^2(M; \mathbb{R}^+)} \left\{ -\left\langle c, \mathcal{D}v(\tilde{k}(t)) \right\rangle + \langle \mathbb{1}, U(c) \rangle \right\} \right) - \left(-\left\langle \tilde{c}(t), \mathcal{D}v(\tilde{k}(t)) \right\rangle + \langle \mathbb{1}, U(\tilde{c}(t)) \rangle \right) \right] \, \mathrm{d}t \geq 0$$

where we used in last step the fact that v is a solution of (20). The last expression gives $v(k_0) - J(k_0, \tilde{c}) \geq 0$ and from the same expression we can also determine that $v(k_0) - J(k_0, c^*) = 0$ (indeed $c^*(\cdot)$ is defined using the feedback defined in (22)). So, for all admissible $\tilde{c}$, $v(k_0) - J(k_0, \tilde{c}) \geq 0 = v(k_0) - J(k_0, c^*)$ and then $J(k_0, \tilde{c}) \leq J(k_0, c^*)$ and then c^* is optimal. In particular, since $v(k_0) = J(k_0, c^*) = 0$ and c^* is an optimal control, $v(k_0)$ is the value function at k_0. This concludes the proof. $\qquad\square$

Proof of Proposition 3.2. It is part of the statement of Proposition A.8, once we read $\langle k_0, \mathbb{1} \rangle^{1-\sigma}$ as $\left(\int_M k_0(x) \, \mathrm{d}x \right)^{1-\sigma}$. $\qquad\square$

Proof of Theorem 3.3. It is a corollary of Proposition A.8. Indeed we have proven that Φ is an optimal feedback so the capital along the optimal trajectory is the solution of the following equation

$$\begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - \Phi(k(t)) \\ k(0) = k_0 \end{cases}$$

that, using (22), is given by

$$\begin{cases} \dot{k}(t) = Gk(t) + Ak(t) - \left(\frac{\rho - A(1-\sigma)}{\sigma \, \mathrm{vol}(M)} \right) \langle k(t), \mathbb{1} \rangle \mathbb{1} \\ k(0) = k_0, \end{cases}$$

that is exactly (10). $\qquad\square$

Proof of Proposition 3.4. From (22), we have $c(t) = \left(\frac{\rho - A(1-\sigma)}{\sigma \, \mathrm{vol}(M)} \right) \langle k(t), \mathbb{1} \rangle \mathbb{1}$ that is

$$(25) \qquad c^*(t, x) = \frac{\rho - A(1-\sigma)}{\mathrm{vol}(M)\sigma} K^*(t),$$

so the aggregate consumption on optimal trajectory is

$$(26) \qquad C^*(t) = \langle c^*(t), \mathbb{1} \rangle = \frac{\rho - A(1-\sigma)}{\sigma} K^*(t).$$

Using such expression in (6) we have

$$\dot{K}^*(t) = AK^*(t) - \frac{\rho - A(1-\sigma)}{\sigma} K^*(t)$$

and then $K^*(t) = K(0)e^{A - \frac{\rho - A(1-\sigma)}{\sigma}} = K(0)e^{\beta t}$ so using again respectively (26) and (25) we have $C^*(t) = \frac{\rho - A(1-\sigma)}{\sigma} K(0)e^{\beta t}$ and $c^*(t, x) = \frac{\rho - A(1-\sigma)}{\mathrm{vol}(M)\sigma} K(0)e^{\beta t}$. This concludes the proof. $\qquad\square$

A.3. **Proof of Theorem 3.5.**

Notation A.9. *The word "convergence" can be confusing. In this paper we always use the word as in the economic growth literature: we have convergence if the (detrended) spatial distribution of the capital tends, in the long run, to equalize across spatial locations. In mathematical terms however, we could say that $k_D(t)(x)$ "converges" to a certain $l(x)$ (when e.g. $t \to \infty$) even if $l(x)$ is non-constant. To avoid this possible confusion we use the expression "tends to" instead of "converges to" for this second meaning.*

Theorem 3.5 is a direct consequence of the following, more detailed, results.

Theorem A.10. *Assume that (17) is satisfied and consider an initial datum $k_0 \in \Omega$ (defined in Proposition A.3). Then:*

1. *If $\rho < A(1-\sigma) + \sigma\lambda_1$ then*

$$\lim_{t \to \infty} e^{-\beta t} k(t) = \frac{\int_M k_0(x) \, \mathrm{d}x}{\mathrm{vol}(M)}, \qquad in \ L^2(M).$$

2. *If $\rho = A(1-\sigma) + \sigma\lambda_1$ then (a part for a set of initial data k_0 spanning a subspace of $L^2(M)$ of co-dimension 1 i.e. a part for a "small" set of initial data)*

$$\lim_{t \to \infty} e^{-\beta t} k(t) = \frac{\int_M k_0(x) \, \mathrm{d}x}{\mathrm{vol}(M)} + \psi_1, \qquad in \ L^2(M),$$

where ψ_1 is an eigenfunction related to the first non-zero eigenvalue λ_1 of the Laplacian.

3. *If $\rho > A(1-\sigma) + \sigma\lambda_1$ then (a part for a set of initial data k_0 spanning a subspace of $L^2(M)$ of co-dimension 1), we have*

$$\lim_{t \to \infty} e^{-(A-\lambda_1)t} k(t) = \psi_1, \qquad in \ L^2(M),$$

where ψ_1 is an eigenfunction related to the first non-zero eigenvalue λ_1 of the Laplacian.

Remark A.11. *In Case 1 particularly, the spatial capital distribution, detrended by the factor $e^{-\beta t}$, tends to a spatially constant distribution. Though in Cases 2 and 3 the spatial capital distribution, detrended respectively by $e^{-\beta t}$ and $e^{-(A-\lambda_1)t}$, tends to a spatially non-constant distribution. Indeed, as already recalled, any eigenfunction ψ_n related to some non-zero eigenvalue λ_n are non-constant. In particular, since $\lambda_1 > 0$, the functions ψ_1 appearing in Cases 2 and 3 of Theorem A.10 are non-constant as function of the space variable. In other words, in Cases 2 and 3 we have capital agglomeration.*

Proof of Theorem A.10. Using the feedback relation (22) into the mild form (15) we have, along the optimal trajectory,

$$k(t) = e^{\tilde{G}t}k_0 - \frac{\rho - A(1-\sigma)}{\sigma}\int_0^t e^{(t-s)\tilde{G}}\langle k(s), \mathbb{1}\rangle \mathbb{1}\, ds.$$

For a given eigenfunction ϕ_n^j associated to some eigenvalue λ_n of the Laplacian, we get

$$(27)\quad \left\langle k(t), \phi_n^j\right\rangle = \left\langle e^{\tilde{G}t}k_0, \phi_n^j\right\rangle - \frac{\rho - A(1-\sigma)}{\sigma\,\mathrm{vol}(M)}\int_0^t \left\langle\langle k(s), \mathbb{1}\rangle\, e^{(t-s)\tilde{G}}\mathbb{1}, \phi_n^j\right\rangle ds$$

$$= \left\langle k_0, e^{\tilde{G}t}\phi_n^j\right\rangle - \frac{\rho - A(1-\sigma)}{\sigma\,\mathrm{vol}(M)}\int_0^t \langle k(s), \mathbb{1}\rangle \left\langle \mathbb{1}, e^{(t-s)\tilde{G}}\phi_n^j\right\rangle ds.$$

Using that $e^{\tilde{G}t}\phi_n^j = e^{(-\lambda_n + A)t}\phi_n^j$ and that, for $n \neq 0$, $\langle\phi_n^j, \mathbb{1}\rangle = 0$, we can see that

$$(28)\qquad \left\langle k(t), \phi_n^j\right\rangle = \left\langle k_0, e^{(A-\lambda_n)t}\phi_n^j\right\rangle = e^{(A-\lambda_n)t}\left\langle k_0, \phi_n^j\right\rangle, \qquad n \neq 0,$$

while, if $n = 0$, we get from Proposition 3.1 that

$$(29)\qquad \langle k(t), \phi_0\rangle = \left\langle k(t), \frac{1}{\sqrt{\mathrm{vol}(M)}}\mathbb{1}\right\rangle = \frac{K(t)}{\sqrt{\mathrm{vol}(M)}} = e^{\beta t}\frac{K(0)}{\sqrt{\mathrm{vol}(M)}}.$$

    Case 1.:

We prove here the first statement, so we assume that $\rho < A(1-\sigma) + \sigma\lambda_1$. Thank to this hypothesis and (17) we can fix a certain $\varepsilon \in \left(0, \frac{A(1-\sigma)+\lambda_1\sigma-\rho}{\sigma}\right)$. We want to prove that $e^{-\beta t}k(t)$ tends in the L^2-norm to $\frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}$. Using (28) and (29) we have

$$\left|e^{-\beta t}k(t) - \frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}\right|_{L^2(M)}^2 = \sum_{n\geq 0}\sum_{j=1}^{\theta_n}\left\langle\left(e^{-\beta t}k(t) - \frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}\right), \phi_n^j\right\rangle^2 = \sum_{n\geq 1}\sum_{j=1}^{\theta_n}e^{2(A-\lambda_n-\beta)t}\left\langle k_0, \phi_n^j\right\rangle^2$$

$$= e^{-2\varepsilon t}\sum_{n\geq 1}\sum_{j=1}^{\theta_n}e^{2(A-\lambda_n-\beta+\varepsilon)t}\left\langle k_0, \phi_n^j\right\rangle^2 \leq e^{-2\varepsilon t}\sum_{n\geq 1}\sum_{j=1}^{\theta_n}\left\langle k_0, \phi_n^j\right\rangle^2 \leq e^{-2\varepsilon t}|k_0|_{L^2(M)}^2 \xrightarrow{t\to\infty} 0,$$

where we used that, for all $n \neq 0$, $(A - \lambda_n - \beta + \varepsilon) = \left(\frac{\sigma(A-\lambda_n)-(A-\rho)+\varepsilon\sigma}{\sigma}\right) < \left(\frac{\sigma(A-\lambda_1)-(A-\rho)+\varepsilon\sigma}{\sigma}\right) < 0.$

    Case 2.:

We analyze now the case in which $\rho = A(1-\sigma) + \sigma\lambda_1$ so that

$$(30)\qquad\qquad\qquad\qquad A - \lambda_1 - \beta = 0.$$

We introduce $\psi_1 := \sum_{i=1}^{\theta_1}\left\langle k_0, \phi_1^j\right\rangle\phi_1^j$. ψ_1 is non-zero for all $k_0 \in \Omega$ except those contained in a subspace of co-dimension $\theta_1 \geq 1$ of $L^2(M)$. To prove that $e^{-\beta t}k(t)$ tends to $\frac{K(0)\mathbb{1}}{\mathrm{vol}(M)} + \psi_1$ in the L^2-norm we first observe that

$$(31)\qquad\qquad\qquad\qquad \left\langle\psi_1, \phi_n^j\right\rangle = 0, \qquad \text{for any } n \neq 1$$

and

$$(32)\qquad\qquad\qquad\qquad \left\langle\psi, \phi_1^j\right\rangle = \left\langle k_0, \phi_1^j\right\rangle, \qquad \text{for any } j \in \{1,..,\theta_1\}.$$

Using these facts together with (28), (29) and (30) we have

$$\left|e^{-\beta t}k(t) - \psi_1 - \frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}\right|_{L^2(M)}^2 = \sum_{n\geq 0}\sum_{j=1}^{\theta_n}\left\langle\left(e^{-\beta t}k(t) - \psi_1 - \frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}\right), \phi_n^j\right\rangle^2$$

$$= \left(\frac{K(0)}{\sqrt{\mathrm{vol}(M)}} - \langle\psi_1, \phi_0\rangle - \left\langle\frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}, \phi_0\right\rangle\right)^2 + \left(\sum_{j=1}^{\theta_1}e^{(A-\lambda_1-\beta)t}\left\langle k_0, \phi_1^j\right\rangle - \left\langle\psi_1, \phi_1^j\right\rangle - \left\langle\frac{K(0)\mathbb{1}}{\mathrm{vol}(M)}, \phi_1^j\right\rangle\right)^2$$

$$+ \sum_{n\geq 2}\sum_{j=1}^{\theta_n}e^{2(A-\lambda_n-\beta)t}\left\langle k_0, \phi_n^j\right\rangle^2 = 0 + \left(\sum_{j=1}^{\theta_1}\left\langle k_0, \phi_1^j\right\rangle - \left\langle k_0, \phi_1^j\right\rangle - 0\right)^2 + \sum_{n\geq 2}\sum_{j=1}^{\theta_n}e^{2(A-\lambda_n-\beta)t}\left\langle k_0, \phi_n^j\right\rangle^2$$

$$= 0 + 0 + \sum_{n\geq 2}\sum_{j=1}^{\theta_n}e^{2(A-\lambda_n-\beta)t}\left\langle k_0, \phi_n^j\right\rangle^2.$$

We fix now $\epsilon \in (0, \lambda_2 - \lambda_1)$. The previous expression equals

$$e^{-2\epsilon t} \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-\beta+\epsilon)t} \left\langle k_0, \phi_n^j \right\rangle^2 \leq e^{-2\epsilon t} \sum_{n \geq 2} \sum_{j=1}^{\theta_n} \left\langle k_0, \phi_n^j \right\rangle^2 \leq e^{-2\epsilon t} \left| k_0 \right|_{L^2(M)}^2 \xrightarrow{t \to \infty} 0,$$

where in the first inequality we used that for all $n \geq 2$, thanks to (11) and (30), $(A - \lambda_n - \beta + \epsilon) \leq (A - \lambda_2 - \beta + \epsilon) < (A - \lambda_2 - \beta + (\lambda_2 - \lambda_1)) = 0$. The last limit holds because $\epsilon > 0$.

Case 3.:

The condition $\rho > A(1 - \sigma) + \sigma\lambda_1$, is equivalent to $(A - \lambda_1 + \beta) > 0$. ψ_1 appearing in the text of the proposition is, as in case 2, given by $\sum_{i=1}^{\theta_1} \left\langle k_0, \phi_1^j \right\rangle \phi_1^j$. Thanks to (29), (31) and then (28) we have

$$\left| e^{-(A-\lambda_1)t} k(t) - \psi_1 \right|_{L^2(M)}^2 = \sum_{n \geq 0} \sum_{j=1}^{\theta_n} \left\langle \left(e^{-(A-\lambda_1)t} k(t) - \psi_1 \right), \phi_n^j \right\rangle^2$$

$$= \left(e^{-(A-\lambda_1+\beta)t} \frac{K(0)}{\sqrt{\text{vol}(M)}} + 0 \right)^2 + \left(\sum_{j=1}^{\theta_1} \left\langle k_0, \phi_1^j \right\rangle - \left\langle \psi_1, \phi_1^j \right\rangle \right)^2 + \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(A-\lambda_n-(A-\lambda_1))t} \left\langle k_0, \phi_n^j \right\rangle^2$$

that, thanks to (32) equals

$$(33) \quad e^{-2(A-\lambda_1+\beta)t} \left(\frac{K(0)}{\sqrt{\text{vol}(M)}} \right)^2 + 0 + \sum_{n \geq 2} \sum_{j=1}^{\theta_n} e^{2(\lambda_1-\lambda_n)t} \left\langle k_0, \phi_n^j \right\rangle^2$$

$$\leq e^{-2(A-\lambda_1+\beta)t} \left(\frac{K(0)}{\sqrt{\text{vol}(M)}} \right)^2 + e^{2(\lambda_1-\lambda_2)t} \sum_{n \geq 2} \sum_{j=1}^{\theta_n} \left\langle k_0, \phi_n^j \right\rangle^2$$

$$\leq e^{-2(A-\lambda_1+\beta)t} \left(\frac{K(0)}{\sqrt{\text{vol}(M)}} \right)^2 + e^{2(\lambda_1-\lambda_2)t} \left| k_0 \right|_{L^2(M)}^2 \xrightarrow{t \to \infty} 0$$

where we used (11) in first inequality and we concluded using that $(A - \lambda_1 + \beta) > 0$ and that $\lambda_1 < \lambda_2$. This concludes the proof. $\square$

Remark A.12. *In the proof of Theorem A.10 we have shown that the limits hold in the $L^2(H)$ sense but with a similar argument we could show that in fact the limits are uniform.*

E-mail address: `giorgio.fabbri@univ-evry.fr`

EPEE, Université d'Evry-Val-d'Essonne (TEPP, FR-CNRS 3126), Département d'Economie, 4 bd. François Mitterrand, 91025 Evry cedex, France.