

The diffusion of economic activity across space: a new approach

Carmen Camacho

Paris School of Economics and CNRS (France)

Agustín Pérez-Barahona

*THEMA, University of Cergy-Pontoise and Ecole Polytechnique (France)**

February 20, 2017

Abstract

Dynamic spatial theory has been a fruitful approach in understanding economic phenomena involving time and space. However, this new field has opened a set of questions still unresolved in the literature. For instance, the identification of the temporal and spatial social optimal allocation of economic activity has not been ensured yet in economic growth. By means of a monotone method we study in this paper the optimal solution of spatial Ramsey-type models. The iterative nature of this approach also allows us to present a new algorithm to simulate the optimal trajectories of the economy. We provide two economic illustrations of our method. Firstly, in order to investigate the importance of capital mobility in economic growth, we consider the spatial Ramsey model. We point out the spatial dynamic implications in social welfare and income inequality. Secondly, under fairly general assumptions, we analytically prove the existence of a unique social optimum in this framework. We also apply the outcome to the spatial growth model and to a set-up for optimal land-use planning, concluding that these problems are well-posed.

*Corresponding author at: THEMA, U. of Cergy-Pontoise, 33 Boulevard du Port, 95011 Cergy-Pontoise cedex France; e-mail: agustin.perez-barahona@u-cergy.fr.

Keywords: Spatial dynamics, Ramsey model, Partial differential equations, Optimal control.

Journal of Economic Literature: C62, O4, R11.

1 Introduction

A recent strand of the literature on spatial economics has introduced diffusion mechanisms in order to understand the allocation of economic activity across time and space. This approach turned out to be useful for explaining a new set of spatial-dynamic phenomena, such as the mobility of production factors (Boucekkine *et al.*, 2013a; and Fabbri, 2016), the dissemination of technology (Desmet and Rossi-Hansberg, 2009, 2014 and 2015), the spatial allocation of natural resources (Brock and Xepapadeas, 2008), or the spread of pollution due to the economic activity (Camacho and Perez-Barahona, 2015). However this framework has come across serious methodological issues, without conclusive answers from either the economic or the mathematical literature. The aim of our paper is to contribute in this regard, providing a general analytical tool to study the solution of spatial Ramsey-type models.

Spatial Ramsey-type models are based on the spatial generalization of the Ramsey-Cass-Koopmans model, first presented in Brito (2004) and Boucekkine *et al.* (2009). This framework considers a forward-looking decision maker that chooses the optimal levels of control variables for a given spatial region and time span. In this problem the social planner also takes into account how the state variables geographically spread. The local level of consumption is the usual example of control, with capital or pollution as state variables. This is indeed an optimal control problem, where the evolution law of the state variables is represented by means of partial differential equations (hereafter PDE). In these equations the partial derivatives, with respect to space and to time, describe the spatio-temporal motion of the corresponding state variable.

The benchmark of this type of set-up is the spatial growth model, which adds spatial structure to the standard Ramsey model. In this framework the policy maker takes into account the welfare or utility of all the individuals distributed in space, considered over a certain period of time. Moreover capital is assumed to be mobile across locations, where the corresponding spatio-temporal law of motion is a non-linear parabolic PDE. Considering this approach, Boucekkine *et al.* (2013a) and Fabbri (2016) point out the role played by space in economic growth, which significantly affects the dynamical properties of the economy. Boucekkine *et al.* (2013a) specifically show that, despite constant returns to capital, transitional dynamics can arise in the AK growth model when we incorporate the spatial structure. In environmental economics, Camacho and Perez-Barahona (2015) provide another example of spatial Ramsey-type models. This paper studies the optimal use of land, encompassing land use activities (*e.g.*, produc-

tion, housing or pollution abatement) and environmental degradation. They focus on the spatial externalities of land use management as drivers of spatial patterns: land is immobile by nature, but local actions affect the whole space since pollution flows across locations resulting in environmental damages. Their stylized model manages to reproduce a great variety of spatial patterns related to the interaction between land use activities and the environment. In particular, in line with Costello and Polasky (2008), they also show that the spatial connectivity (due to the dispersal process), together with the particular characteristics of each location, can economically justify the creation of (temporal or permanent) biological reserves.¹

From a general perspective, spatial Ramsey-type models have been useful for describing the dynamic interaction of economic agents that are spatially distributed. Nevertheless, this framework has opened serious methodological questions. Within the context of social interactions, Mossay and Picard (2015) and Blanchet *et al.* (2016) notice important difficulties in ensuring the existence and uniqueness of equilibrium when we incorporate space into the problem. Considering the spatial Ramsey-type framework, papers such as Boucekkine *et al.* (2009), Boucekkine *et al.* (2013a,b) and Camacho and Perez-Barahona (2015) identified as well this issue when including both time and space dimensions: although a set of necessary conditions can be provided for a general problem, neither the existence nor the uniqueness of solution has been proven so far.²

In this paper we generalize the spatial Ramsey set-up. We then introduce a monotone method, based on Pao (1992), in order to analyze the solution of this type of models. Moreover, the iterative nature of this approach also allows us to present a new algorithm to find numerically the optimal trajectories of the economy. We underline the value of our method with two economic illustrations. One is numerical, the other is analytical.

For the numerical illustration we investigate the importance of capital mobility in economic growth. In order to address this question we apply the spatial growth model, where capital can flow across locations and agents are continuously distributed in the

¹Biological objectives, such as biodiversity conservation, typically give ground for expanding natural reserves. However, without explicitly including the spatial dimension, this is not the case in standard economics.

²One can then say that the problem is ill-posed in the sense of Hadamard (1923). Not being able to ensure the existence of solution is obviously problematic. The concern about non-uniqueness is not less important because it frequently implies indeterminacy of equilibrium in dynamic contexts. In this regard, many economists consider that models with indeterminacy are undesirable (among others, Kehoe *et al.*, 1990; and Acemoglu, 2009).

space.³ Barro *et al.* (1995) present an open-economy neoclassical growth model, without explicitly including spatial structure. Under perfect capital mobility, which is defined as the absence of institutional barriers to spatial capital flows, they observe that transitional dynamics occurs if capital is irreversible together with the presence of adjustment costs (see also Barro and Sala-i-Martin, 2004). We find in our paper that, for plausible values of the parameters, geography can be naturally considered as an adjustment cost for capital. More specifically the spatial structure, which is explicitly represented in our set up by the diffusion process, allows the neoclassical growth model with perfect capital mobility and irreversibility to avoid immediate adjustment. In terms of social welfare, the model shows that capital mobility is beneficial. When capital is mobile the economy manages to reduce the negative effect of decreasing returns to scale. Capital moves indeed to locations where it is less abundant (and, consequently, with higher shadow price), so the global impact of decreasing returns to scale reduces. Finally, our method also allows us to study the effect of capital mobility on the evolution of (spatial) income inequality. We assume that locations are spatially heterogeneous with respect to their initial levels of capital and, consequently, income inequality arises. The simulations show that capital mobility significantly reduces the Gini coefficient of the economy at anytime. In fact, improving capital mobility reduces the importance of the initial conditions because capital diffuses to locations with lower initial endowment. This outcome is in line with the observation made by Galor (1996) or Quah (1996a,b).

Regarding the analytical application, we study the existence and uniqueness of the social optimum in spatial Ramsey-type models. Under fairly general assumptions our method allows us to prove the existence of the unique optimal solution when the time horizon is finite. To the best of our knowledge this is the first paper dealing with the existence and uniqueness of the solution to an optimal control problem governed by a non-linear parabolic PDE. The practicality of this outcome mainly reduces to ensure the existence of *upper* and *lower solutions*. These are a sort of boundaries to the Pontryagin conditions that characterize the social optimum. In our paper we apply this result to the spatial growth model and to the environmental economics example of Camacho and Pérez-Barahona (2015). We conclude that the solution of these two models exists and is unique, so these problems can be considered as well-posed. With this property in hand, we can thus provide appropriate tools for policymaking, and for the implementation of the social optimum in particular, that involves time and spatial dimensions.

³For the importance of considering a continuum of locations see, for instance, Desmet and Rossi-Hansberg (2009 and 2010) and Thisse (2010).

The paper is structured as follows. Section 2 describes a general spatial version of the Ramsey model. Section 3 presents the result of existence and uniqueness of social optimum. In Section 4 we apply our approach to the two examples of spatial Ramsey-type models, introducing as well the numerical method. Finally, we present our conclusions in Section 5.

2 The general problem

The objective of the policy maker is to maximize the aggregated social welfare in a bounded region of space $\Omega \subset \mathbb{R}$ over a finite time period $[0, T]$. To do so, she chooses a time trajectory for a variable c in each location, taking into account the dynamics of a state variable u_1 . The choice of the control c affects in turn the dynamics of the state. More precisely, the policy maker maximizes the welfare provided by c , aggregating felicity over time and space plus a measure the final spatial situation of the economy, which is represented by a scrap value function Ψ , as follows:

$$\max_c \int_0^T \int_{\Omega} U(c(x, t))g(x, t)dxdt + \int_{\Omega} \Psi(u_1(x, T), x)dx \quad (1)$$

subject to:

$$\begin{cases} u_{1,t}(x, t) - u_{1,xx}(x, t) = f_1(u_1, c), \\ \lim_{x \rightarrow \delta\Omega} u_{1,x}(x, t) = 0, \\ u_1(x, 0) = u_{1,0}(x) \geq 0 \text{ given.} \end{cases} \quad (2)$$

We assume in the objective function (1) that $U \in C^2(\mathbb{R}^+)$ is measurable, increasing and concave. The function $g(x, t)$ represents the usual spatial and time discounting. For instance, following Boucekkine *et al.* (2013b), one can consider the example $g(x, t) = \xi(x)e^{-\rho t}$. This functional form assumes that time is exponentially discounted as in the standard growth model, where $\rho > 0$ is the time-discount rate. Moreover, function $\xi(x)$ is assumed to “rapidly decrease” in order to discount space.⁴ As in Camacho and Pérez-Barahona (2015), the spatial discounting can represent the importance that the policy maker gives to each location of the economy. Moreover, the scrap function can also include space and time discounting since the convergence of the corresponding integral

⁴Notice that both time and space discounting can be dropped in our set-up since time period and space are assumed to be bounded. See Section 5 for further remarks about the decision horizon.

term is required as well. We additionally assume that $\Psi \in C^1(D(u_1), \Omega)$, defined on the domain of u_1 , is a positive function with bounded derivative $\Psi'_1 > 0$.

The first equation in (2) describes the spatial and time dynamics of the state variable u_1 . This is a parabolic PDE, where $u_{1,t}$ denotes the partial derivative with respect to time and $u_{1,xx}$ is the second partial derivative with respect to space. As pointed out in the introduction, this expression can represent the spatio-temporal spread of capital in an economic growth framework. Identifying $u_1(x, t)$ with the capital of location x at time t , $k(x, t)$, the law of motion of capital can be written as

$$k_t(x, t) - k_{xx}(x, t) = I(x, t), \quad (3)$$

where $I(x, t)$ represents the net investment of a location x at time t . The main difference with respect to the standard law of motion without space is the term $-k_{xx}(x, t)$. As in Boucekkine *et al.* (2009) and Boucekkine *et al.* (2013a), it reflects that capital flows from regions with lower marginal productivity of capital (*i.e.*, with abundant capital) to the higher ones, with relatively less capital. They actually show that equation (3) can be economically justified as a balance trade equilibrium in a bounded region.

Camacho and Pérez-Barahona (2015) provide another example of this type of equation in a land use set up, where the dynamics of pollution $p(x, t)$ follows

$$p_t(x, t) - p_{xx}(x, t) = E(x, t). \quad (4)$$

$E(x, t)$ summarizes the emissions in time t of a single source located at x . The term $-p_{xx}(x, t)$ above states that a pollutant diffuses from regions where its concentration is high to regions of lower concentration (Fick's law of diffusion). Indeed, equation (4) is the well-known model in physics called the Gaussian plume, which is frequently used to describe the spread of pollution across locations.

The second expression in (2) is known in the literature as the Neumann boundary condition. It characterizes the behaviour of u_1 in the spatial frontier, assuming that there is no flow of the state variable (capital, for instance) at locations that are far away from the origin.⁵ Finally, the last equation is the initial spatial distribution of u_1 , which is assumed to be a known positive function $u_{1,0}(x) \in C(\bar{\Omega})$.

In the optimal control literature few papers deal with the existence of unique solution in this type of problems. Unique solution exists in some particular cases, namely

⁵This condition can be relaxed by assuming a circular space, so $k(0, t) = k(2\pi, t)$. For further discussion see Boucekkine *et al.* (2013a,b).

a linear $f_1(u_1, c)$ (Lions, 1966a,b; and Barbu and Precupanu, 2012) or under specific non-linearities with very limited interaction between the state and the control variables (Lions, 1966a,b and 1972; and Ahmed, 1977). Still, using an appropriated maximum principle, Pontryagin conditions can be associated to general non-linear problems. Fattorini (1990) and Raymond and Zidani (1999) prove that the set of Pontryagin conditions is necessary and sufficient to characterize the solutions of this problem. However, to the best of our knowledge, we have not found any existence result for this type of dynamical systems.

Using the Ekeland (1974) variational principle, we obtain the associated set of Pontryagin necessary conditions as:⁶

$$(O) \left\{ \begin{array}{l} u_{1,t}(x, t) - u_{1,xx}(x, t) = f_1(u_1, c), \\ u_{2,t}(x, t) + u_{2,xx}(x, t) = f_2(u_1, u_2, c), \\ u_2 \frac{\partial f_1}{\partial c} = -U'(c)g(x, t), \\ \lim_{x \rightarrow \delta\Omega} u_{i,x}(x, t) = 0, \quad i = 1, 2, \\ u_1(x, 0) = u_{1,0}(x) \geq 0, \quad \text{given}, \\ u_2(x, T) = \Psi'_1(u_1(x, T), x), \end{array} \right.$$

where $\delta\Omega = \bar{\Omega} \setminus \overset{\circ}{\Omega}$ is Ω 's frontier and u_2 is the corresponding co-state variable (shadow price). As we will see in the next section, ensuring the existence of unique solution to the system (O) is far from being obvious due to presence of two coupled non-linear parabolic PDE, where one of them is time reversed.

Let us also observe that in the literature of spatial economics one finds some pragmatical ways to tackle the problem and to provide a solution. A linear utility in the objective function allows Boucekkine *et al.* (2009) to decouple the Pontryagin conditions, solving each block of PDE separately. Using dynamic programming in Hilbert spaces, Boucekkine *et al.* (2013a) overcome the unicity problem in a spatial version of the AK growth model, where space is a circle and the time horizon is infinite. Finally, other papers find particular solutions (spatially periodic or time-constant solutions) to the spatial Ramsey-type model as Brito (2004) or Brock and Xepapadeas (2008).

⁶See Appendix A for details.

3 Existence and uniqueness of the solution

We prove in this section the existence of the unique solution to our system of non-linear PDE of real functions on $\Omega \times [0, T]$. For the ease of presentation, since c can be expressed as a function of u_1 and u_2 , let us rewrite (O) as:

$$(I) \begin{cases} u_{1,t}(x, t) - u_{1,xx}(x, t) = f_1(u_1, u_2), \\ u_{2,t}(x, t) + u_{2,xx}(x, t) = f_2(u_1, u_2), \\ \lim_{x \rightarrow \delta\Omega} u_{i,x}(x, t) = 0, \quad i = 1, 2, \\ u_1(x, 0) = u_{1,0}(x) \geq 0, \quad \text{given}, \\ u_2(x, T) = \Psi'_1(u_1(x, T), x). \end{cases}$$

This system is a generalization of the set of necessary and sufficient conditions identified in Boucekkine *et al.* (2009). We should observe that (I) is composed of a parabolic PDE for u_1 , a reversed-time parabolic PDE for u_2 , together with conditions on the partial derivatives of u_1 and u_2 at the spatial borders. Besides, $u_{1,0}$ provides the initial distribution for u_1 . Moreover $\Psi'_1(u_1(x, T), x)$ determines the final distribution for u_2 , which is a function of the final distribution for u_1 . This last condition ties the final distribution of the state variable to the final distribution of the associated co-state. As in Camacho *et al.* (2008) it reveals that, at the end of the planning horizon, the shadow price of the state variable equals the marginal social scrap value of the stock.

In this paper we introduce a monotone method based on Pao (1992). These methods were widely used to prove the existence of solution for systems of PDE, of parabolic and hyperbolic type, with both boundary and initial conditions (see, among others, Sattinger, 1972; Changdra *et al.*, 1978; and Bebernes and Schmitt, 1979). Before going into the details of the proof, let us briefly describe the idea of our method. In a nutshell, we build two converging sequences of functions that unambiguously encircle the unique solution to the system. First, we identify an *upper* and a *lower solution* of the original problem (I), which are typically easier to ensure than the solution itself. Starting then from these solutions, we describe sequences of associated linear problems, which have unique solution. We finally show that these associated problems converge to the original one (I) from above (below), starting with the upper (lower) solution.

More formally, let us begin by describing the norms that we will use to size functions and variables in our method. We mix two norms along the proof: the usual norm in \mathbb{R}^2 and the supremum norm. For every $u(x, t) \equiv (u_1(x, t), u_2(x, t)) \in \mathbb{R}^2$ and at every

$(x, t) \in \Omega \times [0, T]$, the usual Euclidean norm in \mathbb{R}^2 is defined as

$$\| u(x, t) \| \equiv | u_1(x, t) | + | u_2(x, t) | .$$

We also make use of the supremum norm $\| \cdot \|_0$, defined for any $u \in \mathbb{R}^2$:

$$\| u \|_0 \equiv \sup_{(x,t) \in \Omega \times [0,T]} \| u \| .$$

As pointed out above, the dynamics of u_2 is described by a reversed-time parabolic PDE. So one of the key ingredients of the proof is to reverse time in the second PDE in (I). We define in this regard the variables w_1 and w_2 as

$$w_1(x, t) \equiv u_1(x, T - t),$$

$$w_2(x, t) \equiv u_2(x, T - t).$$

Therefore, the dynamics of w_2 is described by a parabolic PDE with a known initial distribution $\Psi'_1(w_1(x, 0), x)$:

$$\begin{cases} w_{2,t}(x, t) - w_{2,xx}(x, t) = -f_2(w_1, w_2), \\ w_2(x, 0) = \Psi'_1(w_1(x, 0), x). \end{cases} \quad (5)$$

We aim at extending the monotone method of Pao (1992), considering in our paper systems of mixed types of parabolic PDE. In order to do so, we need to define adequate upper and lower solutions. We say that two functions $\hat{u} = (\hat{u}_1, \hat{u}_2)$ and $\tilde{u} = (\tilde{u}_1, \tilde{u}_2)$ are upper and lower solutions to (I) if they satisfy for every t :

$$\begin{aligned} \tilde{u}_{1,t}(x, t) - \tilde{u}_{1,xx}(x, t) - f_1(\tilde{u}_1, \tilde{u}_2) &\leq 0 \leq \hat{u}_{1,t}(x, t) - \hat{u}_{1,xx}(x, t) - f_1(\hat{u}_1, \hat{u}_2), \\ \hat{u}_{2,t}(x, t) + \hat{u}_{2,xx}(x, t) - f_2(\hat{u}_1, \hat{u}_2) &\leq 0 \leq \tilde{u}_{2,t}(x, t) + \tilde{u}_{2,xx}(x, t) - f_2(\tilde{u}_1, \tilde{u}_2), \\ \lim_{x \rightarrow \delta\Omega} \hat{u}_{i,x}(x, t) = 0, \quad \lim_{x \rightarrow \delta\Omega} \tilde{u}_{i,x}(x, t) = 0, \quad i = 1, 2, \end{aligned} \quad (6)$$

with $\tilde{u}_1 \leq \hat{u}_1$, $\tilde{u}_2 \leq \hat{u}_2$, and

$$\begin{aligned} \tilde{u}_1(x, 0) &< u_1(x, 0) < \hat{u}_1(x, 0), \\ \tilde{u}_2(x, 0) &< u_2(x, 0) < \hat{u}_2(x, 0), \end{aligned} \quad (7)$$

for all $x \in \Omega$.⁷

⁷We say that a function m is larger than another function m' in $\Omega \times [0, T]$, *i.e.*, $m \geq m'$, if and only if $m(x, t) \geq m'(x, t)$ for all $(x, t) \in \Omega \times [0, T]$.

Due to the presence of a reversed-time parabolic PDE, considering (5), the PDE condition for \hat{u}_2 and \tilde{u}_2 in (6) becomes

$$\tilde{w}_{2,t}(x, t) - \tilde{w}_{2,xx}(x, t) + f_2(\tilde{w}_1, \tilde{w}_2) \leq 0 \leq \hat{w}_{2,t}(x, t) - \hat{w}_{2,xx}(x, t) + f_2(\hat{w}_1, \hat{w}_2). \quad (8)$$

That is to say, \hat{w}_2 and \tilde{w}_2 are, respectively, the upper and lower solution of the corresponding parabolic PDE, as it is usually defined in the literature. Moreover,

$$\tilde{w}_2(x, 0) < \Psi'_1(u_1(x, T), x) < \hat{w}_2(x, 0).$$

Notice that $\tilde{u}_i, \hat{u}_i, i = 1, 2$, are required to be $C(\bar{\Omega} \times [0, T]) \cap C^{1,2}(\Omega \times [0, T])$, that is, they are continuous in $\bar{\Omega} \times [0, T]$, continuously differentiable in t and twice continuously differentiable in x for all $(x, t) \in \Omega \times [0, T]$.

Having established the definition of lower and upper solutions, let us also define the set of continuous functions between \tilde{u} and \hat{u} as

$$\langle \tilde{u}, \hat{u} \rangle \equiv \{(u_1, u_2) \in C(\Omega \times [0, T]) : (\tilde{u}_1, \tilde{u}_2) \leq (u_1, u_2) \leq (\hat{u}_1, \hat{u}_2)\}.$$

The interest of $\langle \tilde{u}, \hat{u} \rangle$ is that it defines the precise space where solutions will be confined, *i.e.*, all solutions are limited from below by \tilde{u} and from above by \hat{u} (see Theorem 2 at the end of this section).

As usual in this literature, we need to impose boundedness conditions on initial distributions and functions f_1 and f_2 in order to avoid explosive solutions:

Assumption 1. $u_{1,0}(x) \in C(\bar{\Omega})$ is measurable. The initial distributions verify that

$$\begin{cases} \tilde{u}_1(x, 0) < h_1 e^{h_2|x|^2} \\ u_1(x, 0) < h_1 e^{h_2|x|^2} \\ \hat{u}_1(x, 0) < h_1 e^{h_2|x|^2} \end{cases} \quad \text{and} \quad \begin{cases} \tilde{u}_2(x, 0) < h_1 e^{h_2|x|^2} \\ \hat{u}_2(x, 0) < h_1 e^{h_2|x|^2} \end{cases}$$

for some positive constants h_1, h_2 with $h_2 < \frac{1}{4T}$.

Assumption 2. There exist bounded functions $\underline{a}_i = \underline{a}_i(x, t)$, $\bar{a}_i = \bar{a}_i(x, t)$, $\underline{c}_i = \underline{c}_i(x, t)$ and $\bar{c}_i = \bar{c}_i(x, t)$, with $i = 1, 2$, such that for every $u, v \in \langle \tilde{u}, \hat{u} \rangle$, where $u = (u_1, u_2)$ and $v = (v_1, v_2)$, (f_1, f_2) satisfies

$$-\underline{a}_1(u_1 - v_1) - \underline{a}_2(u_2 - v_2) \leq f_1(u_1, u_2) - f_1(v_1, v_2) \leq \bar{a}_1(u_1 - v_1) + \bar{a}_2(u_2 - v_2),$$

$$-\underline{c}_1(u_1 - v_1) - \underline{c}_2(u_2 - v_2) \leq f_2(u_1, u_2) - f_2(v_1, v_2) \leq \bar{c}_1(u_1 - v_1) + \bar{c}_2(u_2 - v_2),$$

for $u_1 \geq v_1$ and $u_2 \geq v_2$. Furthermore, functions \underline{a}_1 and \bar{c}_2 are Hölder continuous in $\Omega \times [0, T]$.

We restrict f_i domain to sets in which Assumption 3 below is verified. We denote these reduced domains for f_1 and f_2 by J_1 and J_2 , respectively.

Assumption 3. For every $u, v \in \langle \tilde{u}, \hat{u} \rangle$, either f_1 is increasing in both u_1 and u_2 and $\underline{a}_1 \geq 0$, or $\underline{a}_2 \leq 0$. Similarly, either f_2 is decreasing in both u_1 and u_2 and $\bar{c}_2 \geq 0$, or $\bar{c}_1 \leq 0$.

Finally, we include some continuity assumptions on functions f_i :

Assumption 4. For $i = 1, 2$, functions f_i are Hölder continuous in $J_1 \times J_2$ for some subset $J_1 \times J_2 \subseteq \mathbb{R}^2$.⁸

Defining functions M_1 and M_2 as

$$\begin{aligned} M_1 &\equiv \max\{|\underline{a}_1|, |\underline{a}_2|, |\bar{a}_1|, |\bar{a}_2|\}, \\ M_2 &\equiv \max\{|\underline{c}_1|, |\underline{c}_2|, |\bar{c}_1|, |\bar{c}_2|\}, \end{aligned}$$

functions f_1 and f_2 are Lipschitz continuous since $M_i = M_i(x, t)$ are bounded in $\Omega \times [0, T]$. That is to say, there exist bounded functions $M_i = M_i(x, t)$ in $\Omega \times [0, T]$ such that

$$|f_i(u_1, u_2) - f_i(v_1, v_2)| \leq M_i |u - v|,$$

for any u, v in $\langle \tilde{u}, \hat{u} \rangle$.

With functions \underline{a}_1 and \bar{c}_2 from Assumption 2, we build functions F_1 and F_2 as

$$\begin{aligned} F_1(x, t, u_1, u_2) &\equiv f_1(u_1, u_2) + \underline{a}_1(x, t)u_1, \\ F_2(x, t, u_1, u_2) &\equiv f_2(u_1, u_2) - \bar{c}_2(x, t)u_2. \end{aligned} \tag{9}$$

By Assumptions 2 and 4, functions F_i are Hölder continuous in $\Omega \times [0, T] \times \langle \tilde{u}, \hat{u} \rangle$, for $i = 1, 2$. Furthermore, we also define two linear operators of u_1 and u_2 , \mathbb{L}_1 and \mathbb{L}_2 , as

$$\begin{aligned} \mathbb{L}_1 u_1 &\equiv u_{1,t} - u_{1,xx} + \underline{a}_1 u_1, \\ \mathbb{L}_2 u_2 &\equiv u_{2,t} + u_{2,xx} - \bar{c}_2 u_2. \end{aligned}$$

Introducing the functions F_i and the operators \mathbb{L}_i defined above, we can rewrite the two PDE in (I) as

$$\mathbb{L}_i u_i = F_i(u_1, u_2),$$

in $\Omega \times [0, T]$ for $i = 1, 2$.

⁸Our results can be easily extended to functions f_i that depend on x and t , i.e., $f_i(x, t, u_1, u_2)$. In this case, f_i would be Hölder continuous in $\Omega \times [0, T] \times J_1 \times J_2$.

As pointed out before, the PDE corresponding to u_2 is parabolic after reversing time (see Equation 5). So we define an alternative operator to \mathbb{L}_2 in order to consider only parabolic PDEs:

$$\mathbb{L}'_2 w_2 \equiv w_{2,t} - w_{2,xx} + \bar{c}_2 w_2.$$

Taking then the operator \mathbb{L}'_2 , the PDE in (5) becomes

$$\mathbb{L}'_2 w_2 = -F_2(w_1, w_2).$$

This new writing turns out to be extremely useful in the upcoming Theorem 1. It allows us to consider sequences of single linear parabolic PDE that we will use in order to prove the existence of the unique solution for (I). As we briefly explained before, this theorem begins by building two monotonic sequences of functions. We set up, on the one hand, a non-increasing sequence of continuous functions of time and space $\{\bar{u}_1^k, \bar{u}_2^k\}_{k \in \mathbb{N}}$ and, on the other hand, a non-decreasing sequence of continuous functions $\{\underline{u}_1^k, \underline{u}_2^k\}_{k \in \mathbb{N}}$. Without loss of generality, we establish $\{\bar{u}_1^k\}_{k \in \mathbb{N}}$ initiating the sequence with the upper solution $\{\hat{u}_1, \hat{u}_2\}$. At every step, we generate \bar{u}_1^k taking $(\bar{u}_1^{k-1}, \bar{u}_2^{k-1})$. Then, with $\bar{w}_1^k(x, t) = \bar{u}_1^k(x, T - t)$ and $(\bar{w}_1^{k-1}, \bar{w}_2^{k-1})$, we will obtain \bar{w}_2^k . The step ends when we reverse time again to obtain $\bar{u}_2^k(x, t) = \bar{w}_2^k(x, T - t)$. We can repeat this procedure at infinitum to build our sequence. Regarding the sequences $\{\underline{u}_1^k\}_{k \in \mathbb{N}}$ and $\{\underline{u}_2^k\}_{k \in \mathbb{N}}$, we follow a similar strategy, initiating the process with the lower solution $\{\tilde{u}_1, \tilde{u}_2\}$. We then prove that these sequences uniformly converge:

$$\lim_{k \rightarrow \infty} \bar{u}_i^k = \bar{u}_i \quad \text{and} \quad \lim_{k \rightarrow \infty} \underline{u}_i^k = \underline{u}_i, \quad i = 1, 2.$$

Finally we show that $\mathbb{L}^{-1}F$ possesses the contraction property, where we define the diagonal matrix $\mathbb{L} \equiv \text{diag}(\mathbb{L}_1, \mathbb{L}_2)$ and the transpose of the vector of F_i functions $F \equiv (F_1, F_2)'$. So the limits of the above sequences coincide and define the unique solution of (I).

Following these steps, Theorem 1 states that problem (I) has a unique solution in $\Omega \times [0, T]$ under the assumptions established before:

Theorem 1. *Under the Assumptions 1-4, the problem*

$$\left\{ \begin{array}{l} u_{1,t}(x, t) - u_{1,xx}(x, t) = f_1(u_1, u_2), \\ u_{2,t}(x, t) + u_{2,xx}(x, t) = f_2(u_1, u_2), \\ \lim_{x \rightarrow \partial\Omega} u_{i,x}(x, t) = 0, \quad i = 1, 2, \\ u_1(x, 0) = u_{1,0}(x) \quad \text{given}, \\ u_2(x, T) = \Psi'_1(u_1(x, T), x), \end{array} \right. \quad (10)$$

has a unique solution (u_1, u_2) in $\Omega \times [0, T]$. Moreover, $u = (u_1, u_2)$ is the limit of the monotonic sequence $\{u_1^k, u_2^k\}_{k \in \mathbb{N}}$, where $u_2^k(t, x) = w_2^k(x, T - t)$. The sequence $\{u_1^k, w_2^k\}_{k \in \mathbb{N}}$ is generated as

$$\begin{cases} \mathbb{L}_1 u_1^k = F_1(u_1^{k-1}, u_2^{k-1}), & \lim_{x \rightarrow \delta\Omega} u_{1,x}^k(x, t) = 0, \quad \forall t \geq 0, \\ \mathbb{L}'_2 w_2^k = -F_2(w_1^{k-1}, w_2^{k-1}), & \lim_{x \rightarrow \delta\Omega} w_{2,x}^k(x, t) = 0, \quad \forall t \geq 0, \end{cases} \quad (11)$$

for all $k = 1, 2, \dots$. We can choose between two sets of initial conditions and initial distributions. The first set (\hat{S}) of conditions below generates a non-increasing sequence, whereas the second set (\tilde{S}) generates a non-decreasing sequence:

$$\begin{aligned} (\hat{S}) \quad & \begin{cases} (u_1^0, u_2^0) = (\hat{u}_1, \hat{u}_2), \\ u_1^k(x, 0) = \frac{\Upsilon}{k} \hat{u}_1(x, 0) + (1 - \frac{\Upsilon}{k}) u_1(x, 0), \\ w_2^k(x, 0) = \frac{\psi}{k} \hat{w}_2(x, 0) + (1 - \frac{\psi}{k}) \Psi'_1(w_1^{k-1}(x, 0), x); \end{cases} \\ (\tilde{S}) \quad & \begin{cases} (u_1^0, u_2^0) = (\tilde{u}_1, \tilde{u}_2), \\ u_1^k(x, 0) = \frac{\zeta}{k} \tilde{u}_1(x, 0) + (1 - \frac{\zeta}{k}) u_1(x, 0), \\ w_2^k(x, 0) = \frac{\phi}{k} \tilde{w}_2(x, 0) + (1 - \frac{\phi}{k}) \Psi'_1(w_1^{k-1}(x, 0), x), \end{cases} \end{aligned}$$

with $0 < \zeta < \Upsilon < 1$ and $0 < \phi < \psi < 1$.

Proof. See Appendix B. □

We have proven so far that, once a pair of lower and upper solutions is identified, there exists a unique solution u to our original problem between them. A natural question arises, wondering about how many solutions one could find below (above) the specific lower (upper) solution. The answer is that, under the assumptions of the model, we will not find any. So the unique solution of the problem will be unambiguously located between the pair of lower and upper solutions, *i.e.*, $u \in \langle \tilde{u}, \hat{u} \rangle$. We state this final result in the following theorem:

Theorem 2. *Let Assumptions 1-4 hold, and \tilde{u} and \hat{u} be lower and upper solutions to problem (I). Let us further assume that:*

- i) either f_1 is increasing in u_1 and u_2 , with $\underline{a}_1 \geq 0$; or $\underline{a}_2 \leq 0$,*
- ii) either f_2 is decreasing in u_1 and u_2 , with $\bar{c}_2 \geq 0$; or $\bar{c}_1 \leq 0$.*

If u is a solution to (I) then $u \in \langle \tilde{u}, \hat{u} \rangle$.

Proof. See Appendix C. □

4 Applications

Let us apply our method to two examples of spatial-Ramsey type models. We consider first the spatial growth model originally introduced by Brito (2004) and Boucekkine *et al.* (2009). The second example is based on the optimal land-use framework of Camacho and Pérez-Barahona (2015). In both models ill-posedness had not been solved this far, even under the case of finite decision horizon. Boucekkine *et al.* (2013a) is an exception in this respect, but they focus on the case of linear production function (AK model). Camacho and Pérez-Barahona (2015) allow for a non-linear production function. However, their analytical results about the existence of unique equilibrium exclusively apply to the time-invariant solution.

Making use of Theorems 1 and 2, we will conclude that the social optimum exists and is unique. Hence, when the time horizon is finite, we can confirm that these models are well-posed. From a numerical perspective, we will implement the iterative method presented in Theorem 1 in order to simulate the optimal trajectories of the spatial growth model. In particular, we will apply this algorithm to study the role played by capital mobility in economic growth.⁹

4.1 The spatial growth model

Within the general framework of economic growth theory, the spatial Ramsey model considers a policy maker that maximizes utility in a region Ω of space from time 0 to T , subject to the dynamic evolution across time and space of the capital stock.¹⁰ The corresponding optimal control problem is

$$\max_c \int_0^T \int_{\Omega} U(c(x, t)) e^{-\rho t} dx dt + \int_{\Omega} \Psi(k(x, T)) e^{-\rho T} dx \quad (12)$$

⁹Camacho and Pérez-Barahona (2015) already include a numerical analysis of our second example. Their computational setting is alternatively based on a discretization of the Pontryagin conditions, starting from an initial guess for the reversed-time shadow price.

¹⁰Spatial economics usually considers space as a physical characteristic of the problem. However, it is possible to think about alternative interpretations, such as the idea of “economic distance” suggested by papers like Patel (1964) or Quah (1996a).

subject to

$$\begin{cases} k_t(x, t) - k_{xx}(x, t) = Af(k(x, t)) - \delta k(x, t) - c(x, t), \forall (x, t) \in \Omega \times [0, T] \\ \lim_{x \rightarrow \delta\Omega} k_x(x, t) = 0 \quad \forall t \in [0, T], \\ k(x, 0) = k_0(x) > 0 \quad \text{given} \quad \forall x \in \Omega. \end{cases} \quad (13)$$

The policy maker decides on the consumption $c(x, t)$ of the representative household at every location x and time t . In this respect, she takes into account the spatial-dynamic motion of the stock of capital in each location $k(x, t)$. So the first constraint in (13) represents the law of motion of $k(x, t)$. The right hand side of this equation is the net investment, where $\delta \in (0, 1)$ is the rate of capital depreciation and $f(\cdot)$ is the production technology of the economy, with a scale parameter A . As we already observed in Section 2 (equation 3), this expression is a parabolic PDE where $k_{xx}(x, t)$ introduces a new feature with respect to the standard law of motion of capital. This term entirely comes from the spatial mobility of capital, which flows from regions with lower marginal productivity of capital to the higher ones. The second constraint is the (Neumann) boundary condition, which assumes that there is no capital flow at locations that are far away from the origin. Finally $k_0(x) > 0$ is the initial spatial distribution of capital, which is given in the economy.

$\Psi(\cdot) > 0$ measures how much the policy maker cares about the final state of capital, or scrap value. It is a continuously differentiable function with bounded and positive derivative. We assume that U and f are both positive, continuous, differentiable and concave functions. Only settings with measurable k_0 and existing integrals are considered. Following the variational principle in Appendix A, the Pontryagin optimal conditions associated to this problem are

$$\begin{cases} k_t(x, t) - k_{xx}(x, t) = Af(k(x, t)) - \delta k(x, t) - c(x, t), \forall (x, t) \in \Omega \times [0, T], \\ \lambda_t(x, t) + \lambda_{xx}(x, t) = \lambda(\delta + \rho - Af'(k(x, t))), \forall (x, t) \in \Omega \times [0, T], \\ U'(c(x, t)) = \lambda(x, t), \forall (x, t) \in \Omega \times [0, T] \\ \lim_{x \rightarrow \delta\Omega} k_x(x, t) = \lim_{x \rightarrow \delta\Omega} \lambda_x(x, t) = 0, \quad \forall t \in [0, T], \\ k(x, 0) = k_0(x) \quad \text{given}, \quad \forall x \in \Omega, \\ \lambda(x, T) = \Psi'(k(x, T)), \quad \forall x \in \Omega, \end{cases}$$

where λ is the corresponding co-state variable, *i.e.*, the shadow price of capital.

4.1.1 Analytical application

For this illustration we consider the usual CRRA utility function $U(c) = \frac{c^{1-\sigma}}{1-\sigma}$ with $\sigma > 0$, together with a Cobb-Douglas production technology $f(k) = k^\alpha$ where $\alpha \in (0, 1)$. We

assume, for the sake of simplicity, a linear scrapping value $\Psi(k(x, T)) = \chi k(x, T)$ with $\chi > 0$. Using our general formulation (I), $f_1(k, \lambda) = Af(k(x, t)) - \delta k(x, t) - \lambda(x, t)^{-1/\sigma}$ and $f_2(k, \lambda) = \lambda(\delta + \rho - Af'(k(x, t)))$. Let us apply Theorem 1 in order to prove the existence of the unique social optimal solution of the spatial growth model. As we pointed out in the previous section, the main idea of the theorem is to identify upper and lower solutions to the system (I). This will ensure the existence of unique solution within the space that the upper and lower solutions outline. Taking the definition (6)-(8) of this type of solutions, one can verify that the following couple is a lower solution to our problem :

$$(\tilde{k}, \tilde{\lambda}) = \left(\left(\frac{\alpha A}{\delta + \rho} \right)^{\frac{1}{1-\alpha}} - \epsilon, \left(A\tilde{k}^\alpha - \delta\tilde{k} \right)^{-\sigma} + \epsilon \right),$$

with $\epsilon > 0$ a constant. Moreover, the couple below provides an upper solution:

$$(\hat{k}, \hat{\lambda}) = (\varphi e^{\phi t}, \psi e^{\zeta t}),$$

with

$$\begin{aligned} A\alpha\varphi^{\alpha-1} &\leq \delta + \rho - \zeta, \\ e^{\frac{\zeta+\phi}{\sigma}T} &\leq \frac{\psi^{-1/\sigma}}{A\varphi^\alpha - \varphi(1+\delta)}, \end{aligned}$$

for positive constants $\varphi, \psi, \phi, \zeta > 0$.¹¹

As stated in Theorem 1 we also have to ensure several boundary and continuity requirements (Assumptions 1-4). It is easy to see that Assumption 1 holds if we consider non-explosive initial distributions for the capital and the upper and lower solutions. Since f_1 and f_2 are continuously differentiable functions, they are Hölder continuous as well. So Assumption 4 is verified too. With respect to Assumption 2, this is guaranteed if we focus on interior non-explosive solutions. Lastly, notice that f_1 is increasing in both k and λ .¹² One can also prove that f_2 is decreasing in λ . However f_2 is increasing in k since $\lambda > 0$ and the production function is concave in k . We should then choose $\underline{a}_1 \geq 0$ and $\bar{c}_1 \leq 0$ according to Assumption 3. Moreover, since f_2 is increasing in k but decreasing in λ , we can set $\bar{c}_1 = 0$ in order to ensure that Assumption 2 holds.

Summing-up, we have identified an upper and a lower solution. Since the assumptions of Theorem 1 hold for the functional forms of this illustration, we can then confirm that there exist a unique solution to our problem. Considering finally Theorem 2, we

¹¹Notice that these constants can be conveniently set in order to verify the conditions $\tilde{k}(x, 0) < k(x, 0) < \hat{k}(x, 0)$ and $\tilde{\lambda}(x, T) < \Psi'_1(k(x, T), x) < \hat{\lambda}(x, T)$.

¹² $\partial f_1 / \partial \lambda > 0$ because $\lambda(x, t) = U'(c(x, t)) > 0$. For $\partial f_1 / \partial k > 0$ the scale parameter A should be high enough, *i.e.*, $A > \delta(\varphi e^{\phi T})^{1-\alpha} / \alpha$.

can guarantee that every solution lies between the upper and lower solutions, so we are not missing any additional solution. Hence, we can state the following proposition:

Proposition 1. *The spatial growth model (12)-(13) is well-posed.*

4.1.2 Numerical application

Let us consider the initial set (\tilde{S}) or (\hat{S}) in Theorem 1. Each of these sets defines the initialization of an algorithm starting from the lower or the upper solution, respectively. Our numerical method iteratively applies (11) in order to generate the sequences stated in the theorem. Notice that, since we have proved that both sequences converge, the algorithm can be initialized either with the lower or with the upper solution. The details of the algorithm are provided in Appendix D.

For this numerical example we consider the same functional forms as in the analytical application. Our method involves functions F_1 and F_2 defined in (9), so we have to specify \underline{a}_1 and \bar{c}_2 . Taking into account the Assumptions 2 and 3, one can verify that $\underline{a}_1 = \delta$ and $\bar{c}_2 = \delta + \rho$ satisfy all the requirements since we have set $\bar{c}_1 = 0$. All parameters values used in the simulations are displayed in Table 1. For the sake of illustration we set the scale parameter $A = 1$. Regarding the CRRA utility, Attanasio and Weber (1993) and Attanasio and Browning (1995) identify an interval for σ between 1.25 and 3.33 using micro household data. Consistently with this range, Barro *et al.* (1995) assume $\sigma = 2$. Chetty (2006) however observes that, even if macroeconomics tends to suggest values of σ greater than in microeconomic studies, there is an upper bound implying that $\sigma < 2$. So in line with these papers we choose for our application $\sigma = 1.95$. As in Barro *et al.* (1995) and Barro and Sala-i-Martin (2004), we consider a time discount rate $\rho = 0.02$ and a depreciation rate $\delta = 0.05$. We set the output elasticity $\alpha = 0.75$ following the notion of a broad measure that also includes knowledge capital (see, for instance, Barro and Sala-i-Martin, 2004; and Acemoglu, 2009). For the scrapping function, we choose a marginal value $\chi = 0.1$ in order to reduce the direct welfare gains due to the capital left at the end of the decision horizon. Finally we set the constant ϵ of the lower solution to 1. Indeed, as observed before, the iterations can start with either the upper or the lower solutions. Since for this model the lower solutions are simpler (see analytical application), we decide to initialize the algorithm with the set (\tilde{S}) .¹³

¹³As robustness checks we repeated simulations for $\alpha = 1/3$, which is a value frequently assumed for a restricted definition of capital. We also tried alternative marginal scrap values $\chi = \{0.001, 0.01, 0.5, 1\}$.

A	Scale parameter	1
σ	Utility parameter	1.95
ρ	Time discount rate	0.02
δ	Depreciation rate	0.05
α	Output elasticity	0.75
χ	Scrap parameter	0.1
ϵ	Lower solution constant	1

Table 1: Parameters values

For this illustration we assume a decision horizon $T = 10$, considering that the region Ω of space is a real line of length 10. Since our numerical method requires that both time and space are discretized, we divide them into 100 and 20 points respectively. So following the notation of Appendix D, $J = 100$ and $N = 20$ with $Dt = 0.1$ and $Dx = 0.5$. The economic aim of this application is to investigate the role played by capital mobility, paying particular attention to the evolution of spatial inequalities across locations. So we will assume that locations are spatially heterogenous with respect to their capital endowment. We specifically consider that the initial distribution of capital is given by $k_0(x) = 164 - 40 \log(x)$ for x in our discretized space $\Omega \simeq \{Dx, 2 * Dx, \dots, N * Dx\}$. Therefore, the endowment of capital stock is assumed to range from 164 to 44.17 as we move away from the origin.

We first consider the case of perfect capital mobility. For this scenario the algorithm needs 528 iterations to converge, with an Euclidean distance between two last iterations lower than 0.1 for both k and λ . We plot in Figure 1 the optimal trajectories for the capital and its shadow price. We compare this case with the other extreme situation where there is no spatial diffusion. The policy maker still maximises the utility in the whole space. However, capital is not allowed to move from one location to another. For this simulation the algorithm needs 50 iterations to converge and, as in the previous case, the Euclidean distance between two iterations is lower than 0.1. Figure 2 reproduces the corresponding simulations.

We can see from these figures that the economy does not immediately adjust even if capital is perfectly mobile. This outcome contrasts with the open-economy neoclassical growth model introduced in Barro *et al.* (1995). Transitional dynamics is a desir-

In all these inspections we get similar qualitative results. We additionally verified that the simulated trajectories remain identical for different lower solutions, with $\epsilon = \{0.1, 0.5, 1.5, 2\}$.

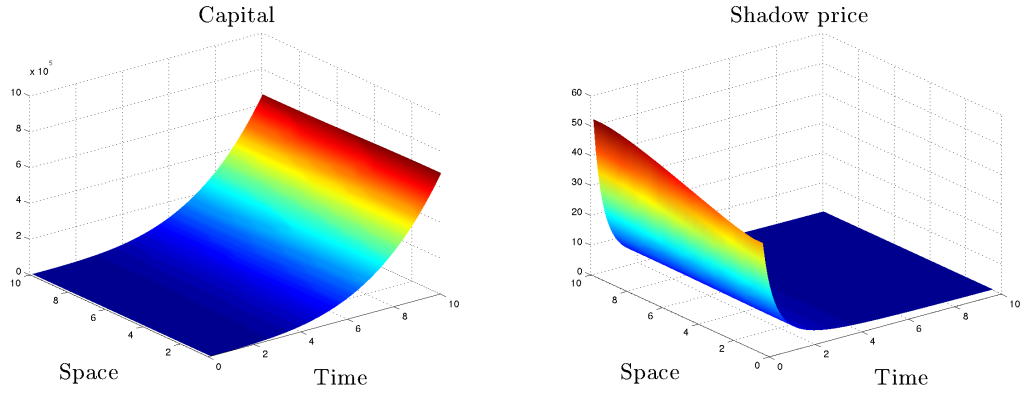


Figure 1: Spatial mobility

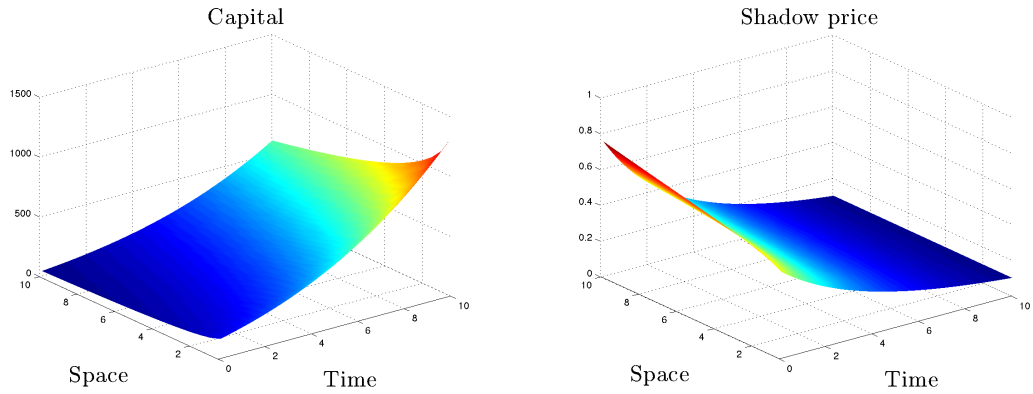


Figure 2: No spatial mobility

able property for a realistic economic model. However they observe that, under the reasonable assumptions of capital irreversibility and economic openness, perfect capital mobility is only compatible with this property if we include capital adjustment costs. Nevertheless, an important limitation of their set-up is that space is not explicitly included. We show in this regard that adding spacial structure, represented in our framework by the diffusion process, avoids immediate adjustment in the growth model even if capital is perfectly mobile. This conclusion is in line with Boucekkine *et al.* (2013a), where transitional dynamics arises in the AK growth model with explicit spatial structure. Within the context of Barro *et al.* (1995)'s observation, geography introduces an additional adjustment cost for capital.

Another remarkable feature that our model points out is the effect of capital mobility on social welfare. Capital mobility allows the economy to reduce the negative effect of decreasing returns to scale. When capital is mobile, it flows to locations with lower abundance (and, consequently, with higher shadow price). So the global impact of decreasing returns to scale reduces. As we can see in our simulations, capital increases with time in the two scenarios. But at the end of the decision horizon, both the minimum and the maximum stock of capital in the space are greater when there is spatial mobility (without mobility, $\min_{|\Omega}(k(x, T)) = 611.93$ and $\max_{|\Omega}(k(x, T)) = 1312.9$; while $\min_{|\Omega}(k(x, T)) = 6.55 \times 10^5$ and $\max_{|\Omega}(k(x, T)) = 6.74 \times 10^5$ with mobility). Taking the evolution of the shadow price, we compute $c(x, t)$ since $U'(c(x, t)) = \lambda(x, t)$. We can then compare the social welfare in each type of economy, considering as well $k(x, T)$ and the corresponding scrap value. We find that capital mobility provides greater social welfare to the economy because capital can be spatially reallocated, thus inducing efficiency gains. This result is consistent with the importance that economic integration gives to capital mobility. Indeed, the elimination of any restrictions on the movement of capital is a primary objective of, for instance, the Economic and Monetary Union of the EU (EU, 2008; and, for a revision of the literature, see Kose *et al.*, 2010).

Let us complete the illustration with the evolution of income inequality in the economy, considering the effect of capital mobility. The income of each location at time t is defined as $y(x, t) = Af(k(x, t))$. Taking the simulated trajectories for $k(x, t)$, we can analyze the dynamics of the spatial income distribution by means of computing the Lorentz curve and the Gini coefficient of the economy in each time t . Figure 3 represents the evolution of spatial inequality, with and without capital mobility.

Since we have assumed an initial distribution of capital that is spatially heterogeneous, Figure 3(a) depicts the resulting income inequality that corresponds to a Gini coefficient of 0.16. In both scenarios, due to the decreasing returns to scale, inequality reduces as capital accumulates (see Figure 3b). However, the spatial mobility of capital allows the economy to achieve lower inequality (blue line). This is also clear if we evaluate the level of inequality in the economy at the end of the decision horizon. Figures 3(c) and (d) show that the area between the line of equality (*i.e.*, 45 degree line) and the Lorentz curve is smaller when capital is spatially mobile. The Gini coefficient is 0.004 when the capital is mobile, which contrasts with a higher value (around 0.1) for the case where spatial mobility is not allowed. Consistently with the statement presented in papers such as Galor (1996) and Quah (1996a,b), our simulations show that improving capital mobility reduces the importance of the initial differences among locations. As

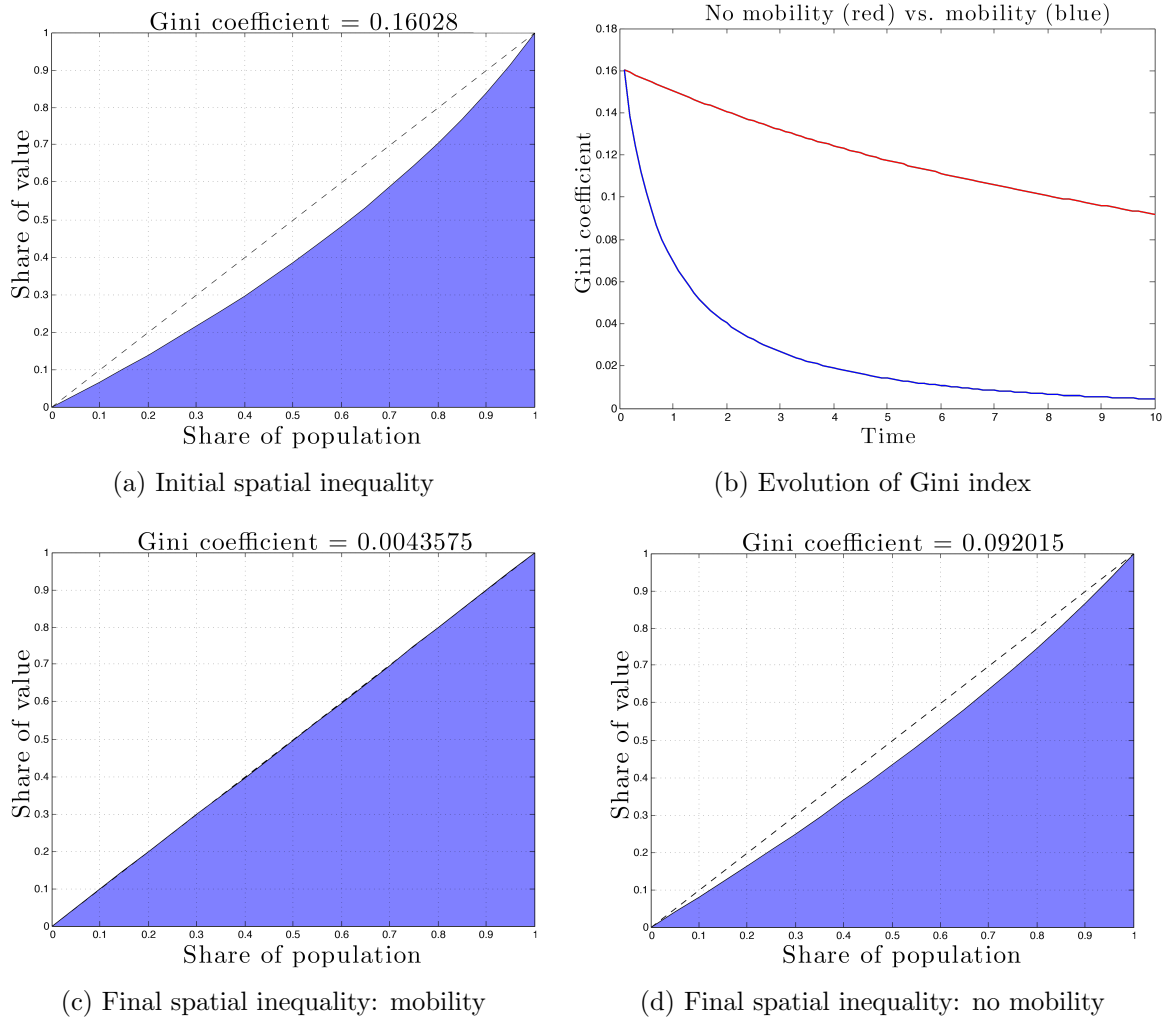


Figure 3: Income distribution

before, the higher social value of capital in locations with lower abundance induces the spatial flow of this input.

4.2 Land use dynamics and the environment

Our last application considers the spatial dynamic model introduced by Camacho and Pérez-Barahona (2015). This paper applies the spatial-Ramsey framework to environmental economics and, in particular, to study the social optimal use of land. It is also consistent with the predictions of spatial models of natural resources such as the harvesting stochastic spatial approach of Costello and Polasky (2008).

We assume an economy where each location has a unit of land, which can be de-

voted to three different activities: production, housing and pollution abatement. For simplicity, the space required for housing at each location is equal to its population density $f(x)$. We also consider no population growth. In this economy there exists a unique consumption good, which only requires land as a production input. We denote the production by $F(l)$. Production generates pollution. However, the remainder of the land can be used as pollution abatement with a technology $G(1 - l - f(x))$. The policy maker chooses the level of consumption *per capita* and the use of land at each location that maximize the discounted social welfare of the entire population. An important feature of this problem is that, even if land is immobile by nature, local actions affect the whole space because pollution flows across locations. We describe the spatial dynamics of pollution by means of the Gaussian plume model presented in Section 2 (equation 4).

The policy maker maximizes the aggregate welfare in a bounded region of space R over a finite time period $[0, T]$, where $c(x, t)$ denotes the consumption *per capita* at location x and time t :

$$\max_{\{c, l\}} \int_0^T \int_R u(c(x, t)) f(x) e^{-\rho t} dx dt + \int_R \psi(p)(x, T) e^{-\rho T} dx \quad (14)$$

subject to

$$\begin{cases} p_t(x, t) - p_{xx}(x, t) = P(p, x)A(x, t)F(l(x, t)) - G(1 - l - f(x)), \\ \int_R c(x, t)f(x)dx = \int_R P(p, x)A(x, t)F(l)dx, \\ \lim_{x \rightarrow \delta R} p_x(x, t) = 0, \\ p(x, 0) = p_0(x) \geq 0. \end{cases} \quad (15)$$

Following the example considered in Camacho and Pérez-Barahona (2015), we introduce two discounts in the objective function (14). The term $e^{-\rho t}$ is the usual temporal discount with $\rho > 0$. The spatial discount $f(x)$ represents the weight that the policy maker gives to each location. This can be identified, for instance, as the population density in order to avoid any subjective spatial preferences. In contrast to the application considered in the previous section, the scrap value represents in this set-up the policy maker's concern about the state of pollution at the end of the planning horizon. Since pollution is assumed to reduce social welfare, the scrap function will be considered as negative.

In this economy pollution damages production. If we denote by $A(x, t)$ the total factor productivity at location x at time t , we have that this location produces

$P(p, x)A(x, t)F(l)$ units of final good when it devotes an amount l of land to production. $1 - P$ represents the share of foregone production due to pollution, with a function $P(p, x)$ decreasing in pollution. The first equation in (15) is the Gaussian plume, where the corresponding emissions $E(x, t)$ represents the instantaneous net emissions of each location, *i.e.*, $PAF(l) - G$. As in Camacho and Pérez-Barahona (2015), the second constraint in (15) allows for the possibility of spatial reallocation of production. So the policy maker collects all the production and reallocate it across locations at no cost. The limit constraint of the problem is the usual boundary condition: it is assumed that there is no pollution flow in the boundaries of the space. Finally $p(0, x)$ represents the initial spatial distribution of pollution.

The Pontryagin conditions of the problem are:

$$\left\{ \begin{array}{l} p_t(x, t) - p_{xx}(x, t) = P(p, x)A(x, t)F(l(x, t)) - G(1 - l - f(x)), \\ q_t(x, t) + q_{xx}(x, t) + P'(p, x)A(x, t)F(l) [u'(c(x, t)) + q(x, t)] + \rho q = 0, \\ [u'(c(x, t)) + q(x, t)] P(p, x)A(x, t)F'(l) + q(x, t)G'(1 - l - f(x)) = 0, \\ \int_R c(x, t)f(x)dx = \int_R P(p, x)A(x, t)F(l)dx, \\ p(x, 0) = p_0(x) \geq 0, \\ \lim_{x \rightarrow \delta R} p_x(x, t) = 0, \quad \lim_{x \rightarrow \delta R} p(x, t)q_x(x, t) = 0, \\ q(x, T) = \psi'(x, T), \end{array} \right. \quad (16)$$

for $(x, t) \in R \times [0, T]$.

As in Boucekkine *et al.* (2009), we should observe that the boundary condition for the shadow price $q(x, t)$ becomes the standard $\lim_{x \rightarrow \delta R} q_x(x, t) = 0$ if we focus on interior solutions. Moreover, the last expression in (16) is the terminal condition of the problem. It states that, at the end of the planning horizon, the shadow price of pollution is equal to the marginal concern of the policy maker about the pollution left behind.

Camacho and Pérez-Barahona (2015) show that, due to the possibility of consumption “imports” (represented in this framework by the consumption reallocation), consumption is spatially homogenous. So it is possible to prove that $c(x, t) = c(t) = \int_R P(p, x)A(x, t)F(l)dx / \int_R f(x)dx$. Moreover, replacing this expression in the equation describing the trade-off between consumption and pollution (third equation in 16), one can prove that the optimal land $l(x, t)$ is a unique function of $p(x, t)$ and $q(x, t)$. Consequently, the spatial dynamic behaviour of this economy is described by a set of equations that correspond to our general system (I) in Section 3. We should notice though that

our general formulation assumes that both state and co-state variables are positive. Therefore, we do the change of variable $\tilde{q}(x, t) = -q(x, t)$ since the shadow price of pollution is negative.

Let us apply Theorems 1 and 2 in order to study the existence of the unique solution to this economy. One can verify first that the assumptions of both theorems hold for the functional forms considered in their paper. We should also observe that f_1 and f_2 are decreasing functions in both arguments p and \tilde{q} . So we have to set $\underline{a}_2 \leq 0$ and $\bar{c}_2 \geq 0$. The last components to be ensured are the upper and lower solutions to the problem. Camacho and Pérez-Barahona (2015) adapt Theorem 3.4 in Pao (1992) to study the existence of a unique time-invariant solution. In order to apply this existence result they identify time-invariant upper and lower solutions. But these stationary solutions are just particular cases of upper and lower solutions. We can then use directly our theorems, concluding that this problem has a unique solution:

Proposition 2. *The land use dynamic problem (14)-(15) is well-posed.*

As it is clear from the last argument above, our paper generalizes Pao (1992)'s stationary method to the full dynamic problem. From this perspective, an additional novelty of our approach is that it is not exclusively constrained to the time-invariant solution.

5 Concluding remarks

We have proposed a general method to study the optimal solution of dynamic models based on the spatial Ramsey framework. The iterative nature of this approach also allowed us to provide a new algorithm to simulate the evolution in time and space of the economy. We included two economic illustrations of our technique. Considering the spatial growth model, we numerically investigated the role played by capital mobility in economic growth. We pointed out as well the effect on social welfare and on the evolution of spatial inequality. From an analytical point of view, we proved the existence of unique optimal solution in spatial Ramsey-type models. As a matter of example, we applied this result to the spatial growth model and the optimal land-use set-up of Camacho and Pérez-Barahona (2015). We concluded that both spatial-dynamic problems are well-posed.

As in the existing mathematical literature, we focused on the case of a finite decision

horizon. Considering a time span that is finite allows us for instance to be able to reverse time, which is a necessary step in this technique. In our approach the finite horizon can be as large as one wishes. However, it is worth paying attention to the case of infinite horizon too. This case, even if beyond the scope of the paper, is particularly relevant to understanding the asymptotic behaviour of the economy. Boucekkine *et al.* (2013b) point out in this respect that infinite time horizon is still an open question in this new field of spatial economics.

Several extensions can be made to our paper. The decentralization of the social optimum in spatial Ramsey-type models has not been explored yet in the literature. This would introduce the possibility of implementing public policy schemes that take into account the geographical characteristics of the economy. From this perspective, and in line with papers considering the interaction between capital mobility and tax coordination (among others, Bental, 1985; and Rodrik and van Ypersele, 2001; and Baldwin and Krugman, 2004), our approach would be also useful to discuss the optimal degree of capital mobility in the economy. Another possible extension would incorporate human capital into the spatial growth model. As observed in the numerical application, we consider a broad definition of capital that includes knowledge as well. This is the case of, for example, tradable knowledge capital through patents. But one could also think about human capital, which is reasonably less mobile (Galor, 1996; and Baldwin and Martin, 2004). Since our set-up admits multiple state variables, we could thus distinguish between physical and human capital. The spatio-temporal motion of each stock would be described by means of a PDE, assuming a lower diffusion degree for the human capital.

Appendices

A Pontryagin optimal conditions

Let us define the value function associated to our problem:

$$\begin{aligned}
 V(c, u_1, u_2) &\equiv \int_0^T \int_{\Omega} U(c(x, t)) g(x, t) dx dt + \int_{\Omega} \Psi(u_1)(x, T) dx \\
 &- \int_0^T \int_{\Omega} u_2(x, t) [u_{1,t}(x, t) - u_{1,xx}(x, t) - f_1(u_1, u_2)] dx dt,
 \end{aligned}
 \tag{17}$$

where u_2 is an auxiliary function. V is a function of c , u_1 and u_2 . If there exists an optimal solution (c^*, u_1^*, u_2^*) , then any other solution to problem (1)-(2) can be written as a deviation from the optimal solution as

$$\begin{aligned} c(x, t) &= c^*(x, t) + \epsilon \kappa(x, t), \\ u_1(x, t) &= u_1^*(x, t) + \epsilon v_1(x, t), \\ u_2(x, t) &= u_2^*(x, t) + \epsilon v_2(x, t), \end{aligned} \tag{18}$$

for $\epsilon \in \mathbb{R}$. We can take the first order derivative of the value function V with respect to ϵ in order to minimize the deviation of the trajectory from the optimal. Beforehand and using integration by parts, we rearrange some integral terms in V :

$$\begin{aligned} \int_0^T \int_{\Omega} u_2(x, t) u_{1,xx}(x, t) dx dt &= \int_0^T u_2(x, t) u_{1,x}(x, t) |_{\delta\Omega} dt - \int_0^T u_{2,x}(x, t) u_1(x, t) |_{\delta\Omega} dt \\ &+ \int_0^T \int_{\Omega} u_{2,xx}(x, t) u_1(x, t) dx dt, \end{aligned} \tag{19}$$

and as usual:

$$\begin{aligned} \int_0^T \int_{\Omega} u_2(x, t) u_{1,t}(x, t) dx dt &= \int_{\Omega} u_1(x, T) u_2(x, T) |_0^T dx - \int_0^T \int_{\Omega} u_1(x, t) u_{2,t}(x, t) dx dt \\ &= \int_{\Omega} u_1(x, T) u_2(x, T) dx - \int_{\Omega} u_1(x, 0) u_2(x, 0) dx - \int_0^T \int_{\Omega} u_2(x, t) u_{2,t}(x, t) dx dt. \end{aligned} \tag{20}$$

We then obtain:

$$\begin{aligned} \frac{\partial V(c, u_1, u_2)}{\partial \epsilon} &= \int_0^T \int_{\Omega} U'(c(x, t)) \kappa(x, t) g(x, t) dx dt + \int_{\Omega} \Psi'_1(u_1(x, T)) v_1(x, T) dx \\ &+ \int_0^T \int_{\Omega} v_1(x, t) [u_{2,t}(x, t) + u_{2,xx}(x, t)] dx dt \\ &- \int_{\Omega} u_2(x, T) v_1(x, T) dx - \int_0^T \int_{\Omega} v_1(x, t) u_{2,x}(x, t) |_{\delta\Omega} dt + \int_0^T \int_{\Omega} u_2(x, t) \left(\frac{\partial f_1}{\partial c} \kappa(x, t) + \frac{\partial f_1}{\partial u_1} v_1 \right) dx dt. \end{aligned}$$

In order to get the necessary conditions, we can group the elements multiplying κ and v_1 , and equate them to zero. If all factors multiplying deviations from optimal values for c and u_1 are equal to zero, then $\frac{\partial V}{\partial \epsilon} = 0$. So we would need:

$$\begin{cases} \kappa : & U'(c)g(x, t) + u_2 \frac{\partial f_1}{\partial c} = 0, \\ v_1 : & u_{2,t} + u_{2,xx} + \frac{\partial f_1}{\partial u_1} u_2 = 0. \end{cases} \tag{21}$$

The first equation in (21) shows that c is an implicit function of u_1 and u_2 . Moreover, we need to add to these conditions the following spatial boundary and transversality conditions:

$$\begin{cases} \lim_{x \rightarrow \delta\Omega} u_{2,x}(x, t) = 0, \\ u_2(x, T) = \Psi'_1(u_1(x, T)). \end{cases}$$

We then obtain the Pontryagin conditions (O) by defining $f_2(u_1, u_2) \equiv -u_2 \frac{\partial f_1}{\partial u_1}$. Finally, taking into account that c can be expressed as an implicit function of u_1 and u_2 , we can rewrite the set of necessary conditions as:

$$\left\{ \begin{array}{l} u_{1,t}(x, t) - u_{1,xx}(x, t) = f_1(u_1, u_2), \\ u_{2,t}(x, t) + u_{2,xx}(x, t) = f_2(u_1, u_2), \\ \lim_{x \rightarrow \delta\Omega} u_{1,x}(x, t) = 0, \\ \lim_{x \rightarrow \delta\Omega} u_{2,x}(x, t) = 0, \\ u_1(x, 0) = u_{1,0}(x), \text{ given,} \\ u_2(x, T) = \Psi'_1(u_1(x, T)). \end{array} \right.$$

For the sake of exposition, we will use this alternative formulation of (O) in Section 3.

B Theorem 1 proof

The final objective of this proof is to demonstrate the contraction property of $\mathbb{L}^{-1}F$, so the sequence

$$\begin{pmatrix} u_1^k \\ u_2^k \end{pmatrix} = \mathbb{L}^{-1}F(u^{k-1})$$

converges to the unique solution $(u_1, u_2)'$. This outcome will require to establish some intermediate results. So let us first summarize our method to build the sequences of the theorem. It mainly consists of four steps, which are developed in detail in the subsequent Proposition B1:

Step 1. We build a sequence of functions $\{\bar{u}_1^k\}_{k \in \mathbb{N}}$ starting from $(\bar{u}_1^0, \bar{u}_2^0) = (\hat{u}_1, \hat{u}_2)$. \bar{u}_1^1 is the solution to

$$\left\{ \begin{array}{l} \mathbb{L}_1 \bar{u}_1^1 = F_1(\bar{u}_1^0, \bar{u}_2^0) = F_1(\hat{u}_1, \hat{u}_2), \quad \forall (x, t) \in \Omega \times [0, T], \\ \lim_{x \rightarrow \delta\Omega} \bar{u}_{1,x}^1(x, t) = 0, \quad \forall t \geq 0, \end{array} \right. \quad (22)$$

with a suitable initial condition for \bar{u}_1^1 verifying that $u_1(x, 0) \leq \bar{u}_1^1(x, 0) \leq \hat{u}_1(x, 0)$. (22) together with the initial condition for \bar{u}_1^1 set up a linear parabolic PDE. Given the structure of the left hand side of the PDE in (22), and since $F_1(\hat{u}_1, \hat{u}_2)$ is a known bounded function of time and space, there exists a unique solution to (22)

from $\bar{u}_1^1(x, 0)$ and subject to the boundary condition, as proven in Theorem 1.7 in Friedman (1959). Furthermore, this solution verifies that

$$\int_0^T \int_{\Omega} |\bar{u}_1^1(x, t)| e^{-\varrho|x|^2} dx dt < \infty,$$

for some positive constant ϱ .

Step 2. We initiate the sequence $\{\bar{w}_2^k\}_{k \in \mathbb{N}}$ as $\bar{w}_2^0(x, t) = \hat{u}_2(x, T - t)$. Using \bar{u}_1^0 as $\bar{w}_1^0(x, t) = \hat{u}_1^1(x, T - t)$, \bar{w}_2^1 is the solution to

$$\begin{cases} \mathbb{L}'_2 \bar{w}_2^1 = -F_2(\bar{w}_1^0, \bar{w}_2^0), \\ \lim_{x \rightarrow \delta\Omega} \bar{w}_2^1(x, t) = 0, \end{cases} \quad (23)$$

and a suitable initial condition for \bar{w}_2^1 verifying $w_2(x, 0) \leq \bar{w}_2^1(x, 0) \leq \hat{w}_2(x, 0)$. As in Step 1, (23) is a linear parabolic PDE, which has a unique positive solution under the boundedness assumption on Ψ and Assumption 1. Furthermore, the solution is bounded as before:

$$\int_0^T \int_{\Omega} |\bar{w}_2^1(x, t)| e^{-\varrho|x|^2} dx dt < \infty,$$

for some positive constant ϱ . We proceed to the next iteration by undoing the change of variable $\bar{u}_2^1(x, t) = \bar{w}_2^1(x, T - t)$. Likewise, building on $(\bar{u}_1^{k-1}, \bar{u}_2^{k-1})$, we compute $(\bar{u}_1^k, \bar{u}_2^k)$ following the procedure just described for $k = 2, 3, \dots$. Moreover, for the subsequent iterations, we also need to state appropriate initial conditions $\bar{u}_1^k(x, 0)$ and $\bar{w}_2^k(x, 0)$ in order to ensure that the initial and terminal conditions in (I) hold as $k \rightarrow \infty$ (notice that the terminal condition for u_2 in (I) is equivalent to the initial condition for w_2 after reversing time).

Step 3. In a similar manner, we can build a sequence $\{\underline{u}_1^k, \underline{u}_2^k\}_{k \in \mathbb{N}}$ and $\{\underline{w}_1^k, \underline{w}_2^k\}_{k \in \mathbb{N}}$ considering the lower solution.

Step 4. We prove that $\{\bar{u}_1^k, \bar{w}_2^k\}_{k \in \mathbb{N}}$ is monotonic non-increasing and that $\{\underline{u}_1^k, \underline{w}_2^k\}_{k \in \mathbb{N}}$ is monotonic non-decreasing. Undoing the change of time variable, we obtain the convergence of sequences $\{\bar{u}_1^k, \bar{u}_2^k\}_{k \in \mathbb{N}}$ and $\{\underline{u}_1^k, \underline{u}_2^k\}_{k \in \mathbb{N}}$ in $\Omega \times [0, T]$. As it will become clear in the proposition, the monotonicity and convergence of the generated sequence hinges on the choice of the initial conditions at each iteration. This choice will depend on the monotonicity characteristics of functions F_i .

We formally enunciate this iterative method as a proposition, together with the monotonicity property of the generated sequences:

Proposition B1. *Let us consider the assumptions of the theorem. Let us introduce variables w as the time-reverse of u , that is, $w_i(x, t) \equiv u_i(x, T-t)$, $\hat{w}_i(x, t) \equiv \hat{u}_i(x, T-t)$ and $\tilde{w}_i(x, t) \equiv \tilde{u}_i(x, T-t)$ for $i = 1, 2$. We build sequences $\{\bar{u}_1^k, \bar{w}_2^k\}_{k \in \mathbb{N}}$ and $\{\underline{u}_1^k, \underline{w}_2^k\}_{k \in \mathbb{N}}$ through the process*

$$\begin{cases} \mathbb{L}_1 \bar{u}_1^k = F_1(\bar{u}_1^{k-1}, \bar{u}_2^{k-1}), \\ \bar{u}_1^k(x, 0) = \frac{\Upsilon}{k} \hat{u}_1(x, 0) + (1 - \frac{\Upsilon}{k}) u_1(x, 0), \\ \lim_{x \rightarrow \delta\Omega} \bar{u}_{1,x}^k = 0, \end{cases}$$

and

$$\begin{cases} \mathbb{L}_1 \underline{u}_1^k = F_1(\underline{u}_1^{k-1}, \underline{u}_2^{k-1}), \\ \underline{u}_1^k(x, 0) = \frac{\zeta}{k} \tilde{u}_1(x, 0) + (1 - \frac{\zeta}{k}) u_1(x, 0), \\ \lim_{x \rightarrow \delta\Omega} \underline{u}_{1,x}^k = 0, \end{cases}$$

with $0 < \zeta < \Upsilon < 1$,

$$\begin{cases} \mathbb{L}'_2 \bar{w}_2^k = -F_2(\bar{w}_1^{k-1}, \bar{w}_2^{k-1}), \\ \bar{w}_2^k(x, 0) = \frac{\psi}{k} \hat{w}_2(x, 0) + (1 - \frac{\psi}{k}) \Psi'_1(\bar{w}_1^{k-1}(x, 0), x), \\ \lim_{x \rightarrow \delta\Omega} \bar{w}_{2,x}^k = 0, \end{cases}$$

and

$$\begin{cases} \mathbb{L}'_2 \underline{w}_2^k = -F_2(\underline{w}_1^{k-1}, \underline{w}_2^{k-1}), \\ \underline{w}_2^k(x, 0) = \frac{\phi}{k} \tilde{w}_2(x, 0) + (1 - \frac{\phi}{k}) \Psi'_1(\underline{w}_1^{k-1}(x, 0), x), \\ \lim_{x \rightarrow \delta\Omega} \underline{w}_{2,x}^k = 0, \end{cases}$$

with $0 < \phi < \psi < 1$. We initialize the sequences as

$$(\bar{u}_1^0, \bar{u}_2^0) = (\hat{u}_1, \hat{u}_2),$$

$$(\bar{w}_1^0, \bar{w}_2^0) = (\hat{w}_1, \hat{w}_2),$$

$$(\underline{u}_1^0, \underline{u}_2^0) = (\tilde{u}_1, \tilde{u}_2),$$

$$(\underline{w}_1^0, \underline{w}_2^0) = (\tilde{w}_1, \tilde{w}_2).$$

Then, the two sequences $\{\bar{u}_1^k, \bar{w}_2^k\}_{k \in \mathbb{N}}$ and $\{\underline{u}_1^k, \underline{w}_2^k\}_{k \in \mathbb{N}}$ obtained above have the monotone property

$$\underline{u}_1^{k-1} \leq \underline{u}_1^k \leq \bar{u}_1^k \leq \bar{u}_1^{k-1},$$

$$\underline{w}_2^{k-1} \leq \underline{w}_2^k \leq \bar{w}_2^k \leq \bar{w}_2^{k-1}.$$

Proof. For the ease of presentation, we divide the proof in five parts. Along this proof it is important to keep in mind that

$$\tilde{u}_1(x, t) \leq \hat{u}_1(x, t),$$

$$\tilde{w}_2(x, t) \leq \hat{w}_2(x, t),$$

and

$$\tilde{u}_1(x, 0) < u_1(x, 0) < \hat{u}_1(x, 0),$$

$$\tilde{w}_2(x, 0) < \Psi'_1(\tilde{u}_1(x, T), x),$$

$$\Psi'_1(\hat{u}_1(x, T), x) < \hat{w}_2(x, 0).$$

i) Let us first construct $\{\bar{u}_1^k\}_{k \in \mathbb{N}}$ and $\{\bar{w}_2^k\}_{k \in \mathbb{N}}$. We initialize the sequences as

$$(\bar{u}_1^0, \bar{w}_2^0) = (\hat{u}_1, \hat{u}_2),$$

$$(\bar{w}_1^0, \bar{w}_2^0) = (\hat{w}_1, \hat{w}_2).$$

As we have observed in the short description of our iterative method, the first element of the sequence $\{\bar{u}_1^k\}_{k \in \mathbb{N}}$ is the solution to

$$\begin{cases} \mathbb{L}_1 \bar{u}_1^1 = F_1(\bar{u}_1^0, \bar{w}_2^0) = F_1(\hat{u}_1, \hat{u}_2), \\ \bar{u}_1^1(x, 0) = \Upsilon \hat{u}_1(x, 0) + (1 - \Upsilon)u_1(x, 0), \\ \lim_{x \rightarrow \delta\Omega} \bar{u}_{1,x}^1 = 0. \end{cases} \quad (24)$$

The PDE in (24) has a unique solution \bar{u}_1^1 given that $F_1(\hat{u}_1, \hat{u}_2)$ is known, \mathbb{L}_1 is linear, and $\bar{u}_1^1(x, 0)$ is strictly positive. Next, we show that $\bar{u}_1^0 \geq \bar{u}_1^1$. Let us construct an auxiliary variable v_1^0 on $\Omega \times [0, T]$:¹⁴

$$v_1^0 \equiv \bar{u}_1^0 - \bar{u}_1^1 = \hat{u}_1 - \bar{u}_1^1.$$

Then,

$$\begin{aligned} \mathbb{L}_1 v_1^0 &= \mathbb{L}_1(\bar{u}_1^0 - \bar{u}_1^1) = \hat{u}_{1,t} - \hat{u}_{1,xx} + \underline{a}_1(x, t)\hat{u}_1 - F_1(\hat{u}_1, \hat{u}_2) \\ &= \hat{u}_{1,t} - \hat{u}_{1,xx} + \underline{a}_1(x, t)\hat{u}_1 - f_1(\hat{u}_1, \hat{u}_2) - \underline{a}_1(x, t)\hat{u}_1 \geq 0. \end{aligned}$$

We know that the expression above is non-negative because (\hat{u}_1, \hat{u}_2) is an upper solution of problem (I). Besides,

$$v_1^0(x, 0) = \bar{u}_1^0(x, 0) - \bar{u}_1^1(x, 0) = \hat{u}_1(x, 0) - \Upsilon \hat{u}_1(x, 0) - (1 - \Upsilon)u_1(x, 0)$$

¹⁴We deliberately abuse of notation in this proof. For $i = 1, 2$, we denote by v_i^0 , v_i^1 and v_i^k different auxiliary variables. Since the usage of these variables is unambiguously restricted to the prove of each intermediate result, we opt for this abuse for the sake of presentation.

$$= (1 - \Upsilon) (\hat{u}_1(x, 0) - u_1(x, 0)) > 0.$$

Then, invoking classical results on positiveness as in Pao (1992), we conclude that $v_1^0(x, t) \geq 0$, which implies that $\bar{u}_1^0(x, t) \geq \bar{u}_1^1(x, t)$ for all $(x, t) \in \Omega \times [0, T]$.

Let us similarly consider \bar{w}_2^1 , solution to

$$\begin{cases} \mathbb{L}'_2 \bar{w}_2^1 = -F_2(\bar{w}_1^0, \bar{w}_2^0) = -F_2(\hat{w}_1, \hat{w}_2), \\ \bar{w}_2^1(x, 0) = \psi \hat{w}_2(x, 0) + (1 - \psi) \Psi'_1(\bar{w}_1^0(x, 0), x), \\ \lim_{x \rightarrow \delta\Omega} \bar{w}_{2,x}^1 = 0, \end{cases}$$

and construct $v_2^0 \equiv \bar{w}_2^0 - \bar{w}_2^1 = \hat{w}_2 - \bar{w}_2^1$, where $\hat{w}_2(x, t) = \hat{u}_2(x, T - t)$ is the time reverse of the second coordinate of the upper solution (\hat{u}_1, \hat{u}_2) . We prove a similar result for v_2^0 :

$$\begin{aligned} \mathbb{L}'_2 v_2^0 &= \mathbb{L}'_2(\hat{w}_2 - \bar{w}_2^1) = \hat{w}_{2,t} - \hat{w}_{2,xx} + \bar{c}_2(x, t) \hat{w}_2 - F_2(\hat{w}_1, \hat{w}_2) \\ &= \hat{w}_{2,t} - \hat{w}_{2,xx} + \bar{c}_2(x, t) \hat{w}_2 + f_2(\hat{w}_1, \hat{w}_2) - \bar{c}_2(x, t) \hat{w}_2 \geq 0, \end{aligned}$$

because of (8). We need to check the sign of $v_2^0(x, 0)$, using the definition of the initial conditions:

$$\begin{aligned} v_2^0(x, 0) &= \hat{w}_2(x, 0) - \bar{w}_2^1(x, 0) = \hat{w}_2(x, 0) - \psi \hat{w}_2(x, 0) - (1 - \psi) \Psi'_1(\bar{w}_1^0(x, 0), x) \\ &= (1 - \psi) (\hat{w}_2(x, 0) - \Psi'_1(\hat{w}_1(x, 0), x)) > 0. \end{aligned}$$

Then $v_2^0(x, t) \geq 0 \Rightarrow \hat{w}_2(x, t) \geq \bar{w}_2^1(x, t)$ for every $(x, t) \in \Omega \times [0, T]$.

ii) We apply the same procedure to the other sequences, starting however with the lower solutions:

$$\begin{aligned} (\underline{u}_1^0, \underline{u}_2^0) &= (\tilde{u}_1, \tilde{u}_2), \\ (\underline{w}_1^0, \underline{w}_2^0) &= (\tilde{w}_1, \tilde{w}_2). \end{aligned}$$

\underline{u}_1^1 is the solution to

$$\begin{cases} \mathbb{L}_1 \underline{u}_1^1 = F_1(\underline{u}_1^0, \underline{u}_2^0) = F_1(\tilde{u}_1, \tilde{u}_2), \\ \underline{u}_1^1(x, 0) = \zeta \tilde{u}_1(x, 0) + (1 - \zeta) u_1(x, 0), \\ \lim_{x \rightarrow \delta\Omega} \underline{u}_{1,x}^1 = 0. \end{cases}$$

Then, defining v_1^0 and v_2^0 as

$$\begin{aligned} v_1^0 &\equiv \underline{u}_1^1 - \underline{u}_1^0 = \underline{u}_1^1 - \tilde{u}_1, \\ v_2^0 &\equiv \underline{w}_2^1 - \underline{w}_2^0 = \underline{w}_2^1 - \tilde{w}_2, \end{aligned}$$

we can prove that $\underline{u}_1^1 \geq \underline{u}_1^0$ and $\underline{w}_2^1 \geq \underline{w}_2^0$. Let us consider in this respect $v_1^0 = \underline{u}_1^1 - \underline{u}_1^0 = \underline{u}_1^1 - \tilde{u}_1$. Applying the operator \mathbb{L}_1 :

$$\begin{aligned}\mathbb{L}_1 v_1^0 &= \mathbb{L}_1(\underline{u}_1^1 - \tilde{u}_1) = F_1(\underline{u}_1^0, \underline{u}_2^0) - (\tilde{u}_{1,t} - \tilde{u}_{1,xx} + \bar{c}_1 \tilde{u}_1) \\ &= f_1(\tilde{u}_1, \tilde{u}_2) + \bar{c}_1 \tilde{u}_1 - \tilde{u}_{1,t} + \tilde{u}_{1,xx} - \bar{c}_1 \tilde{u}_1 \geq 0,\end{aligned}$$

because $(\tilde{u}_1, \tilde{u}_2)$ is a lower solution. Regarding the initial condition:

$$\begin{aligned}v_1^0(x, 0) &= \underline{u}_1^1(x, 0) - \underline{u}_1^0(x, 0) = \underline{u}_1^1(x, 0) - \tilde{u}_1(x, 0) \\ &= \zeta \tilde{u}_1(x, 0) + (1 - \zeta) u_1(x, 0) - \tilde{u}_1(x, 0) = (1 - \zeta)(u_1(x, 0) - \tilde{u}_1(x, 0)) > 0.\end{aligned}$$

So we conclude that $\underline{u}_1^1 \geq \underline{u}_1^0$.

Similarly, let us prove that $\underline{w}_2^0 \leq \underline{w}_2^1$. Taking the definition of v_2^0 , we can show that

$$\begin{aligned}\mathbb{L}'_2 v_2^0 &= \mathbb{L}'_2(\underline{w}_2^1 - \underline{w}_2^0) = -F_2(\underline{u}_1^0, \underline{u}_2^0) - (\underline{w}_{2,t} - \underline{w}_{2,xx} + \bar{c}_2 \underline{w}_2) \\ &= -f_2(\tilde{w}_1, \tilde{w}_2) + \bar{c}_2 \tilde{w}_2 - (\tilde{w}_{2,t} - \tilde{w}_{2,xx} + \bar{c}_2 \tilde{w}_2) \geq 0.\end{aligned}$$

Moreover,

$$\begin{aligned}v_2^0(x, 0) &= \underline{w}_2^1(x, 0) - \underline{w}_2^0(x, 0) \\ &= \underline{w}_2^1(x, 0) - \tilde{w}_2(x, 0) = \phi \tilde{w}_2(x, 0) + (1 - \phi) \Psi'_1(\underline{w}_1^0(x, 0), x) - \tilde{w}_2(x, 0) \\ &= (1 - \phi) (\Psi'_1(\tilde{w}_1(x, 0), x) - \tilde{w}_2(x, 0)) > 0.\end{aligned}$$

We therefore confirm our statement.

iii) We prove next that $\underline{u}_1^1 \leq \bar{u}_1^1$ and $\underline{w}_2^1 \leq \bar{w}_2^1$. To do so, we define the auxiliary variables v_1^1, v_2^1 on $\Omega \times [0, T]$ as

$$v_1^1 \equiv \bar{u}_1^1 - \underline{u}_1^1,$$

$$v_2^1 \equiv \bar{w}_2^1 - \underline{w}_2^1.$$

Let us show first that $\bar{u}_1^1 \leq \underline{u}_1^1$. Applying the operator \mathbb{L}_1 ,

$$\begin{aligned}\mathbb{L}_1 v_1^1 &= \mathbb{L}_1(\bar{u}_1^1 - \underline{u}_1^1) = F_1(\hat{u}_1, \hat{u}_2) - F_1(\tilde{u}_1, \tilde{u}_2) \\ &= f_1(\hat{u}_1, \hat{u}_2) - f_1(\tilde{u}_1, \tilde{u}_2) + \underline{a}_1(\hat{u}_1 - \tilde{u}_1).\end{aligned}$$

Since $\hat{u}_1 \geq \tilde{u}_1$, we can see that $\mathbb{L}_1 v_1^1 \geq 0$ if $\underline{a}_1 \geq 0$ and f_1 increasing in both arguments u_1 and u_2 ; or if alternatively $\underline{a}_2 \leq 0$. In the later case, we would have

$$\mathbb{L}_1 v_1^1 = f_1(\hat{u}_1, \hat{u}_2) - f_1(\tilde{u}_1, \tilde{u}_2) + \underline{a}_1(\hat{u}_1 - \tilde{u}_1) \geq -\underline{a}_2(\hat{u}_2 - \tilde{u}_2) \geq 0.$$

Taking into account that

$$\begin{aligned}
v_1^1(x, 0) &= \bar{u}_1^1(x, 0) - \underline{u}_1^1(x, 0) = \Upsilon \hat{u}_1(x, 0) + (1 - \Upsilon)u_1(x, 0) - (\zeta \tilde{u}_1(x, 0) + (1 - \zeta)u_1(x, 0)) \\
&= \Upsilon \hat{u}_1(x, 0) - \zeta \tilde{u}_1(x, 0) - (\Upsilon - \zeta)u_1(x, 0) \\
&\geq \Upsilon \hat{u}_1(x, 0) - \zeta \hat{u}_1(x, 0) - (\Upsilon - \zeta)u_1(x, 0) = (\Upsilon - \zeta) (\hat{u}_1(x, 0) - u_1(x, 0)) > 0,
\end{aligned}$$

we conclude that $v_1^1(x, t) \geq 0$, $\forall (x, t) \in \Omega \times [0, T]$. Hence, $\underline{u}_1^1 \leq \bar{u}_1^1$.

iv) Let us consider v_2^1 . Applying \mathbb{L}'_2 ,

$$\begin{aligned}
\mathbb{L}'_2 v_2^1 &= \mathbb{L}'_2(\bar{w}_2^1 - \underline{w}_2^1) = -F_2(\hat{w}_1, \hat{w}_2) + F_2(\tilde{w}_1, \tilde{w}_2) \\
&= -f_2(\hat{w}_1, \hat{w}_2) + f_2(\tilde{w}_1, \tilde{w}_2) + \bar{c}_2(\hat{w}_2 - \tilde{w}_2),
\end{aligned}$$

which can be positive if either f_2 decreasing in both w_1 and w_2 and $\bar{c}_2 \geq 0$; or if $\bar{c}_1 \leq 0$ whatever the sign of the partial derivatives of f_2 since

$$\mathbb{L}'_2 v_2^1 \geq -\bar{c}_1(\hat{w}_1 - \tilde{w}_1) \geq 0.$$

Moreover, one can verify that

$$\begin{aligned}
v_2^1(x, 0) &= \bar{w}_2^1(x, 0) - \underline{w}_2^1(x, 0) \\
&= \psi \hat{w}_2(x, 0) + (1 - \psi)\Psi'_1(\bar{w}_1^0(x, 0), x) - \phi \tilde{w}_2(x, 0) - (1 - \phi)\Psi'_1(\underline{w}_1^0(x, 0), x) \\
&\geq (\psi - \phi) (\hat{w}_2(x, 0) - \Psi'_1(\hat{w}_1(x, 0), x)) > 0.
\end{aligned}$$

We therefore conclude that $v_2^1 \geq 0$ and $\bar{w}_2^1 \geq \underline{w}_2^1$.

v) We have proven so far that

$$\underline{u}_1^0 \leq \underline{u}_1^1 \leq \bar{u}_1^1 \leq \bar{u}_1^0,$$

and

$$\underline{w}_2^0 \leq \underline{w}_2^1 \leq \bar{w}_2^1 \leq \bar{w}_2^0.$$

These two conclusions allow us to identify the required base case to show by mathematical induction that our sequences are ordered. Following this strategy, let us assume that

$$\underline{u}_1^{k-1} \leq \underline{u}_1^k \leq \bar{u}_1^k \leq \bar{u}_1^{k-1},$$

and

$$\underline{w}_2^{k-1} \leq \underline{w}_2^k \leq \bar{w}_2^k \leq \bar{w}_2^{k-1}.$$

We have to show, as inductive step, that

$$\underline{u}_1^k \leq \underline{u}_1^{k+1} \leq \bar{u}_1^{k+1} \leq \bar{u}_1^k,$$

and

$$\underline{w}_2^k \leq \underline{w}_2^{k+1} \leq \bar{w}_2^{k+1} \leq \bar{w}_2^k.$$

Let us begin by verifying that the inequality $\bar{u}_1^{k+1} \leq \bar{u}_1^k$ holds. We define in this regard the auxiliary variable v_1^k as

$$v_1^k \equiv \bar{u}_1^k - \bar{u}_1^{k+1}.$$

Applying the operator \mathbb{L}_1 ,

$$\begin{aligned} \mathbb{L}_1 v_1^k &= F_1(\bar{u}_1^{k-1}, \bar{u}_2^{k-1}) - F_1(\bar{u}_1^k, \bar{u}_2^k) \\ &= f_1(\bar{u}_1^{k-1}, \bar{u}_2^{k-1}) - f_1(\bar{u}_1^k, \bar{u}_2^k) + \underline{a}_1 (\bar{u}_1^{k-1} - \bar{u}_1^k) \geq 0. \end{aligned}$$

Since $\bar{u}_1^k \leq \bar{u}_1^{k-1}$ and $\bar{u}_2^k \leq \bar{u}_2^{k-1}$, we can see that $\mathbb{L}_1 v_1^k \geq 0$ if $\underline{a}_1 \geq 0$ and f_1 increasing in both arguments; or if alternatively $\underline{a}_1 \leq 0$. In the later case,

$$\mathbb{L}_1 v_1^k = f_1(\bar{u}_1^{k-1}, \bar{u}_2^{k-1}) - f_1(\bar{u}_1^k, \bar{u}_2^k) + \underline{a}_1 (\bar{u}_1^{k-1} - \bar{u}_1^k) \geq -\underline{a}_2 (\bar{u}_2^{k-1} - \bar{u}_2^k) \geq 0.$$

Moreover,

$$\begin{aligned} v_1^k(x, 0) &= \bar{u}_1^k(x, 0) - \bar{u}_1^{k+1}(x, 0) \\ &= \frac{\Upsilon}{k} \hat{u}_1(x, 0) + \left(1 - \frac{\Upsilon}{k}\right) u_1(x, 0) - \frac{\Upsilon}{k+1} \hat{u}_1(x, 0) - \left(1 - \frac{\Upsilon}{k+1}\right) u_1(x, 0) \\ &= \Upsilon \hat{u}_1(x, 0) \left(\frac{1}{k} - \frac{1}{k+1}\right) - \Upsilon u_1(x, 0) \left(\frac{1}{k} - \frac{1}{k+1}\right), \\ &= \Upsilon \left(\frac{1}{k} - \frac{1}{k+1}\right) (\hat{u}_1(x, 0) - u_1(x, 0)) > 0, \end{aligned}$$

since $\hat{u}_1(x, 0) > u_1(x, 0)$. Hence $v_1^k \geq 0$ and $\bar{u}_1^k \geq \bar{u}_1^{k+1}$.

Similarly, let us prove the inequality $\underline{u}_1^k \leq \underline{u}_1^{k+1}$. We define the auxiliary variable v_1^k as

$$v_1^k \equiv \underline{u}_1^{k+1} - \underline{u}_1^k.$$

Applying again \mathbb{L}_1 ,

$$\begin{aligned} \mathbb{L}_1 v_1^k &= F_1(\underline{u}_1^k, \underline{u}_2^k) - F_1(\underline{u}_1^{k-1}, \underline{u}_2^{k-1}) \\ &= f_1(\underline{u}_1^k, \underline{u}_2^k) - f_1(\underline{u}_1^{k-1}, \underline{u}_2^{k-1}) + \underline{a}_1 (\underline{u}_1^k - \underline{u}_1^{k-1}) \geq 0. \end{aligned}$$

Since $\underline{u}_1^k \geq \underline{u}_1^{k-1}$ and $\underline{u}_2^k \geq \underline{u}_2^{k-1}$, we can see that $\mathbb{L}_1 v_1^1 \geq 0$ if $\underline{a}_1 \geq 0$ and f_1 increasing in both arguments; or if alternatively $\underline{a}_2 \leq 0$. In the later case, we would have

$$\mathbb{L}_1 v_1^k = f_1(\underline{u}_1^k, \underline{u}_2^k) - f_1(\underline{u}_1^{k-1}, \underline{u}_2^{k-1}) + \underline{a}_1(\underline{u}_1^k - \underline{u}_1^{k-1}) \geq -\underline{a}_2(\underline{u}_2^k - \underline{u}_2^{k-1}) \geq 0.$$

One can also verify that

$$\begin{aligned} v_1^k(x, 0) &= \underline{u}_1^{k+1}(x, 0) - \underline{u}_1^k(x, 0) \\ &= \frac{\zeta}{k+1} \tilde{u}_1(x, 0) + \left(1 - \frac{\zeta}{k+1}\right) u_1(x, 0) - \frac{\zeta}{k} \tilde{u}_1(x, 0) + \left(1 - \frac{\zeta}{k}\right) u_1(x, 0) \\ &= \zeta \frac{1}{k(k+1)} \tilde{u}_1(x, 0) + \zeta \frac{1}{k(k+1)} u_1(x, 0) \\ &= \zeta \frac{1}{k(k+1)} (u_1(x, 0) - \tilde{u}_1(x, 0)) > 0. \end{aligned}$$

Therefore $v_1^k \geq 0$ and $\underline{u}_1^k \leq \underline{u}_1^{k+1}$. Next, we need to prove that $\underline{u}_1^{k+1} \leq \bar{u}_1^{k+1}$. Let us define $v_1^{k+1} \equiv \bar{u}_1^{k+1} - \underline{u}_1^{k+1}$, then

$$\mathbb{L}_1 v_1^{k+1} = F_1(\bar{u}_1^k, \bar{u}_2^k) - F_1(\underline{u}_1^k, \underline{u}_2^k) = f_1(\bar{u}_1^k, \bar{u}_2^k) - f_1(\underline{u}_1^k, \underline{u}_2^k) + \underline{a}_1(\bar{u}_1^k - \underline{u}_1^k) \geq 0.$$

Since $\bar{u}_1^k \geq \underline{u}_1^k$ and $\bar{u}_2^k \geq \underline{u}_2^k$, we can see that $\mathbb{L}_1 v_1^{k+1} \geq 0$ if $\underline{a}_1 \geq 0$ and f_1 increasing in both arguments; or if alternatively $\underline{a}_2 \leq 0$. In the later case,

$$\mathbb{L}_1 v_1^{k+1} = f_1(\bar{u}_1^k, \bar{u}_2^k) - f_1(\underline{u}_1^k, \underline{u}_2^k) + \underline{a}_1(\bar{u}_1^k - \underline{u}_1^k) \geq -\underline{a}_2(\bar{u}_2^k - \underline{u}_2^k) \geq 0.$$

We can verify that the initial distribution of v_1^{k+1} is positive:

$$\begin{aligned} v_1^{k+1}(x, 0) &= \bar{u}_1^{k+1}(x, 0) - \underline{u}_1^{k+1}(x, 0) \\ &= \frac{\Upsilon}{k+1} \hat{u}_1(x, 0) + \left(1 - \frac{\Upsilon}{k+1}\right) u_1(x, 0) - \frac{\zeta}{k+1} \tilde{u}_1(x, 0) - \left(1 - \frac{\zeta}{k+1}\right) u_1(x, 0) \\ &= \frac{\Upsilon}{k+1} \hat{u}_1(x, 0) - \frac{\zeta}{k+1} \tilde{u}_1(x, 0) - \frac{\Upsilon - \zeta}{k+1} u_1(x, 0) \geq \frac{\Upsilon - \zeta}{k+1} \hat{u}_1(x, 0) - \frac{\Upsilon - \zeta}{k+1} u_1(x, 0) \\ &= \frac{\Upsilon - \zeta}{k+1} (\hat{u}_1(x, 0) - u_1(x, 0)) > 0, \end{aligned}$$

Hence $v_1^{k+1} \geq 0$ and $\underline{u}_1^{k+1} \leq \bar{u}_1^{k+1}$. Summarizing, we can conclude that

$$\underline{u}_1^k \leq \underline{u}_1^{k+1} \leq \bar{u}_1^{k+1} \leq \bar{u}_1^k.$$

Let us consider the second set of inequalities. Defining $v_2^k \equiv \bar{w}_2^k - \bar{w}_2^{k+1}$, we apply the operator \mathbb{L}'_2 :

$$\mathbb{L}'_2 v_2^k = F_2(\bar{w}_1^k, \bar{w}_2^k) - F_2(\bar{w}_1^{k-1}, \bar{w}_2^{k-1}) = f_2(\bar{w}_1^k, \bar{w}_2^k) - f_2(\bar{w}_1^{k-1}, \bar{w}_2^{k-1}) + \bar{c}_2(\bar{w}_2^{k-1} - \bar{w}_2^k) \geq 0.$$

Since $\bar{w}_1^k \leq \bar{w}_1^{k-1}$ and $\bar{w}_2^k \leq \bar{w}_2^{k-1}$, we can see that $\mathbb{L}'_2 v_2^k \geq 0$ if $\bar{c}_2 \geq 0$ and f_2 decreasing in both arguments; or if alternatively $\bar{c}_1 \leq 0$. In the later case,

$$\begin{aligned}\mathbb{L}'_2 v_2^k &= f_2(\bar{w}_1^k, \bar{w}_2^k) - f_2(\bar{w}_1^{k-1}, \bar{w}_2^{k-1}) + \bar{c}_2(\bar{w}_2^{k-1} - \bar{w}_2^k) \\ &\geq -\bar{c}_1(\bar{w}_1^{k-1} - \bar{w}_1^k) \geq 0.\end{aligned}$$

Moreover,

$$\begin{aligned}v_2^k(x, 0) &= \bar{w}_2^k(x, 0) - \bar{w}_2^{k+1}(x, 0) \\ &= \psi \left(\frac{1}{k} - \frac{1}{k+1} \right) \hat{w}_2(x, 0) + \left(1 - \frac{\psi}{k} \right) \Psi'_1(\bar{w}_1^{k-1}(x, 0), x) - \left(1 - \frac{\psi}{k+1} \right) \Psi'_1(\bar{w}_1^k(x, 0), x).\end{aligned}$$

Since $\bar{w}_1^{k+1}(x, 0) \leq \bar{w}_1^k(x, 0)$, for a convex function Ψ ,

$$\begin{aligned}w_2^k(x, 0) &\geq \psi \left(\frac{1}{k} - \frac{1}{k+1} \right) \hat{w}_2(x, 0) + \left(1 - \frac{\psi}{k} \right) \Psi'_1(\bar{w}_1^{k-1}(x, 0), x) - \left(1 - \frac{\psi}{k+1} \right) \Psi'_1(\bar{w}_1^{k-1}(x, 0), x) \\ &= \psi \left(\frac{1}{k} - \frac{1}{k+1} \right) [\hat{w}_2(x, 0) - \Psi'_1(\bar{w}_1^{k-1}(x, 0), x)] \\ &> \psi \left(\frac{1}{k} - \frac{1}{k+1} \right) [\hat{w}_2(x, 0) - \Psi'_1(\hat{w}_1(x, 0), x)] > 0.\end{aligned}$$

Hence, we conclude that $v_2^k(x, 0) > 0$ and $\bar{w}_2^{k+1} \leq \bar{w}_2^k$.

We prove next that $\underline{w}_2^k \leq \underline{w}_2^{k+1}$. Let us define $v_2^k \equiv \underline{w}_2^{k+1} - \underline{w}_2^k$. Considering the operator \mathbb{L}'_2 we obtain

$$\begin{aligned}\mathbb{L}'_2 v_2^k &= F_2(\underline{w}_1^{k-1}, \underline{w}_2^{k-1}) - F_2(\underline{w}_1^k, \underline{w}_2^k) \\ &= f_2(\underline{w}_1^{k-1}, \underline{w}_2^{k-1}) - f_2(\underline{w}_1^k, \underline{w}_2^k) + \bar{c}_2(\underline{w}_2^k - \underline{w}_2^{k-1}) \geq 0.\end{aligned}$$

Since $\underline{w}_1^{k-1} \leq \underline{w}_1^k$ and $\underline{w}_2^{k-1} \leq \underline{w}_2^k$, we can see that $\mathbb{L}'_2 v_2^k \geq 0$ if $\bar{c}_2 \geq 0$ and f_2 decreasing in both arguments; or if alternatively $\bar{c}_1 \leq 0$. In the later case,

$$\begin{aligned}\mathbb{L}'_2 v_2^k &= f_2(\underline{w}_1^{k-1}, \underline{w}_2^{k-1}) - f_2(\underline{w}_1^k, \underline{w}_2^k) + \bar{c}_2(\underline{w}_2^k - \underline{w}_2^{k-1}) \\ &\geq -\bar{c}_1(\underline{w}_1^k - \underline{w}_1^{k-1}) \geq 0.\end{aligned}$$

We check the initial distribution of v_2^k , for a convex function Ψ :

$$\begin{aligned}v_2^k(x, 0) &= \underline{w}_2^{k+1}(x, 0) - \underline{w}_2^k(x, 0) \\ &= \frac{\phi}{k+1} \tilde{w}_2(x, 0) + \left(1 - \frac{\phi}{k+1} \right) \Psi'_1(\underline{w}_1^k(x, 0), x) - \frac{\phi}{k} \tilde{w}_2(x, 0) - \left(1 - \frac{\phi}{k} \right) \Psi'_1(\underline{w}_1^{k-1}(x, 0), x)\end{aligned}$$

$$\geq \phi \frac{1}{k(k+1)} (\Psi'_1(\underline{w}_1^k(x,0), x) - \tilde{w}_2(x,0)) > \phi \frac{1}{k(k+1)} (\Psi'_1(\tilde{w}_1(x,0), x) - \tilde{w}_2(x,0)) > 0,$$

Therefore $v_2^k(x,0) > 0$ and $\underline{w}_2^k \leq \underline{w}_2^{k+1}$.

Let us finally show that $\underline{w}_2^{k+1} \leq \bar{w}_2^{k+1}$. This time we define $w_2^{k+1} \equiv \bar{w}_2^{k+1} - \underline{w}_2^{k+1}$.

Applying \mathbb{L}'_2 ,

$$\begin{aligned} \mathbb{L}'_2 v_2^{k+1} &= -F_2(\bar{w}_1^k, \bar{w}_2^k) + F_2(\underline{w}_1^k, \underline{w}_2^k) \\ &= -f_2(\bar{w}_1^k, \bar{w}_2^k) + f_2(\underline{w}_1^k, \underline{w}_2^k) + \bar{c}_2(\bar{w}_2^k - \underline{w}_2^k) \geq -\bar{c}_1(\bar{w}_1^k - \underline{w}_1^k) \geq 0. \end{aligned}$$

Since $\bar{w}_2^k \geq \underline{w}_2^k$ and $\bar{w}_1^k \geq \underline{w}_1^k$, we can see that $\mathbb{L}'_2 v_2^{k+1} \geq 0$ if $\bar{c}_2 \geq 0$ and f_2 decreasing in both arguments; or if alternatively $\bar{c}_1 \leq 0$. In the later case,

$$\mathbb{L}'_2 v_2^{k+1} = -f_2(\bar{w}_1^k, \bar{w}_2^k) + f_2(\underline{w}_1^k, \underline{w}_2^k) + \bar{c}_2(\bar{w}_2^k - \underline{w}_2^k) \geq -\bar{c}_1(\bar{w}_1^k - \underline{w}_1^k) \geq 0.$$

Regarding the initial condition $v_2^{k+1}(x,0)$:

$$\begin{aligned} w_2^{k+1}(x,0) &= \\ &= \frac{\psi}{k+1} \hat{w}_2(x,0) + \left(1 - \frac{\psi}{k+1}\right) \Psi'_1(\bar{w}_1^k(x,0), x) - \frac{\phi}{k+1} \tilde{w}_2(x,0) - \left(1 - \frac{\phi}{k+1}\right) \Psi'_1(\underline{w}_1^k(x,0), x) \\ &\geq \frac{\psi - \phi}{k+1} \hat{w}_2(x,0) - \frac{\psi - \phi}{k+1} \Psi'_1(\bar{w}_1^k(x,0), x) = \frac{\psi - \phi}{k} (\hat{w}_2(x,0) - \Psi'_1(\bar{w}_1^k(x,0), x)) \\ &\geq \frac{\psi - \phi}{k} (\hat{w}_2(x,0) - \Psi'_1(\hat{w}_1(x,0), x)) > 0, \end{aligned}$$

for a convex function Ψ . This implies that $v_2^{k+1} \geq 0$ and hence $\underline{w}_2^{k+1} \leq \bar{w}_2^{k+1}$. So we can state that

$$\underline{w}_2^k \leq \underline{w}_2^{k+1} \leq \bar{w}_2^{k+1} \leq \bar{w}_2^k.$$

□

Before going to the next intermediate result, it is important to consider the following observations regarding the generality of Proposition B1:

Remark 1. *The initial conditions for \bar{w}_2^k and \underline{w}_2^k correspond to a function Ψ that is assumed to be convex in w_1 . For Ψ_1 concave we should initialize \bar{w}_2^k and \underline{w}_2^k as*

$$\begin{aligned} \bar{w}_2^k(x,0) &= \frac{\psi}{k} \hat{w}_2(x,0) + \left(1 - \frac{\psi}{k}\right) \Psi'_1(\underline{w}_1^{k-1}(x,0), x), \\ \underline{w}_2^k(x,0) &= \frac{\phi}{k} \tilde{w}_2(x,0) + \left(1 - \frac{\phi}{k}\right) \Psi'_1(\bar{w}_1^{k-1}(x,0), x). \end{aligned}$$

Remark 2. *In the absence of an upper solution, it is still true that $\{\underline{u}_1^k, \underline{u}_2^k\}$ is monotone non-decreasing. Then the sequence either converges to some limit or it becomes unbounded, as $k \rightarrow \infty$. Nevertheless, we cannot prove that the limit is a solution of the original problem. The same holds true for the monotone non-increasing sequence $\{\bar{u}_1^k, \bar{u}_2^k\}$.*

Remark 3. *The initial conditions for our sequences have been set in order to ensure that the iterative process is monotone. Alternative forms might be considered. The main requirement is that they should converge towards the initial and terminal conditions of the original problem when $k \rightarrow \infty$.*

This far we have constructed four sequences of continuous functions $\{\bar{u}_1^k\}_{k \in \mathbb{N}}$, $\{\underline{u}_1^k\}_{k \in \mathbb{N}}$, $\{\bar{w}_2^k\}_{k \in \mathbb{N}}$ and $\{\underline{w}_2^k\}_{k \in \mathbb{N}}$, which are monotone and bounded. Hence, one can conclude that they converge uniformly. So there exist functions \underline{u}_1 , \bar{u}_1 , \underline{w}_2 and \bar{w}_2 defined over $\Omega \times [0, T]$, such that, $\forall (x, t) \in \Omega \times [0, T]$,

$$\lim_{\substack{k \rightarrow \infty \\ k \in \mathbb{N}}} \underline{u}_1^k(x, t) = \underline{u}_1(x, t),$$

$$\lim_{\substack{k \rightarrow \infty \\ k \in \mathbb{N}}} \bar{u}_1^k(x, t) = \bar{u}_1(x, t),$$

and

$$\lim_{\substack{k \rightarrow \infty \\ k \in \mathbb{N}}} \underline{w}_2^k(x, t) = \underline{w}_2(x, t),$$

$$\lim_{\substack{k \rightarrow \infty \\ k \in \mathbb{N}}} \bar{w}_2^k(x, t) = \bar{w}_2(x, t).$$

Notice that $\underline{u}_1(x, 0) = \bar{u}_1(x, 0) = u_1(x, 0)$ and $\underline{w}_2(x, 0) = \bar{w}_2(x, 0) = \Psi'_1(u_1(x, T), x)$ (see Remark 3 above). Undoing the change of variable, we also obtain limit functions $\bar{u}_2(x, t) = \bar{w}_2(x, T - t)$ and $\underline{u}_2(x, t) = \underline{w}_2(x, T - t)$. The next question is whether the limit functions $\underline{u}_1(x, t)$, $\bar{u}_1(x, t)$, $\underline{u}_2(x, t)$ and $\bar{u}_2(x, t)$ coincide, *i.e.*, whether for all $(x, t) \in \Omega \times [0, T]$ it is true that $\underline{u}_1(x, t) = \bar{u}_1(x, t)$ and $\underline{u}_2(x, t) = \bar{u}_2(x, t)$. At this point, we directly focus on u_1 and u_2 . In order to prove that $\underline{u}_1 = \bar{u}_1$ and $\underline{u}_2 = \bar{u}_2$, we will show that the linear operators \mathbb{L}_1 and \mathbb{L}_2 , and the corresponding diagonal matrix $\mathbb{L} \equiv \text{diag}(\mathbb{L}_1, \mathbb{L}_2)$, are invertible. Then, after proving that $\mathbb{L}^{-1}F(u)$ is a contraction where $F(u) \equiv (F_1(u), F_2(u))'$, we can conclude the uniqueness of the limits. Let us first consider the existence of \mathbb{L}^{-1} :

Lemma B1. *There exists a real positive number γ such that the inverse operators of \mathbb{L}_1 and \mathbb{L}_2 , denoted by \mathbb{L}_1^{-1} and \mathbb{L}_2^{-1} respectively, exist and satisfy*

$$|\mathbb{L}_i^{-1}u_i - \mathbb{L}_i^{-1}u'_i|_0 \leq \gamma^{-1} |u_i - u'_i|_0,$$

for any $(x, t) \in \Omega \times [0, T]$, where $|\cdot|_0$ is the supremum norm for functions in $R(\mathbb{L}_i)$, the range of \mathbb{L}_i :

$$\mathbb{L}_i : D(\mathbb{L}_i) \longrightarrow C(\Omega \times [0, T]),$$

for $i = 1, 2$, where \mathbb{L}_i 's domain is

$$\begin{aligned} D(\mathbb{L}_1) &= \{u_1 \in C(\bar{\Omega} \times [0, T]) \cap C^{1,2}(\Omega \times [0, T]), \lim_{x \rightarrow \delta\Omega} u_{1,x} = 0, u_1(x, 0) = u_{1,0}(x), \text{ given}\}, \\ D(\mathbb{L}_2) &= \{u_2 \in C(\bar{\Omega} \times [0, T]) \cap C^{1,2}(\Omega \times [0, T]), \lim_{x \rightarrow \delta\Omega} u_{2,x} = 0, u_2(x, T) = \Psi'_1(u_1(x, T), x)\}. \end{aligned}$$

Proof. We will treat \mathbb{L}_1 and \mathbb{L}_2 separately. We begin by proving lemma's result for \mathbb{L}_1 .

I. Let us take any two $u_1, u'_1 \in D(\mathbb{L}_1)$ verifying $u_1 \neq u'_1$, *i.e.*, they do not coincide in all $(x, t) \in \Omega \times [0, T]$. We build $v \equiv u_1 - u'_1$, and let $(x_1, t_1) \in \Omega \times [0, T]$ be a point such that

$$|v|_0 = |v(x_1, t_1)|.$$

That is to say, v attains the supremum of its absolute value at (x_1, t_1) . Let us show that

$$v(x_1, t_1) (\mathbb{L}_1 u_1(x_1, t_1) - \mathbb{L}_1 u'_1(x_1, t_1)) \geq \gamma_1 |v|_0^2,$$

which would imply, following Pao (1992), the statement of the lemma for the operator \mathbb{L}_1 . The inequality above is obvious if $v(x_1, t_1) = 0$. If $v(x_1, t_1) \neq 0$ then $v(x_1, t_1)$ is a positive maximum or a negative minimum in $\bar{\Omega} \times [0, T]$.

We know that, for $t = 0$ and for any $x \in \Omega$, $v(x, 0) = u_{1,0}(x) - u_{1,0}(x) = 0$. Therefore $t_1 > 0$. Given that t_1 is an extremum,

$$\begin{cases} v_t(x_1, t_1) = 0, & \text{if } t_1 < T, \\ \text{or } v(x_1, t_1)v_t(x_1, t_1) \geq 0, & \text{if } t_1 = T. \end{cases} \quad (25)$$

Regarding the space we also have two alternatives. Rather x_1 is in the interior of the space or it is located in the frontier, *i.e.*, $x_1 \in \overset{\circ}{\Omega}$ or $x_1 \in \delta\Omega$, respectively:

I.1. If $x_1 \in \overset{\circ}{\Omega}$, we have that

$$v(x_1, t_1)v_{xx}(x_1, t_1) \leq 0 \quad (26)$$

because (x_1, t_1) is a extremum. Considering together (25) and (26),

$$v(x_1, t_1) [v_t(x_1, t_1) - v_{xx}(x_1, t_1)] \geq 0$$

or, equivalently, rewriting the inequality above in terms of u_1 and taking into account the definition of \mathbb{L}_1 ,

$$v(x_1, t_1) \{ \mathbb{L}_1 u_1(x_1, t_1) - \mathbb{L}_1 u'_1(x_1, t_1) - \underline{a}_1(x_1, t_1) [u_1(x_1, t_1) - u'_1(x_1, t_1)] \} \geq 0.$$

Therefore, since $\underline{a}_1(x, t)$ is a bounded function (assumption of the theorem),

$$\begin{aligned} v(x_1, t_1) [\mathbb{L}_1 u_1(x_1, t_1) - \mathbb{L}_1 u_1'(x_1, t_1)] &\geq \underline{a}_1(x_1, t_1) |u_1(x_1, t_1) - u_1'(x_1, t_1)|^2 \\ &\geq \gamma_1 |v(x_1, t_1)|^2 = \gamma_1 |v|_0^2 \end{aligned} \quad (27)$$

for any $\gamma_1 \in \mathbb{R}$ such that $0 < \gamma_1 \leq \underline{a}_1(x_1, t_1)$. Hence, as in Pao (1992), since

$$|v|_0 |\mathbb{L}_1 u_1 - \mathbb{L}_1 u_1'|_0 \geq v(x_1, t_1) [\mathbb{L}_1 u_1(x_1, t_1) - \mathbb{L}_1 u_1'(x_1, t_1)], \quad (28)$$

we can conclude that, using (27) and (28),

$$|\mathbb{L}_1 u_1 - \mathbb{L}_1 u_1'|_0 \geq \gamma_1 |v|_0. \quad (29)$$

I.2. Let us consider the case where x_1 is in the frontier. If $x_1 \in \delta\Omega$ then, for any $\epsilon > 0$, there exists a point $(x_\epsilon, t_\epsilon) \in \overset{\circ}{\Omega} \times [0, T]$, sufficiently close to (x_1, t_1) , such that

$$\begin{cases} v(x_1, t_1)v_t(x_\epsilon, t_\epsilon) \geq \frac{-\epsilon}{2} |v|_0^2, \\ v(x_1, t_1)v_{xx}(x_\epsilon, t_\epsilon) \leq \frac{-\epsilon}{2} |v|_0^2, \end{cases}$$

where $t_\epsilon \equiv t_1 \pm \epsilon$ and $x_\epsilon \equiv x_1 \pm \epsilon$. We can state the first inequality because (x_1, t_1) is a maximum of the absolute value. Hence, when we approach to (x_1, t_1) , we must observe that $v_t \rightarrow 0$ by continuity of the trajectories. The same is true for the derivative in space. As a consequence,

$$v(x_1, t_1) [v_t(x_\epsilon, t_\epsilon) - v_{xx}(x_\epsilon, t_\epsilon)] \geq -\epsilon |v|_0^2$$

and, as before,

$$v(x_1, t_1) (\mathbb{L}_1 u_1 - \mathbb{L}_1 u_1')|_{(x_\epsilon, t_\epsilon)} \geq (\gamma_1 - \epsilon) |v|_0^2.$$

The arbitrariness of ϵ therefore ensures that

$$|\mathbb{L}_1 u_1(x_1, t_1) - \mathbb{L}_1 u_1'(x_1, t_1)| \geq \gamma_1 |u_1(x_1, t_1) - u_1'(x_1, t_1)| = \gamma_1 |u_1 - u_1'|_0$$

for $x \in \delta\Omega$ as well. Hence, since

$$|\mathbb{L}_1 u_1 - \mathbb{L}_1 u_1'|_0 \geq |u_1(x_1, t_1) - \mathbb{L}_1 u_1'(x_1, t_1)|,$$

\mathbb{L}_1^{-1} exists on $R(\mathbb{L}_1)$ and satisfies

$$|\mathbb{L}_1^{-1} u_1 - \mathbb{L}_1^{-1} u_1'|_0 \leq \gamma_1^{-1} |u_1 - u_1'|_0.$$

II. We prove next that \mathbb{L}_2^{-1} exists and there is a $\gamma_2 > 0$ such that

$$|\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq \gamma_2 |u_2 - u'_2|_0, \quad (30)$$

for any two $u_2, u'_2 \in D(\mathbb{L}_2)$. We define $v \equiv u_2 - u'_2$, and let (x_2, t_2) be a point in $\Omega \times [0, T]$ such that

$$|v|_0 = |v(x_2, t_2)|.$$

It is important to note that $v(x, T) = 0$ for any $x \in \Omega$, so that $t_2 < T$.

Let us show that

$$v(x_2, t_2) (\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2) \geq \gamma_2 |v|_0^2.$$

If $v(x_2, t_2) = 0$ then the inequality above is trivially verified. If $v(x_2, t_2) \neq 0$, $v(x_2, t_2)$ is rather a positive maximum or a negative minimum in $\Omega \times [0, T]$. Hence,

$$v_t(x_2, t_2) = 0$$

whenever $t_2 < T$. As in point I, x_2 can lie either in $\mathring{\Omega}$ or in $\delta\Omega$:

II.1 If $x_2 \in \mathring{\Omega}$ then

$$v_x(x_2, t_2) = 0$$

and $v(x_2, t_2)v_{xx}(x_2, t_2) \leq 0$. We can then write:

$$-v(x_2, t_2) [v_t(x_2, t_2) + v_{xx}(x_2, t_2)] \geq 0,$$

$$-v(x_2, t_2) [\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2 + \bar{c}_2(u_2 - u'_2)] \geq 0,$$

$$-v(x_2, t_2) [\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2] \geq \bar{c}_2 v(x_2, t_2)(u_2 - u'_2) \geq \gamma_2 |v|_0^2,$$

for any $0 < \gamma_2 \leq \bar{c}_2$. On the other hand,

$$|-v|_0 |\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq -v(x_2, t_2) (\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2)$$

and $|-v|_0 = |v|_0$. Therefore,

$$|\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq \gamma_2 |u_2 - u'_2|_0.$$

Again, as in Pao (1992), this implies that \mathbb{L}_2^{-1} exists and satisfies

$$|\mathbb{L}_2^{-1} u_2 - \mathbb{L}_2^{-1} u'_2|_0 \leq \gamma_2^{-1} |u_2 - u'_2|_0.$$

II.2 If $x_2 \in \delta\Omega$ then, since $\lim_{x \rightarrow \delta\Omega} v_t(x, t) = 0$ for all t , there exists $(x_\epsilon, t_\epsilon) \in \mathring{\Omega} \times [0, T)$ and ϵ sufficiently small such that

$$\begin{cases} v(x_2, t_2)v_t(x_\epsilon, t_\epsilon) \leq \frac{\epsilon}{2} |v|_0^2, \\ v(x_2, t_2)v_{xx}(x_\epsilon, t_\epsilon) \leq \frac{\epsilon}{2} |v|_0^2. \end{cases}$$

In particular, we will restrict ϵ to $\epsilon \leq \bar{c}_2(x_2, t_2)$. Therefore,

$$-v(x_2, t_2) [v_t(x_\epsilon, t_\epsilon) + v_{xx}(x_\epsilon, t_\epsilon)] \geq -\epsilon |v|_0^2,$$

that is,

$$-v(x_2, t_2) [\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2 + \bar{c}_2(u_2 - u'_2)] |_{(x_\epsilon, t_\epsilon)} \geq -\epsilon |v|_0^2. \quad (31)$$

It is also true that

$$|v(x_2, t_2)|_0 |\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq -v(x_2, t_2) [\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2] |_{(x_\epsilon, t_\epsilon)}. \quad (32)$$

Using (31) and (32) we obtain

$$|v(x_2, t_2)|_0 |\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq -\epsilon |v|_0^2 + \bar{c}_2(x_\epsilon, t_\epsilon) |v(x_2, t_2)|_0 (u_2 - u'_2) |_{(x_\epsilon, t_\epsilon)}.$$

Therefore,

$$|\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq -\epsilon |v|_0 + \bar{c}_2(x_\epsilon, t_\epsilon) (u_2 - u'_2) |_{(x_\epsilon, t_\epsilon)}. \quad (33)$$

Besides, by continuity and since $u_{2,x}(x, t) = 0$ for any point in $\delta\Omega$, we also know that

$$\lim_{(x_\epsilon, t_\epsilon) \rightarrow (x_2, t_2)} (u_2 - u'_2)(x_\epsilon, t_\epsilon) = v(x_2, t_2).$$

We can choose any $0 < \gamma_2 \leq \bar{c}_2(x_\epsilon, t_\epsilon) - \epsilon$, and take the limit when $(x_\epsilon, t_\epsilon) \rightarrow (x_2, t_2)$. Consequently,

$$\begin{aligned} -\epsilon |v|_0 + \bar{c}_2(x_\epsilon, t_\epsilon) (u_2 - u'_2) |_{(x_\epsilon, t_\epsilon)} &\rightarrow -\epsilon |v|_0 + \bar{c}_2(x_2, t_2) (u_2 - u'_2) |_{(x_2, t_2)} \\ &= -\epsilon |v|_0 + \bar{c}_2(x_2, t_2) |v|_0 = [-\epsilon + \bar{c}_2(x_2, t_2)] |v|_0. \end{aligned}$$

Hence, (33) implies that

$$|\mathbb{L}_2 u_2 - \mathbb{L}_2 u'_2|_0 \geq [-\epsilon + \bar{c}_2(x_2, t_2)] |v|_0 \geq \gamma_2 |v|_0.$$

□

As a direct consequence of Lemma B1, we can state the following result:

Corollary B1. Taking $\gamma = \min\{\gamma_1, \gamma_2\}$, the operator $\mathbb{L} \equiv (\mathbb{L}_1, \mathbb{L}_2)'$ verifies

$$\| \mathbb{L}^{-1}u - \mathbb{L}^{-1}u' \|_0 \leq \gamma^{-1} \| u - u' \|_0,$$

for $u, u' \in R(\mathbb{L})$.

Under the assumptions of the theorem, there exist functions $K_i = K_i(x, t)$ in $\Omega \times [0, T]$ with values in \mathbb{R} such that, for any u, v in the domain of f_i ,

$$| F_i(u_1, u_2) - F_i(v_1, v_2) | \leq K_i | u - v | .$$

We define $K \equiv K_1 + K_2$. Let \bar{K} be an upper bound for K . As we will see in the next lemma, one can then conclude that $\mathbb{L}^{-1}F(u)$ is a contraction. The limits of our sequences therefore coincide and define the unique solution of (I):

Lemma B2. Let $\gamma > \bar{K}$, and consider $F(u)$ defined in (9) and $\mathbb{L}^{-1}F$ a mapping from $X \subset C(\Omega \times [0, T])$ into $C(\Omega \times [0, T])$. Then $\mathbb{L}^{-1}F$ possesses the contraction property

$$\| \mathbb{L}^{-1}F(u) - \mathbb{L}^{-1}F(u') \|_0 \leq \frac{\bar{K}}{\gamma} \| u - u' \|_0,$$

for all $u, u' \in X$.

Proof. For each $i = 1, 2$ and for each $u \in X \subset C(\Omega \times [0, T])$, the function $F_i(u)$ is in X . This implies that $F_i(u) \in R(\mathbb{L}_i)$, thus $F(u) \in R(\mathbb{L})$. Hence, by Lemma 1, \mathbb{L}^{-1} is well defined on X . Since F_i verifies

$$| F_i(u) - F_i(u') | \leq K_i | u - u' | \leq K_i | u - u' |_0,$$

function F satisfies the condition

$$\| F(u) - F(u') \|_0 \leq \bar{K} \| u - u' \|_0 \tag{34}$$

for all $u, u' \in X$. We also know from Corollary B1 that

$$\| \mathbb{L}^{-1}F(u) - \mathbb{L}^{-1}F(u') \|_0 \leq \gamma^{-1} \| F(u) - F(u') \|_0 . \tag{35}$$

Taking (34) and (35) together, we obtain

$$\| \mathbb{L}^{-1}F(u) - \mathbb{L}^{-1}F(u') \|_0 \leq \gamma^{-1} \bar{K} \| u - u' \|_0,$$

which proves the lemma. □

C Theorem 2 proof

We prove this result by contradiction. Let $u \in C^{2,1}(\Omega \times [0, T])$ be a solution of (I) and assume that $u \not\leq \langle \tilde{u}, \hat{u} \rangle$. We may encounter six different cases:

- i)* u equals the lower bound of $\langle \tilde{u}, \hat{u} \rangle$ in at least one point (x_0, t_0) , that is to say, $u(x_0, t_0) = \tilde{u}(x_0, t_0)$ and $\tilde{u}(x, t) < u(x, t)$ for all other $(x, t) \in \Omega \times [0, T] \setminus (x_0, t_0)$.
- ii)* u equals the upper bound of $\langle \tilde{u}, \hat{u} \rangle$ in at least one point (x_0, t_0) , that is to say $u(x_0, t_0) = \hat{u}(x_0, t_0)$ and $u(x, t) < \hat{u}(x, t)$ for all other $(x, t) \in \Omega \times [0, T] \setminus (x_0, t_0)$.
- iii)* u is smaller than $(\tilde{u}_1, \tilde{u}_2)$ in at least one point.
- iv)* u is larger than (\hat{u}_1, \hat{u}_2) in at least one point.
- v)* u is smaller than $(\tilde{u}_1, \tilde{u}_2)$ in all points.
- vi)* u is larger than (\hat{u}_1, \hat{u}_2) in all points.

We shall focus on points *i)*, *iii)* and *v)*. Similar reasoning can be applied to study the remaining cases.

- i)* There exists a point (x_0, t_0) at which $\tilde{u}(x_0, t_0) = u(x_0, t_0)$ and $\tilde{u}(x, t) < u(x, t)$ for all other $(x, t) \in \Omega \times [0, T] \setminus (x_0, t_0)$. We define $v(x, t) \equiv u(x, t) - \tilde{u}(x, t)$, then (x_0, t_0) is a minimum for v . Therefore,

$$\begin{aligned} v_t(x_0, t_0) &= 0, \\ v_{xx}(x_0, t_0) &> 0. \end{aligned}$$

Consequently, for $v = (v_1, v_2)$, we have that $\mathbb{L}_1 v_1|_{(x_0, t_0)} < 0$ and $\mathbb{L}_2 v_2|_{(x_0, t_0)} > 0$. However

$$\mathbb{L}_1 v_1 = \mathbb{L}_1(u_1 - \tilde{u}_1) = \mathbb{L}_1 u_1 - \mathbb{L}_1 \tilde{u}_1 = f_1(u_1, u_2) - f_1(\tilde{u}_1, \tilde{u}_2) + \underline{a}_1(u_1 - \tilde{u}_1) \geq 0.$$

The expression above is positive under the assumptions because, as we have seen in the proof of Proposition B1, either f_1 is increasing in both arguments, with $\underline{a}_1 \geq 0$ (which yields the result directly), or because $\underline{a}_2 \leq 0$. Notice that in the later case

$$\mathbb{L}_1 v_1(x, t) \geq -\underline{a}_2(u_2 - \tilde{u}_2) \geq 0.$$

Hence, we achieve a contradiction.

Similarly, one can verify that

$$\mathbb{L}_2 v_2 = \mathbb{L}_2(u_2 - \tilde{u}_2) = \mathbb{L}_2 u_2 - \mathbb{L}_2 \tilde{u}_2 = f_2(u_1, u_2) - f_2(\tilde{u}_1, \tilde{u}_2) - \bar{c}_2(u_2 - \tilde{u}_2) \leq 0,$$

which is negative under the theorem assumptions. As before, this result is directly confirmed if f_2 is decreasing in both arguments, with $\bar{c}_2 \geq 0$. Otherwise, we can show that, for $\bar{c}_1 \leq 0$:

$$\mathbb{L}_2 v_2 \leq \bar{c}_1(u_1 - \tilde{u}_1) \leq 0.$$

A contradiction arises again. Then, we finally conclude that u and \tilde{u} cannot coincide in one point.

iii) In the case when u is smaller than the lower bound of $\langle \tilde{u}, \hat{u} \rangle$ in at least one point, by continuity, there must be at least a point (x_0, t_0) at which $u(x_0, t_0) = \tilde{u}(x_0, t_0)$. But we just showed in point *i)* that this is impossible.

v) Notice that u, \tilde{u} are continuous functions on a compact space. The function v , defined as $v \equiv \tilde{u} - u$, achieves then a minimum at a point denoted by (x_0, t_0) . Therefore,

$$\begin{aligned} v_t(x_0, t_0) &= 0, \\ v_{xx}(x_0, t_0) &> 0. \end{aligned}$$

Hence, $\mathbb{L}_1 v_1 |_{(x_0, t_0)} < 0$ and $\mathbb{L}_2 v_2 |_{(x_0, t_0)} > 0$. However, as in case *i)* above,

$$\mathbb{L}_1 v_1 = \mathbb{L}_1(\tilde{u}_1 - u_1) = \mathbb{L}_1 \tilde{u}_1 - \mathbb{L}_1 u_1 = f_1(\tilde{u}_1, \tilde{u}_2) - f_1(u_1, u_2) + \underline{a}_1(u_1 - \tilde{u}_1) \geq 0,$$

which leads to a contradiction under the theorem assumptions. Similarly,

$$\mathbb{L}_2 v_2 = \mathbb{L}_2(\tilde{u}_2 - u_2) = \mathbb{L}_2 \tilde{u}_2 - \mathbb{L}_2 u_2 = f_2(\tilde{u}_1, \tilde{u}_2) - f_2(u_1, u_2) - \bar{c}_2(u_2 - \tilde{u}_2) \leq 0.$$

The contradiction arises because $\mathbb{L}_2 v_2 |_{(x_0, t_0)} > 0$. This finishes the proof.

D A numerical algorithm

In this appendix we develop an algorithm to numerically compute the optimal solution to our optimal control problem. Since the algorithm is based on the iterative method presented in Theorem 1, it will require a lower or an upper solution to start iterating. For illustration purposes let us build a sequence $\{u_1^k, u_2^k\}_{k \in \mathbb{N}}$ with $(u_1^0, u_2^0) = (\tilde{u}_1, \tilde{u}_2)$. The algorithm can be summarized in the following 3 steps:

Step 1: Set a stopping criterion Δ and build a grid to represent the space-time set $\Omega \times [0, T]$, $\Omega \simeq \{Dx, 2 * Dx, \dots, N * Dx\}$ and $[0, T] \simeq \{Dt, 2 * Dt, \dots, J * Dt\}$. We initialize the series $\{u_1^k u_2^k\}_{k \in \mathbb{N}}$ as

$$(\underline{u}_1^0, \underline{u}_2^0) = (\tilde{u}_1, \tilde{u}_2),$$

$$(\underline{w}_1^0, \underline{w}_2^0) = (\tilde{w}_1, \tilde{w}_2).$$

Step 2: For every k and given (u_1^{k-1}, u_2^{k-1}) and $(w_1^{k-1}, w_2^{k-1})(x, t) = (u_1^{k-1}, u_2^{k-1})(x, T - t)$, compute

$$\begin{cases} \mathbb{L}_1 u_1^k = F_1(u_1^{k-1}, u_2^{k-1}), \\ u_1^k(x, 0) = \frac{\zeta}{k} \tilde{u}_1(x, 0) + (1 - \frac{\zeta}{k}) u_1(x, 0), \\ \lim_{x \rightarrow \delta\Omega} u_{1,x}^k = 0, \end{cases}$$

with $0 < \zeta < 1$, and

$$\begin{cases} \mathbb{L}'_2 w_2^k = -F_2(w_1^{k-1}, w_2^{k-1}), \\ w_2^k(x, 0) = \frac{\phi}{k} \tilde{w}_2(x, 0) + (1 - \frac{\phi}{k}) \Psi'_1(w_1^{k-1}(x, 0), x), \\ \lim_{x \rightarrow \delta\Omega} w_{2,x}^k = 0, \end{cases}$$

with $0 < \psi < 1$.

Step 3: After every k , we compute the distance between (u_1^{k-1}, u_2^{k-1}) and (u_1^k, u_2^k) . If the distance is smaller than Δ , then STOP. If not, set $k = k + 1$ and go to step 2.

References

- Acemoglu D. (2009), *Introduction to Modern Economic Growth*, Princeton University Press, New Jersey.
- Ahmed N. (1977), “Optimal Control of a Class of Strongly Nonlinear Parabolic Systems”, *Journal of Mathematical Analysis and Applications*, 61, 188-207.
- Attanasio O. and M. Browning (1995), “Consumption over the Life Cycle and over the Business Cycle”, *American Economic Review*, 85(5), 1118-1137.
- Attanasio O. and G. Weber (1993), “Consumption, the Interest Rate and Aggregation”, *Review of Economic Studies*, 60(3), 631-649.

- Baldwin R. and P. Martin (2004), “Agglomeration and regional growth”, in *Handbook of Regional and Urban Economics*, vol.4, V. Henderson and J. Thisse (eds.), North-Holland, Amsterdam, The Netherlands.
- Baldwin R. and P. Krugman (2004), “Agglomeration, Integration and Tax Harmonisation”, *European Economic Review*, 48(1), 1-23.
- Barbu V. and T. Precupanu (2012), *Convexity and optimization in Banach spaces*, Springer Science & Business Media, New York City.
- Barro R., and X. Sala-i-Martin (2004), *Economic Growth*, 2nd Edition, The MIT Press, Cambridge, Massachusetts.
- Barro R., G. Mankiw, and X. Sala-i-Martin (1995), “Capital Mobility in Neoclassical Models of Growth”, *American Economic Review*, 85(1), 103-115.
- Bebernes J. and K. Schmitt, K (1979), “On the existence of maximal and minimal solutions for parabolic partial differential equations”, *Proceedings of the American Mathematical Society*, 73(2), 211-218.
- Bental B. (1985), “Is Capital Mobility Always Desirable? A Welfare Analysis of Portfolio Autarky in a Growing Economy”, *International Economic Review*, 26(1), 203-212.
- Blanchet A., P. Mossay, and P. Santambrogio (2016), “Existence and uniqueness of equilibrium for a spatial model of social interactions”, *International Economic Review*, 57, 31-59.
- Boucekkine R., C. Camacho, and B. Zou (2009), “Bridging the gap between growth theory and the new economic geography: The spatial Ramsey model”, *Macroeconomic Dynamics*, 13, 20-45.
- Boucekkine R., C. Camacho, and G. Fabbri (2013a), “Spatial dynamics and convergence: The spatial AK model”, *Journal of Economic Theory*, 148, 2719-2736.
- Boucekkine R., C. Camacho, and G. Fabbri (2013b), “On the optimal control of some parabolic partial differential equations arising in economics”, *Serdica Mathematical Journal*, 39, 1001-1024.
- Brito P. (2004), “The dynamics of growth and distribution in a spatially heterogenous world”, WP 13/2004, ISEG Working Papers.

- Brock W. and A. Xepapadeas (2008), “Diffusion-induced instability and pattern formation in infinite horizon recursive optimal control”, *Journal of Economic Dynamics and Control*, 32, 2745-2787.
- Camacho C. and A. Perez-Barahona (2015), “Land use dynamics and the environment”, *Journal of Economic Dynamics and Control*, 52, 96-118.
- Camacho C., Zou B. and Briani M. (2008), “On the dynamics of capital accumulation across space”, *European Journal of Operational Research*, 186(2), 451-465.
- Chandra J., V. Lakshmikantham and S. Leela (1978), “A monotone method for infinite systems of nonlinear boundary value problems”, *Archive for Rational Mechanics and Analysis*, 68(2), 179-190.
- Chetty R. (2006), “A New Method of Estimating Risk Aversion”, *American Economic Review*, 96(5), 1821-1834.
- Costello C. and S. Polasky (2008), “Optimal harvesting of stochastic spatial resources” *Journal of Environmental Economics and Management*, 56, 1-18.
- Desmet K. and E. Rossi-Hansberg (2009), “Spatial Growth and Industry Age”, *Journal of Economic Theory*, 144(6), 2477-2502.
- Desmet K. and E. Rossi-Hansberg (2010), “On spatial dynamics”, *Journal of Regional Science*, 50(1), 43-63.
- Desmet K. and E. Rossi-Hansberg (2014), “Spatial development”, *American Economic Review*, 104(4), 1211-1243.
- Desmet K. and E. Rossi-Hansberg (2015), “On the spatial economic impact of global warming”, *Journal of Urban Economics*, 88, 16-37.
- Ekeland I. (1974), “On the variational principle”, *Journal of Mathematical Analysis and Applications*, 47, 324-353.
- European Union (2008), “Consolidated version of the Treaty on the Functioning of the European Union. Part 3: Union policies and internal actions. Title IV: Free movement of persons, services and capital. Chapter 4: Capital and payments. Article 63 (ex Article 56 TEC)”, *The Official Journal of the European Union* 115, 71-71.
- Fabbri G. (2016), “Geographical structure and convergence: A note on geometry in spatial growth models”, *Journal of Economic Theory*, 162, 114-136.

- Fattorini H., (1990), "Some remarks on Pontryagin's maximum principle for infinite dimensional control problems", in *Perspectives in Control Theory*, Springer Science+Business Media, 12-25.
- Friedman A. (1959), "Convergence of solutions of parabolic equations to a steady state", *Journal of Mathematics and Mechanics*, 8, 57-76.
- Galor O. (1996), "Convergence? Inferences from theoretical models", *The Economic Journal*, 106, 1056-1069.
- Hadamard J., (1923), *Lectures on the Cauchy Problem in Linear Partial Differential Equations*. Yale University Press.
- Kehoe T., D. Levine, and P. Romer (1990), "Determinacy of Equilibria in Dynamic Models with Finitely Many Consumers", *Journal of Economic Theory*, 50 (1), 1-21.
- Kose A., E. Prasad, K. Rogoff, and S. Wei (2010), "Financial Globalization and Economic Policies", in *Handbook of Development Economics*, vol.5, D. Rodrik and M. Rosenzweig (eds.), North-Holland, Amsterdam, The Netherlands.
- Lions J. (1966a), "Optimisation pour certaines classes d'équations d'évolution non linéaires", *Annali di Matematica Pura ed Applicata*, 72(1), 275-294.
- Lions J. (1966b), "Sur quelques problèmes d'optimisation dans les equations d'evolution linéaires de type parabolique", in *Applications of functional Analysis to Optimization*, Caianiello (ed.), Academic Press, New York.
- Lions J. (1972), "Some aspects of the optimal control of distributed parameter systems", *Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, PA.
- Mossay P. and P. Picard (2011), "On spatial equilibria in a social interaction model", *Journal of Economic Theory*, 146, 2455-2477.
- Pao C. (1992), *Nonlinear Parabolic and Elliptic Equations*, Plenum Press, New York.
- Patel S. (1964), "The Economic Distance Between Nations: Its Origin, Measurement and Outlook", *The Economic Journal*, 74(293), 119-131.
- Quah D. (1996a), "Convergence Empirics Across Economies with (Some) Capital Mobility", *Journal of Economic Growth*, 1, 95-124.

- Quah D. (1996b), “Twin peaks: growth and convergence in models of distribution dynamics”, *The Economic Journal*, 106, 1045-1055.
- Raymond J. and H. Zidani (1999), “Pontryagin’s principle for Time-Optimal Problems”, *Journal of Optimization Theory and Applications*, 101(2), 375-402.
- Rodrik D. and T. van Ypersele (2001), “Captial mobility, distributive conflict and international tax coordination”, *Journal of International Economics*, 54, 57-73.
- Sattinger D. (1972), “Monotone Methods in Nonlinear Elliptic and Parabolic Boundary Value Problems”, *Indiana University Mathematics Journal*, 21(11), 979-1000.
- Thisse J. (2010), “Towards a unified theory of economic geography and urban economics”, *Journal of Regional Science*, 50(1), 281-296.