Managing spatial linkages and geographic heterogeneity in dynamic models with transboundary pollution

Published Version:
Managing spatial linkages and geographic heterogeneity in dynamic models with transboundary pollution / Boucekkine R.; Fabbri G.; Federico S.; Gozzi F.. - In: JOURNAL OF MATHEMATICAL ECONOMICS. - ISSN 0304-4068. - ELETTRONICO. - 98:(2022), pp. 102577.1-102577.15. [10.1016/j.jmateco.2021.102577]

This version is available at: https://hdl.handle.net/11585/942146 since: 2023-09-17

Published:

DOI: http://doi.org/10.1016/j.jmateco.2021.102577

(Article begins on next page)

MANAGING SPATIAL LINKAGES AND GEOGRAPHIC HETEROGENEITY IN DYNAMIC MODELS WITH TRANSBOUNDARY POLLUTION

RAOUF BOUCEKKINE[a], GIORGIO FABBRI[b], SALVATORE FEDERICO[c], AND FAUSTO GOZZI[d]

Abstract. We construct a spatiotemporal frame for the study of optimal growth under transboundary pollution. Space is continuous and polluting emissions originate in the intensity of use of the production input. Pollution flows across locations following a diffusion process. The objective functional of the economy is to set the optimal production policy over time and space to maximize welfare from consumption, taking into account a negative local pollution externality and the diffusive nature of pollution. Our framework allows for space and time dependent preferences and productivity, and does not restrict diffusion speed to be space-independent. Accordingly, we develop a methodology to investigate the environmental and economic implications of geographic heterogeneity. We propose a method for an analytical characterization of the optimal paths and the asymptotic spatial distributions. Our method enables us to enucleate a deep economic concept of spatiotemporal welfare effect of pollution, making it definitely useful for economic analysis. An application to first nature causes of geographic externalities, namely technological spillovers, is proposed for illustration.

Key words: Spatiotemporal modelling, geographic heterogeneity, transboundary pollution, infinite dimensional optimal control, technological spillovers.

JEL classification: Q53, R11, C61, R12, O41.

Date: October 20, 2023.

[a]Aix-Marseille University (IMéRA and AMSE), CNRS, EHESS and Ecole Centrale de Marseille. Corresponding member, IRES, Université catholique de Louvain.
[b]Univ. Grenoble Alpes, CNRS, INRA, Grenoble INP, GAEL, 38000 Grenoble.
[c]Università degli Studi di Siena, Dipartimento di Economia Politica e Statistica, Siena, Italy.
[d]Dipartimento di Economia e Finanza, LUISS *Guido Carli*, Rome, Italy.

1. Introduction

It is readily recognized in the economic literature, and notably in the economic geography and urban economics areas, that space matters in the shape and mitigation of pollution damages (see for example, Arnott et al., 2008). Several important theoretical contributions have been made within full-fledged spatial frameworks to identify optimal spatial allocations when pollution externalities are internalized and also the subsequent decentralization policies; this includes, among others, the early works of Henderson (1977) and Hochman and Ofek (1979), and more recently the above mentioned paper by Arnott et al. (2008). A large majority of papers in this topic remains however empirical (see for example, Gibson and Carnovale, 2015, or Henderson, 1996).

Moreover, it should be observed that beside being based on static models, the related theoretical work is usually based on simplifying *ad hoc* assumptions on the level of pollution at given location (this fact is for instance totally transparent in Arnott et al., 2008[1]). This is of course acceptable in certain contexts, and it is even more acceptable if the alternative more accurate specifications are intractable. This paper is a methodological contribution to this important area of economics at the intersection of geographic and environmental economics. Since admittedly spatial heterogeneity is key in shaping (local) pollution and the induced spatial distribution of pollution, we propose a methodology which closely incorporates a large set of spatial heterogeneities as explained below while the diffusion process of pollution over time and space is accurately modeled (through a diffusion partial differential equation). With this focus and objective in mind, we are able to build up a spatiotemporal framework for pollution control, which produces closed-form solutions for any of these spatial heterogeneity traits, and analytical characterizations of the induced long-term spatial distributions as well. This is done at the cost of a simplification in the shape of the production function (linearity assumption) but the economic insight from the

[1]See page 392, in particular.

analytical exploration of spatial heterogeneity impact on economic decision making and pollution outcomes well outweigh the latter cost.

As one can infer from above, we are firstly interested in transboundary (diffusive) pollution like air or water pollution. The problem of pollution control is deeply intricate in this case. It raises several key issues. One has to do with the strategic ingredients of the problem either at the international level (refer to the successive failures or at least questioning of internationals agreements to control global warming, from the Kyoto protocol in 1991 to the Paris agreement in 2016) or at the regional scale. We abstract away from these considerations here. Several papers have been written in the last decade on this topic from 2-country setting (Boucekkine et al., 2011) to continuous space modelling (see in particular, de Frutos and Martin-Herran, 2019) through multi-country frameworks (for example, Dockner and Long, 1993). We rather concentrate on a second set of issues, those related to the fact that the impact of air pollution is first of all local, and its magnitude and persistence depend pretty much on the local conditions. If a central planner at a country or international level has to set a pollution control policy for the benefit of all the individuals concerned, then she should take as much as possible into account the heterogeneity across locations, in addition to the fact that air pollution, being transboundary, requires the internalization of spatial externalities.

There are several spatial heterogeneity features to account for. Obvious ones are technological heterogeneity and heterogeneity in preferences (which covers cultural discrepancies with respect to the environment among others). But there are also more geographic and ecological differences, which matter a lot both in the diffusion of pollution across locations and in its local impact. The self-cleaning capacity of Nature may vary from a region to a close one, the local topography, land use and infrastructures may speed up pollution diffusion or slow it down... etc. These issues are quite known across disciplines (see Tiwari and Closs, 2010, for an excellent book on air pollution), and, as indicated in the beginning of this introduction, they are also very much recognized in the literature for decades. Several recent papers have been devoted to transboundary pollution.[2] None of these papers however poses the latter problem in a full-fledged analytical spatiotemporal frame incorporating the above mentioned technological, preference, geographic and ecological spatial discrepancies.

Clearly, taking up the challenge would involve plenty of technical problems, with a very likely lack of tractability and the forced use of numerical solutions. Contrary to other disciplines (like quantitative geography, climate science or ecology) in which the use of black-box disaggregated models is routine, the economists, being more interested in identifying mechanisms, are more keen to develop parsimonious models. In this paper, we build up a spatiotemporal optimal control framework allowing to encompass the heterogeneity traits outlined above in the presence of transboundary pollution, and still producing a comprehensive analytical characterization. In particular, with respect to the aforementioned recent literature on transboundary pollution, we considerably enrich the geography of the model. On the technological ground, we allow productivity both time and space-dependent: any pattern of technology diffusion across time and space can be accommodated. As for preferences, we make them time and space-dependent as well. Finally, to account for the local geographic and ecological conditions discrepancy, we allow for a space dependent self-cleaning capacity and pollution diffusion speed.

This generality has a methodological implication, since the heterogeneities in space and time make very difficult to use the dynamic programming method (see e.g. Boucekkine et al., 2019a, 2019b) or the maximum principle (e.g. Brito, 2004, or more recently, Ballestra, 2016) in order to

[2]For example, Camacho and Perez-Barahona (2015) studied the problem of atmospheric transboundary pollution in the context of an optimal land use problem. Other authors have also provided valuable contributions to the analysis of the spatiotemporal deep nature of the transboundary pollution control problem; see for instance Augeraud-Véron et al. (2017, 2019b), Boucekkine et al. (2019a) or Grass and Ueckner (2017), and the recent survey in Augeraud-Véron et al., 2019a).

produce analytical solutions. Instead, we apply a functional transformation technique observing that the objective functional can be rewritten in a way that allows for a direct maximization method, ultimately finding the explicit form of the optimal control[3]. An extremely appealing feature of this method is that it builds on a pivotal spatial function (denoted $\alpha(\cdot)$ in Section 3), which admits a neat economic interpretation: it corresponds at any location x to the discounted sum of future disutility stream of a unit of pollutant initially located at x. Indeed, in our spatiotemporal framework with transboundary pollution, a unit of pollutant has a different effect on social welfare depending on where it is initially located and how it is going to spread over space in the future. Our alternative method has the virtue to put forward this deep economic concept, which makes it definitely transparent and useful for economic analysis.

To provide with a first illustration of our method, we will use it to study, following the terminology of Krugman (1993), a first nature cause of spatial externalities: **given** geographic differences in terms of efficiency either at production or at pollution abatement. We rely on the non-spatial technology evolution model specified in Nocco (2005) and adapt it to our spatial setting. A key feature of our adaptation is that technological spillovers tend to decrease with distance to a given technological center. This is totally consistent with empirical evidence. For instance, Deltas and Karkalakos (2013) found that in increase in distance between the center and the recipient by 500 km reduces spillovers by 55-70%. It is important to note here that our method can also in principle support the study of second nature causes. In a companion paper, we show, in a simplified setting where we drop heterogeneity in diffusion and elasticity, how the introduction of an additional control variable, that is abatement, does not at all break down the analytical solution scheme despite the nonlinearities conveyed by this addition (Boucekkine et al., 2020). That speaks a lot about the flexibility of our approach.

[3] A somehow related method is used, in a different economic context, by Barucci and Gozzi (2001) but the problem there did not include a diffusive process.

The paper is organized as follows. Section 2 presents the basic economic problem and defends its economic relevance. Section 3 develops the method used to solve the generic spatiotemporal optimal growth model under transboundary pollution. In particular, closed-form solutions of the optimal strategies are identified. We show also further analytical results on transition dynamics and asymptotics (that's the computation of asymptotic spatial distributions). Section 4 gives an illustration of the method in the presence of technological discrepancy across space using the spatiotemporal spillover mechanism explained just above. Of course, our aim in this ultimate section is not to exploit entirely the richness of our setting but only to indicate how far the analytical study can go. Section 5 concludes. Appendix A contains the formal proofs.

2. The basic economic problem

The basic problem incorporates some elements of the models by Boucekkine et al. (2011), Camacho and Perez Barahona (2015), Boucekkine et al. (2019a), and de Frutos and Martin-Herran (2019). The originality of our contribution is to provide a full-fledged analytical spatiotemporal setting yet incorporating, as described in detail below, space and time-dependence of both preferences and production technology, as well as space-dependence of the pollution diffusion coefficient and of the self-cleaning capacity.

Let us now describe the basic economic problem. Consider a continuum of locations, say along the circle in $\mathbb{R}^2$. The choice of the circle is made for simplicity.[4] Call it S^1:

$$S^1 := \left\{x \in \mathbb{R}^2 : \ |x|_{\mathbb{R}^2} = 1\right\}. \tag{1}$$

[4]Our approach allows generalizations to compact finite dimensional manifolds without boundary (see, e.g., Fabbri, 2016).

Each location uses a linear (Leontief) production function: at any location x in time $t \geq 0$, production is

$$y(t,x) = a(t,x)\, i(t,x), \tag{2}$$

where $y(t,x)$ is the output, $i(t,x)$ is the capital input, and $a(t,x)$ is productivity at location x in time t. A few comments are in order already at this stage. First, we allow productivity per location to be both generically space-dependent and time-dependent, so that our setting can include for example the typical exponential exogenous technological progress in neoclassical growth theory. Actually, our modelling allows for much more: as argued in the introduction section, $a(t,x)$ can be specified to model possible technological spillovers across locations, uneven technological development over space (that's barriers to technological diffusion) and the like. Admittedly, the latter are key features in regional development. Second, it's worth noticing that our production technology mimics the AK technology, which is a basic ingredient in endogenous growth theory (see Barro and Sala-i-Martin, 2004, Chapter 4), but with full depreciation of capital. This is essential to get our closed-form solutions, and actually this is a well known trick in optimal growth theory either to generate analytical solutions and/or to simplify the analysis (reduction of the dimension of the dynamic systems involved) and focus on other state variables (see again Barro and Sala-i-Martin, 2004, Chapter 6). In our case, both arguments hold: given the complexity of our problem (infinite dimensional optimization) and the economic target (economic performance in the heterogenous space with transboundary pollution), we find it convenient to shut down capital accumulation to be able to develop a comprehensive enough analytical spatiotemporal framework to approach the latter economic objectives.

At any location, output is produced, consumed and locally invested (no trade across locations), implying:

$$c(t,x) + i(t,x) = y(t,x), \tag{3}$$

where $c(t,x)$ is consumption at location x at time t. The unique link among locations is transboundary pollution. We target air pollution and consider a broad specification incorporating ecological efficiency at any location x and time t. Precisely, the accumulation of pollution spatial profile is assumed to evolve according to the following parabolic partial differential equation (PDE):

(4)
$$\begin{cases} \dfrac{\partial p}{\partial t}(t,x) = \dfrac{\partial}{\partial x}\left(\sigma(x)\dfrac{\partial p}{\partial x}(t,x)\right) - \delta(x)p(t,x) + \eta(t,x)i(t,x), \quad (t,x) \in \mathbb{R}^+ \times S^1, \\ \\ p(0,x) = p_0(x), \quad x \in S^1, \end{cases}$$

where $p(t,x)$ is the pollution stock at location x and time t.

First, notice the general shape of the pollutants' emissions term $\eta(t,x)i(t,x)$ in the pollution spatiotemporal dynamics depicted above. Indeed, the unusual function $\eta(t,x)$ is meant to reflect that the pollution impact of emissions arising from the use of one unit of input may not be the same over time and space. It can be readily observed that: (i) taking $\eta(t,x) = 1$ brings to consider the case where polluting emissions at location x are exactly equal to input use intensity; (ii) if we specify $\eta(t,x) = a(t,x)\phi(t,a)$, the term $\eta(t,x)i(t,x)$ reads as $a(t,x)\phi(t,x)i(t,x) = \phi(t,x)y(t,x)$, and we are able to give a specification of the model where emissions depend on output rather than on input used; (iii) the temporal dependence allows us, in general, to incorporate exogenous technological progresses in ecological efficiency (such as those conveyed by abatement activities); (iv) finally, independent spatial heterogeneities can also be taken into account with this more general specification.

Second, we underline the space-dependent form of the transboundary pollution diffusion term, that is $\frac{\partial}{\partial x}(\sigma(x)\frac{\partial p}{\partial x}(t,x))$, where $\sigma(x)$ is the pollution diffusion speed at location x. Indeed, there is a large bunch of works documenting the role of local conditions in air pollution diffusion (see Tiwary and Colls, 2010, Chapter 1). Therefore, this generalization significantly increases the relevance of our analytical frame. Moreover, we notice that the pollution diffusion also depends on nature local regeneration capacity (the term $\delta(x)p(t,x)$), and on current emissions (as captured by the term $\eta(t,x)i(t,x)$).

Finally, since our setting is spatiotemporal, an initial spatial distribution of pollution is needed, and it is given by function $p_0(x)$ defined on S^1. The whole state variable dynamics then follows the PDE (4). The infinite dimensional nature of the involved optimization problem derives from the latter characteristic of the state dynamics.

On the side of preferences, we choose the instantaneous per location utility to be of CRRA type with respect to consumption and linear with respect to pollution:

$$U\left(c(t,x),p(t,x)\right)=\frac{c(t,x)^{1-\gamma(t,x)}}{1-\gamma(t,x)}-w(x)p(t,x),$$

where $\gamma(t,x)\in(0,1)\cup(1,\infty)$ measures the inverse of the elasticity of intertemporal substitution in consumption at location x and time t, and $w(x)$ measures, for instance, local environmental awareness at location x. Observe that our negative pollution externality is local. Our framework is not designed primarily to study global warming and therefore global pollution, but the diffusion of air pollutants with local health impact like particles for example (again a comprehensive account of air pollutants can be found in Tiwari and Colls, 2010, Chapter 1). The local impact of pollution is also captured via $w(x)$, which can be indeed interpreted as local awareness and sensitivity to environmental problems. It can also reflect specific priorities of the planner to cope with particular local conditions. Notice also that our frame allows for time and space varying preferences through the parameter $\gamma(t,x)$. We consider a planner problem who has to maximize the following social welfare under the above specified technological constraints of the economy and the transboundary pollution faced:

$$J\big(p_0;i\big):=\int_0^\infty e^{-\rho t}\left(\int_{S^1}\left(\frac{c(t,x)^{1-\gamma(t,x)}}{1-\gamma(t,x)}-w(x)p(t,x)\right)\mathrm{d}x\right)\mathrm{d}t,$$

where ρ is the parameter at which the planner discounts time. By using equations (1) and (2), we can rewrite the functional in terms of the control variable $i(t,x)$:

(5)

$$J\big(p_0;i\big)=\int_0^\infty e^{-\rho t}\left(\int_{S^1}\left(\frac{\big((a(t,x)-1)i(t,x)\big)^{1-\gamma(t,x)}}{1-\gamma(t,x)}-w(x)p(t,x)\right)\mathrm{d}x\right)\mathrm{d}t.$$

Notice we do not incorporate population density in our analysis, and notably in the functional, nor do we introduce mortality (possibly related to pollution). We fundamentally focus on handling spatial heterogeneity abstracting away from demography, which is an already daunting task. At the minute, one could simply interpret the social welfare function above as a Benthamite welfare function summing individual welfare over locations and time with one infinite-lived individual at each location.

3. Theoretical analysis

In this section we give a precise description of our results, specifying hypotheses and formalizing the statements. To increase the readability of the text we postpone the proof in Appendix A and we divide the section in four parts.

To apply the functional transformation technique that we exploit to solve the optimal control problem (Theorem 3.6), we start by rewriting the problem in a suitable Hilbert space formalism (Subsection 3.1), then we identify a spatial function α that will be essential in the transformed expression of the functional (Subsection 3.2), and finally we will characterize the optimal control and the corresponding social welfare (Subsection 3.3). Subsection 3.4 contains transitional and long-run analysis of the dynamics via series expansions.

3.1. **Infinite dimensional formulation and preliminary results.** On the space support S^1 introduced in (1) we consider the metrics induced by the Euclidean metrics of $\mathbb{R}^2$. In this way S^1 can be isometrically identified with $2\pi\mathbb{R}/\mathbb{Z}$ and the (class of) functions $S^1\to\mathbb{R}$ with 2π-periodic function

$\mathbb{R} \to \mathbb{R}$; differentiation of functions $S^1 \to \mathbb{R}$ is defined according to this identification. Consequently, the initial pollution distribution and the space dependent parameters δ, σ and w are measurable functions

$$p_0, \delta, \sigma, w : S^1 \to \mathbb{R}^+;$$

similarly the time and space dependent parameters γ, a, η are measurable functions

$$\gamma : \mathbb{R}^+ \times S^1 \to (0,1) \cup (1,+\infty), \quad a : \mathbb{R}^+ \times S^1 \to (1,+\infty), \quad \eta : \mathbb{R}^+ \times S^1 \to (0,+\infty).$$

We proceed now to our infinite dimensional reformulation of the problem. We will use the framework of Lebesgue and Sobolev spaces, for more details we refer to Brezis (2011). The infinite dimensional space H, where we will reformulate our maximization, is the Lebesgue space $L^2(S^1;\mathbb{R})$, i.e.[5]

$$H := L^2(S^1;\mathbb{R}) := \left\{ f : S^1 \to \mathbb{R} \text{ measurable} : \int_{S^1} |f(x)|^2 \mathrm{d}x < \infty \right\},$$

endowed with the usual inner product $\langle f, g\rangle = \int_{S^1} f(x)g(x)\mathrm{d}x$, which makes it a Hilbert space. We denote by $\|\cdot\|$ the associated norm, by H^+ the nonnegative cone of H, i.e.

$$H^+ := \{f \in H : \ f \geq 0\},$$

and by $\mathbf{1}$ the constant function equal to 1 on S^1. Moreover, we introduce the Sobolev space[6]

$$(6) \quad W^{2,2}(S^1;\mathbb{R}) := \big\{ f \in L^2(S^1;\mathbb{R}) : \ f \text{ is twice weakly differentiable and } f', f'' \in L^2(S^1;\mathbb{R}) \big\}.$$

Some degree of regularity of the parameters will be necessary in the analysis. We will work with the following assumptions.

[5]Actually, rather than a space of functions, $L^2(S^1;\mathbb{R})$ is a space of equivalence classes of functions, with the equivalence relation identifying functions which are equal *almost everywhere*, i.e. out of a null Lebesgue measure set. For details we refer again to Brezis (2011).

[6]We refer to Brezis (2011) for the notion of *weak* differentiability.

Assumption 3.1.

(i) $p_0 \in L^2(S^1;\mathbb{R}^+)$, $\delta \in C(S^1;\mathbb{R}^+)$, $\sigma \in C^1(S^1;(0,+\infty))$, $w \in C(S^1;(0,+\infty))$;

(ii) the function $\gamma : \mathbb{R}^+ \times S^1 \to (0,1) \cup (1,+\infty)$ *is measurable and there exists* $\kappa \in (0,1)$ *such that, for every* $(t,x) \in \mathbb{R}^+ \times S^1$,

$$\text{either (Case (A))} \qquad \kappa \;\leq\; \gamma(t,x) \;\leq\; 1-\kappa$$
$$\text{or (Case (B))} \qquad 1+\kappa \;\leq\; \gamma(t,x) \;\leq\; \frac{1}{\kappa}.$$

(iii) There exist $L > 0$, *and* $g \geq 0$ *such that, for every* $(t,x) \in \mathbb{R}^+ \times S^1$,

$$\left(\frac{a(t,x)-1}{\eta(t,x)}\right)^{\frac{1}{\gamma(t,x)}-1} \;\leq\; Le^{gt};$$

(iv) $\rho > g$.

Notice that the above Assumption 3.1 (iii) and (iv), which may seem quite technical, are indeed the translation, in this case, of the standard assumptions needed to guarantee "well posedness" (i.e. finiteness of the value function and existence/uniqueness of the optimal strategy) exactly as in the standard Ramsey-type AK model, see e.g. Freni et al (2006) on this.

Hereafter, our arguments will make use of the theory of unbounded linear operators and semigroups of linear operators, for which we refer to Engel and Nagel (1995). Denote by $L(H)$ the space of bounded linear operators on H. We consider the differential operator $\mathcal{L} : D(\mathcal{L}) \subset H \to H$, where

$$D(\mathcal{L}) = W^{2,2}(S^1;\mathbb{R}); \qquad (\mathcal{L}\varphi)(x) = (\sigma\varphi')'(x) - \delta(x)\varphi(x), \quad \varphi \in D(\mathcal{L}).$$

Proposition 3.2. *Let Assumption 3.1 hold. Then* $\mathcal{L}$ *generates a strongly continuous contraction semigroup* $(e^{t\mathcal{L}})_{t\geq 0} \subset L(H)$. *Moreover,* ρ *belongs to the resolvent set of* $\mathcal{L}$, *i.e.* $\rho - \mathcal{L} : D(\mathcal{L}) \to H$ *is invertible with bounded inverse* $(\rho-\mathcal{L})^{-1} : H \to D(\mathcal{L})$ *and*

$$(\rho-\mathcal{L})^{-1}h = \int_0^\infty e^{-(\rho-\mathcal{L})t}h\,\mathrm{d}t \qquad \forall h \in H. \tag{7}$$

Proof. See Appendix A ☐

Given $i : \mathbb{R}^+ \times S^1 \to \mathbb{R}^+$, define

$$I : \mathbb{R}^+ \to H^+, \qquad I(t) := i(t, \cdot).$$

Morever, define

$$\Psi : \mathbb{R}^+ \to H^+, \quad \Psi(t) := \eta(t, \cdot).$$

Finally, given $h, k \in H$, define $(hk)(x) := h(x)k(x)$. Then, with the identification $P(t) = p(t, \cdot)$, we reformulate (4) in H as

$$\begin{cases} P'(t) = \mathcal{L}P(t) + \Psi(t)I(t), & t \geq 0, \\ P(0) = p_0 \in H, \end{cases} \tag{8}$$

According to Definition 3.1(v), Chapter 1, Part II, of Bensoussan et al. (2007), given $I \in L^1_{loc}(\mathbb{R}^+; H^+)$, we define the *mild solution* to (8) as

$$P(t) = e^{t\mathcal{L}} p_0 + \int_0^t e^{(t-s)\mathcal{L}} \Psi(s) I(s) \mathrm{d}s, \quad t \geq 0. \tag{9}$$

Setting $A(t) := a(t, \cdot)$, $\Gamma(t) := \gamma(t, \cdot)$, and

$$\left[\frac{\big((A(t) - \mathbf{1})I(t)\big)^{1-\Gamma(t)}}{1 - \Gamma(t)} \right](x) := \frac{\big((a(t,x) - 1)i(t,x)\big)^{1-\gamma(t,x)}}{1 - \gamma(t,x)}, \quad x \in S^1,$$

the functional (5) is rewritten in this formalism as

$$J(p_0, I) = \int_0^\infty e^{-\rho t} \left[\left\langle \frac{\big((A(t) - \mathbf{1})I(t)\big)^{1-\Gamma(t)}}{1 - \Gamma(t)}, \mathbf{1} \right\rangle - \langle w, P(t) \rangle \right] \mathrm{d}t. \tag{10}$$

We introduce the following set of admissible controls

$$\mathcal{A} := \left\{ I \in L^1_{loc}(\mathbb{R}^+; H^+) : \int_0^\infty e^{-\rho t} \|\Psi(t)I(t)\| \, \mathrm{d}t < \infty \right\}.$$

The following result shows that the functional is well defined on $\mathcal{A}$.

Proposition 3.3. *Let Assumption 3.1 hold. The functional $J(p_0, I)$ is well defined for all $p_0 \in H$ and $I \in \mathcal{A}$.*

Proof. See Appendix A. □

Finally, we define the value function as the optimal value of J over $\mathcal{A}$, i.e.

$$v(p_0) := \sup_{I \in \mathcal{A}} J(p_0; I).$$

Note that this function may possibly be infinite. The function

$$\alpha := (\rho - \mathcal{L})^{-1} w \in H. \tag{11}$$

will play a key role in the transformation of the functional J that we will perform: it represents the core of the solution. Moreover, as recalled in the introduction, it admits a clear economic interpretation: it corresponds at any location x to the discounted sum of future disutility stream of a unit of pollutant initially located at x. Hence in the next subsection we will investigate some properties of it and clarify its economic interpretation.

3.2. The function α and its properties.

3.2.1. *The equation for α and its mathematical properties.* By definition α is the unique solution in $W^{2,2}(S^1;\mathbb{R})$ of the abstract ODE

$$(\rho - \mathcal{L})\,\alpha = w. \tag{12}$$

More explicitly, α, as defined in (11), is the unique solution in $W^{2,2}(S^1;\mathbb{R})$ to

$$\rho\alpha(x) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\sigma(x)\frac{\mathrm{d}}{\mathrm{d}x}\alpha(x)\right) + \delta(x)\alpha(x) = w(x), \quad x \in S^1, \tag{13}$$

meaning that it verifies (13) pointwise almost everywhere in S^1. The latter ODE can be viewed as on ODE on the interval $(0, 2\pi)$ with zero-order and first-order periodic boundary conditions[7], that is

$$\begin{cases} \rho\alpha(x) - \dfrac{\mathrm{d}}{\mathrm{d}x}\left(\sigma(x)\dfrac{\mathrm{d}\alpha}{\mathrm{d}x}(x)\right) + \delta(x)\alpha(x) = w(x), \quad x \in (0, 2\pi), \\ \alpha(0) = \alpha(2\pi), \quad \alpha'(0) = \alpha'(2\pi), \end{cases}$$

[7]Falling into the Sturm-Liouville theory with periodic boundary conditions (see Coddington and Levinson, 1955).

where $\alpha'(0)$ and $\alpha'(2\pi)$ are, respectively, the right derivative at 0 and the left derivative at 2π of α.

As better argued below the value of α at a certain spatial point x has the meaning of the sum of all future (discounted) disutility of a unit of pollutant initially located at x. By Sobolev embedding $W^{2,2}(S^1;\mathbb{R}) \subset C^1(S^1;\mathbb{R})$, so $\alpha \in C^1(S^1;\mathbb{R})$. With a little more subtle analysis something more can be said about the properties of α as shown in next proposition.

Proposition 3.4. *Let Assumption 3.1 hold. Then* $\alpha \in C^2(S^1;\mathbb{R})$ *and*

$$0 < \min_{S^1} \frac{w}{\rho + \delta} \leq \alpha(x) \leq \max_{S^1} \frac{w}{\rho + \delta} \quad \forall x \in S^1.$$

Proof. See Appendix A. □

We have the following interesting result on the dependence of α on the diffusion coefficient σ when the latter is constant over space.

Proposition 3.5. *Let Assumption 3.1 hold. Denote by* α_{σ^o} *the solution to* (13) *when* $\sigma(\cdot) \equiv \sigma^o > 0$*. We have*

(14)
$$\lim_{\sigma^o \to 0^+} \alpha_{\sigma^o}(x) = \frac{w(x)}{\rho + \delta(x)}, \qquad \lim_{\sigma^o \to +\infty} \alpha_{\sigma^o}(x) = \frac{\int_{S^1} w(x)dx}{\int_{S^1} (\rho + \delta(x))dx}, \quad \forall x \in S^1.$$

Proof. See Appendix A. □

As will be clearer shortly (Subsection 3.3), the function α has a key role both in expressing the functional in its transformed form and in describing the optimal behavior of the planner. For this reason, understanding its behavior is interesting to describe the behavior of the model.

3.2.2. *The economic interpretation of* α. We now exploit the equation for α and the Propositions 3.4-3.5 to describe the economic intuition about the function α. We said that $\alpha(x)$ has the meaning of the sum of all future (discounted) disutilities of a unit of pollutant initially located at x. First of all observe that, if we take the non-spatial version of the model or, equivalently, a specification of the model where all the parameters are constant in space, the value of such disutility is $\frac{w}{\rho+\delta}$. Indeed in this case the sum of all future (discounted) disutilities of a unit of pollutant would simply be (we take the integral since we are in the continuous time case)

$$\int_0^{+\infty} e^{-\rho t} e^{-\delta t} w \, dt, \tag{15}$$

where we recall that δ is the natural decay of the pollution.

This result, not surprisingly, would be the same if we consider a specification of the model where w and δ are not anymore are constant in space, but no diffusion is present (i.e. $\sigma \equiv 0$). Indeed this corresponds to the case where the pollution does not move among the locations and accumulates in the production site. As recalled above, in the model the only link among the locations is the transboundary pollution. Letting the diffusivity to zero means to reset this channel of interdependence and therefore the model reduces to an independent (uni-dimensional) optimization problem to each point of the space. Since the pollutant which is produced at the location x remains there forever then, similarly to (15), the sum of all future (discounted) disutilities of a unit of pollutant would be

$$\int_0^{+\infty} e^{-\rho t} e^{-\delta(x) t} w(x) \, dt = \frac{w(x)}{\rho + \delta(x)}$$

The above, as expected, is also the pointwise expression appearing in Proposition 3.5, equation (14), when $\sigma^o \to 0^+$.

When we add the space to the model a second order term, depending on σ, appears in the equation which defines α, see (13). This comes of course from the spatial diffusion process of the pollution. The solution of that equation (except when both w and δ do not depend on time) is given by a space-heterogeneous function α. This heterogeneity is due to the fact that, in the model, a unit of pollutant has different effect on the social utility depending on where it is located and how it is going to spread in the future. In terms of pollution-disutility, locations are indeed different for two reasons: the different decay $\delta(x)$ of pollutions and the different unitary instantaneous disutilities $w(x)$. In the general spatial case, the function α at a point x is the (total/social) future discounted disutility of a unit of pollutant initially located at point x. To explain this fact we denote by $\Delta_{\{x\}}$

the Dirac delta in a certain spatial point x, and observe that (see equation (37) in the Appendix where we substitute p_0 with $\Delta_{\{x\}}$) we formally have:

$$\alpha(x) = \left\langle \alpha, \Delta_{\{x\}} \right\rangle = \int_0^\infty e^{-\rho t} \left\langle w, e^{\mathcal{L}t} \Delta_{\{x\}} \right\rangle \mathrm{d}t$$

i.e.

$$\alpha(x) = \int_0^\infty e^{-\rho t} \left(\int_{S^1} w(\xi)\varphi(t,\xi;x)\, \mathrm{d}\xi \right) \mathrm{d}t \tag{16}$$

where $\varphi(t,\xi;x) = \left(e^{\mathcal{L}t}\Delta_{\{x\}}\right)(\xi)$ is the solution of the parabolic equation

$$\begin{cases} \frac{\partial \varphi}{\partial t}(t,\xi) = \frac{\partial}{\partial \xi}\left(\sigma(\xi)\frac{\partial \varphi}{\partial \xi}(t,\xi)\right) - \delta(\xi)\varphi(t,\xi), \\ \varphi(0,\xi) = \Delta_{\{x\}}(\xi) \end{cases},$$

i.e. the spatial density (with respect to the variable x) at time t of a pollutant initially concentrated at point x, once one takes into account the diffusion process and the natural decay. Thus, the term $\int_{S^1} w(\xi)\varphi(t,\xi;x)\,\mathrm{d}\xi$ measures the instantaneous disutility all over the space and the whole expression in the right hand side of (16) is the total spatial (temporally discounted) future social disutility of a unit of pollutant initially concentrated at x.

Finally the second limit of Proposition 3.5, equation (14), correspond to the infinite diffusivity benchmark that is the case where, at each moment, the speed of the diffusion process is so fast that the pollution is instantaneously redistributed uniformly throughout the space. For this reason, whatever the specific value of δ or of w in the precise point of the emission is not relevant but only global averages of the parameters matters.

3.3. **Characterization of the optimal control.** We describe now how previous results can be used: first to rewrite the functional in a transformed form; second to explicitly find the optimal solution of the problem, the related optimal trajectory and the welfare function. The main results are described first (Theorem 3.6) in the Hilbert space formalism introduced above and then restated (Corollary 3.7) using a more readable PDE notation.

Theorem 3.6. *Let Assumption 3.1 hold.*

(i) The functional (10) can be rewritten as

(17)

$$J\big(p_0;I\big) = -\langle \alpha, p_0\rangle + \int_0^\infty e^{-\rho t}\left[\left\langle \frac{\big((A(t)-\mathbf{1})I(t)\big)^{1-\Gamma(t)}}{1-\Gamma(t)}, \mathbf{1}\right\rangle - \langle \alpha, \Psi(t)I(t)\rangle\right] \mathrm{d}t.$$

(ii) The control I^* given by

$$I^*(t)(x) := (\eta(t,x)\alpha(x))^{-\frac{1}{\gamma(t,x)}}(a(t,x)-1)^{\frac{1}{\gamma(t,x)}-1} \tag{18}$$

belongs to $\mathcal{A}$ and is the unique optimal control of the problem.

(iii) The optimal state at time $t \geq 0$, that is $P^*(t)$, is given by

$$P^*(t) := e^{t\mathcal{L}}p_0 + \int_0^t e^{(t-s)\mathcal{L}}\Psi(s)I^*(s)ds. \tag{19}$$

(iv) The value function is finite and affine in p_0; more precisely,

$$v(p_0) = J\big(p_0;I^*\big) = -\langle \alpha, p_0\rangle + q,$$

where

$$q := \int_0^\infty e^{-\rho t}\left[\left\langle \frac{\big((A(t)-\mathbf{1})I^*(t)\big)^{1-\Gamma(t)}}{1-\Gamma(t)}, \mathbf{1}\right\rangle - \langle \alpha, \Psi(t)I^*(t)\rangle\right] \mathrm{d}t.$$

Proof. See Appendix A. □

In the following corollary we summarize the results we have obtained so far rephrasing them in the PDE setting, where we use the identification of S^1 with the real interval $[0, 2\pi]$ with the identification of the extremes 0 and 2π.

Corollary 3.7. *Let Assumption 3.1 hold.*

(i) The optimal investment production input is given by

$$i^*(t,x) = (\eta(t,x)\alpha(x))^{-\frac{1}{\gamma(t,x)}}(a(t,x)-1)^{\frac{1}{\gamma(t,x)}-1}. \tag{20}$$

where α is the unique solution to the following ODE

$$\begin{cases} \rho\alpha(x) - \dfrac{\mathrm{d}}{\mathrm{d}x}\left(\sigma(x)\dfrac{\mathrm{d}\alpha}{\mathrm{d}x}(x)\right) + \delta(x)\alpha(x) = w(x), \quad x \in (0, 2\pi), \\ \alpha(0) = \alpha(2\pi), \quad \alpha'(0) = \alpha'(2\pi). \end{cases}$$

(ii) The dynamics of the pollution profile p^* along the optimal path is the unique solution to the following parabolic PDE

(21)

$$\begin{cases} \dfrac{\partial p^*}{\partial t}(t,x) = \dfrac{\partial}{\partial x}\left(\sigma(x)\dfrac{\partial p^*}{\partial x}(t,x)\right) - \delta(x)p^*(t,x) + \alpha(x)^{-\frac{1}{\gamma(t,x)}}\left(\dfrac{a(t,x)-1}{\eta(t,x)}\right)^{\frac{1}{\gamma(t,x)}-1}, \\ p^*(t,0) = p^*(t,2\pi), \quad \dfrac{\partial p^*}{\partial x}(t,0) = \dfrac{\partial p^*}{\partial x}(t,2\pi), \qquad t \geq 0, \\ p^*(0,x) = p_0(x), \quad x \in [0,2\pi]. \end{cases}$$

(iii) The social welfare is

$$v(p_0) = -\int_0^{2\pi} \alpha(x)p_0(x)\mathrm{d}x + \int_0^{\infty} e^{-\rho t}\left(\int_0^{2\pi} \frac{\gamma(t,x)}{1-\gamma(t,x)}\left(\frac{a(t,x)-1}{\eta(t,x)\alpha(x)}\right)^{\frac{1}{\gamma(t,x)}-1}\mathrm{d}x\right)\mathrm{d}t, \tag{22}$$

In the previous statements we have seen the explicit solution of the optimal problem of the planner. A first, eminently technical, observation concerns the transformation of the functional. In fact, the first result of Theorem 3.6 is that the functional (10) can be rewritten in the form (17). The new form is particularly useful as, in this expression, the state P no longer appears and it is only given in terms of the control I; this fact greatly simplifies the analysis. Thanks to this transformation, it is indeed much easier to find the expression of the optimal investment and the subsequent results.

Looking at the expressions that appear in Corollary 3.7, it is immediately evident that all the heterogeneities of the problem enter directly and in a non-trivial way in the solution. Partly, they appear explicitly in the expressions of the optimal spatial profile of the investment or of the parabolic equation describing the evolution of the spatial distribution of pollution and partly they contribute to these expressions through the expression of the function α. In this way the model delivers optimal outcomes that are sensitive to heterogeneities of different nature: environmental heterogeneities (the ability to regenerate of the ecological context measured in each place by δ), productive heterogeneities (the productivity which depends exogenously on both the location and the time) and preferences heterogeneities (both through the spatial heterogeneity of the disutility of pollution w and through the spatial heterogeneity of the elasticity of intertemporal substitution).

This is clear for example in the optimal spatiotemporal path of investment given in the corollary above: it depends directly on technological efficiency both at production and at depollution ($a(t, x)$ and $\eta(t, x)$ respectively), and it depends also indirectly through the α function (and equation (13)) on other local conditions like self-cleaning capacity of Nature ($\delta(x)$), proenvironmental awarness ($w(x)$) and pollution diffusion ($\sigma(x)$). Last but not least, as any optimal outcome of an intertemporal problem, investment also depends on the typical inherent parameters, the time discount parameter (ρ) and the elasticity of intertemporal substitution (through $\gamma(t, x)$), which happens to be time and space-dependent in our framework.

Another remarkable implication of our setting shows up in the optimal dynamics of pollution, equation (20). As one can see, the technological determinants of the latter entirely play through a single "sufficient" statistic, the ratio $\frac{a(t,x)-1}{\eta(t,x)}$. Of course, the larger productivity at production, the larger pollution accumulation, while depollution efficiency works in the opposite direction. In our model, the two technological channels interact directly in the expression of optimal pollution stock dynamics. A technological evolution in which the dynamics of productivity at production (here captured by the dynamics of $a(t, x) - 1$) are one-to-one offset by those of depollution will leave the law of motion of pollution stock unaffected, and therefore entirely determined by non-technological factors, which are far from innocuous. One of the virtues of our framework is precisely to highlight the role of these non-technological factors. In particular, all the factors affecting pollution absorption capacity (that is, $\delta(x)$) are increasingly important over time: they point at the crucial importance of land use policies over the world (not only at the Amazon).

As in study in sustainable development under pollution, we should not only look at short-term effect but also at the large. We show here below that our setting also allows for extracting some interesting asymptotic convergence results in the same analytical manner.

3.4. **Transitional and long-run analysis of P^*.** In this section we analyze both the transitional dynamics of $P^*(t)$ and its limit behavior as $t \to \infty$.

3.4.1. *Transitional dynamics through series expansion.* Recall that a non identically zero function $\phi \in \mathcal{D}(\mathcal{L})$ is called *eigenfunction* of $\mathcal{L}$ if there exists a real number (called *eigenvalue*) λ such that $\mathcal{L}\phi = \lambda\phi$.

Proposition 3.8. *Let Assumption 3.1 hold. There exists a decreasing sequence* $\{\lambda_n\}_{n\in\mathbb{N}} \subset (-\infty, 0]$ *such that* $\lambda_n \to -\infty$ *and an orthonormal basis* $\{\mathbf{e}_n\}_{n\in\mathbb{N}} \subset H$ *such that*

$$\mathbf{e}_n \in \mathcal{D}(\mathcal{L}) \quad \textit{and} \quad \mathcal{L}\mathbf{e}_n = \lambda_n \mathbf{e}_n \quad \forall n \in \mathbb{N}.$$

Then the pollution profile along the optimal trajectory can then be expressed as a convergent series in H*:*

$$P^*(t) = \sum_{n\in\mathbb{N}} p_n^*(t)\mathbf{e}_n, \qquad \textit{where} \;\; p_n^*(t) := \langle P^*(t), \mathbf{e}_n \rangle \tag{23}$$

and the expressions of the coefficients $p_n^*(t)$ *can be explicitly given in the following form*

$$p_n^*(t) := \langle p_0, \mathbf{e}_n \rangle e^{\lambda_n t} + \int_0^t e^{\lambda_n (t-s)} \xi_n^*(s) \mathrm{d}s, \qquad \forall t \geq 0, \; \forall n \in \mathbb{N}, \tag{24}$$

where

$$\xi_n^*(t) := \langle \Psi(t) I^*(t), \mathbf{e}_n \rangle, \qquad \forall t \geq 0, \; \forall n \in \mathbb{N}. \tag{25}$$

Proof. See Appendix A. □

The previous result allows to express the solution of the equation (21) along the optimal path in terms of a series of space functions not dependent on time multiplied by time-dependent coefficients. This expression, which can be used in general to simulate the model, takes a particularly familiar form in some specific cases. For example, if the diffusivity σ and the natural decay of pollution δ are constant in the space variable, a standard Fourier series is obtained, and the functions $\mathbf{e}_n$ are sins and cosines.

Observe also that (23) can also be used to express the total pollution $\int_{S^1} p^*(t,x)dx$ as a function of time, namely,

$$\int_{S^1} p^*(t,x)dx = \sum_{n\in\mathbb{N}} \langle \mathbf{e}_n, \mathbf{1}\rangle \, p_n^*(t), \qquad t \geq 0.$$

where $p_n^*(t)$ is given in (24). Indeed, again using this result, an even more precise description of the total pollution at time t can be given whenever the following when δ is constant in space as shown by the following proposition.

Proposition 3.9. *Let Assumption 3.1 hold and assume that the function δ is constant, i.e. $\delta(\cdot) \equiv \delta_0 \geq 0$. Then*

$$(26) \quad \int_{S^1} p^*(t,x)\mathrm{d}x = \left(\int_{S^1} p_0(x)\mathrm{d}x\right) e^{-\delta_0 t} + \int_0^t e^{-\delta_0(t-s)} \left(\int_{S^1} \eta(s,x) i^*(s,x)\mathrm{d}x\right) \mathrm{d}s.$$

Proof. See Appendix A. □

3.4.2. *Limit behaviour in the time-homogeneous case.* We consider now the special case when the productivity coefficients and the ecological efficiency of the production process are time-independent: $a(t,x) = a(x)$, $\eta(t,x) = \eta(x)$; similarly for the inverse of the elasticity of intertemporal substitution, i.e. $\gamma(t,x) = \gamma(x)$. In this case, the expressions of the optimal control are time independent

$$(27) \qquad I^*(t)(x) \equiv \bar{I}^*(x) = (\eta(x)\alpha(x))^{-\frac{1}{\gamma(x)}} (a(x)-1)^{\frac{1-\gamma(x)}{\gamma(x)}},$$

and we have a direct characterization of the long-run profile of the pollution stock along the optimal path as described in the following proposition.

Proposition 3.10. *Let Assumption 3.1 hold. Assume that the coefficients a, γ, η are time-independent and that $\delta \not\equiv 0$. Then we have*

$$\lim_{t\to\infty} P^*(t) = p^*_\infty \quad \text{in } H,$$

where p^*_∞ is the unique solution to the ODE

$$(28)\quad \begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}x}\left(\sigma(x)\dfrac{\mathrm{d}p^*_\infty}{\mathrm{d}x}(x)\right) - \delta(x)p^*_\infty(x) + \alpha(x)^{-\frac{1}{\gamma(x)}}\left(\dfrac{a(x)-1}{\eta(x)}\right)^{\frac{1}{\gamma(x)}-1} = 0, \\ p^*_\infty(0) = p^*_\infty(2\pi), \quad \dfrac{\mathrm{d}p^*_\infty}{\mathrm{d}x}(0) = \dfrac{\mathrm{d}p^*_\infty}{\mathrm{d}x}(2\pi). \end{cases}$$

Proof. See Appendix A. □

4. An illustration: Technology evolution over space and time

We now propose an application of our method to technology diffusion across space and time. In our setting, two technological exogenous variables are considered: one represents the typical efficiency of the production process in the final good sector, that is $a(t, x)$, and the second reflects essentially the advances in the depollution activity, as reflected by $\eta(t, x)$. We will concentrate on the former for simplicity, given the illustrative purpose of this section. We could have performed similar work on the latter. More precisely, we look at the economic and environmental implications of spatiotemporal diffusion of technology using the North-South diffusion advocated by Nocco (2005).[8]

Consider an economy with a technological core (say the "North") at $x = \pi$, which is therefore the technological center. Concretely, we assume an initial distribution for productivity in the final sector, $a(0, x)$, with a peak at $x = \pi$.[9] As argued above, we suppose that knowledge spillovers from the

[8]Again, we could have chosen alternative specifications for technology diffusion. See a short survey in the 2005 Nocco's paper. The important common point to these specifications is that distance to the technological center matters in the strength of spillovers, which is an empirically established feature, as outlined in the introduction.

[9]These initial distributions are represented in the second graphics of Figures 1, 2, 3. To make the point we keep the same initial distribution in all our simulations.

technological center are more difficult for peripheral locations. We capture this fact by defining a *learning capabilities* function ψ with the following linear form

$$\psi(x) = \begin{cases} a - b|x-\pi| & \text{if } |x-\pi| < a/b \\ 0 & \text{if } |x-\pi| \geq a/b \end{cases} \tag{29}$$

where a and b are two non-negative constants and $|x-\pi|$ is the distance between x and π on S^1.

We normalize, at all times, the level of the technology in the center π to a certain fixed level $a_\pi > 1$,[10] and following Nocco (2005), we suppose that the dynamics of the technology at any point x has the form[11]

$$\dot{a}_x(t) = (a_x(t) - a_\pi)^3 + \psi(x)(a_\pi - a_x(t)) \tag{30}$$

with the initial value $a_x(0)$ such that $1 < a_x(0) < a_\pi$. This equation is made of two parts. The first part $(a_x(t) - a_\pi)^3$ describes the fact that, in absence of knowledge spillovers, the differences in technology between the center and the periphery tends to increase due to the accumulation process. The second part $\psi(x)(a_\pi - a_x(t))$ represents knowledge spillovers which are higher when technological differences are larger and depends on the distance with the center through the learning capabilities function ψ introduced above.

We represent three possible dynamics of the system depending on the strength and on the form of the learning capability function and then on the knowledge spillovers intensity. The choice of other parameters is the same in all the figures. We keep the value of various parameters constant constant in space to emphasize the (possible) endogenous divergence dynamics driven by technological accumulation process. Their values are the following: the time discount ρ is 0.03, the diffusion coefficient σ is 0.5, the

[10]It is $a_\pi = 1.2$ in the numerical simulations.

[11]The solution to (30) might become smaller than 1 depending on the value of $\psi(x)$; to be consistent with our results, we bound the Nocco's construction in order to remain in the range allowed by our hypotheses, in the illustrations which follow, we actually choose a to be the maximum among a given constant belonging to the interval $(1, a_\pi)$ and the solution of (30).

nature local regeneration capacity δ is 0.2, the pollution impact of emissions η is constant in time (and space) and equal to 1 as the local environmental awareness w. Finally the inverse of the elasticity of intertemporal substitution in consumption γ is time and space independent and equal to 0.5. The values of these parameters do of course matter in the shape of the optimal spatiotemporal dynamics calculated, but our here aim is to compare the outcomes of three different learning capabilities schemes, and this comparison is unaffected by the latter parameters' choices.

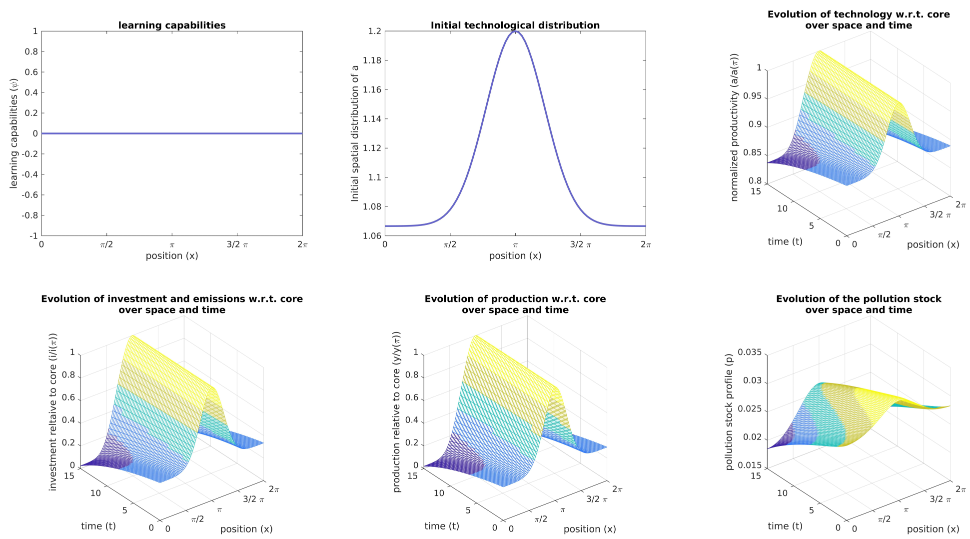

FIGURE 1. The no knowledge spillover case. Evolution of technological level, investment and output (with respect to the technological center) and pollution stock in time and space.

In Figure 1 we study the no-learning benchmark. In this case we take a (and b) equal to 0 in (29) so that no knowledge spillovers is at work. This specific form of ψ is represented in the first image of Figure 1. As already said, the second graphic represents the initial spatial distribution of the technology. The third graphic displays the evolution of the technological level with respect to the center's, that is for any location x, the ratio $a_x(t)/a_\pi$ where $a_x(t)$ follows the dynamics described by (30). In the absence of spillover effects a strong divergence dynamics arises and the relative distance among technologies in the core and in the periphery increases over time. This kind of behavior is reproduced in the investment and output dynamics (fourth and fifth graphics of Figure 1). The sixth graphic of Figure 1 gives the dynamics of the stock of pollution[12]. Not surprisingly the regions where input and then emissions are higher are the more polluted and differences accentuate over time due to the divergence process in input (and then emission) among locations over time. Observe that since pollution is transboundary, the relative difference between core and periphery in terms of pollution stock is milder than the relative difference of emissions.

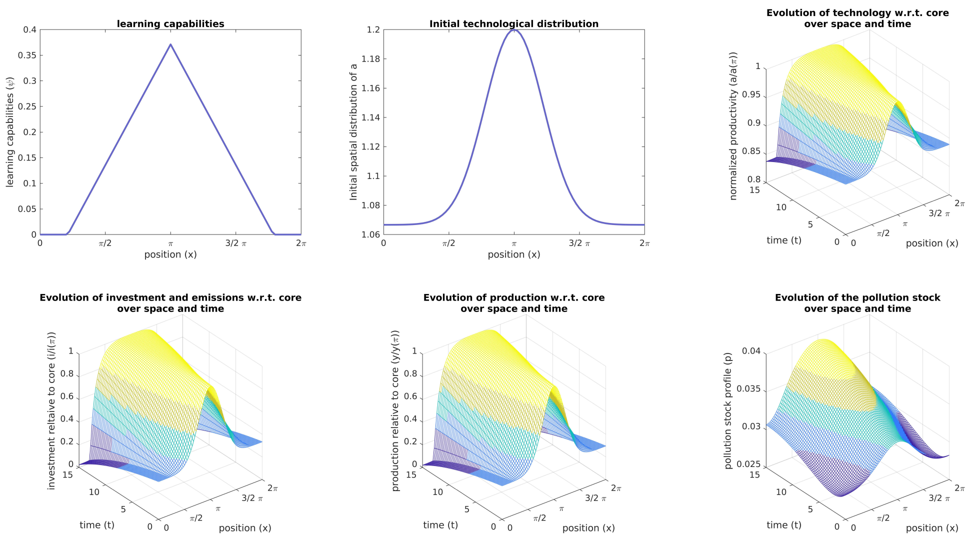

FIGURE 2. The low knowledge spillover case. Evolution of technological level, investment and output (with respect to the technological center) and pollution stock in time and space.

In Figure 2 we have a *low-spillover scenario*, a situation where the learning/spillover dynamics is active but it is not very strong. In particular, for regions which are too far from the core the learning function ψ turns to be zero. More precisely in the definition of ψ given in (29) we fix now $a = 0.371$ and b equal to 0.15. The corresponding learning function is represented in the first graphic of Figure 2. We observe that the dynamics is now qualitatively different compared to the benchmark. In some locations,

[12]The initial datum for the pollution stock profile is found by considering an history of past emissions always equal to the initial emission profile appearing in the fifth graphic.

the magnitude of learning capabilities, ψ, and the initial technological level are high enough to give rise to a convergence dynamics. Other regions, in particular the more peripheral ones with zero learning capabilities, remain far from the frontier, ultimately experiencing divergence from the center. So differently from the benchmark, we have a different picture: some of the non-core locations converge to the center, others no.

In Figure 3 we consider the *high-spillover scenario*: knowledge spillover are strong in all locations.[13] Now the knowledge spillovers are strong enough to offset the divergent forces at work in all locations. We see this fact in the spatial dynamics of all the variables: relative level of technology, input and production and also in the dynamics of the spatial profile of the pollution stock which, in the long run converges toward the same value in all the regions. The same set-up is replicated in Figure 4 where we look at the dynamics for a longer time interval.

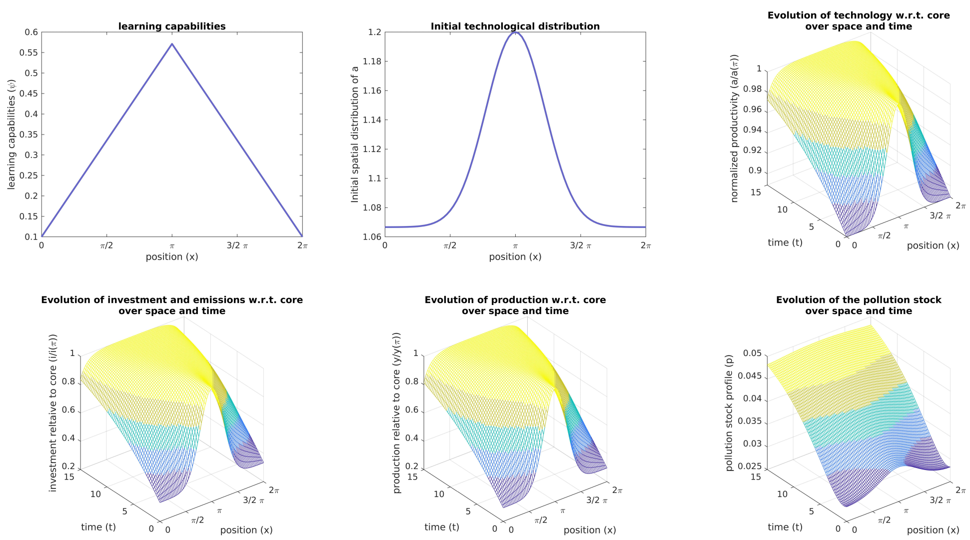

FIGURE 3. The high knowledge spillover case. Evolution of technological level, investment and output (with respect to the technological center) and pollution stock in time and space.

[13]We set $a = 0.571$ and b equal to 0.15, the corresponding learning function is represented in the first graphic of Figure 3) as before.

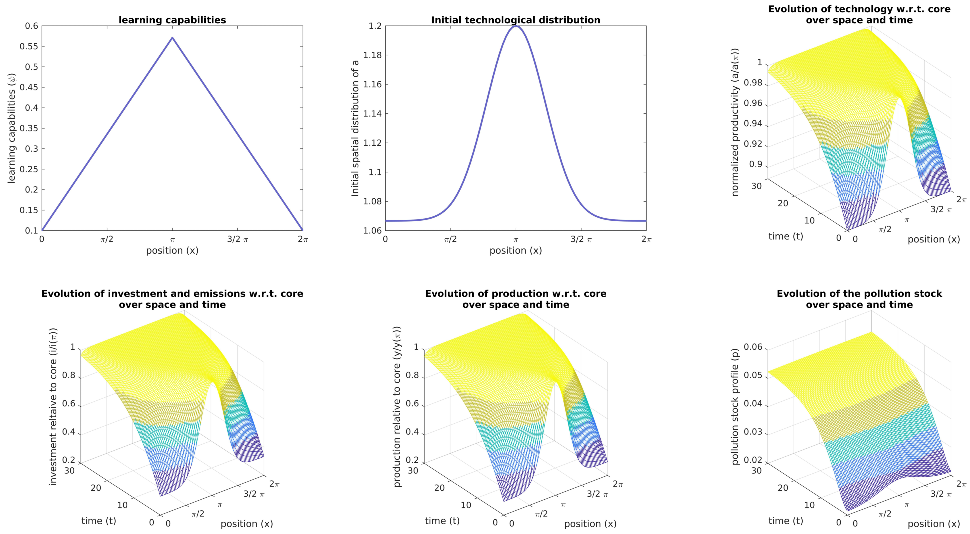

FIGURE 4. The high knowledge spillover case with a longer time evolution.

5. CONCLUSION

In this paper, we have provided with a generic spatiotemporal non-linear-quadratic framework for transboundary pollution control. The objective functional to be maximized is a Benthamite social welfare function depending on the intertemporal stream of consumption at any location, and internalizing the spatial externalities resulting from pollution diffusion. The essential contribution of this work is to identify optimal pollution control policies with a very large account of geographic heterogeneity: (i) heterogeneity in productivity and in ecological efficiency of the production process, which also includes the broad spatio-temporal characteristics of the exogenous technological process; (ii) heterogeneity in preferences, notably in the intertemporal elasticity of substitution and in the disutility from the pollution, and finally: (iii) the heterogeneity in the environmental/ecological context, in particular in terms of speed of diffusion of pollutants and local regeneration capacity.

Despite the complexity of the problem, we have been able to produce a solution method which has two unexpected virtues (given the complexity of the task). First, it allows for closed-form solutions, and second, the solutions produced are based on a neatly singled out spatial function with a clear economic interpretation. We do believe that such a framework can be used in a large set of applications given the generality of most of the specifications. Clearly, one can still visualize a number of possible future extensions (for example the incorporation of demographic dynamics with space-dependent mortality depending on local pollution) but we firstly believe that the next step should be the exploitation of the variety of applications allowed by this framework.

References

Arnott, R., Hochman, O. and Rausser, G. (2008). Pollution and land use: Optimum and decentralization. Journal of Urban Economics, 64, 390-407.

Augeraud-Véron, E., Boucekkine, R. and Veliov, V. M. (2019a). Distributed optimal control models in environmental economics: a review. Mathematical Modelling of Natural Phenomena, 14(1), Article 106.

Augeraud-Véron, E., Choquet, C. and Comte, E. (2017). Optimal control for a groundwater pollution ruled by a convection–diffusion–reaction problem. Journal of Optimization Theory and Applications, 173(3), 941-966.

Augeraud-Véron, E., Choquet, C. and Comte, E. (2019). Existence, uniqueness and asymptotic analysis of optimal control problems for a model of groundwater pollution. ESAIM: Control, Optimisation and Calculus of Variations, 25, n. 53.

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

Barro R. and Sala-i-Martin X. (2004). *Economic Growth.* 2nd edition. Cambridge: MIT Press.

Barucci, E. and Gozzi, F. (2001). Technology adoption and accumulation in a vintage-capital model. Journal of economics, 74(1), 1-38.

Bensoussan A., Da Prato G., Delfour M.C. and Mitter S.K. (2006). *Representation and Control of Infinite Dimensional Systems.* 2nd edition. Basel: Birkhäuser.

Boucekkine R., Fabbri G., Federico S. and Gozzi F. (2020). From firm to global-level pollution control: the case of transboundary pollution. Working Paper n. 818, Universita di Siena.

Boucekkine, R., Fabbri, G., Federico, S. and Gozzi, F. (2019a). Geographic environmental Kuznets curves: The optimal growth linear-quadratic case. Mathematical Modelling of Natural Phenomena, 14(1), 105.

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

Boucekkine R., Krawczyk J. and Vallée T. (2011). Environmental quality versus economic performance: A dynamic game approach. Optimal Control Applications and Methods, 32, 29-46.

Brezis H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. New York: Springer.

Brito, P. (2004). The dynamics of growth and distribution in a spatially heterogeneous world. Working Papers Department of Economics 2004/14. ISEG, University of Lisbon

Camacho C. and Perez Barahona A. (2015). Land use dynamics and the environment. Journal of Economic Dynamics and Control, 52, 96-118.

Crandall, M. G., Ishii, H. and Lions, P. L. (1992). User's guide to viscosity solutions of second order partial differential equations. Bulletin of the American mathematical society, 27(1), 1-67.

Coddington, E. A. and Levinson, N. (1955). Theory of ordinary differential equations. New York: McGraw-Hill.

Deltas, G., Karkalakos S. (2013). Similarity of R&D activities, physical proximity and R&D spillovers. Regional Science and Urban Economics, 43, 124-131.

Dockner, E. J. and Van Long, N. (1993). International pollution control: cooperative versus noncooperative strategies. Journal of Environmental Economics and Management, 25(1), 13-29.

Engel K.J., Nagel R. (1995). *One-parameter Semigroups for Linear Evolution Equations.*, Graduate Texts in Mathematics 194, Berlin: Springer.

de Frutos F., Martin-Herran G. (2019). Spatial vs. non-spatial transboundary pollution control in a class of cooperative and non-cooperative dynamic games. European Journal of Operational Research, 276(1), 379-394

Fabbri, G. (2016). Geographical structure and convergence: A note on geometry in spatial growth models. Journal of Economic Theory, 162(1), 114-136.

Freni G., Gozzi F., Salvadori N., (2006). Existence of optimal strategies in linear multisector models. Economic Theory, 29, 25–48.

Gibson M., Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. Journal of Urban Economics, 89, 62-73.

Grass, D. and Uecker, H. (2015). Optimal management and spatial patterns in a distributed shallow lake model. Electronic Journal of Differential Equations, 2017(01), 1-21.

Henderson, V. (1977). Externalities in a spatial context. Journal of Public Economics, 7, 89-110.

Henderson, V. (1996). Effects of air quality regulation. American Economic Review, 86, 789-813.

Hochman, O., Ofek, H. (1979). A theory on the behavior of municipal governments: The case of internalizing pollution externalities. Journal of Urban Economics, 6, 416-431.

Krugman, P. (1993). First nature, second nature, and the metropolitan location. Journal of Regional Science, 33(2), 129-144.

Lunardi A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Basel: Birkhäuser.

Nocco A. (2005). The rise and fall of regional inequalities with technological differences and knowledge spillovers. Regional Science and Urban Economics, 35, 542-569.

Tiwary A., Colls J. (2010), *Air Pollution.* Abingdon-on-Thames: Routledge.

Appendix A. Proofs

Proof of Proposition 3.2. Due to Assumption 3.1, $\mathcal{L}$ is a closed, densely defined, unbounded linear operator on the space H (see, e.g. Lunardi, 1995, p. 71-75, Sections 3.1 and 3.1.1). A core for it is the space $C^\infty(S^1;\mathbb{R})$ (see, e.g., Engel and Nagel, 1995, pages 69-70). Let $\varphi \in C^\infty(S^1;\mathbb{R})$. Integration by parts yields

$$\begin{aligned} \langle \mathcal{L}\varphi, \varphi\rangle &= \int_{S^1} (\mathcal{L}\varphi)(x)\varphi(x)\mathrm{d}x \\ &= -\int_{S^1} \sigma(x)|\varphi'(x)|^2\mathrm{d}x - \int_{S^1} \delta(x))|\varphi(x)|^2\mathrm{d}x \le 0 \end{aligned} \tag{31}$$

Since $C^\infty(S^1;\mathbb{R})$ is a core for $\mathcal{L}$, (32) extends to all functions $\varphi \in D(\mathcal{L})$, showing that the operator $\mathcal{L}$ is dissipative. Similarly, a double integration by parts shows that

$$\langle \mathcal{L}\varphi, \eta\rangle = \langle \varphi, \mathcal{L}\eta\rangle, \quad \forall \varphi, \eta \in C^\infty(S^1;\mathbb{R}). \tag{32}$$

Again, since $C^\infty(S^1;\mathbb{R})$ is a core for $\mathcal{L}$, (32) extends to all couples of functions in $D(\mathcal{L})$, showing that $\mathcal{L}$ is self-adjoint, i.e. $\mathcal{L} = \mathcal{L}^\star$, where $\mathcal{L}^\star$ denotes the adjoint of $\mathcal{L}$. Therefore, by Engel and Nagel (1995) (in particular, Chapter II), $\mathcal{L}$ generates a strongly continuous contraction semigroup $(e^{t\mathcal{L}})_{t\ge 0} \subset L(H)$; in particular, since $\rho > 0$, by standard theory of strongly continuous semigroup in Banach spaces (see, e.g. pages 82-83, Chapter II and Theorem 1.10, Chapter II of Engel and Nagel 1995), it follows that ρ belongs to the resolvent set of $\mathcal{L}$ and that (7) holds. □

Proof of Proposition 3.3. First of all we observe that, by Assumption 3.1(ii), the first term in the functional is always positive (Case (A)) or always negative (Case (B)), possibly infinite. Hence to prove the claim it is enough to show that, given any $p_0 \in H$, the term $\int_0^\infty e^{-\rho t}\langle w, P(t)\rangle \mathrm{d}t$ is well defined and finite for every $I \in \mathcal{A}$. We have

$$\int_0^\infty e^{-\rho t}\langle w, P(t)\rangle \mathrm{d}t = \int_0^\infty e^{-\rho t}\Big\langle w, e^{t\mathcal{L}}p_0 + \int_0^t e^{(t-s)\mathcal{L}}\Psi(s)I(s)\mathrm{d}s\Big\rangle \mathrm{d}t \tag{33}$$

Now, since w is bounded and $e^{t\mathcal{L}}$ is a contraction, the integral $\int_0^\infty e^{-\rho t}\langle w, e^{t\mathcal{L}}p_0\rangle \mathrm{d}t$ is finite. Moreover, for all $T > 0$ we get, by Fubini-Tonelli's Theorem

$$\begin{aligned} &\int_0^T \left(\int_0^t e^{-\rho t}\left\langle w, e^{(t-s)\mathcal{L}}\Psi(s)I(s)\right\rangle \mathrm{d}s\right)\mathrm{d}t \\ &= \int_0^T \left(\int_0^t e^{-\rho s}\left\langle w, e^{-(\rho-\mathcal{L})(t-s)}\Psi(s)I(s)\right\rangle \mathrm{d}s\right)\mathrm{d}t \\ &= \int_0^T e^{-\rho s}\left\langle w, \int_s^T e^{-(\rho-\mathcal{L})(t-s)}\Psi(s)I(s)\mathrm{d}t\right\rangle \mathrm{d}s \end{aligned}$$

Using again the fact that $e^{(t-s)\mathcal{L}}$ is a contraction and Assumption 3.1, we have, for each $s \ge 0$, $T \ge 0$

$$\left\| \int_s^T e^{-(\rho-\mathcal{L})(t-s)}\Psi(s)I(s)\mathrm{d}t \right\| \le \int_s^\infty e^{-\rho(t-s)}\|\Psi(s)I(s)\|\mathrm{d}t \le \frac{1}{\rho}\|\Psi(s)I(s)\|.$$

Hence, by definition of $\mathcal{A}$, the claim follows sending T to $+\infty$. □

Proof of Proposition 3.4. The fact that α solves (13) and the fact that, by Assumption 3.1, we have $\sigma(\cdot) > 0$ yield

$$\alpha''(x) = \frac{1}{\sigma(x)} \left[(\rho + \delta(x))\alpha(x) - \sigma'(x)\alpha'(x) - w(x)\right], \quad \text{for a.e. } x \in S^1.$$

Since $\alpha \in C^1(S^1;\mathbb{R})$, it follows, by Assumption 3.1, that $\alpha \in C^2(S^1;\mathbb{R})$.

Now, let $x_* \in S^1$ be a minimum point of α over S^1. Then $\alpha''(x_*) \geq 0$. Plugging this into (13) we get

$$(\rho + \delta(x_*))\alpha(x_*) = \sigma(x_*)\alpha''(x_*) + w(x_*) \geq w(x_*),$$

and the estimate from below follows. The estimate from above can be obtaind symmetrically. □

Proof of Proposition 3.5. Case $\sigma \to 0^+$. First, notice that under the above assumptions (13) reads as

$$\rho\alpha_{\sigma^o}(x) - \sigma^o\alpha''_{\sigma^o}(x) + \delta(x)\alpha_{\sigma^o}(x) = \widehat{w}(x), \quad x \in S^1, \tag{34}$$

By Proposition 3.4 we have

$$\alpha_*(x) := \liminf_{\overline{\sigma^o}\to 0^+} \left\{\alpha_{\sigma^o}(\zeta) : \ \sigma^o \leq \overline{\sigma^o},\ \zeta \in S^1,\ |\zeta - x| \leq 1/\overline{\sigma^o}\right\} \geq \min_{S^1} \frac{w}{\rho+\delta},$$

$$\alpha^*(x) := \limsup_{\overline{\sigma^o}\to 0^+} \left\{\alpha_{\sigma^o}(\zeta) : \ \sigma^o \leq \overline{\sigma^o},\ \zeta \in S^1,\ |\zeta - x| \leq 1/\overline{\sigma^o}\right\} \leq \max_{S^1} \frac{w}{\rho+\delta}.$$

Clearly $\alpha^* \geq \alpha_*$. By stability of viscosity solutions (see e.g. Crandall et al., 1992), the latter functions are, respectively, (viscosity) super- and sub-solution to the limit equation

$$\rho\alpha_0(x) + \delta(x)\alpha_0(x) = w(x),$$

whose unique solution is

$$\alpha_0(x) = \frac{w(x)}{\rho + \delta(x)}.$$

By standard comparison of viscosity solutions one has $\alpha_* \geq \alpha_0 \geq \alpha^*$. It follows that

$$\exists \lim_{\sigma^o\to 0^+} \alpha_{\sigma^o}(x) = \alpha_*(x) = \alpha^*(x) = \alpha_0(x) \quad \forall x \in S^1.$$

Case $\sigma \to +\infty$. First, we rewrite (34) as

$$\alpha''_{\sigma^o}(x) = \frac{1}{\sigma^o}\left[\rho\alpha_{\sigma^o}(x) + \delta(x)\alpha_{\sigma^o}(x) - w(x)\right], \quad x \in S^1, \tag{35}$$

From this and from Proposition 3.4, we see that α_{σ^o} is equi-bounded and equi-uniformly continuous with respect to $\sigma^o \geq 1$. Hence, by Ascoli-Arzelà Theorem we have that, from each sequence $\sigma_n \to +\infty$, we can extract a subsequence σ_{n_k} such that

$$\lim_{k\to+\infty} \alpha_{\sigma_{n_k}} = \alpha_\infty \quad \text{uniformly on } x \in S^1,$$

for some $\alpha_\infty \in C(S^1;\mathbb{R})$. Again by stability viscosity solutions, α_∞ must solve, in the viscosity sense, the limit equation

$$\alpha''_\infty(x) = 0, \qquad x \in S^1.$$

Hence, it must be $\alpha_\infty \equiv c_0$ for some $c_0 \geq 0$. To find the value of c_0 we may integrate (34) over S^1 getting

$$\int_{S^1} (\rho + \delta(x))\alpha_{\sigma^o}(x)\mathrm{d}x = \int_{S^1} w(x)\mathrm{d}x.$$

Letting $\sigma^o \to +\infty$ above, we get

$$c_0 = \frac{\int_{S^1} w(x)\mathrm{d}x}{\int_{S^1}(\rho + \delta(x))\mathrm{d}x}.$$

As this value does not depend on the sequence σ_n chosen, the claim follows. $\square$

Proof of Theorem 3.6. (i) Using (9) it is possible to rewrite the second part of (10). We first set

$$e^{-(\rho-\mathcal{L})t} := e^{-\rho t}e^{t\mathcal{L}}, \quad t \geq 0,$$

and we write

$$\begin{aligned}
\int_0^\infty e^{-\rho t}\langle w, P(t)\rangle \mathrm{d}t &= \int_0^\infty e^{-\rho t}\Big\langle w, e^{t\mathcal{L}}p_0 + \int_0^t e^{(t-s)\mathcal{L}}\Psi(s)I(s)ds\Big\rangle \mathrm{d}t \\
&= \left\langle w, \int_0^\infty e^{-(\rho-\mathcal{L})t}p_0\, \mathrm{d}t\right\rangle + \int_0^\infty e^{-\rho t}\left\langle w, \int_0^t e^{(t-s)\mathcal{L}}\Psi(s)I(s))\mathrm{d}s\right\rangle \mathrm{d}t
\end{aligned} \tag{36}$$

Note that the first term of the right hand side is the only one which depends on the initial datum and, by (7), it can be rewritten as

$$\left\langle w, \int_0^\infty e^{-(\rho-\mathcal{L})t}p_0\, \mathrm{d}t\right\rangle = \big\langle w, (\rho-\mathcal{L})^{-1}p_0\big\rangle = \big\langle (\rho-\mathcal{L})^{-1}w, p_0\big\rangle = \langle \alpha, p_0\rangle, \tag{37}$$

where α is defined in (11).

We look now at the last term of the last line of (36). It can be rewritten by exchanging the integrals as follows:

$$\begin{aligned}
&\int_0^\infty \left(\int_0^t e^{-\rho t}\left\langle w, e^{(t-s)\mathcal{L}}\Psi(s)I(s)\right\rangle \mathrm{d}s\right)\mathrm{d}t \\
&= \int_0^\infty \left(\int_0^t e^{-\rho s}\left\langle w, e^{-(\rho-\mathcal{L})(t-s)}\Psi(s)I(s)\right\rangle \mathrm{d}s\right)\mathrm{d}t \\
&= \int_0^\infty e^{-\rho s}\left\langle w, \int_s^\infty e^{-(\rho-\mathcal{L})(t-s)}\Psi(s)I(s)dt\right\rangle \mathrm{d}s \\
&= \int_0^\infty e^{-\rho s}\big\langle w, (\rho-\mathcal{L})^{-1}\Psi(s)I(s)\big\rangle \mathrm{d}s \\
&= \int_0^\infty e^{-\rho s}\big\langle (\rho-\mathcal{L})^{-1}w, \Psi(s)I(s)\big\rangle \mathrm{d}s
\end{aligned}$$

Hence, we can finally rewrite (10) as (17).

(ii) After writing explicitely the inner products in (10), the integral can be optimized pointwisely. We end up, for $(t,x) \in \mathbb{R}^+ \times S^1$ fixed, with the optimization

$$\sup_{\iota \geq 0} \left\{ \frac{\big((a(t,x)-1)\iota\big)^{1-\gamma(t,x)}}{1-\gamma(t,x)} - \alpha(x)\eta(t,x)\iota \right\};$$

so, we easily get the claimed expression (18) of the candidate unique optimal control I^*. On the other hand, we need to verify that $I^* \in \mathcal{A}$. Indeed, by Assumption 3.1, we have

$$\begin{aligned}\Psi(t)(x)I^*(t)(x) &= \alpha(x)^{-\frac{1}{\gamma(t,x)}}\eta(t,x)^{1-\frac{1}{\gamma(t,x)}}(a(t,x)-1)^{\frac{1}{\gamma(t,x)}-1} \\ &= \alpha(x)^{-\frac{1}{\gamma(t,x)}}\left(\frac{a(t,x)-1}{\eta(t,x)}\right)^{\frac{1}{\gamma(t,x)}-1}.\end{aligned}$$

Since α is bounded from above and from below by positive constants, then so is $\alpha(x)^{-\frac{1}{\gamma(t,x)}}$ by Assumption 3.1(ii). Consequently, by Assumption 3.1(iii), we get, for some $C_0 > 0$

$$0 \leq \Psi(t)(x)I^*(t)(x) \leq C_0 e^{gt}, \quad \forall x \in S^1.$$

Since $\rho > g$ by Assumption 3.1(iv), we get $I^* \in \mathcal{A}$.

(iii)-(iv) These claims immediately follow by straightforward computations. □

Proof of Proposition 3.8. Consider the operator $(1-\mathcal{L})^{-1} \in L(H)$. The range of this operator is the space $D(\mathcal{L}) = W^{2,2}(S^1;\mathbb{R})$, which by Kondrachov's Theorem is embedded in $H = L^2(S^1;\mathbb{R})$ with compact embedding. It follows that $(1-\mathcal{L})^{-1}$ is compact; being also self-adjoint, by standard spectral theory in Hilbert spaces, there exists an orthonormal basis of eigenvectors for it, hence also for $\mathcal{L}$. Hence, considering also that $\mathcal{L}$ is dissipative and unbounded, there exists a decreasing sequence $\{\lambda_n\}_{n\in\mathbb{N}} \subset (-\infty, 0]$ such that $\lambda_n \to -\infty$ and an orthonormal basis $\{\mathbf{e}_n\}_{n\in\mathbb{N}} \subset H$ such that

$$\mathbf{e}_n \in \mathcal{D}(\mathcal{L}) \quad \text{and} \quad \mathcal{L}\mathbf{e}_n = \lambda_n \mathbf{e}_n \quad \forall n \in \mathbb{N}. \tag{38}$$

Consider the Fourier series expansion

$$P^*(t) = \sum_{n\in\mathbb{N}} p_n^*(t)\mathbf{e}_n, \qquad \text{where} \;\; p_n^*(t) := \langle P^*(t), \mathbf{e}_n\rangle.$$

We can write explicitly the Fourier coefficients $p_n^*(t)$ by the following argument. By Proposition 3.2, Chapter 1, Part II of Bensoussan et al. (2007), the function P^* defined in (19) is also a *weak solution* to (8) with $I = I^*$, i.e., taking into account that $\mathcal{L}$ is self-adjoint, i.e. $\mathcal{L} = \mathcal{L}^\star$, it holds

$$\langle P^*(t), \varphi\rangle = \langle p_0, \varphi\rangle + \int_0^t \Big(\langle P^*(s), \mathcal{L}\varphi\rangle + \langle \Psi(s)I^*(s), \varphi\rangle\Big)\mathrm{d}s \quad \forall t \geq 0, \;\; \forall \varphi \in D(\mathcal{L}).$$

In particular, taking into account (38), we have

$$p_n^*(t) = \langle P^*(t), \mathbf{e}_n \rangle = \langle p_0, \mathbf{e}_n \rangle + \int_0^t \Big(\lambda_n \langle P^*(s), \mathbf{e}_n \rangle + \langle \Psi(s) I^*(s), \mathbf{e}_n \rangle \Big) \mathrm{d}s$$
$$= \langle p_0, \mathbf{e}_n \rangle + \int_0^t \Big(\lambda_n p_n^*(s) + \xi_n^*(s) \Big) \mathrm{d}s \qquad \forall t \geq 0, \quad \forall n \in \mathbb{N}.$$

Then

$$\int_{S^1} p^*(t,x) \mathrm{d}x = \langle P^*(t), \mathbf{1} \rangle = \Big\langle \sum_{n \in \mathbb{N}} p_n^*(t) \mathbf{e}_n, \mathbf{1} \Big\rangle = \sum_{n \in \mathbb{N}} \langle \mathbf{e}_n, \mathbf{1} \rangle \, p_n^*(t), \qquad t \geq 0,$$

as claimed. □

Proof of Proposition 3.9. First we observe that, in this case, $\mathbf{e}_0(\cdot) \equiv \frac{1}{\sqrt{2\pi}}$, $\lambda_0 = -\delta_0$. Hence

$$\int_{S^1} p^*(t,x) \mathrm{d}x = \langle P^*(t), \mathbf{1} \rangle = \sqrt{2\pi} \langle P^*(t), \mathbf{e}_0 \rangle.$$

From the mild form of P given in (9) we now get, for $t \geq 0$,

$$\begin{aligned}
\langle P^*(t), \mathbf{e}_0 \rangle &= \langle e^{t\mathcal{L}} p_0, \mathbf{e}_0 \rangle + \int_0^t \langle e^{(t-s)\mathcal{L}} \Psi(s) I^*(s), \mathbf{e}_0 \rangle \mathrm{d}s \\
&= \langle p_0, e^{t\mathcal{L}} \mathbf{e}_0 \rangle + \int_0^t \langle \Psi(s) I^*(s), e^{(t-s)\mathcal{L}} \mathbf{e}_0 \rangle \mathrm{d}s \\
&= \langle p_0, e^{-\delta_0 t} \mathbf{e}_0 \rangle + \int_0^t \langle \Psi(s) I^*(s), e^{-\delta_0 (t-s)} \mathbf{e}_0 \rangle \mathrm{d}s,
\end{aligned}$$

where we used that $e^{-\delta_0 t}$ is the eigenvalue of $e^{t\mathcal{L}}$ associated to $\mathbf{e}_0$. The claim immediately follows. □

Proof of Proposition 3.10. In this case $I^*(\cdot) \equiv \bar{I}^* \in H$ is time independent too. Since $\delta \not\equiv 0$, we have $\lambda_0 < 0$. Let us write

$$\mathcal{L} = \mathcal{L}_0 - \lambda_0, \quad \text{where} \ \ \mathcal{L}_0 := \mathcal{L} + \lambda_0,$$

and note that $\mathcal{L}_0$ is dissipative by definition, hence $e^{s\mathcal{L}_0}$ is a contraction. Then, setting $\bar{\Psi} := \eta(\cdot) \in H$, we can rewrite

$$P^*(t) = e^{\lambda_0 t} e^{t\mathcal{L}_0} p_0 + \int_0^t e^{\lambda_0 (t-s)} e^{(t-s)\mathcal{L}_0} \bar{\Psi} \bar{I}^* \mathrm{d}s = e^{\lambda_0 t} e^{t\mathcal{L}_0} p_0 + \int_0^t e^{\lambda_0 t} e^{t\mathcal{L}_0} \bar{\Psi} \bar{I}^* \mathrm{d}s,$$

and take the limit above when $t \to \infty$. Since $e^{s\mathcal{L}_0}$ is a contraction, the first term of the right hand side converges to 0, whereas the second one converges to

$$P_\infty^* := \int_0^\infty e^{-\lambda_0 s} e^{s\mathcal{L}_0} \bar{\Psi} \bar{I}^* \mathrm{d}s \in H.$$

Then, the limit state $P_\infty^* \in H$ can be expressed using again Proposition 3.14, page 82 and Theorem 1.10, Chapter II of Engel and Nagel (1995) as

$$P_\infty^* = (\lambda_0 - \mathcal{L}_0)^{-1} \bar{\Psi} \bar{I}^*,$$

i.e. P^*_∞ is the solution to

$$(\lambda_0 - \mathcal{L}_0)P^*_\infty = \bar{\Psi}\bar{I}^*,$$

equivalently

$$\mathcal{L}P^*_\infty + \bar{\Psi}\bar{I}^* = 0,$$

i.e., in the PDE formalism, $p^*_\infty(\cdot) := P^*_\infty$ solves (28). □

ACKNOWLEDGEMENTS

The work of Giorgio Fabbri is supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02) and in that of the center of excellence LABEX MME-DII (ANR-11-LABX-0023-01).

R. Boucekkine, IMéRA. 2 Place Le Verrier, 13004 Marseille, France.

Email address: raouf.boucekkine@univ-amu.fr

G. Fabbri, Univ. Grenoble Alpes, CNRS, INRA, Grenoble INP, GAEL - CS 40700 - 38058 Grenoble CEDEX 9, France.

Email address: giorgio.fabbri@univ-grenoble-alpes.fr

S. Federico, Università degli Studi di Siena, Dipartimento di Economia Politica e Statistica. Piazza San Francesco 7-8, 53100 Siena , Italy.

Email address: salvatore.federico@unisi.it

F. Gozzi, Dipartimento di Economia e Finanza, Libera Università degli Studi Sociali *Guido Carli*, Roma

Email address: fgozzi@luiss.it