From firm to global-level pollution control: The case of transboundary pollution

Raouf Boucekkine, Giorgio Fabbri, Salvatore Federico, Fausto Gozzi

Working paper GAEL n° 08/2020

GAEL
Grenoble Applied Economic Laboratory
Consumption – Energy - Innovation

From firm to global-level pollution control:

The case of transboundary pollution

Boucekkine, Raouf
Fabbri, Giorgio
Federico, Salvatore
Gozzi, Fausto

July, 2020

JEL codes: Q53,R12,E60,C61,Q52

https://gael.univ-grenoble-alpes.fr/accueil-gael

contact : agnes.vertier@inrae.fr

FROM FIRM TO GLOBAL-LEVEL POLLUTION CONTROL: THE CASE OF TRANSBOUNDARY POLLUTION

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

Abstract. We study the joint determination of optimal investment and optimal depollution in a spatiotemporal framework where pollution is transboundary. Pollution is controlled at a global level. The regulator internalizes that: (i) production generates pollution, which is bad for the wellbeing of population, and that (ii) pollution flows across space driven by a diffusion process. We solve analytically for the optimal investment and depollution spatiotemporal paths and characterize the optimal long-term spatial distribution when relevant. We finally explore numerically the variety of optimal spatial distributions obtained using a core/periphery model where the core differs from the periphery either in terms of input productivity, depollution efficiency, environmental awareness or self-cleaning capacity of nature. We also compare the distributions with and without diffusion. Key aspects in the optimal policy of the regulator are the role of aversion to inequality, notably leading to smoothing consumption across locations, and the control of diffusive pollution adding another smoothing engine.

Key words: Decision analysis, pollution control, geography, transboundary pollution, infinite dimensional optimal control problems

---

Date: July 21, 2020.

[a] Aix-Marseille University, CNRS, EHESS, Ecole Centrale de Marseille and AMSE (France).

[b] Univ. Grenoble Alpes, CNRS, INRA, Grenoble INP, GAEL, 38000 Grenoble (France).

[c] Dipartimento di Economia Politica e Statistica, Università degli Studi di Siena (Italy).

[d] Dipartimento di Economia e Finanza, LUISS *Guido Carli*, Roma (Italy). Corresponding author.

1. Introduction

Pollution control has been the target of a huge bulk of research in many fields, including management, operational research and economics. In the management literature (see for example the early work of Cohen and Hurter, 1974, and references therein), the minimization of the cost of pollution control at the firm and industry levels has been the object of numerous studies. While the usual direct costs (say those related to equipment control and to adoption of cleaner technologies or inputs) were of course taken into account, this literature has also incorporated into the analysis for a long time less direct factors affecting the profitability of firms and industries. One has to do with the subsequent impact on demand due to rising consumer price indexes (see again Cohen and Hurter, 1974). An even more important indirect cost, according to this terminology, comes from environmental regulation, in particular taxation (see Bawa, 1975). An interesting related question is whether taxation is more efficient compared to direct pollution control instruments, which we label later pollution abatement controls. For example, Bawa (1975) argued that environmental problems, being often characterized by infrequent serious crises, make direct controls more appropriate than tax policy.

In this paper, we do not contribute to this old but still highly relevant question. If any, the most recent trend in the related ongoing debate is the increasingly shared belief that the most urgent environmental problems are global, and in any case hardly restricted to the borders of a city, region or country. This does not disqualify the firm and industry levels because even though the problems are global, actions have to be taken also locally in all respects. For example, while the global warming problem has to do with global CO_2 pollution, ancillary local greenhouse gas mitigation actions might well be useful as excellently argued by Ostrom (2016) (see also Davis et al., 2001). Our point here is that a large part of the serious environmental problems faced can be hardly formulated at the level of a local industry. This is due to the fact that the latter problems derive from pollution processes which are essentially transboundary. This is true for air and water pollution both for global or local pollutant diffusion. Indeed, there is an increasing evidence that an ancillary carbon reduction benefit can be achieved through the introduction of SO_2 control policies, a typical local air pollutant. These issues and many others related have been intensively tackled in the operational research literature, either in applied work (see, e.g., Kerl et al., 2015, on U.S., Xu and Masui, 2009, and Zhao et al., 2013, on China, or the earlier work of Stam et al., 1992, on Europe) or in more stylized settings (see Bertinelli et al., 2014, or de Frutos et al., 2019a,b).

Needless to say, one cannot capture all the essential implications of transboundary pollution without modelling it as precisely as possible. Highly stylized models of transboundary pollution, like those using the two-country setting (see for example, Bertinelli et al., 2014), may capture some of the significant implications of transboundary pollution, but they cannot by construction deliver all the spatiotemporal dynamics inherent in the latter. Here, we rely on an analytical framework developed by Boucekkine et al. (2019a) to inquire about optimal pollution abatement policy over continuous space and continuous time when pollution is transboundary, driven by a diffusion equation. Key contribution of Boucekkine et al. (2019a) is to provide with a spatiotemporal structure allowing for a large set of geographic heterogeneity features while still providing closed-form solutions. We only use some of the heterogeneity traits allowed but, importantly enough, we introduce into the latter framework space and time dependent pollution abatement controls (Boucekkine et al., 2019a, do not consider pollution control policies)[1]. More precisely, we shall consider three types of geographic discrepancy: discrepancy in productivity, in abatement efficiency, and in the nature self-cleaning capacity. Of course, the literature (in particular in Operation Research) is not short of contributions accounting for geographic discrepancy in a context of pollution mitigation. One of the most recent ones, due to Leibowicz (2020), develops a quite compelling theoretical spatial framework for sustainable urban land use and transportation planning, taking into account the damages caused by greenhouse gas emissions. The framework developed is static. Here we are concerned with the spatiotemporal implications of transboundary pollution, and in particular on the inherent long-term sustainabilty issues. Moreover, our extended model allows to generate closed-form spatiotemporal optimal paths, and to identify the (optimal) corresponding long-term geographic distributions.[2]

Importantly enough, we consider a central planner problem, that is the case of a regulator who has to design the optimal spatiotemporal production paths for the whole spatial economy (here space is the unit circle in $\mathbb{R}^2$). More precisely, she has to internalize two facts: (i) production generates pollution, which is bad for the welfare of the inhabitants, and (ii) pollution flows across space driven by a diffusion process. The regulator is given a pollution control instrument to help her to tackle the problem, a direct one we label abatement. It is worth pointing out here that by internalizing (ii), the regulator is actually preventing potential free-riding: as pollution

[1] In a stimulating contribution Camacho and Perez-Barahona (2015) present a spatial model which includes the possibility of abatement in terms of land devoted to pollution absorption.

[2] Typical related exercises run by economists (see the seminal paper by Stokey, 1999) only characterize optimal time paths and the corresponding pointwise stationary solutions.

is transboundary, non-cooperative producers located at different areas may choose the intensity and location of their production facilities in an opportunistic way. We shut down this channel and abstract away from any strategic ingredient (see de Frutos et al., 2019a, 2019b for an excellent exploration of these aspects, and La Torre et al., 2019, for the global vs local dimension of policies.).

Differently from all the operational research contributions already mentioned, our approach to the solution of the dynamic optimization problem does not include the use of dynamic programming or maximum principle methods. Indeed it is based on a reformulation technique of the objective functional where the spatial distribution of the instantaneous disutility of the emissions in the various locations appears (see Section 3 for details).

The functional reformulation approach that we use to solve the problem allows, in particular, to include in our setting a series of spatial heterogeneities which could not be managed in the previous contributions that explicitly model the process of spatio-temporal diffusion of pollution, in particular those of Frutos et al., 2019a, 2019b. In fact, thanks to the linearity of the disutility of pollution in the expression of the agent's payoff, once transformed, the functional objective can be maximized pointwise (in time and space). This fact allows us to solve a general class of problems without having to solve complex infinite-dimensional Hamilton-Jacobi-Bellman equations and finally getting the optimal control "directly" (see in particular Theorem 3.3 and its proof in Appendix A).

Our model allows, in particular, to take account of different levels of productivity in different locations and, moreover, to consider time-dependent productivity parameters and different levels of environmental awareness in different locations. All of them are arguably essential to understand the impact of spatial technological and behavioral discrepancy. As already underlined our model also allows a double channel of decision making: investment and abatement. A more detailed discussion of the properties of our methodology and its applicability for real world explorations is reported in Section 5.

The rest of the paper is organized as follows. Section 2 briefly presents the model. Section 3 delivers the main results, the associated technical and methodological details being relegated in the appendix. Section 4 provides a sample of the optimal spatiotemporal paths generated depending notably on the type of geographic heterogeneity selected. In Section 5 we discuss the results while Section 6 concludes.

2. The model

We now describe briefly the model. We consider a central planning problem of a spatial economy. There is only one good in this economy: it is consumed, used in production as input (or invested), used in pollution abatement, and, of course, it is produced at any location. Furthermore, we postulate that this good is not traded across locations. Only pollution is transboundary. This is of course made for simplification, in order to get the closed-form solutions needed. Modelling trade across locations requires typically another adequate diffusion equation as in Boucekkine et al. (2019b), which would make the problem definitely much more intricate than those treated separately in the two previous papers. As to the spatial support, we choose to work on the circle in $\mathbb{R}^2$:

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

The space structure we use is the simplest possible which allows to treat the problem in a way that the movements of the pollutants cannot push them out of the space. So in the model the total mass of the pollutant can only be reduced by the effect of the decay and of the depollution policies (see below). The set S^1 can be thought as the segment $[0, 2\pi]$ with the identification of the two extremes 0 and 2π. As a sub-manifold of $\mathbb{R}^2$, the space S^1 has no boundary and this is crucial to guarantee the above property.

At time t, at any location $x \in S^1$, there is a single individual consuming $c(t,x)$, investing $i(t,x)$, depolluting $b(t,x)$, and producing $y(t,x)$.[3] The production technology is linear in the capital input, $i(t,x)$, that is:

$$y(t,x) = a(t,x)\ i(t,x),$$

where $a(t,x)$ is productivity at location x and time t. Again for simplification, to deal with a single diffusion equation in the optimal control problem, we assume that capital inputs do not accumulate over time nor are they exchanged across space. The resource constraint at any location x is:

$$c(t,x) + i(t,x) + b(t,x) = y(t,x),$$

[3]Notice that we do not account for heterogenous population density in this model, contrary to Boucekkine et al. (2019b). That is because we are more concerned with other types of geographic heterogeneities, in particular the technological ones, as argued in the introduction. We could of course introduce a general population density as in Boucekkine et al. (2019b) without any change in the methodology. Allowing people to move voluntarily is more difficult to insert without inducing nontrivial methodological problems but conveniently chosen ad hoc migration schemes can be easily accommodated.

which yields : $c(t,x) = (a(t,x) - 1)\, i(t,x) - b(t,x)$. This is a very simple and flexible productive structure, which has the invaluable advantage to accommodate most types of spatial heterogeneity (so far, heterogeneity in productivity), including the solution step. The link with pollution is simple too: we assume that pollution one-to-one arises from the use of input, $i(t,x)$.

We consider the following control problem with infinite time horizon in S^1. Let

$$p_0, \delta : S^1 \to \mathbb{R}, \qquad \varphi : \mathbb{R}^+ \times S^1 \to \mathbb{R},$$

be given measurable functions. At each time $t \in \mathbb{R}^+$ and location $x \in S^1$, the planner chooses the control variables investment, $i(t,x)$, and the pollution abatement, $b(t,x)$, knowing that the law of motion of pollution, $p(t,x)$, is given by the following parabolic PDE

$$(1) \qquad \begin{cases} p_t(t,x) = \sigma p_{xx}(t,x) - \delta(x)p(t,x) + i(t,x) - \varphi(t,x)b(t,x)^\theta, \quad (t,x) \in \mathbb{R}^+ \times S^1, \\ \\ p(0,x) = p_0(x), \quad x \in S^1, \end{cases}$$

where $\sigma > 0$, $\theta \in (0,1)$. Here, p_t, p_x, p_{xx} denote, respectively, the derivative with respect to time, and the first and second derivative with respect to space. The right-hand side of the PDE above gives the component of pollution stock variation at location x: it depends on transboundary pollution, $\sigma p_{xx}(t,x)$, where σ measures the strength of diffusion; on nature self-cleaning capacity, $\delta(x)p(t,x)$, where $\delta(x)$ is the rate of self-cleaning at location x; on input, $i(t,x)$; and, finally, on abatement, $\varphi(t,x)b(t,x)^\theta$, where $\varphi(t,x)$ is the efficiency (or productivity) of abatement, and θ is the return to scale of the abatement technology ($0 < \theta \leq 1$). As the law of pollution accumulation is written, one can read abatement as applying to the whole pollution stock. However, we can rewrite the same equation as

$$p_t(t,x) = \sigma p_{xx}(t,x) - \delta(x)p(t,x) + (1 - \eta(t,x))\; i(t,x),$$

where $\eta(t,x) = \frac{\varphi(t,x)b(t,x)^\theta}{i(t,x)}$. The term $\eta(t,x)$ can be immediately interpreted as the abatement rate of new emissions (as the latter are indexed on input use, $i(t,x)$), that is abatement at the source of pollution. That is the way we interpret abatement in our model as it transpires from our quantitative exercises in Section 4. In Boucekkine et al (2019a), there is no pollution abatement control, i.e. $\varphi \equiv 0$, and no way to distinguish between gross emissions and net (after abatement) emissions.

The payoff functional of the regulator internalizes, as outlined in the introduction, the negative externality exercized by (local) pollution on population wellbeing (notably health) and also the

transboundary nature of pollution via the state equation above. Precisely, the functional writes as

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

where $\rho > 0$, $\gamma \in (0,1)\cup(1,+\infty)$, and

$$a : \mathbb{R}^+ \times S^1 \to \mathbb{R}^+, \qquad w : S^1 \to \mathbb{R}^+,$$

are given measurable functions with $a \geq 1$. The function p in the integral above is the solution to (1) corresponding to the initial datum p_0 and to the control (i,b). The functional above takes into account two aspects of human well being: consumption and health (via pollution negative externality). Note that the functional is strictly concave in the former and linear in pollution. The latter linearity assumption is needed for the analytical solution to work. The spatial function $w(x)$ can be interpreted as the degree of pro-environmental culture (or environmental awareness) at location x. We shall not play on it in this paper.

Finally note that because of absence of trade and the technological assumptions made above, we can rewrite the objective functional in terms of the investment and abatement controls as announced above:

$$J\big(p_0;(i,b)\big) := \int_0^\infty e^{-\rho t}\left(\int_{S^1}\left(\frac{\big((a(t,x)-1)i(t,x)-b(t,x)\big)^{1-\gamma}}{1-\gamma} - w(x)p(t,x)\right)\mathrm{d}x\right)\mathrm{d}t. \tag{2}$$

We shall work with this functional hereafter.

3. Analytic results

In this section we describe the analytical results we get for the model we introduced in Section 2. For each of them we will refer the reader to the proof in the appendix where all the mathematical setting necessary to study the problem is developed.

We will make use of the following

Assumption 3.1.

(i) $\int_{S^1} |p_0(x)|^2 \mathrm{d}x < \infty$, $\delta \in C(S^1;\mathbb{R}^+)$, *and* $w \in C(S^1;(0,+\infty))$;

(ii) There exists $L > 0$ *and* $g \geq 0$ *such that*

$$(a(t,x)-1)^{\frac{1-\gamma}{\gamma}} + \varphi(t,x)(a(t,x)-1)^{\frac{\theta}{1-\theta}} \;\leq\; Le^{gt}, \qquad \forall (t,x) \in \mathbb{R}^+ \times S^1;$$

(iii) $\rho > g$.

We also specify the set of the admissible controls: the planner aims at maximizing (2) over the measurable functions (i, b) in the set

$$\mathcal{A} = \Bigg\{ (i,b)\colon \mathbb{R}^+ \times S^1 \to \mathbb{R}^+ \times \mathbb{R}^+ : \int_0^\infty e^{-(\rho-g)t} \left(\int_{S^1} |i(t,x) - \varphi(t,x) b(t,x)^\theta|^2 \mathrm{d}x \right)^{1/2} \mathrm{d}t < \infty \tag{3}$$
$$\text{and } \; c(t,x) = (a(t,x)-1)i(t,x) - b(t,x) \geq 0 \quad \forall (t,x) \in \mathbb{R}^+ \times S^1 \Bigg\}.$$

First, we observe, in the following proposition, that Assumption 3.1 is enough to ensure the well-posedness of the functional J.

Proposition 3.2. *Let Assumption 3.1 hold. Then the functional $J(p_0, (i,b))$ is well-defined for all $(i,b) \in \mathcal{A}$, possibly equal to $+\infty$ or $-\infty$ (depending, respectively, on the occurences $\gamma \in (0,1)$ and $\gamma \in (1, +\infty)$, respectively).*

Proof. See Appendix A.1. □

The planner aims at solving the optimization problem

$$v(p_0) := \sup_{(i,b)\in\mathcal{A}} J(p_0; (i,b)). \tag{4}$$

The function v denotes the value function of the optimization problem and a couple (i^*, b^*) such that $J(p_0; (i^*, b^*)) = v(p_0)$ is said to be an optimal control for the problem starting at p_0.

As announced the planner optimization problem can be explicitly solved. To state the result it is necessary to introduce a function $\alpha : S^1 \to \mathbb{R}$, which will represent the core of our solution. The latter is a function of the spatial variable x and, as argued below, it can be suitably interpreted in terms of the disutility of the pollution. It is defined (see Appendix A.2 and, notably, Proposition A.2) as the unique solution in $C^2(S^1; \mathbb{R})$ to the ODE

$$\rho\alpha(x) - \sigma\alpha''(x) + \delta(x)\alpha(x) = w(x), \quad x \in S^1. \tag{5}$$

This ODE can be viewed as an ODE on the interval $[0, 2\pi]$ with periodic boundary zero-order and first-order boundary conditions, i.e.

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

falling into the Sturm-Liouville theory with periodic boundary conditions (see Coddington and Levinson, 1955). We have

$$0 < \min_{S^1} \frac{w(\cdot)}{\rho + \delta(\cdot)} \leq \alpha(x) \leq \max_{S^1} \frac{w(\cdot)}{\rho + \delta(\cdot)} \qquad \forall x \in S^1. \tag{6}$$

Such function represents the core of the solution and has a natural interpretation: indeed one way to formally rewrite α is the following integral expression

$$\alpha(x) = \int_0^\infty e^{-\rho t} \left(\int_{S^1} w(\xi) z(t, \xi; x) \, \mathrm{d}\xi \right) \mathrm{d}t \tag{7}$$

where $z(t, \xi; x)$ is the solution of the parabolic equation

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

$z(t, \cdot; x)$ defined here is indeed the spatial density (with respect to the variable x) at time t of a unit of pollutant initially concentrated at point x (here $\Delta_{\{x\}}(\xi)$ is the so-called Dirac's delta at x), once one takes into account the diffusion process and the natural decay. Thus, the term $\int_{S^1} w(\xi) z(t, \xi; x) \, \mathrm{d}\xi$ measures the instantaneous disutility all over the space and the whole expression in the right hand side of (7) is the total spatial (temporally discounted) future social disutility of a unit of pollutant initially concentrated at x.

The following theorem contains the core of our results. We are able to explicitly find the input level and the cleaning effort which maximize the social welfare function. Consequently, we can also find the expression of the optimal net emissions, the optimal consumption and the the maximal attainable social welfare.

Theorem 3.3. *Let Assumption 3.1 hold. Then the couple* (i^*, b^*) *given by*

$$i^*(t, x) := \alpha(x)^{-\frac{1}{\gamma}} (a(t, x) - 1)^{\frac{1-\gamma}{\gamma}} + (\varphi(t, x)\theta)^{\frac{1}{1-\theta}} (a(t, x) - 1)^{\frac{\theta}{1-\theta}}, \tag{8}$$

$$b^*(t, x) := (\theta \varphi(t, x)(a(t, x) - 1))^{\frac{1}{1-\theta}}, \tag{9}$$

belongs to $\mathcal{A}$ *and is optimal for the problem* (4). *The optimal net emissions flow is*

(10)

$$n^*(t, x) := i^*(t, x) - \varphi(t, x) b^*(t, x)^\theta = \alpha(x)^{-\frac{1}{\gamma}} (a(t, x) - 1)^{\frac{1-\gamma}{\gamma}} + \theta^{\frac{1}{1-\theta}} (1 - \theta^{-1}) \varphi(t, x)^{\frac{1}{1-\theta}} (a(t, x) - 1)^{\frac{\theta}{1-\theta}}$$

and the optimal consumption flow is

$$c^*(t, x) := (a(t, x) - 1) i^*(t, x) - b^*(t, x) = \left(\frac{a(t, x) - 1}{\alpha(x)} \right)^{\frac{1}{\gamma}}. \tag{11}$$

The evolution of the optimal pollution profile p^* over time is the unique solution of the following parabolic equation:

$$\begin{cases} p_t(t,x) = \sigma p_{xx}(t,x) - \delta(x)p(t,x) + n^*(t,x), & (t,x) \in \mathbb{R}^+ \times S^1, \\ p(0,x) = p_0(x), & x \in S^1. \end{cases} \tag{12}$$

Finally, the maximal social welfare (value function) is affine in p_0:

$$\begin{aligned} v(p_0) \;=\; J\big(p_0;(i^*,b^*)\big) \;=\; & -\int_{S^1}\alpha(x)p_0(x)\mathrm{d}x + \int_0^\infty e^{-\rho t}\left(\int_{S^1}\frac{\gamma}{1-\gamma}\Big(\frac{a(t,x)-1}{\alpha(x)}\Big)^{\frac{1-\gamma}{\gamma}}\mathrm{d}x\right)\mathrm{d}t \\ & - \theta^{\frac{1}{1-\theta}}\int_0^\infty e^{-\rho t}\left(\int_{S^1}\alpha(x)\big(\varphi(t,x)(a(t,x)-1)\big)^{\frac{\theta}{1-\theta}}\mathrm{d}x\right)\mathrm{d}t. \end{aligned}$$

Proof. See Appendix A.4. □

In our model, the local depollution effort depends on local productivities (at production and at depollution), and is not impacted by the transboundary nature of pollution. In contrast, investment (and therefore consumption and production) do depend on this the latter: the regulator has to account for the fact not only the local technological characteristics but also the implication of investing at a certain location on the neighboring ones in terms of pollution. Finally notice that local investment is not necessarily increasing with local productivity at production, $a(t,x)$: a higher local productivity might lead to lower investment i, in such a way that local emissions decrease at the expense of a slightly lower production. All these properties are illustrated in the numerical section below.

In the particular case where the coefficients a and φ do not depend on time, the optimal values of i and b described in (8) and (9) are also time-independent and the pollution profile converges in the long-run to a steady state, as described in the following proposition.

Proposition 3.4. *Let Assumption 3.1 hold. Assume that the coefficients a and φ are time independent, i.e. $a(t,x) = a(x)$ and $\varphi(t,x) = \varphi(x)$, and that $\delta(\cdot) \neq 0$. Then*

$$\lim_{t\to\infty}\int_{S^1}|p^*(t,x) - p^*_\infty(x)|^2\,\mathrm{d}x = 0,$$

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

$$\sigma p''(x) - \delta(x)p(x) + n^*(x) = 0, \quad x \in S^1. \tag{13}$$

Proof. See Appendix A.5. □

4. Numerical exploration of the long-term spatial distribution

In this section we shall use the analytical results of the previous section, notably Proposition 3.4, to explore the properties of the optimal spatial distributions, and in particular the optimal long term spatial distribution of pollution.

In the pictures, we represent the space S^1 as the segment $[0, 2\pi]$ with the additional requirement that the pollution which "exits" from the right boundary 2π automatically "enters" from the left boundary 0 (and vice-versa).

All the simulations of this section take all the coefficients (notably a and φ) as time-independent. This is indeed the situation where we can apply Proposition 3.4 to characterize the long-run optimal spatial distributions. Observe indeed that, whenever the coefficients are time-dependent the convergence toward a stationary state is not at all ensured (and it arises only for very particular specifications)[4].

The contribution of this section is specifically to highlight the implications of the type of geographic heterogeneity considered. We shall study four of them: the heterogeneity in productivity, the heterogeneity in abatement efficiency, the heterogeneity in environmental awareness, and finally, the discrepancy in the rate of natural self-cleaning. For a given benchmark calibration to be displayed below, we first compute the spatial distributions induced corresponding to each type of discrepancy represented in the form of a core/periphery configuration.[5] We also study briefly the implications of combining some of these heterogeneities. Second, we show how some structural parameters of the model are important in the shape of distributions obtained.

4.1. **Benchmark calibration.** The level of the parameter a is calibrated in order to have an investment (input) - GDP ratio (which corresponds here to the value of $1/a$) in a typical range of $15\% - 40\%$ (see for instance IMF, 2019). Future discount is typically considered to be around $1\% - 5\%$ (see Barro, 2015), it is taken here equal to 3%. The inverse of intertemporal elasticity

[4]In particular, if we would have dealt with a more general time-dependent setting, where the parameters exogenously change over time, we could again use the results of Theorem 3.3 to identify and represent the optimal contemporaneous values of i, b and n, but we could not use Proposition 3.4 to characterize and represent any long-run steady state. The dynamics of i, b and n with time-dependent parameters can be richer but, since their values at the optimum at a fixed time t only depend on the value of parameters at t, their behavior follows exactly the same mechanisms we describe in our time-independent parameter examples.

[5]We use this New Economic Geography terminology to designate unimodal spatial distributions which capture the idea of economic/demographic/technological centers. In this literature, second nature (endogenous) agglomeration mechanisms are usually studied. Here we consider first nature agglomeration: we take agglomeration as given, and start from given unimodal distributions for productivity and other exogenous variables.

of substitution is 6 that is consistent for instance with the data by Barsky et al. (1997). The natural decay of pollutants strongly depends on the type of pollution (air, water, soil,...) and the nature of pollutant. Here, though our theory would apply, as we wrote, to different types of pollution, we focus on air pollution. The related empirical literature is quite abundant: depending on the pollutant (including, e.g., SO_2 and NO_2), the value of an annual decay ranges from 30% to 50% (see Versino and Angeletti, 2012, or Perry and Tabor, 1962, among others). As to the diffusion parameter σ, we shall comment on it in the discussion Section 5.2 (see references therein). Broadly speaking, for a rigourous calibration of the speed and extent of pollution on real data, we'd need a more comprehensive model including, among others, advection to capture the role of winds and currents (in the case of water pollution). An essential point we make in this paper is the implications of departing from the no-diffusion case ($\sigma = 0$). Indeed, the counterpart no-diffusion cases are represented in our figures below. A not-so-big departure from the no-diffusion case will be shown to be enough to make readily clear the long-term outcomes of diffusion as to the spatial distribution of pollution notably.

It is more difficult to calibrate other parameters: the efficiency of the depollution technology (i.e., the values of φ and θ), the unitary disutility from pollution w, and as we have written just above, the diffusion parameter σ. We take them respectively equal to 0.05, 0.2, 1 and 0.5. In the last subsection, we shall consider some departures from these benchmark values to illustrate how the shapes of spatial distributions are sensitive to some parameters. In particular, the value of γ will be shown to be crucial. As explained in Boucekkine et al. (2019b) in a different spatiotemporal context, γ measures a mixture of individual behaviour (here, intertemporal substitution in consumption) and the regulator's aversion to inequality in consumption as well. Therefore, the larger γ, the more equal should be the spatial consumption distribution.

Last but not least, it should be noted that the linearity in local pollution in the objective functional (for analytical reasons outlined above) shuts down this specific aversion to inequality channel in what concerns pollution. This does not mean that the regulator will not take care of the spatial distribution of pollution: as it will be clear in the figures displayed below, where the optimal long-term pollution spatial distributions are shown both with and without diffusion, the regulator does optimally internalize the transboundary nature of pollution but smoothing out the distribution of pollution across space. In a sense, our exercise allows to disentangle properly the action of the regulator in the face of transboundary pollution. One could expect that, with strictly convex disutility of pollution in the objective functional, the optimal shapes would have

been smoother. Still, given the specification of the model we are able to solve, we cannot provide a proof of this fact.

Heterogeneity in the technological parameter a

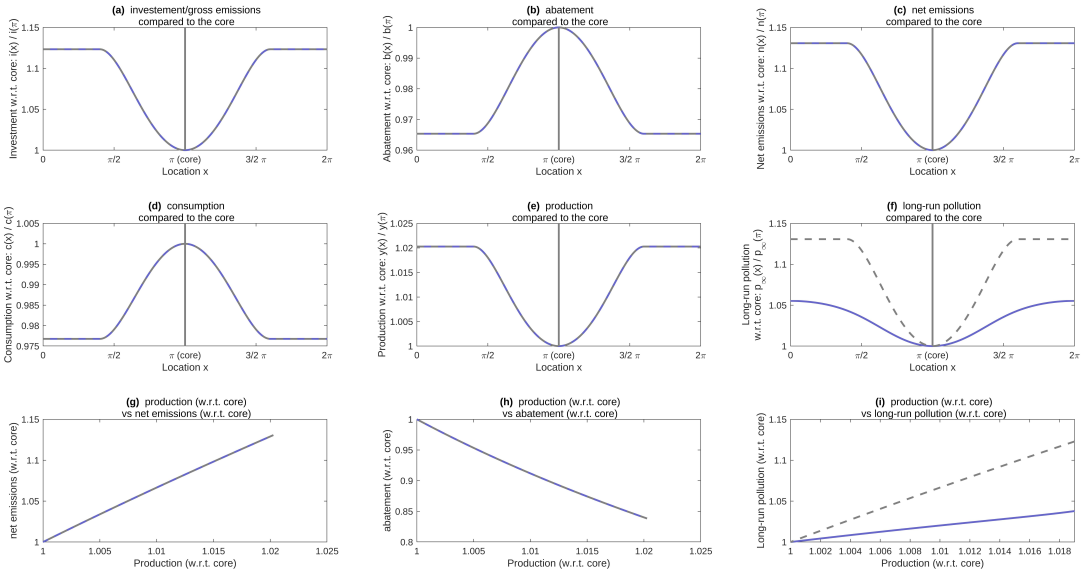

FIGURE 1. Numerical illustration of the situation in which input productivity is higher in a technologically more developed core and lower in the periphery, the peak in productivity is 10% higher than the floor value (the value of a is 3.3 at the core $x = \pi$ and 3 at the periphery). All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.4$, $w = 1$, $\gamma = 6$, $\varphi = 0.05$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

4.2. **Spatial distributions by type of geographic heterogeneity.** In this and in the following subsection we present the results of a series of quantitative exercises using the calibrated model. There are eight associated figures. Each figure contains the graphs of the same nine variables. The first five graphs represent, for the given exercise, the optimal spatial profiles (hence the variable in the horizontal axis is the space variable $x \in [0, 2\pi]$) for: investment (Subfigure (a)), abatement (b), net emission (c), consumption (d), and production (e). Since, as already mentioned, the parameters a and φ are kept constant over time (but not in space) the profiles of all these variable are also constant over time (see Theorem 3.3): by construction, they also represent, of course, the long run spatial distributions of these variables. The sixth graph (Subfigure

High heterogeneity in the technological parameter a

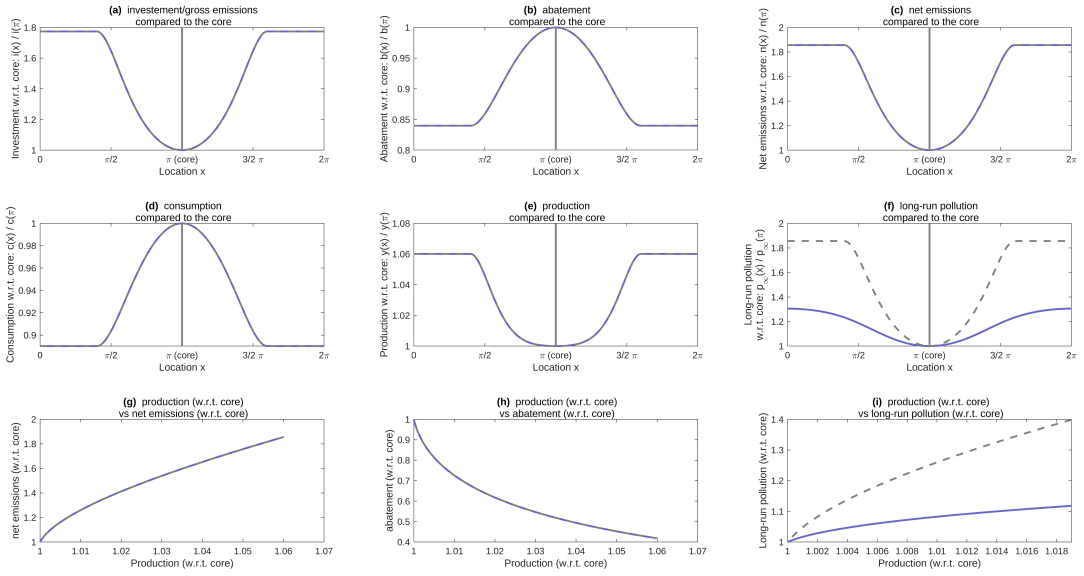

FIGURE 2. Numerical illustration of the situation in which input productivity is higher in a technologically more developed core and lower in the periphery, the peak in productivity is 66% higher than the floor value (the value of a is 5 at the core $x = \pi$ and 3 at the periphery). All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.4$, $w = 1$, $\gamma = 6$, $\varphi = 0.05$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

(f)) represents the long-run distribution of the pollutants that is the steady state characterized in Proposition 3.4. Observe that, differently from the variables we have just mentioned above, the pollution profile does have a dynamics over time and here we just represent its long-run distribution (hence, again, the variable in the horizontal axis is the space variable $x \in [0, 2\pi]$). In such six graphs the values of the plotted variables (hence the variable in the vertical axis) are expressed in terms of their value at the core so the vertical coordinate at any given point in the space is the value of the variable (e.g. investment) divided by its value at the core $x = \pi$. The core location is emphasized by a gray vertical line.

The last three graphs are about the relations between three couples of variables: production vs net emission (Subfigure (g)), production vs abatement (h), and production vs long-run pollution (i). In all three graphs the horizontal axis expresses the value of the production/income divided

by its value at the core. Hence, when the core is a minimum point for the production with respect to space (Fig. 1, 2, 4, 5), the values on the horizontal axis are smaller than 1 while, when the core is a maximum point for the production with respect to space (Fig. 3, 6, 7, 8) the values on the horizontal axis are bigger than 1. The variable in the vertical axis, in the three graphs is, respectively: net emissions (g), abatement (h), long-run pollution (i); again it is displayed divided by its value at the core. Hence, as for the horizontal variable, the scale may differ depending on the fact that the core is a minimum/maximum[6] point of the corresponding variable. Due to this choice in all such graphs, when production is equal to 1, also the other variables are equal to 1. The interest here is to see the trends of pollution related variables when production changes: in all the figures but Figure 8 we get monotonic relations between the couples of variables under examination while in Figure 8 we show that, in a case where two heterogeneities are at work, also non monotonic relations may arise.

In each figure, in each graph, we visualize two situations: the first is a case with positive diffusion $\sigma > 0$, the second is the no-diffusion benchmark ($\sigma = 0$) to show the role of the spatial dynamics we have in the model. Indeed the no-diffusion benchmark corresponds to point-by-point dynamic optimization while the positive diffusion case takes into account the complex spatio-temporal dynamics of the pollutants we have in the model. Observe that in some of the figures the profiles of investment, abatement, net emission, consumption, and production are the same for the two situations but the spatial long-run profiles of the pollutants are different. This is due to the presence of the diffusion dynamics of the pollutants when $\sigma > 0$. This process is not at work when $\sigma = 0$.

Figure 1 and Figure 2 illustrate the outcomes of a **spatial discrepancy in input productivity** through a core/periphery configuration. The peak value of productivity is 10% (Resp. 66%) higher than the periphery floor value in the former (Resp. latter) figure. Both deliver however the same qualitative picture. For the benchmark calibration chosen, both show that the regulator will invest less, depollute more, produce less and therefore pollute less in the core than in the periphery. This goes at odds with the typical picture generated by economic growth models à la Stokey (1999). Typically, these models are concerned with the evolution over time of a growing economy: as the economy develops, it eventually starts depolluting when it becomes rich enough, ultimately leading to curb pollution without breaking down growth. Here, looking

[6]This is true in all the figures but Figure 8, where the graphs of net emissions and long-run pollution with respect to space have, in the core, only a local, but not global, minimum: hence the scales of the vertical variable in the seventh and ninth graph contain both values less and greater than 1.

at the spatial (long-term) distributions resulting from optimal spatiotemporal dynamics[7], we see that the picture can be indeed very different from the Stokey story. The optimal outcome would be indeed less pollution in the core (Subfigures 1(f) and 2(f)) than in the periphery but at the same time, more production in the periphery (Subfigures 1(e) and 2(e)). Incidentally, consumption is higher in the core (Subfigures 1(d) and 2(d)) and pollution is smoothed out with respect to the no-diffusion case, illustrating the two main engines driving the regulator's action: on one hand, she aims at lowering inequalities in consumption, which in this case requires to produce more and to pollute more in the periphery (as no trade in the good is allowed across locations); on the other hand, she internalizes both the negative pollution externality at any location and pollution diffusion, leading to smooth out the pollution spatial shape with respect to the no-diffusion case. We will see in the last sub-section that the value of parameter γ is crucial in the shape of the relationships generated between long-term pollution, production and abatement.

When it comes to **heterogeneity in depollution efficiency across space**, the picture is different, as shown in Figure 3. Quite naturally, abatement is larger in the core (Subfigure 3(b)), and this goes with a larger optimal investment and a larger production (Subfigures 3(a) and 3(e)). More interesting, though production is larger in the core, its long-term optimal pollution (Subfigure 3(f)) is lower thanks to its technological superiority in abatement. The picture is analogous to Stokey's typical outcome on time paths. But we do generate it here putting spatial heterogeneity in depollution efficiency, not in input productivity (when $\gamma = 6$). If we combine both heterogeneities, as in Figure 4 with same benchmark calibration, we get qualitatively the same outcomes as in the case of a single discrepancy in input productivity, that's the above described mechanisms associated with the latter dominate. Again, this is only true at the benchmark calibration as we will show below. Many more pictures can arise when we depart from the latter, and we will show one particular and striking case therein.

We now shed light on the consequences of **spatial discrepancy in self-cleaning capacity** parameter, that is through the function $\delta(x)$, we get a peculiar and interesting picture (see Figure 5). First of all, and in contrast to the previous figures in which diffusion and no-diffusion paths coincide except for pollution, here diffusion (that's the parameter σ) matters in optimal investment, production and consumption (Subfigures 5(a), 5(e) and 5(d))). By Theorem 3.3, $\delta(x)$

[7]As observed in the first footnote of Section 4 we expect that the main mechanisms should be preserved looking at the transitional dynamics, also for time variable data a and φ.

Heterogeneity in the depollution efficiency φ

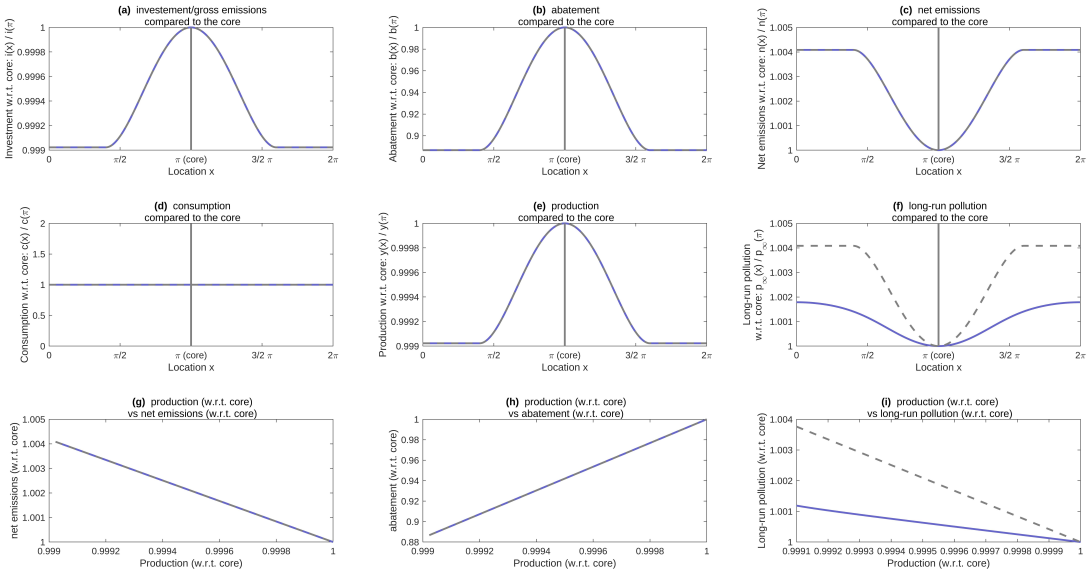

FIGURE 3. Numerical illustration of the situation in which depollution efficiency is higher in a certain core and lower in the periphery, the peak in depollution efficiency is 10% higher than the floor value (the value of ϕ is 0.055 at the core $x = \pi$ and 0.05 at the periphery). All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $a = 3$, $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.4$, $w = 1$, $\gamma = 6$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

enters the expression of optimal investment — given in (8) — indirectly through function $\alpha(x)$, which itself depends on the diffusion parameter σ via the solution of the differential equation (5). Second, optimal abatement (Subfigure 5(b)) — given in (9) — does not depend on $\delta(x)$ neither directly nor indirectly. As a result, as self-cleaning capacity is larger in the core, investment, production and consumption (Subfigures 5(a), 5(e), 5(d)) are bigger at the core, while pollution is still lower than in the periphery. We get again the Stokey classical picture where pollution goes down with production (or income) across location (Subfigures 5(i)). Moreover, for all variables, the regulator "uses" diffusion to smooth pollution across space.

Finally when we consider the **spatial heterogeneity in the environmental awareness** parameter w. In Figure 6 all the parameters are space-independent but $w(x)$ has the core/periphery unimodal shape: higher in a region (around $x = \pi$, where the peak is reached) and lower in the periphery. The picture we get is in line with the intuition: in the zones where the environmental

Heterogeneity in productivity and depollution efficiency

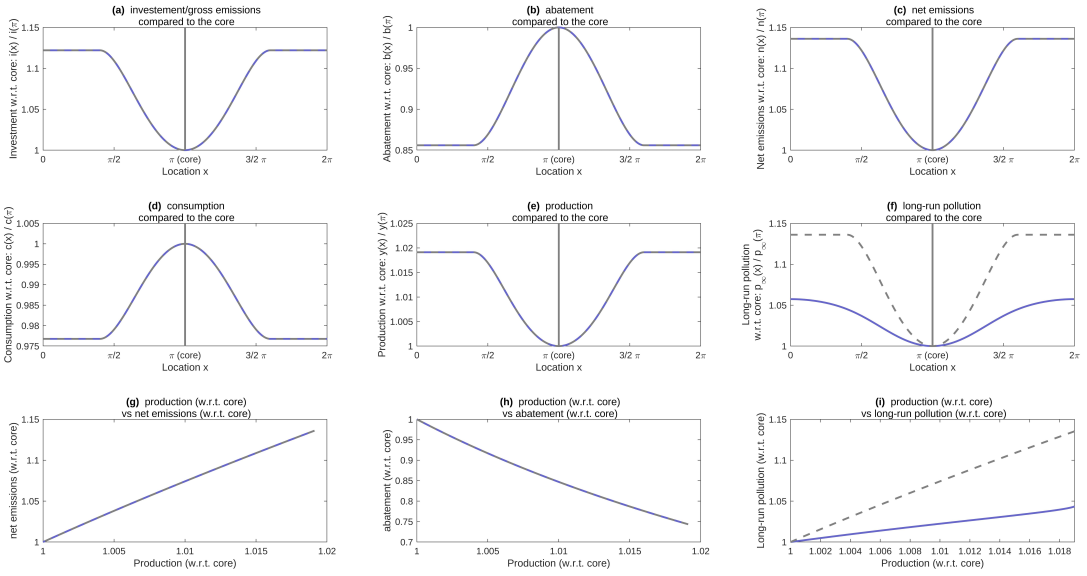

FIGURE 4. Numerical illustration of the situation in which both input productivity and depollution efficiency are higher in a technologically more developed core and lower in the periphery. Both the peaks in the productivity the depollution efficiency are 10% higher than the floor values (the value of a is 3.3 at the core $x = \pi$ and 3 at the periphery, the value of ϕ is 0.055 at the core and 0.05 at the periphery). All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.4$, $w = 1$, $\gamma = 6$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

awareness is higher the disutility for one unit of pollution is higher and then the planner decides to reduce the polluting investments and consequently to reduce both production (Subfigure 6(e)) and consumption (Subfigure 6(d)) in these locations. The behavior of the abatement (Subfigure 6(b)) can seem surprising because it does not peak in the area where the awareness is higher but this is in line with the expression of the optimal spatial distribution of b (see (9)) which does not take into account the distribution of w.

4.3. **The role of structural parameters.** In this subsection, we will provide two striking examples of how the spatial distributions can be strongly impacted by changes in the structural parameters. In Figure 7, we report the optimal distributions obtained when we decrease γ from 6 to 0.5 under a core/periphery configuration structure on input productivity (all the

Heterogeneity in the natural decay parameter δ

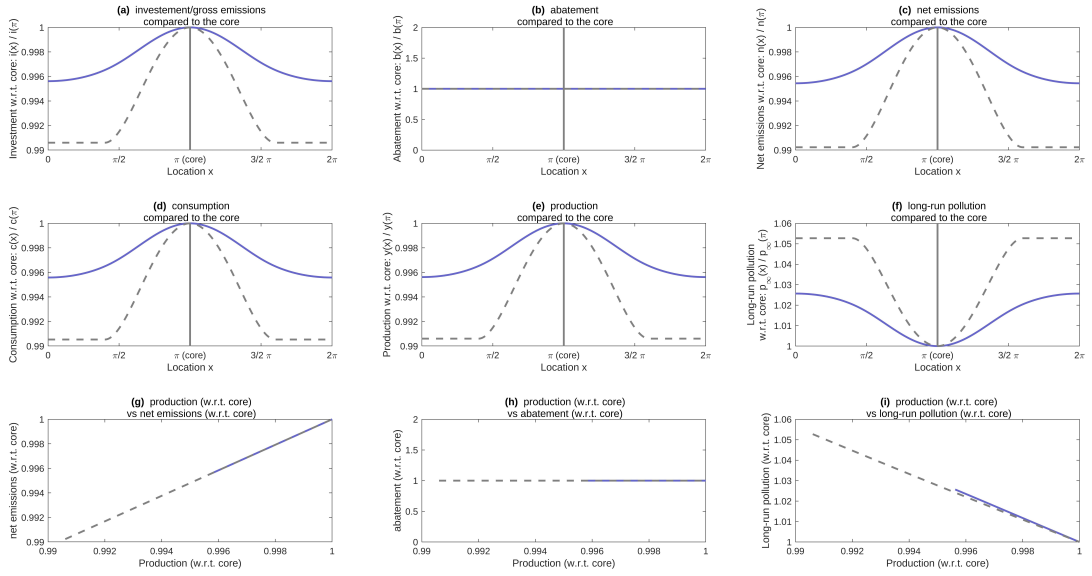

FIGURE 5. Numerical illustration of the situation in which natural decay is higher in a region and lower in the remaining part of the surface, the peak is in $x = \pi$ where the value is 0.425 and it decreases in peripheral where its value is 0.4. All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $a = 3$, $\rho = 0.03$, $\sigma = 0.5$, $w = 1$, $\gamma = 6$, $\varphi = 0.05$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

other parameter values are unchanged). Compared to the benchmark, here Figure 1, there are notable changes: investment in the core is increased (Subfigures 7(a)), and not reduced, and despite an increase in depollution effort in the core (Subfigures 7(b)), net emissions (Subfigures 7(c)) and long-term pollution (Subfigures 7(f)) are higher than in the periphery, again in sharp contrast to the benchmark case. The reason behind this fact is that lowering so strongly γ is reflected in a much lower aversion to inequality: as a result, the regulator does not need to increase investment more in the periphery to compensate the disadvantage in input productivity. Eventually, consumption (Subfigures 7(d)) is less smooth than in the benchmark.

So far, we have generated a rich set of spatial distributions shapes but with a monotonic relationship between production and pollution (or emissions). Our model can generate non-monotonic relationships between these two variables by combining different types of geographic discrepancies and adjusting the values of some critical parameters like γ. Figure 8 provides an

High heterogeneity in the environmental awareness parameter w

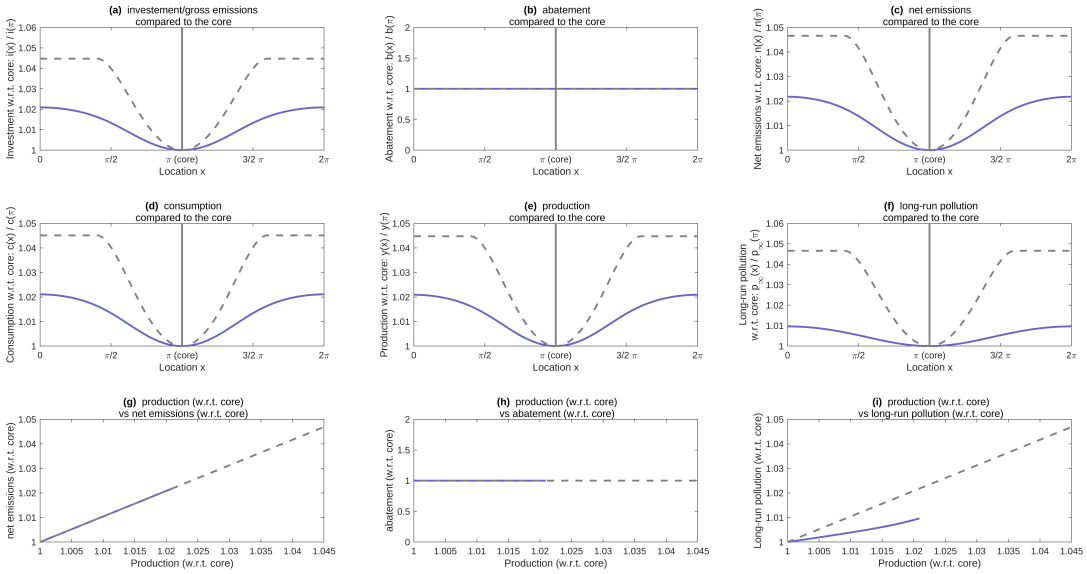

FIGURE 6. Numerical illustration of the situation in which environmental awareness is higher in a region and lower in the remaining part of the surface, the peak is in $x = \pi$ where the value is 1.3 and it decreases in peripheral where its value is 1. All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $a = 3$, $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.4$, $\gamma = 6$, $\varphi = 0.05$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

example. In this picture, as in Figure 4, we both consider heterogeneities in input productivity and depollution efficiency but here the level of the depollution efficiency is significantly higher (equal to 0.595 in the periphery) and, as in Figure 7, $\gamma = 0.5$. We get a non-monotonic relation between production and emissions (Subfigures 8(i)) in the spirit of the the so-called Environmental KC (see Stokey, 1999).

5. Discussion of results

5.1. **Heterogeneity in the model.** The topic of pollution management and dynamics, especially in its transboundary terms, is certainly not new in the operations research literature. We can for instance mention the contributions by Stam et al., (1992), Xu and Masui (2009), Zhao et al. (2013), Bertinelli et al. (2014), Ouardighi et al. (2018), Ferrari and Koch (2019) and many others.

Heterogeneity in input productivity a with low γ

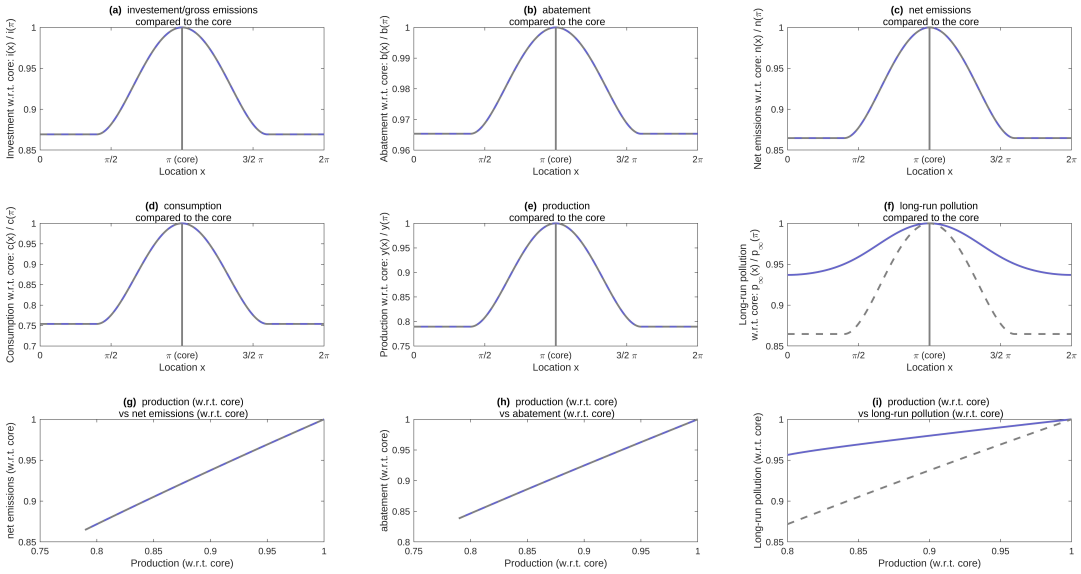

FIGURE 7. Numerical illustration of the situation in which input productivity is higher in a region and lower in the remaining part of the surface. The peak in productivity is 10% higher than the floor value (the value of a is 3.3 at the core $x = \pi$ and 3 at the periphery). All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.4$, $w = 1$, $\gamma = 0.5$, $\varphi = 0.05$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

The contribution of this work is situated in a new and promising strand of literature in which pollution management issues are considered in a model where a continuous number of locations explicitly appears. This makes it possible to directly take into account the spatial externalities of pollution and its distribution. To our knowledge, the only two works that have considered such a setting are those of de Frutos et al., (2019a,b) and, in the different context of land-use dynamics, Camacho and Perez-Barahona (2015).

In terms of the ability to take into account the complexity of the problem and the heterogeneity observed from one site to another, the work presented here improves the setting of de Frutos et al. (2019a,b) (which is, notably, in a game set-up) in different ways:

- Our model allows to specify different levels of productivity in different locations. This means in particular to be able to take into account heterogeneous levels of technological

Heterogeneity in productivity and depollution efficiency with low γ

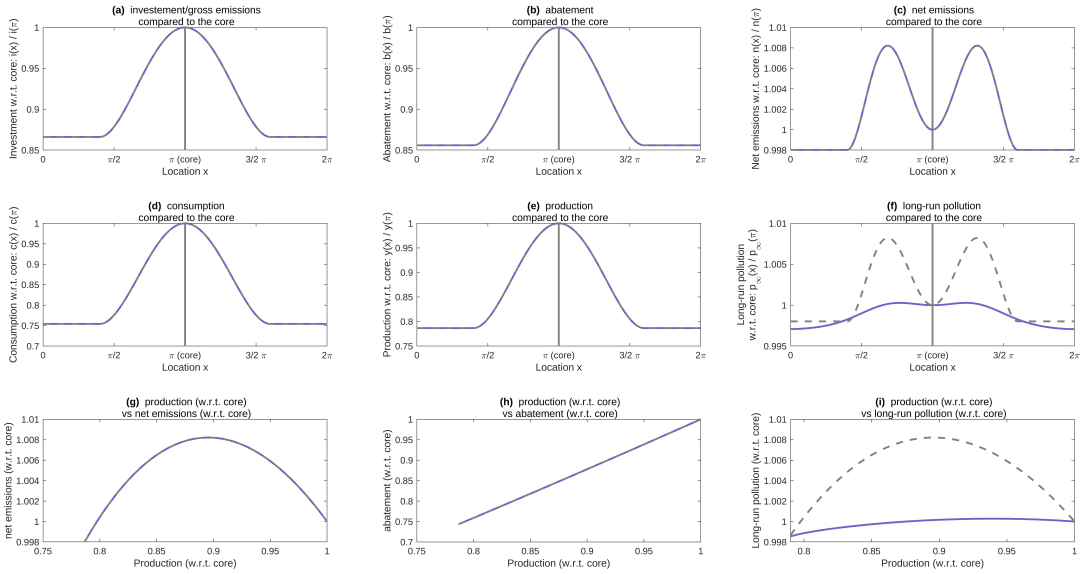

FIGURE 8. Numerical illustration of the situation in which both input productivity and depollution efficiency are higher in a region and lower in the remaining part of the surface. The value of a is 3.3 at the core $x = \pi$ and 3 at the periphery, the value of ϕ is 0.65 at the core and 0.59 at the periphery. All other exogenous parameters are homogeneous in space. The spatial optimal distribution of economic and environmental relevant variables, in terms of their value at the core, the relations between production and net emissions and production and abatement are represented. The values of other parameters (constant over space) are: $\rho = 0.03$, $\sigma = 0.5$, $\delta = 0.2$, $w = 1$, $\gamma = 0.5$, $\theta = 0.2$. Dashed lines are related to the no-diffusion benchmark.

development and adoption of technologies in different areas and then to use the model to study core-periphery situations as the ones we simulated in Section 4.

- Our formulation also makes it possible to consider time-dependent productivity parameters. This allows to take into account (exogenously) the process of growth and technological development.
- Our setting allows us to consider different levels of environmental awareness in different locations. This type of heterogeneity is, arguably, very important to explain a number of differences in individual behavior that we observe across the different countries of the world.
- We are able to specify a more general utility function with a generic inter-temporal substitution parameter.

In addition to all this, our model, unlike previous contributions in which the spatial dimension and the pollutant spatio-temporal diffusion process are explicitly modeled, allows a double channel of decision making: on the one hand the control i, the investment (and therefore the production and the emissions); on the other hand the control b, the abatement.

5.2. **On the link of the model with available data.** Our model is designed and presented in stylized way to provide a simple and effective description of the main qualitative results, as can be seen in the results of Section 3 and in the simulations of Section 4. It is however important to note that, with a simple mathematical effort (that we avoided here to make our qualitative analysis more clear), our model can be extended to more general set-ups, that could be even more useful to test it on real data on pollution management across space.

First of all we observe that all the main results of Section 3 can be extended, without big effort, to the case when the spatial variable x lives in the sphere S^2 (the surface of planet earth) or in a given a subset of it (a given country or region). This can be done along the way used in Fabbri (2016), using the results of Grigoryan (2012) and could be useful to use real spatial data to calibrate our model.

Concerning calibration on the parameters of our model, we already observed, at the beginning of Subsection 4.1, that the parameters a, ρ, δ and γ can be reasonably calibrated from the existing empirical literature. With a some more effort the other parameters can also be deduced from the actual data. For example:

- Concerning w one can use indicators of environmental awareness such as, for instance, those described in the report by European Commission (2015) which uses, in particular, data of land use and waste recycling.
- One could use classical empirical studies to determine the value of σ (see for example Birney and Durham, 1971 or Desalu et al., 1974). Such values evidently depend on the type of pollutant considered.

Other possible extensions of the model include σ depending on x and an advection term p reflecting the average flow ("wind") direction in the model. In this way other aspects of the problem could be quantified in the context of our analysis. We are working on an extension of the model taking account of such parameters.

6. Conclusions

In this paper, we construct a spatiotemporal model where optimal investment and optimal depollution can be jointly determined where pollution is transboundary. Beside the nice feature of keeping the analytical nature of the solution paths when including pollution control, we have been able to uncover many aspects of optimal policy under transboundary pollution and geographic heterogeneities. We have studied three types of spatial discrepancies (input productivity, depollution efficiency and nature self-cleaning capacity) and we could have studied more given the flexibility of our framework. A rich set of optimal spatial distributions has been identified and we have also clarified the mechanisms leading the regulator to choose these distributions. Key aspects in the optimal policy of the regulator are indeed the role of aversion to inequality, notably leading to smoothing consumption across locations, and the control of diffusive pollution adding another smoothing engine.

Clearly, while clarifying the latter aspects, several issues remain open. A few are due to the analytical solution which requires linearity in the state variable within the objective functional. Allowing for the strict convexity of the disutility from pollution is a natural extension, beside generalization. It will allow to incorporate aversion to environmental inequality in the frame, which is an increasingly important normative aspect.

Appendix A. Proofs

NOTE: Throughout the Appendix, Assumption 3.1 will hold and will not be repeated in the statement of the results.

A.1. **Formulation in Hilbert spaces.** Here we give a rigorous formulation of the problem formulated in Section 3 by embedding it in an infinite dimensional setting.

On S^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 functions $S^1 \to \mathbb{R}$ with 2π-periodic function $\mathbb{R} \to \mathbb{R}$; differentiaton of functions $S^1 \to \mathbb{R}$ is defined according to this identification.

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 procedure is similar to Boucekkine et al. (2019a), but their results do not cover the problem we are studying here because of the presence of the abatment.

The infinite dimensional space H, where we will reformulate our maximization, is the Lebesgue space $L^2(S^1;\mathbb{R})$, i.e.

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

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. 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). 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 — we refer to Brezis (2011) for the notion of *weak* differentiability:

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

Consider the differential operator $\mathcal{L} : D(\mathcal{L}) \subset H \to H$, where

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

Due to Assumption 3.1, the latter 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 $\psi \in C^\infty(S^1;\mathbb{R})$. Integration by parts yields

$$\langle \mathcal{L}\psi, \psi\rangle \quad = \quad \int_{S^1} ([\mathcal{L}\psi](x))\,\psi(x)\mathrm{d}x = -\int_{S^1} \sigma|\psi'(x)|^2\mathrm{d}x - \int_{S^1} \delta(x))|\psi(x)|^2\mathrm{d}x \leq 0. \tag{14}$$

Since $C^\infty(S^1;\mathbb{R})$ is a core for $\mathcal{L}$, (15) extends to all functions $\psi \in D(\mathcal{L})$, showing that the operator $\mathcal{L}$ is dissipative. Similarly, a double integration by parts shows that

$$\langle \mathcal{L}\psi_1, \psi_2\rangle = \langle \psi_1, \mathcal{L}\psi_2\rangle, \quad \forall \psi_1, \psi_2 \in C^\infty(S^1;\mathbb{R}). \tag{15}$$

Again, since $C^\infty(S^1;\mathbb{R})$ is a core for $\mathcal{L}$, (15) extends to all couples of functions in $D(\mathcal{L})$, showing that $\mathcal{L}$ is self-adjoint, i.e. $\mathcal{L} = \mathcal{L}^*$, where $\mathcal{L}^*$ denotes the adjoint of $\mathcal{L}$. Therefore, by Engel and Nagel (1995) (see in particular Chapter II), $\mathcal{L}$ generates a strongly continuous contraction semigroup $(e^{t\mathcal{L}})_{t\geq 0} \subset L(H)$; in particular, since $\mathcal{L}$ is dissipative and $\rho > 0$, by standard theory of strongly continuous semigroup in Banach spaces (see, e.g. Engel and Nagel, 1995, Ch. II, p. 82-83 and Ch. II, Th. 1.10, p. 55]), it follows that ρ belongs to the resolvent set of $\mathcal{L}$, i.e.

$$\rho - \mathcal{L} :\ D(\mathcal{L}) \longrightarrow H$$

is invertible with bounded inverse $(\rho - \mathcal{L})^{-1} : H \to H$ and the resolvent formula hold: for every $\rho > 0$

$$(\rho - \mathcal{L})^{-1}h = \int_0^\infty e^{-(\rho-\mathcal{L})t}h\,\mathrm{d}t \qquad \forall h \in\ H. \tag{16}$$

Given $i, b : \mathbb{R}^+ \times S^1 \to \mathbb{R}^+$ with functions $I, B : \mathbb{R}^+ \to H^+$, provided by

$$I(t) = i(t,\cdot), \quad B(t) = b(t,\cdot).$$

Moreover, define

$$\Phi : \mathbb{R}^+ \to H^+, \quad \Phi(t) := \varphi(t, \cdot),$$

and

$$[\Phi(t)B(t)^\theta](x) := \Phi(t)(x)(B(t)(x))^\theta = \varphi(t,x)b(t,x)^\theta, \quad t \geq 0,\ x \in S^1.$$

Then, defining also the function (net emissions)

$$N : \mathbb{R}^+ \to H, \quad N(t) := I(t) - \Phi(t)B(t)^\theta,$$

the set $\mathcal{A}$ is rewritten as

$$\mathcal{A} = \Bigg\{ (I,B) : \mathbb{R}^+ \to H^+ \times H^+ : \int_0^\infty e^{-(\rho-g)t}\|N(t)\|\mathrm{d}t < \infty,$$
$$(A(t)(x) - 1)I(t)(x) - B(t)(x) \geq 0 \quad \forall (t,x) \in \mathbb{R}^+ \times S^1 \Bigg\}.$$

Hence, given $(I,B) \in \mathcal{A}$ and with the identification $P(t) = p(t,\cdot)$, we can reformulate (1) in H as an abstract evolution equation:

$$\begin{cases} P'(t) = \mathcal{L}P(t) + N(t), \quad t \geq 0, \\ P(0) = p_0 \in H. \end{cases} \tag{17}$$

According to Part. II, Ch. 1. Def. 3.1(v) of Bensoussan et al. (2007) we define the *mild solution* to (17) as

$$P(t) = e^{t\mathcal{L}}p_0 + \int_0^t e^{(t-s)\mathcal{L}}N(s)\mathrm{d}s, \quad t \geq 0. \tag{18}$$

The formula (18) provides our notion of solution to (1). Now we go on by reformulating the objective functional. Set $A(t) := a(t,\cdot)$ and

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

The functional (2) is rewritten in this formalism as

$$J(p_0,(I,B)) = \int_0^\infty e^{-\rho t}\left[\left\langle \frac{\big((A(t) - \mathbf{1})I(t) - B(t)\big)^{1-\gamma}}{1-\gamma}, \mathbf{1}\right\rangle - \langle w, P(t)\rangle\right]\mathrm{d}t. \tag{19}$$

Proposition A.1. *$J(p_0,(I,B))$ is well defined for all $p_0 \in H$ and $(I,B) \in \mathcal{A}$.*

Proof. The term $\dfrac{((A(t) - \mathbf{1})I(t) - B(t))^{1-\gamma}}{1-\gamma}$ in (19) is always either positive (if $\gamma \in (0,1)$) or negative (if $\gamma > 1$). So, it suffices to show that $\int_0^\infty e^{-\rho t}\langle w, P(t)\rangle\mathrm{d}t$ is well defined and finite. We have

$$\int_0^\infty e^{-\rho t}\big\langle w, P(t)\big\rangle\mathrm{d}t = \int_0^\infty e^{-\rho t}\Big\langle w, e^{t\mathcal{L}}p_0 + \int_0^t e^{(t-s)\mathcal{L}}N(s)\mathrm{d}s\Big\rangle\mathrm{d}t. \tag{20}$$

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

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

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

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

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

Proof of Proposition 3.2. The claim of Proposition 3.2 is the counterpart of the claim of Proposition A.1 above, where the rigorous formulation of the optimal control problem in the infinite dimensional space H is considered.

A.2. **The function α.** We define a function α, which will represent the core of the solution. Set

$$\alpha := (\rho - \mathcal{L})^{-1}w = \int_0^\infty e^{-(\rho-\mathcal{L})t}w\,\mathrm{d}t, \tag{21}$$

where the equality above is due to (16). By definition, α is the unique solution in $W^{2,2}(S^1;\mathbb{R})$ of the abstract ODE

$$(\rho - \mathcal{L})\,\alpha = w. \tag{22}$$

More explicitly, α, as defined in (21), is the unique solution in the class $W^{2,2}(S^1;\mathbb{R})$ to

$$\rho\alpha(x) - \sigma\alpha''(x) + \delta(x)\alpha(x) = w(x), \quad x \in S^1, \tag{23}$$

meaning that it verifies (23) pointwise almost everywhere in S^1. The latter ODE can be viewed as an ODE on the interval $[0, 2\pi]$ with periodic boundary zero-order and first-order boundary conditions, i.e.

$$\begin{cases}
\rho\alpha(x) - \sigma\alpha''(x) + \delta(x)\alpha(x) = w(x), \quad x \in [0, 2\pi], \\
\alpha(0) = \alpha(2\pi), \quad \alpha'(0) = \alpha'(2\pi),
\end{cases}$$

falling into the Sturm-Liouville theory with periodic boundary conditions (see Coddington and Levinson, 1955). Recall that we are dealing with the topology induced by the topology of $\mathbb{R}^2$ on S^1 and we are identifying functions on S^1 with 2π-periodic functions on $\mathbb{R}$. By Sobolev embedding $W^{2,2}(S^1;\mathbb{R}) \subset C^1(S^1;\mathbb{R})$, so $\alpha \in C^1(S^1;\mathbb{R})$.

Proposition A.2. *We have $\alpha \in C^2(S^1;\mathbb{R})$ and*

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

Proof. The fact that α solves (23) and the fact that $\sigma > 0$ yield

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

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 (23) 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 obtained symmetrically. □

A.3. **Rewriting the objective functional.** Using (18) it is possible to rewrite the second part of the functional (2) in a more convenient way. Setting

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

first, we rewrite

$$\begin{aligned}
&\int_0^\infty e^{-\rho t}\left(\int_{S^1} w(x)p(t,x)\mathrm{d}x\right)\mathrm{d}t = \int_0^\infty e^{-\rho t}\big\langle w, P(t)\big\rangle \mathrm{d}t \\
&= \int_0^\infty e^{-\rho t}\Big\langle w, e^{t\mathcal{L}}p_0 + \int_0^t e^{(t-s)\mathcal{L}}N(s)ds\Big\rangle \mathrm{d}t \\
&= \left\langle w, \int_0^\infty e^{-(\rho-\mathcal{L})t}p_0\,\mathrm{d}t\right\rangle + \int_0^\infty e^{-\rho t}\left\langle w, \int_0^t e^{(t-s)\mathcal{L}}N(s))ds\right\rangle \mathrm{d}t.
\end{aligned} \tag{24}$$

Note that the first term of the right hand side is the only one which depends on the initial datum p_0. We now devote some space to rewrite and study (in Propositions A.2) such term. Then we show how to rewrite the whole functional, including second term (Proposition A.3). First of all, by (16), it can be rewritten as

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

We now rewrite also the last term of the last line of (24) getting the following result.

Proposition A.3. *We have*

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

Proof. Using the definition of α given in (21), the last term of the last line of (24) can be rewritten by exchanging the integrals as follows:

$$\begin{aligned}
&\int_0^\infty \left(\int_0^t e^{-\rho t} \left\langle w, e^{(t-s)\mathcal{L}} N(s) \right\rangle \mathrm{d}s \right) \mathrm{d}t \\
&= \int_0^\infty \left(\int_0^t e^{-\rho s} \left\langle w, e^{-(\rho-\mathcal{L})(t-s)} N(s) \right\rangle \mathrm{d}s \right) \mathrm{d}t \\
&= \int_0^\infty e^{-\rho s} \left\langle w, \int_s^\infty e^{-(\rho-\mathcal{L})(t-s)} N(s) dt \right\rangle \mathrm{d}s \qquad (25) \\
&= \int_0^\infty e^{-\rho s} \left\langle w, (\rho-\mathcal{L})^{-1} N(s) \right\rangle \mathrm{d}s \\
&= \int_0^\infty e^{-\rho s} \left\langle (\rho-\mathcal{L})^{-1} w, N(s) \right\rangle \mathrm{d}s.
\end{aligned}$$

The claim immediately follows. □

A.4. **Solution of the problem.**

Theorem A.4. *The couple (I^*, B^*) given by*

$$B^*(t)(x) = (\theta\varphi(t,x)(a(t,x)-1))^{\frac{1}{1-\theta}}, \qquad (26)$$

$$I^*(t)(x) = \alpha(x)^{-\frac{1}{\gamma}} (a(t,x)-1)^{\frac{1-\gamma}{\gamma}} + (\theta\varphi(t,x))^{\frac{1}{1-\theta}} (a(t,x)-1)^{\frac{\theta}{1-\theta}}. \qquad (27)$$

belongs to $\mathcal{A}$ and is optimal starting at each p_0. The corresponding optimal net emissions flow is

$$N^*(t) := I^*(t) - \Phi(t) B^*(t)^\theta = \alpha(x)^{-\frac{1}{\gamma}} (a(t,x)-1)^{\frac{1-\gamma}{\gamma}} + \theta^{\frac{1}{1-\theta}} (1-\theta^{-1}) \varphi(t,x)^{\frac{1}{1-\theta}} (a(t,x)-1)^{\frac{\theta}{1-\theta}}, \qquad (28)$$

and the optimal consumption flow is

$$C^*(t) = (A(t) - \mathbf{1}) I^*(t) - B^*(t) = \Big(\frac{a(t,x)-1}{\alpha(x)} \Big)^{\frac{1}{\gamma}}. \qquad (29)$$

The optimal pollution flow is

$$P^*(t) := e^{t\mathcal{L}} p_0 + \int_0^t e^{(t-s)\mathcal{L}} N^*(t) \mathrm{d}s. \qquad (30)$$

Finally, the value function is affine in p_0:

$$\begin{aligned}
v(p_0) \;=\; J\big(p_0; (I^*, B^*)\big) \;=\; & -\langle \alpha, p_0 \rangle + \int_0^\infty e^{-\rho t} \left(\int_{S^1} \frac{\gamma}{1-\gamma} \Big(\frac{a(t,x)-1}{\alpha(x)} \Big)^{\frac{1-\gamma}{\gamma}} \mathrm{d}x \right) \mathrm{d}t \\
& - \theta^{\frac{1}{1-\theta}} \int_0^\infty e^{-\rho t} \left(\int_{S^1} \alpha(x) \big(\varphi(t,x)(a(t,x)-1) \big)^{\frac{\theta}{1-\theta}} \mathrm{d}x \right) \mathrm{d}t.
\end{aligned}$$

Proof. First of all, we need to check that $(I^*, B^*) \in \mathcal{A}$. We have

$$\big((A(t) - \mathbf{1}) I^*(t) - B^*(t) \big)(x) = \Big(\frac{a(t,x)-1}{\alpha(x)} \Big)^{\frac{1}{\gamma}} \geq 0 \quad \forall (t,x) \in \mathbb{R}^+ \times S^1.$$

Morever, by the expression of $N^*(t)$ provided by (28) and by (6), and considering Assumtpion 3.1(ii), we get the existence of some constant $C_0 > 0$ such that

$$0 \leq N^*(t)(x) \leq C_0 e^{gt} \quad \forall (t,x) \in \mathbb{R}^+ \times S^1.$$

We conclude that $(I^*, B^*) \in \mathcal{A}$ by Assumption 3.1(iii).

Concerning optimality, after writing explicitely the inner products in the expression of J provided by Proposition A.3, the integrals can be optimized pointwisely, getting the expression of the optimizers. Indeed, fix $(t,x) \in \mathbb{R}^+ \times S^1$. By strict concavity of the integrand function with respect to $\iota := I(t)(x)$ and $\beta := B(t)(x)$, the unique maximum point can be found just by first order optimality conditions. The resulting system is

$$\begin{cases} \big((a(t,x)-1)\iota - \beta\big)^{-\gamma}(a(t,x)-1) - \alpha(x) = 0, \\ \\ -\big((a(t,x)-1)\iota - \beta\big)^{-\gamma} + \alpha(x)\varphi(t,x)\theta\beta^{\theta-1} = 0. \end{cases} \tag{31}$$

The claim on the optimal control then follows by solving the above system. The remaining claims immediately follow from straightforward computations. □

Proof of Theorem 3.3. It follows by noticing that the claims of Theorem 3.3 are just the rephrasing in the original PDE setting of the claims expressed in Theorem A.4 in the infinite-dimensional formulation.

A.5. **Limit behaviour in the time-homogeneous case.** We consider now the special case when the coefficient are time-independent, i.e. $a(t,x) = a(x)$, etc.. In this case the expressions of the optimal controls are time independent, too:

$$B^*(t)(x) \equiv B^*(x) := [\varphi(x)(a(x)-1)\theta]^{\frac{1}{1-\theta}}, \tag{32}$$

$$I^*(t)(x) \equiv I^*(x) := \alpha(x)^{-\frac{1}{\gamma}}(a(x)-1)^{\frac{1-\gamma}{\gamma}} + (\varphi(x)\theta)^{\frac{1}{1-\theta}}(a(x)-1)^{\frac{\theta}{1-\theta}}. \tag{33}$$

Finally, we prove the following proposition.

Proposition A.5. *Let Assumption 3.1 hold and assume, furthermore, that $\delta(\cdot) \neq 0$. We have*

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

where $P^*_\infty \in W^{2,2}(S^1;\mathbb{R})$ is the unique solution in H to the abstract ODE

$$\mathcal{L}P + N^* = 0.$$

Proof. Let λ_0 be the spectral bound of $\mathcal{L}$ (see Chapter IV of Engel and Nagel (1995)). Since $\delta \not\equiv 0$, the operator $\mathcal{L}$ is strictly disspative, hence $\lambda_0 < 0$. 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, we can rewrite

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

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

$$P^*_\infty := \int_0^\infty e^{-\lambda_0 s} e^{s\mathcal{L}_0} N^* \mathrm{d}s \in H.$$

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 $P^*_\infty = (\lambda_0 - \mathcal{L}_0)^{-1} N^*$, i.e. P^*_∞ is the solution to $(\lambda_0 - \mathcal{L}_0)P = N^*$ or, equivalently, to $\mathcal{L}P + N^* = 0$. □

Proof of Proposition 3.4. The claim of Proposition 3.4 is the counterpart of the claim of Proposition A.5 above. Indeed, (13) is the ODE counterpart of the abstract equation $\mathcal{L}P^*_\infty + N^* = 0$.

References

Barro, R. J. (2015). Environmental protection, rare disasters and discount rates. *Economica*, 82(325), 1-23.

Barsky, R. B., Juster, F. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, 112(2), 537-579.

Bawa, V.S. (1975). On optimal pollution control policies. *Management Science*, 21, 1397-1404.

Bensoussan, A., Da Prato, G., Delfour, M.C. and Mitter, S.K. (2007). *Representation and Control of Infinite Dimensional Systems*. 2nd edition. Basel: Birkhäuser.

Bertinelli, L., Camacho, C., and Zou, B. (2014). Carbon capture and storage and transboundary pollution: A differential game approach, *European Journal of Operational Research*, 237 (2), 721-728.

Boucekkine, R., Fabbri, G., Federico, S. and Gozzi, F. (2019a). A spatiotemporal framework for the analytical study of optimal growth under transboundary pollution. Working Paper 2019-26, AMSE.

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 (6), 1287–1318.

Fish B. R. and Durham J. L. (1971) Diffusion Coefficient of SO_2 in Air, *Environmental Letters*, 2(1):13-21.

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

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

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

Cohen, A., and Hurter, A. P. (1974). An input-output analysis of the costs of air pollution control. *Management Science*, 21 (4), 453-461.

Davis, D., Krupnick, A., and McGlynn, G. (Eds.), 2001. *Ancillary benefits and costs gas mitigation*. In: Proceedings of an IPCC Co-Sponsored Workshop,27–29 March 2000. Organization for Economic Co-Operation and Development,Washington, DC

Desalu, A., Gould, L. and Schweppe, F. (1974). Dynamic estimation of air pollution. *IEEE Transactions on Automatic Control*, 19(6), pp.904-910.

El Ouardighi, F., Kogan, K., Gnecco, G., and Sanguineti, M. (2018). Transboundary pollution control and environmental absorption efficiency management. *Annals of Operations Research*, 1-29.

European Commission (2015) Attitudes of Europeans towards biodiversity. Special Eurobarometer 436.

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

Ferrari, G. and Koch, T., (2019). On a strategic model of pollution control. *Annals of Operations Research*, 275(2), pp.297-319.

de Frutos F., Martin-Herran G. (2019a). 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

de Frutos, J., López, P. and Martin-Herran, G. (2019b). Equilibrium strategies in a multiregional transboundary pollution differential game with spatially distributed controls. arXiv preprint arXiv:1912.00481.

Grigoryan, A. (2012). Heat Kernel and Analysis on Manifolds. American Mathematical Society.

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

IMF (2019). World economic outlook database.

Kerl P.Y., Zhang W., Moreno-Cruz J.B., Nines A., Realff M.J., Russell A.G., Sokol J., Thomas V.M. New approach for optimal electricity planning and dispatching with hourly time-scale air quality and health considerations. *Proceedings of the National Academy of Sciences*, 112 (35) (2015), 10884-10889.

La Torre D., Liuzzi D., Marsiglio, S (2019). Transboundary Pollution Externalities: Think Globally, Act Locally?. arXiv preprint arXiv:1910.04469.

Leibowicz, B.D (2020). Urban land use and transportation planning for climate change mitigation: A theoretical framework. *European Journal of Operational Research*, 284, 604-616.

Ostrom E (2016). A Polycentric approach for coping with climate change. Policy Research Working Paper 5095, The World Bank.

Perry, W. H. and Tabor, E. C. (1962). National air sampling network measurement of SO_2 and NO2. *Archives of Environmental Health: An International Journal*, 4(3), 254-264.

Stam, A., Kuula, M., Cesar, H. (1992). Transboundary air pollution in Europe: An interactive multicriteria tradeoff analysis. *European Journal of Operational Research*, 56, 263-277.

Stokey N. (1999). Are there limits to growth? *International Economic Review*, 39, 1-31.

Versino, B. and Angeletti, G. (Eds.). (2012). Physico-Chemical Behaviour of Atmospheric Pollutants: Proceedings of the Third European Symposium. Springer Science and Business Media.

Xu, Y., and Masui, T. (2009). Local air pollutant emission reduction and ancillary carbon benefits of SO_2 control policies: Application of AIM/CGE model to China. *European Journal of Operational Research*, 198(1), 315-325.

Zhao, L., Li, C., Huang, R., Si, S., Xue, J., Huang, W., Hu, Y. (2013). Harmonizing model with transfer tax on water pollution across regional boundaries in a China's lake basin. *European Journal of Operational Research*, 225, 377-382.

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 of the center of excellence LABEX MME-DII (ANR-11-LBX-0023-01)

R. Boucekkine, 5-9 Boulevard Bourdet, 13001 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