Economic Theory (2024) 77:349–411
https://doi.org/10.1007/s00199-023-01517-w

RESEARCH ARTICLE

Epidemic outbreaks and the optimal lockdown area: a spatial normative approach

Davide La Torre[1] · Danilo Liuzzi[2] · Simone Marsiglio[3]

Received: 15 March 2022 / Accepted: 28 July 2023 / Published online: 17 August 2023

Abstract

Infectious diseases generate heterogeneous economic and health impacts within countries, thus it is essential to account for the spatial dimension in the design of epidemic management programs. We analyze the optimal regional policy to contain the spread of a communicable disease in a spatial framework with endogenous determination of the regional borders characterizing which policy regime will prevail. Specifically, the social planner needs to choose how to split the entire spatial economy in a number of regions in which a different combination of lockdown and treatment measures will be employed: in some region the only mitigation instrument will be treatment, while in some other treatment will be accompanied by a partial lockdown. We characterize the optimal solution both in an early and an advanced epidemic setting, showing that according to the circumstances it may be convenient either to partition the spatial economy in multiple regions with differentiated policies or to consider it a unique region subject to the same policy measure. Moreover, we show that from a normative perspective it is rather difficult to understand how to effectively determine the optimal size of a lockdown area (and thus of the lockdown intensity) since this critically depends on a number of factors, including the initial spatial distribution of disease

We are grateful to two anonymous referees for their constructive comments on an earlier draft. All remaining errors and omissions are our own sole responsibility. Simone Marsiglio's research has been supported by the Italian Ministry of University and Research as a part of the PRIN 2017 (Grant Protocol Number 2017FKHBA8), and the PRIN 2022 (Grant Protocol Number 20227339KC) programs.

✉ Davide La Torre
davide.latorre@skema.edu

Danilo Liuzzi
danilo.liuzzi@unimi.it

Simone Marsiglio
simone.marsiglio@unipi.it

1 SKEMA Business School and Universite Cote d'Azur, Sophia Antipolis, Valbonne, France

2 Department of Economics, Management and Quantitative Methods, University of Milan, Milan, Italy

3 Department of Economics and Management, University of Pisa, Pisa, Italy

prevalence, the amount of resources diverted from one region to the other, and the possible spatio-temporal evolution of the disease.

Keywords Macroeconomic-epidemiological model · Optimal lockdown area · Regional policy · Spatio-temporal dynamics

JEL Classification C60 · E20 · I10

1 Introduction

The ongoing COVID-19 pandemic is ravaging the world showing more clearly than ever than infectious diseases may generate devastating consequences both in industrialized and developing countries. By affecting the labor market, education attainment, fertility, life expectancy and civil unrest, communicable diseases impact economic activities through a variety of microeocnomic and macroeconomic channels (Bloom et al. 2004; Acemoglu and Johnson 2007; Bleakley 2007; Adda 2016; Cervellati et al. 2017; Klasing and Milionis 2020). Such effects generate important short run implications determining the availability of resources to finance private and public investments, ultimately harming the long run prospects of economic growth and development (Boucekkine et al. 2009a, b; Lopez et al. 2006; WHO 2009). This explains why the eradication of infectious diseases is a primary social objective, as confirmed by the inclusion in the UN's 2030 sustainable development goals of specific targets related to the elimination or reduction of HIV, tuberculosis, hepatitis, malaria, water-borne and tropical illnesses (UN 2015). In order to contain the spread of communicable diseases and mitigate their economic consequences a range of publicly-funded pharmaceutical measures, broadly categorized as forms of prevention (aiming to reduce disease incidence) and treatment (aiming to reduce disease prevalence), have been employed almost everywhere in the world. The role of such control policies along with their effectiveness as mitigation strategies have been extensively discussed in the economic epidemiology literature (Goldman and Lightwood 2002; Gersovitz and Hammer 2004; Anderson et al. 2010; La Torre et al. 2023; see also the surveys by Philipson 2000; Gersovitz and Hammer 2003). Less known is the impact of non-pharmaceutical policy tools, recently employed on a large scale to fight the COVID-19 pandemic, which represent a special form of preventive measures that do not only improve epidemiological outcomes but also deteriorate economic conditions.[1]

COVID-19 is a highly contagious virus-induced communicable disease for which human-to-human transmission occurs though unprotected contacts between infective and suceptible individuals (WHO 2020a). The epidemic has origin in China in late 2019 and the disease has spread rapidly both within and between countries reaching a pandemic status in a few months, resulting thus far (in March 2023) in over

[1] Non-pharmaceutical interventions have probably been employed for the first time during the sixth century in the Roman empire to contain the buboinic plague epidemic (also known as Justinian's plague), when isolation and quarantine measures have been introduced to fight the diffusion of the disease (Sarris 2022). Since then non-pharmaceutical interventions have been frequently used everywhere in the world to limit the spread of deadly infectious diseases, such as the plague, cholera, ebola virus, SARS.

430 million confirmed cases and nearly 6 million deaths globally (Dong et al. 2020). On top of the traditional prevention and treatment measures, a variety of alternative non-pharmaceutical policy actions have been taken all over the world to contain the epidemic, including the introduction of social and physical distancing, the imposition of mobility constraints, and the adoption of voluntary isolations and lockdowns (Cheng et al. 2020). Despite their beneficial effects in reducing the disease incidence by decreasing the probability of transmission, such policy measures have also generated dramatic economic effects by imposing stringent restrictions on individuals and firms' behavior resulting in sharp reductions in GDP and sharp increases in job losses worldwide. The impacts of the disease and the policy measures implemented to contain it have been highly diverse between countries and industries, with those extensively relying on contact-intensive activities suffering the most (OECD 2020a). Understanding thus how relying on non-pharmaceutical intervention tools to balance the needs to preserve health conditions and to support economic activities in the shadow of the COVID-19 pandemic is a current priority for policymakers, and our goal in this paper is to contribute to this important issue by developing a normative approach to characterize how policy measures should be determined in order to account for the geographical features of economic and epidemiological outcomes.

Indeed, the impacts of COVID-19 have been highly heterogeneous not only between countries but also within countries, showing large differences both at regional and local levels (Thomas et al. 2020; Amdaoud et al. 2021; Francetic and Munford 2021). For example, in Italy, one of the countries suffering the most in Europe, the northern regions have been hit hardest, with Lombardy (the wealthiest region in the country) registering the largest number of cases and deaths (Bourdin et al. 2021). Therefore, in order to limit the detrimental consequences on national economic activity, it is essential to understand how to discriminate the form and intensity of mitigation strategies at subnational level, as the spatial dimension plays an important role in the design of effective epidemic management programs (Della Rossa et al. 2020; Desmet and Wacziarg 2021). Despite initially the policy response to the growing spread of the disease has been uniformly applied at the national level, in an attempt to balance the competing economic and epidemiological needs policymakers have opted more and more frequently for policy heterogeneity (Kraemer et al. 2020; OECD 2020b). Several measures, including mobility restrictions and lockdowns, have been implemented with variable intensity and timing at the different subnational levels, trying to closely relate their severity and duration with the local level of disease prevalence (Cheng et al. 2020; Keystone 2020).[2] All over the world, the implemented policies have been differentiated at various (state, interregional, regional, municipal) levels and policymakers have been facing the critical problem to determine at which specific geographical level the alternative policy measures should be applied (OECD 2020b). In order to address

[2] For example, in Italy the national territory has been divided in four main (red, orange, yellow and white) areas, with the severity of the mitigation measures changing according to which zone a specific region or municipality belongs to, such that while in some areas businesses have been allowed to regularly run and individuals' mobility to normally occur (apart from the need to ensure physical distancing and to enforce the use of personal protective equipments), in others unnecessary businesses have been forced to closure and individuals' mobility completely forbidden (Sanfelici 2020).

such an important research and policy question we develop a spatial macroeconomic-epidemiological framework to characterize how policy measures should vary between regions by endogenously determining the regional borders characterizing which policy regime is applied within the different regions.

Specifically, we analyze a spatially structured optimal control problem in which the national economy is represented by a continuum of locations distributed over a closed interval, and the global policymaker chooses the regional borders delimiting the area (i.e., the set of locations) in which a specific policy mix will prevail. In particular, consistent with the COVID-19 policy response implemented in several countries (such as Italy and China, just to mention two countries which have been deeply exposed to the disease), the social planner needs to split the entire spatial economy in a number of regions in which a different combination and intensity of lockdown and treatment measures will be employed: in some region the only mitigation instrument will be treatment, while in some other treatment will be accompanied by a partial lockdown whose intensity needs to be optimally determined. Moreover, restrictions on individuals' mobility are in place to prevent agents to travel between regions, such that the different regions are isolated from an epidemiological perspective since disease transmission may occur only within a single region. However from an economic point of view the different regions are interconnected as the need to finance extra-treatment in the regions with the largest degree of disease prevalence is fulfilled by partly relying on the tax revenue collected in the other regions. By determining the optimal size of the different regions the social planner affects not only the number of locations in which the lockdown is applied but also the amount of resources available for treatment in all locations. In this setting we characterize the optimal lockdown policy along with the optimal size of the different regions, analyzing how the results may change in early and advanced epidemic frameworks. We show that according to the specific circumstances, it may be convenient either to partition the spatial economy in multiple regions with differentiated policies or to consider it as a unique region subject to the same policy measures. Moreover, we show that from a normative perspective it is rather difficult to understand how to effectively determine the optimal size of a lockdown area (and thus the lockdown intensity) since this critically depends on a number of factors, including the initial spatial distribution of disease prevalence, the amount of resources diverted from one region to the other, and the possible spatio-temporal evolution of the disease.

Our paper mainly contributes to two different literature streams. Clearly our work is closely related to the macroeconomic epidemiology literature which aims to analyze the mutual implications between epidemics and aggregate economic activity (Goenka and Liu 2012, 2019; Goenka et al. 2014; La Torre et al. 2020), in particular in the light of the ongoing COVID-19 pandemic (Acemoglu et al. 2021; Alvarez et al. 2021; Atkeson 2020; La Torre et al. 2021a, 2022; Eichenbaum et al. 2021). Most of these studies adopt a normative point of view to determine the optimal policy to contrast the spread of a communicable disease by accounting for the macroeconomic consequences of different policies including the fact that the availability of resources to finance them depends on the level of disease prevalence. With a few exceptions which analyze the spatial implications of such epidemiological-macroeconomic issues abstracting from the determination of the optimal dimension of a lockdown area (La

Torre et al. 2022; Camacho et al. 2023), the discussion of the optimal policy at different geographical levels has never been explored thus far, which is instead a distinctive feature of our work aiming to characterize how mitigation policy should be differentiated between regions and locations.[3] Such a spatial focus closely relates our work also to the macroeconomic geography literature which discusses the geographical implications of economic activities, with particular emphasis on capital accumulation (Camacho et al. 2008; Boucekkine et al. 2009a, b, 2013, 2019; La Torre et al. 2015) and environmental problems (Brock and Xepapadeas 2008, 2010, 2017; de Frutos and Martin-Herran 2019a, b; La Torre et al. 2019a, b, 2021b). These studies characterize how the optimal economic policy may change across space to account for the eventual heterogeneity in the features of different locations and for the spatial externalities associated with cross-location flows of resources, distinguishing eventually between global and local approaches characterized by the presence or absence of regional coordination, respectively. To the best of our knowledge, none of them allows for an endogenous determination of the borders of the different geographical areas which instead are assumed to be exogenously given, while the possibility to determine the size of the geographical areas in which a specific policy is implemented is an important peculiarity of our setup.

The paper proceeds as follows. Section 2 discusses a spatial extension of the traditional SIS epidemiological model pointing out the differences arising from spatial and a-spatial settings. Section 3 presents our baseline two-regions spatial macroeconomic-epidemiological model in which policymakers determine not only the optimal intensity of the policy measures implemented but also the geographical areas in which specific policies are employed. Section 4 focuses on the early stage of the epidemic in which the disease prevalence is rather limited and the disease dynamic equation can be approximated by a linear partial differential equation, allowing thus for a closed-form characterization of the solution. Section 5 focuses on the advanced stage of the epidemic in which such an approximation is not applicable since disease prevalence is no longer negligible, and thus the disease dynamics is described by a quadratic partial differential equation, precluding the possibility to obtain closed-form solutions. Section 6 generalizes our analysis by considering three regions and different combinations of lockdown and non-lockdown areas, showing how increasing the number of regions in which to partition the entire spatial economy may affect our conclusions. Section 7 presents some extensions of our baseline two-region model to account for the effects of the policy measures implemented in order to control the spread of COVID-19 and for some of its epidemiological peculiarities, showing that our framework can account also for important realistic consequences of mitigation policies along with realistic features of the disease. Sect. 8 as usual concludes and suggests directions for

[3] A limited number of studies introduces a spatial dimension in the analysis of optimal lockdown strategies by considering that the economy is structured in a network and the nodes represent different local units (Birge et al. 2020; Bisin and Moro 2020; Bognanni et al. 2020; Cuñat and Zymek 2020; Fajgelbaum et al. 2020; Giannone et al. 2020). These works focus on how the interactions between geographical units affect the epidemic dynamics, showing that lockdowns targeted at the different local levels are more cost-effective than uniform lockdowns as they allow to reduce the economic losses associated with mitigation policies. Different from these papers in which the analysis is carried out in a purely numerical setting precluding the possibility to understand the different mechanisms at work, our approach allows for an analytical characterization of the lockdown intensity at different local levels clarifying the role of spatial interactions.

future research. All technicalities are presented either in Appendix A, or in the online appendix.

2 The epidemiological framework

The SIS model is one of the simplest and most widely used framework in mathematical epidemiology (Kermack and McKendrick 1927; Hethcote 2000), allowing to describe the evolution of a number of infectious diseases which do not confer permanent immunity after recovery. It can thus be used to analyze the spread of common diseases, such as the seasonal flu and the common cold, but also of emerging diseases, such as COVID-19 since thus far there exists no compelling evidence that *"people who have recovered from COVID-19 and have antibodies are protected from a second infection"* (WHO 2020b). We introduce a SIS setting into a spatial context in which the economy develops along a line (Hotelling 1929) and the points on the line represent different geographical units, such that the population may move across different locations and thus the diseases may spread across the different geographical units (Wang 2014).

We denote with $N_{x,t}$, $S_{x,t}$ and $I_{x,t}$, respectively the population, the susceptibles and the infectives in the position x at date t, in a compact interval $[x_a, x_b] \subset \mathbb{R}$, and we assume that the Neumann boundary conditions hold true, that is there are no mobility flows through the borders of $[x_a, x_b]$ namely the directional derivatives are null, $\frac{\partial N_{x,t}}{\partial x} = \frac{\partial S_{x,t}}{\partial x} = \frac{\partial I_{x,t}}{\partial x} = 0$, at $x = x_a$ and $x = x_b$. Every individual in the population in each location can be either suceptible or infective, $N_{x,t} = S_{x,t} + I_{x,t}$: susceptibles become infectives by interacting with other infectives and $\alpha \geq 0$ measures the infectivity rate (i.e., the average number of contacts required in order to give rise to a new infection), and infectives become susceptible after recovering from the disease and $\delta \geq 0$ is the recovery rate. In each location the evolution of susceptibles and infectives can be described through the following system of partial differential equations:

$$\frac{\partial S_{x,t}}{\partial t} = d\frac{\partial^2 S_{x,t}}{\partial x^2} + \delta I_{x,t} - \alpha S_{x,t}\frac{I_{x,t}}{N_t} \tag{1}$$

$$\frac{\partial I_{x,t}}{\partial t} = d\frac{\partial^2 I_{x,t}}{\partial x^2} + \alpha S_{x,t}\frac{I_{x,t}}{N_t} - \delta I_{x,t}. \tag{2}$$

where $N_t = \frac{1}{x_b - x_a}\int_{x_a}^{x_b} N_{x,t}dx$ represents the average population size within the whole spatial economy. In the above equations the term $S_{x,t}\frac{I_{x,t}}{N_t}$ states that human interactions do not change with the spread of the disease, and thus what determines the disease transmission is the share of infected population, $\frac{I_{x,t}}{N_t}$, rather than the total number of infectives, $I_{x,t}$. Moreover, the diffusion terms $d\frac{\partial^2}{\partial x^2}$ captures the existence of spatial externalities, describing how individuals' mobility within the spatial economy may lead the disease to spread geographically even in locations far away from the location of the initial outbreak. In particular, diffusion describes the effects of the demographic changes associated with migration which leads individuals to permanently move from

one location to the next (La Torre et al. 2022). The parameter $d = d_I = d_S \geq 0$ measures the speed at which such cross-location effects take place, which for the sake of simplicity is assumed to be the same for both infectives and susceptibles.

Since $N_{x,t} = S_{x,t} + I_{x,t}$, by summing the above equations it follows that the population in location x satisfies the classical heat equation:

$$\frac{\partial N_{x,t}}{\partial t} = d\frac{\partial^2 N_{x,t}}{\partial x^2} \tag{3}$$

subject to the Neuman boundary conditions and $N_{x,0} = I_{x,0} + S_{x,0}$. This implies that the average population size within the whole spatial economy is constant over time since:

$$\frac{d}{dt}N_t = \frac{1}{x_b - x_a}\int_{x_a}^{x_b}\frac{\partial}{\partial t}N_{x,t}dx = \frac{1}{x_b - x_a}d\int_{x_a}^{x_b}\frac{\partial^2}{\partial x^2}N_{x,t}dx = 0, \tag{4}$$

thus in the following we shall denote the average population size with the constant N, that is $N_t = N, \forall t$.

Equation (3) states that human population moves from locations more populated to those less populated, and so do susceptibles and infectives (see Eqs. (1) and (2)) which are two sub-population groups which make up the whole population.[4] Migration towards less densely populated locations, on the one hand, decreases the probability of infection of the single individual and, on the other hand, increases the degree of infection in the locations receiving such migrants (La Torre et al. 2022). Therefore, such a characterization of population mobility and infection spread is well suitable to describe the spatial disease dynamics during an epidemic outbreak, in which individuals' attempt to reduce their disease exposure increases the probability of infection of other individuals.[5]

To analyze the disease dynamics it is convenient to recast the model in terms of the share of infectives, $i_{x,t} = \frac{I_{x,t}}{N}$, and the share of suscetibles, $s_{x,t} = \frac{S_{x,t}}{N}$, which is

[4] We assume that the mobility patterns of infectives and susceptibles coincide, thus also infectives can move from one location to the next. This is generally the case with common diseases (such as the seasonal flu or the common cold) while with highly contagious and deadly diseases (such as COVID-19), infectives' mobility might be precluded by hospitalization or limited by public regulation. In reality though, even if public regulations might require non-hospitalized infectives to be quarantined or self-isolating, individual compliance is not perfect and many individuals (especially asymptomatics) may be unaware of their infectivity status, meaning that infectives' spatial mobility represents an important determinant of disease spreading. It seems interesting thus in our baseline framework to focus on the implications of spatial mobility on epidemic dynamics and the mitigation policies in a context in which both susceptibles and infectives move across locations. We postpone to later the discussion of how results may change when the mobility patterns of infectives and susceptibles differ (see Sect. 7).

[5] Modeling mobility in a different way to allow for population to get more and more concentrated in given locations would give rise to the counterintuive result in which because of their own migration decisions people end up increasing their individual probability of infection. The implication of such an alternative modeling approach is clearly not consistent with real world experiences during major epidemic outbreaks, thus our modeling approach seems the most convenient to characterize the spatial spread of a communicable

Footnote 5 continued

disease. Note that a similar framework is typically employed also in mathematical epidemiology (Martcheva 2009; Anita and Capasso 2017).

summarized by the following dynamic equations with boundary and initial conditions:

$$\frac{\partial s_{x,t}}{\partial t} = d\frac{\partial^2 s_{x,t}}{\partial x^2} + \delta i_{x,t} - \alpha s_{x,t} i_{x,t}, \quad x \in (x_a, x_b),\ t > 0 \tag{5}$$

$$\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t} i_{x,t} - \delta i_{x,t}, \quad x \in (x_a, x_b),\ t > 0 \tag{6}$$

$$\frac{\partial s_{x,t}}{\partial x} = \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b\}, \quad t > 0 \tag{7}$$

$$s_{x,0} = s_0(x), \quad i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \tag{8}$$

By assuming that s_0 and i_0 are continuous on $[x_a, x_b]$, $s_0(x) > 0, i_0(x) > 0$ for any $x \in [x_a, x_b]$, the theory for parabolic equations (Lieberman 1996) ensures that the above boundary value problem admits a unique classical solution $s_{x,t}, i_{x,t} \in C^{2,1}([x_a, x_b] \times (0, \infty))$. Moreover, the strong maximum principle and the Hopf boundary lemma for parabolic equations (Protter and Weinberger 1984) establish that both $s_{x,t}$ and $i_{x,t}$ are nonnegative for all $x \in [x_a, x_b]$ and $t > 0$ (Wang et al. 2015). Indeed, by being defined in terms of the average population size in the spatial economy, the above shares of infectives and suscetibles turn out to be only nonnegative and not naturally limited by one.

Since a new disease outbreak does not lead to an immediate policy response, for a certain period of time the disease spreads freely within the population and thus the share of infectives tends to grow at a constant (or even decreasing) rate. In order to characterize such a specific situation several studies analyze the disease dynamics in an early epidemic setting by assuming that the evolution of the share of infectives is described by a linear equation (Chowell et al. 2016). In our spatial context such an early epidemic framework translates into assuming that the share of susceptibles is approximately constant, namely $s_{x,t} \simeq s, \forall x, t$. This implies that it is possible to characterize the model's outcome by analyzing only the evolution of the disease prevalence, which is given by the following equation:

$$\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s i_{x,t} - \delta i_{x,t} \tag{9}$$

As discussed in literature the above approximation is applicable only in situations in which the disease prevalence does not get too large, which in a spatial context means that the total prevalence within the whole spatial economy, $i_t = \int_{x_a}^{x_b} i_{x,t} dx$, remains small over time. Therefore, an early epidemic setting can be used to describe the disease dynamics only in contexts in which the total prevalence shows a nonincreasing pattern converging eventually to a situation in which eradication will be achieved in the spatial economy.

After an early stage in which the disease spreads freely within the population in the absence of policymakers' containment efforts, the disease prevalence may get too large for the early epidemic approximation to keep holding true. In such an advanced epidemic stage, in order to analyze the model's outcome, we need to accompany the dynamics of the disease prevalence with that of the share of population in location x. Indeed, since $n_{x,t} = s_{x,t} + i_{x,t}$, the shares of susceptibles is automatically determined

once the disease prevalence and the local population are known. By exploiting the fact that $s_{x,t} = n_{x,t} - i_{x,t}$, we need to analyze the following system of partial differential equations:

$$\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(n_{x,t} - i_{x,t})i_{x,t} - \delta i_{x,t} \tag{10}$$

$$\frac{\partial n_{x,t}}{\partial t} = d\frac{\partial^2 n_{x,t}}{\partial x^2}. \tag{11}$$

where $n_{x,t}$ is simply $N_{x,t}$ scaled down for the constant N. The expression of $n_{x,t}$ satisfies the classical heat equation with Neumann boundary conditions, and its solution is well known (see James et al. 2020) and given by the following expression:

$$n_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{x_b - x_a}\right)^2 dt} \cos\left[\frac{n\pi(x - x_a)}{x_b - x_a}\right], \tag{12}$$

where $B_0 = \frac{1}{x_b - x_a}\int_{x_a}^{x_b} n_{x,0}dx$ and $B_n = \frac{2}{x_b - x_a}\int_{x_a}^{x_b} n_{x,0}\cos\left[\frac{n\pi(x - x_a)}{x_b - x_a}\right]dx$ are the Fourier coefficients. In particular, the closed-form expression above implies that $n_{x,t} \to B_0$ whenever $t \to \infty$, which allows us to determine the equilibria of the system (10) - (11). It is straightforward to show that the system above admits two homogeneous equilibria $\mathcal{E} = (\bar{i}, \bar{n})$, given by the two pairs:

$$\begin{aligned} \mathcal{E}^F &: \bar{i}^F = 0, \quad \bar{n}^F = 1 \\ \mathcal{E}^E &: \bar{i}^E = \frac{\alpha - \delta}{\alpha}, \quad \bar{n}^E = 1 \end{aligned} \tag{13}$$

The former represents a disease-free equilibrium as the disease prevalence is null ($\bar{i}^F = 0$) in each location in the spatial economy and it exists for all parameter values. The latter instead is an endemic equilibrium since the disease prevalence is strictly positive ($\bar{i}^E = \frac{\alpha - \delta}{\alpha} > 0$) and it exists only if $\alpha > \delta$. Moreover, when only $\mathcal{E}^F$ exists this equilibrium is asymptotically stable, while when both equilibria exist $\mathcal{E}^F$ loses its stability and $\mathcal{E}^E$ becomes asymptotically stable. Therefore, according to the parameters configuration the economy may alternatively converge to either a disease-free equilibrium (if $\alpha \leq \delta$) or an endemic equilibrium ($\alpha > \delta$). This conclusion is consistent with the traditional result in mathematical epidemiology stating that the prevailing equilibrium depends on the value of the "basic reproduction number", $\mathcal{R}_0$, quantifying the average number of secondary infections produced by a typical infectious individual introduced into a completely susceptible population. The basic reproduction number is given by the ratio between the infectivity rate and the recovery rate as follows:

$$\mathcal{R}_0 \equiv \frac{\alpha}{\delta} \tag{14}$$

Also in our spatial framework, exactly as in an a-spatial context, the threshold value of $\mathcal{R}_0$ determining which equilibrium will emerge is equal to 1: if $\mathcal{R}_0 \leq 1$ the prevailing equilibrium will be disease-free while if $\mathcal{R}_0 > 1$ it will be endemic. Intuitively, any single infection needs to generate a multiplier effect in order to allow the disease to become endemic persisting thus in the long run. As extensively discussed both in the mathematical epidemiology and economic epidemiology literature, since pharmaceutical and non-pharmaceutical mitigation tools affect either the infectivity rate (i.e., prevention) or the recovery rate (i.e., treatment) they may be effectively employed in order to bring and maintain the basic reproduction number below unity allowing thus to achieve disease eradication in the long run.

Before introducing our economic framework, some further comments on the assumptions underlying our epidemiological setting are needed. First of all, even if the SIS framework can be used to describe the spread of COVID-19 in a simplified and intuitive way, it clearly does not provide a precise representation of its dynamics since infection from COVID-19 seems to confer some temporary (few-months long) immunity which would require to add a further group of individuals (i.e., the recovereds) who, different from the susceptibles, cannot become infectives for a certain period of time. Moreover, the fact the spatial spread of the disease is driven entirely by demographic dynamics may represent well the spatial pattern of infections in developing and less developed countries where people's movements are limited, while it may not represent well the situation in industrialized countries in which individuals regularly commute between locations for business or leisure favoring thus the spatial propagation of a disease. In order to take these issues into account we would need to introduce the presence of an additional sub-population group and local effects which partly drive the spatial spread of infection. Since these additional features will substantially complicate the model's structure precluding the possibility to obtain closed-form solutions, it seems convenient to present our setup in its simplest possible form first, postponing to later a formal discussion of how our baseline model might need to be modified in order to describe more rigorously some epidemiological features of COVID-19 in industrialized countries, including the possibility of temporary immunity and disease spreading faster than demography (see Sect. 7).

3 The economic framework

We now extend the above spatial SIS framework to account for the role of public policies and to analyze its relation with macroeconomic outcomes. We thus present a spatial macroeconomic-epidemiological setup in which the disease prevalence drives output production and both non-pharmaceutical (i.e., lockdown) and pharmaceutical (i.e., treatment) measures are used to manage the epidemic. While lockdowns do not require funding, treatment is funded via income taxation which however depends on disease prevalence, since by affecting output it determines the availability of resources to finance pharmaceutical interventions. Such feedback effects between health and macroeconomic outcomes are complicated by the spatial interactions within and between different geographical units.

We analyze a short time horizon setting in which the social planner decides which policy measures to implement in order to reduce the spread of a communicable disease, along with the geographical area of applicability of such policies, in order to minimize the social cost associated with the epidemic management program. We consider therapeutic treatment, $0 < v_{x,t} < 1$ which increases the speed of recovery of the infectives decreasing disease prevalence, and a lockdown which limits the social contacts by a percentage $0 < u_{x,t} < 1$ reducing thus disease incidence. Over a short time horizon saving and capital accumulation are irrelevant while fiscal policy can be taken as exogenously given because of its implementation delay, thus we simply assume that in each location individuals entirely consume their disposable income as follows: $c_{x,t} = (1-\tau)y_{x,t}$, where $c_{x,t}$ denotes consumption, $y_{x,t}$ income and $0 < \tau < 1$ the exogenous and constant tax rate. The tax revenue is allocated to finance treatment, thus pharmaceutical policy interventions are completely publicly provided by maintaining a balanced budget position. Output is produced through a linear production function by susceptibles as follows: $q_{x,t} = s_{x,t}$.

The planner needs to choose how to split the entire spatial economy in a number of regions in which different combinations of the pharmaceutical and non-pharmaceutical interventions will be implemented. For the sake of expositional simplicity we consider only two regions for the time being. In one region (which we refer to as region $A = [x_a, \xi]$ for expositional simplicity) both treatment and lockdown will be used, while in the other (region $B = [\xi, x_b]$) only treatment will be used. In region A, which develops from x_a to ξ, only a certain share of the social contacts, $1-u_{x,t}$, is allowed to regularly occur, thus output net of lockdown measures is given by: $y_{x,t} = (1-u_{x,t})q_{x,t}$. The tax revenue, $\tau y_{x,t}$, is entirely used to finance treatment locally, and some extra resources from treatment are collected from region B, $r_{x,t}$, such that $v_{x,t} = \tau y_{x,t} + r_{x,t}$. The disease dynamics is described by a SIS equation as follows: $\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1-u_{x,t})s_{x,t}i_{x,t} - \delta(1+\omega v_{x,t})i_{x,t}$, where $\omega > 0$ measures the effectiveness of treatment in speeding up recovery. In region B, which develops from ξ to x_b, there is no lockdown thus output in each location is determined by the unconstrained supply: $y_{x,t} = q_{x,t}$. The tax revenue, $\tau y_{x,t}$, is entirely used to finance treatment, but only a part $0 \leq \beta \leq 1$ is employed locally while the remaining part is allocated to finance extra treatment in the region A. The disease dynamics is described by a SIS equation as follows: $\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t}i_{x,t} - \delta(1+\omega v_{x,t})i_{x,t}$, where $v_{x,t} = \beta\tau q_{x,t}$. The total amount of tax revenues diverted from region B, $(1-\beta)\tau \int_B q_{x,t}dx$, is equally split within region A thus each location $x \in [x_a, \xi]$ receives a share $\frac{1}{\xi - x_a}$ of the total: $r_t = \frac{(1-\beta)\tau}{\xi - x_a}\int_B q_{x,t}dx$. In both regions the disease is characterized by the same features and we assume that the infectivity rate is larger than the recovery rate (i.e., $\alpha > \delta$) such that the basic reproduction number in the absence of public intervention is larger than one (i.e., $\mathcal{R}_0 > 1$) and thus public policy is effectively needed in order to eventually achieve long run disease eradication. We also assume that the initial prevalence is strictly positive over the whole spatial domain (i.e., $i_{x,0} = i_0(x) > 0$), which thanks to evolution operator theory guarantees that disease prevalence is well defined (that is $i_{x,t} > 0, \forall x, t$).

The social planner needs to determine the size of the two regions in order to minimize the social cost of the epidemic management program, which is given by the

discounted sum ($\rho > 0$ is the discount rate) of the instantaneous losses in regions A and B during the duration of the program, augmented for the final damage in the entire spatial economy at the end of the program. In region A, the instantaneous loss function depends on the spread of the disease and the output lost due to the lockdown measure, and takes a quadratic non-separable form as follows: $\ell(i_{x,t}, u_{x,t}q_{x,t}) = \frac{i_{x,t}^2(1+u_{x,t}^2 q_{x,t}^2)}{2}$, penalizing deviations from the disease-free status and from the no-production-loss scenario. In region B, as there is no lockdown the instantaneous loss function depends only on the level of disease prevalence: $\ell(i_{x,t}) = \frac{i_{x,t}^2}{2}$. Since in region B the public health intervention is rather limited and some of its resources are diverted to region A, the instantaneous losses in region B are weighted by its size-adjusted importance with respect to region A's, given by $\frac{\mu}{\xi - x_a}$ where $\mu > 0$ and $\xi - x_a > 0$ measure the relative importance and size of the region B, respectively. The final damage function depends only of the level of disease prevalence at the end of the epidemic management program and takes a quadratic form as follows: $d_T = \frac{i_{x,T}^2}{2}$. The relative weight of the final damage in terms of the instantaneous losses is measured by the degree of health concerns, $\phi > 0$.

The social planner needs also to determine the size of the regions by determining their boundary point $\xi \in (x_a, x_b)$, thus the optimization problem can be stated as follows: Find $\xi \in (x_a, x_b)$ which minimizes the optimal value:

$$
\begin{aligned}
\mathcal{C}(\xi) = \min_{u_{x,t} \in \mathcal{U}} & \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2[1 + u_{x,t}^2 s_{x,t}^2]}{2} e^{-\rho t} dx dt \\
& + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i_{x,t}^2}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\
s.t. \quad & \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1 - u_{x,t}) s_{x,t} i_{x,t} \\
& - \delta \left[1 + \omega\tau(1 - u_{x,t}) s_{x,t} + \frac{(1-\beta)\omega\tau}{\xi - x_a} \right. \\
& \left. \int_{\xi}^{x_b} s_{x,t} dx \right] i_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t} i_{x,t} - \delta[1 + \beta\omega\tau s_{x,t}] i_{x,t}, \quad x \in [\xi, x_b] \\
& \frac{\partial n_{x,t}}{\partial t} = d \frac{\partial^2 n_{x,t}}{\partial x^2} \\
& s_{x,t} = n_{x,t} - i_{x,t} \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& \frac{\partial n_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& i_{x,0} = i_0(x) > 0, \quad x \in [x_a, x_b] \\
& n_{x,0} = n_0(x) > 0, \quad x \in [x_a, x_b]
\end{aligned} \tag{15}
$$

where the set of admissible controls is given by:

$$\mathcal{U} = \{u_{x,t} : [x_a, \xi] \times [0, T] \to \mathcal{R} : u_{x,t} \text{ is continuous}, 0 < u_{x,t} < 1\} \quad (16)$$

Since there is no control over the region $B = [\xi, x_b]$ we can extend any element in $\mathcal{U}$ to be zero over such a region. We assume that $i_{x,t}$ is piecewise smooth and integrable, that is it belongs to $C^{2,1}((x_a, \xi) \times (0, T)) \cap L^1([x_a, \xi] \times [0, T])$ and $C^{2,1}((\xi, x_b) \times (0, T)) \cap L^1([\xi, x_b] \times [0, T])$, and similarly that also both $i_0(x)$ and $n_0(x)$ are piecewise continuous and integrable, that is they are in $C((x_a, \xi) \cap L^1([x_a, \xi])$ and $C((\xi, x_b) \cap L^1([\xi, x_b])$. By applying the same argument used for the system (5) - (8), it is possible to conclude that the boundary value problems over both the regions $A = [x_a, \xi]$ and $B = [\xi, x_b]$ are well-posed and admit a unique positive solution $i_{x,t}, s_{x,t}, n_{x,t}$ whenever $\alpha > \delta\beta\omega\tau$ and $\alpha > \delta\omega\tau$, respectively. This follows from the considerations similar to those discussed earlier along with a comparison theorem for reaction–diffusion equations (Friedman 2008). Since $\beta \in [0, 1]$, it follows that $\delta\omega\tau > \delta\beta\omega\tau$ and, therefore, both boundary value problems for the state equations are well-posed under the common condition $\alpha > \delta\omega\tau$ across the two regions, which we will assume to hold true in the rest of the paper.

The strict convexity of the objective function in $\mathcal{C}(\xi)$ implies that the solution of the above optimal control problem exists and is unique for any fixed $\xi \in (x_a, x_b)$ (Casas et al. 2018). Note that the boundedness of the function $i_{x,t}$ (due to the fact that $i_{x,t} \leq n_{x,t}$ and $n_{x,0}$ is continuous) implies that the the middle integral goes to $+\infty$ when $\xi \to x_a^+$. Furthermore, thanks to the continuity of both $i_{x,t}$ and $u_{x,t}$ we have that the integrands in the first two integrals are bounded (and then integrable). By using the property that the Lebesgue integral of an integral function is continuous and the well-posedness of the state equations, we get that $\mathcal{C}(\xi)$ is a continuous function for any $\xi \in (x_a, x_b]$. Therefore $\mathcal{C}(\xi)$ attains a global minimum on each compact interval $[k, x_b]$ with $k > x_a$. The fact the optimal solution may need to be found for values of $\xi > x_a$ implies that in our setting at least a minimal lockdown area is required to contain the spread of the disease ($\xi = x_b$ represents the case in which the region B does not exist).

Note also that by choosing the boundary point between the two regions, the social planner endogenously determines the weight to attach to the instantaneous losses experienced by region B and the point at which mobility between locations is interrupted. Because of this interruption in cross-regional movements, at the boundary point between regions some extra Neumann boundary conditions are applied (i.e., $\frac{\partial i_{x,t}}{\partial x} = \frac{\partial n_{x,t}}{\partial x} = 0$), which means that the two regions are a smaller-scale copy of the entire spatial economy. Note also that the problem above can be seen as a particular type of Stefan problems or variable boundary problems (Stefan 1899; Rubinstein 1971; Friedman 1968), which represent general frameworks in which the boundary between regions changes over time. In our formulation we consider the separating boundary between two regions to be endogenously determined and constant over the period $[0, T]$,[6] thus in our setup the boundary does not change over time but it is

[6] Our model can be easily extended to consider a multi-period framework, in which there exists a sequence of time horizons over which the planner determines the boundary point between the two regions, eventually

determined optimally by the social planner, and to the best of our knowledge a similar problem has never been analyzed in economics thus far.

Before proceeding with our analysis, let us add a few remarks on some assumptions underlying our economic framework. We have assumed that the planner wishes to partition the economy in two regions only, and clearly this is a mere simplification of reality in which instead a different (possibly large) number of regions may be considered as well. In order to set up a baseline framework for our analysis, it seems convenient to present first the case in which the number of regions is limited to two, postponing to later the discussion of how our results may change in a multi-regional context (see Sect. 6 where we will consider three regions, which requires to analyze what happens whether this extra region is either a lockdown or non-lockdown area). We have also assumed the diffusion coefficients capturing the speed of the spatial spread of the epidemics is homogeneous across population groups and across regions, while it may be reasonable to consider that the recommendations for self-isolation in the case of infection and the introduction of lockdown measures reduce mobility among infectives and within the lockdown area respectively, such that both the population subpopulation groups and the regions might be characterized by different diffusion parameters. Given the short time horizon nature of the problem this will not modify our qualitative conclusions, thus also in this case it seems convenient to present our model in its simplest possible form discussing later how results may change under heterogeneity in mobility patterns across groups and across regions (see Sect. 7).

4 The early epidemic stage

We first analyze the above problem in an early epidemic setting in which, recalling what we have discussed in Sect. 2, the share of susceptibles can be thought to be constant over time and homogeneous across space, $s_{x,t} \simeq s, \forall x, t$. Under this approximation the problem boils down to the following: Find $\xi \in (x_a, x_b)$ which minimizes the optimal value:

$$\begin{aligned}
\mathcal{C}(\xi) = \min_{u_{x,t}} \; & \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2[1+u_{x,t}^2 s^2]}{2} e^{-\rho t} dx dt + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i_{x,t}^2}{2} e^{-\rho t} dx dt \\
& + \phi \int_{x_a}^{x_b} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\
s.t. \; & \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1-u_{x,t}) s i_{x,t} \\
& - \delta \left[1 + \omega\tau(1-u_{x,t})s + \frac{(1-\beta)\omega\tau}{\xi - x_a} \int_{\xi}^{x_b} s dx \right] i_{x,t}, \quad x \in [x_a, \xi]
\end{aligned}$$

revising his decisions moving from one time horizon to the next. In this case the model would consist of multiple optimization problems to be solved over a finite number N of disjoint time intervals $[0, T_1], [T_1, T_2], \ldots, [T_{N-1}, T_N]$, such that ξ might change between time intervals. This setup allows to characterize the true problem faced by policymakers in the real world in which the lockdown decisions (i.e., the lockdown intensity and the lockdown area) are determined in advanced for a certain period of time, to then be revised when new information regarding disease prevalence become available.

$$\begin{aligned}
&\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s i_{x,t} - \delta[1+\beta\omega\tau s] i_{x,t}, \quad x \in [\xi, x_b] \\
&\frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
&i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
&n_{x,0} = n_0(x) > 0 \quad x \in [x_a, x_b]
\end{aligned} \tag{17}$$

Note that in this case the well-posedness of the boundary value problems for the state equations and the strict positivity of their solution is ensured by the linearity of the epidemic equations. Moreover, since the solution can be obtained in closed-form, the value of objective functional can be calculated explicitly in order to find the value of ξ which minimizes it. Therefore, in order to determine the optimal solution of the problem (17) we first assume that the boundary point $\xi \in (x_a, x_b)$ is fixed and solve the problems in the regions $[x_a, \xi]$ and $[\xi, x_b]$, and then we determine the optimal boundary point by using the solution obtained in the previous step to compute the value of the functional and identify the value of ξ which minimizes it. Finally, note that because of the convexity of the objective functional the above optimal control problem is also well-posed (Dontchev and Zolezzi 1993).

Let us start thus from the case in which ξ is fixed. The following result (Lemma 1) presents the elements of the exponential matrix which will be used to determine the closed-form solution as a function of ξ (Theorem 1). The proofs of all the theorems, propositions, corollaries and lemmas are presented in appendix A.

Lemma 1 *By defining the following aggregated parameters:*

$$\eta_1 := s(\alpha - \delta\,\tau\omega) - \delta - \frac{\delta\,(x_b - \xi)\,(1-\beta)\,\tau\omega\,s}{\xi - x_a} \tag{18}$$

$$\begin{aligned}
\eta_2 &:= 4\,\delta^2\tau^2\omega^2 - 8\,\alpha\,\delta\omega\,\tau + 4\,\alpha^2 + 4\,{\eta_1}^2 - 4\,\eta_1\,\rho + \rho^2 \\
&= 4(\alpha - \delta\omega\tau)^2 + (\rho - 2\eta_1)^2
\end{aligned} \tag{19}$$

and considering the matrix:

$$\Theta = \begin{bmatrix} s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right) & -(\alpha - \delta\tau\omega)^2 \\ -1 & \rho - s(\alpha - \delta\tau\omega) + \delta + \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right) \end{bmatrix}, \tag{20}$$

then the entries of the exponential matrix $e^{\Theta t}$ are given by:

$$e_{11}^{\Theta t} = -\mathrm{e}^{\frac{\rho t}{2}} \frac{\left(\rho - 2\,\eta_1\right)\sinh\left(\frac{\sqrt{\eta_2}t}{2}\right) - \sqrt{\eta_2}\cosh\left(\frac{\sqrt{\eta_2}t}{2}\right)\Big)}{\sqrt{\eta_2}} \tag{21}$$

$$e_{12}^{\Theta t} = -2\mathrm{e}^{\frac{\rho t}{2}}\,(-\delta\,\tau\omega + \alpha)^2\,\frac{\sinh\left(\frac{\sqrt{\eta_2}t}{2}\right)}{\sqrt{\eta_2}} \tag{22}$$

$$e_{21}^{\Theta t} = -2\mathrm{e}^{\frac{\rho t}{2}}\,\frac{\sinh\left(\frac{\sqrt{\eta_2}t}{2}\right)}{\sqrt{\eta_2}} \tag{23}$$

$$e_{22}^{\Theta t} = \mathrm{e}^{\frac{\rho t}{2}} \frac{\left(\rho - 2\,\eta_1\right) \sinh\left(\frac{\sqrt{\eta_2} t}{2}\right) + \sqrt{\eta_2} \cosh\left(\frac{\sqrt{\eta_2} t}{2}\right)\Big)}{\sqrt{\eta_2}} \tag{24}$$

Let us notice that in the region B, where lockdown measures are not employed, the lockdown intensity is null (i.e., $u_{x,t} \equiv 0, \forall x \in [\xi, x_b]$). This allows us to express the closed-form solution in a more compact form. In order to ease readibility in the statement of our results we shall denote the lockdown intensity and the share of infectives in regions A and B with the superscript $j = \{A, B\}$. Furthermore, throughout the paper the convergence of the Fourier series is intended in Lebesgue sense.

Theorem 1 *Assuming that $\xi \in (x_a, x_b)$ is fixed. Under the regularity assumptions on the state and the control variables presented in the model statement, the optimal pair $(i_{x,t}, u_{x,t})$ solving problem (17) satisfies the following optimality conditions:*

- $i_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as*

$$i_{x,t} = i_{x,t}^A \chi_{[x_a,\xi]} + i_{x,t}^B \chi_{[\xi,x_b]} \tag{25}$$

  where χ_S is the indicator function of the set S ($\chi_S(x) = 1$ for any $x \in S$, 0 otherwise)
- $u_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as*

$$u_{x,t} = u_{x,t}^A \chi_{[x_a,\xi]} \tag{26}$$

- *The function $i_{x,t}^B$ solves the following boundary value problem:*

$$\begin{cases} \frac{\partial i_{x,t}^B}{\partial t} = d \frac{\partial^2 i_{x,t}^B}{\partial x^2} + (\alpha s - \delta - \delta\beta\tau\omega s)\, i_{x,t}^B & x \in [\xi, x_b] \\ \frac{\partial i_{x,t}^B}{\partial x} = 0 & x \in \{\xi, x_b\} \\ i_{x,0}^B = i_0(x) & x \in [\xi, x_b] \end{cases} \tag{27}$$

  whose solution is:

$$i_{x,t}^B = \sum_{n \geq 0} C_n e^{-d\left(\frac{n\pi}{x_b - \xi}\right)^2 t} e^{(\alpha s - \delta - \delta\beta\tau\omega s)t} \cos\left[\frac{n\pi(x - \xi)}{x_b - \xi}\right] \tag{28}$$

  where:

$$\begin{aligned} C_0 &= \frac{1}{x_b - \xi} \int_\xi^{x_b} i_0(x) dx, \\ C_n &= \frac{2}{x_b - \xi} \int_\xi^{x_b} i_0(x) \cos\left[\frac{n\pi(x - \xi)}{x_b - \xi}\right] dx \end{aligned} \tag{29}$$

- *The pair* $(i^A_{x,t}, u^A_{x,t})$ *solves the following optimality conditions:*

$$\begin{cases} \frac{\partial i^A_{x,t}}{\partial t} = d\frac{\partial^2 i^A_{x,t}}{\partial x^2} \\ \quad + \left[s(\alpha - \delta\omega\tau) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] i^A_{x,t} - s u^A_{x,t} i^A_{x,t}(\alpha - \delta\tau) \\ \frac{\partial \lambda^A_{x,t}}{\partial t} = \rho\lambda^A_{x,t} - d\frac{\partial^2 \lambda^A_{x,t}}{\partial x^2} - i^A_{x,t} \\ \quad -\lambda^A_{x,t}\left[s(\alpha - \delta\omega\tau) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] \\ u^A_{x,t} = \frac{\lambda^A_{x,t}(\alpha - \delta\tau\omega)}{s i^A_{x,t}} \\ \frac{\partial i^A_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ \frac{\partial \lambda^A_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ i^A_{x,0} = i_0(x) \quad x \in [x_a, \xi] \\ \lambda^A_{x,T} = \phi i^A_{x,T} \quad x \in [x_a, \xi] \end{cases} \tag{30}$$

where $\lambda^A_{x,t}$ *is the costate variable. The closed-form solution is given by:*

$$i^A_{x,t} = \sum_{n\geq 0}\left(e^{\Theta t}_{11} A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 t} + e^{\Theta t}_{12} B_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 (T-t)}\right)\cos\left(n\pi\left[\frac{x - x_a}{\xi - x_a}\right]\right) \tag{31}$$

$$u^A_{x,t} = \frac{\sum_{n\geq 0}(\alpha - \delta\tau\omega)\left(e^{\Theta t}_{21} A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 t} + e^{\Theta t}_{22} B_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 (T-t)}\right)\cos\left(n\pi\left[\frac{x - x_a}{\xi - x_a}\right]\right)}{\sum_{n\geq 0}\left(e^{\Theta t}_{11} A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 t} + e^{\Theta t}_{12} B_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 (T-t)}\right)\cos\left(n\pi\left[\frac{x - x_a}{\xi - x_a}\right]\right)} \tag{32}$$

where:

$$B_n = \left[\frac{e^{\Theta T}_{21} - \phi_2 e^{\Theta T}_{11}}{\phi_2 e^{\Theta T}_{12} - e^{\Theta T}_{22}}\right] A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 T} \tag{33}$$

$$A_0 = \frac{1}{\xi - x_a}\int_{x_a}^{\xi} i_0(x)dx,$$

$$A_n = \frac{2}{\xi - x_a}\int_{x_a}^{\xi} i_0(x)\cos\left(n\pi\left[\frac{x - x_a}{\xi - x_a}\right]\right)dx \tag{34}$$

and $e^{\Theta t}_{11}$, $e^{\Theta t}_{12}$, $e^{\Theta t}_{21}$, *and* $e^{\Theta t}_{22}$ *are the entries of the exponential matrix* $e^{\Theta t}$ *presented in Lemma* 1.

Note that, since we are assuming that ξ is fixed for the time being, provided that $i_{x,t} > 0$ and $0 < u_{x,t} < 1$ the conditions stated in Theorem 1 are also sufficient thanks to the convexity of the objective functional. In particular, the theorem determines in an explicit form the spatio-temporal dynamic path of the lockdown intensity (in region A) and disease prevalence (in both regions A and B), which are expressed in terms of cosine Fourier expansions. We can notice that both $u^j_{x,t}$ and $i^j_{x,t}$ are in general

non-constant over time and non-homogeneous across space. However, their analytical expressions are particularly cumbersome and thus it is not possible to perform any comparative statics exercise in order to understand how they depend on the different parameters. Nevertheless, they still allow us to derive some interesting results.

Corollary 1 *The following result holds true:*

$$\|i_{x,t}\|_{L^2}^2 = \|i_{x,t}^A\|_{L^2}^2 + \|i_{x,t}^B\|_{L^2}^2 \tag{35}$$

and:

$$\frac{2}{\xi - x_a}\|i_{x,t}^A\|_{L^2}^2 = 2\left(e_{11}^{\Theta t}A_0 + e_{12}^{\Theta t}B_0\right) + \sum_{n\geq 1}\left(e_{11}^{\Theta t}A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 t} + e_{12}^{\Theta t}B_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 (T-t)}\right)^2 \tag{36}$$

$$\frac{2}{x_b - \xi}\|i_{x,t}^B\|_{L^2}^2 = 2C_0^2 e^{2(\alpha s - \delta - \delta\beta\tau\omega s)t} + \sum_{n\geq 1} C_n^2 e^{2t\left[\alpha s - \delta - \delta\beta\tau\omega s - d\left(\frac{n\pi}{x_b - \xi}\right)^2\right]} \tag{37}$$

Corollary 1 provides an upper bound to control the L^2 norm of the disease prevalence in both regions A and B, and thus in the entire spatial economy. The following proposition refines the above result by providing sufficient conditions which guarantee that the epidemic management program is effective, that is the total disease prevalence decreases over time in both regions A and B.

Proposition 1 *Assume that* $\beta > \frac{\alpha s - \delta}{s\delta\omega\tau}$, *then it follows that* $\frac{d}{dt}\int_{\xi}^{x_b} i_{x,t}^B dx = \frac{d}{dt}\|i_{x,t}^B\|_{L^1([\xi, x_b])} < 0$ *and* $\frac{d}{dt}\int_{x_a}^{\xi} i_{x,t}^A dx = \frac{d}{dt}\|i_{x,t}^A\|_{L^1([x_a, \xi])} < 0$, *which imply that the total disease prevalence over the spatial domain (i.e.,* $i_t = \int_{x_a}^{x_b} i_{x,t} dx = \|i_{x,t}\|_{L^1([x_a, x_b])}$*), decreases over time.*

Proposition 1 states that if treatment in region B is effective (i.e., $\beta > \frac{\alpha s - \delta}{s\delta\tau\omega}$), that is the amount of resources kept within region B to finance local treatment is large enough to reverse the disease growth pattern, then the total disease prevalence will decrease over time not only in region B but also in region A, and thus over the whole spatial domain.[7] The result is intuitive: since all economic and epidemiological parameters are exactly the same in both regions, if the cross-subsization of treatment from region B to region A is not too large, then region B will have enough resources to contain the disease regionally and the same will be true in region A, where all resources are kept within the region. Regional cross-subsization of treatment, despite limited, works jointly with the lockdown intensity: if less resources are diverted from region B, a more stringent lockdown to achieve effective disease containment may be needed. Overall, provided that $\beta > \frac{\alpha s - \delta}{s\delta\tau\omega}$, the treatment and lockdown policy mix in

[7] Note that the sufficient conditions outlined in Proposition 1 also imply that the early epidemic stage approximation is suitable to describe the disease dynamics as prevalence does not show an explosive behavior but it always remain bounded from above.

the two regions allows for a monotonic reduction in the total disease prevalence in the whole spatial economy. Despite the epidemic management program is effective as it allows to decrease the total disease prevalence, consistent with what shown in La Torre et al. (2021a) in an a-spatial setting, complete spatial eradication is not possible over a finite time horizon, while it may be possible only asymptotically. Moreover, as discussed in Sect. 2, if the condition in Proposition 1 holds true, then our early epidemic approximation allows us to well characterize the spatio-temporal disease dynamics within the entire spatial economy.

Thus far we have simply assumed that the boundary point between the two regions is exogenously given, while this needs to be optimally determined by the social planner. In order to determine the optimal ξ, which we shall denote with ξ_{opt}, we plug the analytical expressions of $u^j_{x,t}$ and $i^j_{x,t}$ provided by Theorem 1 into the functional cost in (17) to assess how it depends on ξ. Specifically, we need to minimize the following functional cost $\mathcal{C}(\xi)$:

$$\mathcal{C}(\xi) = \int_0^T \int_{x_a}^{\xi} \frac{i^2_{x,t}[1+u^2_{x,t}s^2]}{2} e^{-\rho t} dx dt + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i^2_{x,t}}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i^2_{x,T}}{2} e^{-\rho T} dx \tag{38}$$

with respect to ξ over the interval (x_a, x_b). As we have already observed the function $\mathcal{C}(\xi)$ goes to $+\infty$ when $x \to x_a^+$ and it is continuous over any compact interval $[k, x_b]$ with $k > x_a$, thus thanks to Weierstrass's theorem this implies the existence of a global minimizer over the interval $[k, x_b]$ which may be either an interior or a corner solution. Since determining analytically the optimal boundary point is not possible we will need to proceed through numerical analysis which confirms that in different circumstances ξ_{opt} may be an interior or a corner point.

In order to visually illustrate the implications of our analysis we now present a calibration of our model based on the Italian COVID-19 experience during its first epidemic wave (February–June 2020). We set the parameter values as in La Torre et al.'s (2021a) baseline a-spatial model, distinguishing between an early epidemic stage at the national level (where the disease prevalence has remained relatively small) and an advanced epidemic stage at the regional level in Bergamo (where the prevalence has reached a sizeable portion of the local population). Specifically, we set $\alpha = 0.1328$ and $\delta = 0.0476$ (implying $\mathcal{R}_0 = 2.79$), along with $\rho = 0.04/365$, $\omega = 8.23$ and $\tau = 0.3$. The parameters related to the peculiarities of our spatial framework are instead set arbitrarily as follows: $x_b = 1$, $x_a = -1$, $d = 0.01$, $s = 0.98$, while $i_0(x)$ and β are varied to see how they impact the solution. Figures 1 and 2 depict the spaio-temporal evolution of the lockdown intensity (left panels) and share of infectives (right panels) under different configurations of the initial distribution of the disease prevalence, distinguishing between situations in which $\beta = 0.2$ (top panels) and $\beta = 0.8$ (bottom panels). In order to allow comparability, at time 0 the average disease prevalence over the whole spatial domain is assumed to be the same and in particular to be equal to 0.03, while disease prevalence is relatively higher in the left than in the right locations such that effectively the leftmost region may require more

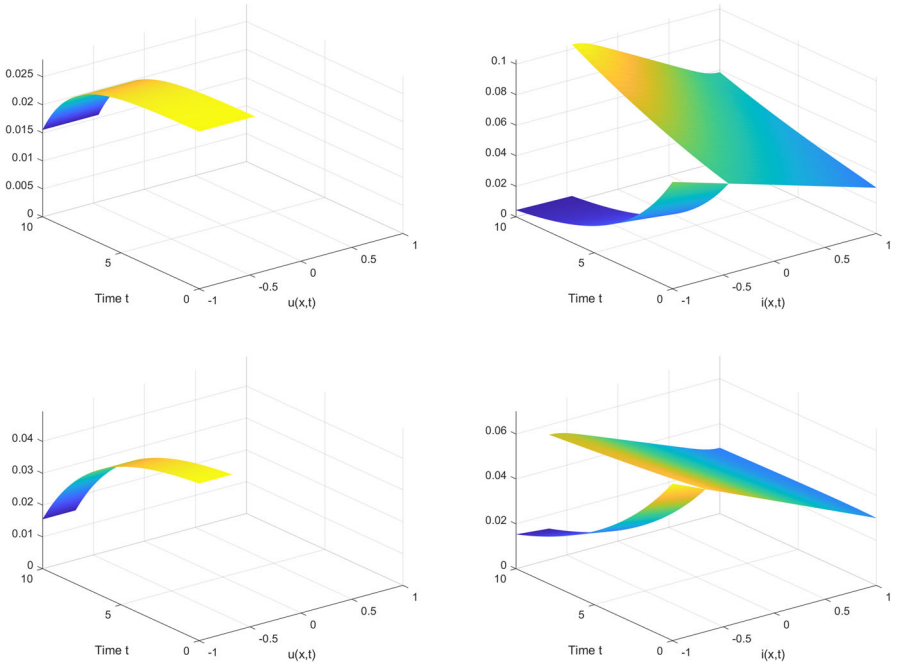

Fig. 1 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with $i_{x,0} = -\frac{0.04}{2}(x+1)+0.05$, and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = -0.5253$ with $\mathcal{C}_{opt} = 0.0233$ (top) and $\xi_{opt} = -0.7273$ with $\mathcal{C}_{opt} = 0.0124$ (bottom)

intervention than the rightmost one, justifying thus the adoption of a lockdown policy at least somewhere in the spatial economy. Note that under our parametrization the condition in Proposition 1 ($\beta > \frac{\alpha s-\delta}{s\delta\omega\tau} = 0.72$) holds true only in the $\beta = 0.8$ case and thus the epidemic management program gives rise to a reduction in the total disease prevalence in the spatial economy only in such a scenario, while in the $\beta = 0.2$ case the total disease prevalence may even increase over time.

Figure 1 represents the outcome in a scenario in which the initial distribution is monotonically decreasing over the spatial domain, and specifically it is given by $i_{x,0} = -\frac{0.05}{2}(x+1)+0.05$. We can observe that in region A the lockdown intensity decreases over time and this initially high value is large enough to reverse the growth pattern of the disease leading the disease prevalence to substantially decrease over time; in region B, despite the absence of lockdown, as enough (not enough) resources are allocated to local treatment, the disease prevalence decreases (increases) over time whenever $\beta = 0.8$ (whenever $\beta = 0.2$). It is interesting to note that the lockdown intensity is homogeneous across space in region A despite the fact that each location in the region is characterized initially by a different level of disease prevalence. This is due to the role of spatial diffusion which tends to increase prevalence in locations in which it is initially lower leading thus to a homogenization of prevalence across locations over time (La Torre et al. 2021b), thus from the social planner's perspective it is convenient to internalize such dynamic effects applying the same lockdown intensity

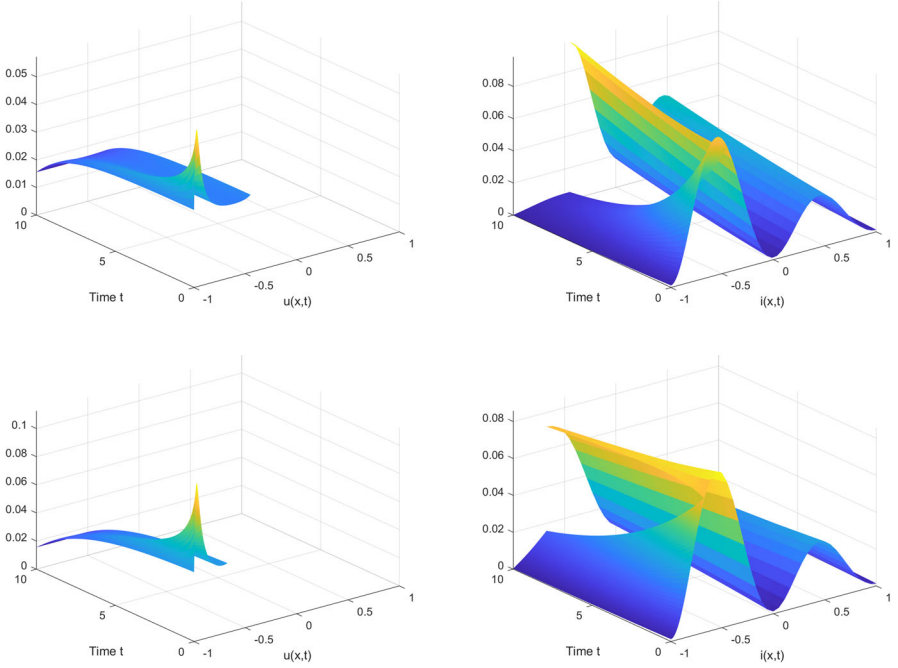

Fig. 2 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with $i_{x,0} = k(\sin(\pi x)^2 e^{-x-1} + 0.01)$ where $k = \frac{0.03(4\pi^2+1)}{(1-e^{-2})\pi^2+0.01(4\pi^2+1)}$, and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = -0.4243$ with $\mathcal{C}_{opt} = 0.0069$ (top) and $\xi_{opt} = -0.6667$ with $\mathcal{C}_{opt} = 0.0052$ (bottom)

everywhere in the region A (La Torre et al. 2021b). Even if from a qualitative point of view the behavior of the variables is independent of the value of β, this parameter critically impacts the optimal boundary point ξ and the locdkdown intensity. Indeed, β determines which share of the tax revenue collected in region B is spent locally in the region, thus a higher β implies that region B subsidizes treatment in region A by a smaller extent. We can observe that with a lower value of β (top panels) region A covers a larger area than what happens with a higher β (bottom panels): as intervention is less cross-subsidized by region B it is optimal to reduce the size of region A and increase the lockdown intensity in order to compensate for the reduced resources for treatment and to allow for a reduction in disease prevalence within the region. It is possible to show that this result is robust and the larger β the smaller ξ_{opt} (i.e., ξ_{opt} moves leftward) such that region A gets smaller.

Figure 2 repeats our previous analysis in the case in which the initial distribution of the disease prevalence is sinusoidal and given by the following expression: $i_{x,0} = k(\sin(\pi x)^2 e^{-x-1} + 0.01)$ with $k = \frac{0.03(4\pi^2+1)}{(1-e^{-2})\pi^2+0.01(4\pi^2+1)}$. From a qualitative point of view the results are similar to those discussed earlier: disease prevalence decreases in region A, but the high heterogeneity of the initial distribution within the region leads the lockdown intensity to be initially higher in the locations in which prevalence is initially higher. Despite the physical mechanism of diffusion tends to increase prevalence in

locations in which it is initially lower and to decrease it in those in which it is initially higher, it is optimal from the social planner's perspective to intervene more severely in the central locations of the region. It is interesting to observe that the ξ_{opt} does not coincide with $x = 0$, that is the point associated with the lowest initial prevalence level between the peaks occurring at $x = -0.5$ and $x = 0.5$, but it is shifted leftmost. Moreover, ad before different values of β do not change the qualitative behavior of the main variables but the parameter plays an important role in determining the optimal boundary point. In particular, the larger β the smaller ξ (i.e., ξ moves lefttward) such that region A gets smaller, which is exactly the same result we have discussed in Fig. 1.

Overall, these figures show that from a normative perspective it is not simple to understand how to effectively determine the optimal size of a lockdown area since the chosen regional size determines the availability of resources to finance pharmaceutical interventions in different regions and thus it critically affects the eventual success of national epidemic management programs. In particular, the optimal choice of the lockdown area (and thus of the lockdown intensity) depends on the initial spatial distribution of disease prevalence and on the amount of resources diverted from one region to the other, thus policymakers' task to determine the optimal lockdown area and intensity is all but simple as it requires a large amount of information on the characteristics of the disease across the whole spatial economy.

4.1 The homogeneous case

A particular case of our general model is represented by the situation in which the initial distribution of the disease prevalence is spatially homogeneous (i.e., $i_0(x) = i_0$). In this specific situation the leftmost locations are characterized by exactly the same level of disease prevalence as the rightmost locations, thus intervening more severily (with both lockdown and treatment measures) in the left region may not be a sensible approach. Nevertheless, this special case allows us to derive substantially simpler analytical expressions for $u^j_{x,t}$ and $i^j_{x,t}$ and to understand more clearly how the optimal boundary point is determined, thus it represents an interesting and instructive example. The following proposition summarizes our results under homogeneity of the initial distribution of the disease prevalence.

Corollary 2 *Suppose $i_0(x) = i_0$ for any $x \in [x_a, x_b]$. Then, assuming that ξ is fixed, the optimal pair $(i_{x,t}, u_{x,t})$ solving problem (17) is given by the following expressions:*

$$i_{x,t} = \begin{cases} i^A_{x,t} = i_0 \left(e^{\Theta t}_{11} + e^{\Theta t}_{12} \left[\frac{e^{\Theta T}_{21} - \phi_2 e^{\Theta T}_{11}}{\phi_2 e^{\Theta T}_{12} - e^{\Theta T}_{22}} \right] \right) & x \in [x_a, \xi] \\ i^B_{x,t} = i_0 e^{(\alpha s - \delta - \delta\beta\tau\omega s)t} & x \in [\xi, x_b] \end{cases} \tag{39}$$

$$u_{x,t} = \begin{cases} u^A_{x,t} = \dfrac{(\alpha-\delta\tau\omega)\left(e^{\Theta t}_{21}+e^{\Theta t}_{22}\left[\frac{e^{\Theta T}_{21}-\phi_2 e^{\Theta T}_{11}}{\phi_2 e^{\Theta T}_{12}-e^{\Theta T}_{22}}\right]\right)}{\left(e^{\Theta t}_{11}+e^{\Theta t}_{12}\left[\frac{e^{\Theta T}_{21}-\phi_2 e^{\Theta T}_{11}}{\phi_2 e^{\Theta T}_{12}-e^{\Theta T}_{22}}\right]\right)} & x \in [x_a, \xi] \\ u^B_{x,t} = 0 & x \in [\xi, x_b] \end{cases} \tag{40}$$

Corollary 2 determines the dynamics of the lockdown intensity and disease prevalence in the case in which the initial prevalence distribution is homogeneous across space, showing that $u^j_{x,t}$ and $i^j_{x,t}$ are spatially homogeneous as well. This result is due to the fact that all economic and epidemiological parameters are spatially homogeneous and thus in the absence of any source of heterogeneity the variables behavior needs to be exactly the same in every location within each region. In the case of a homogeneous initial distribution of the disease prevalence, the functional form of $C(\xi)$ simplifies and it can be written as:

$$C(\xi) = C_1(\xi) + C_2(\xi) + C_3(\xi) \tag{41}$$

where:

$$C_1(\xi) = \frac{(\xi - x_a)i_0^2}{2}\int_0^T \left(e^{\Theta t}_{11} + Le^{\Theta t}_{12}\right)^2 e^{-\rho t}dt + \frac{(\xi - x_a)i_0^2 s^2(\alpha - \delta\omega\tau)^2}{2}\int_0^T \left(e^{\Theta t}_{11} + Le^{\Theta t}_{12}\right)^2 e^{-\rho t}dt \tag{42}$$

$$C_2(\xi) = \frac{\mu(x_b - \xi)i_0^2}{2(\xi - x_a)}\left(\frac{e^{2(\alpha s-\delta-\delta\beta\tau\omega s)T}-1}{2(\alpha s-\delta-\delta\beta\tau\omega s)}\right) \tag{43}$$

$$C_3(\xi) = \frac{\Phi(x_b - x_a)i_0^2}{2}\left(\left(e^{\Theta T}_{11} + Le^{\Theta T}_{12}\right)^2 + 2e^{2(\alpha s-\delta-\delta\beta\tau\omega s)T}\right) \tag{44}$$

and $L = \frac{e^{\Theta T}_{21}-\phi_2 e^{\Theta T}_{11}}{\phi_2 e^{\Theta T}_{12}-e^{\Theta T}_{22}}$. The integrals in the above expression can be solved in closed-form but the expression of $C(\xi)$ turns out to be quite cumbersome in general, and thus it is not possible to explicitly determine the optimal ξ. However, in a special case in which the tax revenue collected from region B is employed entirely locally within the region (i.e., $\beta = 1$) it is possible to obtain in closed-form such an optimal boundary point, as stated in the following corollary.

Corollary 3 *Suppose that $\beta = 1$, and the following condition holds:*

$$\frac{2\mu(\alpha s-\delta-\delta\tau\omega s)}{\left(e^{2(\alpha s-\delta-\delta\tau\omega s)T}-1\right)}\left[\int_0^T (e^{\Theta t}_{11} + Le^{\Theta t}_{12})^2 e^{-\rho t}dt\right]^{-1} < x_b - x_a$$

Then $C'(\xi) = 0$ if and only if the following result holds:

$$\xi = x_a + \sqrt{\frac{2\mu(\alpha s - \delta - \delta\tau\omega s)(x_b - x_a)}{(1 + s^2(\alpha - \delta\omega\tau)^2)\left(e^{2(\alpha s - \delta - \delta\tau\omega s)T} - 1\right)\int_0^T (e_{11}^{\Theta t} + Le_{12}^{\Theta t})^2 e^{-\rho t} dt}} \tag{45}$$

where $\eta_1 = s(\alpha - \delta\,\tau\omega) - \delta$ *and* $\eta_2 = 4(\alpha - \delta\omega\tau)^2 + (\rho - 2\eta_1)^2$. *Furthermore* ξ *is a global minimizer of* C *when* $x > x_a$ *and* $\xi \in (x_a, x_b)$.

Corollary 3 demonstrates that, under a quite general parameter condition the function C has a global minimizer, which is an interior point of the domain (x_a, x_b), which suggests that also under a homogeneous initial disease prevalence distribution the optimal ξ may be characterized by an interior solution. It is not possible though to infer how this optimal boundary point depends on the main parameters, as its expression is extremely complicated due to the presence of the entries of the exponential matrix which depend nonlinearly on the model parameters. Apart from what happens in this particular case, in general the behavior of $C(\xi)$ shows a vertical asymptote at $\xi = x_a$ and then a decreasing trend followed by an increasing one, which suggests that also under a homogeneous initial disease prevalence distribution the optimal ξ may be characterized by an interior or a corner solution, exactly as it happens whenever the initial disease prevalence distribution is heterogeneous. Therefore, as before we proceed with numerical simulations to illustrate the behavior of the optimal solution.

Fig. 3 shows the spatio-temporal evolution of the lockdown intensity (left panel) and disease prevalence (right panel) in the homogeneous initial prevalence distribution case. In particular, we set $i_{x,0} = 0.03$ such that the total share of infectives within the whole spatial economy is the same as in Figs. 1 and 2. We can observe that the results are qualitative identical to those earlier discussed in the case of a heterogeneous initial distribution: the initially high value of the lockdown intensity in region A (which covers a small portion of the entire spatial domain) allows to reverse the disease growth pattern in the region, while prevalence increases or decreases in region B according to the size of β. The most noticeable difference with respect to what have seen earlier is related to the impact of β: independently of the value of this parameter the optimal boundary point is always located in the same position: as there are no sources of heterogeneity between regions (the parameters along with the initial prevalence are the same in the two regions) the redistribution of resources between regions affects only the intensity of the lockdown intensity in region A (and clearly the possibilities for treatment in both regions A and B) but not the optimal size of the lockdown area.

5 The advanced epidemic stage

We now analyze the problem in an advanced epidemic setting in which the share of susceptibles is not necessarily either constant or homogeneous, $s_{x,t} \neq s$. In this context, along the lines of what we have seen in Sect. 2, by exploiting the fact that $s_{x,t} = n_{x,t} - i_{x,t}$ the problem can be rewritten as follows: Find $\xi \in (x_a, x_b]$ which minimizes the following functional:

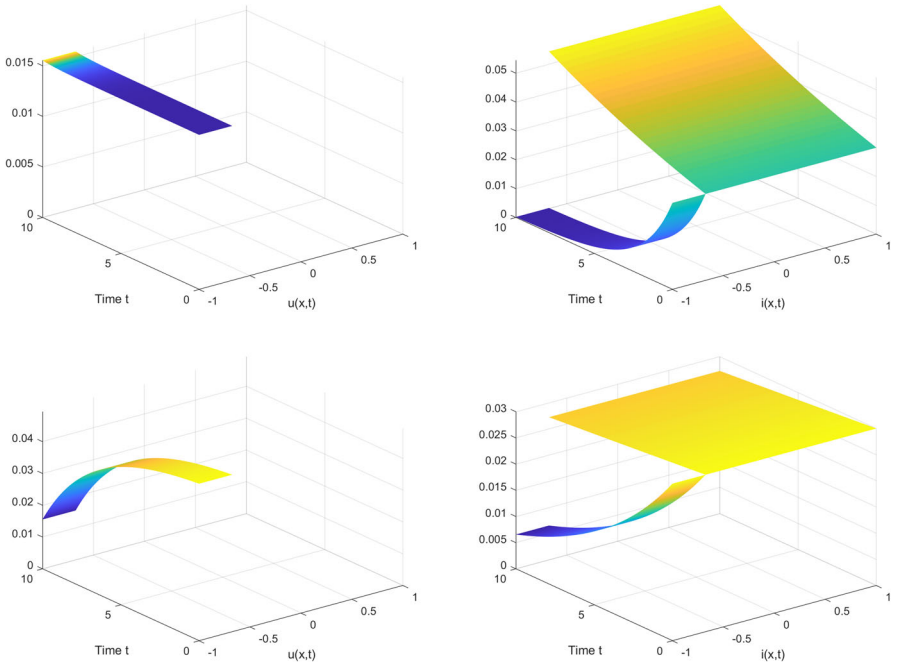

Fig. 3 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with $i_{x,0} = 0.03$ and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = -0.7273$ with $\mathcal{C}_{opt} = 0.0091$ (top) and $\xi_{opt} = -0.7273$ with $\mathcal{C}_{opt} = 0.0045$ (bottom)

$$
\begin{aligned}
\mathcal{C}(\xi) = \min_{u_{x,t}} \; & \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2[1 + u_{x,t}^2(n_{x,t} - i_{x,t})^2]}{2} e^{-\rho t} dx dt \\
& + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i_{x,t}^2}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\
s.t. \; & \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1 - u_{x,t})(n_{x,t} - i_{x,t}) i_{x,t} \\
& -\delta \left[1 + \omega\tau(1 - u_{x,t})(n_{x,t} - i_{x,t}) + \frac{(1 - \beta)\omega\tau}{\xi - x_a} \right. \\
& \left. \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] i_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(n_{x,t} - i_{x,t}) i_{x,t} - \delta[1 + \beta\omega\tau(n_{x,t} - i_{x,t})] i_{x,t}, \\
& \quad x \in [\xi, x_b] \\
& \frac{\partial n_{x,t}}{\partial t} = d \frac{\partial^2 n_{x,t}}{\partial x^2} \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\}
\end{aligned}
$$

$$\begin{aligned}
&\frac{\partial n_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
&i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
&n_{x,0} = n_0(x) > 0 \quad x \in [x_a, x_b]
\end{aligned} \tag{46}$$

In this case, even after a variable change the problem does not take a linear-quadratic form but nevertheless the objective function is convex and it is possible to prove existence and uniqueness of the solution by using the fact that the optimal control is unique along with the fact that the state and costate equations are bounded, meaning that the associated system of differential equations has a Lipschitz structure (Fleming and Rishel 1975). Also in this case because of the convexity of the objective functional the optimal control problem turns out to be well-posed (Dontchev and Zolezzi 1993). Exactly as in the previous section we proceed by steps, considering ξ as exogenously given first, and determining ξ_{opt} endogenously then. However, different from the previous section in which it has been possible to obtain the optimal solution in closed-form, now the highly nonlinear structure of the problem precludes us from deriving explicitly the dynamic path of the lockdown intensity and disease prevalence. Therefore, apart from the characterization of the optimality conditions (Theorem 2), we will need to mostly rely on a numerical analysis.

Theorem 2 *Assuming that $\xi \in (x_a, x_b)$ is fixed. Under the regularity assumptions on the state and the control variables presented in the model statement, then the optimal pair $(i_{x,t}, u_{x,t})$ solving problem (46) satisfies the following optimality conditions:*

- $i_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as* $i_{x,t} = i^A_{x,t}\chi_{[x_a,\xi]} + i^B_{x,t}\chi_{[\xi,x_b]}$
- $u_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as* $u_{x,t} = u^A_{x,t}\chi_{[x_a,\xi]}$
- *The function $i^B_{x,t}$ solves the following boundary value problem:*

$$\begin{cases}
\frac{\partial i^B_{x,t}}{\partial t} = d\frac{\partial^2 i^B_{x,t}}{\partial x^2} + (\alpha - \delta\beta\tau\omega) i^B_{x,t}\left(n^B_{x,t} - i^B_{x,t}\right) - \delta i^B_{x,t}, & x \in [\xi, x_b] \\
\frac{\partial i^B_{x,t}}{\partial x} = 0, & x \in \{\xi, x_b\} \\
i_{x,0} = i_0(x) & x \in [\xi, x_b]
\end{cases} \tag{47}$$

where:

$$n^B_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{x_b-\xi}\right)^2 dt} \cos\left[\frac{n\pi(x-\xi)}{x_b-\xi}\right] \tag{48}$$

$$B_0 = \frac{1}{x_b-\xi}\int_\xi^{x_b} n_{x,0}dx,$$

$$B_n = \frac{2}{x_b-\xi}\int_\xi^{x_b} n_{x,0}\cos\left[\frac{n\pi(x-\xi)}{x_b-\xi}\right]dx \tag{49}$$

- *The pair* $(i^A_{x,t}, u^A_{x,t})$ *solves the following optimality conditions:*

$$\begin{cases} \frac{\partial i^A_{x,t}}{\partial t} = d\frac{\partial^2 i^A_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega) i^A_{x,t}(n^A_{x,t} - i^A_{x,t})(1 - u^A_{x,t}) \\ \quad -\delta\left[1 + \frac{(1-\beta)\tau\omega}{\xi - x_a}\int_\xi^{x_b}\left(n^B_{x,t} - i^B_{x,t}\right)dx\right] i^A_{x,t} \\ \frac{\partial \lambda^A_{x,t}}{\partial t} = \rho\lambda_{x,t} - d\frac{\partial^2 \lambda^A_{x,t}}{\partial x^2} - i^A_{x,t} \\ \quad -\lambda^A_{x,t}\left[(\alpha - \delta\tau\omega)\left(n^A_{x,t} - 2i^A_{x,t}\right) - \delta - \frac{(1-\beta)\tau\omega}{\xi - x_a}\int_\xi^{x_b}(n^B_{x,t} - i^B_{x,t})dx\right] \\ u^A_{x,t} = \frac{\lambda^A_{x,t}(\alpha - \delta\tau\omega)}{(n^A_{x,t} - i^A_{x,t})i^A_{x,t}} \\ \frac{\partial i^A_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ \frac{\partial \lambda^A_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ i^A_{x,0} = i_0(x) \quad x \in [x_a, \xi] \\ \lambda^A_{x,T} = \phi i^A_{x,T} \quad x \in [x_a, \xi] \end{cases} \tag{50}$$

where $\lambda^A_{x,t}$ *is the costate variable and*

$$n^A_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{\xi - x_a}\right)^2 dt} \cos\left[\frac{n\pi(x - x_a)}{\xi - x_a}\right] \tag{51}$$

$$B_0 = \frac{1}{\xi - x_a}\int_{x_a}^{\xi} n_{x,0}dx,$$

$$B_n = \frac{2}{\xi - x_a}\int_{x_a}^{\xi} n_{x,0}\cos\left[\frac{n\pi(x - x_a)}{\xi - x_a}\right]dx \tag{52}$$

Theorem 2 determines the necessary conditions for an optimum of our optimization problem. Since we are assuming that ξ is fixed for the time being, provided that $0 < i_{x,t} < n_{x,t}$ and $0 < u_{x,t} < 1$, the conditions stated in Theorem 2 are also sufficient. In fact, despite the Hamiltonian function is non-convex, it is possible to show that the optimal control and the state and costate equations are bounded, meaning that the derived system of forward–backward differential equations has a Lipschitz structure (Jung et al. 2002; La Torre et al. 2020). These specific properties of the model ensure that the solution that we are able to characterize by analyzing the system of first order conditions is effectively the unique optimal solution of our minimization problem. However, the absence of a closed-form solution does not allow us to infer anything about the behavior of the lockdown intensity and disease prevalence over time and across space. Nevertheless, similar to what we have seen in an early epidemic setting, it is possible to derive some sufficient conditions guaranteeing that the epidemic management program is effective, allowing thus a monotonic reduction in the total disease prevalence in the entire spatial economy.

Proposition 2 *Assume that* $\frac{\alpha\tilde{s} - \delta}{\tilde{s}\delta\omega\tau} < \beta < \frac{\alpha}{\omega\delta\tau}$ *where* $\tilde{s} = \max_{x,t}(n_{x,t} - i_{x,t}) \geq 0$*; then it follows that* $\frac{d}{dt}\int_\xi^{x_b} i^B_{x,t}dx = \frac{d}{dt}\|i^B_{x,t}\|_{L^1([\xi, x_b])} < 0$ *and* $\frac{d}{dt}\int_{x_a}^{\xi} i^A_{x,t}dx =$

$\frac{d}{dt}\|i^A_{x,t}\|_{L^1([x_a,\xi])} < 0$, *which imply that the total disease prevalence over the spatial domain (i.e.,* $i_t = \int_{x_a}^{x_b} i_{x,t}dx = \|i_{x,t}\|_{L^1([x_a,x_b])}$*), decreases over time.*

Proposition 2 states that if the amount of resources kept within region B to finance local treatment is large enough to reverse the disease growth pattern (i.e., $\beta > \frac{\alpha\tilde{s}-\delta}{\tilde{s}\delta\omega\tau}$) but not excessively large to allow for enough cross-subsidization of treatment in region A (i.e., $\beta < \frac{\alpha}{\omega\delta\tau}$), then the total disease prevalence will decrease over time in both regions A and B, and thus also in the entire spatial domain. Note that, different from what we have seen in an early epidemic setting, the minimal amount of resources required to finance local treatment in region B depends on $\tilde{s}$, which represents the maximum of the susceptibles $s_{x,t} = n_{x,t} - i_{x.t}$ over time and space within the spatial economy. This means that ensuring that the epidemic management program is effective requires to forecast the possible spatio-temporal disease dynamics in order to determine which share of the resources needs to be maintained within region B and which share can be diverted to region A. Apart from this difference, comments similar to those discussed in an early epidemic setting apply.

As already discussed in the early epidemic scenario, in order to determine ξ_{opt} we proceed by minimizing the functional cost $\mathcal{C}(\xi)$ given by:

$$\mathcal{C}(\xi) = \int_0^T \int_{x_a}^{\xi} \frac{i^2_{x,t}[1 + u^2_{x,t}(n_{x,t} - i_{x,t})^2]}{2} e^{-\rho t} dx dt + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i^2_{x,t}}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i^2_{x,T}}{2} e^{-\rho T} dx \tag{53}$$

with respect to ξ over the interval (x_a, x_b). The absence of a closed form expression for the dynamic path of the control and state variables does not allow us to state anything about the behavior of the cost functional and thus we need to proceed via numerical analysis in order to determine the optimal spatio-temporal dynamic path of $u^j_{x,t}$ and $i^j_{x,t}$ and the optimal boundary point (see the online appendix for a discussion of the numerical method employed in our analysis). We keep relying on the same parameters we have employed in the previous section based on the COVID-19 experience during its first epidemic wave in Italy. In order to characterize an advanced epidemic setting in which the disease prevalence is no longer negligible as in an early epidemic stage, we now increase the average prevalence value over the whole spatial domain to 0.1. Figures 4 and 5 depict the spatio-temporal evolution of the lockdown intensity (left panels) and disease prevalence (right panels) under different configurations of the initial prevalence distribution, distinguishing between the case in which $\beta = 0.2$ (top panels) and $\beta = 0.8$ (bottom panels). Note that under our parametrization and numerical simulations the condition in Proposition 2 holds true ($\frac{\alpha\tilde{s}-\delta}{\tilde{s}\delta\omega\tau} < \beta < \frac{\alpha}{\omega\delta\tau}$) only in the $\beta = 0.8$ case and thus the epidemic management program gives rise to a reduction in the total disease prevalence in the spatial economy only in such a scenario, while in the $\beta = 0.2$ case the total disease prevalence may even increase over time.

Under a monotonically decreasing initial prevalence distribution (see Fig. 4) we can see that the results are qualitatively similar to those discussed in the previous section, apart from the fact that the lockdown intensity increases over time. We can

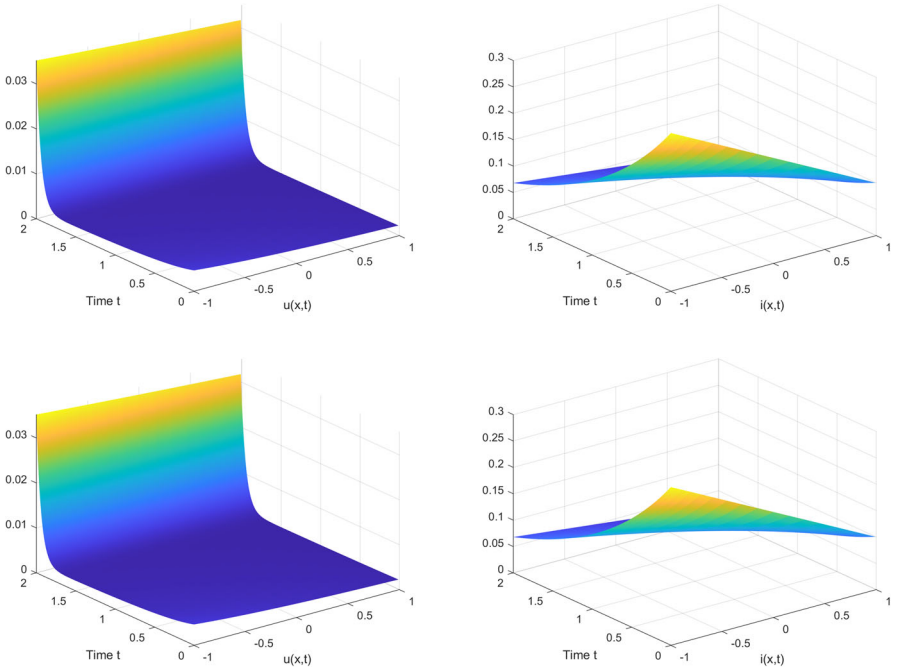

Fig. 4 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with $i_{x,0} = -\frac{0.2}{2}(x+1) + 0.21$, and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = 1$ with $C_{opt} = 0.0143$ (top) and $\xi_{opt} = 1$ with $C_{opt} = 0.0274$ (bottom)

note that in an advanced epidemic stage the size of region A is larger than in an early epidemic setting (ξ_{opt} is located rightmost), and actually the region A covers the entire spatial economy. As initially disease prevalence is higher in the whole spatial domain, the need for intervention is more stringent in every location and thus it is convenient to increase the size of region A (where more stringent lockdown measures can be applied) in order to effectively reverse the disease growth pattern. In this case β does not play any role as independently of the value of the parameter the optimal boundary point is located at its upper extreme.

Similar comments apply also under a sinusoidal initial prevalence distribution (see Fig. 5). Similar to what happens in an early epidemic stage, initially the lockdown intensity is higher in locations characterized by a higher level of disease prevalence, Moreover, exactly as what we have seen in the previous figure in the case of a monotonically decreasing prevalence distribution, we can note that with a sinusoidal distribution in an advanced epidemic stage the size of region A is larger than in an early epidemic setting (ξ_{opt} is located rightmost). This is again due to the role of a higher disease prevalence in the whole spatial domain, which thus requires to intervene more severely (with both lockdown and treatment measures) in a larger number of locations. As the lockdown region does not cover the entire spatial economy, the effects of β on the optimal boundary point are noticeable in this case. Specifically, different from what we have discussed under a monotonically decreasing initial

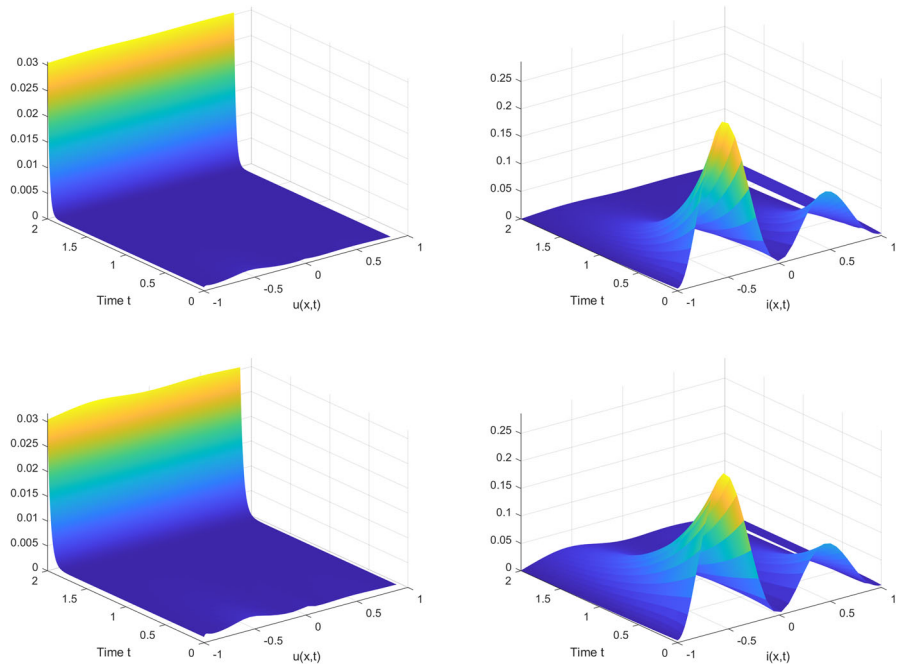

Fig. 5 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with $i_{x,0} = k(\sin(\pi x)^2 e^{-x-1} + 0.01)$ where $k = \frac{0.1(4\pi^2+1)}{(1-e^{-2})\pi^2+0.01(4\pi^2+1)}$, and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = 0.8284$ with $\mathcal{C}_{opt} = 0.0043$ (top) and $\xi_{opt} = 0.8890$ with $\mathcal{C}_{opt} = 0.0069$ (bottom)

prevalence distribution, the optimal boundary point shifts rightward as β increases. The availability of less resources from region B to subsidize treatment in region A requires to intervene more severely (through lockdown measures) in a larger number of locations.

Overall, these figures confirm what we have earlier discussed in an early epidemic setting, and in particular from a normative perspective it is not simple to understand how to effectively determine the optimal size of a lockdown area since this critically depends on the initial spatial distribution of disease prevalence and the amount of resources diverted from one region to the other. Moreover, different from what we have discussed in an early epidemic stage, ensuring the effectiveness of the epidemic management program requires to forecast the possible spatio-temporal disease dynamics in order to determine which share of the resources needs to be maintained within one region for local treatment and which share can be diverted to cross-subsidizing treatment in the other region.

5.1 The homogeneous case

Different from what we have seen in an early epidemic setting in which the case of a spatially homogeneous initial distribution of the disease prevalence (i.e., $i_0(x) = i_0$)

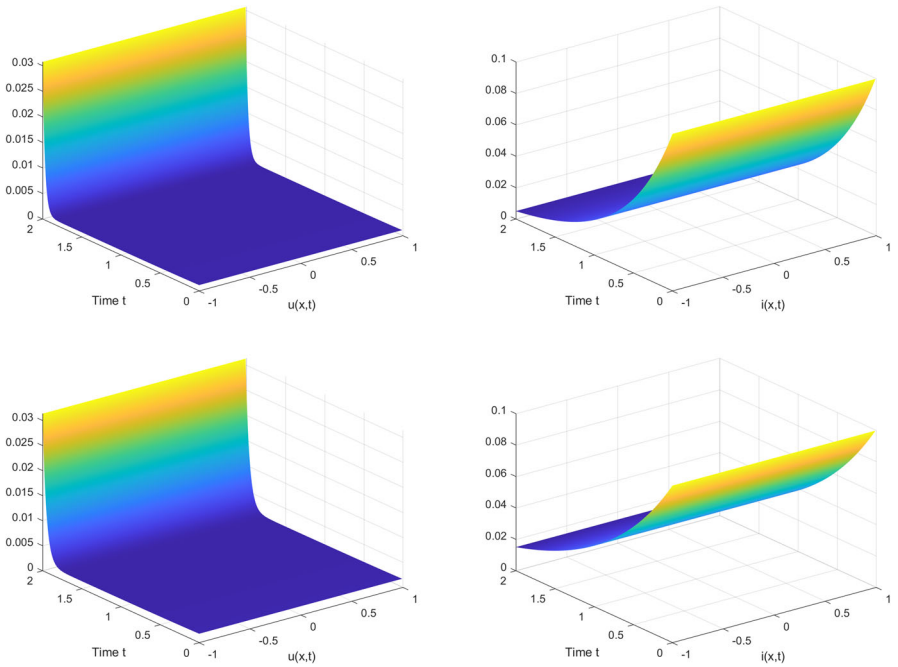

Fig. 6 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with $i_{x,0} = 0.03$ and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = 1$ with $\mathcal{C}_{opt} = 0.0032$ (top) and $\xi_{opt} = 1$ with $\mathcal{C}_{opt} = 0.0053$ (bottom)

allows for a more clear characterization of the optimal solution and of the determination of the optimal boundary point, in an advanced epidemic stage such advantages deriving from the homogeneity assumption disappear because the spatio-temporal dynamic path of the state and control variables cannot be derived in closed-form. Therefore, we necessarily need to rely on a numerical analysis which however confirms what we have already discussed in an early epidemic stage: in our previous parametrization based on Italian COVID-19 experience, the cost functional is resembles an U-shaped parabola in ξ and thus the optimal boundary point may coincide with either one of the extremes of the spatial domain or with an interior point.

Figure 6shows the spaio-temporal evolution of the lockdown intensity (left panels) and disease prevalence (right panels) in the homogeneous initial prevalence distribution case. We can observe that the results are qualitative similar to those earlier discussed in an early epidemic stage, apart from the fact that, exactly as in the previous monotonically decreasing initial prevalence scenario, the lockdown area covers the entire spatial domain independently of the value of β. Different from what we have seen in an early stage setting, in an advanced epidemic stage because of the lack of heterogeneity between regions it is convenient from the social planner's point of view to apply the same policy mix (based both on lockdown and treatment measures) everywhere in the spatial economy.

6 Multiple regions

Thus far we have presented our analysis in a two-regions framework as a matter of expositional simplicity, but our setup can be straightforwardly generalized to multiple regions. In order to exemplify how the addition of further regions may affect our results we focus only on a three-regions case in which two boundary points (rather than one as in a two-regions context) need to be optimally determined. Specifically, we consider two alternative scenarios differing on the nature of a third region (which we shall refer to as region C): in the first one the third region is not subject to lockdown and part of the tax revenues collected in the region are diverted to the lockdown region to finance extra treatment (i.e., the extra region resembles region B); in the second one the third region is subject to lockdown such that the intensity of the lockdown needs to be determined in two regions (i.e., the extra region resembles region A). These two alternative scenarios are not exhaustive of all the possible ways in which a continuous spatial domain may be partitioned in order to control the spread of an infectious disease in real world situations, but they well exemplify the generality of our setup and how it can be modified in order to deal with the different needs of policymakers. In our following discussion we will focus only on the advanced epidemic stage, but similar to what we have shown in the previous sections the early epidemic stage may be analyzed as well. For the benefit of the reader and simplify the exposition, in the following we will not stress the technical mathematical assumptions but rather we will focus on the economic and policy implications of our model extension. However, precise definitions of the set of admissible controls as well as the regularity hypotheses on the state and the control variables can be stated similar to those presented in the previous sections.

6.1 An extra non-lockdown region

The planner needs to choose how to split the entire spatial economy in three regions in which different combinations of the pharmaceutical and non-pharmaceutical interventions will be implemented. In one region (region $A = [\xi_1, \xi_2]$) both treatment and lockdown will be used, while in the others (regions $C = [x_a, \xi_1]$ and $B = [\xi_2, x_b]$) only treatment will be used (with $x_a < \xi_1 < \xi_2 < x_b$). In region A, which develops from ξ_1 to ξ_2, output net of lockdown measures is given by: $y_{x,t} = (1 - u_{x,t})q_{x,t}$, while the tax revenue, $\tau y_{x,t}$, is entirely used to finance treatment locally, and some extra resources from treatment are collected from regions B and C, $r^B_{x,t}$ and $r^C_{x,t}$, such that $v_{x,t} = \tau y_{x,t} + r^B_{x,t} + r^C_{x,t}$. The disease dynamics is described by a SIS equation as follows: $\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1 - u_{x,t})s_{x,t}i_{x,t} - \delta(1 + \omega v_{x,t})i_{x,t}$. In region B and C, which develop from ξ_2 to x_b and from x_a to ξ_1 respectively, output in each location is determined by the unconstrained supply: $y_{x,t} = q_{x,t}$, while a part of the tax revenue, $\beta^i \tau y_{x,t}$ where $0 \leq \beta^i \leq 1$ with $i = \{B, C\}$ is employed locally to finance treatment and the remaining part is allocated to finance extra treatment in the region A. The disease dynamics is described by a SIS equation as follows: $\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t}i_{x,t} - \delta(1+\omega v_{x,t})i_{x,t}$, where $v_{x,t} = \beta^i \tau q_{x,t}$. The total amount of

tax revenues diverted from regions B and C, $(1-\beta^B)\tau \int_B q_{x,t}dx+(1-\beta^C)\tau \int_C q_{x,t}dx$, is equally split within region A thus each location $x \in [\xi_1, \xi_2]$ receives a share $\frac{1}{\xi_2-\xi_1}$ of the total: $r^B_{x,t} + r^C_{x,t} = \frac{\tau}{\xi_2-\xi_1}[(1-\beta^B)\int_B q_{x,t}dx + (1-\beta^C)\int_C q_{x,t}dx]$. The social planner needs to determine the size of the three regions in order to minimize the social cost of the epidemic management program, accounting also for the instantaneous loss in region C. In region C, exactly as in region B, as there is no lockdown the instantaneous loss function depends only on the level of disease prevalence: $\ell(i_{x,t}) = \frac{i^2_{x,t}}{2}$. Since in region B and C the public health intervention is rather limited and some of its resources are diverted to region A, the instantaneous losses in these regions are weighted by their size-adjusted importance with respect to region A's, given by $\frac{\mu^B}{x_b-\xi_2}$ with $\mu^b > 0$ for region B and $\frac{\mu^C}{\xi_1-x_a}$ with $\mu^C > 0$ for region C. The planner's optimization problem reads thus as follows: Find $\xi_1, \xi_2 \in (x_a, x_b)$ with $\xi_2 > \xi_1$ which minimizes the optimal value:

$$
\begin{aligned}
\mathcal{C}(\xi_1,\xi_2) = \min_{u_{x,t}} \; & \frac{\mu}{\xi_1 - x_a}\int_0^T\int_{x_a}^{\xi_1}\frac{i^2_{x,t}}{2}e^{-\rho t}dxdt \\
& + \int_0^T\int_{\xi_1}^{\xi_2}\frac{i^2_{x,t}[1+u^2_{x,t}s^2_{x,t}]}{2}e^{-\rho t}dxdt \\
& + \frac{\mu}{x_b-\xi_2}\int_0^T\int_{\xi_2}^{x_b}\frac{i^2_{x,t}}{2}e^{-\rho t}dxdt + \phi\int_{x_a}^{x_b}\frac{i^2_{x,T}}{2}e^{-\rho T}dx \\
s.t. \; & \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t}i_{x,t} - \delta[1+\beta^C\omega\tau s_{x,t}]i_{x,t}, \quad x \in [x_a, \xi_1] \\
& \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1-u_{x,t})s_{x,t}i_{x,t} \\
& \quad -\delta\left\{1+\omega\tau(1-u_{x,t})s_{x,t} + \frac{\omega\tau}{\xi_2-\xi_1}\right. \\
& \quad \left.\left[(1-\beta^B)\int_{\xi_2}^{x_b}s_{x,t}dx + (1-\beta^C)\int_{x_a}^{\xi_1}s_{x,t}dx\right]\right\}i_{x,t}, \quad x \in [\xi_1, \xi_2] \\
& \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t}i_{x,t} - \delta[1+\beta^B\omega\tau s_{x,t}]i_{x,t}, \quad x \in [\xi_2, x_b] \\
& \frac{\partial n_{x,t}}{\partial t} = d\frac{\partial^2 n_{x,t}}{\partial x^2} \\
& s_{x,t} = n_{x,t} - i_{x,t} \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi_1, \xi_2\} \\
& \frac{\partial n_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi_1, \xi_2\} \\
& i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
& n_{x,0} = n_0(x) > 0 \quad x \in [x_a, x_b]
\end{aligned} \tag{54}
$$

By following the same approach we have adopted in our baseline two-regions framework, we can analytically determine the optimality conditions by taking the

boundary points ξ_1 and ξ_2 as given and then numerically determine the optimal boundary points which minimize the social cost associated with the epidemic management program. Similar to what earlier discussed, the following theorem states the optimality conditions for the model above.

Theorem 3 *Assuming that ξ_1, ξ_2 are fixed, $x_a < \xi_1 < \xi_2 < x_b$. The optimal pair $(i_{x,t}, u_{x,t})$ solving problem (54) satisfies the following optimality conditions:*

- $i_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as* $i_{x,t} = i^C_{x,t}\chi_{[x_a,\xi_1]} + i^A_{x,t}\chi_{[\xi_1,\xi_2]} + i^B_{x,t}\chi_{[\xi_2,x_b]}$
- $u_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as* $u_{x,t} = u^A_{x,t}\chi_{[\xi_1,\xi_2]}$
- *The function $i^B_{x,t}$ solves the following boundary value problem:*

$$\begin{cases} \frac{\partial i^B_{x,t}}{\partial t} = d\frac{\partial^2 i^B_{x,t}}{\partial x^2} + (\alpha - \delta\beta\tau\omega) i^B_{x,t}\left(n^B_{x,t} - i^B_{x,t}\right) - \delta i^B_{x,t}, & x \in [\xi_2, x_b] \\ \frac{\partial i^B_{x,t}}{\partial x} = 0, & x \in \{\xi_2, x_b\} \\ i_{x,0} = i_0(x) & x \in [\xi_2, x_b] \end{cases} \tag{55}$$

where:

$$n^B_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{x_b - \xi_2}\right)^2 dt} \cos\left[\frac{n\pi(x - \xi_2)}{x_b - \xi_2}\right] \tag{56}$$

$$B_0 = \frac{1}{x_b - \xi_2}\int_{\xi_2}^{x_b} n_{x,0} dx,$$

$$B_n = \frac{2}{x_b - \xi_2}\int_{\xi_2}^{x_b} n_{x,0}\cos\left[\frac{n\pi(x - \xi_2)}{x_b - \xi_2}\right] dx \tag{57}$$

- *The function $i^C_{x,t}$ solves the following boundary value problem:*

$$\begin{cases} \frac{\partial i^C_{x,t}}{\partial t} = d\frac{\partial^2 i^C_{x,t}}{\partial x^2} + (\alpha - \delta\beta\tau\omega) i^C_{x,t}\left(n^C_{x,t} - i^C_{x,t}\right) - \delta i^C_{x,t}, & x \in [x_a, \xi_1] \\ \frac{\partial i^C_{x,t}}{\partial x} = 0, & x \in \{x_a, \xi_1\} \\ i_{x,0} = i_0(x) & x \in [x_a, \xi_1] \end{cases} \tag{58}$$

where:

$$n^C_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{\xi_1 - x_a}\right)^2 dt} \cos\left[\frac{n\pi(x - x_a)}{\xi_1 - x_a}\right] \tag{59}$$

$$B_0 = \frac{1}{\xi_1 - x_a}\int_{x_a}^{\xi_1} n_{x,0} dx,$$

$$B_n = \frac{2}{\xi_1 - x_a}\int_{x_a}^{\xi_1} n_{x,0}\cos\left[\frac{n\pi(x - x_a)}{\xi_1 - x_a}\right] dx \tag{60}$$

- *The pair* $(i^A_{x,t}, u^A_{x,t})$ *solves the following optimality conditions:*

$$\begin{cases} \frac{\partial i^A_{x,t}}{\partial t} = d\frac{\partial^2 i^A_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega) i^A_{x,t}(n^A_{x,t} - i^A_{x,t})(1 - u^A_{x,t}) \\ \quad -\delta\left[1 + \frac{\omega\tau}{\xi_2 - \xi_1}\left((1-\beta^B)\int_{\xi_2}^{x_b} s^B_{x,t}dx + (1-\beta^C)\int_{x_a}^{\xi_1} s^C_{x,t}dx\right)\right] i^A_{x,t} \\ \frac{\partial \lambda^A_{x,t}}{\partial t} = \rho\lambda_{x,t} - d\frac{\partial^2 \lambda^A_{x,t}}{\partial x^2} - i^A_{x,t} - \lambda^A_{x,t}\left[(\alpha - \delta\tau\omega)\left(n^A_{x,t} - 2i^A_{x,t}\right)\right. \\ \quad \left. -\delta - \frac{\omega\tau}{\xi_2 - \xi_1}\left((1-\beta^B)\int_{\xi_2}^{x_b} s^B_{x,t}dx + (1-\beta^C)\int_{x_a}^{\xi_1} s^C_{x,t}dx\right)\right] \\ u^A_{x,t} = \frac{\lambda^A_{x,t}(\alpha - \delta\tau\omega)}{(n^A_{x,t} - i^A_{x,t}) i^A_{x,t}} \\ \frac{\partial i^A_{x,t}}{\partial x} = 0, \quad x \in \{\xi_1, \xi_2\} \\ \frac{\partial \lambda^A_{x,t}}{\partial x} = 0, \quad x \in \{\xi_1, \xi_2\} \\ i^A_{x,0} = i_0(x) \quad x \in [\xi_1, \xi_2] \\ \lambda^A_{x,T} = \phi i^A_{x,T} \quad x \in [\xi_1, \xi_2] \end{cases} \tag{61}$$

where $\lambda^A_{x,t}$ *is the costate variable and*

$$n^A_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{\xi_2 - \xi_1}\right)^2 dt} \cos\left[\frac{n\pi(x - \xi_1)}{\xi_2 - \xi_1}\right] \tag{62}$$

$$B_0 = \frac{1}{\xi_2 - \xi_1}\int_{\xi_1}^{\xi_2} n_{x,0}dx,$$

$$B_n = \frac{2}{\xi_2 - \xi_1}\int_{\xi_1}^{\xi_2} n_{x,0}\cos\left[\frac{n\pi(x - \xi_1)}{\xi_2 - \xi_1}\right]dx \tag{63}$$

In order to illustrate the implications of a third non-lockdown region on the partition of the whole spatial economy, we rely on the same parameter values employed in our Italian COVID-19 calibration by setting $\beta^B = \beta^C = 0.2$. Figure 7 shows the spatio-temporal evolution of the lockdown intensity (left panels) and infectives share (right panel) in a situation in which the initial disease prevalence distribution is linear (top panels) and sinusoidal (bottom panels). We can note that in both scenarios the results are qualitatively similar to those discussed in the previous section, that is disease prevalence decreases in the lockdown region while it tends to increase in the non-lockdown areas. It is interesting to observe that the possibility to partition the spatial economy in a further region decreases the size of the lockdown area such that a lower number of locations is subject to lockdown measures, which turn out to be more (less) stringent than in a two-regions setting at the beginning (at the end) of the planning horizon.

6.2 An extra lockdown region

We now consider a situation in which the third region (i.e., region C) is subject to lockdown measures. As before, the planner needs to choose how to split the entire spatial economy in three regions in which different combinations of the pharmaceutical

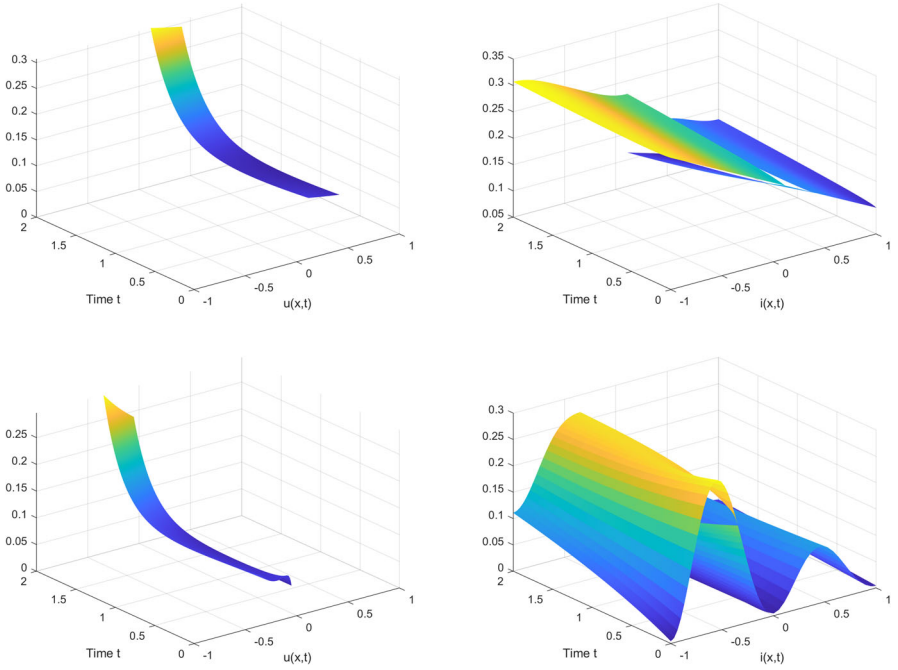

Fig. 7 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with either $i_{x,0} = -\frac{0.2}{2}(x+1) + 0.21$ (top) or $i_{x,0} = k(\sin(\pi x)^2 e^{-x-1} + 0.01)$ where $k = \frac{0.1(4\pi^2+1)}{(1-e^{-2})\pi^2+0.01(4\pi^2+1)}$ (bottom) in the case of two non-lockdown regions, Optimal boundary points: $\xi_1^{opt} = 0.1125$ and $\xi_2^{opt} = 0.4151$ (top), with $\mathcal{C}_{opt} = 0.1112$ and $\xi_1^{opt} = -0.3467$ and $\xi_2^{opt} = 0.0520$ (bottom), with $\mathcal{C}_{opt} = 0.0235$

and non-pharmaceutical interventions will be implemented. In two regions (region $A = [x_a, \xi_1]$ and $C = [\xi_2, x_b]$) both treatment and lockdown will be used, while in the other (region $B = [\xi_1, \xi_2]$) only treatment will be used (with $x_a < \xi_1 < \xi_2 < x_b$). In regions A and C, which develop from x_a to ξ_1 and from ξ_2 to x_b respectively, output net of lockdown measures is given by: $y_{x,t} = (1 - u_{x,t})q_{x,t}$, while the tax revenue, $\tau y_{x,t}$, is entirely used to finance treatment locally, and some extra resources from treatment are collected from region B, $r_{x,t}$, such that $v_{x,t} = \tau y_{x,t} + r_{x,t}$. The disease dynamics is described by a SIS equation as follows: $\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1 - u_{x,t})s_{x,t}i_{x,t} - \delta(1 + \omega v_{x,t})i_{x,t}$. In region B, which develop from ξ_1 to ξ_2, output in each location is determined by the unconstrained supply: $y_{x,t} = q_{x,t}$, while a part of the tax revenue, $\beta\tau y_{x,t}$ is employed locally to finance treatment and the remaining part is allocated to finance extra treatment in regions A and C. The disease dynamics is described by a SIS equation as follows: $\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t}i_{x,t} - \delta(1+\omega v_{x,t})i_{x,t}$, where $v_{x,t} = \beta\tau q_{x,t}$. The total amount of tax revenues diverted from region B, $(1-\beta)\tau \int_B q_{x,t}dx$, is split between regions A and C in proportions $0 < \theta < 1$ and $1-\theta$ respectively, and within in each region the amount received is equally split between locations, thus each location $x \in [x_a, \xi_1]$ receives an amount equal to $r^A_{x,t} = \frac{(1-\beta)\theta\tau}{\xi_1 - x_a}\int_B q_{x,t}dx$, while each location

$x \in [x_2, x_b]$ an amount equal to $r^C_{x,t} = \frac{(1-\beta)(1-\theta)\tau}{x_b - \xi_2} \int_B q_{x,t} dx$. The social planner needs to determine the size of the three regions in order to minimize the social cost of the epidemic management program, accounting also for the instantaneous loss in region C. In region C, exactly as in region A, as there is lockdown the instantaneous loss function depends on the spread of the disease and the output lost due to the lockdown measure as follows: $\ell(i_{x,t}, u_{x,t} q_{x,t}) = \frac{i^2_{x,t}(1+u^2_{x,t}q^2_{x,t})}{2}$. Since in region B the public health intervention is rather limited and some of its resources are diverted to regions A and C, the instantaneous losses in this regions are weighted by its size-adjusted importance with respect to regions A and C's, given by $\frac{\mu}{\xi_1 - x_a + x_b - \xi_2}$. The planner's optimization problem reads thus as follows: Find $\xi_1, \xi_2 \in (x_a, x_b)$ with $\xi_2 > \xi_1$ which minimizes the optimal value:

$$
\begin{aligned}
\mathcal{C}(\xi_1, \xi_2) = \min_{u_{x,t}} \; & \int_0^T \int_{x_a}^{\xi_1} \frac{i^2_{x,t}[1 + u^2_{x,t} s^2_{x,t}]}{2} e^{-\rho t} dx dt \\
& + \frac{\mu}{\xi_1 - x_a + x_b - \xi_2} \int_0^T \int_{\xi_1}^{\xi_2} \frac{i^2_{x,t}}{2} e^{-\rho t} dx dt \\
& + \int_0^T \int_{\xi_2}^{x_b} \frac{i^2_{x,t}[1 + u^2_{x,t} s^2_{x,t}]}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i^2_{x,T}}{2} e^{-\rho T} dx \\
s.t. \; & \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1 - u_{x,t}) s_{x,t} i_{x,t} \\
& - \delta \left[1 + \omega\tau(1 - u_{x,t}) s_{x,t} + \frac{(1-\beta)\omega\tau\theta}{\xi_1 - x_a} \int_{\xi_1}^{\xi_2} s_{x,t} dx \right] i_{x,t}, \\
& \quad x \in [x_a, \xi_1] \\
& \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t} i_{x,t} - \delta[1 + \beta\omega\tau s_{x,t}] i_{x,t}, \quad x \in [\xi_1, \xi_2] \\
& \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1 - u_{x,t}) s_{x,t} i_{x,t} \\
& - \delta \left[1 + \omega\tau(1 - u_{x,t}) s_{x,t} + \frac{(1-\beta)\omega\tau(1-\theta)}{x_b - \xi_2} \int_{\xi_1}^{\xi_2} s_{x,t} dx \right] i_{x,t}, \\
& \quad x \in [\xi_2, x_b] \\
& \frac{\partial n_{x,t}}{\partial t} = d \frac{\partial^2 n_{x,t}}{\partial x^2} \quad x \in [x_a, x_b] \\
& s_{x,t} = n_{x,t} - i_{x,t} \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi_1, \xi_2\} \\
& \frac{\partial n_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi_1, \xi_2\} \\
& i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
& n_{x,0} = n_0(x) > 0 \quad x \in [x_a, x_b]
\end{aligned}
\tag{64}
$$

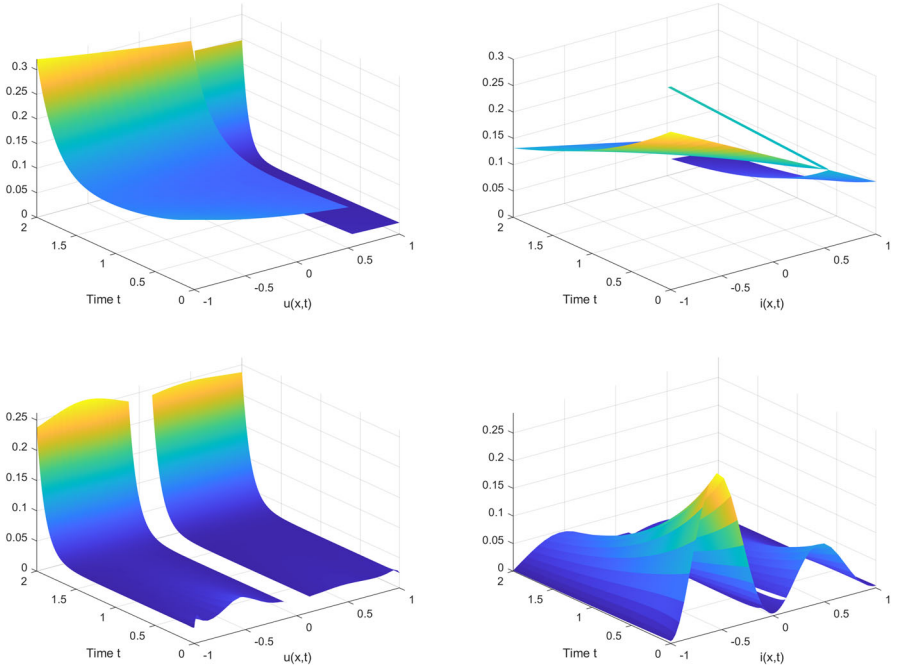

Fig. 8 Spatio-temporal evolution of $u_{x,t}$ (left) and $i_{x,t}$ (right) with either $i_{x,0} = -\frac{0.2}{2}(x+1) + 0.21$ (top) or $i_{x,0} = k(\sin(\pi x)^2 e^{-x-1} + 0.01)$ where $k = \frac{0.1(4\pi^2+1)}{(1-e^{-2})\pi^2+0.01(4\pi^2+1)}$ (bottom) in the case of two lockdown regions, Optimal boundary points: $\xi_1^{opt} = 0.5055$ and $\xi_2^{opt} = 0.5384$ (top), with $\mathcal{C}_{opt} = 0.0502$ and $\xi_1^{opt} = -0.1837$ and $\xi_2^{opt} = 0.0968$ (bottom), with $\mathcal{C}_{opt} = 0.0101$

Exactly as before, it is possible to analytically derive the optimality conditions taking the boundary points as given (Theorem 4) and numerically derive the optimal boundary points (Fig. 8),

Theorem 4 *Assuming that ξ_1, ξ_2 are fixed and $x_a < \xi_1 < \xi_2 < x_b$. Under the regularity assumptions on the state and the control variables, then the optimal pair $(i_{x,t}, u_{x,t})$ solving problem (46) satisfies the following:*

- $i_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as* $i_{x,t} = i^A_{x,t}\chi_{[x_a,\xi_1]} + i^B_{x,t}\chi_{[\xi_1,\xi_2]} + i^C_{x,t}\chi_{[\xi_2,x_b]}$
- $u_{x,t} : [x_a, x_b] \times [0, T] \to R_+$ *and it is defined as* $u_{x,t} = u^A_{x,t}\chi_{[x_a,\xi_1]} + u^C_{x,t}\chi_{[\xi_2,x_b]}$
- *The function $i^B_{x,t}$ solves the following boundary value problem:*

$$\begin{cases} \frac{\partial i^B_{x,t}}{\partial t} = d\frac{\partial^2 i^B_{x,t}}{\partial x^2} + (\alpha - \delta\beta\tau\omega) i^B_{x,t}\left(n^B_{x,t} - i^B_{x,t}\right) - \delta i^B_{x,t}, & x \in [\xi_1, \xi_2] \\ \frac{\partial i^B_{x,t}}{\partial x} = 0, & x \in \{\xi_1, \xi_2\} \\ i_{x,0} = i_0(x) & x \in [\xi_1, \xi_2] \end{cases} \tag{65}$$

where:

$$n^B_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{\xi_2-\xi_1}\right)^2 dt} \cos\left[\frac{n\pi(x-\xi_1)}{\xi_2-\xi_1}\right] \tag{66}$$

$$B_0 = \frac{1}{\xi_2-\xi_1}\int_{\xi_1}^{\xi_2} n_{x,0}dx,$$

$$B_n = \frac{2}{\xi_2-\xi_1}\int_{\xi_1}^{\xi_2} n_{x,0}\cos\left[\frac{n\pi(x-\xi_1)}{\xi_2-\xi_1}\right]dx \tag{67}$$

- *The pair* $(i^A_{x,t}, u^A_{x,t})$ *solves the following optimality conditions:*

$$\begin{cases} \frac{\partial i^A_{x,t}}{\partial t} = d\frac{\partial^2 i^A_{x,t}}{\partial x^2} + (\alpha-\delta\tau\omega)i^A_{x,t}(n^A_{x,t}-i^A_{x,t})(1-u^A_{x,t}) \\ \quad -\delta\left[1+\frac{(1-\beta)\tau\omega\theta}{\xi_1-x_a}\int_{\xi_1}^{\xi_2}(n^B_{x,t}-i^B_{x,t})dx\right]i^A_{x,t} \\ \frac{\partial \lambda^A_{x,t}}{\partial t} = \rho\lambda_{x,t} - d\frac{\partial^2\lambda^A_{x,t}}{\partial x^2} - i^A_{x,t} \\ \quad -\lambda^A_{x,t}\left[(\alpha-\delta\tau\omega)(n^A_{x,t}-2i^A_{x,t}) - \delta - \frac{(1-\beta)\tau\omega\theta}{\xi_1-x_a}\int_{\xi_1}^{\xi_2}(n^B_{x,t}-i^B_{x,t})dx\right] \\ u^A_{x,t} = \frac{\lambda^A_{x,t}(\alpha-\delta\tau\omega)}{i^A_{x,t}(n^A_{x,t}-i^A_{x,t})} \\ \frac{\partial i^A_{x,t}}{\partial x} = 0, \quad x\in\{x_a,\xi_1\} \\ \frac{\partial \lambda^A_{x,t}}{\partial x} = 0, \quad x\in\{x_a,\xi_1\} \\ i^A_{x,0} = i_0(x) \quad x\in[x_a,\xi_1] \\ \lambda^A_{x,T} = \phi i^A_{x,T} \quad x\in[x_a,\xi_1] \end{cases} \tag{68}$$

where $\lambda^A_{x,t}$ *is the costate variable and*

$$n^A_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{\xi_1-x_a}\right)^2 dt} \cos\left[\frac{n\pi(x-x_a)}{\xi_1-x_a}\right] \tag{69}$$

$$B_0 = \frac{1}{\xi_1-x_a}\int_{x_a}^{\xi_1} n_{x,0}dx,$$

$$B_n = \frac{2}{\xi_1-x_a}\int_{x_a}^{\xi_1} n_{x,0}\cos\left[\frac{n\pi(x-x_a)}{\xi_1-x_a}\right]dx \tag{70}$$

- *The pair* $(i^C_{x,t}, u^C_{x,t})$ *solves the following optimality conditions:*

$$\begin{cases} \frac{\partial i^C_{x,t}}{\partial t} = d\frac{\partial^2 i^C_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega) i^C_{x,t}(n^C_{x,t} - i^C_{x,t})(1 - u^C_{x,t}) \\ \quad -\delta\left[1 + \frac{(1-\beta)\tau\omega(1-\theta)}{x_b - \xi_2}\int_{\xi_1}^{\xi_2}(n^B_{x,t} - i^B_{x,t})dx\right] i^C_{x,t} \\ \frac{\partial \lambda^C_{x,t}}{\partial t} = \rho\lambda_{x,t} - d\frac{\partial^2 \lambda^C_{x,t}}{\partial x^2} - i^C_{x,t} \\ \quad -\lambda^C_{x,t}\left[(\alpha - \delta\tau\omega)(n^C_{x,t} - 2i^C_{x,t}) - \delta - \frac{(1-\beta)\tau\omega(1-\theta)}{x_b - \xi_2}\int_{\xi_1}^{\xi_2}(n^B_{x,t} - i^B_{x,t})dx\right] \\ u^C_{x,t} = \frac{\lambda^C_{x,t}(\alpha - \delta\tau\omega)}{(n^C_{x,t} - i^C_{x,t}) i^C_{x,t}} \\ \frac{\partial i^C_{x,t}}{\partial x} = 0, \quad x \in \{\xi_2, x_b\} \\ \frac{\partial \lambda^C_{x,t}}{\partial x} = 0, \quad x \in \{\xi_2, x_b\} \\ i^C_{x,0} = i_0(x) \quad x \in [\xi_2, x_b] \\ \lambda^C_{x,T} = \phi i^C_{x,T} \quad x \in [\xi_2, x_b] \end{cases} \tag{71}$$

where $\lambda^C_{x,t}$ *is the costate variable and*

$$n^C_{x,t} = \sum_{n\geq 0} B_n e^{-\left(\frac{n\pi}{x_b - \xi_2}\right)^2 dt} \cos\left[\frac{n\pi(x - \xi_2)}{x_b - \xi_2}\right] \tag{72}$$

$$B_0 = \frac{1}{x_b - \xi_2}\int_{\xi_2}^{x_b} n_{x,0}dx,$$

$$B_n = \frac{2}{x_b - \xi_2}\int_{\xi_2}^{x_b} n_{x,0}\cos\left[\frac{n\pi(x - \xi_2)}{x_b - \xi_2}\right]dx \tag{73}$$

Figure 8 shows the outcome of our analysis in the case of our Italian COVID-19 calibration by setting $\theta = 0.5$ and $\beta = 0.2$. Also in this case the results are qualitative similar to those discussed in a two-regions context, that is disease prevalence decreases in the two lockdown regions while it tends to increase in the non-lockdown area. The two lockdown areas are subject to different lockdown intensities, which intuitively results to be higher (on average) in the region in which the level of disease prevalence is higher. It is interesting to observe that also in this case, the possibility to partition the spatial economy in a further region (despite this additional region is a lockdown region) decreases the size of the lockdown area such that a lower number of locations is subject to lockdown measures, which are again more (less) stringent than in a two-regions setting at the beginning (at the end) of the planning horizon.

6.3 How many regions subject to lockdown?

Figures 7 and 8 show that by increasing the number of regions it is possible to decrease the size of the area subject to the more stringent policy measures (i.e., lockdowns). This suggests that by allowing the spatial economy to be partitioned in a larger number of regions allows to intervene with more precision in the areas which are more in need of policy support. However, we cannot assess whether it may be most convenient to

Table 1 Social cost of disease containment strategies in different scenarios

	One lockdown area	Two lockdown areas
Linear initial conditions	0.1112	0.0502
Nonlinear initial conditions	0.0235	0.0101

partition the spatial economy in one or two lockdown areas, thus in order to comment on this we need to quantify the social cost associated with the epidemic management program.

Table 1 reports the social cost of the epidemic management program in our three-regions setup in different scenarios, associated with a different number of lockdown areas (one or two) and different initial prevalence conditions (linearly decreasing or nonlinear sinusoidal). We can straightforwardly note that independently of the initial spatial distribution of disease prevalence the social cost is significantly lower when the spatial economy is partitioned in two lockdown areas: the possibility to split the economy in two distinct areas where the most severe policy intervention tools are applied allows to cater the lockdown intensity for the specific needs of such areas, resulting in a most effective way to reverse to disease growth pattern in specific areas and in the entire spatial economy as well.

7 Extensions

We now return to our two-regions baseline setup to extend it to account for some of the epidemiological peculiarities of COVID-19 and the effects of some policy measures implemented in order to control its spread, which we have not considered in our baseline model for the sake of analytical tractability. We first consider how results change if mobility patterns captured by the diffusion term are different between regions and between population sub-groups, then whether the possibility to adjust fiscal policy changes our conclusions, and finally how a more rigorous characterization of the COVID-19 epidemics modifies our analysis. As in the previous section, we focus only on an advanced epidemic setting, and for the sake of expositional simplicity we present only the results of our numerical simulations while the optimality conditions and further technicalities can be found in the online appendix. As in the previous section, we will not stress the technical mathematical assumptions but we will focus on the economic and policy implications of the different model extensions. Once again, precise definitions of the set of admissible controls as well as the regularity hypotheses on the state and the control variables can be easily stated similar to those presented for our benchmark model.

7.1 The economic setting

An important consequence of the policy measures implemented to contain the spread of COVID-19 worldwide consists of affecting heterogeneously individual mobility

across different population groups and regions. Indeed, since infectives have been subject to either hospitalized or self-imposed confinement rules their spatial mobility has been much lower than susceptibles' who instead have been able to freely circulate within the national or regional borders, thus the speed of diffusion cannot be simply assumed to be the same between the two population sub-groups. Moreover, due to lockdown regulations the mobility patterns of both susceptibles and infectives have largely reduced in lockdown regions compared to those occurring in non-lockdown ares, thus the speed of diffusion cannot be assumed to be the same between regions. Therefore, we modify our baseline model along two directions: (i) we allow for the population groups to be characterized by different diffusion parameters, d_I and d_S for infectives and susceptibles respectively, with $d_I \leq d_S$; (ii) and we allow for the two regions to be characterized by different diffusion coefficients, d_A and d_B in regions A and B respectively, with $d_A \leq d_B$.

By abstracting from mitigation policy and denoting with $d_i^j \geq 0$ the diffusion parameter for group $i = \{I, S\}$ in region $j = \{A, B\}$, our extended model in each region j can be described by the following system of partial differential equations:

$$\frac{\partial s_{x,t}}{\partial t} = d_S^j \frac{\partial^2 s_{x,t}}{\partial x^2} + \delta i_{x,t} - \alpha s_{x,t} i_{x,t} \tag{74}$$

$$\frac{\partial i_{x,t}}{\partial t} = d_I^j \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t} i_{x,t} - \delta i_{x,t} \tag{75}$$

It is straightforward to observe that the above equations are no longer symmetric and thus the equation for the share of population in location x, $n_{x,t} = s_{x,t} + i_{x,t}$, turns out to take the following form:

$$\frac{\partial n_{x,t}}{\partial t} = d_S^j \frac{\partial^2 n_{x,t}}{\partial x^2} + \left(d_I^j - d_S^j\right) \frac{\partial^2 i_{x,t}}{\partial x^2}, \tag{76}$$

which clearly does no longer coincide with the classical heat equation but rather depends endogenously on the share of infectives. Therefore, it is not possible to substitute $s_{x,t} = n_{x,t} - i_{x,t}$ in (75) in order to obtain a unique endogenous state equation paired with an exogenously given variable as we have done in the previous sections, but we need to carry on in our analysis both the equations for the susceptibles and infectives shares, given by (74) and (75) respectively. This makes the model slightly more complicated than the previous formulation due to the presence of two endogenous state variables for each region. Note that whenever $d^j = d_I^j = d_S^j$ (76) returns being the classical heat equation and thus we can follow the same approach we have employed in the previous sections.

Taking this complication into account, in our extended model with heterogeneous diffusion between groups and regions, the social planner's problem can be stated as follows: Find $\xi \in (x_a, x_b)$ which minimizes the following functional:

$$\mathcal{C}(\xi) = \min_{u_{x,t}} \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2 [1 + u_{x,t}^2 s_{x,t}^2]}{2} e^{-\rho t} dx dt + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i_{x,t}^2}{2} e^{-\rho t} dx dt$$
$$+ \phi \int_{x_a}^{x_b} \frac{i_{x,T}^2}{2} e^{-\rho T} dx$$

$$
\begin{aligned}
s.t. \quad & \frac{\partial s_{x,t}}{\partial t} = d_S^A \frac{\partial^2 s_{x,t}}{\partial x^2} - \alpha(1-u_{x,t}) s_{x,t} i_{x,t} \\
& + \delta \left[1 + \omega\tau(1-u_{x,t}) s_{x,t} + \frac{(1-\beta)\omega\tau}{\xi - x_a} \int_{\xi}^{x_b} s_{x,t} dx \right] i_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial i_{x,t}}{\partial t} = d_I^A \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1-u_{x,t}) s_{x,t} i_{x,t} \\
& - \delta \left[1 + \omega\tau(1-u_{x,t}) s_{x,t} + \frac{(1-\beta)\omega\tau}{\xi - x_a} \int_{\xi}^{x_b} s_{x,t} dx \right] i_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial s_{x,t}}{\partial t} = d_S^B \frac{\partial^2 s_{x,t}}{\partial x^2} - \alpha s_{x,t} i_{x,t} + \delta[1 + \beta\omega\tau s_{x,t}] i_{x,t}, \quad x \in [\xi, x_b] \\
& \frac{\partial i_{x,t}}{\partial t} = d_I^B \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha s_{x,t} i_{x,t} - \delta[1 + \beta\omega\tau s_{x,t}] i_{x,t}, \quad x \in [\xi, x_b] \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& \frac{\partial s_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
& s_{x,0} = s_0(x) \geq 0 \quad x \in [x_a, x_b]
\end{aligned} \tag{77}
$$

In an advanced epidemic setting we can determine the optimal size of the lockdown area and the optimal lockdown intensity only numerically, exactly as in the previous sections. We continue relying on the same parameter values we have employed in our COVID-19 Italian calibration, and in order to allow for heterogeneity in the speed of diffusion between groups we set $d_I^j = 0.01$ and $d_S^j = 0.05$ while to allow for heterogeneity between regions we set $d_i^A = 0.01$ and $d_i^B = 0.1$, showing how such a lack of diffusion homogeneity affects our results. Figure 9 shows the spatio-temporal evolution of the lockdown intensity (left panels), susceptibles share (central panels) and disease prevalence (right panels) under homogeneity between groups and regions (top panels), homogeneity between groups but heterogeneity between regions (mid panels), and heterogeneity between groups but homogeneity between regions (bottom panels). We can observe that in all three scenarios the results are qualitatively identical: the lockdown intensity monotonically decreases over time in order to allow disease prevalence to decrease within the lockdown region, while the absence of lockdown in the non-lockdown region does not allow for a reduction in prevalence within the region.[8] Heterogeneity in the mobility patterns affects quantitatively the results, and in particular the lockdown intensity and the size of the lockdown area are reduced by the presence of heterogeneity either between regions or between groups. This is due to the fact that a faster diffusion of a population group or within a region increases the speed at which the disease naturally tends to die out in a single location by spreading geographically across locations, reducing thus the need to employ stringent and widely spread lockdown measures.

Note also that in order to visualize the effects of heterogeneity we have analyzed what happens when the degree of heterogeneity (i.e., the gap between the diffusion

[8] Note that we cannot compare our results with those presented earlier in Fig. 4 since the evolution of the share of population residing in a given location (and thus that of the share of susceptibles) is different in the two frameworks.

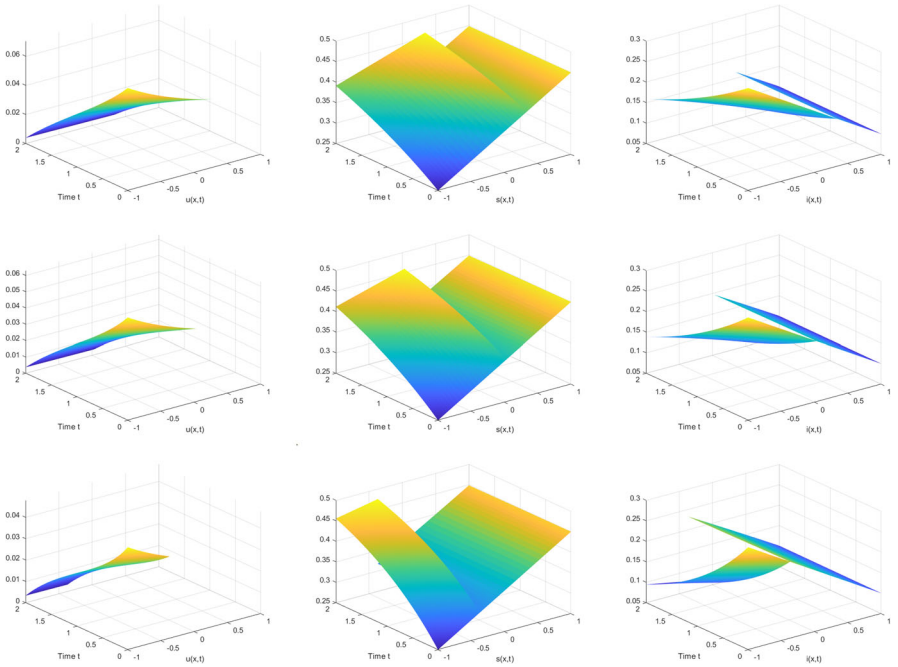

Fig. 9 Spatio-temporal evolution of $u_{x,t}$ (left), $s_{x,t}$ (center) and $i_{x,t}$ (right) with $i_{x,0} = -\frac{0.2}{2}(x+1)+0.21$ and $s_{x,0} = \frac{0.2}{2}(x+1)+0.29$ with $d_I^A = d_I^B = d_S^A = d_S^B = 0.01$ (top) or $d_I^A = d_I^B = 0.01$, $d_S^A = d_S^B = 0.05$ (mid) and $d_I^A = d_S^A = 0.01$, $d_I^B = d_I^B = 0.1$ (bottom). Optimal boundary point: $\xi_{opt} = 0.3409$ (top), with $\mathcal{C} = 0.0530$, $\xi_{opt} = 0.0298$ (center), with $\mathcal{C} = 0.0561$ and $\xi_{opt} = -0.369$ (bottom), with $\mathcal{C} = 0.0846$

parameters) is substantially large. In fact, it is possible to show that if the gap between the diffusion parameters is small differences in the degree of heterogeneity will not lead to noticeable differences in the evolution of the main variables (unless one of the two diffusion parameters becomes particularly large, at least one order of magnitude larger than the other epidemiological parameters). This suggests that under realistic parameter values we can safely analyze the implications of our economic-epidemiological framework under the assumption of homogeneous mobility patterns between groups and between regions.

7.2 The role of fiscal policy

Given our previous conclusion regarding the limited role of heterogeneity in mobility patterns under a realistic model parametrization, we now return to our baseline setup with homogeneous mobility between groups and between regions to extend it to account for the possibility to optimally determine fiscal policy, which provides policymakers with an additional mitigation instrument. Indeed, in our analysis thus far we have assumed that fiscal policy is exogenously given and the tax rate takes a constant

value such that lockdowns are the only form of policy intervention. In reality policymakers can try to mitigate the health-economic consequences of infectious diseases through fiscal policy adjustments, as confirmed by the experience of several countries during the ongoing COVID-19 epidemic in which different tax cuts, subsidies and other fiscal measures have been implemented to support households and firms' income. We now try to reassess our previous conclusions in the light of these considerations, and specifically we do so by endogeneizing the tax rate which now represents an additional control variable. By optimally determining the tax rate, policymakers can apply heterogenous tax rates between the lockdown and the non-lockdown regions choosing whether financing the extra treatment needs of the lockdown region through additional taxation either in the same region or in the non-lockdown region.

In a setting in which fiscal policy is endogenous, the policymaker needs to choose the optimal tax rate $\tau_{x,t}$ accounting for its social cost since a higher tax rate increases the instantaneous losses (in both regions A and B) associated with the epidemic management program. In particular, in region A, the instantaneous loss function depends no longer only on the spread of the disease and the output lost due to the lockdown measure but also on the tax rate, and by maintaining our quadratic formulation assumption it reads as follows: $\ell(i_{x,t}, u_{x,t}q_{x,t}, \tau_{x,t}) = \frac{i_{x,t}^2(1+u_{x,t}^2q_{x,t}^2)+\tau_{x,t}^2}{2}$, where the last term penalizes deviations from the no-tax scenario. Similarly, in the non-lockdown region B the instantaneous loss function depends not only on the level of disease prevalence but also on the tax rate as follows: $\ell(i_{x,t}, \tau_{x,t}) = \frac{i_{x,t}^2+\tau_{x,t}^2}{2}$. Note that even if in region B there is no lockdown, the social planner needs to determine the intensity of the tax rate in the region, thus also in region B they face an optimal control problem. And by optimally choosing the tax rate in each location within the two regions the social planner determines not only how many resources to employ locally in a specific location to finance treatment, but also how many resources to divert from region B to finance extra treatment in region A. Specifically, the total amount of tax revenues diverted from region B, $(1-\beta)\int_B \tau_{x,t}q_{x,t}dx$, is equally split within region A thus each location $x \in [x_a, \xi]$ receives a share $\frac{1}{\xi - x_a}$ of the total: $r_t = \frac{(1-\beta)}{\xi - x_a}\int_B \tau_{x,t}q_{x,t}dx$. In this more complicated setting, the social planner's optimization problem becomes: Find $\xi \in (x_a, x_b)$ which minimizes the following functional:

$$
\begin{aligned}
\mathcal{C}(\xi) = \min_{u_{x,t},\tau_{x,t}} \; & \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2[1+u_{x,t}^2(n_{x,t}-i_{x,t})^2]+\tau_{x,t}^2}{2} e^{-\rho t} dx dt \\
& + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i_{x,t}^2+\tau_{x,t}^2}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\
s.t. \; & \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1-u_{x,t})(n_{x,t}-i_{x,t})i_{x,t} \\
& -\delta\left[1+\omega\tau_{x,t}(1-u_{x,t})(n_{x,t}-i_{x,t}) + \frac{(1-\beta)\omega}{\xi - x_a}\right. \\
& \left. \int_{\xi}^{x_b} \tau_{x,t}(n_{x,t}-i_{x,t})dx\right] i_{x,t}, \quad x \in [x_a, \xi]
\end{aligned}
$$

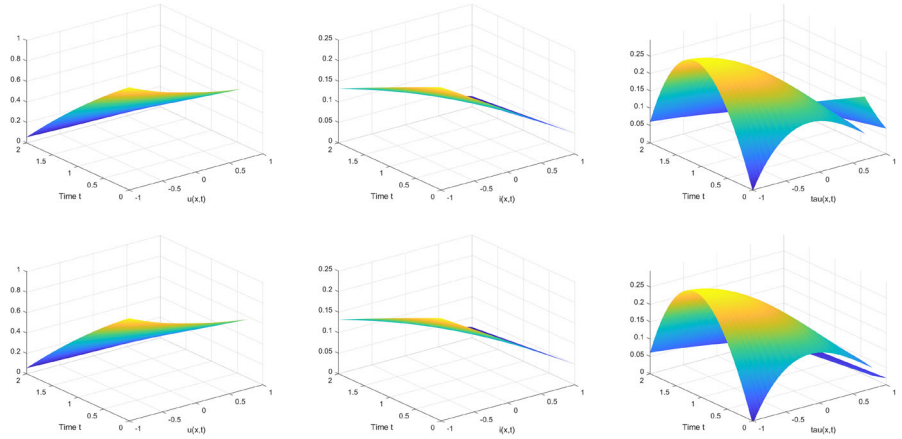

Fig. 10 Spatio-temporal evolution of $u_{x,t}$ (left), $i_{x,t}$ (center) and $\tau_{x,t}$ (right) with $i_{x,0} = -\frac{0.2}{2}(x+1)+0.21$ and either $\beta = 0.2$ (top) or $\beta = 0.8$ (bottom). Optimal boundary point: $\xi_{opt} = 0.6750$ with $\mathcal{C}_{opt} = 0.0909$ (top) and $\xi_{opt} = 0.7897$ with $\mathcal{C}_{opt} = 0.1106$ (bottom)

$$
\begin{aligned}
&\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(n_{x,t} - i_{x,t})i_{x,t} - \delta[1 + \beta\omega\tau_{x,t}(n_{x,t} - i_{x,t})]i_{x,t}, \\
&\quad x \in [\xi, x_b] \\
&\frac{\partial n_{x,t}}{\partial t} = d\frac{\partial^2 n_{x,t}}{\partial x^2} \\
&\frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
&\frac{\partial n_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
&i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
&n_{x,0} = n_0(x) > 0 \quad x \in [x_a, x_b]
\end{aligned} \tag{78}
$$

In an advanced epidemic setting we can numerically determine not only the optimal size of the lockdown area and the optimal lockdown intensity as in the previous sections, but also the optimal intensity of the tax rate. We continue relying on the same parameter values we have employed in our COVID-19 Italian calibration, and our results are illustrated in Fig. 10, which shows the spatio-temporal evolution of the lockdown intensity (left panels), disease prevalence (central panels) and tax rate (right panels) whenever either $\beta = 0.2$ (top panels) or $\beta = 0.8$ (bottom panels).

We can observe that the results are qualitatively similar to those we have discussed in our baseline model: the lockdown intensity monotonically decreases over time in order to allow disease prevalence to decrease within the lockdown region, while the absence of lockdown in the non-lockdown region does not allow for a reduction in prevalence within the region. The tax rate instead shows a different pattern in the lonckdown and non-lockdown regions: in the non-lockdown region it is monotonically decreasing mimicking the behavior of the lockdown intensity, while in the lockdown area it is non-monotonic (increasing first and decreasing then) and highly heterogeneous (being

initially higher in the central locations of the region than in the lateral ones). Compared to what we have discussed in our baseline setup where we have assumed the tax rate being homogeneous across space and constant over time (i.e., $\tau_{x,t} = 0.3, \forall x, t$), its optimal value results to be lower everywhere in space and time generating a slower reduction in disease prevalence along with a reduction in the size of the lockdown area (see Fig. 4). A higher β generates a rightward shift of the optimal boundary point, that is when the amount of resources diverted from region B to region A decreases, it is convenient to increase the tax rate within region A to allow for effective treatment and for a larger size of the lockdown area. Comments similar to those presented in our baseline setup apply, thus apart from some quantitative difference in the optimal values of the boundary point and the lockdown intensity, qualitatively speaking the endogenous determination of the optimal local tax rate does not modify our main conclusions and thus can be safely ignored in our following analysis.

7.3 The epidemiological setting

Given our previous conclusion regarding the limited role of the optimal determination of the tax rate on our qualitative results, we now return to our baseline setup with a constant and homogeneous tax rate to extend it to account for some of the epidemiological peculiarities of COVID-19 which we have not considered thus far for the sake of analytical tractability. Indeed, since recovery from COVID-19 infection provides temporary immunity from the disease its dynamics cannot be simplistically described by a SIS model. Moreover, the spatial spread of the disease seems to be much faster than the spatial pace of the demographic changes associated with migration, thus diffusion cannot be the only element of spatial propagation of the disease. Therefore, we modify our baseline model along two directions: (i) we consider a SIRS framework in which upon recovery individuals become immune before returning susceptible to the disease again when the acquired temporary immunity dies out; (ii) we introduce local effects conveyed by a spatial integral term which captures short-lived individuals' spatial movements due to personal and business trips.

By abstracting from mitigation policy, our extended SIRS epidemiological model can be described as follows. Every individual in the population in each location can be suceptible, infective, or recovered, $R_{x,t}$. Susceptibles become infectives by interacting with other infectives within the entire spatial economy (and not only those located in the same venue, as in our baseline framework); infectives become immune after recovering from the disease; recovereds return being susceptible again after the temporary immunity gained from recovery dies out, and $\epsilon > 0$ measures the speed of immunity loss. In each location the evolution of the susceptibles, infectives and recovereds shares where $r_{x,t} = \frac{R_{x,t}}{N}$ can be described through the following system of partial differential equations:

$$\frac{\partial s_{x,t}}{\partial t} = d\frac{\partial^2 s_{x,t}}{\partial x^2} + \epsilon r_{x,t} - \alpha \int_{\Omega} s_{x',t} i_{x',t} \psi_{x',x} dx \tag{79}$$

$$\frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha \int_{\Omega} s_{x',t} i_{x',t} \psi_{x',x} dx - \delta i_{x,t} \tag{80}$$

$$\frac{\partial r_{x,t}}{\partial t} = d\frac{\partial^2 r_{x,t}}{\partial x^2} + \delta i_{x,t} - \epsilon r_{x,t} \tag{81}$$

In the equations above, the spatial integral term $\int_\Omega s_{x',t} i_{x',t} \psi_{x',x} dx$ captures the infections generated by the social interactions generated by commuting and business trips which take place on a daily basis (La Torre et al. 2022). Because of such intra-day movements individuals get in contact with a number of individuals originally located even far away from their origin location, thus all the contacts between infectives and susceptibles contribute to determine the spread of the disease in a given location. The kernel $\varphi_{x',x}$ where $\int_\Omega \varphi_{x',x} dx = 1$ measures the extent to which these cross-locations contacts determine the level of disease prevalence in a given location. While the diffusion term captures the effects of migration and thus describes a dynamic externality affecting disease dynamics over time, the integral term represents a static externality affecting disease dynamics instantaneously (La Torre et al. 2022).

Similar to what we have discussed in our baseline SIS model, it is possible to show that in the absence of the integral term the system above admits two homogeneous equilibria, one of which is disease-free, $\mathcal{E}^F = (\bar{i}^F, \bar{s}^F, \bar{r}^F)$, and the other endemic, $\mathcal{E}^E = (\bar{i}^E, \bar{s}^E, \bar{r}^E)$, characterized as follows:

$$\begin{aligned} &\mathcal{E}^F : \bar{i}^F = 0, \quad \bar{s}^F = 1, \quad \bar{r}^F = 0 \\ &\mathcal{E}^E : \bar{i}^E = \frac{\epsilon(\alpha - \delta)}{\alpha(\delta + \epsilon)}, \quad \bar{s}^E = \frac{\delta}{\alpha}, \quad \bar{r}^E = \frac{\delta(\alpha - \delta)}{\alpha(\delta + \epsilon)} \end{aligned} \tag{82}$$

Exactly as in our baseline SIS framework, the disease-free equilibrium exists for all parameter values while the endemic one only whenever $\alpha > \delta$, and the system converges to a disease-free situation whenever $\alpha \leq \delta$ or to an endemic situation whenever $\alpha > \delta$. Therefore, also in this case the basic reproduction number is given by (14), thus the same comments discussed in our baseline model still apply. However, note that these conclusions hold true only whenever the integral term is absent (i.e., the kernel takes the form of the Dirac's delta function), while whenever the integral term is present the results will be partially different. In particular, a disease-free and an endemic equilibrium will still exist and will still represent alternative outcomes, but these equilibria will become spatially heterogeneous (i.e., no longer homogeneous) and will be no longer possible to characterize them explicitly. Despite these differences, most our qualitative conclusions still apply and according to the intensity of mitigation policies it may be possible to achieve a disease-free or a endemic outcome, either locally or globally within the spatial economy.

Such an extended SIRS epidemiological framework with local effects requires us to take into account also the role of recovereds in our economic framework. Indeed, since they are healthy individuals we now assume that production depends linearly not only on susceptibles but also on recovereds as follows: $q_{x,t} = s_{x,t} + r_{x,t}$. Taking this into account and recalling that $s_{x,t} = n_{x,t} - i_{x,t} - r_{x,t}$, it follows that the social planner's optimization problem becomes: Find $\xi \in (x_a, x_b)$ which minimizes the following

functional:

$$
\begin{aligned}
\mathcal{C}(\xi) = \min_{u_{x,t}} \; & \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2[1+u_{x,t}^2(n_{x,t}-i_{x,t})^2]}{2} e^{-\rho t} dx dt \\
& + \frac{\mu}{\xi - x_a} \int_0^T \int_{\xi}^{x_b} \frac{i_{x,t}^2}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{x_b} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\
s.t. \; & \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha(1-u_{x,t}) \int_{x_a}^{\xi} (n_{x',t} - i_{x',t} - r_{x',t}) i_{x',t} \psi_{x',x} dx' \\
& -\delta \left[1 + \omega\tau(1-u_{x,t})(n_{x,t} - i_{x,t} - r_{x,t}) + \frac{(1-\beta)\omega\tau}{\xi - x_a} \right. \\
& \left. \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] i_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial r_{x,t}}{\partial t} = d \frac{\partial^2 r_{x,t}}{\partial x^2} \\
& +\delta \left[1 + \omega\tau(1-u_{x,t})(n_{x,t} - i_{x,t} - r_{x,t}) + \frac{(1-\beta)\omega\tau}{\xi - x_a} \right. \\
& \left. \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] i_{x,t} \\
& -\epsilon r_{x,t} \quad x \in [x_a, \xi] \\
& \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + \alpha \int_{\xi}^{x_b} (n_{x',t} - i_{x',t} - r_{x',t}) i_{x',t} \psi_{x',x} dx \\
& -\delta[1 + \beta\omega\tau(n_{x,t} - i_{x,t})] i_{x,t}, \quad x \in [\xi, x_b] \\
& \frac{\partial r_{x,t}}{\partial t} = d \frac{\partial^2 r_{x,t}}{\partial x^2} + \delta[1 + \beta\omega\tau(n_{x,t} - i_{x,t})] i_{x,t} - \epsilon r_{x,t} \quad x \in [\xi, x_b] \\
& \frac{\partial n_{x,t}}{\partial t} = d \frac{\partial^2 n_{x,t}}{\partial x^2} \\
& s_{x,t} = n_{x,t} - i_{x,t} - r_{x,t} \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& \frac{\partial r_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& \frac{\partial n_{x,t}}{\partial x} = 0, \quad x \in \{x_a, x_b, \xi\} \\
& i_{x,0} = i_0(x) > 0 \quad x \in [x_a, x_b] \\
& r_{x,0} = r_0(x) > 0 \quad x \in [x_a, x_b] \\
& n_{x,0} = n_0(x) > 0 \quad x \in [x_a, x_b]
\end{aligned} \tag{83}
$$

In an advanced epidemic setting we can numerically determine the optimal size of the lockdown area and the optimal lockdown intensity exactly as in the previous sections. We continue relying on the same parameter values we have employed in our COVID-19 Italian calibration, and in order to model temporary immunity and local

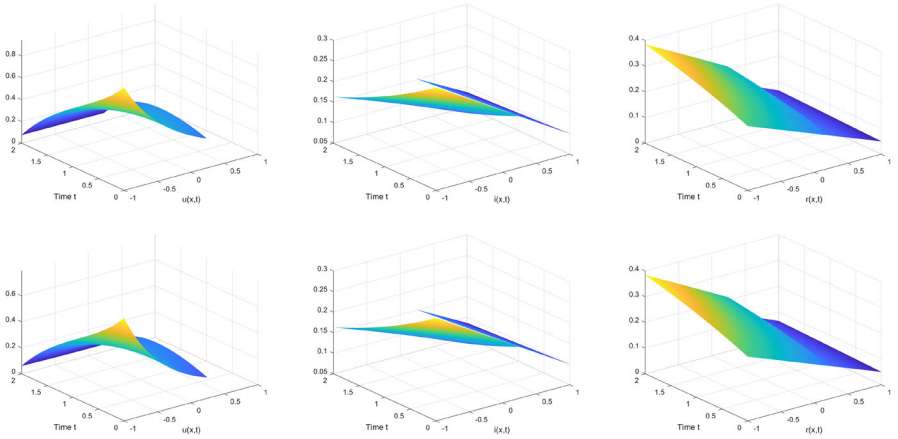

Fig. 11 Spatio-temporal evolution of $u_{x,t}$ (left), $i_{x,t}$ (center) and $r_{x,t}$ (right) with $i_{x,0} = -\frac{0.2}{2}(x+1)+0.21$ and $r_{x,0} = 0.5 i_{x,0}$ with either $\sigma = 1.5$ (top) or $\sigma = 2$ (bottom). Optimal boundary point: $\xi_{opt} = 0.2343$ (top), with $\mathcal{C} = 0.0536$ and $\xi_{opt} = 0.2434$ (bottom), with $\mathcal{C} = 0.0532$

effects we set $\epsilon = 0.0056$ and $\varphi_{x,x'} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x'-x}{\sigma})^2}$ where $\sigma \geq 0$ measures the standard deviation of the kernel and we consider different values of this parameter to understand how it affects our results (La Torre et al. 2022).

Figure 11 shows the spatio-temporal evolution of the lockdown intensity (left panels), infectives share (mid panels) and recovereds share (right panels) under a monotonically decreasing initial disease prevalence distribution in the case in which $\sigma = 0.5$ (top panels) or $\sigma = 1$ (bottom panels). While the evolution of the infectives resembles what we have already discussed in the previous sections and that of the recovereds is intuitive, the most noticeable difference with respect to what we have seen in our baseline model is related to the behavior of the lockdown intensity. Indeed, we can observe that despite the initial condition is linear, the lockdown intensity is highly nonlinear to account for the effects of the integral term, which quantifies how cross-location proximity effects impact disease prevalence in a specific location. In particular, the central (lateral) locations within the lockdown area are the ones which are affected by the influence of prevalence from a larger (smaller) number of surrounding locations. However, the extent to which such cross-location effects impact prevalence in a specific location x depends on the value of σ. A higher standard deviation increases the total number of locations with non-negligible impact on x, but only those within the lockdown area effectively matter (there is no flow in and out of the boundary point ξ), thus the higher σ the smaller the number of relevant proximity effects for the epidemic dynamics in location x. This is the reason why the lockdown intensity is on average higher and the lockdown area is smaller when the standard deviation is smaller. A smaller standard deviation by increasing the number of relevant proximity effects requires to reduce the size of the lockdown area and to increase the lockdown intensity in order to account for the augmented disease incidence induced by cross-location proximity effects.

8 Conclusions

The ongoing COVID-19 pandemic has shown more clearly than ever that the consequences of infectious diseases on macroeconomic activity may be particularly dramatic. It has also shown that such effects are to a large extent heterogeneous between and within countries, which thus requires policymakers to differentiate the intensity and the type of intervention tools employed at local levels in order to contain the spread of the disease without excessively compromising economic activity. In order to shed some light on how such a policy differentiation may need to be implemented, our paper characterizes from a normative perspective the optimal regional policy to contain the spread of a communicable disease in a spatial framework with endogenous determination of the regional borders characterizing which policy regime will prevail in a given region. Specifically, the social planner needs to choose how to split the entire spatial economy in a number of regions in which a different combination of lockdown and treatment measures will be employed: in some region the only mitigation instrument will be treatment, while in some other treatment will be accompanied by a partial lockdown. We characterize the optimal solution both in an early and an advanced epidemic setting in which the disease prevalence in the total population is and is not negligible, respectively. We show that according to the specific circumstances, it may be convenient either to partition the spatial economy in multiple regions with differentiated policies or to consider it as a unique region subject to the same policy measures. Moreover, we show that from a normative perspective it is rather difficult to understand how to effectively determine the optimal size of a lockdown area (and thus the optimal lockdown intensity in the area) since this critically depends on a number of factors, including the initial spatial distribution of disease prevalence, the amount of resources diverted from one region to the other, and the possible spatio-temporal evolution of the disease. We present a calibration based on the COVID-19 experience in Italy during its first epidemic wave showing how the prescribed solution may change under alternative initial distributions of the disease prevalence across space.

To the best of our knowledge, this is the first paper characterizing analytically the optimal lockdown area in a spatial setting and also the first paper analyzing the optimal determination of the borders between regions in a spatial setting. Given the complexity of the problem under investigation we have tried to maintain the setup as simple as possible in order to understand the mechanisms underlying the formation of such optimal choices. However, this has precluded us the possibility to account for some important features observed in the real world during the COVID-19 epidemic. In particular, lockdown measures have generated important effects not only on disease prevalence and economic activity but also on the natural environment, thus introducing some mutual links among epidemics, economies and pollution may allow us to describe more accurately different dimensions of the COVID-19 experience in industrialized countries. Moreover, lockdown decisions have often been made not at a centralized national level but at a decentralized regional level giving often rise to free-riding behavior between regions, thus introducing some strategic interactions between different geographical units may allow us to characterize the externalities imposed by free riding on disease and macroeconomic dynamics and thus to determine how to eventually decentralize

the social optimum. Extending the analysis along these directions is left for future research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00199-023-01517-w.

A Proofs of the main results

A.1 Proof of Lemma 1

The proof is a straightforward application of Sylvester's formula (see Bhatia 1997).

A.2 Proof of Theorem 1

For any fixed $\xi \in (x_a, x_b)$, the model can be analyzed via a two-stage approach. First, we determine the solution in region B which is independent of region A's, and then we plug this solution into region A's problem determining its solution. The share of infectives $i_{x,t}$ over the region $[\xi, x_b]$ solves the boundary value problem:

$$\begin{cases} \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + (\alpha s - \delta - \delta\beta\tau s\omega)\, i_{x,t} & x \in [\xi, x_b] \\ \frac{\partial i_{x,t}}{\partial x} = 0 & x \in \{\xi, x_b\} \\ i_{x,0} = i_0(x) & x \in [\xi, x_b] \end{cases} \tag{84}$$

The above problem admits a closed-form solution given by:

$$i_{x,t} = e^{(\alpha s - \delta - \delta\beta\tau\omega s)t} h_{x,t}, \tag{85}$$

where $h_{x,t}$ is the well-known classical solution of the heat equation with Neumann boundary conditions, given by the following expression:

$$h_{x,t} = \sum_{n\geq 0} C_n e^{-d\left(\frac{n\pi}{x_b - x_a}\right)^2 t} \cos\left[\frac{n\pi(x-\xi)}{x_b - x_a}\right] \tag{86}$$

where:

$$C_0 = \frac{1}{x_b - \xi}\int_{\xi}^{x_b} i_0(x)dx \tag{87}$$

$$C_n = \frac{2}{x_b - \xi}\int_{\xi}^{x_b} i_0(x)\cos\left[\frac{n\pi(x-\xi)}{x_b - \xi}\right]dx \tag{88}$$

It is then straightforward to show the following:

$$i_{x,t} = \sum_{n\geq 0} C_n e^{-d\left(\frac{n\pi}{x_b - \xi}\right)^2 t} e^{(\alpha s - \delta - \delta\beta\tau\omega s)t} \cos\left[\frac{n\pi(x-\xi)}{x_b - \xi}\right], \tag{89}$$

from which, with some algebraic computations and the integration term by term of the Fourier series of $i_{x,t}^2$, it is possible to determine the value of the following terms:

$$\begin{aligned}\Omega_1 &= \int_0^T \int_\xi^{x_b} \frac{i_{x,t}^2}{2} e^{-\rho t} dx dt = \frac{1}{2} \sum_{n\geq 0} C_n^2 \int_0^T e^{2\left[-d\left(\frac{n\pi}{x_b-\xi}\right)^2+(\alpha s-\delta-\delta\beta\tau\omega s)\right]t} dt \\ &= \sum_{n\geq 0} C_n^2 \left[\frac{1-e^{2\left[-d\left(\frac{n\pi}{x_b-\xi}\right)^2+\alpha s-\delta-\delta\beta\tau\omega s\right]T}}{-d\left(\frac{n\pi}{x_b-\xi}\right)^2+\alpha-\delta-\delta\beta\tau\omega s} \right] \qquad (90)\end{aligned}$$

$$\Omega_2 = \int_\xi^{x_b} \frac{i_{x,T}^2}{2} dx = \frac{1}{2} \sum_{n\geq 0} C_n^2 e^{2\left[-d\left(\frac{n\pi}{x_b-\xi}\right)^2+(\alpha s-\delta-\delta\beta\tau\omega s)\right]T} \qquad (91)$$

Once the solution over the region $[\xi, x_b]$ has been determined and plugged into, the model in the region $[x_a, \xi]$ can be restated as follows:

$$\begin{aligned}&\min_{u_{x,t}\in\mathcal{U}} \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2[1+s^2u_{x,t}^2]}{2} e^{-\rho t} dx dt + \frac{\mu}{\xi-x_a}\Omega_1 + \phi\Omega_2 + \phi \int_{x_a}^{\xi} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\ &s.t.\ \ \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + (1-u_{x,t})s i_{x,t}(\alpha-\delta\tau\omega) \\ &\qquad -\delta\left(1+\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi-x_a}\right) i_{x,t}, \quad x\in[x_a,\xi] \\ &\qquad \frac{\partial i_{x,t}}{\partial x} = 0, \quad x\in\{x_a,\xi\} \\ &\qquad i_{x,0} = i_0(x) \quad x\in[x_a,\xi] \qquad (92)\end{aligned}$$

where:

$$\mathcal{U} = \{u_{x,t} : [x_a,\xi]\times[0,T]\to\mathcal{R} : u_{x,t} \text{ is continuous}, 0 < u_{x,t} < 1\} \qquad (93)$$

By noticing that $\frac{\mu}{\xi-x_a}\Omega_1+\phi\Omega_2$ is just a translation term, the above model is totally equivalent to the following:

$$\begin{aligned}&\min_{u\in\mathcal{U}} \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2+(su_{x,t}i_{x,t})^2}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{\xi} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\ &s.t.\ \ \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \left[s(\alpha-\delta\tau\omega)-\delta-\delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi-x_a}\right)\right] i_{x,t} \\ &\qquad -su_{x,t}i_{x,t}(\alpha-\delta\tau\omega), \quad x\in[x_a,\xi] \\ &\qquad \frac{\partial i_{x,t}}{\partial x} = 0, \quad x\in\{x_a,\xi\} \\ &\qquad i_{x,0} = i_0(x) \quad x\in[x_a,\xi] \qquad (94)\end{aligned}$$

In order to determine the optimality conditions we rely on the method based on the current value Hamiltonian function in (x, t) (Troltzsch 2010), which reads as follows:

$$\mathcal{H} = \frac{i_{x,t}^2 + (su_{x,t}i_{x,t})^2}{2} + \lambda_{x,t}\left[d\frac{\partial^2 i_{x,t}}{\partial x^2} + \left[s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b - \xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] i_{x,t} - su_{x,t}i_{x,t}(\alpha - \delta\tau\omega)\right],$$

The FOCs for a minimum are given by the following expressions:

$$\begin{cases} \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \left[s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] i_{x,t} - su_{x,t}i_{x,t}(\alpha - \delta\tau\omega) \\ \frac{\partial \lambda_{x,t}}{\partial t} = \rho\lambda_{x,t} - d\frac{\partial^2 \lambda_{x,t}}{\partial x^2} - i_{x,t} - \lambda_{x,t}\left[s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] \\ u_{x,t} = \frac{\lambda_{x,t}(\alpha - \delta\tau\omega)}{si_{x,t}} \\ \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ \frac{\partial \lambda_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ i_{x,0} = i_0(x) \quad x \in [x_a, \xi] \\ \lambda_{x,T} = \phi i_{x,T} \quad x \in [x_a, \xi] \end{cases} \tag{95}$$

By replacing the control variable $u_{x,t}$, the system of FOCs reads as follows:

$$\begin{cases} \frac{\partial i_{x,t}}{\partial t} = d\frac{\partial^2 i_{x,t}}{\partial x^2} + \left[s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] i_{x,t} - \lambda_{x,t}(\alpha - \delta\tau\omega)^2 \\ \frac{\partial \lambda_{x,t}}{\partial t} = -d\frac{\partial^2 \lambda_{x,t}}{\partial x^2} - i_{x,t} - \lambda_{x,t}\left[-\rho + s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right)\right] \\ \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ \frac{\partial \lambda_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ i_{x,0} = i_0(x) \quad x \in [x_a, \xi] \\ \lambda_{x,T} = \phi i_{x,T} \quad x \in [x_a, \xi] \end{cases} \tag{96}$$

By recalling the definition of the matrix Θ and by introducing the matrix D as it follows:

$$\Theta := \begin{bmatrix} s(\alpha - \delta\tau\omega) - \delta - \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right) & -(\alpha - \delta\tau\omega)^2 \\ -1 & \rho - s(\alpha - \delta\tau\omega) + \delta + \delta\left(\frac{(x_b-\xi)(1-\beta)\tau\omega s}{\xi - x_a}\right) \end{bmatrix} \tag{97}$$

$$D := \begin{bmatrix} d & 0 \\ 0 & -d \end{bmatrix} \tag{98}$$

and the vector:

$$z_{x,t} := \begin{bmatrix} i_{x,t} \\ \lambda_{x,t} \end{bmatrix} \tag{99}$$

the optimality conditions in vector form can be restated as:

$$\begin{cases} \frac{\partial z_{x,t}}{\partial t} = D\frac{\partial^2 z_{x,t}}{\partial x^2} + \Theta z_{x,t} \\ \frac{\partial z_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\ z^1_{x,0} = i_0(x) \quad zx \in [x_a, \xi] \\ z^2_{x,T} = \phi z^1_{x,T} \quad x \in [x_a, \xi] \end{cases} \tag{100}$$

By introducing the exponential matrix $e^{\Theta t}$ and the variable $\tilde{z}_{x,t} = e^{-\Theta t} z_{x,t}$, we get:

$$\frac{\partial \tilde{z}_{x,t}}{\partial t} = D\frac{\partial^2 \tilde{z}_{x,t}}{\partial x^2}, \tag{101}$$

where $\tilde{z}_{x,t}$ is the solution to the classical heat equation given by:

$$\tilde{z}_{x,t} = \begin{bmatrix} \sum_{n\geq 0} A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 t} \cos\left(n\pi \left[\frac{x-x_a}{\xi - x_a}\right]\right) \\ \sum_{n\geq 0} B_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 (T-t)} \cos\left(n\pi \left[\frac{x-x_a}{\xi - x_a}\right]\right) \end{bmatrix} \tag{102}$$

We now wish to determine the expressions of A_n and B_n. If we plug $t = 0$ we get that $e^{\Theta 0} = I$ and then $\tilde{z}_{x,0} = I z_{x,0}$, thus the first component of z boils down to:

$$\tilde{z}^1_{x,0} = i_0(x) = \sum_{n\geq 0} A_n \cos\left(n\pi \left[\frac{x - x_a}{\xi - x_a}\right]\right),$$

which implies that A_0 and A_n are the Fourier coefficients of i_0, that is:

$$A_0 = \frac{1}{\xi - x_a}\int_{x_a}^{\xi} i_0(x)dx$$
$$A_n = \frac{2}{\xi - x_a}\int_{x_a}^{\xi} i_0(x)\cos\left(n\pi \left[\frac{x - x_a}{\xi - x_a}\right]\right)dx$$

In order to determine the expression of B_n as function of A_n, let us consider the terminal condition. By plugging $t = T$ into the expression of $z_{x,t}$ we obtain:

$$z_{x,T} = e^{\Theta T}\begin{bmatrix} \tilde{z}^1_{x,T} \\ \tilde{z}^2_{x,T} \end{bmatrix} = e^{\Theta T}\begin{bmatrix} \sum_{n\geq 0} A_n e^{-d\left(\frac{n\pi}{\xi - x_a}\right)^2 T} \cos\left(n\pi \left[\frac{x-x_a}{\xi - x_a}\right]\right) \\ \sum_{n\geq 0} B_n \cos\left(n\pi \left[\frac{x-x_a}{\xi - x_a}\right]\right) \end{bmatrix},$$

and by using the terminal condition:

$$z^2_{x,T} = \phi z^1_{x,T}$$

we get the following system:

$$\frac{1}{\phi}=\frac{e_{11}^{\Theta T}\left(\sum_{n\geq 0}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T}\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)\right)+e_{12}^{\Theta T}\left(\sum_{n\geq 0}B_n\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)\right)}{e_{21}^{\Theta T}\left(\sum_{n\geq 0}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T}\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)\right)+e_{22}^{\Theta T}\left(\sum_{n\geq 0}B_n\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)\right)}$$

which can be transformed into

$$\frac{1}{\phi}=\frac{\sum_{n\geq 0}\left(e_{11}^{\Theta T}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T}+e_{12}^{\Theta T}B_n\right)\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)}{\sum_{n\geq 0}\left(e_{21}^{\Theta T}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T}+e_{22}^{\Theta T}B_n\right)\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)}$$

and then:

$$\begin{aligned}&\left\{\sum_{n\geq 0}\left(e_{21}^{\Theta T}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T}+e_{22}^{\Theta T}B_n\right)\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)\right\}\\&=\phi\left\{\sum_{n\geq 0}\left(e_{11}^{\Theta T}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T}+e_{12}^{\Theta T}B_n\right)\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)\right\},\end{aligned}$$

from which we finally get the expressions of B_n in terms of A_n by noticing the following:

$$B_n=\left[\frac{e_{21}^{\Theta T}-\phi e_{11}^{\Theta T}}{\phi_2 e_{12}^{\Theta T}-e_{22}^{\Theta T}}\right]A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 T} \tag{103}$$

From this we easily get the expression of $i_{x,t}^A$ which is given by:

$$\begin{aligned}i_{x,t}^A=&\sum_{n\geq 0}\left(e_{11}^{\Theta t}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 t}+e_{12}^{\Theta t}B_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 (T-t)}\right)\\&\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right),\end{aligned} \tag{104}$$

and by using the expression linking the control $u_{x,t}$ with the costate variable $\lambda_{x,t}$ we obtain:

$$u_{x,t}^A=\frac{\sum_{n\geq 0}(\alpha-\delta\tau\omega)\left(e_{21}^{\Theta t}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 t}+e_{22}^{\Theta t}B_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 (T-t)}\right)\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)}{\sum_{n\geq 0}\left(e_{11}^{\Theta t}A_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 t}+e_{12}^{\Theta t}B_n e^{-d\left(\frac{n\pi}{\xi-x_a}\right)^2 (T-t)}\right)\cos\left(n\pi\left[\frac{x-x_a}{\xi-x_a}\right]\right)} \tag{105}$$

A.3 Proof of Corollary 1

The proof of this result is quite straightforward and it follows by classical results in Fourier series theory. In fact, it is enough to use Parseval's identity applied to $i^A_{x,t}$ and $i^B_{x,t}$ over the two separate regions $A = [x_a, \xi]$ and $B = [\xi, x_b]$ and then recombine the two L^2 norms to get the thesis.

A.4 Proof of Proposition 1

The proof is straightforward and it follows by taking the spatial integrals of both sides, by using Neumann's conditions, and by using classical comparison results for ordinary differential equations.

A.5 Proof of Corollary 2

The proof follows by replacing i_0 into the Fourier coefficients and recalling that the cosine basis is orthogonal.

A.6 Proof of Corollary 3

The proof is quite straightforward and it follows by taking the classical first order derivative of

$$C(\xi) = C_1(\xi) + C_2(\xi) + C_3(\xi) \tag{106}$$

with respect to ξ. The hypothesis on β, $\beta = 1$, makes most of the terms involved in the expression of $C(\xi)$ to be independent from ξ. In particular, this is true for η_1 and η_2. The condition on the parameters implies that $\xi < x_b$ and standard calculus arguments allow to conclude that ξ is an internal minimizer.

A.7 Proof of Theorem 2

The problem can be analyzed by following the same steps discussed in the early epidemic setting. However, in this case, it is possible to provide only a characterization of the optimal solution in terms of optimality conditions as the nonlinearity prevents the possibility to determine a closed-form solution. The share of infectives $i_{x,t}$ over $[\xi, x_b]$ solves the equation:

$$\begin{cases} \frac{\partial i^B_{x,t}}{\partial t} = d\frac{\partial^2 i^B_{x,t}}{\partial x^2} + (\alpha - \delta\beta\tau) i^B_{x,t}(n^B_{x,t} - i^B_{x,t}) - \delta i^B_{x,t}, & x \in [\xi, x_b] \\ \frac{\partial i^B_{x,t}}{\partial x} = 0, & x \in \{\xi, x_b\} \\ i^B_{x,0} = i_0(x) & x \in [\xi, x_b] \end{cases} \tag{107}$$

This equation admits a unique solution $i^B_{x,t}$ which is defined over the interval $[\xi, x_b]$. By plugging this expression into the integral and the objective function, we can determine the value of the two terms:

$$\Omega_1 = \frac{1}{2}\int_0^T \int_\xi^{x_b} \frac{\left(i^B_{x,t}\right)^2}{2} e^{-\rho t} dx dt \tag{108}$$

$$\Omega_2 = \int_\xi^{x_b} \frac{\left(i^B_{x,T}\right)^2}{2} e^{-\rho T} dx \tag{109}$$

We need thus to solve the following problem:

$$\begin{aligned}
\min \ & \int_0^T \int_{\xi_a}^{\xi} \frac{i^2_{x,t}[1+(1-i^A_{x,t})^2 (u^A_{x,t})^2]}{2} e^{-\rho t} dx dt + \phi \int_{x_a}^{\xi} \frac{(i^A_{x,T})^2}{2} e^{-\rho T} dx \\
& + \frac{\mu}{\xi - x_a} \Omega_1 + \phi \Omega_2 \\
s.t. \ & \frac{\partial i^A_{x,t}}{\partial t} = d \frac{\partial^2 i^A_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega)(1-u^A_{x,t}) i_{x,t}(n^A_{x,t} - i^A_{x,t}) \\
& -\delta \left[1 + \frac{(1-\beta)\tau\omega}{s-a} \int_\xi^{x_b} (n^B_{x,t} - i^B_{x,t}) dx\right] i^A_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial n^A_{x,t}}{\partial t} = d \frac{\partial^2 n^A_{x,t}}{\partial x^2} \\
& \frac{\partial i^A_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\
& \frac{\partial n^A_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\
& i^A_{x,0} = i_0(x) \quad x \in [x_a, \xi] \\
& n^A_{x,0} = n_0(x) \quad x \in [x_a, \xi]
\end{aligned} \tag{110}$$

Note that the expression of $n_{x,t}$ over the interval $[x_a, x_b]$ is known and its closed-form is provided by:

$$n_{x,t} = \begin{cases} \sum_{n\geq 0} B^1_n e^{-\left(\frac{n\pi}{\xi - x_a}\right)^2 dt} \cos\left[\frac{n\pi(x-x_a)}{\xi - x_a}\right] & x \in [x_a, \xi] \\ \sum_{n\geq 0} B^2_n e^{-\left(\frac{n\pi}{x_b - \xi}\right)^2 dt} \cos\left[\frac{n\pi(x-\xi)}{x_b - \xi}\right] & x \in [\xi, x_b] \end{cases} \tag{111}$$

where $B^1_0 = \frac{1}{\xi - x_a}\int_{x_a}^{\xi} n_{x,0} dx$ and $B^1_n = \frac{2}{\xi - x_a}\int_{x_a}^{\xi} n_{x,0} \cos\left[\frac{n\pi(x-x_a)}{\xi - x_a}\right] dx$ and $B^2_0 = \frac{1}{x_b - \xi}\int_\xi^{x_b} n_{x,0} dx$ and $B^2_n = \frac{2}{x_b - \xi}\int_\xi^{x_b} n_{x,0} \cos\left[\frac{n\pi(x-\xi)}{x_b - \xi}\right] dx$.

By dropping the translation term, the model reads as:

$$\begin{aligned}
\min \quad & \int_0^T \int_{x_a}^{\xi} \frac{i_{x,t}^2 + (n_{x,t} - i_{x,t})^2 u_{x,t}^2 i_{x,t}^2}{2} e^{-\rho t} dx dt + \phi_2 \int_{x_a}^{\xi} \frac{i_{x,T}^2}{2} e^{-\rho T} dx \\
s.t. \quad & \frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega) i_{x,t}(n_{x,t} - i_{x,t})(1 - u_{x,t}) \\
& \quad -\delta \left[1 + \frac{(1-\beta)\tau\omega}{\xi - x_a} \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] i_{x,t}, \quad x \in [x_a, \xi] \\
& \frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\
& i_{x,0} = i_0(x) \quad x \in [x_a, \xi]
\end{aligned} \tag{112}$$

The current value Hamiltonian function in (x, t) reads as:

$$\begin{aligned}
\mathcal{H} = & \frac{i_{x,t}^2 + (n_{x,t} - i_{x,t})^2 u_{x,t}^2 i_{x,t}^2}{2} + \lambda_{x,t} \left\{ d \frac{\partial^2 i_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega) i_{x,t}(n_{x,t} - i_{x,t})(1 - u_{x,t}) \right. \\
& \left. - \delta \left[1 + \frac{(1-\beta)\tau\omega}{\xi - x_a} \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] i_{x,t} \right\}
\end{aligned}$$

The optimal solution over the region $[x_a, \xi]$ can be characterized by means of the FOCs which read as (Troltzsch 2010):

$$\begin{cases}
\frac{\partial i_{x,t}}{\partial t} = d \frac{\partial^2 i_{x,t}}{\partial x^2} + (\alpha - \delta\tau\omega) i_{x,t}(n_{x,t} - i_{x,t})(1 - u_{x,t}) \\
\quad -\delta \left[1 + \frac{(1-\beta)\tau\omega}{\xi - x_a} \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] i_{x,t} \\
\frac{\partial \lambda_{x,t}}{\partial t} = \rho \lambda_{x,t} - d \frac{\partial^2 \lambda_{x,t}}{\partial x^2} - i_{x,t} \\
\quad -\lambda_{x,t} \left[(\alpha - \delta\tau\omega)(n_{x,t} - 2 i_{x,t}) - \delta - \delta \frac{(1-\beta)\tau\omega}{\xi - x_a} \int_{\xi}^{x_b} (n_{x,t} - i_{x,t}) dx \right] \\
u_{x,t} = \frac{\lambda_{x,t}(\alpha - \delta\tau\omega)}{i_{x,t}(n_{x,t} - i_{x,t})} \\
\frac{\partial i_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\
\frac{\partial \lambda_{x,t}}{\partial x} = 0, \quad x \in \{x_a, \xi\} \\
i_{x,0} = i_0(x) \quad x \in [x_a, \xi] \\
\lambda_{x,T} = \phi_2 i_{x,T} \quad x \in [x_a, \xi]
\end{cases} \tag{113}$$

A.8 Proof of Proposition 2

The proof is straightforward. It follows by taking the spatial integrals of both sides, by using Neumann's conditions, and by using classical comparison results for ordinary differential equations.

A.9 Proof of Theorem 3

The proof is very similar to the one presented for Theorem 2. For any fixed pair of ξ_1 and ξ_2, the first step is to determine the amount of infectives over the regions B and C which can be done by solving two logistic-type partial differential equations. After substituting both of them into the model for the region A, the second step involves the characterization of the optimal solution by means of the FOCs.

A.10 Proof of Theorem 4

The proof is very similar to the one presented for Theorem 2. For any fixed pair of ξ_1 and ξ_2, the first step is to determine the amount of infectives over the region B which can be done by solving a logistic-type partial differential equation. After substituting it into the model for the regions A and C, the second step involves the characterization of the optimal solutions over A and C by means of the FOCs.

References

Acemoglu, D., Chernozhukov, V., Werning, I., Whinston, M.D.: A multi-risk SIR model with optimally targeted lockdown. Am. Econ. Rev. Insights **3**, 487–502 (2021)
Acemoglu, D., Johnson, S.: Disease and development: the effect of life expectancy on economic growth. J. Polit. Econ. **115**, 925–985 (2007)
Adda, J.: Economic activity and the spread of viral diseases: evidence from high frequency data. Q. J. Econ. **131**, 891–941 (2016)
Alvarez, F.E., Argente, D., Lippi, F.: A simple planning problem for COVID-19 lockdown. Am. Econ. Rev. Insights **3**, 367–382 (2021)
Anita, S., Capasso, V.: Reaction–diffusion systems in epidemiology (2017). arXiv:1703.02760
Amdaoud, M., Arcuri, G., Levratto, N.: Are regions equal in adversity? A spatial analysis of spread and dynamics of COVID-19 in Europe. Eur. J. Health Econ. forthcoming (2021)
Anderson, S.T., Laxminarayan, R., Salant, S.W.: Diversify or focus? Spending to combat infectious diseases when budgets are tight. J. Health Econ. **31**, 658–675 (2010)
Atkeson, A.: What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios. In: NBER Working Paper 26867 (2020)
Bhatia, R.: Matrix analysis. In: Graduate Texts in Mathematics, vol. 169. Springer. ISBN 978-0-387-94846-1 (1997)
Bleakley, H.: Disease and development: evidence from hookworm eradication in the American South. Q. J. Econ. **122**, 73–117 (2007)
Birge, J., Candogan, O., Feng, Y.: Controlling epidemic spread: reducing economic losses with targeted closure. In: Working Paper 2020-57. Becker Friedman Institute, University of Chicago (2020)
Bisin, A., Moro, A.: Learning epidemiology by doing: the empirical implications of a spatial-SIR model with behavioral responses. In: NBER Working Paper 27590 (2020)
Bognanni, M., Hanley, D., Kolliner, D., Mitman, K.: Economic activity and COVID–19 transmission: evidence from an estimated economic–epidemiological model. In: IZA Discussion Paper 13797 (2020)
Bourdin, S., Ludovic, J., Nadou, F., Noiret, G.: Does lockdown work? A spatial analysis of the spread and concentration of COVID-19 in Italy. Reg. Stud. forthcoming (2021)
Bloom, D.E., Canning, D., Sevilla, J.: The effect of health on economic growth: a production function approach. World Dev. **32**, 1–13 (2004)
Boucekkine, R., Desbordes, R., Latzer, H.: How do epidemics induce behavioral changes? J. Econ. Growth **14**, 233–264 (2009)
Boucekkine, R., Camacho, C., Zou, B.: Bridging the gap between growth theory and economic geography: the spatial Ramsey model. Macroecon. Dyn. **13**, 20–45 (2009)

Boucekkine, R., Camacho, C., Fabbri, G.: Spatial dynamics and convergence: the spatial AK model. J. Econ. Theory **148**, 2719–2736 (2013)
Boucekkine, R., Fabbri, G., Federico, S., Gozzi, F.: Growth and agglomeration in the heterogeneous space: a generalized AK approach. J. Econ. Geogr. **19**, 1287–1318 (2019)
Brock, W.A., Xepapadeas, A.: Diffusion-induced instability and pattern formation in infinite horizon recursive optimal control. J. Econ. Dyn. Control **32**, 2745–2787 (2008)
Brock, W.A., Xepapadeas, A.: Pattern formations, spatial externalities and regulation in a coupled economic-ecological system. J. Environ. Econ. Manag. **59**, 149–164 (2010)
Brock, W.A., Xepapadeas, A.: Climate change policy under polar amplification. Eur. Econ. Rev. **94**, 263–282 (2017)
Camacho, C., Zou, B., Briani, M.: On the dynamics of capital accumulation across space. Eur. J. Oper. Res. **186**(2), 451–465 (2008)
Camacho, C., Desbordes, R., La Torre, D.: A time-space integro-differential economic model of epidemic control. Econ. Theor. (2023). https://doi.org/10.1007/s00199-023-01506-z
Casas, E., Ryll, C., Tröltzsch, F.: Optimal control of a class of reaction–diffusion systems. Comput. Optim. Appl. **70**, 677–707 (2018)
Cervellati, M., Sunde, U., Valmori, S.: Pathogens, weather shocks and civil conflict. Econ. J. **127**, 2581–2616 (2017)
Cheng, C., Barceló, J., Hartnett, A.S., Kubinec, R., Messerschmidt, L.: COVID-19 government response event dataset (CoronaNet vol 1.0). Nat. Hum. Behav. **4**, 756–768 (2020)
Chowell, G., Sattenspiel, L., Bansal, S., Viboud, C.: Mathematical models to characterize early epidemic growth: a review. Phys. Life Rev. **18**, 66–97 (2016). https://doi.org/10.1016/j.plrev.2016.07.005. Epub 2016 Jul 11. PMID: 27451336; PMCID: PMC5348083
Cuñat, A., Zymek, R.: The (structural) gravity of epidemics. Covid Econ. Vetted Real Time Pap. **17**, 153–173 (2020)
de Frutos, J., Martin-Herran, G.: Spatial effects and strategic behavior in a multiregional transboundary pollution dynamic game. J. Environ. Econ. Manag. **97**, 182–207 (2019)
de Frutos, F., Martin-Herran, G.: Spatial vs. non-spatial transboundary pollution control in a class of cooperative and non-cooperative dynamic games. Eur. J. Oper. Res. **276**, 379–394 (2019)
Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., Calabrese, C., Guarino, A., Cardona-Rivera, R., De Lellis, P., Liuzza, D., Lo Iudice, F., Russo, G., di Bernardo, M.: A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. Nat. Commun. **11**, 5106 (2020)
Desmet, K., Wacziarg, R.: Understanding spatial variation in COVID-19 across the United States. J. Urban Econ. forthcoming (2021)
Dong, E., Du, H., Gardner, L.: An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect. Dis. **20**, 533–534 (2020)
Dontchev, A.L., Zolezzi, T.: Well-posedness in optimal control. In: Morel, J.-M., Teissier, B. (eds) Well-Posed Optimization Problems, Lecture Notes in Mathematics, vol. 1543, pp. 176–229. Springer, Berlin (1993)
Eichenbaum, M., Rebelo, S., Trabandt, M.: The macroeconomics of epidemics. Rev. Financ. Stud. **34**, 5149–5187 (2021)
Fajgelbaum, P.D., Khandelwal, A., Kim, W., Mantovani, C., Schaal, E.: Optimal lockdown in a commuting network. In: NBER Working Paper 27441 (2020)
Fleming, W.H., Rishel, R.W.: Deterministic and Stochastic Optimal Control. Springer, New York (1975)
Francetic, I., Munford, L.: Corona and coffee on your commute: a spatial analysis of COVID-19 mortality and commuting flows in England in 2020. Eur. J. Public Health, ckab072 (2021)
Friedman, A.: One dimensional Stefan problems with non monotone free boundary. Trans. Am. Math. Soc. **133**, 89–114 (1968)
Friedman, A.: Partial Differential Equations of Parabolic Type. Dover Publications, Mineola (2008)
Gersovitz, M., Hammer, J.S.: Infectious diseases, public policy and the marriage of economics and epidemiology. World Bank Res. Obs. **18**, 129–157 (2003)
Gersovitz, M., Hammer, J.S.: The economical control of infectious diseases. Econ. J. **114**, 1–27 (2004)
Giannone, E., Paixao, N., Pang, X.: The geography of pandemic containment. Covid Econ. Vetted Real Time Pap. **52**, 68–95 (2020)
Goenka, A., Liu, L.: Infectious diseases and endogenous fluctuations. Econ. Theor. **50**, 125–149 (2012). https://doi.org/10.1007/s00199-010-0553-y

Goenka, A., Liu, L.: Infectious diseases, human capital and economic growth. Econ. Theor. **70**, 1–47 (2019). https://doi.org/10.1007/s00199-019-01214-7
Goenka, A., Liu, L., Nguyen, M.H.: Infectious diseases and economic growth. J. Math. Econ. **50**, 34–53 (2014)
Goldman, S.M., Lightwood, J.: Cost optimization in the SIS model of infectious disease with treatment. Top. Econ. Anal. Policy **2**, Article 4 (2002)
Hethcote, H.W.: The mathematics of infectious diseases. SIAM Rev. **42**, 599–653 (2000)
Hotelling, H.: Stability in competition. Econ. J. **39**, 41–57 (1929)
Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. Ser. A **115**, 700–721 (1927)
Keystone: Coronavirus city and county non-pharmaceutical intervention rollout date dataset. https://www.keystonestrategy.com/coronavirus-covid19-intervention-dataset-model/ (2020)
Klasing, M.J., Milionis, P.: The international epidemiological transition and the education gender gap. J. Econ. Growth **25**, 37–86 (2020)
Kraemer, M.U.G., Yang, C.H., Gutierrez, B., Wu, C.H., Klein, B., Pigott, D.M., Open COVID-19 Data Working Group, du Plessis, L., Faria, N.R., Li, R., Hanage, W.P., Brownstein, J.S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O.G., Scarpino, S.V: The effect of human mobility and control measures on the COVID-19 epidemic in China. Science **368**, 493–497 (2020)
James, G., Dyke, P.P.G., Searl, J., Craven, M., Wei, Y.: Modern Engineering Mathematics. Pearson, UK (2020)
Jung, E., Lenhart, S., Feng, Z.: Optimal control of treatments in a two-strain 96 tuberculosis model. Discrete Contin. Dyn. Syst. (Ser. B) **2**, 473–482 (2002)
La Torre, D., Liuzzi, D., Marsiglio, S.: Pollution diffusion and abatement activities across space and over time. Math. Soc. Sci. **78**, 48–63 (2015)
La Torre, D., Liuzzi, D., Marsiglio, S.: Population and geography do matter for sustainable development. Environ. Dev. Econ. **24**, 201–223 (2019)
La Torre, D., Liuzzi, D., Marsiglio, S.: The optimal population size under pollution and migration externalities: a spatial control approach. Math. Modell. Nat. Phenom. **14**, 104 (2019)
La Torre, D., Liuzzi, D., Marsiglio, S.: Epidemics and macroeconomic outcomes: social distancing intensity and duration. J. Math. Econ. **93**, 102473 (2021)
La Torre, D., Liuzzi, D., Marsiglio, S.: Transboundary pollution externalities: think globally, act locally? J. Math. Econ. **96**, 102511 (2021)
La Torre, D., Liuzzi, D., Marsiglio, S.: Geographical heterogeneities and externalities in an epidemiological-macroeconomic framework. J. Public Econ. Theory **24**, 1154–1181 (2022)
La Torre, D., Malik, T., Marsiglio, S.: Optimal control of prevention and treatment in a basic macroeconomic-epidemiological model. Math. Soc. Sci. **108**, 100–108 (2020)
La Torre, D., Marsiglio, S., Mendivil, F., Privileggi, F.: Stochastic disease spreading and containment policies under state-dependent probabilities. Econ. Theory (2023). https://doi.org/10.1007/s00199-023-01496-y
Lieberman, G.M.: Second Order Parabolic Differential Equations. World Scientific, River Edge (1996)
Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T., Murray, C.J.L.: Global Burden of Disease and Risk Factors. Oxford University Press, New York (2006)
Martcheva, M.: A non-autonomous multi-strain SIS epidemic model. J. Biol. Dyn. **3**, 235–251 (2009)
OECD: OECD Economic Outlook, vol. 2020, issue 2 (2020a)
OECD: The territorial impact of COVID-19: managing the crisis across levels of government (2020b). https://www.oecd.org/coronavirus/policy-responses/the-territorial-impact-of-COVID-19-managing-the-crisis-across-levels-of-government-d3e314e1/
Philipson, T.: Economic epidemiology and infectious disease. In: Cuyler, A.J., Newhouse, J.P. (eds) Handbook of Health Economics, vol. 1B, pp. 1761–1799. North Holland, Amsterdam (2000)
Rubinstein, L.I.: The Stefan Problem, Translations of Mathematical Monographs, vol. 27. American Mathematical Society, Providence (1971)
Protter, M.H., Weinberger, H.F.: Maximum Principles in Differential Equations. Springer, Berlin (1984)
Sanfelici, M.: The Italian response to the COVID-19 crisis: lessons learned and future direction in social development. Int. J. Commun. Soc. Dev. **2**, 191–210 (2020)
Sarris, P.: Viewpoint new approaches to the 'Plague of Justinian'. Past Present **254**, 315–346 (2022)
Stefan, J.: Uber einige prob1eme der theorie der warmeleitung. Sitzungsberichte der Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche Klasse **2**, 173–484 (1899)

Thomas, L.J., Huang, P., Yin, F., Luo, X.I., Almquist, Z.W., Hipp, J.R., Butts, C.T.: Spatial heterogeneity can lead to substantial local variations in COVID-19 timing and severity. Proc. Natl. Acad. Sci. **117**, 24180–24187 (2020)

Troltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods and Applications. American Mathematical Society, Providence (2010)

United Nations: Resolution adopted by the General Assembly on 25 September 2015 (2015). http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

Wang, T.: Dynamics of an epidemic model with spatial diffusion. Physica A **409**, 119–129 (2014)

Wang, B.G., Li, W.T., Wang, Z.C.: A reaction–diffusion SIS epidemic model in an almost periodic environment. Z. Angew. Math. Phys. **66**, 3085–3108 (2015)

World Health Organization: World Health Statistics 2009 (2009). http://www.who.int/gho/publications/world_health_statistics/EN_WHS09_Full.pdf

World Health Organization: Report of the WHO - China joint mission on coronavirus disease 2019 (COVID-19) (2020a). https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf

World Health Organization: Immunity passports in the context of COVID-19: scientific brief (2020b). https://www.who.int/news-room/commentaries/detail/immunity-passports-in-the-context-of-covid-19