WORKING PAPERS

N° 1111

June 2020

"Optimal epidemic suppression under an ICU constraint"

Laurent Miclo, Daniel Spiro
And Jörgen W. Weibull

Toulouse School of Economics

Optimal epidemic suppression under an ICU constraint*

Laurent Miclo[†], Daniel Spiro[‡] and Jörgen Weibull[§]

June 16, 2020

Abstract

How much and when should we limit economic and social activity to ensure that the health-care system is not overwhelmed during an epidemic? We study a setting where ICU resources are constrained and suppression is costly. Providing a fully analytical solution we show that the common wisdom of "flattening the curve", where suppression measures are continuously taken to hold down the spread throughout the epidemic, is suboptimal. Instead, the optimal suppression is discontinuous. The epidemic should be left unregulated in a first phase and when the ICU constraint is approaching society should quickly lock down (a discontinuity). After the lockdown regulation should gradually be lifted, holding the rate of infected constant, thus respecting the ICU resources while not unnecessarily limiting economic activity. In a final phase, regulation is lifted. We call this strategy "filling the box".

Keywords: Epidemic; Optimal control; Health; Suppression; Infection; Corona.

JEL codes: C61; D6; I18.

---

*We wish to thank Atahan Afsar, Tommy Andersson, Jérôme Bolte, Tom Britton, Avinash Dixit, Christian Gollier, Bengt Holmström, Niko Jaakkola, Hannes Malmberg, Lars-Göran Mattsson, Jean Tirole and Robert Östling for valuable comments. LM acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future program (Investissements d'Avenir, grant ANR-17-EURE-0010).

†Toulouse Institute of Mathematics, Toulouse School of Economics and CNRS. Email: laurent.miclo@math.cnrs.fr.

‡Department of Economics, Uppsala University. Email: daniel.spiro.ec@gmail.com.

§Department of Economics, Stockholm School of Economics. Email: jorgen.weibull@hhs.se.

1 Introduction

Amid the Covid-19 health and economic crisis one question stood at the centre of professional opinion: How much and when should we limit economic and social activity to ensure that the health-care system is not overwhelmed? This question embodies two simultaneous goals when fighting a pandemic: (1) To ensure that each infected person gets the best possible care, we need to ensure that the capacity of the health-care system (henceforth the ICU constraint) is never breached. Under Covid-19 the ICU constraint is essentially the number of available respirators, indeed a scarce resource in most countries. It was perhaps best epitomized by the UK slogan "Protect the NHS" and by the Imperial College report (Ferguson et al., 2020). (2) The more one is suppressing the spread the costlier it is since, absent a vaccine, suppression boils down to keeping people away from each other thus limiting economic and social life.

This paper extends the standard S.I.R. model (Kermack and McKendrick, 1927) with those two extensions to provide an *analytical* answer to the above question. Our answer departs from common wisdom. During the Covid-19 pandemic, authorities, news reporting and policy makers popularized the ideal policy as "flattening the curve",[1] i.e., imposing continuous limitations to lower the number of simultaneously infected in all time periods. This would ensure that the peak of the curve never crossed the ICU constraint. We show that this policy is suboptimal. Instead, the optimal policy can characterized as "filling the box" involving a discontinuous suppression. More precisely, it prescribes (Theorem 1 and Figure 2) leaving the spread unregulated during a first phase. As the number of infected approaches the ICU constraint we enter a second phase where harsh suppression measures are imposed at once (a discontinuity) but afterwards gradually relaxed. The aim of policy in this second phase is to precisely stop the number of infected from exceeding the ICU constraint and keep it constant at that level. The discontinuous tightening followed by gradual relaxation of suppression is optimal since the underlying growth of infections is highest in the beginning of this phase. In a third phase, once the underlying growth of infections subsides, no suppression measures are taken.

The logic behind this result is simple, but bears relevance for a disease spreading such as Covid-19. When access to a vaccine is not realistic within a sufficiently near future and pinpointing each infectious person is not feasible, which is implicitly assumed in our model, full eradication

[1]See, e.g., the Imperial College report (Ferguson et al., 2020), Branswell (2020), Time (2020), Pueyo (2020), even Donald Trump (The Sun, 2020) and many more.

is not possible. What remains then, is letting the infection spread in the population but ensuring that each person gets best possible care, i.e., ensuring the available health-care resources are sufficient at all times. But there is no point in leaving some of the respirators idle (if considering risk one can view the ICU constraint as being the number of respirators with a margin). Hence, early suppression is unnecessary and costly. Once the number of infections reaches the ICU constraint, drastic suppression has to be installed to keep it below. But also here it is unnecessarily costly to suppress the whole curve as it leaves idle respirators. Hence, the aim during the second phase is to precisely fill the ICU capacity. Once the infection rate goes down so that the ICU constraint is no longer binding – the third phase – suppression can be lifted. The number of respirators and the time axis can essentially be thought of as a box. "Filling the box" then simply means respecting the ICU constraint while not incurring costs to leave idle resources. In the concluding remarks we further discuss how various enrichments may change this result.

Apart from the policy implication, our main contribution is methodological. We develop and show how to fully analytically solve an epidemic-economic model for the optimal suppression policy. Importantly, the suppression policy is allowed to be fully time varying. Our approach is thus clearly distinguished from a large number of recent papers (not least in economics) that analyze policies numerically (e.g., Wearing et al., 2005; Iacoviello and Liuzzi, 2008; Lee et al., 2011; Kar and Batabyal, 2011; Iacoviello and Stasio, 2013; Giamberardino and Iacoviello, 2017; Gollier, 2020; Wang, 2020; Farboodi et al, 2020; Eichenbaum et al., 2020; Alvarez et al, 2020). In order to make analytical headway we abstract from many nuances that such, numerical, papers consider including the possibility of testing (Gollier, 2020; Wang, 2020, Berger et al., 2020), the arrival of a vaccine (Zaman et al., 2008; Iacoviello and Liuzzi, 2008; Lee et al., 2011; Kar and Batabyal, 2011; Giamberardino and Iacoviello, 2017; Farboodi et al, 2020), treatment and education (e.g., Bakare et al., 2014) group heterogeneity (e.g., Shim, 2013; Sjödin, 2020), contact tracing (see, e.g., Wearing et al., 2005; McCaw and McVernon, 2007; Britton and Malmberg, 2020; and references therein), time delays (Zaman et al., 2009), network effects (e.g., Gourdin, 2011), regional breakdown (Favero et al., 2020) and individual decision making (Farboodi et al, 2020; Eichebaum et al., 2020). Our exercise is a stepping stone for considering also such aspects in future work. To our knowledge ours is the first paper to at all consider optimal policy under an ICU constraint.[2]

[2]In a rich numerical model Favero et al. (2020) analyze ways to restart the Italian economy. They do take into account that ICU capacity is limited among many other

In the epidemiology literature there exist other papers with analytical solutions for optimal policy.[3] For a literature review on the early research see Wickwire (1977). Many papers model vaccinations (Morton and Wickwire, 1974; Ledzewicz and Schättler, 2011; Hu and Zou, 2014, Laguzet and Turinici, 2015; Maurer and de Pinho, 2015), some model screening (Ainseba & Iannelli, 2012). The previous papers focusing on suppression (or quarantine) either restrict the policy (e.g., diLauro et al., 2020, see also Nowzari 2016 for a review) or abstract from the fact that increasing the suppression is costly, obviously a key aspect of any economic analysis.[4] A complementary paper to ours is by Toxvaerd (2020). He has a different research question and, indeed, a different model than us, focusing on individual choice of social distancing (a form of self-suppression). He does not analyze social optimality and abstracts from the ICU capacity which are the focus of our paper. Yet, interestingly, Toxvaerd (2020) finds that individuals may, by their own individual choices, collectively create an infection spread which graphically looks similar to the one that we show is optimal. The paper closest to ours is an elegant analysis by Kruse and Strack (2020). They also look at optimal suppression with costs which are increasing in suppression. They show existence of an optimizer for a rather general health-cost function but only solve for the optimizer in the special case where the health costs are linear in the number of currently infected.[5] This is equivalent to assuming that the total number of deaths (over time) is proportional to the total number of infected (the linearity assumption

---

things and analyze a subset of policy options (containing not only suppression) but not global optimality.

[3]There is a much larger literature studying epidemics without controls, of course, see for instance Dickison et al (2012) and Brauer and Castillo-Chavez (2011), Pastor-Satorras et al. (2015) and references therein.

[4]Many papers analytically solve for a suppression policy while respecting a budget constraint (Hansen and Day, 2011; Bolzoni et al., 2019) or a time constraint (so that the suppression cannot be too long, Morris et al 2020) but disregarding that more suppression within a time period is costlier than less suppression (Bolzoni et al., 2017; Piunovskiy et al., 2019). This is isomorphic to restricting the suppression policy to be binary since, once there is suppression within a time period, it may as well be at full force. We allow the suppression policy to take any value within a period and change in any way between time periods. Grigorieva et al. (2016) and Grigorieva and Khailov (2014) analyze an objective of minimizing the number of infectious during or at some end period, but the control bears no cost. Abakus (1973) and Behncke (2000) analyze an objective of minimizing the total (over time) number of infected (see Behncke, 2000, Section 3) but the cost of putting a person in quarantine is only taken once so is independent of the length of quarantine. Finally, Gonzales-Eiras and Niepelt (2020) analyze an S.I. model, finding, just like some of the papers above, that the optimal control is binary.

[5]They also show existence of an optimizer for when a vaccine can arrive.

implies bang-bang solutions for suppression) so it does not (directly) matter how many are infected at the same time like is the focus of our paper (in that sense, their paper is similar to Grigorieva et al., 2016, and Grigorieva and Khailov, 2014). Our contribution is thus complementary to theirs since we study a health cost which specifically captures the overwhelming of the health-care system.

2 Model

Our model setup closely follows the canonical Susceptible-Infectious-Removed model (Kermack and McKendrick 1927; see also Brauer and Castillo-Chavez, 2011, for an excellent overview). At any time $t \geqslant 0$, let $x(t)$ be the population share of individuals who at time t are susceptible to the infection, and let $y(t)$ be the population share of individuals who are infected at time t. All infected individuals are assumed to be contagious, and population shares are defined with respect to the initial population size, N. Let $\lambda(t)$ be the rate at time t of pairwise meetings between susceptible and infected, and let $q(t)$ be the probability of contagion when an infected person meets a susceptible person at time t. Write $b(t) = \lambda(t) q(t)$. Infected individuals are removed from the population at rate $\alpha > 0$.[6] This may be either because they get immune or because they die. An important assumption is that those who are infected never again become susceptible.

The population dynamic is then defined by the following simple system of ordinary differential equations:

$$\begin{cases} \dot{x}(t) = -b(t) y(t) x(t) \\ \dot{y}(t) = b(t) y(t) x(t) - \alpha y(t) \end{cases} \tag{1}$$

The initial condition is $x(0) = 1 - \varepsilon$ and $y(0) = \varepsilon$, for some $\varepsilon \in (0,1)$. That is, the infection enters the population at time zero in a population share $\varepsilon > 0$. The state space of this dynamic is $\Delta = \left\{(x,y) \in \mathbb{R}_+^2 : x + y \leqslant 1\right\}$. The only difference from the standard S.I.R. model is that the propagation coefficient $b(t)$, instead of being a constant over time, say, $b(t) = \beta > 0$ for all $t \geqslant 0$, we here allow it to vary over time.[7]

Indeed, we will view $b : \mathbb{R}_+ \to \mathbb{R}_+$ as a function in the hands of a social planner who strives to minimize the economic and social costs of shutting down parts of the economy and social life in the population,

[6] Implicitly this assumes that the duration of the infection in an individual is an i.i.d. exponentially distributed random variable with mean value $1/\alpha$.

[7] If $z(t)$ denotes the population share of removed individuals in a standard S.I.R. model, then its dynamic is $\dot{z}(t) = \alpha y(t)$, and $x(t) + y(t) + z(t) = 1$ at all times $t \geqslant 0$.

while never letting the population share of infected individuals, $y(t)$, exceed an exogenously given level γ. The latter is interpreted as the capacity of the health-care system to treat infected patients. We refer to it as the *ICU capacity* or *constraint*.[8] It is meant to capture a situation, such as under Covid-19, where if the number of simultaneously and seriously ill exceeds the number of respirators implies instant death. Not breaching the ICU capacity thus ensures that all get the best possible care.[9]

We assume that the cost of keeping $b(t)$ below its *natural*, or *unregulated level* β is a linear function of the difference, while there is no cost of moving $b(t)$ above β. The latter assumption is made to "tilt the table" against us in the subsequent analysis, where we will show that it is suboptimal to enhance the propagation of the infection even if this can be done at no cost. Formally, the cost function $\mathcal{C} : \mathcal{B} \to \mathbb{R}_+$ is defined by $\mathcal{C}(b) = \int_0^\infty [\beta - b(t)]_+ \, dt$, and the social planner faces the optimization program

$$\min_{b \in \mathcal{B}_\gamma} \mathcal{C}(b), \tag{2}$$

where $\mathcal{B}$ is the class of piecewise continuous functions $b : \mathbb{R}_+ \to \mathbb{R}_+$ that have finitely many points of discontinuity (including no discontinuity at all), and $\mathcal{B}_\gamma$, for any given $\gamma > 0$, is the subset of functions in $\mathcal{B}$ for which $y(t) \leqslant \gamma$ at all times $t \geqslant 0$.[10]

We focus on situations in which $\varepsilon < \gamma$, that is, when the initial infection level is below the ICU capacity constraint. Moreover, we assume that $\beta > \alpha$. Otherwise the population share of infected individuals does not increase from its initial value, which would imply herd immunity already from the outset, and thus the social planner's optimization program then has a trivial solution; *laissez-faire*, that is, $b(t) \equiv \beta$.

In our setting the *basic reproduction number* is $R_0 = \beta/\alpha$.

---

[8]If, for example, on average 20% of those infected need intense care and the number of ICUs is C in a population of size N, then $\gamma = 5C/N$. To allow for risk, the ICU constraint can of course also include a margin to the actual limit.

[9]In practice this of course is no guarantee against fatalities. We implicitly assume that those that pass away despite getting the best care are not within the control of the policy maker. See, e.g., Kruse and Strack (2020), Grigorieva et al. (2016) and Grigorieva and Khailov (2014) for models where the objective is to minimize the number of infected.

[10]To be more precise, we require that there is a finite set $T \subset \mathbb{R}_+$ such that the function $b : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous at all other points, and that it is everywhere left-continuous and has a right limit. We also require that b is positive except on at most finitely many connected components.

3 Analysis

We first comment on the set of policies allowed by the optimization program. For functions $b \in \mathcal{B}$, it can be shown that (1) defines a unique solution trajectory through any given state $(x(t_0), y(t_0)) \in \Delta$ and time $t_0 \geqslant 0$.[11] Trivially all constant functions $b : \mathbb{R}_+ \to \mathbb{R}_+$, with $b(t) = \delta$ for some $\delta > 0$, belong to $\mathcal{B}$. However, they do not all belong to $\mathcal{B}_\gamma$, i.e., they may violate the ICU constraint. It is easy to show that such constant policies belong to $\mathcal{B}_\gamma$ if δ is sufficiently low, for any given $\gamma > 0$. Thus, to choose δ as high as possible, while keeping $y(t) \leqslant \gamma$ for all $t \geqslant 0$, is a feasible policy (belongs to $\mathcal{B}_\gamma$), and can be called *flattening the curve*. However, such a policy incurs an infinite cost if $\delta < \beta$, since it lasts forever. An alternative feasible control function, with finite cost, is to only temporarily keep $b(t)$ at a constant level $\delta < \beta$, where δ is such that $y(t) \leqslant \gamma$ for all $t \geqslant 0$. However, as will be shown below, also such "temporary constant shut down" policies are suboptimal. Before turning to the formal statement of our main result, we analyze some general properties of the dynamic induced by (1).

3.1 The dynamic

Some well-known properties of the solutions to standard S.I.R. models hold also here (see Brauer and Castillo-Chavez, 2011). A key such property is that the population share of susceptible individuals, $x(t)$, is non-increasing over time t. Roughly speaking, this follows from the first equation in (1), since $b(t)$ is always non-negative and $y(t)$ is positive at all times $t \geqslant 0$. Being bounded from below by zero, $x(t)$ necessarily has a limit value as $t \to \infty$, which we denote x_∞. According to (1), also the sum $y(t) + x(t)$ is strictly decreasing over time t, and hence also this sum has a limit value, $x_\infty + y_\infty$. By standard arguments, it is easily verified that this implies that $y_\infty = 0$.[12] In other words, in the very long run, the population share of infected individuals tends to zero. Denoting by $z_\infty = \lim_{t\to\infty} z(t)$ the total population share of removed individuals during the whole epidemic, we thus have $z_\infty = 1 - x_\infty$, and Nz_∞ is approximately (for large N), the total number of infected individuals during the epidemic.

Let us now consider the solution to (1) through any given state $(x(t_0), y(t_0)) \in \Delta$ and time $t_0 \geqslant 0$, where $0 < x(t_0) < 1$ and $0 < y(t_0) < 1$. Dividing both sides of the first equation in (1) by $x(t) > 0$

[11] See Appendix for a proof. For this class of functions b, the time derivatives in (1) represent left derivatives. The solution trajectories $(x(t), y(t))_{t>0}$ are then uniquely determined and are continuous in t.

[12] If $y_\infty > 0$, then $x(t) + y(t) \to -\infty$.

and integrating, one obtains

$$\ln(x(t)) = \ln(x(t_0)) - \int_{t_0}^{t} b(s)\, y(s)\, ds \quad \forall t \geqslant t_0.$$

Moreover, integrating the sum of the two equations in (1), we obtain

$$x(t) + y(t) = x(t_0) + y(t_0) - \alpha \int_{t_0}^{t} y(s)\, ds \quad \forall t \geqslant t_0.$$

3.1.1 Constant policy

In particular, if $b(t) = \delta > 0$ for all $t \geqslant t_0$, for some $\delta > 0$, then for all $t \geqslant t_0$:

$$\ln \frac{x(t)}{x(t_0)} = -\delta \int_{t_0}^{t} y(s)\, ds = \frac{\delta}{\alpha} \left[x(t) - x(t_0) + y(t) - y(t_0) \right],$$

or

$$y(t) = y(t_0) + \frac{\alpha}{\delta} \ln \left(\frac{x(t)}{x(t_0)} \right) - x(t) + x(t_0) \quad \forall t \geqslant t_0. \tag{3}$$

This equation is well-known for S.I.R. models. Moreover, (3) implies that $(x(t), y(t)) \to (x_\infty, 0) \in \Delta$, where x_∞ by continuity solves (3) for $y(t) = 0$, so

$$x_\infty = \frac{\alpha}{\delta} \ln \frac{x_\infty}{x(t_0)} + x(t_0) + y(t_0). \tag{4}$$

Since $x(t)$ is strictly decreasing, $x_\infty < x(t_0)$. It is easily verified that the fixed-point equation (4) has a unique solution $x_\infty \in (0, x(t_0))$.[13] Equation (4) will later be used to calculate the population share that were infected during the whole epidemic under the optimal suppression policy.

The value

$$\hat{y} = \sup_{t \geqslant 0} y(t)$$

is the peak infection level. It obtains when $\dot{y}(t) = 0$, or, equivalently (by (1)), when $x(t) = \alpha/\delta$. From (3) we obtain

$$\hat{y} = 1 + \frac{\alpha}{\delta} \ln \left(\frac{\alpha}{\delta(1-\varepsilon)} \right) - \frac{\alpha}{\delta}. \tag{5}$$

The right-hand side is a strictly decreasing function of the ratio α/δ. Thus, the infection peak is higher the larger δ is and the smaller α is.

[13] The right-hand side of (4) is a continuous and strictly increasing functions $f : (0, x(t_0)) \to \mathbb{R}$ of x_∞. Moreover, $f(x_\infty) \to -\infty$ as $x_\infty \downarrow 0$ and $f(x(t_0)) = x(t_0) + y(t_0) > x(t_0)$, $f' > 0$ and $f'' < 0$, so there exists a unique fixed point in $(0, x(t_0))$.

Once the population share $x(t)$ of susceptible individuals has fallen below the level α/δ, achieved precisely when $y(t) = \hat{y}$, flock immunity is obtained; the population share $y(t)$ of infected individuals falls. In particular, the limit state as $t \to \infty$ is Lyapunov stable. That is, there is no risk of a second infection wave, since after any small perturbation of the limit population state $(x_\infty, 0) \in \Delta$, obtained by exogenously inserting a small population share of infected individuals, the population share of infected individuals will fall gradually back towards zero, while the population share of susceptible individuals gradually moves towards a somewhat lower, new limit value.

Equation (5) is particularly relevant for the case when $\delta = \beta$, that is, under laissez-faire. Because if the peak of the infection wave then does not exceed the ICU capacity constraint, that is, if

$$1 + \frac{\alpha}{\beta} \ln\left(\frac{\alpha}{\beta(1-\varepsilon)}\right) - \frac{\alpha}{\beta} \leqslant \gamma, \tag{6}$$

then laissez-faire is optimal; $b^*(t) \equiv \beta$ solves (2) at no cost. Since $R_0 = \beta/\alpha$, we note that inequality (6) holds if and only if R_0 is small enough. When the initial infection constitutes an infinitesimal population fraction ($\varepsilon \to 0$), (6) can be written as $R_0/(1 + \ln(R_0)) \leqslant 1/(1-\gamma)$. For $\gamma \geqslant 1$, the condition is trivially met (then the health sector has the capacity to receive the whole population). For $\gamma < 1$, equation (6) gives an upper bound on R_0. But if the peak is above the ICU constraint, regulation has to be implemented. This is the topic of the next subsection.

3.2 Optimization

To the best of our knowledge, the optimization program (2) has not been analyzed before. We summarize below our main result, which treats all cases when laissez-faire is suboptimal. If (6) does not hold, which we henceforth assume, then the solution orbit (3) under laissez-faire intersects the capacity constraint $y(t) = \gamma$ twice. Let $\tau_1 > 0$ be the first such time and let $x(\tau_1)$ be the population share of susceptible individuals at that time. Then[14]

$$\tau_1 = \min\left\{t \geqslant 0 : x(t) = 1 - \gamma + \frac{\alpha}{\beta} \ln\left(\frac{x(t)}{1-\varepsilon}\right)\right\} \tag{7}$$

where $x(t)$ is solved for according to (1) when $b(t) \equiv \beta$, and $x(\tau_1)$ is the larger of the two solutions to the associated fixed-point equation in x,

$$x = 1 - \gamma + \frac{\alpha}{\beta} \ln\left(\frac{x}{1-\varepsilon}\right). \tag{8}$$

[14] An analytic expression for τ_1 is provided at the end of the appendix.

We note that $x(\tau_1) > \alpha/\beta$.[15] Let

$$\tau_2 = \tau_1 + \frac{1}{\alpha\gamma}\left(x(\tau_1) - \frac{\alpha}{\beta}\right)$$

Theorem 1 *Suppose that $\varepsilon < \gamma$, $\alpha < \beta$ and (6) does not hold. There exists a solution to program (2), one of which is the policy $b^* \in \mathcal{B}_\gamma$ defined by*

$$b^*(t) = \begin{cases} \beta & \text{for } t \leqslant \tau_1 \\ \frac{\beta}{1+\beta\gamma(\tau_2-t)} & \text{for } \tau_1 < t \leqslant \tau_2 \\ \beta & \text{for } t > \tau_2 \end{cases}$$

Every optimal policy $b \in \mathcal{B}_\gamma$ agrees with b^* on $[0, \tau_2]$ and satisfies $b(t) \geqslant \beta$ for all $t > \tau_2$.

We note that the optimal policy is laissez-faire both before time τ_1 and after time τ_2. We also note that the optimal policy has exactly one discontinuity, namely, a sudden shut-down of society at time τ_1; then $b^*(t)$ falls from $b^*(\tau_1) = \beta$ to

$$\lim_{t\downarrow\tau_1} b^*(t) = \frac{\beta}{1+\beta\gamma(\tau_2-\tau_1)} = \frac{\alpha}{x(\tau_1)}.$$

From time τ_1 on, $b^*(t)$ rises continuously until time τ_2, at which point $b^*(t)$ reaches the level β. In the mean-time, between times τ_1 and τ_2, the population share $y(t)$ of infected individual remains constant, at the capacity level γ, while the population share $x(t)$ falls linearly over time to the level α/β, reached at time τ_2.

One obtains the following expression for the minimized cost (see the proof of Theorem 1 in the appendix for the details):

$$\mathcal{C}(b^*) = \frac{1}{\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1 + \frac{\beta}{\alpha} - \ln(1-\varepsilon)\right) - \frac{\beta}{\alpha} \tag{9}$$

From this expression it is clear that the lower the ICU capacity (γ) is the more costly implementing the policy will be. Similarly, it can be shown that the cost is increasing in the basic level of infection spread (β). Moreover, since in this model $R_0 = \beta/\alpha$, the expression for the minimal cost can be rewritten entirely in terms of three fundamental parameters, the ICU capacity constraint, γ, the initial infection size, ε, and R_0. For small $\varepsilon > 0$, the expression for $\mathcal{C}(b^*)$, the total cost of suppression, when minimized, takes the simple form

[15] This follows from the observation that the derivative of the right-hand side of (8) is less than unity at $x = x(\tau_1)$.

$$\mathcal{C}(b^*) = \frac{1}{\gamma}(R_0 - 1 - \ln(R_0)) - R_0$$

This is a continuous function of R_0, taking the value zero when R_0 is such that the infection peak, $\hat{y}$, in the absence of regulation, equals γ. For R_0 above this critical level (see equation (6)), $\mathcal{C}(b^*)$ is strictly increasing in R_0.

The proof of Theorem 1 is mathematically involved, and is given in the Appendix. It uses measure theory and views the minimization as taking place in phase space (much in line with equation (3)). For a rich enough measure space, existence of a solution to (2) is obtained by topological arguments. Invoking the Picard-Lindelöf theorem, it is shown that the differential equations (1) indeed uniquely define solutions. The next step in the proof is to show that the minimizer measure is absolutely continuous (with respect to Lebesgue measure). This brings us back to functions $b \in \mathcal{B}$, now viewed as transforms of Radon-Nikodyn derivatives of the measures in question. The rest of the proof consists in verifying that the above function, b^*, indeed corresponds to an optimal measure, and that it is unique in the sense stated. In particular, one needs to show that it is neither worthwhile to slow down nor speed up the infection in its early phase (before time τ_1).

The result is illustrated in Figure 1, where the solid kinked curve is the solution orbit induced by (1) under the optimal control function b^*. The dotted curve is the infection orbit under laissez-faire ($b(t) \equiv \beta$).

Figure 2 depicts the optimal policy as a function of time in comparison to a strategy of flattening the curve, here assumed to take the form: keep $b(t)$ at the level $\delta < \beta$ for which $\hat{y} = \gamma$ (see (5)) until the infection wave has passed its peak, and then return to laissez-faire, $b(t) = \beta$ (outside the time range of the figure). The upper panel shows the dynamics of infections and the lower panel the policy $b(t)$.

As can be seen, and as expressed by the theorem, the optimal policy is characterized by leaving the spread unregulated initially, then a sudden shut-down of society (a discontinuity at τ_1), followed by gradual (continuous) opening of society, until τ_2, from which onwards the propagation is not regulated. The time axis and the ICU constraint create a square – a box. The economic logic behind the optimal policy is essentially to ensure that we do not close down society while leaving idle ICU resources – "filling the box". This implies that whenever the natural spread is not threatening the constraint, it should go unregulated. This holds in the early phase when only few have been infected, and in the last phase, when many have already been infected but most

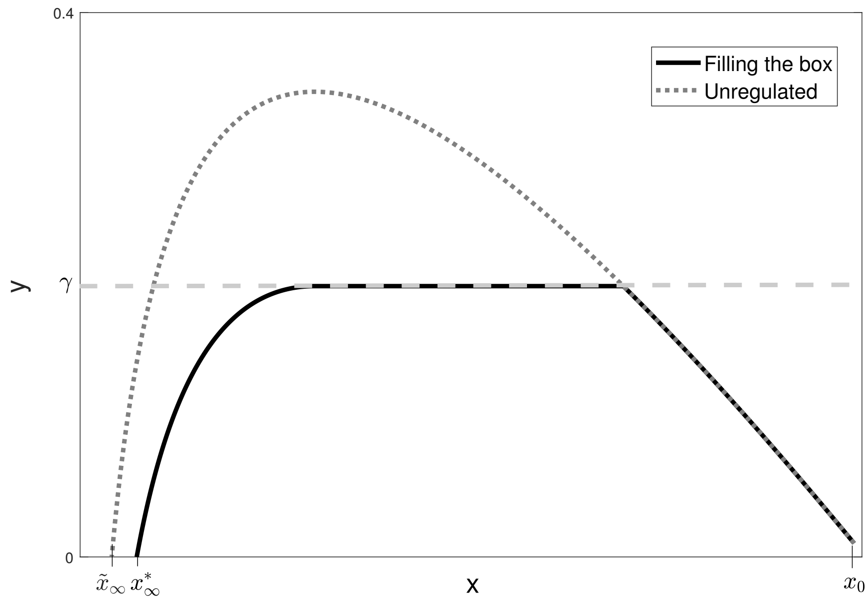


Figure 1: The solution orbit (solid) in the (x, y)-plane under the optimal policy b^*, and the solution orbit under unregulated spread (dotted) . Parameter values used: $\alpha = 0.3$, $\beta = 1$, $\gamma = 0.2$, and $\varepsilon = 0.01$.

of them also have recovered. It is only when the epidemic may breach the ICU constraint – the second phase – that it should be regulated. In order to ensure that the constraint is not breached, strong suppression has to be imposed when reaching the ICU constraint – a sudden shutdown. The reason for the abruptness of this policy (the discontinuity) is that the natural infection is progressing very quickly at that point, so a sudden break is needed to stop it. This can be seen in the lower panel by the drop at τ_1. After that, b^* gradually increases. The reason for this is that the suppression only needs to keep the infection just below the ICU constraint. Then since over time the number of susceptible (x) is falling, the number of infected (y) is held endogenously constant and since new infections depends on their product $(b(t)\, y(t)\, x(t))$ it follows that b^* is increasing during the second phase. The policy as a function of the population share simply is $b^*(t) = \alpha / x(t)$, i.e., recovery (α) determines what share of the susceptible population that can be allowed to be infected. A few further remarks about optimal policies are now in place.

It may be noted that the optimal policy never attempts to fully eradicate the spread. In our model, like in all standard S.I.R. models, this is since y only asymptotically goes to zero. Hence, full eradication (a form of extreme corner solution) would imply locking down forever.

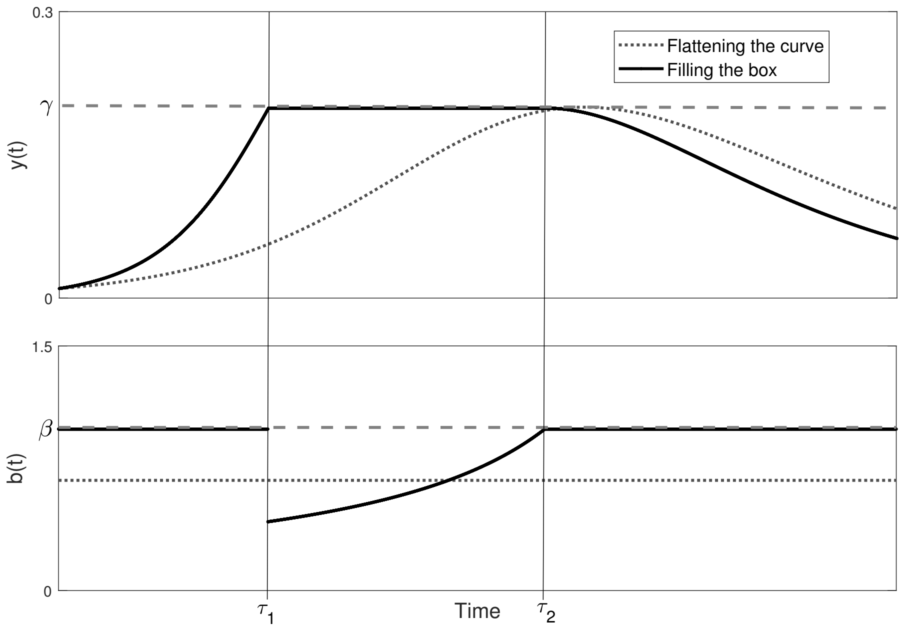


Figure 2: Upper panel: The share of infected over time under the optimal policy (solid) and flattening the curve (dotted). The horizontal dashed line represents the ICU constraint γ. Lower panel: Optimal suppression (solid) and flattening-the-curve suppression (dotted). The horizontal dashed line represents the baseline spread β. Parameter values used: $\alpha = 0.3$, $\beta = 1$, $\gamma = 0.2$, and $\varepsilon = 0.01$.

We discuss this further in the conclusions.

Furthermore, the optimal policy is unique during the first and the second phase but not during the third. The uniqueness during the first phase is not obvious. To see this note that here there is no reason to hold back the spread. Then, given that $b(t) > \beta$ has been assumed to be costless, why would accelerating the spread not be optimal? The answer is that, if one does that, then the ICU capacity is reached at a high speed of infection hence it would require hitting the breaks very hard. This is not optimal. The multiplicity of optimal strategies during the third phase is due to the same assumption – acceleration is free. Hence, not only laissez-faire is optimal, but also acceleration (of which one can think as stimulus for economic interaction). The acceleration cannot be too fast, however, as it may then breach the ICU constraint. Naturally, should we assume that there is a cost of acceleration (even the slightest) this multiplicity disappears and a unique optimal policy emerges also in the third phase – laissez-faire.

Compared with the optimal policy, "flattening the curve" implies costs that lead to idle resources. This is visible in the upper panel of Figure 2, where costs are incurred without the spread posing a threat to the health system – both before and after the peak, suppression costs are incurred for no reason. The additional cost of flattening the curve (instead of filling the box) can be seen in the lower panel by comparing the rectangle between β and the dashed-dotted line on the one hand with the area between β and the solid line on the other. It is potentially very large, in particular if the policy maker continues to flatten the curve long after the peak.

3.3 Infection in the long run

Our results allow analysis of the final number that were ever infected z_∞.[16] To conserve on space, we do this under the assumption that the initial number of infected is very small: $\varepsilon \to 0$.[17] Recall that $z_\infty = 1 - x_\infty$. Equation (4) is very useful as it implicitly determines x_∞ for any given value of susceptible $x(t_0)$ and infected $y(t_0)$ given some constant policy δ. Denote by x_∞^* the limit share of susceptible individuals under the optimal policy and by x_∞^{LF} and x_∞^{FTC} the equivalent under laissez-faire and flattening the curve respectively.

Under the optimal policy b^*, the solution orbit at time τ_2 passes through the population state $x(\tau_2) = \alpha/\beta$, $y(\tau_2) = \gamma$, after which the spread is unregulated, i.e. $b^*(t) = \beta, \forall t \geqslant \tau_2$. Using $x(t_0) = x(\tau_2)$, $y(t_0) = y(\tau_2)$ and $\delta = \beta$ in (4), one obtains

$$x_\infty^* = \frac{\alpha}{\beta} \ln x_\infty^* + \frac{\alpha}{\beta} - \frac{\alpha}{\beta} \ln\left(\frac{\alpha}{\beta}\right) + \gamma. \tag{10}$$

Under laissez-faire, $b(t) = \beta, \forall t \geqslant 0$, and letting $x(t_0) \to 1$ and $y(t_0) \to 0$ in (4), one obtains

$$x_\infty^{LF} = \frac{\alpha}{\beta} \ln x_\infty^{LF} + 1. \tag{11}$$

Under the maintained assumption that $\gamma < \hat{y}$ under laissez-faire, that is,

$$\frac{\alpha}{\beta} - \frac{\alpha}{\beta} \ln\left(\frac{\alpha}{\beta}\right) + \gamma < 1$$

the RHS of (10) lies below the RHS of (11), and hence $x_\infty^* > x_\infty^{LF}$ as can also be seen in Figure 1.

[16]Recall that our objective function does not put any weight on the total number of infected, so the results presented here essentially show an "unintended" side effect of the social planner's choice. See Kruse and Strack (2020) for optimal policy focusing on the total number of infected.

[17]The qualitative results hold also when $\varepsilon > 0$.

Finally we consider flattening the curve. However, since in its basic form it is infinitely costly in terms of eternal suppression, we consider a somewhat improved version of it, essentially prescribing to release regulation at the precise moment where doing so avoids a second wave. More precisely, let $b(t) = \tilde{\delta}$ (where $\tilde{\delta}$ is determined to precisely respect the ICU constraint) until the time τ_3 when $x(t)$ has fallen to the value α/β, since then "herd immunity" holds under laissez-faire. At τ_3, $y(t)$ has fallen to

$$y(\tau_3) = \frac{\alpha}{\tilde{\delta}} \ln\left(\frac{\alpha}{\beta}\right) - \frac{\alpha}{\beta} + 1$$

which is less than γ. After τ_3 regulation ceases, i.e., $b(t) = \beta$, $\forall t \geqslant \tau_3$. Using (4) with these values one obtains

$$x_\infty^{FTC} = \frac{\alpha}{\beta} \ln x_\infty^{FTC} + 1 - \left(\frac{\alpha}{\beta} - \frac{\alpha}{\tilde{\delta}}\right) \ln \frac{\alpha}{\beta}. \tag{12}$$

Using the definition for $\tilde{\delta}$ in (5) with $\hat{y} = \gamma$ and $\varepsilon \to 0$ one gets that $x_\infty^{FTC} > x_\infty^*$. Hence:

$$z_\infty^{FTC} < z_\infty^* < z_\infty^{LF}.$$

That is, the optimal policy leads to a smaller share of ultimately infected than under laissez-faire but a larger share than under flattening the curve.

4 Concluding discussion

This paper provides an *analytical* answer to the question: What is the optimal time-varying suppression policy to avoid a collapsed health-care system when suppression is costly? We have shown that the general recommendation of "flattening the curve" is suboptimal. Instead the optimal policy essentially prescribes "filling the box". In a first phase, the spread is unregulated until the number of infected reaches the ICU constraint. A second phase begins by discontinuous suppression followed by gradual relaxation until, in a third phase, the spread is left unregulated again.

A contribution of the paper is methodological, showing how to obtain a fully analytical solution to an S.I.R. model with economic costs which are increasing in suppression. Our paper also contributes to the policy debate on how to fight an epidemic. We discuss here the robustness of this policy to various perturbations.

In our model (like in all standard S.I.R models) attempting for a complete wipe out of the spread is never optimal. In fact, it is not even feasible. Technically this is since the number of infected only asymptotically goes to zero hence a wipe out would require suppression to be in place for the infinite future. Naturally, if there exists a maximal duration within which all recover, then it could be optimal to go for a full wipe out right away. What makes that hard in many cases is the practical difficulty in identifying all infected and since in practice the cost of full suppression ($b = 0$) is virtually infinite – after all, people need to access food and medical services. Furthermore, unless countries are closed more or less indefinitely, a very costly prospect, under a pandemic such as Covid-19 one risks importing new cases.

In the model we have assumed that the only medical harm is if violating the ICU constraint. If two assumptions were added – medical harm from the aggregate number of infected and existence of a vaccine within a reasonable time frame – then suppressing the spread more than what our policy prescribes could be optimal. Likewise if the number of simultaneously infected would cause harm more gradually. This, however, does not seem to be case for Covid-19 where the bottleneck in most countries is the number of respirators – respecting the ICU constraint is the main issue. Another factor that *could* suggest early suppression is if the ICU constraint can be expanded (for Covid-19 equivalent to an increased number of respirators or development of a cure or improved treatment). Another possibility is that one learns about the parameters. However, for that to motivate regulation early, one has to assume that the suppression itself does not distort the signal. Finally, if the cost of suppression were convex, in particular so that small suppression is very cheap, then that would motivate some suppression early on. However, it would still most probably be optimal to discontinuously increase suppression to ensure that the number of infected is constant and does not surpass the ICU constraint.[18]

We hope our analysis provides a stepping stone for future analyses of more complex and richer models.

References

[1] Brauer and Castillo-Chavez (2011): *Mathematical Models in Population Biology and Epidemiology.* Second Edition, Springer Verlag.∪

[2] Abakus, A., 1973. An optimal isolation policy for an epidemic. *Journal of Applied Probability*, 10(2), pp.247-262.

[3] Ainseba, B. and Iannelli, M., 2012. Optimal screening in structured SIR epidemics. *Mathematical Modelling of Natural Phenomena*, 7(3), pp.12-27.

[4] Alvarez, F.E., Argente, D. and Lippi, F., 2020. A simple planning problem for covid-19 lockdown (No. w26981). National Bureau of

[18] See Appendix for a discussion of how the cost function can be generalized.

Economic Research.

[5] Bakare, E.A., Nwagwo, A. and Danso-Addo, E., 2014. Optimal control analysis of an SIR epidemic model with constant recruitment. *International Journal of Applied Mathematics Research*, 3(3), p.273.

[6] Behncke, H., 2000. Optimal control of deterministic epidemics. *Optimal control applications and methods*, 21(6), pp.269-285.

[7] Berger, D.W., Herkenhoff, K.F. and Mongey, S., 2020. An seir infectious disease model with testing and conditional quarantine (No. w26901). National Bureau of Economic Research.

[8] Bolzoni, L., Bonacini, E., Soresina, C. and Groppi, M., 2017. Time-optimal control strategies in SIR epidemic models. *Mathematical biosciences*, 292, pp.86-96.

[9] Bolzoni, L., Bonacini, E., Della Marca, R. and Groppi, M., 2019. Optimal control of epidemic size and duration with limited resources. *Mathematical biosciences*, 315, p.108232.

[10] Branswell. H., 2020 Why 'flattening the curve' may be the world's best bet to slow the coronavirus. In: STAT News (Mar. 2020). url: https://www.statnews.com/2020/03/11/flattening-curve-coronavirus/.

[11] Francis Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264 of *Graduate Texts in Mathematics*. Springer, London, 2013.

[12] Di Giamberardino, P. and Iacoviello, D., 2017. Optimal control of SIR epidemic model with state dependent switching cost index. *Biomedical signal processing and control*, 31, pp.377-380.

[13] Di Lauro, F., István Z.K., and Mille, J., 2020. The timing of one-shot interventions for epidemic control. medRxiv (2020).

[14] Dickison, M., Havlin, S. and Stanley, H.E., 2012. Epidemics on interconnected networks. *Physical Review* E, 85(6), p.066109.

[15] Eichenbaum, M.S., Rebelo, S. and Trabandt, M., 2020. The macroeconomics of epidemics (No. w26882). National Bureau of Economic Research.

[16] Farboodi, M., Jarosch, G. and Shimer, R., 2020. Internal and External Effects of Social Distancing in a Pandemic. University of Chicago, Becker Friedman Institute for Economics Working Paper, (2020-47).

[17] Favero, C.A., Ichino, A. and Rustichini, A., 2020. Restarting the economy while saving lives under Covid-19. CEPR Discussion Paper No. DP14664.

[18] Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunuba Perez, Z.U.L.M.A., Cuomo-Dannenburg, G. and Dighe, A., 2020. Report

9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.

[19] Gollier, C., 2020, Cost-benefit analysis of deconfinement strategies, mimeo, Toulouse School of Economics.

[20] Gonzalez-Eiras, M. and Niepelt, D., 2020. On the Optimal" Lockdown" During an Epidemic (No. 20.01).

[21] Gourdin, E., Omic, J. and Van Mieghem, P., 2011, October. Optimization of network protection against virus spread. In 2011 8th International Workshop on the Design of Reliable Communication Networks (DRCN) (pp. 86-93). IEEE.

[22] Grigorieva, E.V. and Khailov, E.N., 2014. Optimal vaccination, treatment, and preventive campaigns in regard to the SIR epidemic model. *Mathematical Modelling of Natural Phenomena*, 9(4), pp.105-121.

[23] Grigorieva, E.V., Khailov, E.N. and Korobeinikov, A., 2016. Optimal control for a SIR epidemic model with nonlinear incidence rate. *Mathematical Modelling of Natural Phenomena*, 11(4), pp.89-104.

[24] Hansen, E. and Day, T., 2011. Optimal control of epidemics with limited resources. *Journal of mathematical biology*, 62(3), pp.423-451.

[25] Hu, Q. and Zou, X., 2013. Optimal vaccination strategies for an influenza epidemic model. *Journal of Biological Systems*, 21(04), p.1340006.

[26] Iacoviello, D. and Liuzzi, G., 2008, June. Optimal control for SIR epidemic model: A two treatments strategy. In 2008 16th Mediterranean Conference on Control and Automation (pp. 842-847). IEEE.

[27] Iacoviello, D. and Stasio, N., 2013. Optimal control for SIRC epidemic outbreak. *Computer methods and programs in biomedicine*, 110(3), pp.333-342.

[28] JHU 2020, "Has the curve flattened?" https://coronavirus.jhu.edu/data/new-cases

[29] Kar, T.K. and Batabyal, A., 2011. Stability analysis and optimal control of an SIR epidemic model with vaccination. *Biosystems*, 104(2-3), pp.127-135.

[30] Kermack, W.O. and McKendrick, A.G., 1927. A contribution to the mathematical theory of epidemics. Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character, 115(772), pp.700-721.

[31] Kruse, T. and Strack, P., 2020. Optimal Control of an Epidemic through Social Distancing. SSRN WP

[32] Laguzet, L. and Turinici, G., 2015. Global optimal vaccination in

the SIR model: properties of the value function and application to cost-effectiveness analysis. *Mathematical biosciences*, 263, pp.180-197.

[33] Ledzewicz, U. and Schättler, H., 2011. On optimal singular controls for a general SIR-model with vaccination and treatment. *Discrete and continuous dynamical systems*, 2, pp.981-990.

[34] Lee, S., Chowell, G. and Castillo-Chávez, C., 2010. Optimal control for pandemic influenza: the role of limited antiviral treatment and isolation. *Journal of Theoretical Biology*, 265(2), pp.136-150.

[35] Daniel Liberzon. *Calculus of variations and optimal control theory.* Princeton University Press, Princeton, NJ, 2012. A concise introduction.

[36] Malmberg, H. and Britton, T., 2020. Inflow restrictions can prevent epidemics when contact tracing efforts are effective but have limited capacity. medRxiv.

[37] Maurer, H. and De Pinho, M.D.R., 2014. Optimal control of epidemiological seir models with L1-objectives and control-state constraints.

[38] McCaw, J.M. and McVernon, J., 2007. Prophylaxis or treatment? Optimal use of an antiviral stockpile during an influenza pandemic. *Mathematical biosciences*, 209(2), pp.336-360.

[39] Morris, D.H., Rossine, F.W., Plotkin, J.B. and Levin, S.A., 2020. Optimal, near-optimal, and robust epidemic control. arXiv preprint arXiv:2004.02209.

[40] Morton, R. and Wickwire, K.H., 1974. On the optimal control of a deterministic epidemic. *Advances in Applied Probability*, 6(4), pp.622-635.

[41] Nowzari, C., Preciado, V.M. and Pappas, G.J., 2016. Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems Magazine,* 36(1), pp.26-46.

[42] Pastor-Satorras, R., Castellano, C., Van Mieghem, P. and Vespignani, A., 2015. Epidemic processes in complex networks. *Reviews of modern physics*, 87(3), p.925.

[43] Piunovskiy, A., Plakhov, A. and Tumanov, M., 2019. Optimal impulse control of a SIR epidemic. *Optimal Control Applications and Methods*.

[44] Pueyo, T., 2020. Coronavirus: The Hammer and the Dance. https://medium.com/@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56

[45] Walter Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[46] Schneider, C.M., Mihaljev, T., Havlin, S. and Herrmann, H.J.,

2011. Suppressing epidemics with a limited amount of immunization units. *Physical Review E*, 84(6), p.061911.
[47] Shim, E., 2013. Optimal strategies of social distancing and vaccination against seasonal influenza. *Mathematical Biosciences & Engineering*, 10(5&6), pp.1615-1634.
[48] Sjödin, H., Johansson, A.F., Brännström, Å., Farooq, Z., Kriit, H.K., Wilder-Smith, A., Åström, C., Thunberg, J. and Rocklöv, J., 2020. Covid-19 health care demand and mortality in Sweden in response to non-pharmaceutical (NPIs) mitigation and suppression scenarios. MedRxiv https://doi. org/10.1101/2020.03, 20.
[49] The Sun, 2020. "'WE'LL BOOM AGAIN' Donald Trump says US coronavirus curve is 'FLATTENING' after almost 2,000 die in just 24 hours in deadliest day" https://www.the-sun.com/news/655595/trump-coronavirus-flattening-die-24-hours-boom/
[50] Time 2020, The U.S. Has Flattened the Curve. Next Up Is 'Squashing' It — and That's Not Going Well https://time.com/5827156/squashing-squash-curve-coronavirus-covid19/
[51] Toxvaerd, F.M.O., 2020. Equilibrium social distancing, Cambridge-INET WP No: 2020/08
[52] Wang, Y., 2020. An Analytical SIR model of Epidemics and A Sustainable Suppression Policy: Testing. Available at SSRN 3573979.
[53] Wearing, H.J., Rohani, P. and Keeling, M.J., 2005. Appropriate models for the management of infectious diseases. *PLoS medicine*, 2(7).
[54] Wickwire, K., 1977. Mathematical models for the control of pests and infectious diseases: a survey. *Theoretical population biology*, 11(2), pp.182-238.
[55] Zaman, G., Kang, Y.H. and Jung, I.H., 2008. Stability analysis and optimal vaccination of an SIR epidemic model. *BioSystems*, 93(3), pp.240-249.
[56] Zaman, G., Kang, Y.H. and Jung, I.H., 2009. Optimal treatment of an SIR epidemic model with time delay. *BioSystems*, 98(1), pp.43-50.

5 Appendix

This appendix provides a proof of Theorem 1.

Our strategy consists of the following steps:

- The optimization problem is written in the phase space $\triangle$.
- The new formulation admits a natural extension on a signed-measure space.
- Topological properties of this measure space and of the functional deduced from $\mathcal{C}$ imply the existence of a global minimizer.
- A priori a global minimizer is a general signed measure, but it turns out to be an absolutely continuous, bringing us back to a functional setting.
- Calculus of variation arguments show that the minimizer is uniquely determined until the time when x reaches the level α/β, and this leads to Theorem 1.

To the best of our knowledge, there is no such result in calculus of variations or optimal control theory (see the books of Clarke [11] or Liberzon [35]). We therefore give a direct and self-contained proof (only requiring a first knowledge of measure theory, as can be found e.g. in Rudin [45]). For the optimization problem at hand, our extension to measure spaces seems natural, and we believe it is original. The Euler-Lagrange equations will not be satisfied and the constraints will play a more important role. This is related to the fact that we consider a cost of suppression that is linear in downwards deviations, and zero for upwards deviations. If one is interested in more general costs of the form

$$\widetilde{\mathcal{C}}(b) := \int_0^{+\infty} F(\beta - b(t))\, dt \tag{13}$$

where the mapping $F : \mathbb{R} \to \mathbb{R}_+$ is e.g. a strictly convex function attaining its minimum at 0, then the Euler-Lagrange equations admit solutions leading to optimal policies different from b^* (but b^* remains a minimizer for certain functionals F, see Remark 9 at the end of this appendix). When F is close to the mapping $(\cdot)_+$ considered here, for instance if F is given by

$$\forall\ x \in \mathbb{R}, \qquad F(x) := \begin{cases} x^{1+\epsilon} & \text{, when } x \in \mathbb{R}_+ \\ \epsilon|x|^{1+\epsilon} & \text{, when } x \in \mathbb{R}_- \end{cases}$$

where $\epsilon > 0$ is small, we expect that the corresponding solution will be close to $b*$. In particular a jump will still occur. We plan to investigate more precisely this situation in future work.

Let us now move toward the proof of Theorem 1 according to the above strategy. We assume in the sequel that we are in the "interesting" case where $y_0 = \varepsilon < \gamma$ and where the *laissez-faire* policy $b \equiv \beta$ leads y to take values strictly large than γ, that is, (6) does not hold. This hypothesis will be referred to as the *underlying assumption*.

5.1 The phase space $\triangle$

We begin by rewriting the constrained control problem of minimizing $\mathcal{C}$ on $\mathcal{B}_\gamma$ as a optimization problem in the associated phase space $\triangle$.

Let us be more precise: $\mathcal{B}$ is the set of piecewise continuous functions $b : \mathbb{R}_+ \to \mathbb{R}_+$ with a finite number of discontinuities and such that $\{t \geqslant 0 : b(t) = 0\}$ has a finite number of connected components. Let $(x_0, y_0) \in \triangle$, with $y_0 \in (0,1)$, as well as $b \coloneqq (b(t))_{t\geqslant 0} \in \mathcal{B}$ be given and consider $(x,y) \coloneqq (x(t), y(t))_{t\geqslant 0}$ the maximal (over time) solution of the S.I.R. ODE (1) starting from $(x(0), y(0)) = (x_0, y_0)$ (this is a slight generalization of the setting of Theorem 1, where $x_0 = 1 - y_0$, the important hypothesis is that the underlying assumption holds). The existence and uniqueness of this solution is a consequence of the Picard-Lindelöf or Cauchy-Lipschitz theorem, extended to a time-dependent vector field that is left continuous with right limits (instead of continuous). The important fact being that the r.h.s. of (1) is locally Lipschitz with respect to $(x(t), y(t))$. By $\mathcal{B}_\gamma$, we designate the set of $b \in \mathcal{B}$ such that y always remains below γ.

We begin with a simple observation.

Lemma 1 *When $y_0 \in (0,1)$, the solution (x,y) is defined for all times $t \geqslant 0$ and $(x(t), y(t)) \in (0,1)^2 \cap \triangle$.*

Proof

As already mentioned, the solution (x,y) is locally unique. From this it follows that if x reaches 0, then it will stay there forever afterward, since the r.h.s. of (1) is zero if $x(t) = 0$. Similarly, if y reaches 0, then it will stay there forever afterward. It follows that x and y will stay non-negative. As a consequence, x is non-increasing and thus will stay below x_0 and never hit 1. From the identity

$$\dot{x}(t) + \dot{y}(t) = -\alpha y(t) \ \leqslant \ 0$$

we deduce that $x + y$ will stay below $y_0 + x_0 \leqslant 1$ and in particular y will stay below 1. The first equation of (1) then implies that

$$\dot{x}(t) \geqslant -b(t)x(t)$$

and Gronwall lemma shows

$$x(t) \geqslant x_0 \exp\left(-\int_0^t b(s)\, ds\right)$$

so x remains positive. The inequality $x+y \leqslant 1$ then insures that y never reaches 1. Finally, the second equation of (1) then implies that

$$y(t) \geqslant -\alpha y(t)$$

so that $y(t) \geqslant y_0 \exp(-\alpha t)$ and y cannot reach 0 in finite time.

Since (x, y) stays in the compact square $[0,1]^2$, the solution of (1) is defined for all times.

■

Remark 1 In particular, since for all $t \geqslant 0$ we have $y(t) > 0$, a part of the population will always remain infectious, whatever the choice of the policy b: it is impossible to entirely eliminate the disease. This feature is due to the fact we are considering continuous populations, it would not be true for approximating finite random populations.

□

Introduce $\mathcal{B}^+$ the set of $b \in \mathcal{B}$ that are everywhere positive (nevertheless, if t is a discontinuity point of b, we can have $b(t+) = 0$).

When $b \in \mathcal{B}^+$, Lemma 1 and the first equation of (1) imply that x is decreasing, so x admits a limit $x_\infty \geqslant 0$ that it will never reach. Another consequence is:

Lemma 2 *Assume that $b \in \mathcal{B}^+$. There exists a unique function φ : $(x_\infty, x_0] \to (0,1)$ such that*

$$\forall\ t \geqslant 0, \qquad y(t) = \varphi(x(t)) \tag{14}$$

The function φ is piecewise $\mathcal{C}^1$, its left and right derivatives exist everywhere and differ only at a finite number of points. Denoting the right derivative by φ', we have

$$\forall\ r \in (x_\infty, x_0], \qquad \varphi'(r) > -1$$

Proof

As observed above, for $b \in \mathcal{B}^+$, x is decreasing from $\mathbb{R}_+$ to $(x_\infty, x_0]$. Since x it continuous, it is a homeomorphism between $\mathbb{R}_+$ to $(x_\infty, x_0]$. Denote by τ its inverse, so that

$$\forall\ u \in (x_\infty, x_0], \qquad x(\tau(u)) = u \tag{15}$$

Let $t \in \mathbb{R}_+$ be a time where b is continuous. Let $u \in (x_\infty, x_0]$ be such that $\tau(u) = t$. We can differentiate (15) at u to get that

$$\dot{\tau}(u) = \frac{1}{\dot{x}(\tau(u))} = -\frac{1}{b(\tau(u))uy(\tau(u))}$$

Considering discontinuity time t of b, we see that the above relation also holds, if $\dot{\tau}(u)$ is seen as a right derivative (recall $\dot{x}$ is a left derivate). Furthermore, taking into account that τ is decreasing, we have the existence of the left limit:

$$\lim_{v \to u_-} \dot{\tau}(v) = -\frac{1}{b(\tau(u)+)uy(\tau(u))}$$

It leads us to define φ via

$$\forall\, u \in (x_\infty, x_0], \qquad \varphi(u) := y(\tau(u))$$

since this is indeed equivalent to (14). Its left and right derivatives exist everywhere as a consequence of the differentiability properties of y and τ. These left and right derivatives do not coincide only on a finite number of points, those of the form $x(t)$, where $t \in \mathbb{R}_+$ is a discontinuity time of b.

Our conventions insure that (14) can be left differentiated everywhere and that

$$\forall\, t \geqslant 0, \qquad \dot{y}(t) = \varphi'(x(t))\dot{x}(t)$$

(recall $\dot{y}$ is a left derivate), namely

$$\forall\, t \geqslant 0, \qquad \varphi'(x(t)) = \frac{\dot{y}(t)}{\dot{x}(t)} = -1 + \frac{\alpha}{b(t)x(t)} > -1$$

Remark 2 From the knowledge of φ it is possible to reconstruct b, at least when φ' is Lipschitzian. Indeed, (1) can be written

$$\begin{cases} \dot{x}(t) = -b(t)x(t)\varphi(x(t)) \\ \varphi'(x(t))\dot{x}(t) = b(t)x(t)\varphi(x(t)) - \alpha\varphi(x(t)) \end{cases}$$

which implies that

$$\dot{x}(t) = -\varphi'(x(t))\dot{x}(t) + \alpha\varphi(x(t))$$

i.e.

$$\dot{x}(t) = \alpha\frac{\varphi(x(t))}{1 + \varphi'(x(t))}$$

So when φ' is Lipschitzian, we can solve this ODE to reconstruct $x \coloneqq (x(t))_{t\geqslant 0}$. The trajectory $y \coloneqq (y(t))_{t\geqslant 0}$ is then obtained as $(\varphi(x(t)))_{t\geqslant 0}$ and b via the formula

$$\forall\ t \geqslant 0, \qquad b(t) = -\frac{\dot{x}(t)}{x(t)y(t)}$$

□

This inequality can be translated into $\varphi' > -1$ on (x_∞, x_0).

■

To any function φ as in the previous lemma, associate the quantity

$$\mathcal{J}(\varphi) \coloneqq \int_{x_\infty}^{x_0} L(\xi, \varphi(\xi), \varphi'(\xi))\, d\xi$$

where for any $(\xi, \chi, \chi') \in (x_\infty, x_0] \times (0,1) \times (-1, +\infty)$,

$$L(\xi, \chi, \chi') \coloneqq \frac{\beta}{\alpha}\left(\frac{1+\chi'}{\chi} - \frac{\alpha}{\beta\xi\chi}\right)_+$$

The interest of these definitions is to enable us to write the cost functional $\mathcal{C}$ in terms of φ:

Lemma 3 *For $b \in \mathcal{B}^+$ and with the notations of Lemma 2, we have*

$$\mathcal{C}(b) = \mathcal{J}(\varphi)$$

Proof

Equation (16) enables us to recover b in terms of φ and x:

$$\forall\ t \geqslant 0, \qquad b(t) = \frac{\alpha}{x(t)(1 + \varphi'(x(t)))} \tag{16}$$

and we deduce that

$$\begin{aligned}
\mathcal{C}(b) &= \int_0^\infty \left(\beta - \frac{\alpha}{x(t)(1+\varphi'(x(t)))}\right)_+ dt \\
&= -\int_0^\infty \left(\beta - \frac{\alpha}{x(t)(1+\varphi'(x(t)))}\right)_+ \frac{1}{b(t)x(t)y(t)}\, \dot{x}(t)dt \\
&= \int_{x_\infty}^{x_0} \left(\beta - \frac{\alpha}{u(1+\varphi'(u))}\right)_+ \frac{1}{b(\tau(u))u\varphi(u)}\, du
\end{aligned}$$

where we used the change of variable $t = \tau(u)$, the mapping τ being defined in (15).

Let us remove the term $b(\tau(u))$ in the latter integral. Replacing t by $\tau(u)$, we get from (16)

$$b(\tau(u)) = \frac{\alpha}{u(1+\varphi'(u))}$$

so that

$$\begin{aligned}\mathcal{C}(b) &= \int_{x_\infty}^{x_0} \left(\beta - \frac{\alpha}{u(1+\varphi'(u))}\right)_+ \frac{1+\varphi'(u)}{\alpha\varphi(u)}\, du \\ &= \frac{\beta}{\alpha}\int_{x_\infty}^{x_0} \left(\frac{1+\varphi'(u)}{\varphi(u)} - \frac{\alpha}{\beta}\frac{1}{u\varphi(u)}\right)_+ du = \mathcal{J}(\varphi)\end{aligned}$$

■

Let us now extend the above transformation to a policy $b \in \mathcal{B}$ which can take the value 0. More precisely consider two times $0 \leqslant t_1 \leqslant t_2$ such that $[t_1, t_2]$ or $(t_1, t_2]$ is a connected component of the set $\{t \geqslant 0 : b(t) = 0\}$. We recall that this set is assumed to be finite union of such intervals, if it is not empty.

Let us first suppose that $t_1 \neq t_2$. On $(t_1, t_2]$, (1) is transformed into

$$\begin{cases} \dot{x}(t) = 0 \\ \dot{y}(t) = -\alpha y(t) \end{cases}$$

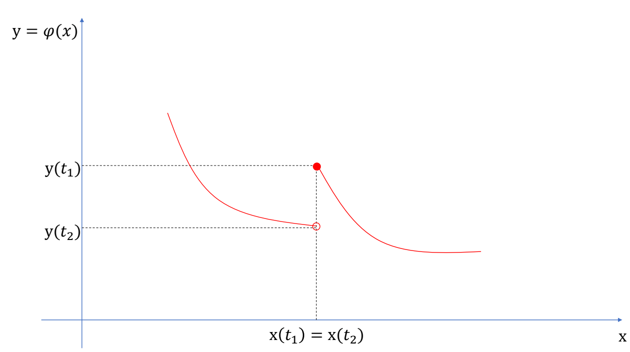
namely

$$\forall\ t \in (t_1, t_2], \qquad \begin{aligned} x(t) &= x(t_1) \\ y(t) &= y(t_1)\exp(-\alpha(t-t_1)) \end{aligned}$$

Since x remains constant and y is changing, one cannot represent y as a function φ of x. To circumvent this difficulty, we allow φ to jump at $x(t_1)$, taking

$$\begin{aligned} \varphi(x(t_1)) &\coloneqq y(t_1) \\ \varphi(x(t_1)-) &\coloneqq y(t_2) \quad = \quad y(t_1)\exp(-\alpha(t_2-t_1)) \end{aligned}$$

This is illustrated in the following diagram:



The contribution of the period $(t_1, t_2]$ to the cost $\mathcal{C}(b)$ is

$$\int_{t_1}^{t_2} (\beta - 0)_+ \, dt = \beta(t_2 - t_1) = \frac{\beta}{\alpha} \ln\left(\frac{y(t_1)}{y(t_2)}\right) = \frac{\beta}{\alpha} \ln\left(\frac{\varphi(x(t_1)}{\varphi(x(t_1)-)}\right)$$

The above observations are also valid, but trivial, when $t_2 = t_1$.

A priori, our definition of $\mathcal{B}$ does not exclude the fact that $t_2 = +\infty$, namely that b ends up vanishing identically after t_1. By convention in this case $\varphi(x(t_1)-) = 0$ and the above formula gives an infinite contribution to the cost, which is coherent with the fact that $\int_{t_1}^{+\infty} (\beta - b(t))_+ \, dt = \int_{t_1}^{+\infty} \beta \, dt = +\infty$. Since this situation is not interesting for our optimization problem, we exclude it from our considerations and from now on the connected components of $\{b = 0\}$ are assumed to be bounded.

Repeating the above treatment to all the connected components of $\{b = 0\}$ and extending Lemmas 2 and 3 to the connected components of $\{b > 0\}$, we can associate to any $b \in \mathcal{B}$ a function φ satisfying

(H1): φ is defined on $(x_\infty, x_0]$, takes values in $(0, 1)$ and $\varphi(x_0) = y_0$,

(H2): φ has at most a finite number of discontinuity points and is right continuous and admits a positive left limit at them,

(H3): at any discontinuity point u, $\varphi(u) > \varphi(u-)$,

(H4): φ admits a right derivative φ', as well as a left derivative, outside a finite number of points (which includes the discontinuity points of φ, but also some points corresponding to the case $t_2 = t_1$ described above, since due to (16), at these points φ' may diverge to $+\infty$),

(H5): $\varphi' > -1$ where it is defined,

But the most important feature is that

$$\mathcal{C}(b) = \mathcal{J}(\varphi)$$

where the functional is defined by

$$\mathcal{J}(\varphi) := \int_{(x_\infty, x_0)} L(\xi, \varphi(\xi), \varphi'(\xi)) \, d\xi + \frac{\beta}{\alpha} \sum_{u \in (x_\infty, x_0] \,:\, \varphi(u) \neq \varphi(u-)} \ln\left(\frac{\varphi(u)}{\varphi(u-)}\right)$$

(x_∞ is excluded from the last sum, since we don't allow b to end up vanishing identically).

There is a point which is not satisfactory in the above construction of φ: its definition domain $(x_\infty, x_0]$ still depends explicitly on b via x_∞. It is possible to erase this drawback with the convention that φ vanishes on

$[0, x_\infty]$ Accordingly, L has to be extended to $[0, x_0] \times [0, 1) \times (-1, +\infty)$ via the convention that $L(\xi, \chi, \chi') = 0$ if $\xi \in [0, x_\infty]$ (or equivalently if $\chi = 0$).

Remarking that the condition $y \leqslant \gamma$ can be translated into $\varphi \leqslant \gamma$, we have embedded the problem of the global minimization of $\mathcal{C}$ over $\mathcal{B}_\gamma$ into the problem of the global minimization of $\mathcal{J}$ over $\widetilde{\mathcal{F}}_\gamma$, the set of functions φ that satisfy the requirements (H1a), (H2), (H3), (H4) and (H5), where (H1) has been replaced by

(H1a): φ is defined on $[0, x_0]$, takes values in $[0, \gamma]$, $\varphi(0) = 0$, $\varphi(x_0) = y_0$ and if $\varphi(u) > 0$ for some $u \in (0, x_0]$, then $\varphi(v) > 0$ and $\varphi(v-) > 0$ for all $v \in [u, x_0]$.

As it was explained in the main text, the most important part of the above optimization problem concerns the contribution of the restriction of the function φ to the interval $[\alpha/\beta, x_0]$, since once x has reached α/β, the *laissez-faire* policy is cost-free and induces y to be non-increasing. Let us furthermore assume the function φ satisfies

$$\varphi(\alpha/\beta) = \gamma \tag{17}$$

This restriction will be justified a posteriori (see Subsection 5.6), but it can also be understood a priori according to the following heuristic. Let us come back to the temporal description given by (b, x, y). To get $x_\infty \geqslant \alpha/\beta$ is extremely costly and requires in fact $\mathcal{C}(b) = +\infty$ (think for instance to the case where $x_\infty = x_0$ which asks that $b \equiv 0$), so for the optimization problem at hand, we can dismiss this possibility and assume $x_\infty < \alpha/\beta$. Then consider τ_+ the first time x reaches α/β, we have $y(\tau_+) \in (0, \gamma]$. As already observed, taking $b = \beta$ after time τ_+ leads to a non-increasing evolution of y on $[\tau_+, +\infty)$. In particular, $y(t)$ stays below γ for all times $t \in [\tau_+, +\infty)$, while this part of the trajectory does not participate positively to the cost. Due to our underlying assumption, the *laissez-faire* policy $b \equiv \beta$ leads y to hit γ strictly before x reaches α/β. Then if $y(\tau_+) < \gamma$, it means that for some times $t \in [0, \tau_+)$, we have $b(t) < \beta$ and $y(t) < \gamma$. Increasing a little b at such times, will have the effect of increasing $y(\tau_+)$ while keeping y below γ. Furthermore this operation will decrease a little the cost. As a consequence, in order to find a minimizing policy b for $\mathcal{C}$, we can assume that $y(\tau_+) = \gamma$. This amounts to assuming (17).

These considerations and Assumption (17) lead us to modify the functional set $\widetilde{\mathcal{F}}_\gamma$ into $\mathcal{F}_\gamma$, replacing its first requirement (H1a) by

(H1b): φ is defined on $[\alpha/\beta, x_0]$, takes values in $(0, \gamma]$, the left limits of φ are positive, $\varphi(\alpha/\beta) = \gamma$ and $\varphi(x_0) = y_0$

Of course the cost functional on such functions φ is given by

$$\begin{aligned} \mathcal{J}(\varphi) \coloneqq & \int_{(\alpha/\beta,x_0)} L(\xi,\varphi(\xi),\varphi'(\xi))\,d\xi \\ & +\frac{\beta}{\alpha}\sum_{u\in(\alpha/\beta,x_0]\,:\,\varphi(u)\neq\varphi(u-)} \ln\left(\frac{\varphi(u)}{\varphi(u-)}\right) \end{aligned} \tag{18}$$

Our ultimate goal is to prove that $\mathcal{J}$ admits a unique minimizer φ^* over $\mathcal{F}_\gamma$ and that this minimizer is obtained from b^* by the operation described above. Up to the justification of the reduction of (H1a) to (H1b), given at the end of Subsection 5.6, Theorem 1 will then be proven.

5.2 Extension of $\mathcal{J}$ to measures

To prove the existence of the minimizer φ^* of $\mathcal{J}$ on $\mathcal{F}_\gamma$, we begin by generalizing this optimization problem by replacing $\mathcal{F}_\gamma$ by a set of measures.

To any given $\varphi \in \mathcal{F}_\gamma$, we associate three measures μ, ψ and ν on $I \coloneqq [\alpha/\beta, x_0]$ via

$$\begin{aligned} \mu(dx) &\coloneqq \frac{\varphi'(x)}{\varphi(x)}\,dx + \sum_{u\in(\alpha/\beta,x_0]\,:\,\varphi(u)\neq\varphi(u-)} \ln\left(\frac{\varphi(u)}{\varphi(u-)}\right)\delta_u(dx) \\ \psi(dx) &\coloneqq \frac{1}{\varphi(x)}\,dx \\ \nu(dx) &\coloneqq \frac{1}{\varphi(x)}\left(1-\frac{\alpha}{\beta x}\right)\,dx \end{aligned}$$

(where δ_u stands for the Dirac mass at u).

Note that ψ and ν are non-negative measures, but μ is a signed measure. Denote F_μ the repartition function associated to μ via

$$\forall\ x \in I, \qquad F_\mu(x) \coloneqq \mu([\alpha/\beta, x])$$

Recall that φ, as well as its left limits, are positive on I, it follows that φ is bounded below by a positive constant on I. This observation enables us to compute F_μ:

$$\forall\ x \in I, \qquad F_\mu(x) = \ln(\varphi(x)) - \ln(\varphi(\alpha/\beta)) = \ln(\varphi(x)/\gamma) \tag{19}$$

We deduce that

$$\forall\ x \in I, \qquad \varphi(x) = \gamma\exp(F_\mu(x))$$

and as a consequence

$$\psi(dx) \coloneqq \frac{\exp(-F_\mu(x))}{\gamma}dx$$
$$\nu(dx) \coloneqq \left(1-\frac{\alpha}{\beta x}\right)\frac{\exp(-F_\mu(x))}{\gamma}dx$$

On these expressions, it appears that ψ and ν only depend on μ (in addition to the constants α, β, γ), so they will be denoted ψ_μ ν_μ from now on.

From (19) we get that the total weight $\mu(I)$ of μ is given by $F_\mu(x_0) = \ln(y_0/\gamma)$. Let us estimate the total variation $\|\mu\|_{\mathrm{tv}}$ of μ:

Lemma 4 *We have*

$$\|\mu\|_{\mathrm{tv}} \leqslant \frac{2}{\epsilon}(x_0-\alpha/\beta)+\ln(y_0/\gamma)$$

where

$$\epsilon \coloneqq \inf\{\varphi(x)\ :\ x\in I\} \quad = \quad \min\{\varphi(x)\wedge\varphi(x-)\ :\ x\in I\}$$

Proof

Recall that any signed measure m on I can be decomposed into $m_+ - m_-$, where m_- and m_+ are two non-negative measures mutually singular. The total variation is given by $\|m\|_{\mathrm{tv}} = m_-(I) + m_+(I)$.

Coming back to μ, we have

$$\mu_-(dx) = \frac{|\varphi'(x)|}{\varphi(x)}\mathbb{1}_{\{\varphi'(x)>0\}}\,dx$$

so that

$$\mu_-(I) \leqslant \int_I \frac{1}{\varphi(x)}\,dx \leqslant \frac{1}{\epsilon}\int_I 1\,dx = \frac{x_0-\alpha/\beta}{\epsilon}$$

It follows that

$$\|\mu\|_{\mathrm{tv}} = \mu_-(I)+\mu_+(I) = 2\mu_-(I)+\mu(I) \leqslant \frac{2}{\epsilon}(x_0-\alpha/\beta)+\ln(y_0/\gamma)$$

■

The quantity $\epsilon > 0$ associated to φ in the previous lemma can be estimated in terms of $\mathcal{J}(\varphi)$:

Lemma 5 *We have*

$$\epsilon \geqslant y_0\exp(-\beta\mathcal{J}(\varphi)/\alpha)$$

Fix an arbitrary point $x \in I$. We have

$$\begin{aligned}
\mathcal{J}(\varphi) &\geqslant \frac{\beta}{\alpha}\int_I L(u,\varphi(u),\varphi'(u))\,du \geqslant \frac{\beta}{\alpha}\int_x^{x_0} L(u,\varphi(u),\varphi'(u))\,du\\
&= \frac{\beta}{\alpha}\int_x^{x_0}\left(\frac{\varphi'(u)}{\varphi(u)}+\frac{\alpha}{\beta\varphi(u)}\left(1-\frac{\alpha}{\beta u}\right)\right)_+ du\\
&\geqslant \frac{\beta}{\alpha}\int_x^{x_0}\frac{\varphi'(u)}{\varphi(u)}+\frac{\alpha}{\beta\varphi(u)}\left(1-\frac{\alpha}{\beta u}\right) du\\
&\geqslant \frac{\beta}{\alpha}\int_x^{x_0}\frac{\varphi'(u)}{\varphi(u)}\,du = \frac{\beta}{\alpha}\ln\left(\frac{\varphi(x_0)}{\varphi(x)}\right)
\end{aligned}$$

where in the last-but-one inequality, we took into account that $1-\frac{\alpha}{\beta u} \geqslant 0$ for $u \geqslant x \geqslant \alpha/\beta$. The above bound can be written

$$\varphi(x) \geqslant \varphi(x_0)\exp(-\alpha\mathcal{J}(\varphi)/\beta) = y_0\exp(-\alpha\mathcal{J}(\varphi)/\beta)$$

By taking the infimum over all $x \in I$, we get the desired result.

■

Consider an element φ_0 of $\mathcal{F}_\gamma$ such that $\mathcal{J}(\varphi_0) < +\infty$, for instance the function constructed from b^* as in the previous subsection. Let

$$M \coloneqq \frac{2\exp(\alpha(1+\mathcal{J}(\varphi_0))/\beta)}{y_0}(x_0-\alpha/\beta)+\ln(y_0/\gamma) \tag{20}$$

The global minimization of $\mathcal{J}$ on $\mathcal{F}_\gamma$ is equivalent of the global minimization of $\mathcal{J}$ on $\{\varphi \in \mathcal{F}_\gamma : \mathcal{J}(\varphi) \leqslant \mathcal{J}(\varphi_0)\}$. So according to Lemma 4, we can restrict our attention to measure μ satisfying

(C1): $\|\mu\|_{\mathrm{tv}} \leqslant M$.

Furthermore note that the fact that φ belongs to $\mathcal{F}_\gamma$ implies three conditions on μ:

(C2): $\mu(I) = \ln(y_0/\gamma)$,

(C3): $F_\mu \leqslant 0$,

(C4): $\mu + \psi_\mu \geqslant 0$.

Denote by $\mathcal{M}_\gamma$ the set of signed measures μ on I which satisfy the conditions (C1), (C2), (C3) and (C4).

Up to now, we did not use ν_μ, its interest comes from the fact that for $\varphi \in \mathcal{F}_\gamma$, the cost functional writes

$$\mathcal{J}(\varphi) = \frac{\beta}{\alpha}(\mu+\nu_\mu)_+(I)$$

This observation leads us to define for any $\mu \in \mathcal{M}_\gamma$,

$$\mathcal{K}(\mu) = \frac{\beta}{\alpha}(\mu + \nu_\mu)_+(I)$$

since the global minimization of $\mathcal{J}$ on $\{\varphi \in \mathcal{F}_\gamma \,:\, \mathcal{J}(\varphi) \leqslant \mathcal{J}(\varphi_0)\}$ can be embedded in the global minimization of $\mathcal{K}$ on $\mathcal{M}_\gamma$.

5.3 Existence of a global minimizer of $\mathcal{K}$ on $\mathcal{M}_\gamma$

The successive extensions of our initial minimization problem worked out in the two previous subsections will be justified here by showing the existence of a global minimizer, via abstract topological arguments. It won't be clear that such a minimizer from $\mathcal{M}_\gamma$ corresponds to an element of $\mathcal{B}_\gamma$ via the transformation of Subsection 5.1: this will be investigated in the next subsections.

Let us endow the set of (signed) measures on I with the weak topology, i.e. a sequence $(\mu_n)_{n\in\mathbb{N}}$ of such measures converges toward a measure μ if and only if for any continuous function $g \,:\, I \to \mathbb{R}$, we have

$$\lim_{n\to\infty} \mu_n[g] = \mu[g]$$

The existence of a global minimizer of $\mathcal{K}$ on $\mathcal{M}_\gamma$ is a consequence of a version of Weierstrass' maximum theorem through the two following results.

Proposition 3 *The set $\mathcal{M}_\gamma$ is compact.*

Proposition 4 *The mapping $\mathcal{K} \,:\, \mathcal{M}_\gamma \to \mathbb{R}$ is lower semi-continuous.*

Proof of Proposition 3

Consider the ball (with respect to the strong topology) $B(M)$ consisting of the signed measures whose total variation is smaller or equal to M. It is well-known that $B(M)$ is weakly compact. So it sufficient to show that the sets

$$\begin{aligned} S_1 &\coloneqq \{\mu \in B(M) \,:\, \mu(I) = \ln(y_0/\gamma)\} \\ S_2 &\coloneqq \{\mu \in B(M) \,:\, F_\mu \leqslant 0\} \\ S_3 &\coloneqq \{\mu \in B(M) \,:\, \mu + \psi_\mu \geqslant 0\} \end{aligned}$$

are closed.

• Concerning S_1, this is obvious, since $\mu(I) = \mu[\mathbb{1}_I]$, where $\mathbb{1}_I$ is the continuous function on I always taking the value 1.

• For S_2, consider a sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures from S_2 converging toward a signed measure μ. We have to check that $\mu \in S_2$. Consider $\mathcal{A}$

the set of atoms of μ, $\mathcal{A}$ is at most denumerable and for $x \in I\backslash\mathcal{A}$, we have

$$\lim_{n\to\infty} F_{\mu_n}(x) = F_\mu \tag{21}$$

We deduce

$$\forall\ x \in I\backslash\mathcal{A}, \qquad F_\mu(x) \leqslant 0$$

It remains to use that F_μ is right-continuous and that $I\backslash\mathcal{A}$ is dense in I to extend the validity of the above inequality to the whole set I.

• For S_3, consider a sequence $(\mu_n)_{n\in\mathbb{N}}$ of measures from S_3 converging toward a signed measure μ. We have to check that $\mu + \psi_\mu \geqslant 0$. Consider g a continuous function on I. We begin by showing that

$$\lim_{n\to\infty} \psi_{\mu_n}[g] = \psi_\mu[g] \tag{22}$$

Indeed, for any $n \in \mathbb{N}$, we have

$$\psi_{\mu_n}[g] = \frac{1}{\gamma}\int_I g(x)\exp(-F_{\mu_n}(x))\,dx$$

We have seen in (21) that F_{μ_n} is almost everywhere converging to F_μ. Furthermore, we have for all $n \in \mathbb{N}$,

$$\forall\ x \in I, \qquad |F_{\mu_n}(x)| \leqslant \|\mu_n\|_{\mathrm{tv}} \leqslant M$$

It follows that dominated convergence can be invoked to conclude to (22).

We deduce that

$$\lim_{n\to\infty}(\mu_n + \psi_{\mu_n})[g] = (\mu + \psi_\mu)[g]$$

Now assume furthermore that $g \geqslant 0$. The above convergence implies that

$$(\mu + \psi_\mu)[g] \geqslant 0$$

Since this is true for all non-negative continuous function g, we get that $\mu + \psi_\mu \geqslant 0$, as desired.

■

Proof of Proposition 4

To get that the mapping $\mathcal{K} : \mathcal{M}_\gamma \to \mathbb{R}$ is lower semi-continuous, it is sufficient to write it as the supremum of (weakly) continuous functions on $\mathcal{M}_\gamma$. By definition, we have

$$\mathcal{K}(\mu) = \frac{\beta}{\alpha}(\mu + \nu_\mu)_+(I) = \frac{\beta}{\alpha} \sup_{g \in \mathcal{C}(I)\,:\,0 \leqslant g \leqslant 1} (\mu + \nu_\mu)[g]$$

where $\mathcal{C}(I)$ is the set of continuous functions on I.

So it remains to check that for any fixed $g \in \mathcal{C}(I)$ with $0 \leqslant g \leqslant 1$, the mapping

$$\mathcal{M}_\gamma \ni \mu \mapsto (\mu + \nu_\mu)[g]$$

is continuous. The proof of this continuity is similar to the closure of S_3 in the proof of Proposition 3: both the mappings $\mathcal{M}_\gamma \ni \mu \mapsto \psi_\mu$ and $\mathcal{M}_\gamma \ni \mu \mapsto \nu_\mu$ are continuous (taking values in the set of non-negative measures on I endowed with the weak topology).

■

5.4 Reduction to absolutely continuous measures

Let μ^* be a minimizer of $\mathcal{K}$ on $\mathcal{M}_\gamma$. We will show here that μ^* is absolutely continuous with respect to λ, the Lebesgue measure on I.

Recall that any signed measure μ on I can be uniquely decomposed into a sum $\mu_{\mathrm{a}} + \mu_s + \mu_{\mathrm{c}}$, where μ_{a} is atomic, μ_{c} is diffuse and singular with respect to λ and μ_{c} is absolutely continuous with respect to λ. Due to Radon-Nikodym theorem, μ_{c} admits a (signed) density $f : I \to \mathbb{R}$ with respect to λ, this relation will be denoted $\mu_{\mathrm{c}} = f \cdot \lambda$.

Using this decomposition, the cost functional $\mathcal{K}$ can be written under the form

$$\mathcal{K}(\mu) = \frac{\beta}{\alpha}\left(\mu_{\mathrm{a}}(I) + \mu_s(I) + \int_I (f + \nu_\mu)_+ \, d\lambda\right)$$

where we identified ν_μ with its density with respect to λ, namely

$$\forall\ x \in I, \qquad \nu_\mu(x) = \frac{1}{\gamma}\left(1 - \frac{\alpha}{\beta x}\right)\exp(-F_\mu(x))$$

Our goal here is to show that $\mu^*_{\mathrm{a}} = 0 = \mu^*_{\mathrm{s}}$, by perturbative arguments. Both proofs will follow the same pattern, but the deduction of $\mu_{\mathrm{a}} = 0$ is simpler, so for pedagogical reasons we will insist on this one.

Proof of $\mu^*_{\mathrm{a}} = 0$

The argument is by contradiction. Assume that $\mu^*_{\mathrm{a}} \neq 0$, then there exist $x_1 \in I$ and $\epsilon > 0$ such that $\mu^*_{\mathrm{a}} \geqslant \epsilon\delta_{x_1}$. Since $F_{\mu^*}(x_1) - F_{\mu^*}(x_1-) \geqslant \epsilon$ and that $F_{\mu^*}(x_1) \leqslant 0$, we have $F_{\mu^*}(x_1-) \leqslant -\epsilon$ and we can find $x_2 \in (\alpha/\beta, y_1)$ such that

$$\forall\ x \in [x_2, x_1), \qquad F_{\mu^*}(x) \leqslant -\frac{\epsilon}{2}$$

It leads us to consider for $r > 0$, the perturbation μ_r^* defined by

$$\mu_r^* \coloneqq \mu + r\frac{\mathbb{1}_{[x_2,x_1)}}{x_1 - x_2} \cdot \lambda - r\delta_{x_1}$$

(where $\mathbb{1}_{[x_2,x_1)}$ is the indicator function of $[x_2, x_1)$ and the middle term is an absolutely continuous measure).

Let us check that $\mu_r \in \mathcal{M}_\gamma$ for $r > 0$ small enough, namely that (C1), (C2), (C3) and (C4) are satisfied by μ_r^*.

• $\|\mu_r^*\|_{\mathrm{tv}} \leqslant M$:

Since μ^* is also a minimizer of $\mathcal{K}$ on the space of signed measures on I, we have $\mathcal{K}(\mu^*) \leqslant \mathcal{L}(\varphi_0)$, where φ_0 was defined above (20). It follows from Lemmas 4 and 5, taking into account the definition of M in (20), that $\|\mu^*\|_{\mathrm{tv}} < M$. Triangular inequality with respect to the total variation norm implies

$$\|\mu_r^*\|_{\mathrm{tv}} \leqslant \|\mu^*\|_{\mathrm{tv}} + \left\| r\frac{\mathbb{1}_{[x_2,x_1)}}{x_1 - x_2} \cdot \lambda - r\delta_{x_1} \right\|_{\mathrm{tv}} = \|\mu^*\|_{\mathrm{tv}} + 2r$$

insuring that for $r \leqslant (M - \|\mu^*\|_{\mathrm{tv}})/2$, we have $\|\mu^*\|_{\mathrm{tv}} \leqslant M$.

• $\mu_r^*(I) = \ln(y_0/\gamma)$:

The total weight of μ_r^* is always the same as that of μ^*, since

$$\left(\frac{\mathbb{1}_{[x_2,x_1)}}{x_1 - x_2} \cdot \lambda - \delta_{x_1}\right)(I) = 1 - 1 \ = \ 0$$

• $F_{\mu_r^*} \leqslant 0$:

Note that outside (x_2, x_1), $F_{\mu_r^*}$ coincides with F_{μ^*}, so we just need to check that $F_{\mu_r^*}(x) \leqslant 0$ for all $x \in (x_1, x_2)$. Indeed, we have

$$\forall\ x \in (x_1, x_2), \qquad F_{\mu_r^*}(x) = F_{\mu^*}(x) + r\frac{x - x_2}{x_1 - x_2} \leqslant -\frac{\epsilon}{2} + r \leqslant 0$$

as soon as $r \leqslant \epsilon/2$.

• $\mu_r^* + \psi_{\mu_r^*} \geqslant 0$:

Note again that outside $(x_2, x_1]$, we have $\mu_r^* + \psi_{\mu_r^*} = \mu^* + \psi_{\mu^*}$, so we just need to check that the measure $\mu_r^* + \psi_{\mu_r^*}$ is non-negative on $(x_2, x_1]$. The diffusive singular part of $\mu_r^* + \psi_{\mu_r^*}$ is the same as that of $\mu^* + \psi_{\mu^*}$ and the atomic ones only differ at x_1. Note that $\mu_r^*(\{x_1\}) = \mu_s(\{x_1\}) - r \geqslant$

$\epsilon - r$ and this is non-negative, as soon as $r \leqslant \epsilon$. Concerning the absolutely continuous part, denote f^* (respectively f_r^*) the density of μ^* (resp. μ_r^*) with respect to λ. We have a.e. for $x \in (x_2, x_1)$,

$$f_r^*(x) = f^*(x) + \frac{r}{x_1 - x_2} \tag{23}$$

Identify $\psi_{\mu_r^*}$ with its density, so that we can write (a.e.)

$$\forall\ x \in (x_2, x_1), \qquad \psi_{\mu_r^*}(x) = \frac{\exp(-F_{\mu_r^*}(x))}{\gamma}$$

Note that

$$\forall\ x \in (x_2, x_1), \qquad \left|F_{\mu_r^*}(x) - F_{\mu^*}(x)\right| = \left|\int_{x_2}^{x} \frac{r}{x_1 - x_2}\, d\lambda\right| = \frac{r(x - x_2)}{x_1 - x_2} \ \leqslant\ r$$

Let

$$C := \frac{\exp(1 + \max_{x \in [x_2, x_1]} |F_\mu(x)|)}{\gamma}$$

For any $r \in [0, 1]$, we have

$$\forall\ x \in (x_2, x_1), \qquad \left|\psi_{\mu_r^*}(x) - \psi_{\mu^*}(x)\right| \leqslant Cr$$

Comparing with (23), we get that if x_2 has been chosen so that $x_1 - x_2 \leqslant C^{-1}$, then a.e. for $x \in (x_2, x_1)$,

$$f_r^*(x) + \psi_{\mu_r^*}(x) \geqslant f^*(x) + \psi_{\mu^*}(x) \ \geqslant\ 0$$

This ends the proof that $\mu_r \in \mathcal{M}_\gamma$, as soon as $r > 0$ is small enough and x_2 is sufficiently close to x_1.

Let us evaluate $\mathcal{K}(\mu_r^*)$. We have

$$\mathcal{K}(\mu_r^*) = \mathcal{K}(\mu^*) - r + \int_{x_2}^{x_1} \left(f + \frac{r}{x_1 - x_2} + \nu_{\mu_r^*}\right)_+ - (f + \nu_\mu)_+\, d\lambda \tag{24}$$

and for almost all x belonging to (x_2, x_1),

$$\left(f(x) + \frac{r}{x_1 - x_2} + \nu_{\mu_r^*}(x)\right)_+ - (f(x) + \nu_\mu(x))_+ \leqslant \frac{r}{x_1 - x_2} + \nu_{\mu_r^*}(x) - \nu_\mu(x)$$

Note that for $x \in (x_2, x_1)$, we have $F_{\mu_r^*}(x) > F_{\mu^*}(x)$, so that $\nu_{\mu_r^*}(x) < \nu_{\mu^*}(x)$. Thus we get

$$\left(f(x) + \frac{r}{x_1 - x_2} + \nu_{\mu_r^*}(x)\right)_+ - (f(x) + \nu_\mu(x))_+ < \frac{r}{x_1 - x_2}$$

It follows that that for $r > 0$:

$$\int_{x_2}^{x_1} \left(f + \frac{r}{x_1 - x_2} + \nu_{\mu_r^*}\right)_+ - (f + \nu_\mu)_+ \, d\lambda < r$$

and (24) implies the contradiction $\mathcal{K}(\mu_r^*) < \mathcal{K}(\mu^*)$.

■

Remark 5 In Subsection 5.2, we have seen that the atomic part of a measure corresponds to imposing the drastic policy $b = 0$ for some time, as a partial attempt toward eradication of the disease (according to Remark 1 this goal cannot be fully attained). The significance of $\mu_{\mathrm{a}}^* = 0$ is that such attempts are sub-optimal. From the above proof we see that it is better to replace such attempts by future softer policies, replacing a (partial) Dirac mass at x_1 by a density before x_1 (recall that x and time go in reverse directions).

□

Proof of $\mu_{\mathrm{s}}^* = 0$

The pattern of the proof is the same as that for $\mu_{\mathrm{a}}^* = 0$, except that we have to "thicken a little" x_1. Indeed, if $\mu_{\mathrm{s}}^* \neq 0$, we can find $x_1 \in (\alpha/\beta, x_0]$ and $\epsilon' > 0$ such that

$$\begin{aligned} & \frac{\alpha}{\beta} < x_1 - \epsilon' \\ & \mu_{\mathrm{s}}^*((x_1 - \epsilon', x_1 + \epsilon')) > 0 \\ \forall\, \epsilon'' \in (0, \epsilon'], \qquad & \mu_{\mathrm{s}}^*((x_1 - \epsilon'', x_1 + \epsilon'')) \geqslant 2\mu_{\mathrm{c}}^*((x_1 - \epsilon'', x_1 + \epsilon'')) \end{aligned}$$

In comparison with the previous proof, the restriction of μ_{s}^* to $(x_1 - \epsilon', x_1 + \epsilon')$ plays the role of $\epsilon\delta_{x_1}$ with $\mu_{\mathrm{s}}^*((x_1 - \epsilon', x_1 + \epsilon'))$ playing the role of ϵ.

It is now possible to find $x_2 \in (\alpha/\beta, x_1 - \epsilon')$ such that for $r > 0$ small enough, the measure

$$\mu_r^* \coloneqq \mu + r\frac{\mathbb{1}_{[x_2, x_1 - \epsilon')}}{x_1 - \epsilon' - x_2} \cdot \lambda - \frac{r}{\mu_{\mathrm{s}}^*([x_1 - \epsilon', x_1 + \epsilon'])}\mathbb{1}_{(x_1 - \epsilon', x_1 + \epsilon')} \cdot \mu_{\mathrm{s}}^*$$

belongs to $\mathcal{M}_\gamma$ and

$$\mathcal{K}(\mu_r^*) < \mathcal{K}(\mu^*)$$

This is the desired contradiction. The details of the adaptation of the arguments of the proof of $\mu_{\mathrm{a}}^* = 0$ are left to the reader.

■

We have shown that any minimizer μ^* of $\mathcal{K}$ can be written in the form $f^* \cdot \lambda$. As a consequence, such a density f^* is a minimizer of a suitably formulated optimization problem, to which we now turn..

For any integrable function f on I, let associate to f the notions that were previously associated to the measure $f \cdot \lambda$: for any $x \in I$,

$$\begin{aligned} F(x) \coloneqq F_{f\cdot\lambda}(x) &= \int_{\alpha/\beta}^{x} f(u)\,du \\ \psi_F(x) &\coloneqq \frac{1}{\gamma}\exp(-F(x)) \\ \nu_F(x) &\coloneqq \left(1-\frac{\alpha}{\beta x}\right)\exp(-F(x)) \end{aligned}$$

We introduce $\mathcal{D}_\gamma$ the set of integrable functions f such that

$$F(x_0) = \ln(y_0/\gamma), \quad F \leqslant 0, \quad f + \psi_F \geqslant 0$$

and we consider on $\mathcal{D}_\gamma$ the functional

$$\mathcal{G}(f) \coloneqq \mathcal{K}(f\cdot\lambda) = \frac{\beta}{\alpha}\int_I (f+\nu_F)_+\,d\lambda$$

We have shown the optimization problem of $\mathcal{G}$ on $\mathcal{D}_\gamma$ (which is an extension of the optimization problem of $\mathcal{J}$ on $\mathcal{F}_\gamma$) admits global minimizers: they are exactly the functions f^* such that $f^* \cdot \lambda \in \mathcal{M}_\gamma$ is a global minimizer of $\mathcal{K}$.

5.5 Characterization of the minimizer of $\mathcal{G}$ on $\mathcal{D}_\gamma$

In this subsection, a minimizer of $\mathcal{G}$ on $\mathcal{D}_\gamma$ will be denoted f and our goal is to show that it is a.e. equal to the function f^* described below.

Define x^* as the unique solution belonging to $[\alpha/\beta, x_0]$ of the equation

$$x^* - \frac{\alpha}{\beta}\ln(x^*) = 1 - \gamma - \frac{\alpha}{\beta}\ln(x_0) \tag{25}$$

(the existence of this solution is due to our underlying assumption, note that x^* coincides with $x(\tau_1)$ defined above (8)), and take

$$\forall\ x \in [\alpha/\beta, x_0],$$
$$f^*(x) \coloneqq \begin{cases} 0 & \text{, if } x \leqslant x^* \\ -\left(\gamma - \left(x - x^* - \frac{\alpha}{\beta}\ln(x/x^*)\right)\right)^{-1}\left(1-\frac{\alpha}{\beta x}\right) & \text{, if } x > x^* \end{cases} \tag{26}$$

Theorem 2 *The function f^* is the unique minimizer of $\mathcal{G}$ on $\mathcal{D}_\gamma$, up to modifications on subsets of I with Lebesque measure zero.*

The proof of this result proceeds by way of a few lemmas.

We associate to f, a minimizer of $\mathcal{G}$ on $\mathcal{D}_\gamma$, the notions defined at the end of the previous subsection. In addition, we define

$$\begin{aligned} v &\coloneqq \max\{x \in I \;:\; F(x) = 0\} \\ \mathcal{G}_1(f) &\coloneqq \int_{\alpha/\beta}^{v} (f + \nu_F)_+ \, d\lambda \\ \mathcal{G}_2(f) &\coloneqq \int_{v}^{x_0} (f + \nu_F)_+ \, d\lambda \end{aligned}$$

Note that $\mathcal{G}(f) = (\beta/\alpha)(\mathcal{G}_1(f) + \mathcal{G}_2(f))$. We will investigate separately $\mathcal{G}_1(f)$ and $\mathcal{G}_2(f)$. The computation of the former will be a consequence of:

Lemma 6 *The function F vanishes everywhere on $[\alpha/\beta, v]$.*

Proof

We have

$$\mathcal{G}_1(f) \geqslant \int_{\alpha/\beta}^{v} f + \nu_F \, d\lambda = F(v) + \int_{\alpha/\beta}^{v} \nu_F \, d\lambda = \int_{\alpha/\beta}^{v} \nu_F \, d\lambda$$

since by continuity of F, $F(v) = 0$.

The bound $F \leqslant 0$ implies that

$$\forall\; x \in I, \qquad \nu_F(x) = \frac{1}{\gamma}\left(1 - \frac{\alpha}{\beta x}\right)\exp(-F(x)) \geqslant \frac{1}{\gamma}\left(1 - \frac{\alpha}{\beta x}\right) \tag{27}$$

and we deduce

$$\mathcal{G}_1(f) \geqslant \frac{1}{\gamma}\int_{\alpha/\beta}^{v} 1 - \frac{\alpha}{\beta x} \, dx = \mathcal{G}_1(\widetilde{f})$$

where the function $\widetilde{f}$ is given by

$$\forall\; x \in I, \qquad \widetilde{f}(x) \coloneqq \begin{cases} 0 & , \text{ if } x \leqslant v \\ f(x) & , \text{ if } x > v \end{cases}$$

Since $F(v) = 0$, we get that for any $x \in [v, x_0]$, that

$$\widetilde{F}(x) \coloneqq \int_{\alpha/\beta}^{x} \widetilde{f}(u) \, du = \int_{v}^{x} \widetilde{f}(u) \, du = \int_{v}^{x} f(u) \, du = F(x)$$

We deduce that $\widetilde{f} + \nu_{\widetilde{F}}$ coincides with $f + \nu_F$ on $[v, x_0]$ and thus $\mathcal{G}_2(\widetilde{f}) = \mathcal{G}_2(f)$, which implies

$$\mathcal{G}(\widetilde{f}) \leqslant \mathcal{G}(f)$$

Taking into account that $\widetilde{f} \in \mathcal{D}_\gamma$ and that f is a minimizer of $\mathcal{G}$ on $\mathcal{D}_\gamma$, we must have $\mathcal{G}(\widetilde{f}) = \mathcal{G}(f)$ and $\mathcal{G}_1(\widetilde{f}) = \mathcal{G}_1(f)$. In particular, (27) must be an equality a.e. on $[\alpha/\beta, v]$. This means that F vanishes a.e. on $[\alpha/\beta, v]$. The continuity of F then implies that F vanishes identically on $[\alpha/\beta, v]$.

■

As announced, we deduce the value of $\mathcal{G}_1(f)$:

Lemma 7

$$\mathcal{G}_1(f) = \frac{1}{\gamma}\left(v - \frac{\alpha}{\beta}\ln(v)\right) + \frac{\alpha}{\beta\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1\right)$$

and the r.h.s. is increasing with respect to v.

Proof

From Lemma 6, we deduce that $f = 0$ on $[\alpha/\beta, v]$ (a.e., as all statements about densities, in the sequel we will no longer mention it). Recalling that $\nu_F \geqslant 0$, we get

$$\begin{aligned}
\mathcal{G}_1(f) &= \int_{\alpha/\beta}^{v} \nu_F \, d\lambda = \int_{\alpha/\beta}^{v} \nu_0 \, d\lambda = \frac{1}{\gamma}\int_{\alpha/\beta}^{v} 1 - \frac{\alpha}{\beta x}\, dx \\
&= \frac{1}{\gamma}\left[x - \frac{\alpha}{\beta}\ln(x)\right]_{\alpha/\beta}^{v} = \frac{1}{\gamma}\left(v - \frac{\alpha}{\beta}\ln(v)\right) + \frac{\alpha}{\beta\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1\right)
\end{aligned}$$

■

We now come to the study of $\mathcal{G}_2$. Our main step in this direction is:

Proposition 6 *We have $f + \nu_F \leqslant 0$ on $[v, x_0]$.*

Proof

Consider $x_1 \in (v, x_0]$. Since x_1 can be arbitrary close to v, it is sufficient to show that $f + \nu_F \leqslant 0$ on $[x_1, x_0]$. The advantage of considering such a x_1 is that we can find $\eta > 0$ such that

$$\forall\ x \in [x_1, x_0], \qquad F(x) \leqslant -\eta \tag{28}$$

This property will be important for the perturbations of f we are to consider. More precisely they will be of the form $f_r := f + rg$, with $r > 0$ sufficiently small and where g is an appropriate bounded function on $[\alpha/\beta, x_0]$ and satisfying $g = 0$ on $[\alpha/\beta, x_1]$.

Defining for $x \in [\alpha/\beta, x_0]$

$$F_r(x) := \int_{\alpha/\beta}^{x} f_r \, d\lambda$$
$$G(x) := \int_{\alpha/\beta}^{x} g \, d\lambda \;=\; \int_{v}^{v \vee x} g \, d\lambda$$

we have $F_r = F + rG$ and it is clear from (28) that the inequality $F_r \leqslant 0$ will be satisfied on $[v, x_0]$ for $r > 0$ sufficiently small. Since F_r and F coincide on $[\alpha/\beta, v]$, we will get $F_r \leqslant 0$ on I.

It will be easy to impose that $G(x_0) = 0$, to get $F_r(x_0) = F(x_0) = \ln(y_0/\gamma)$.

It will be more tricky to insure that $f_r + \psi_{F_r} \geqslant 0$ (on $[x_1, x_0]$, since it is trivial on $[\alpha/\beta, x_1]$ where $f_r + \psi_{F_r}$ coincides with $f + \psi_F$) and we have to be very careful about the choice of g. The construction of such appropriate g is given as follows.

First note that if $x \in [x_1, x_0]$ is such that $f(x) + \psi_F(x) = 0$, then

$$f(x) + \nu_F(x) = f(x) + \psi_F(x) - \frac{\alpha}{\beta\gamma x}\exp(-F(x)) \leqslant -\epsilon$$

with

$$\epsilon := \min\left\{ \frac{\alpha}{\beta\gamma x}\exp(-F(x)) \;:\; x \in [x_1, x_0] \right\}$$

Denote

$$A := \{x \in [x_1, x_0] \;:\; f(x) + \nu_F(x) \leqslant -\epsilon/2\}$$
$$B := \{x \in [x_1, x_0] \;:\; f(x) + \nu_F(x) > 0\}$$

Let h be a bounded and measurable function defined on B. Consider the two functions ξ and χ given on $[x_1, x_0]$ by

$$\xi(x) := \begin{cases} 1 & \text{, if } x \in A \\ h(x) & \text{, if } x \in B \\ 0 & \text{, otherwise} \end{cases}$$
$$\chi(x) := \begin{cases} \psi_F(x) & \text{, if } x \in A \\ \nu_F(x) & \text{, otherwise} \end{cases}$$

Solve on $[x_1, x_0]$ the weak ODE in G:

$$\begin{cases} G(x_1) = 0 \\ \dot{G} = \chi G + \xi \end{cases} \tag{29}$$

(this is always possible, even with irregular ξ and χ, see (30) below).

Next extend G to $[\alpha/\beta, x_1]$ by imposing that G vanishes there (equivalently, keep solving (29) with $\xi = \chi = 0$ there) and define $g \coloneqq \dot{G}$.

Note that on A, we have $g - \psi_F G = 1$, i.e.

$$g - \gamma^{-1} e^{-F} G = 1$$

It follows that for $r \geqslant 0$ sufficiently small, say $r \in [0, r_0)$, with some $r_0 > 0$, we have

$$g - \gamma^{-1} e^{-F_r} G \geqslant 1/2 \qquad \text{on } A$$

(r_0 depends on h through (29) via a bound on G).

The latter inequality can be written as

$$\partial_r f_r + \partial_r \psi_{F_r} \geqslant \frac{1}{2} \qquad \text{on } A$$

and we deduce that for $r \in [0, r_0)$,

$$f_r + \psi_{F_r} \geqslant 0 \qquad \text{on } A$$

Due to the definition of ϵ and A, $f + \psi_F$ is bounded below by $\epsilon/2$ on $[x_1, x_0]\backslash A$. It follows that up to diminishing r_0, we can insure that $f_r + \psi_{F_r} \geqslant 0$ on $[x_1, x_0]$ for all $r \in [0, r_0)$.

Up to imposing $G(x_0) = 0$, this is the type of perturbations f_r we are to consider.

Note that (29) can be solved explicitly:

$$\forall\ x \in [x_1, x_0], \qquad G(x) = e^{H(x)} \int_{x_1}^{x} e^{-H(u)} \xi(u)\, du \tag{30}$$

where

$$\forall\ x \in [x_1, x_0], \qquad H(x) = \int_{x_1}^{x} \chi(u)\, du \tag{31}$$

Thus the condition $G(x_0) = 0$ writes

$$\int_{x_1}^{x_0} e^{-H} \xi\, d\lambda = 0$$

namely

$$\int_A e^{-H}\, d\lambda + \int_B e^{-H} h\, d\lambda = 0 \tag{32}$$

Once this condition is satisfied, we have $f_r \in \mathcal{D}_\gamma$ for $r \in [0, r_0)$. It leads us to investigate $\mathcal{G}_2(f_r)$. Let us differentiate this quantity at $r = 0$. First note that

$$\partial_r|_{r=0}(f_r + \nu_{F_r}) = g - \nu_F G$$

If $x \in [x_1, x_0]$ is such that $f(x) + \nu_F(x) = 0$, then x does not belong to $A \sqcup B$, so $g - \nu_F G = 0$. Taking into account the definition of B, we obtain by differentiation under the integral

$$\partial_r|_{r=0}\mathcal{G}_2(f_r) = \int_B g - \nu_F G \, d\lambda = \int_B h \, d\lambda$$

Since f is a global minimizer of $\mathcal{G}$ on $\mathcal{D}_\gamma$ and that $\mathcal{G}(f_r) - \mathcal{G}(f) = \mathcal{G}_2(f_r) - \mathcal{G}_2(f)$, we must have

$$\int_B h \, d\lambda \geqslant 0 \tag{33}$$

So we have shown that if h is such that (32) is true, then (33) holds.

To finish the proof, it remains to see that this property implies that $\lambda(D) = 0$.

We proceed by contradiction, assuming $\lambda(D) > 0$. Consider $x_2 \in (x_1, x_0)$ such that

$$\lambda(D_-) = \lambda(D_+) \tag{34}$$

with

$$D_- := D \cap [x_1, x_2), \qquad D_+ := D \cap [x_2, x_0)$$

Find a bounded and measurable function h_0 on B such that

$$\int_B \exp(-H) h_0 \, d\lambda = -\int_A \exp(-H) \, d\lambda \tag{35}$$

Consider next

$$h = h_0 + t\exp(H)\mathbb{1}_{D_-} - t\exp(H)\mathbb{1}_{D_+}$$

with $t \geqslant 0$ to be chosen later.

Due to (34) and (35), (32) holds.

However, as can be seen in (31), H is strictly increasing on B. It follows that

$$\forall\ x' \in D_-, \forall\ x'' \in D_+, \qquad H(x') < H(x'')$$

and as a consequence

$$\int_D h\,d\lambda = \int_D h_0\,d\lambda + t\left(\int_{D_-} e^H\,d\lambda - \int_{D_+} e^H\,d\lambda\right)$$

diverges toward $-\infty$ as t goes to $+\infty$ (the latter assertion follows from $\lambda(D_-) = \lambda(D_+) = \lambda(D)/2 > 0$). However, this leads to a contradiction with (33), so we must have $\lambda(D) = 0$, as desired.

■

We can now come to the

Proof of Theorem 2

From Proposition 6, we deduce that $\mathcal{G}_2(f) = 0$ and so $\mathcal{G}(f) = (\beta/\alpha)\mathcal{G}_1(f)$. Lemma 7 tells us that $\mathcal{G}(f)$ will be minimal if v is as small as possible.

From Proposition 6, we also get that for $x \in [v, x_0]$,

$$f(x) \leqslant -\frac{1}{\gamma}\left(1 - \frac{\alpha}{\beta x}\right)\exp(-F(x)) \tag{36}$$

inequality which can rewritten under the form

$$\frac{d}{dx}\exp(F(x)) \leqslant -\frac{1}{\gamma}\frac{d}{dx}\left(x - \frac{\alpha}{\beta}\ln(x)\right)$$

(where d/dx corresponds to a weak derivative).

Integrating this bound we obtain

$$e^{F(x_0)} - e^{F(v)} \leqslant -\frac{1}{\gamma}\left(x_0 - v - \frac{\alpha}{\beta}\ln(x_0/v)\right)$$

Recalling that $F(v) = 0$ and $F(x_0) = \ln(y_0/\gamma)$, we deduce

$$\frac{y_0}{\gamma} - 1 \leqslant -\frac{1}{\gamma}\left(x_0 - v - \frac{\alpha}{\beta}\ln(x_0/v)\right)$$

i.e.

$$1 - \gamma - \frac{\alpha}{\beta}\ln(x_0) \leqslant v - \frac{\alpha}{\beta}\ln(v)$$

The r.h.s. is a quantity which is increasing with v. Thus we must have $v \geqslant x^*$, where x^* was defined in (25).

The equality $v = x^*$ is realized if and only if (36) is an equality (a.e.), namely

$$f(x) = -\frac{1}{\gamma}\left(1 - \frac{\alpha}{\beta x}\right)\exp(-F(x)) \tag{37}$$

Integrating this equality as before, we get

$$\forall\ x \in [x^*, x_0], \qquad e^{F(x)} - 1 = -\frac{1}{\gamma}\left(x - x^* - \frac{\alpha}{\beta}\ln(x/x^*)\right)$$

Replacing this value of $e^{F(x)}$ in (37), we get the function announced in (26).

■

In the above arguments also enable to compute the minimal value of $\mathcal{G}$ on $\mathcal{D}_\gamma$ (which is also the minimal value of $\mathcal{K}$ on $\mathcal{M}_\gamma$ according to Subsection 5.4).

Corollary 1

$$\begin{aligned}\min_{\mathcal{D}_\gamma} \mathcal{G} = \mathcal{G}(f^*) &= \frac{\beta}{\alpha\gamma}\left(x^* - \frac{\alpha}{\beta}\ln(x^*)\right) + \frac{1}{\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1\right)\\ &= \frac{1}{\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1 + \frac{\beta}{\alpha} - \ln(x_0)\right) - \frac{\beta}{\alpha}\end{aligned}$$

Proof

In the proof of Theorem 2 we have seen that $\mathcal{G}(f^*) = (\beta/\alpha)\mathcal{G}_1(f^*)$. Using the expression given in Lemma 7, where v is replaced by x^*, we get

$$\mathcal{G}_1(f) = \frac{1}{\gamma}\left(x^* - \frac{\alpha}{\beta}\ln(x^*)\right) + \frac{\alpha}{\beta\gamma}\left(\ln\left(\frac{\alpha}{\beta}\right) - 1\right)$$

It remains to take into account the characterization (25) of x^*.

■

Remember this value of $\min_{\mathcal{D}_\gamma} \mathcal{G}$ is only valid under our underlying assumption, otherwise this minimum is simply 0, as it is attained at the *laissez-faire* policy. It should be noted that $\min_{\mathcal{D}_\gamma} \mathcal{G}$ is decreasing with respect to γ, as long as our underlying assumption is satisfied (recall (6) in the main text). This observation will be useful in the next subsection.

5.6 Back to Theorem 1

Finally we come to the proof of Theorem 1. But first we have to return to a rigorous justification of the restriction to (17), which was only heuristically discussed at the end of Subsection 5.1. We will also provide an analytical value for τ_1 and present an extension of Theorem 1 in Remark 9.

In the setting of Subsection 5.1, consider a function $\varphi : [\alpha/\beta, x_0] \to [0, \gamma]$ with $\varphi(\alpha/\beta) < \gamma$, and satisfying (H2), (H3), (H4) and (H5). Due to the fact that φ is right continuous and only jumps upward, this function attains its maximum, say at $x_1 \in [\alpha/\beta, x_0]$. Define a new function $\widetilde{\varphi}$ via

$$\forall\ x \in [\alpha/\beta, x_0], \qquad \widetilde{\varphi}(x) := \begin{cases} \varphi(x_1)\,, \text{ if } x \leqslant x_1 \\ \varphi(x)\quad, \text{ if } x > x_1 \end{cases}$$

Note that $\widetilde{\varphi}$ still satisfies (H2), (H3), (H4) and (H5). Recall the functional $\mathcal{J}$ defined in (18).

Lemma 8 *We have*

$$\mathcal{J}(\widetilde{\varphi}) \leqslant \mathcal{J}(\varphi)$$

and the inequality is strict if $\widetilde{\varphi} \neq \varphi$.

Proof

The argument is similar to the proof of Lemma 6 and is based on the following computation:

$$\begin{aligned}
\int_{\alpha/\beta}^{x_1} L(u, \varphi(u), \varphi'(u))\,du &\geqslant \frac{\beta}{\alpha}\int_{\alpha/\beta}^{x_1} \frac{1+\varphi'(u)}{\varphi(u)} - \frac{\alpha}{\beta u \varphi(u)}\,du \\
&= [\ln(\varphi(u)]_{\alpha/\beta}^{x_1} + \frac{\beta}{\alpha}\int_{\alpha/\beta}^{x_1} \frac{1}{\varphi(u)}\left(1 - \frac{\alpha}{\beta u}\right)du \\
&\geqslant [\ln(\varphi(u)]_{\alpha/\beta}^{x_1} + \frac{\beta}{\alpha\varphi(x_1)}\int_{\alpha/\beta}^{x_1} 1 - \frac{\alpha}{\beta u}\,du \\
&= \ln(\varphi(x_1)/\varphi(\alpha/\beta)) + \frac{\beta}{\alpha\varphi(x_1)}\left[u - \frac{\alpha}{\beta}\ln(u)\right]_{\alpha/\beta}^{x_1} \\
&\geqslant \frac{\beta}{\alpha\varphi(x_1)}\left[u - \frac{\alpha}{\beta}\ln(u)\right]_{\alpha/\beta}^{x_1}
\end{aligned}$$

If we replace φ by $\widetilde{\varphi}$ in the above computations, all the inequalities become equalities, so the last term is in fact equal to

$$\int_{\alpha/\beta}^{x_1} L(u, \widetilde{\varphi}(u), \widetilde{\varphi}'(u))\,du$$

We have furthermore

$$\sum_{u\in(\alpha/\beta,x_1]\,:\,\varphi(u)\neq\varphi(u-)} \ln\left(\frac{\varphi(u)}{\varphi(u-)}\right) \geqslant 0 = \sum_{u\in(\alpha/\beta,x_1]\,:\,\widetilde{\varphi}(u)\neq\widetilde{\varphi}(u-)} \ln\left(\frac{\widetilde{\varphi}(u)}{\widetilde{\varphi}(u-)}\right)$$

and the respective contributions of φ and $\widetilde{\varphi}$ to the costs $\mathcal{J}(\varphi)$ and $\mathcal{J}(\widetilde{\varphi})$ are the same on $(x_1, x_0]$. It follows that

$$\mathcal{J}(\widetilde{\varphi}) \leqslant \mathcal{J}(\varphi)$$

The equality holds if in the above computation of the integral $\int_{\alpha/\beta}^{x_1} L(u, \varphi(u), \varphi'(u))\, du$, all inequalities are equalities and we get that $\varphi(x) = \varphi(x_1)$ for a.e. $x \in (\alpha/\beta, x_1)$, namely $\widetilde{\varphi} = \varphi$.

■

It follows that in the perspective of minimizing $\mathcal{J}$, we can restrict our attention to functions φ attaining their maximum at α/β, i.e. we can replace (H1a) by

(H1c): φ is defined on $[\alpha/\beta, x_0]$, takes values in $[0, \gamma]$, $\varphi(\alpha/\beta) = \max_{[\alpha/\beta, x_0]} \varphi$, $\varphi(x_0) = y_0$ and the left limits of φ are positive.

To go further toward (H1b), let φ a function satisfying (H1c), (H2), (H3), (H4) and (H5), denote $\eta \coloneqq \varphi(\alpha/\beta)$ and assume that $\eta < \gamma$. From the development of Subsections 5.2 to 5.5, we get that

$$\mathcal{J}(\varphi) \geqslant \min_{\mathcal{D}\eta} \mathcal{G}$$

According to Corollary 1 (see also its following paragraph), we have

$$\min_{\mathcal{D}\eta} \mathcal{G} > \min_{\mathcal{D}_\gamma} \mathcal{G}$$

This observation ends up the justification of the replacement of (H1a) by (H1b), relatively to the search of a global minimizer of $\mathcal{L}$.

We can now come to the

Proof of Theorem 1

One direct way would be to check that the procedure described in Subsection 5.1 transform b^* into f^*. There is even a faster way, as it is sufficient to check that

$$\mathcal{C}(b^*) = \mathcal{G}(f^*) \tag{38}$$

To do so, let us come back to (9) and provide the missing computations:

$$\begin{aligned}
\mathcal{C}(b^*) &= \beta \int_{\tau_1}^{\tau_2} \frac{\beta\gamma(\tau_2 - t)}{1 + \beta\gamma(\tau_2 - t)} dt \\
&= \frac{1}{\gamma} \int_0^{\beta\gamma(\tau_2 - \tau_1)} \left(1 - \frac{1}{1+s}\right) ds \\
&= \frac{1}{\gamma} \left[\beta\gamma(\tau_2 - \tau_1) - \ln(1 + \beta\gamma(\tau_2 - \tau_1))\right] \\
&= \frac{\beta}{\alpha\gamma} \left(x(\tau_1) - \frac{\alpha}{\beta} \ln(x(\tau_1))\right) + \frac{1}{\gamma} \left(\ln\left(\frac{\alpha}{\beta}\right) - 1\right) \\
&= \frac{1}{\gamma} \left(\ln\left(\frac{\alpha}{\beta}\right) - 1 + \frac{\beta}{\alpha} - \ln(1 - \varepsilon)\right) - \frac{\beta}{\alpha}
\end{aligned}$$

Comparing this value with that of Corollary 1 (recall that $x^* = x(\tau_1)$), we get the validity of (38).

This argument shows that b^* is a global minimizer of $\mathcal{C}$ over $\mathcal{B}_\gamma$. To see that any global minimizer coincides with b^* on the time interval $[0, \tau_1]$, take into account Remark 2: since f^* is Lipschitzian, there is only one policy leading to f^* (defined on $[\alpha/\beta, x_0]$).

$\blacksquare$

The expression (26) of f^* enables us to give an analytical expression for τ_1 defined in (7). It needs the dilogarithm function Li_2, which is defined on the complex plane via

$$\forall\ z \in \mathbb{C}, \qquad \mathrm{Li}_2(z) = \int_1^{1-z} \frac{\ln(u)}{1-u}\, du$$

Proposition 7

$$\tau_1 = -\frac{1}{\beta}\left(x^* - x_0 + \ln\left(\frac{\beta x^* - \alpha}{\beta x_0 - \alpha}\right)\left[\frac{\alpha}{\beta} - x^* + \frac{\alpha}{\beta}\ln\left(\frac{x^*\beta}{\alpha}\right)\right] \right.$$
$$\left. + \frac{\alpha}{\beta}\left[\mathrm{Li}_2\left(1 - \frac{\beta x^*}{\alpha}\right) - \mathrm{Li}_2\left(1 - \frac{\beta x_0}{\alpha}\right)\right]\right)$$

Proof

On $[0, \tau_1]$, we have $\dot{x} = -\beta x y$, with $y = f^*(x)$. Recalling the expression of f^* given in (26), we get

$$\left(x - x^* - \frac{\alpha}{\beta}\ln\left(\frac{x}{x^*}\right)\right)\frac{\dot{x}}{x - \alpha/\beta} = -\beta$$

By integration, we deduce

$$\begin{aligned}\tau_1 &= -\frac{1}{\beta}\int_{x_0}^{x^*}\left(x - x^* - \frac{\alpha}{\beta}\ln\left(\frac{x}{x^*}\right)\right)\frac{dx}{x - \alpha/\beta}\\ &= -\frac{1}{\beta}(I_1 + I_2)\end{aligned}$$

with

$$I_1 \coloneqq \int_{x_0}^{x^*}\frac{x - x^* + (\alpha/\beta)\ln(x^*)}{x - \alpha/\beta}\, dx\,, \quad I_2 \coloneqq -\frac{\alpha}{\beta}\int_{x_0}^{x^*}\frac{\ln(x)}{x - \alpha/\beta}\, dx$$

Let us compute separately these two integrals:

$$\begin{aligned}I_1 &= \int_{x_0}^{x^*} 1 + \frac{\alpha/\beta - x^* + (\alpha/\beta)\ln(x^*)}{x - \alpha/\beta}\, dx\\ &= x^* - x_0 + [\alpha/\beta - x^* + (\alpha/\beta)\ln(x^*)]\ln\left(\frac{\beta x^* - \alpha}{\beta x_0 - \alpha}\right)\end{aligned}$$

$$
\begin{aligned}
I_2 &= -\int_{x_0}^{x^*} \frac{\ln(x)}{(\beta/\alpha)x-1}\,dx \\
&= -\frac{\alpha}{\beta}\int_{\beta x_0/\alpha}^{\beta x^*/\alpha} \frac{\ln((\alpha/\beta)x)}{x-1}\,dx \\
&= -\frac{\alpha}{\beta}\int_{\beta x_0/\alpha}^{\beta x^*/\alpha} \frac{\ln(\alpha/\beta)}{x-1}+\frac{\ln(x)}{x-1}\,dx \\
&= -\frac{\alpha}{\beta}\ln\left(\frac{\alpha}{\beta}\right)\ln\left(\frac{\beta x^*-\alpha}{\beta x_0-\alpha}\right)+\frac{\alpha}{\beta}\left[\int_1^{\beta x^*/\alpha}\frac{\ln(x)}{1-x}\,dx-\int_1^{\beta x_0-\alpha}\frac{\ln(x)}{1-x}\,dx\right] \\
&= -\frac{\alpha}{\beta}\ln\left(\frac{\alpha}{\beta}\right)\ln\left(\frac{\beta x^*-\alpha}{\beta x_0-\alpha}\right)+\frac{\alpha}{\beta}\left[\mathrm{Li}_2\left(1-\frac{\beta x^*}{\alpha}\right)-\mathrm{Li}_2\left(1-\frac{\beta x_0}{\alpha}\right)\right]
\end{aligned}
$$

Putting together these two expressions, we end up with the announced result.

■

Let us end this subsection with several observations.

Remark 8 Corollary 1 provides in fact a quantitative formulation of our underlying assumption, since it does correspond to $\min_{\mathcal{D}_\gamma}\mathcal{G}>0$ and we get

$$
\gamma<\frac{\alpha}{\beta}\left(\ln\left(\frac{\alpha}{\beta}\right)-1+\frac{\beta}{\alpha}-\ln(x_0)\right)
$$

(in particular the r.h.s. must be positive). We recover (6) in the main text.

□

Remark 9 Consider a cost functional of the form (13), where F coincides with $(\cdot)_+$ on $\mathbb{R}_+$ and is non-negative on $(-\infty,0)$. Then we get that $\widetilde{\mathcal{C}}\geqslant\mathcal{C}$. Note nevertheless that $\widetilde{\mathcal{C}}(b^*)=\mathcal{C}(b^*)$. It follows that b^* is also a global minimizer of $\widetilde{\mathcal{C}}$. In particular if F is positive on $(-\infty,0)$, then b^* is the unique minimizer of $\widetilde{\mathcal{C}}$.

□

Remark 10 Our extension of the optimization problem to measure spaces suggests that the S.I.R. ODE (1) could itself be generalized into

$$
\begin{cases}
dx=-xy\,dB \\
dy=xy\,dB-\alpha y\,dt
\end{cases}
\tag{39}
$$

where B is a Radon signed measure on $\mathbb{R}_+$ (in (1), it is given by $B([0,t]) = \int_0^t b(s)\,ds$, for all $t \geqslant 0$). Equation (39) is to be understood in the Stieltjes sense: for any $t \geqslant 0$,

$$\begin{cases} x(t) = x(0) - \int_{[0,t]} x(s)y(s)\,dB(s) \\ y(t) = y(0) + \int_{[0,t]} x(s)y(s)\,dB(s) - \int_{[0,t]} \alpha y(s)\,ds \end{cases}$$

(where x and y are themselves only right-continuous with left limits, in fact they should be seen as repartition functions of measures). It would be modeling very erratic policies and Theorem 2 would imply that even among them, b^* is a minimizer of the extension of $\mathcal{C}$ similar to $\mathcal{J}$ (as one would guess).

□