WORKING PAPERS

2023-06

Lockdown policies and the dynamics of a pandemic: foresight, rebounds and optimality

Hubert Kempf, stéphane Rossignol

CEPS
Centre for Economics at Paris-Saclay

école normale supérieure paris—saclay
université evry val d'essonne
université PARIS-SACLAY
université PARIS-SACLAY
Maison des Sciences de l'Homme PARIS-SACLAY

Lockdown policies and the dynamics of a pandemic: foresight, rebounds and optimality.

Hubert Kempf* Stéphane Rossignol†

May 16, 2023

Abstract

We study optimal lockdown decisions taken by a policymaker facing a pandemic modelled according to the standard SIR deterministic model. The policymaker trades off the economic costs and the mortality record of the pandemic which depend on the severity and duration of the lockdown. We contrast the shortsightedness versus the farsightedness of the policymaker. Policy-related peaks and rebounds are characterized and explain why a zero-Covid policy is self-defeating. When the ICU constraint is present and the policymaker is shortsighed, there is a large intermediate range of 'values of life' for which the optimal lockdown consists in exactly saturating this constraint. A farsighted policy is not too severe so as to avoid a rebound. The shortest duration consistent with a given health goal is not the less costly. In contrast with the case of shortsightness, a farsighted policy taking into account the ICU constraint sets successive lockdowns of decreasing severity. We address the impact of vaccination on the optimal choice of a lockdown.

Keywords: Pandemic; lockdown policy; Covid-19.

JEL classification: D61; H51; I18.

* *Université Paris-Saclay, ENS Paris-Saclay, CEPS, 91190, Gif-sur-Yvette, France. E-mail :* hubert.kempf@ens-paris-saclay.fr

† *Corresponding author. LED, Université Paris 8 Saint-Denis, 2 rue de la Liberté, 93526 Saint-Denis Cedex, France. E-mail :* strossignol@gmail.com

1 Introduction

The Covid-19 pandemic which started in 2019 has raised the issue of the best way to tackle it and in particular the extent of lockdown policy as the way to stem the dissemination of the virus within a given population. There are two main strategies with respect to lockdown: suppression or mitigation. The suppression strategy (aka "zero Covid") does not compromise and aims at eradicating the pandemic by means of an extreme lockdown policy disregarding the immediate economic costs so generated. The mitigation strategy (aka "living with Covid") aims at finding a compromise between the objective of limiting the number of fatal casualties generated by the pandemic and the objective of mitigating the economic negative consequences of lockdown measures.

In this paper we investigate the mitigation strategy from a theoretical point of view and tackle the determination of optimal lockdown policy decided by a policymaker confronted with a pandemic which is modelled according to the SIR model used in epidemiology.[1] The propositions resulting from this analytical effort provide illuminating insights on the interplay between the dynamics of a pandemic and the policy measures adopted to control it, such as lockdowns. Useful recommendations on the strictness and duration of a lockdown may be derived from these propositions and should be useful for policymakers confronted with a pandemic such as the Covid-19. Results derived from a model which does not overly simplify the dynamics laws of a pandemic by eliminating state variables or making extreme assumptions appear more general and robust than those obtained when simpler models of a pandemic's dynamics are used or when simulation exercises based on calibrated versions of the standard SIR model are performed. On the whole, they support the view that a society must "live with Covid-19" (or any epidemic) through a mitigation lockdown policy rather than attempting to follow a "zero Covid-19" policy (that is, the desire to strongly fight the dissemination of the virus as soon as possible by means of a very strict lockdown policy). The latter policy is unsustainable as the epidemic will surge again sooner or later, once the extreme lockdown measures are lifted, and cannot appear as an optimal policy.[2]

[1] Kermack & McKendrick (1927), Murray (2007).

[2] The covid-19 pandemic proved complex and the simple SIR model does not fully take account of

We first study the dynamics of a pandemic without any lockdown policy. This configuration provides us with a benchmark: our analysis crucially hinges on the result that the peak of a pandemic is attained when the proportion of susceptible agents is equal to the inverse of the "natural" reproduction number.

An active lockdown policy aims at increasing social distancing during a certain time period (duration) by restricting freedom of behavior, including freedom of moves: wearing masks, forbidding certain act, limiting access to some activities. Social distancing limits individual interactions and thus the dissemination of the virus. A lockdown policy therefore consists in choosing the reproduction number of the pandemic over a given period. The policymaker faces a standard dilemma between economic and health objectives. On the one hand, a lockdown inflicts economic losses that the government wants to limit; on the other hand, it reduces social interaction, thus the spreading of the virus and the induced loss of lifes. This dilemma is formalized by a welfare function depending on these two arguments where the relative weight given to the mortality argument captures the implicit "value of life" as assessed by the policymaker.

We contrast a short-term and a long-term perspective. We define a short-term perspective as a single setting of a policy-chosen reproduction number when the time horizon of the policymaker is limited to the duration of the lockdown policy. In contrast, a long-term perspective takes into account the future consequences of a fixed-duration lockdown policy. The first one is likely to be more in line with the behavior of some policymakers observed during the Covid-19 pandemic, insofar as most of them had a limited time horizon in mind. The second one is more consistent with economic rationality. We characterize the dynamics of the pandemic in both cases. In particular we study the impact of the "value of life" parameter on the chosen reproduction number.

For each perspective a series of propositions illuminates the solution of this dilemma. The importance of the various timings related to lockdown, the timing of action, the termination date of the pandemic and their interplay with the marginal impact of the degree of lockdown, is highlighted. Considering the policy duration as given, we show

it. To fully model this precise pandemic would require a complex version of the model, introducing variants, etc. Yet the basic message is not altered when these modelling refinements are included.

that the optimal lockdown degree (the optimal reproduction number chosen by the policymaker) negatively depends on the "value of life" and the fatality parameters.

Using the long-term perspective we show that a post-policy pandemic rebound may happen if the lockdown policy has been too strict and/or its duration too short: it happens when controlled pandemic has not passed its peak, that is when the end-of-policy proportion of susceptible agents is above the inverse of the natural reproduction number. This explains why a zero-Covid policy is self-defeating. Realistically considering that it cannot last forever, there will be a rebound once this policy is lifted and this rebound may lead to a very high number of deaths and a low proportion of end-of-time susceptible agents. Taking into account the possibility of rebounds explains why the optimal reproduction number chosen in the long-term perspective may be higher than the reproduction number chosen in the short-term perspective.

The dynamics of the pandemic may be affected by the existence of hospital capacity constraints. As these constraints impact on the mortality record, not trespassing the hospital capacity has clearly been an objective common to many different governments during the Covid-19 pandemic, eventhough they may have differed in their relative assessment of the various costs of a mitigation policy. Tackling this issue in our theoretical model, we prove that stricly keeping to this limit is indeed consistent with differing values of life and fatality parameters. In a long-term perspective, respecting this limit may imply successive lockdown measures, decreasing in intensity and implying multiple rebounds. Again this is consistent with casual observations of what happened during the Covid-19 pandemic.

When addressing the role of the duration of a lockdown policy, we show that the shortest policy able to reach a given mortality number, implying a stricter lockdown policy (the choice of a lower reproduction number), is not economically the less costly: a less severe lockdown policy extending over a longer period of time generates less economic losses. Relaxing the assumption of a given lockdown duration and endogenizing duration, we prove that there exists an optimal couple of duration and reproduction number solving the policy trade-off. Finally we address the impact of a gradual vaccination policy. The end-of-time susceptible proportion is null in the presence of vaccination.

We prove that a vaccination policy cannot prevent a rebound. However, the intensity of the rebound depends negatively on the rate of vaccination. Over time, a more massive vaccination policy through a higher vaccination rate reduces mortality.

Literature review

The Covid-19 pandemic has generated a flurry of papers aiming at finding the proper lockdown policy either through calibration exercises or analytical approaches.[3].

Focusing on SIR models, Federico & Ferrari (2021) explore the impact of a lockdown assuming that the transmission rate follows a diffusive stochastic process and resort to numerical techniques to investigate the properties of the optimal policy. Caulkins *et al.* (2021) use simulation techniques to investigate optimal lockdown strategies within a SIR model but do not tackle the analytical solution of the problem. Camera & Gioffré (2021) analytically study the economic impact of a sequence of short-lived but extreme lockdowns in a model based on the theory of random matching, which makes explicit how epidemics spread through economic activity. They do not study the extent of lockdown and therefore do not address the issue of optimal lockdowns. Lastly Gonzalez-Eiras & Niepelt (2020b) resort to two simplified and tractable versions of the modified SIR model developed by Bohner *et al.* (2019) in order to study an optimal lockdown policy[4]. The first one neglects the death burden and the distinction between infected and recovered. The second one rests on the assumption of full mortality for the infected and of equal productivity of the susceptible and the infected. The lockdown variable adjusts continuously. These models are calibrated. In a companion paper, Gonzalez-Eiras & Niepelt (2020a) building upon Bailey (1975) simplify the SIR model and restrict it to a single state variable, eliminating the possibility of recovery after infection and thus mortality; this model is calibrated searching for the optimal lockdown trajectory. These various simplifying assumptions drastically reduce the scope of a lockdown policy. The impact of ICU constraints has been addressed by Loertscher & Muir (2021) as

[3]For a survey on the economics of Covid-19, cf. Brodeur *et al.* (2020). For papers using calibration techniques, see for example Eichenbaum, Rebelo, and Trabandt (2020a), Eichenbaum, Rebelo, and Trabandt (2020b), Alvarez, Argente, and Lippi (2020), Piguillem and Shi (2020), Acemoglu et al. (2021)).

[4]They justifiy their choice by writing: "SIR models of various avours feature two endogenous epidemiological state variables; this makes it difficult to embed economic choices in those frameworks without sacricing analytical tractability, transparency, and generality." Gonzalez-Eiras & Niepelt (2020b)

they analyze a lockdown policy in a SIR theoretical model which maximizes output subject to the constraint that contagion is contained so that total hospitalizations do not exceed the health capacity constraint both with a homogeneous population and an heteregenous population. Miclo *et al.* (2022) assume that the ICU cannot be overpassed (contrarily to what we assume) and obtain a lockdown policy which is decreasing over a limited interval of time. Lastly, several papers use alternative models such as SIS models to study lockdown policies in the context of endemic diseases (See Bosi *et al.* (2021), Atolia *et al.* (2021), La Torre *et al.* (2021)).

2 The model.

We consider a closed society which is affected by a pandemic. There is no shock in this setting and according to the standard deterministic SIR model, the dynamics of the pandemic is given by the following set of equations for any $t \in \mathbb{R}$:

$$\frac{ds}{dt} = -\beta_0 i(t) s(t) \tag{1}$$

$$\frac{di}{dt} = \beta_0 i(t) s(t) - \gamma i(t) \tag{2}$$

$$\frac{dr}{dt} = \gamma i(t) \tag{3}$$

where $s(t)$ is the proportion of individuals susceptible of being infected in the population at a given instant t, $i(t)$ the proportion in the population of infected individuals and $r(t)$ the proportion of removed (or recovered) individuals, with $s(t) + i(t) + r(t) = 1$ at every instant t.[5] We set $\mathcal{R}_0 \equiv \frac{\beta_0}{\gamma}$ the natural (initial) reproduction number.[6] The parameter β_0 refers to social interactions and controls the spreading of the pandemic as it affects the variation of the size of the "susceptible" agents. It is specific to a pandemic and captures the physical impact of social interactions within society on the

[5] For a brief presentation of the SIR model, see Avery *et al.* (2020).

[6] aka "basic" reproduction number. As noted by Avery *et al.* (2020), this number "embodies both the underlying biological ability of the pathogen to jump from person to person in various types of interactions as well as the number of interactions of each type that people have in the ordinary course of their daily lives" (p.84) and may partially result from self-interested voluntary measures of social distancing.

dynamics of the pandemic. This structural parameter is related to social habits and collective mores. The parameter γ is positive and corresponds to the rate of infected individuals recovering in a given unit of time. "Recovering" means either returning to perfect health or death. A fraction δ of the "recovered" dies from the pandemic. This parameter δ is the infection fatality rate. It is assumed here that once someone recovers from the virus, he or she is never infected again: recovery is permanent[7]. We shall return to this point in the conclusion. This model has been used by Rowthorn and Maciejovski (2020) for simulation exercises related to the Covid-19 pandemic. Here we shall analytically solve it, under various lockdown policy configurations. The "natural" reproduction number may capture the rearrangement of the production process such as teleworking, and more generally, the changes of voluntary behavior induced by the advent of the pandemic. We abstract from investigating this issue.

We first study the dynamics of the pandemic when there is no lockdown policy imposed by a public authority.[8] The pandemic develops freely according to the reproduction number $\mathcal{R}_0$ and dies away when a sufficient fraction of the population has recovered and does not transmit the virus any more. This policy has been dubbed a "collective immunity" strategy. In this case, the pandemic eventually vanishes through herd immunity: the number of recovered people is large enough so that the virus does not find a significant number of "susceptible" individuals and does not reproduce itself anymore. We shall use this configuration as a benchmark against which the various lockdown policies may be compared. We assume $\mathcal{R}_0 > 1$, otherwise the pandemic cannot start. Following Krœger and Schlickeiser (2020), we assume the following boundary conditions: $s(-\infty) = 1$, $i(-\infty) = 0$ and $r(-\infty) = 0$.

Collective immunity is reached when the pandemic is extinct. Given the deterministic nature of the model, it is reached at the "end of time". Formally it is defined as $(s_\infty, 0, r_\infty)$: there are no more infected people, the proportion of susceptible s_∞ is positive and the proportion of recovered r_∞ is equal to $1 - s_\infty$. We refer to s_∞ as

[7]This assumption doen not properly reflect the Covid-19 pandemic. Yet the model captures its basic characteristics when new variants are neglected. Variants can be introduced in the model at the cost of increasing complexity.

[8]Boris Johnson, UK prime minister: "We should all basically just go about our normal daily lives.", 30 March 2020.

the end-of-pandemic (or "terminal") susceptible proportion. It corresponds to herd immunity.

Assuming that the reproduction number does not vary over time, we have the following proposition on the dynamics of the SIR model, that is the "natural" laws of motion of the three key variables $s(t), i(t)$ and $r(t)$ of the pandemic (when driven by $\mathcal{R}_0$)[9].

Proposition 1. (i) *The dynamics of the pandemic for $t \in \mathbb{R}$ is given by*

$$r(t) = -\frac{1}{\mathcal{R}_0} \ln s(t) \tag{4}$$

$$i(t) = 1 - s(t) + \frac{1}{\mathcal{R}_0} \ln s(t) \tag{5}$$

$$\int_{s(t)}^{s(0)} \frac{1}{\beta_0 s \left[1 - s + \frac{1}{\mathcal{R}_0} \ln s\right]} ds = t \ . \tag{6}$$

$t \mapsto s(t)$ is a decreasing function, and $t \mapsto r(t)$ is an increasing function.

(ii) *The proportion of infected $i(t)$ is first increasing, then decreasing. It is maximal when $s(t) = \frac{1}{\mathcal{R}_0}$ and equal to $i_{\max} = 1 - \frac{1}{\mathcal{R}_0}[1 + \ln(\mathcal{R}_0)]$.*

(iii) *At the end of the pandemic, we have $(s, i, r) = (s_\infty, 0, r_\infty)$, with $r_\infty = 1 - s_\infty$ and s_∞ given by*

$$\mathcal{R}_0 = -\frac{\ln(s_\infty)}{1 - s_\infty}, 0 < s_\infty < 1. \tag{7}$$

Proof. See Appendix A.1. □

This proposition explicits the functional dynamics of the three variables of interest, the numbers of susceptible, infected and recovered people from the SIR model and establishes some properties of these dynamics. *(i)* details the interdependence between the dynamics of these variables. Eq.(4) shows that the proportion of recovered is a decreasing function of the proportion of susceptible; Eq.(5) shows that the proportion of

[9] For a similar result see Harko *et al.* (2014), p.187.

infected is a non-monotone function of the proportion of susceptible. Eq.(6) shows that the proportion of susceptible varies with time depending on β_0 and γ. As expected, the proportion of susceptible decreases with time and the proportion of recovered increases with time.

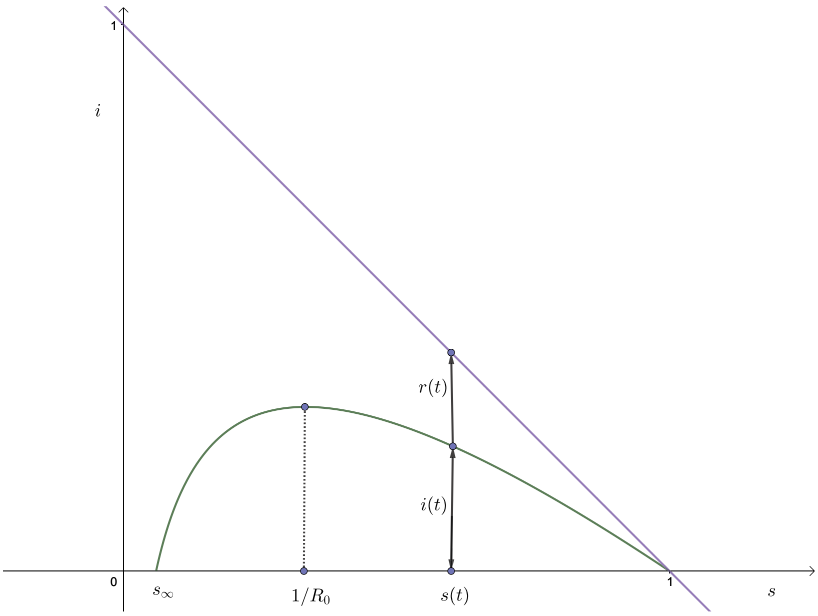

Figure 1: Variations of i with respect to s in the plane (s, i)

Figure 1 gives a graphical representation of the dynamics of the pandemic in the plane (s, i). At the very beginning (when $t \to -\infty$) the point $(s(t), i(t))$ is close to $(1; 0)$. When t increases $s(t)$ decreases, and $i(t)$ first increases then decreases. At the end of the pandemic (when $t \to +\infty$) the point $(s(t), i(t))$ is close to $(s_\infty; 0)$.

(ii) proves that the relation between the proportion of infected and the proportion of susceptible is non-monotone and reaches a maximum when $s(t)$ is equal to $\frac{1}{\mathcal{R}_0}$. The higher the natural reproduction number, the higher the peak of the pandemic. As we shall see later, the reproduction number plays a critical role in the dynamics of the pandemic once a lockdown policy is put in place. Once having recovered, one cannot be infected again. Thus the evolution of the proportion of infected depends on both the reproduction number and the evolution of the pool of susceptible. At the beginning of the pandemic for a given $\mathcal{R}_0$, the pool of susceptible is large and the number of infections increases since newly infected agents easily spread the virus into a large population of

susceptible. But the pool (hence the proportion) of susceptible necessarily declines once the pandemic has started. Over time newly infected spread the virus into a smaller and smaller population even with the same reproduction number. This negative effect curbs down the rate of new infections and the proportion of infected declines. Thus the proportion of infected is a non-monotonous function of the proportion of susceptible. The maximum rate is achieved when the proportion of susceptible is just equal to the inverse of the natural reproduction number. The higher this number, the lower the susceptible proportion after which the number of infected people starts declining. At $s(t) = 1/\mathcal{R}_0$, there is a "herd immunity" threshold: the expected number of people that a newly infected person will directly infect is equal to 1.

Finally *(iii)* characterizes the final (end-of-time) proportion of susceptible denoted by s_∞, where s_∞ is a decreasing function of $\mathcal{R}_0$: the higher is $\mathcal{R}_0$, the more violent and deadly is the pandemic. All this is consistent with intuition and observations of actual pandemics.

3 Shortsighted lockdown policies.

We now assume that the "government" ruling this society (*aka* a public authority able to impose some lockdown policy) is able to act at a given period t_0.[10] Social interactions, namely "social distancing", can be modified by law and public punishments (fines, etc.) decided by the government. Here we assume that a lockdown policy amounts to replace the parameter β_0 by β lower than β_0 over a limited and definite interval of time: it is supposed to be put in place in t_0 and ends in period T. The policy duration is therefore the interval (t_0, T). It negatively affects the reproduction number which becomes $\mathcal{R} = \frac{\beta}{\gamma}$ instead of $\mathcal{R}_0 = \frac{\beta_0}{\gamma}$ during this interval. After T, the reproduction number is back to its "natural" value. The stricter the lockdown policy, the lower β and the lower the reproduction number. It is thus controlled by the public authority and represents its

[10] This is in line with Schlickeiser and Kröger (2021). Schlickeiser and Kröger assume that the pandemic, while in infancy, is taken into account from some given date denoted by "0", equivalent to our t_0: at this date there is a small but positive number of infected people. This amounts to consider the pandemic at "its infancy". The susceptible proportion is close to 1 and the infection number is close to 0.

lockdown policy instrument. We define a "short-term" policy by the following property: the consequences of this policy after T until the end of the pandemic are not taken into account by the policymaker. Said in other words, this policymaker is "short-sighted".

3.1 The dynamics of the pandemic with lockdown policy.

Assume that a decision of lockdown or general social distancing is taken at date $t_0 \in \mathbb{R}$, before the epidemic attains its maximum, i.e. $s(t_0) > \frac{1}{\mathcal{R}_0}$. A lockdown policy decided in instant t_0 consists in setting a new reproduction number $\mathcal{R}$ which applies to the period from t_0 onwards up to T. There are two periods to be distinguished in the evolution of the pandemic until the end of the lockdown:

– Before t_0, the reproduction number is $\mathcal{R}_0$ and the dynamics is governed by Eqs. (1)-(3).

– In the interval between t_0 and T, the reproduction number is $\mathcal{R} = \beta/\gamma$, with $\mathcal{R} < \mathcal{R}_0$, the dynamic system capturing the dynamics of the pandemic after t_0 becomes

$$\frac{ds}{dt} = -\beta i(t) s(t) \tag{8}$$

$$\frac{di}{dt} = \beta i(t) s(t) - \gamma i(t) \tag{9}$$

$$\frac{dr}{dt} = \gamma i(t) \tag{10}$$

with $\beta < \beta_0$.

In short, the dynamics of the pandemic starts as in Figure 1, being governed by $\mathcal{R}_0$. When $s(t_0)$ is reached, it bifurcates as it is then governed by $\mathcal{R}$, up to T. $s(t)$, $i(t)$, $r(t)$ will be denoted by $s_{\mathcal{R}}(t)$, $i_{\mathcal{R}}(t)$, $r_{\mathcal{R}}(t)$ for $t \in]t_0, T]$ when it will be necessary to stress their dependency to $\mathcal{R}$. The laws of motion of the three variables $s(t), i(t)$ and $r(t)$ under a temporary lockdown policy are specified in the following

Proposition 2. *Assume that the reproduction number is $\mathcal{R}_0 = \frac{\beta_0}{\gamma}$ on $t < t_0$ and $\mathcal{R} = \frac{\beta}{\gamma}$ on $t \in [t_0; T]$, with $\mathcal{R} < \mathcal{R}_0$, and $s(t_0) > \frac{1}{\mathcal{R}_0}$, i.e. t_0 is before the natural peak of the epidemic is reached. Assume also that $s(T) < \frac{1}{\mathcal{R}}$, i.e. T is after the lockdown-related peak.*

(i) The dynamics of the pandemic until the end of the lockdown is given by the following sets of equations:

For $t < t_0$, the dynamics is given by eqs. (4)-(6).

For $t \in [t_0, T]$,

$$r(t) = r(t_0) + \frac{1}{\mathcal{R}} \ln s(t_0) - \frac{1}{\mathcal{R}} \ln s(t) \tag{11}$$

$$i(t) = i(t_0) + s(t_0) - s(t) + \frac{1}{\mathcal{R}} \ln s(t) - \frac{1}{\mathcal{R}} \ln s(t_0) \tag{12}$$

$$\int_{s(t)}^{s(t_0)} \frac{1}{\beta s \left[i(t_0) + s(t_0) - s + \frac{1}{\mathcal{R}} \ln s - \frac{1}{\mathcal{R}} \ln s(t_0) \right]} ds = t - t_0. \tag{13}$$

$t \mapsto s(t)$ is a decreasing function, and $t \mapsto r(t)$ is an increasing function.

(ii) If $s(t_0) \geq \frac{1}{\mathcal{R}}$, then on $t \leq T$ the proportion of infected individuals $i(t)$ is first increasing, then decreasing. It is maximal when $s(t) = \frac{1}{\mathcal{R}}$.

If $s(t_0) < \frac{1}{\mathcal{R}}$, then $i(t)$ is maximal at $t = t_0$. It is decreasing on $t \in [t_0, T]$.

(iii) For $\mathcal{R}' > \mathcal{R}$, the curve $(s_{\mathcal{R}'}(t), i_{\mathcal{R}'}(t))_{t \geq t_0}$ is strictly above $(s_{\mathcal{R}}(t), i_{\mathcal{R}}(t))_{t \geq t_0}$ in the plane (s, i), except a unique common point at $(s(t_0), i(t_0))$.

$\mathcal{R} \mapsto s(t) = s_{\mathcal{R}}(t)$ is a decreasing function of $\mathcal{R}$.

Proof. See Appendix A.2. □

(i) shows the dynamic impact of the lockdown policy. It obviously has no impact over the period before t_0. Given the number of susceptible $s(t_0)$, the dynamics of the pandemic during the policy period is governed by a similar set of equations than in the previous period but depending on $\mathcal{R}$. Figure 2 gives a graphical representation of the dynamics of the pandemic with a lockdown of reproduction number $\mathcal{R}$, beginning at t_0 and ending at T.

(ii) tackles the impact of the policy reproduction number $\mathcal{R}$ over the interval $[t_0, T]$ on the dynamics of $i(t)$. It is either non-monotone or decreasing, depending on $\mathcal{R}$. If, given the initial number of susceptible in t_0, the reproduction number $\mathcal{R}$ is not too low, $i(t)$ is first increasing, then decreasing. The explanation is similar as the

one given above for the dynamics driven by $\mathcal{R}_0$. If $\mathcal{R}$ is sufficiently high, higher than $1/s\left(t_0\right)$, the contamination policy is too lenient to allow for an immediate decrease in the number of infected. This one first continues to grow and reaches a maximum when the susceptible proportion attains $1/\mathcal{R}$. It is important to note that this value does not depend on the other parameters of the model, in particular depends neither on t_0 nor on T. If, on the other hand, the policy number $\mathcal{R}$ is low enough (lower than $1/s\left(t_0\right)$), $i(t)$ is immediately decreasing. The social distancing is strong enough, $\mathcal{R}$ being sufficient small, to overcome the existence of a relatively large pool of agents susceptible of becoming infected and thus the easy spreading of the pandemic, so as to trigger an immediate decrease in the number of infected. The lower $\mathcal{R}$, the steeper the slope of the increasing function relating the proportion of susceptible and the proportion of infected. Such a strong lockdown policy can be dubbed a "(almost-)zero Covid policy": the policymaker wants to see the pandemic decreasing immediately and forcefully, despite the immediate negative economic consequences of this policy with the hope of getting rid of the pandemic and be able to resume a "normal" life with no lockdown and a low number of deaths. Notice that the higher is t_0, the lower is $s\left(t_0\right)$, the initial pool of susceptible, making the spreading of the pandemic more difficult.

According to *(iii)*, the number of susceptible in a given instant is a decreasing function of $\mathcal{R}$. This is consistent with intuition: the pool of susceptible agents decreases more rapidly when the reproduction number is high as the virus spreads more rapidly: the proportion of susceptible $s_{\mathcal{R}'}\left(t\right)$ related to $\mathcal{R}'$ is consistently below $s_{\mathcal{R}}\left(t\right)$ when $\mathcal{R}' > \mathcal{R}$. Consequently $s_{\mathcal{R}}\left(T\right) > s_{\mathcal{R}'}\left(T\right)$: a higher policy-chosen reproduction number leads to a lower end-of-lockdown proportion of susceptible and a higher mortality record. Furthermore there is no "catching-up" effect: a lenient lockdown policy (corresponding to $\mathcal{R}'$) cannot reach a combination (s, i) reached by a stricter policy (corresponding to $\mathcal{R}$). Since the curves $\left(s_{\mathcal{R}'}\left(T\right), i_{\mathcal{R}'}\left(T\right)\right)_{t\geq t_0}$ and $\left(s_{\mathcal{R}}\left(T\right), i_{\mathcal{R}}\left(T\right)\right)_{t\geq t_0}$ with $\mathcal{R} \neq \mathcal{R}'$ have a unique common point at $\left(s\left(t_0\right), i\left(t_0\right)\right)$, a given pair (s, i) can be attained by one lockdown policy only and at one instant T only.

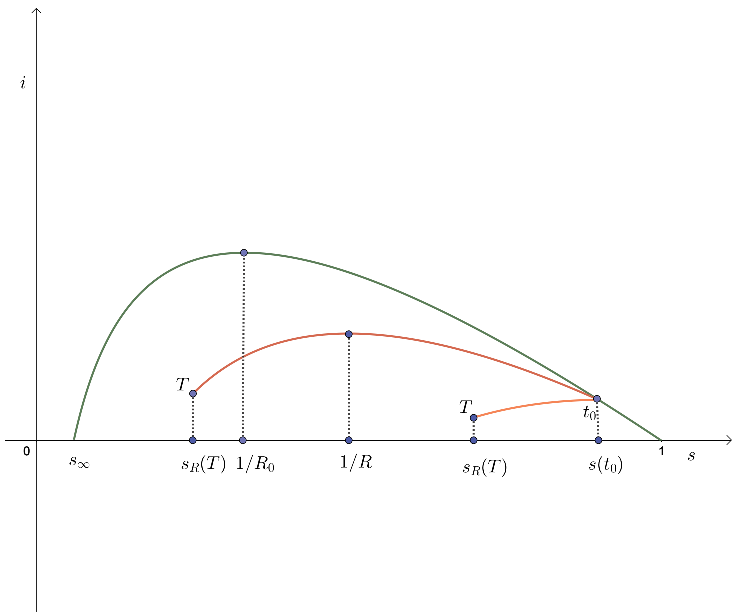

Figure 2: Variations of i with respect to s in the plane (s, i) with decision at date t_0

3.2 Shortsighted optimal lockdown policies

The previous section allowed us to understand how a lockdown policy aiming at reducing the reproduction number of the pandemic affects its dynamics and the proportion of recovered individuals at any period. We now tackle the dilemma facing the policymaker: a lockdown policy mitigates the health consequences of the pandemic but negatively affects the economy by restricting the social interactions between economic agents. In other words, there is a trade-off between economic and health objectives and a responsible government must solve this trade-off: how to optimally live with a pandemic?

In order to investigate the issue of defining the optimal lockdown policy in the presence of economic costs, we take for given the starting instant t_0 and ending instant T. Fixing the duration of a lockdown policy makes sense. The policymaker may be constrained by political conditions such as the term of her mandate or subject to economic pressures to end the policy before a certain date. She may be unable to act after a certain period for constitutional reasons or she may be unable or unwilling to anticipate the entire future of the pandemic. Above all, it corresponds to the very plausible case when the policymaker is myopic and does not foresee beyond a given date. It is also true that the dynamics of a pandemic is such that the most of the contamination happens in a very short period and not much is lost by fixing this period. The marginal gains of increasing the policy period beyond a plausible duration may be limited. Lastly addressing the optimal lockdown problem with two instruments in a non-linear system such as the one governing a pandemic is quite complex and likely to obscure the picture for little analytical gains. We will address this issue in Section 5.1, once important results, easy to understand and empirically relevant, have been obtained. In brief, it is reasonable to first reason with T given and focus on the lockdown policy-determined reproduction number: it is the parameter which attracts the most attention and is critical in the dynamics of the pandemic as proven above.

We will consider a shortsighted government which limits its time horizon to the end of its lockdown policy T. In the following section we will consider a farsighted government which considers the entire future and therefore the after-lockdown dynamics of the pandemic, knowing that her lockdown policy affects this dynamics. This distinction is relevant. When observing the behavior of governments during the Covid-19 pandemic, one notices that policymakers regularly adopted a position of advocating before public opinion that the decided lockdown policy put in place would be sufficient to return to "normal life" and the economy would soon "pass the corner". Public opinion too seems to be oblivious of the long-term duration of a pandemic and such an attitude affects the decision process of the policymaker.

3.2.1 Optimal lockdown policy with no hospital constraints.

We first consider the simple case where the pandemic develops without meeting any other barrier than the social distancing measure adopted by the government. In particular, there is no health constraint such as the availability of properly equipped hospital beds and there is no change in the therapy against it. The health system is structurally able to "properly" deal with patients. Formally, we assume that the infection fatality rate $\delta > 0$ is constant, independent of circumstances and exogenously given.

The cumulated mortality rate in the population is $m_T = \delta(r(T)+i(T)) = \delta(1-s(T))$ where $1 - s(T)$ denotes the fraction of the population which eventually recovers (or dies) from the pandemic after having being infected before T. Denoting by N the total population, the final number of deaths due to contaminations before T is $M_T = Nm_T$. A lockdown policy consists in adopting a series of compulsory social distancing measures so as to affect the reproduction number of the pandemic at any given period of time. Assuming that the inverse link between social distancing and this number is deterministic amounts to say that the government controls the reproduction number $\mathcal{R}$. Again we assume that a decision of lockdown or general social distancing is taken at date t_0, before the epidemic attains its maximum, i.e., $s(t_0) > \frac{1}{\mathcal{R}_0}$ and consists in setting a single reproduction rate which applies to any period from t_0 onwards: $\mathcal{R}(t) = \mathcal{R}, \forall t \in [t_0; T]$. As we abstract from any shock, including on the biological characteristics of the virus, and any change in government, this assumption is reasonable. M_T depends on $\mathcal{R}$, so we write $M_T = M_T(\mathcal{R})$.

It is reasonable to assume that a benevolent policymaker is willing to limit the amount of casualties of the pandemic by means of an active control of the reproduction number. Given her shortsightedness, the policymaker takes into account the cumulated mortality $M_T(\mathcal{R})$ due to contaminations occurring up to T. This mortality record is affected by her choice of $\mathcal{R}$. It is defined as

$$M_T(\mathcal{R}) \equiv \delta N (i_{\mathcal{R}}(T) + r_{\mathcal{R}}(T)) = \delta N (1 - s_{\mathcal{R}}(T))$$

A lockdown policy also incurs economic losses: social distancing affects both the supply and the demand sides of the economy. On the one hand, some firms cannot open, some workers cannot work as efficiently as in "normal" times, or are out of work. On the whole the capacity to produce goods and services is impaired. On the other hand, some goods are not demanded because the social distancing prevent their consumption. Consumption and investment are depressed and the well-being of individuals is negatively affected by the desire to control the pandemic by means of social distancing measures. The goverment trades off the economic losses and the sanitary adverse consequences of the pandemic. Formally the welfare function of the decision-maker is assumed to be

$$V_T(\mathcal{R}) = (T - t_0)\, y(\mathcal{R}) - \lambda M_T(\mathcal{R}) = (T - t_0)\, y(\mathcal{R}) - \lambda \delta N \left(1 - s_{\mathcal{R}}(T)\right) \tag{14}$$

where $y(\mathcal{R})$ is an aggregate output index such as GDP per unit of time, depending on the lockdown parameter $\mathcal{R}$, and $\lambda \in [0; \infty)$ is the weight put on mortality relative to economic activity (as measured by y). It captures the "value of life" as assessed by the policymaker relative to the economic target of boosting the economy[11]. The function $y(\mathcal{R})$, in addition to the reduction of activity directly due to the reduction in mobility, may capture the change in the production process decided when the lockdown is imposed. Here we abstract from distinguishing the various channels which shape the relationship between y and $\mathcal{R}$ and directly reason on the reduced form given by $y(\mathcal{R})$.

During the pandemic, the economic index y is an increasing concave function of $\mathcal{R}$: the more lenient the lockdown policy, the higher the aggregate output index. We assume that the marginal gain of relaxing the lockdown policy (increasing the reproduction number) is diminishing with this number and an immediate impact of $\mathcal{R}$ on y without lagged effects. The lowest value of y is obtained when the social distancing is at its maximum, that is, when $\mathcal{R}$ is equal to 0: economic losses are at their maximum. We consider that $y(0)$ is equal to 0 since then all activities, including productive ones, are frozen[12]. Thus $y(\mathcal{R})$ can be seen as the gain from relaxing the lockdown parameter from 0 to $\mathcal{R}$. When $\mathcal{R}$ increases above 0 (the social distancing constraint is relaxed), the economy partially recovers and losses are reduced. We assume y is a strongly concave function of $\mathcal{R}$, and even that $y'(0) = +\infty$. When $\mathcal{R}$ equals $\mathcal{R}_0$, this corresponds to the "hands-off" regime characterized by the natural reproduction number and there are no economic losses.

In the sequel, when we want to get a more precise analytical understanding of the trade-off between the (negative) economic and (positive) sanitary consequences of a

[11] This welfare function, displaying an economic argument and a "loss of life" one, is similar to the functions used by Acemoglu et al. (2021) and Rowthorn and Maciejowski (2020).

[12] We neglect household productions.

lockdown decision, we shall use the following specification:

$$y(\mathcal{R}) = A\mathcal{R}^{\alpha} \tag{15}$$

with $A > 0$ and $0 < \alpha < 1$.

We assume that $V_T(\mathcal{R})$ is a concave function of $\mathcal{R}$ (i.e. y is sufficiently concave relative to $M_T(\mathcal{R})$ to have $V_T(\mathcal{R})$ concave for $\mathcal{R} \leq \mathcal{R}_0$). The first term in (14) corresponds to the cumulative economic effect of the decision $\mathcal{R}$ from t_0 up to T. The second term corresponds to the health cost of the pandemic, measured in the total of deaths due to the pandemic up to date T. The optimal policy consists in choosing $\mathcal{R}^{opt}$ maximizing the welfare function, that is

$$\mathcal{R}^{opt} = \arg\max_{\mathcal{R} \leq \mathcal{R}_0} V_T(\mathcal{R}) \tag{16}$$

This optimal value generates a mortality record, an economic loss and thus a given level of welfare. On the whole, the global configuration with optimal policy is characterized by $(\mathcal{R}^{opt}, M_T(\mathcal{R}^{opt}), y(\mathcal{R}^{opt}), V_T(\mathcal{R}^{opt}))$. We could equivalently say that the decision-maker wants to minimize the short-term loss due to the pandemic and the lockdown

$$L_T(\mathcal{R}) = (T - t_0)(y(\mathcal{R}_0) - y(\mathcal{R})) + \lambda M_T(\mathcal{R}). \tag{17}$$

We are able to offer the following

Proposition 3.
Setting $\lambda_0 = \frac{(T-t_0)y'(\mathcal{R}_0)}{-\delta N \left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\mathcal{R}_0}} \geq 0$ *we have:*

(i) $\mathcal{R}^{opt}$ *is equal to* $\mathcal{R}_0$ *for* $\lambda \in [0; \lambda_0]$ *and is a decreasing function of* λ *for* $\lambda \in [\lambda_0; \infty)$, *with* $\lim_{\lambda \to \infty} \mathcal{R}^{opt} = 0$.

(ii) *For* $\lambda \geq \lambda_0$, $\mathcal{R}^{opt}$ *is a decreasing function of* t_0 *if* $y'(\mathcal{R}^{opt}) > \lambda \delta N \left(\frac{\partial^2 s_{\mathcal{R}}(T)}{\partial \mathcal{R} \partial t_0}\right)_{\mathcal{R}=\mathcal{R}^{opt}}$ *and an increasing function of* t_0 *otherwise.*

(iii) *For* $\lambda \geq \lambda_0$, $\mathcal{R}^{opt}$ *is an increasing function of* T *if* $y'(\mathcal{R}^{opt}) > -\lambda \delta N \left(\frac{\partial^2 s_{\mathcal{R}}(T)}{\partial \mathcal{R} \partial T}\right)_{\mathcal{R}=\mathcal{R}^{opt}}$ *and a decreasing function of* T *otherwise.*

(iv) *When (15) applies, for $\lambda \geq \lambda_0$, $\mathcal{R}^{opt}$ is always an increasing function of A and is an increasing function of α on $1 + \alpha \ln(\mathcal{R}^{opt}) > 0$, i.e. on $\mathcal{R}^{opt} > e^{-1/\alpha}$.*

Proof. See Appendix A.3. □

(i) states that the optimal lockdown policy consists in having a low reproduction number when the "value of life" weight in the welfare function is sufficiently high. When the value of life weight is close to 0, the lockdown optimal policy is the laissez-faire policy, corresponding to no imposition of social distancing measures. This comes from the fact that the marginal loss of imposing a stricter policy is high while the marginal welfare gain in terms of saved life is negligible, given that λ is low. When the weight is sufficiently high, the optimal value of $\mathcal{R}^{opt}$ starts decreasing. The more the policymaker cares about the mortality record of the pandemic relative to the adverse economic consequences of social distancing (the higher λ), the stricter is the chosen lockdown. When λ tends to infinity, $\mathcal{R}^{opt}$ tends to 0. When the value of life is arbitrarily large, the best policy is to neglect the economic costs for the sake of saving lifes by imposing an extremely severe lockdown. By a similar reasoning, notice that $\mathcal{R}^{opt}$ is a decreasing function of the fatality rate[13] δ.[14]

According to *(ii)*, the optimal decision depends on the time of action t_0. It is not possible to state that it always increases or decreases monotonically with this date. Delaying action (increasing t_0) has two effects on the optimal value of the lockdown parameter. First, if the time of action is delayed from t_0 to $t_0' > t_0$, the impact of the natural reproduction number $\mathcal{R}_0$ is increased and the number of susceptible at the time of action is significantly reduced because during the interval $(t_0' - t_0)$ the reproduction number is $\mathcal{R}_0 > \mathcal{R}^{opt}$, thus $s(t_0') < s(t_0)$. Keeping constant the optimal reproduction number associated to t_0 is not adequate. It becomes economically too costly because it amounts to lockdown the entire population to protect a reduced pool of susceptible. This effect induces an increase in the optimal reproduction number. But a second effect

[13] The Ebola epidemics which is highly deadly leads to the most extreme lockdown measures.

[14] Rowthorn and Maciejovski (2020), based on simulation of the model, find that a 10-week lockdown is optimal if the value of life for Covid-19 victims exceeds £10m.

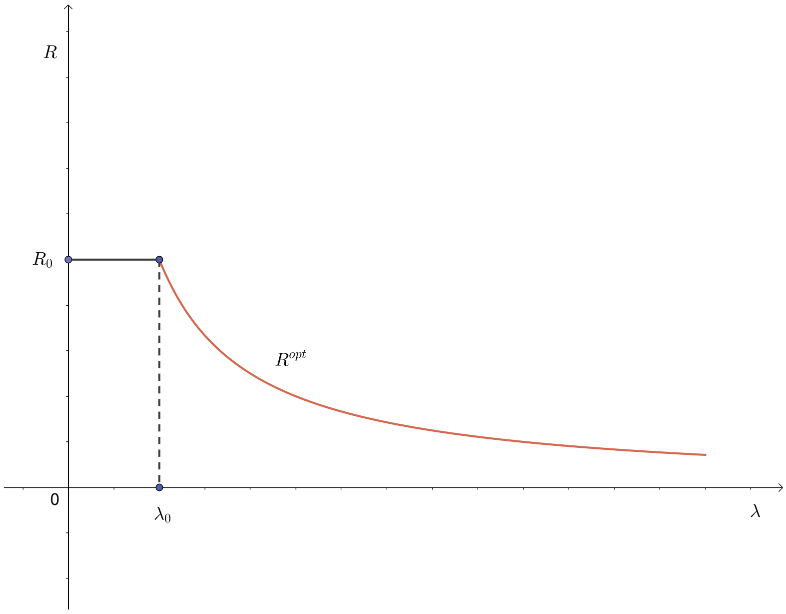

Figure 3: Variations of $\mathcal{R}^{opt}$ with respect to λ

is at work. If $t_0' > t_0$, the duration of lockdown is reduced: $(T - t_0') < (T - t_0)$. The lockdown policy is effective during a shorter period. This leads the policymaker to tighten the reproduction number in order to compensate this shorter period of action on the health (mortality) record. This latter effect counters the former one and this explains the ambiguity result stated in *(ii)*. Given these opposed effects, the impact of a higher lag in action on the optimal decision is ambiguous. Postponing action may lead either to an increase in the lockdown effort or a reduction. The second effect interacts with the value of $\mathcal{R}$ since the function $s(T) = s_{\mathcal{R}}(T)$ derived from (13) is highly non linear. Thus the impact on mortality of a variation of $\mathcal{R}^{opt}$ coincidental on the increase in t_0 depends on the cross-derivative of the function $s_{\mathcal{R}}(T)$ evaluated at this time of decision for this precise value of $\mathcal{R}$.

(iii) gives similar results for the impact of T on the optimal choice of the policy rate. It has an ambiguous impact because two opposite effects are at work. A lengthier policy duration from T to $T' > T$ is beneficial in terms of control of the pandemic and generates a lower mortality ratio, assuming $\mathcal{R}^{opt}$ associated to T to be kept constant. This pleads for a relaxation of the policy-chosen reproduction number for economic reasons: mitigating the adverse negative impact of an increased duration by reducing

the strictness of the lockdown, given the positive impact on mortality of the increased duration. However increasing the term of a short-term lockdown policy means that the policymaker lengthens its planning horizon and therefore takes into consideration a higher number of infected agents ceteris paribus. This leads then to strengthen its lockdown policy (reducing the reproduction number). Therefore there are two contradictory effects on welfare of an increase in the reproduction number. Because of the non-linearity of the function $s_{\mathcal{R}}(T)$, it is impossible for one effect to systematically dominate the other one given that the cross-derivatives of this function matter. If they are equal to (close to) 0, an extension of the policy period by increasing T or by decreasing t_0 leads to an increase of $\mathcal{R}^{opt}$: the marginal economic effect is sufficiently strong so as to allow a relaxing of the lockdown rate for economic gains.

(iv) assesses the sensitivity of the optimal policy rate to the parameters of the economic function (15). A being a multiplier, its marginal effect is constant for $\mathcal{R}$ given. If it increases, the marginal gain of a relaxing of the social distancing measures increases and this leads to an increase in the optimal policy reproduction number. On the contrary an increase in the parameter α has an impact on this marginal gain which varies with the value of α. We face a non-linearity issue again.

3.2.2 Optimal lockdown policy with hospital capacity constraints.

We now introduce a hospital capacity constraint in the SIR model and study the optimal lockdown policy in the presence of such a constraint. We model this constraint as follows. There is a limit to the hospital capacity in intensive care, so that beyond a certain number of patients the infection fatality rate increases sharply. This is because when the intensive care units are full, the extra sick cannot be treated properly. The proportion of infected beyond which mortality increases for this reason is denoted by $\bar{i}$. Once this limit is reached, additional infected patients cannot be treated in intensive care units (ICU) and are subject to higher death hazard. The ICU saturation exemplifies hospital capacity constraints. Therefore, due to these constraints and the induced dual treatment of patients, the infection fatality rate is $\delta' > 0$ for the patients treated outside intensive care units, that is for the proportion of patients $\left(i - \bar{i}\right)$ when i is higher than

$\bar{i}$, with $\delta' > \delta > 0$. This affects the mortality record of the pandemic. It is assumed here that $\bar{i} \in [i(t_0); i_{\max}(\mathcal{R}_0)]$. The decision is made on date t_0 before the intensive care units are full, but knowing that they will be saturated if no lockdown decision is made. More precisely, modifying previous equations, we get

$$\text{For } \mathcal{R} \leq \overline{\mathcal{R}}, \text{ we have } m_T(\mathcal{R}) = \delta\left(i_{\mathcal{R}}(T) + r_{\mathcal{R}}(T)\right) = \delta\left(1 - s_{\mathcal{R}}(T)\right) \tag{18}$$

$$\text{For } \mathcal{R} > \overline{\mathcal{R}}, \text{ we have } m_T(\mathcal{R}) = \delta(1 - s_{\mathcal{R}}(T)) + (\delta' - \delta)\, z(\mathcal{R}) \tag{19}$$

The excess mortality rate is the difference between the number of deaths when ICU are saturated during a certain interval (and the mortality rate jumps to δ') and the number of deaths that would be generated by the pandemic in the absence of ICU saturation over the interval (t_0, T). This rate, for a given policy reproduction ratio $\mathcal{R}$, is detailled in the following

Proposition 4. *The excess mortality rate due to ICU saturation over (t_0, T) is equal to $(\delta' - \delta)\, z\left(\mathcal{R}, \bar{i}\right)$, where*

$$z\left(\mathcal{R}, \bar{i}\right) = \left(r_{\mathcal{R}}(T_2) - r_{\mathcal{R}}(T_1)\right) - \gamma \bar{i}\,(T_2 - T_1)$$

for $\mathcal{R} > \overline{\mathcal{R}}$, and $z\left(\mathcal{R}, \bar{i}\right) = 0$ for $\mathcal{R} \leq \overline{\mathcal{R}}$, with $\overline{\mathcal{R}}$ such that $i_{\max}\left(\overline{\mathcal{R}}\right) = \bar{i}$ where $i_{\max}\left(\overline{\mathcal{R}}\right) = i(t_0) + s(t_0) - \frac{1}{\overline{\mathcal{R}}} - \frac{\ln(\overline{\mathcal{R}})}{\overline{\mathcal{R}}} - \frac{1}{\overline{\mathcal{R}}} \ln(s(t_0))$. T_1 is the first instant t such that $i(t) \geq \bar{i}$ and T_2 is the last instant t such that $i(t) \geq \bar{i}$.

The excess mortality rate is an increasing function of $\mathcal{R}$ and a decreasing function of $\bar{i}$.

Proof. See Appendix A.4. □

It is assumed that $t_0 < T_1 < T_2 < T$, i.e. the lockdown is decided before the saturation of ICU and the lockdown is lifted after the end of the saturation. ICU saturation is reached at T_1. As the proportion of infected continues to grow, the infection fatality rate becomes δ' for the patients treated outside ICU. After the pandemic peak (when $s(t) = 1/\mathcal{R}$), the number of deaths declines and reaches again the ICU limit

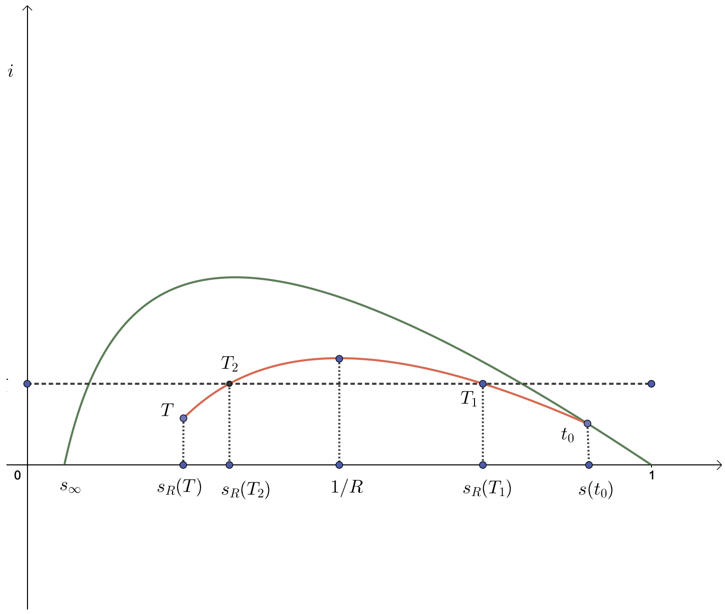

Figure 4: Excess mortality between T_1 and T_2 if $\mathcal{R} > \overline{\mathcal{R}}$.

at T_2. Afterwards the infection fatality rate returns to δ. The excess mortality rate depends positively on the difference between δ' and δ and the duration of the period of ICU saturation $(T_2 - T_1)$. The higher the policy reproduction number, the higher the excess mortality rate. If $\mathcal{R}$ is high, the pandemic spreads quickly after t_0 and reaches more rapidly the ICU limit $\bar{i}$. Consequently the excess mortality rate is an increasing function of $\mathcal{R}$. On the other hand, for a given $\mathcal{R}$, if the ICU constraint $\bar{i}$ increases, it is reached later (T_1 is increased) and the excess mortality period is shortened (T_2 is decreased).

The optimal policy $\mathcal{R}_h^{opt}$ with hospital capacity constraints is studied in the following

Proposition 5.

Let $\mathcal{R}_h^{opt}$ be the optimal value of $\mathcal{R}$ with hospital capacity constraints. Then setting $\lambda_1 = \frac{(T-t_0)y'(\mathcal{R}_0)}{-\delta' N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\mathcal{R}_0} + (\delta'-\delta)Nz'(\mathcal{R}_0)} \geq 0,$ *and* $\lambda_2 = \frac{(T-t_0)y'\left(\overline{\mathcal{R}}\right)}{-\delta' N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\overline{\mathcal{R}}} + (\delta'-\delta)Nz'_+\left(\overline{\mathcal{R}}\right)} \geq 0$ *and* $\lambda_3 = \frac{(T-t_0)y'\left(\overline{\mathcal{R}}\right)}{-\delta N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\overline{\mathcal{R}}}} \geq 0,$ *with* $0 < \lambda_1 < \lambda_2 < \lambda_3$, *we have:*

(i) If $\lambda \in [0; \lambda_1]$, $\mathcal{R}_h^{opt}$ *is equal to* $\mathcal{R}_0$.

(ii) If $\lambda \in [\lambda_1; \lambda_2]$, $\mathcal{R}_h^{opt}$ *is a decreasing function of* λ.

(iii) If $\lambda \in [\lambda_2; \lambda_3]$, $\mathcal{R}_h^{opt} = \overline{\mathcal{R}}$ *is a constant function of* λ *and a decreasing function*

of t_0.

(iv) If $\lambda \in [\lambda_3; +\infty)$, $\mathcal{R}_h^{opt}$ *is a decreasing function of* λ.

Proof. See Appendix A.5. $\square$

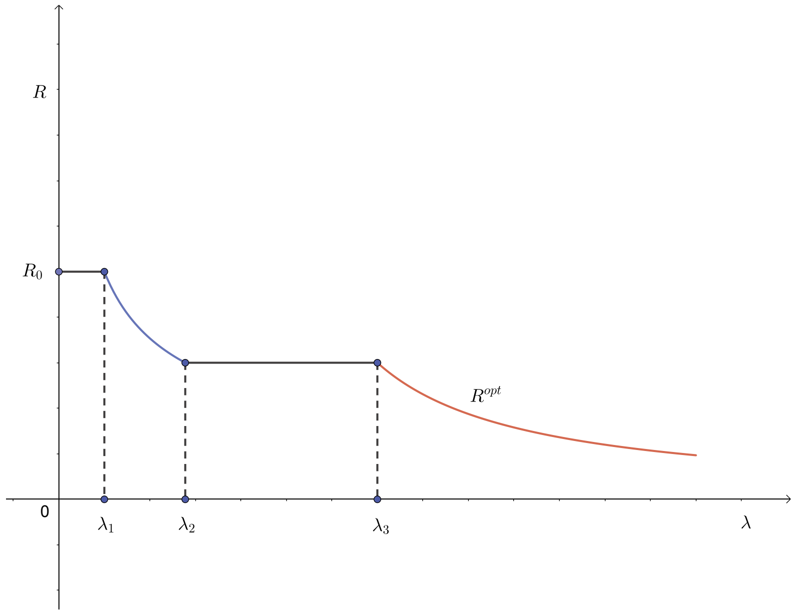

Figure 5: Variations of $\mathcal{R}_h^{opt}$ with respect to λ. The case of hospital capacity constraint.

Proposition 5 contrasts with Proposition 3. The impact of λ on the optimal choice of the policy reproduction number is more complex due to the existence of ICU saturation. The striking result of this proposition is that there is an intermediate range of "values of life" weight λ for which the optimal lockdown policy consists in exactly saturating the hospital ICU capacity. This comes from the discontinuity in the infection fatality rate linked to the ICU capacity.

For λ higher than λ_3, the standard trade-off applies and there is a decreasing relationship between $\mathcal{R}_h^{opt}$ and λ. λ_3 is the "value of live" weight for which the ICU capacity is met but does not constrain the lockdown decision.

For λ lower than λ_3, the ICU capacity binds and affects this decision. For these values, the policymaker is confronted to the dilemma of avoiding or not the extra mortality associated with δ' when the health system is oversaturated and thus beyond the ICU capacity. For λ lower than λ_3 but higher than λ_2, she adopts the reproduction number $\overline{\mathcal{R}}$. Decreasing λ below λ_3 would lead to an increase in the chosen reproduction number $\mathcal{R}$. But for λ higher than λ_2, the implied extra-mortality would augment the health argument in the welfare function and generate higher health costs than the economic gains associated with such an increase in $\mathcal{R}$. Thus it is preferable to keep the chosen $\mathcal{R}$ at $\overline{\mathcal{R}}$. For values lower than λ_2, the policymaker prefers trespassing the ICU capacity and having some infected people dying at a rate δ' so as not to incur the economic costs associated with strictly meeting the ICU limit. This amounts to choose an optimal reproduction number higher than $\overline{\mathcal{R}}$.

Obviously the welfare level reached at the end of the policy period is higher in the absence of hospital contraints (because for each value of $\mathcal{R}$, the welfare is higher without hospital constraints). In the same logic, this welfare level is an increasing function of $\bar{\imath}$.

4 Farsighted lockdown policies.

We now investigate the pandemic and its relation wich lockdown policy in a long-term perspective. Compared to a short-term perspective, three differences can be introduced. The first one is that the delayed consequences of a lockdown policy are taken into account: Given the dynamic nature of the problem, anything happening in a given time interval impacts on the subsequent evolution of the pandemic. Secondly, we may relax the assumption made before that solely one lockdown with one reproduction number is fixed by the policymaker. Lastly, it may happen that the post-policy reproduction number differs from the initial one: it is equal to $\mathcal{R}_0'$ instead of $\mathcal{R}_0$. In the sequel, we focus on the first difference: the policymaker deciding in t_0 takes into account the delayed impact after the end-of-policy instant T of her lockdown policy. Said in other words, she may be qualified as “far-sighted”. We maintain the assumption that the policymaker chooses a unique reproduction number[15] and we assume that the post-policy reproduction number is the “natural” one: $\mathcal{R}_0' = \mathcal{R}_0$.

[15] We will relax this assumption in Section 4.3.

4.1 Long-term dynamics of the pandemic with lockdown policy.

We characterize the impact of a lockdown policy with a constant $\mathcal{R}$ over the duration period: $\mathcal{R} = \frac{\beta}{\gamma}$ on $t \in [t_0; T]$. For $t < t_0$, the dynamics is given by eqs. (4)-(6). For $t \in [t_0, T]$, it is given by eqs. (11)-(13). For $t > T$, it is given by the following equations

$$r(t) = r(T) + \frac{1}{\mathcal{R}_0} \ln s(T) - \frac{1}{\mathcal{R}_0} \ln s(t) \tag{20}$$

$$i(t) = i(T) + s(T) - s(t) + \frac{1}{\mathcal{R}_0} \ln s(t) - \frac{1}{\mathcal{R}_0} \ln s(T) \tag{21}$$

$$\int_{s(t)}^{s(T)} \frac{1}{\beta_0 s \left[i(T) + s(T) - s + \frac{1}{\mathcal{R}_0} \ln s - \frac{1}{\mathcal{R}_0} \ln s(T) \right]} ds = t - T. \tag{22}$$

$s(t)$, $i(t)$, $r(t)$ will be denoted by $s_{\mathcal{R},T}(t)$, $i_{\mathcal{R},T}(t)$, $r_{\mathcal{R},T}(t)$ for $t > T$ when it will be necessary to stress their dependency to $\mathcal{R}$ and T. Eqs. (20)-(22) are similar to (11)-(13) with the crucial difference that the notations $r(T)$, $i(T)$ and $s(T)$ are introduced[16]. (20)-(22) depend on $s(T)$ and $i(T)$, that is on the outcome of the policy fixing $\mathcal{R}$ over the interval $t \in [t_0, T]$. This proves the delayed consequences of a lockdown policy after it has stopped. In the two following propositions, we investigate the dynamics of the pandemic after T.

After the end of a lockdown policy, either the proportion of infected pursues its decline (at a different pace) or it reverts to increasing again. The latter case refers to a "rebound" which is a very common feature in actual pandemics. For example, in many countries the Covid-19 pandemic was characterized by several rebounds, not all due to the advent of variants of the original virus. Thus it is important to understand under which circumstances such a reversal happens in the absence of a renewed source of infection such as a new virus or a variant of the current one. This is answered in the following

[16] They replace $r(t_0)$, $i(t_0)$ and $s(t_0)$ respectively, and $\mathcal{R}$ is replaced by $\mathcal{R}_0$.

Proposition 6.

(i) If $s\,(T) > \frac{1}{\mathcal{R}_0}$*, there is a rebound of the epidemic after* T *and* $i(t)$ *is maximal on* $t \geq T$ *when* $s(t) = \frac{1}{\mathcal{R}_0}$.

If $s\,(T) \leq \frac{1}{\mathcal{R}_0}$*, there is no rebound of the epidemic after* T *and* $i(t)$ *is maximal on* $t \geq T$ *when* $t = T$.

(ii) There exists a value $\widetilde{\mathcal{R}}_{t_0}$ *such that*

1. *For* $\mathcal{R} > \widetilde{\mathcal{R}}_{t_0}$*, there is no rebound after* T *if* T *is sufficiently high.*

2. *For* $\mathcal{R} \leq \widetilde{\mathcal{R}}_{t_0}$*, there is necessarily a rebound after* T *for any value of* T*.*

Proof. See Appendix A.6. □

Figure 6 illustrates the case of dynamics including a rebound, Figure 7 the case without a rebound.

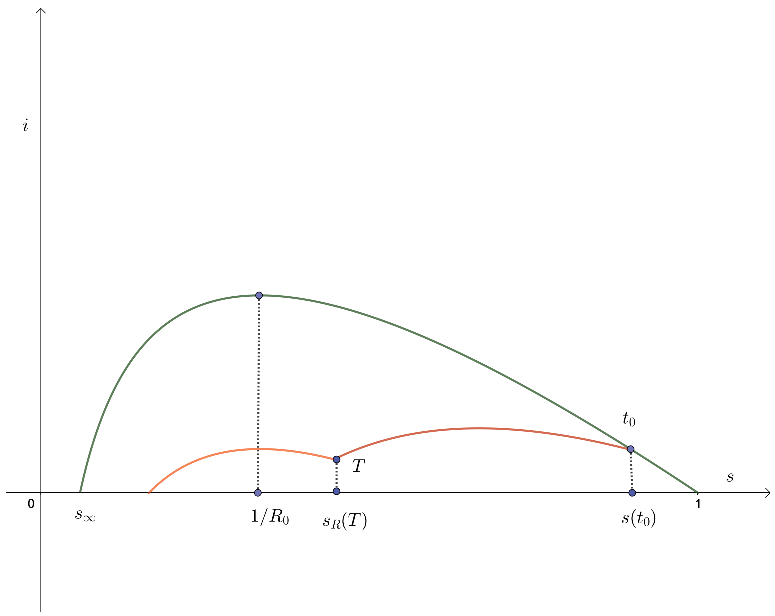

Figure 6: Rebound

(i) focuses on the impact of $s\,(T) = s_{\mathcal{R}}\,(T)$ and therefore implicitly of $\mathcal{R}$. It proves that if the susceptible proportion $s(T)$ at the end of the lockdown is above the "natural" peak of the pandemic, there will be a rebound when the reproduction number switches back to $\mathcal{R}_0$. The pair $(\mathcal{R}, T)$ is such that the control over the pandemic is not sufficient

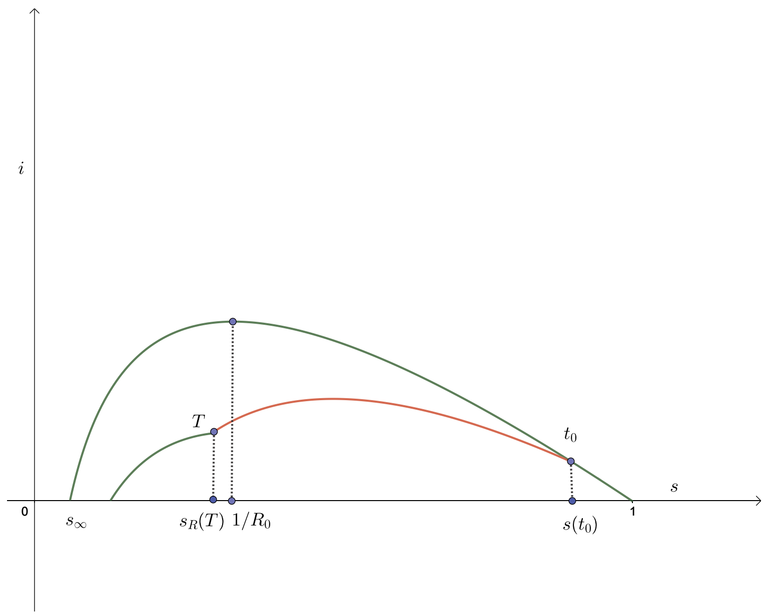

Figure 7: No rebound

to pass this peak. The pool of people susceptible to be infected after the lockdown is too large and the contamination process starts increasing again after T: a rebound occurs. Notice that the post-T function between $s(t)$ and $i(t)$ shares the same property as in the case without lockdown: its peak is at $\frac{1}{\mathcal{R}_0}$. Therefore if $s(T)$ is higher than $\frac{1}{\mathcal{R}_0}$, it implies that the number of infected people increases after T when the susceptible proportion pursues its decline. It is solely if $s(T)$ is smaller than $\frac{1}{\mathcal{R}_0}$ that the two numbers decline together: no rebound occurs.

(ii) focuses on the impact of $\mathcal{R}$. If the reproduction number is higher than a critical value denoted by $\widetilde{\mathcal{R}}_{t_0}$, a rebound is avoided for T given and sufficiently high: both parameters conjugate to avoid a rebound after the lockdown policy. However, if $\mathcal{R}$ is too low, the control of the pandemic whatever the duration length is insufficient and it rebounds. This is due to the fact that the end-of-policy susceptible proportion (the pool of people available for infection) is high enough so as to let the natural reproduction number have a huge impact on the number of infected and lead to a rebound. This counter-intuitive result casts doubt on a policy whose severity is meant to control efficiently the pandemic. This is true in the short-term but eventually it will undo itself. This is particularly true to a "zero-Covid policy" which cannot last

for ever given its economic costs. Sooner or later, an extremely severe lockdown policy (implying a very low reproduction number) will stop before the complete eradication of the virus (which can only happen at the end-of-time, that is, when t is arbitrarily large). At the end of the severe lockdown period, the pool of susceptible will be close to $s(t_0)$, and the dynamics of the pandemic governed by the natural reproduction number $\mathcal{R}_0$ and given by Proposition 1.[17] On the whole, society suffers from a large economic cost due to the severity of the lockdown policy without much impact on long-term mortality. This is unknown (or neglected) by the short-sighted policymaker who solely looks at the outcome at the end-of-policy instant T but this unpleasant conclusion appears clearly when a long-term perspective is adopted.

Turning to the eventual impact of a lockdown policy on the end-of-pandemic susceptible proportion we offer the following

Proposition 7.

(i) At the end of the pandemic, we have $(s, i, r) = (s_\infty(\mathcal{R}, T), 0, r_\infty(\mathcal{R}, T))$, *with* $r_\infty(\mathcal{R}, T) = 1 - s_\infty(\mathcal{R}, T)$ *and* $s_\infty(\mathcal{R}, T)$ *given by*

$$\mathcal{R}_0 = \frac{\ln(s(T)) - \ln(s_\infty(\mathcal{R}, T))}{i(T) + s(T) - s_\infty(\mathcal{R}, T)}, \qquad 0 < s_\infty(\mathcal{R}, T) < 1. \tag{23}$$

We have $s_\infty(\mathcal{R}, T) < \frac{1}{\mathcal{R}_0}$, *and* $s_\infty(\mathcal{R}, T)$ *is an increasing function of* T *and a decreasing function of* $\mathcal{R}_0$.

(ii) The end-of-pandemic susceptible proportion $s_\infty(\mathcal{R}, T)$ *is always higher than* $s_\infty(\mathcal{R}_0)$. *If* T *is sufficiently large,* $s_\infty(\mathcal{R}, T)$ *is a non-monotonic function of* $\mathcal{R}$, *it is increasing on* $\mathcal{R} < \widetilde{\mathcal{R}}_{t_0}$, *and decreasing on* $\mathcal{R} > \widetilde{\mathcal{R}}_{t_0}$.

Proof. See Appendix A.7. □

(i) makes clear that a lockdown policy always has an impact on herd immunity, that is, the end-of-pandemic susceptible proportion. It suffices to compare (23) with (7): in

[17] The case of Australia which pursued a zero-Covid strategy against Covid-19 is exemplary. Its prime minister Scott Morrison has declared on August 23rd 2021: "This is not a sustainable way to live in this country". See https://www.economist.com/asia/2021/08/28/australia-is-ending-its-zero-covid-strategy?utm_campaign=coronavirus-special-edition&utm_medium=newsletter&utm_source=salesforce-marketing-cloud&utm_term=2021-08-28&utm_content=article-link-1&etear=nl_special_1

the former equation $i(T) = i_{\mathcal{R}}(T)$ and $s(T) = s_{\mathcal{R}}(T)$ which depend on the policy stance $(\mathcal{R}, T)$ now appear. The longer lasts the lockdown policy, the better it is in terms of herd immunity, which is consistent with intuition. This is due to the fact that both $i(T)$ and $s(T)$ decline with T, when T is large enough. It is logically a decreasing function of $\mathcal{R}_0$ as this number governs the post-policy dynamics: a higher reproduction number applying after T leads to a worsening of the pandemic in the post-lockdown period and eventually a higher fatality record.

(ii) proves that any lockdown policy, however light (a high $\mathcal{R}$) and/or short (a small T), leads to an improvement in the eventual mortality record due to the pandemic. There is never a perverse long-term effect of an active policy. Yet it does not mean that the terminal susceptible proportion is a monotone function of $\mathcal{R}$. Actually this is due to the possible presence of rebounds. As we have seen above, a tight lockdown policy ($\mathcal{R}$ low) may lead to a large rebound whereas a not so tight policy leads to a small rebound. The ending of the large rebound may thus be at the left of the ending of the small rebound. Actually the relationship of $s_\infty(\mathcal{R}, T)$ is a non-monotone function of $\mathcal{R}$ peaking at $\widetilde{\mathcal{R}}_{t_0}$ if T is large. In the absence of rebound ($\mathcal{R} > \widetilde{\mathcal{R}}_{t_0}$), the relationship is decreasing: a stricter lockdown policy improves the terminal susceptible proportion. With rebound ($\mathcal{R} < \widetilde{\mathcal{R}}_{t_0}$), the relationship is increasing: a stricter lockdown generates a higher rebound.

Notice that these two propositions can easily be adapted for the case $\mathcal{R}_0' \neq \mathcal{R}_0$.

4.2 Farsighted optimal lockdown policy without hospital constraints.

Given the delayed consequences of a policy stance $(\mathcal{R}, T)$, a far-sighted policymaker (adopting a long-term perspective on the pandemic) takes them into account. The total economic cost of a lockdown policy $(\mathcal{R}, T)$ is equal to $(T - t_0)(y(\mathcal{R}_0) - y(\mathcal{R}))$. This impact is subject to two opposite forces: a longer duration increases economic losses whereas a higher reproduction number $\mathcal{R}$ decreases them. After T, since the reproduction number returns to $\mathcal{R}_0$ and no lagged economic effect of a policy fixed lockdown is assumed, there is no economic loss due to lockdowns. On the contrary, the health impact of the pandemic still goes on, based on $s(T)$ and $\mathcal{R}_0$, as shown in (20)-(22).

Extending (17), the decision-maker's objective is now to minimize losses over the entire future (for T given):

$$L_\infty(\mathcal{R}) = (T - t_0)\left(y(\mathcal{R}_0) - y(\mathcal{R})\right) + \lambda M_\infty(\mathcal{R}) \tag{24}$$

with $M_\infty(\mathcal{R}) = N\delta r_\infty(\mathcal{R}) = N\delta\left(1 - s_\infty(\mathcal{R}, T)\right)$. The properties of the optimal decision of a farsighted policymaker are given in

Proposition 8.

Let $\mathcal{R}_\infty^{opt}$ be the value of $\mathcal{R}$ minimizing the long-term loss $L_\infty(\mathcal{R})$ on $0 \leq \mathcal{R} \leq \mathcal{R}_0$ (for T given). Then setting $\lambda_0' = \frac{(T-t_0)y'(\mathcal{R}_0)}{-\delta N s_\infty'(\mathcal{R}_0)} \geq 0$ we have:

(i) $\mathcal{R}_\infty^{opt}$ is equal to $\mathcal{R}_0$ for $\lambda \in [0; \lambda_0']$ and $\mathcal{R}_\infty^{opt}$ is a decreasing function of λ for $\lambda \in [\lambda_0'; \infty)$. Moreover $\mathcal{R}_\infty^{opt} \geq \widetilde{\mathcal{R}}_{t_0}$ if T is sufficiently high.

(ii) When (15) applies, for $\lambda \geq \lambda_0'$, $\mathcal{R}_\infty^{opt}$ is always an increasing function of A and an increasing function of α on $1 + \alpha \ln(\mathcal{R}_\infty^{opt}) > 0$, i.e. on $\mathcal{R}_\infty^{opt} > e^{-1/\alpha}$.

(iii) For T and λ high enough, $\mathcal{R}_\infty^{opt} \geq \widetilde{\mathcal{R}}_{t_0} > \mathcal{R}^{opt}$.

Proof. See Appendix A.8. □

(i) and *(ii)* generalize the first and fourth points in proposition 3 and are similarly explained. The impact of t_0 and T on $\mathcal{R}_\infty^{opt}$ are similar to the ones obtained in proposition 3[18] and are not repeated here. Yet there is a crucial difference about the impact of the value of life parameter λ. Assuming T sufficiently high, the long-term optimal reproduction number is above a positive value $\widetilde{\mathcal{R}}_{t_0}$ for any value of λ and thus does not tend to 0 when λ tends to infinity, unlike the short-term optimal number (See Proposition 3). This results from the desire of the farsighted policymaker to avoid a rebound after T, an event which is not anticipated by a shortsighted policymaker. A rebound leads to an increase in mortality and a lower herd immunity level. Avoiding

[18] with s_∞ replacing $s(T)$

a rebound makes sense especially when the life argument in the loss function is given a higher weight. According to Proposition 6 (ii), if T is sufficiently high, then there is a rebound if $\mathcal{R} \leq \widetilde{\mathcal{R}}_{t_0}$, and there is no rebound if $\mathcal{R} > \widetilde{\mathcal{R}}_{t_0}$. This implies that the far-sighted policymaker will choose a reproduction number $\mathcal{R}_\infty^{opt}$ higher than $\widetilde{\mathcal{R}}_{t_0}$. In other words, an optimal lockdown policy is such that it rules out rebounds.

(iii) directly derives from Proposition 6. The comparison between $\mathcal{R}^{opt}$ and $\mathcal{R}_\infty^{opt}$ leads to a counter-intuitive result: A far-sighted policymaker may adopt a more lenient lockdown policy than a short-sighted one. This is due to the presence of rebounds. For λ high, for a given T (sufficiently large), a short-term perspective leads to a low reproduction number such that there is a rebound ($s_{\mathcal{R}^{opt}}(T) > 1/\mathcal{R}_0$) whereas $\mathcal{R}_\infty^{opt}$ generates no rebound ($s_{\mathcal{R}_\infty^{opt}}(T) < 1/\mathcal{R}_0$). This implies that $\mathcal{R}_\infty^{opt} > \mathcal{R}^{opt}$. If the far-sighted policymaker adopts a relatively lenient lockdown policy such that $s_{\mathcal{R}_\infty^{opt}}(T) < 1/\mathcal{R}_0$, it suffers from lower economic losses during the policy period and does not witness a rebound. She may reach a higher end-of-time mortality rate $s_\infty(\mathcal{R}_\infty^{opt})$ than the short-sighted policymaker: $s_\infty(\mathcal{R}_\infty^{opt}) > s_\infty(\mathcal{R}^{opt})$. On the whole, despite the more lenient long-term choice compared to the short-term choice, economic costs and the mortality record are lower. Again, this is in line with the result that a too strict lockdown policy may be harmful in the long-term.

4.3 Farsighted optimal lockdown policy with hospital capacity constraints.

We now introduce hospital capacity constraints as in section 3.2.2. In particular we make the same assumptions. But we relax the assumption (which defined the short-sighted perspective) that there is a unique lockdown, with a unique policy-chosen reproduction number and we may allow the policymaker to choose instead a sequence of lockdowns, with different reproduction numbers. We assume that δ' and λ are so high that the policymaker never wants to trespass the ICU limit $\bar{\imath}$. Either the jump in the fatality rate (from δ to δ') is very large or the political cost of a jump in mortality is such that the relative value of life is high. During the Covid-19 pandemic, clearly many

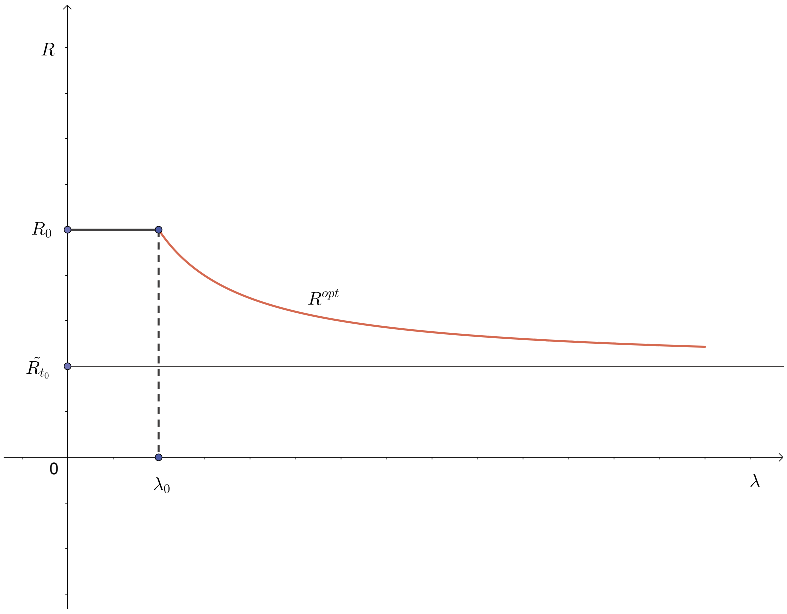

Figure 8: Variations of $\mathcal{R}_\infty^{opt}$ with respect to λ.

policymakers designed their policy with this objective in mind. The issue is how to reach it. We answer this question in the following

Proposition 9.

Assume that at each date t with $i(t) = i(t_0)$ a new lockdown begins such that during this lockdown $\max i = \bar{i}$. Then

If $\bar{i}$ is sufficiently high, one lockdown is sufficient to achieve herd immunity.

If $\bar{i}$ is sufficiently low, several successive lockdowns are necessary to achieve herd immunity. We obtain a sequence of lockdowns $\mathcal{R}_1$, $\mathcal{R}_2$, ..., $\mathcal{R}_n$ such that $\mathcal{R}_1 < \mathcal{R}_2 < ... < \mathcal{R}_n$, i.e. they are less and less severe.

Proof. See Appendix A.9.

□

Clearly if $\bar{i}$ is sufficiently high, a unique policy is sufficient. Suppose that it is arbitrarily high, it is then never binding for any value of the reproduction number, including the natural number. By continuity, there is a unique policy pair when it is lower but close to the value of $\bar{i}$ where the hospital capacity constraint binds for a given reproduction number. If $\bar{i}$ is low enough, the optimal rate chosen in the absence of ICU

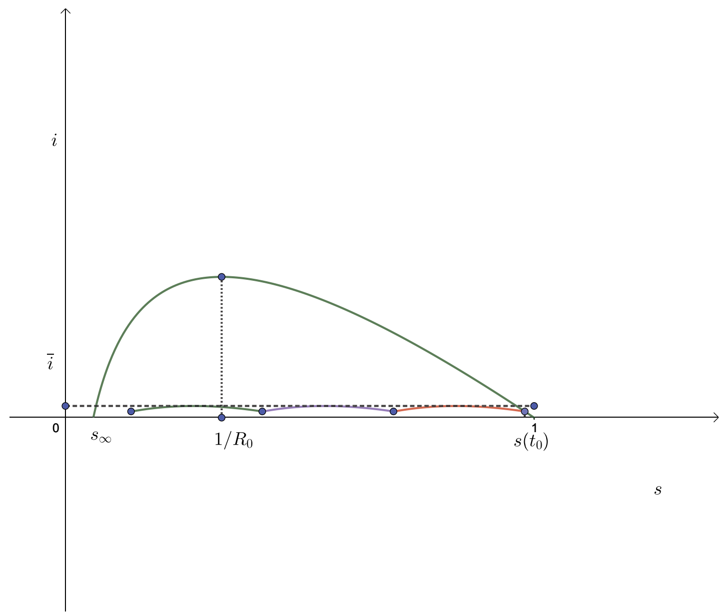

Figure 9: Lockdowns consistent with ICU.

saturation cannot be implemented as it would generate excess mortality. Therefore it is necessary to reduce the policy reproduction number and choose in t_0 a lower value $\mathcal{R}_1$ such that $i = \bar{i}$ when $s(t) = 1/\mathcal{R}_1$. Given the stronger social distancing, once the contamination peak is passed, the infection ratio decreases and the level $i\,(t_0)$ is met in t_1. If this lockdown is stopped (at t_1 or after t_1), there will be a rebound since herd immunity is not achieved. Thus it is necessary to adopt a new lockdown to avoid ICU saturation. But given that $s\,(t_1)$ is lower than $s\,(t_0)$, to reach a peak corresponding to $\bar{i}$ during this new lockdown necessitates a higher reproduction number. The second lockdown is less severe than the first one. Choosing a reproduction number $\mathcal{R}_2$ higher than $\mathcal{R}_1$ will allow to get closer to herd immunity. Repeating the argument until the last chosen reproduction number $\mathcal{R}_n$ generates the series of inequalities given in Proposition 9. This is consistent with the evidence that the successive lockdowns imposed during the Covid-19 pandemic were of decreasing severity in advanced countries at least.

5 Extensions.

Choosing a lockdown policy consists in choosing a value $\mathcal{R} \in (0, \mathcal{R}_0)$ and a value $T \in]t_0, \infty[$. Up to now, we considered the policy duration $T - t_0$ as given. It is interesting to relax this assumption as interrogations about the duration of lockdowns were rife during the Covid-19 pandemic. Lastly, introducing the case of a gradual vaccination policy starting at a given date (T for simplicity), we investigate the consequences of the rate of vaccination over the dynamics of the pandemic.

5.1 On policy duration

To shed some light on the role of policy duration, we consider the following problem. Supposing the policymaker wishes to attain collective immunity consistent with a given mortality M_∞, i.e. equivalently with a certain final susceptible proportion $s_\infty = \frac{1}{\mathcal{R}_0} - \varepsilon$, which pair $(\mathcal{R}, T)$ does she choose? By considering an objective in terms of reaching a given collective immunity ratio we do not oppose here the health objective to the economic one. Instead we focus on a possible trade-off between the duration of a lockdown policy and its stringency. It may be argued that the collective immunity level should be reached in the minimal time through a "tough" lockdown policy, given the impatience of the people to get rid of the pandemic as soon as possible, rather than applying a more lenient lockdown policy (a higher $\mathcal{R}$) on a longer period. Is it true? Is the shortest duration policy optimal? We answer this problem in the following proposition. We denote by $(\mathcal{C}_\varepsilon)$ the curve in the plane (s, i) representing the end of lockdowns $(s_\mathcal{R}(T), i_\mathcal{R}(T))$ which lead after release of lockdown to $s_\infty(\mathcal{R}, T) = \frac{1}{\mathcal{R}_0} - \varepsilon$.

Proposition 10. *Let* $\varepsilon \in \left(0; \frac{1}{\mathcal{R}_0}\right)$ *be given.*

(i) There exist an infinity of couples $(\mathcal{R}, T)$ *such that* $s_\infty(\mathcal{R}, T) = \frac{1}{\mathcal{R}_0} - \varepsilon$. *More precisely, there exist* $\mathcal{R}_1$ *and* $\mathcal{R}_2$, $\mathcal{R}_2 > \mathcal{R}_1$ *and a function* T_ε *such that*

$$s_\infty(\mathcal{R}, T) = \frac{1}{\mathcal{R}_0} - \varepsilon \Leftrightarrow \mathcal{R} \in (\mathcal{R}_1, \mathcal{R}_2) \text{ and } T = T_\varepsilon(\mathcal{R}).$$

Moreover, $\lim_{\mathcal{R} \to \mathcal{R}_1 +} T_\varepsilon(\mathcal{R}) = \lim_{\mathcal{R} \to \mathcal{R}_2 -} T_\varepsilon(\mathcal{R}) = +\infty$.

(ii) We denote by $\left(\widehat{\mathcal{R}}_\varepsilon, \widehat{T}_\varepsilon\right)$ *the policy pair which generates the minimal economic cost and by* $\left(\mathcal{R}_\varepsilon^\circ, T_\varepsilon^\circ\right)$ *the policy which allows to reach* $s_\infty = \frac{1}{\mathcal{R}_0} - \varepsilon$ *in the minimal time. Then* $\widehat{\mathcal{R}}_\varepsilon > \mathcal{R}_\varepsilon^\circ$ *and* $\widehat{T}_\varepsilon > T_\varepsilon^\circ$*.*

Proof. See Appendix A.10. ☐

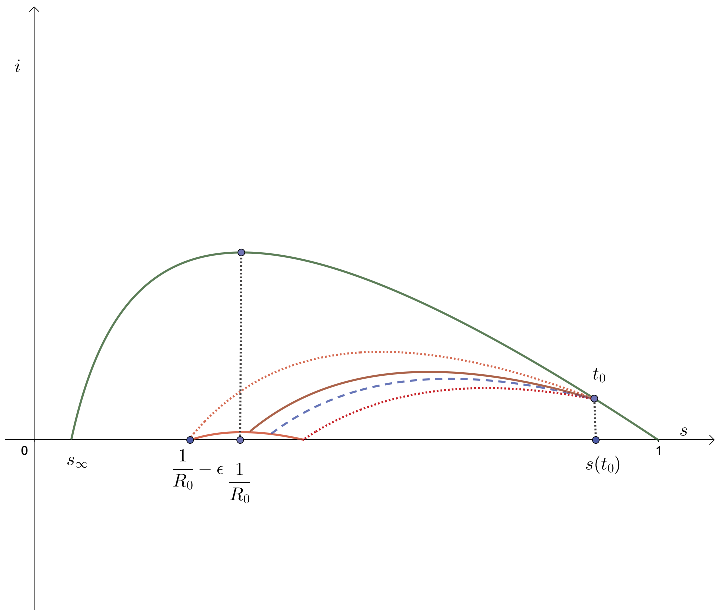

Figure 10: Minimal economic cost to reach $s_\infty = \frac{1}{R_0} - \varepsilon$

(i) states that the objective can be attained by an infinite number of combinations of duration and reproduction number (but not all reproduction numbers $\mathcal{R}$ are admissible). This is due to the adverse consequences of lengthening the policy interval (increasing T) and lightening the lockdown intensity (increasing $\mathcal{R}$). These different lockdown policies cannot be determined according to health considerations only, since they do not have the same economic impact. This raises the question of which policy is best from an economic point of view. *(ii)* shows that the shortest lockdown policy consistent with a health objective generates a higher economic cost than is necessary as it is linked to a relatively severe lockdown policy, i.e. a small reproduction number. Despite the fact that this one does not last long it harms too much economic activity. The cost-minimizing policy implies a higher reproduction number imposed over a longer duration:

patience is rewarding. The conclusion is that living with Covid-19 consists in balancing policy duration and severity of a lockdown policy.

Figure 10 illustrates this proposition. The two dotted curves represent the curves associated with $\mathcal{R}_1$ and $\mathcal{R}_2$. The lower curve reaching $\frac{1}{\mathcal{R}_0} - \varepsilon$ represents $(\mathcal{C}_\varepsilon)$. Any trajectory corresponding to a lockdown policy with $s_\infty(\mathcal{R}, T) = \frac{1}{\mathcal{R}_0} - \varepsilon$ must terminate on $(\mathcal{C}_\varepsilon)$ at the end of the lockdown. Afterwards, the pandemic is governed by $\mathcal{R}_0$ and follows $(\mathcal{C}_\varepsilon)$. The dashed curve corresponds to the time-minimizing trajectory and the continuous curve to the cost minimizing one. Since it is above the dashed curve, it corresponds to a higher reproduction number (from Proposition 2(iii)). It reaches the same mortality record, but it lasts longer.

Proposition 10 tells us that indeed there is an optimal policy stance to reach a given collective immunity level. We can use this result to prove the following

Proposition 11. *There is an optimal policy stance* $(\mathcal{R}^*, T^*)$ *such that* $L_\infty(\mathcal{R}, T)$ *is minimized.*

Proof. See Appendix A.11. □

Proposition 11 claims that the policymaker, acting at t_0, is able to play on both parameters of the policy stance so as to minimize the long-term consequences of the pandemic. Characterizing more precisely this optimal pair is mathematically overly complex and unlikely to convey much information. Addressing the optimal lockdown problem with two instruments in a non-linear system such as the SIR model, while not impossible, is quite complex and likely to obscure the picture for very little analytical gains. It is reasonable to reason with T given and focus on the lockdown policy rate which attracts the most attention and is critical in the dynamics of the pandemic as proven above.

5.2 Vaccination.

We now turn to the issue of vaccination which has proven crucial to stem the dynamics of the Covid-19 pandemic. We assume that vaccination begins at date T, i.e. just at the end of the lockdown period. We denote by $v(t)$ the proportion of vaccinated people at date t, with $t \geq T$. We assume moreover that the rate of vaccination is constant, i.e. there exists $\rho > 0$ such that $\frac{dv}{dt} = \rho s(t)$. The dynamics for $t \geq T$ becomes

$$\frac{ds}{dt} = -\beta_0 i(t) s(t) - \rho s(t) \tag{25}$$

$$\frac{dv}{dt} = \rho s(t) \tag{26}$$

$$\frac{di}{dt} = \beta_0 i(t) s(t) - \gamma i(t) \tag{27}$$

$$\frac{dr}{dt} = \gamma i(t) \tag{28}$$

with $s(t) + v(t) + i(t) + r(t) = 1$, for all $t \geq T$. We denote by $I_\rho(s)$ the proportion of infected people when the proportion of susceptible is s and the rate of vaccination is ρ. The impact of this vaccination policy is given in the following

Proposition 12. *With a constant rate of vaccination ρ, the dynamics of $i(t)$ with respect to $s(t)$ is given for $t \geq T$ by:*

$$H_\rho(s(t), i(t)) = H_\rho(s(T), i(T))$$

where H_ρ is defined by $H_\rho(s, i) = i + s + \frac{\rho}{\beta_0} \ln(i) - \frac{1}{\mathcal{R}_0} \ln(s)$.

If $s(T) > \frac{1}{\mathcal{R}_0}$, there is a rebound after T, even with vaccination. The higher the rate of vaccination ρ, the lower the rebound.

If $t \to +\infty$, then $(s(t), v(t), i(t), r(t))$ tends to $(s_\infty, v_\infty, i_\infty, r_\infty) = (0, v_\infty, 0, r_\infty)$. The proportion of susceptible vanishes when $t \to +\infty$.

The final proportion of vaccinated people is v_∞ given by

$$v_\infty = \int_0^{s(T)} \frac{ds}{1 + \frac{\beta_0}{\rho} I_\rho(s)}$$

where the function I_ρ is defined by: $i = I_\rho(s) \Leftrightarrow H(s, i) = H(s(T), i(T))$. The end-of-time number of deaths $M_\infty = \delta N r_\infty = \delta N (1 - v_\infty)$ is decreasing in the rate of vaccination ρ.

Proof. See Appendix A.12. □

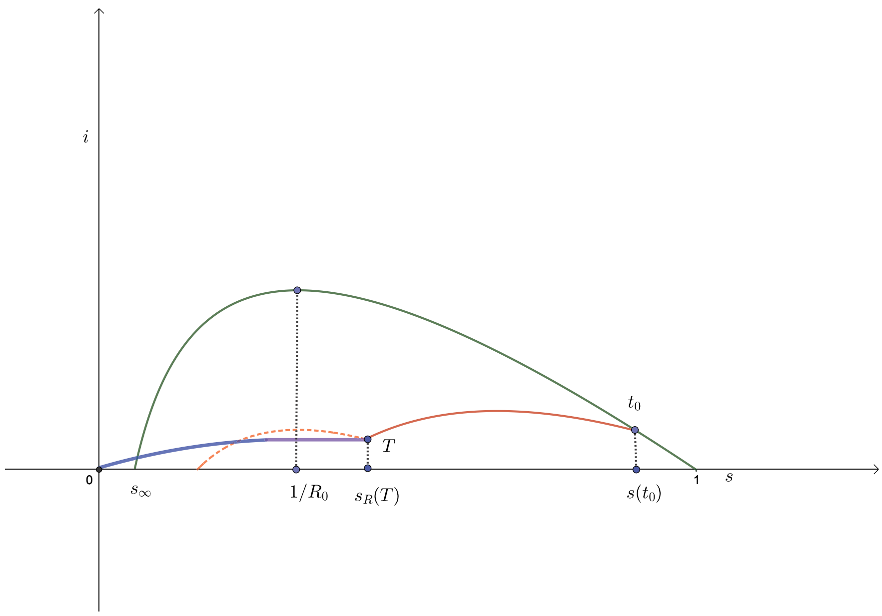

Figure 11: The impact of vaccination

The existence of a rebound depends on the final number of susceptible at the end-of-policy instant T therefore before the vaccination policy starts. However its severity depends on the rate of vaccination ρ. A higher rate limits the extent of the rebound because it depresses more the pool of susceptible after T. Since the vaccination process extends to the end-of-time, agents get vaccinated as long as the pool of susceptible is not null and tends to exhaust this pool. Thus the end-of-time proportion of susceptible tends to zero (see Figure 11[19]). Given that the number of vaccinated is ultimately increasing in the vaccination rate, the end-of-time number of recovered (and thus the number of deaths) is decreasing in this rate. A vaccination policy is fruitful, insofar as the vaccine is efficient.

6 Conclusion.

This paper offers a theoretical analysis of the optimal decision in social distancing taken by a policymaker confronted with a pandemic and facing a dilemma between reducing the economic costs of lockdown and minimizing the mortality rate through social distancing.

[19]The thick line corresponds to the post T dynamics of the vaccination.

Using the workhorse model of epidemiology, namely a deterministic version of the SIR model, we first look at the dynamics of the pandemic in the absence of any policy action aiming at controlling the pandemic. We obtain a non-monotone relation between the infected proportion and the susceptible proportion which peaks at a susceptible proportion equal to the inverse of the reproduction number. This very simple expression plays a critical role in the understanding of the pandemic dynamics when a lockdown policy is put in place.

A lockdown policy is defined by three parameters: the instant of decision, the extent of lockdown which affects the reproduction number and the duration of the lockdown. We show that the dynamics of the pandemic is strongly affected by these variables. Such a policy is designed so as to trade off the health benefits and the economic losses due to a longer and a more stringent lockdown policy. Solving this trade-off amounts to search for the optimal lockdown policy to be followed. This issue is the core of this paper. We distinguish between a short-term perspective, when the policymaker is short-sighted and limits her time horizon to the ending of the lockdown, and a long-term perspective, when she is far-sighted and takes into account the posterior consequences on the dynamics of the pandemic after this ending, up to infinity. We focus the analysis on the reproduction number (inversely related to the extent of social distancing) which is the policy instrument.

We show that there can be rebounds in the pandemic happening either when the policy duration is too short and/or when the lockdown policy reproduction number is too low. A far-sighted policy takes into account these rebounds. The existence of rebounds explains why a "zero-Covid" policy, however its duration, is unsustainable as it leads eventually to huge rebounds with a very low terminal susceptible proportion as well as huge economic costs. Yet, even with rebounds, any policy stance leads to a higher collective immunity relative to a non-interventionist position and a weak improvement. The policymaker may choose not to intervene by means of imposing some social distancing measure if the relative value of life is sufficiently low. When the parameter capturing this relative value in the welfare function characterizing the policymaker is above a certain threshold, the magnitude of the lockdown depends positively on this parameter: the chosen reproduction number is a decreasing function of this parameter. This is true in both perspectives. The optimal long-term optimal reproduction number taking into account the possibility of rebounds mitigates the strictness of lockdown in order to avoid these rebounds or limit their amplitude. If the duration period is sufficiently long, the optimal reproduction number is large enough so as to avoid rebounds. It may therefore be that the long-term optimal lockdown is more lenient than the short-term optimal one, for a given lockdown duration.

In the same vein, it is not true that the shortest duration combined with a strict lockdown measure consistent with reaching a given collective immunity target is optimal. A more lenient policy (a chosen higher reproduction number) enforced over a longer duration period dominates such a policy as it rules out rebounds and thus lowers the final mortality record and generates a lower total economic cost. Living with the pandemic may mean enduring a milder lockdown over a longer period. There exists an optimal pair of reproduction number and duration.

Taking into account hospital constraints turns out to be critical. Both in a short-term and a long-term perspectives, this considerably affects the optimal policy stance and its sensitivity to structural parameters such as the relative value of life. When it is supposed that the costs of trespassing these constraints are arbitrarily large, the optimal policy consists in generating a series of lockdown decisions which are less and less stringent. Addressing the vaccination process, we prove that a constant rate of vaccination beginning at the end of the lockdown cannot prevent a rebound but can limit its amplitude. In the very long-term, there are no more susceptible people: everyone is vaccinated or ”removed” (i.e. recovered or dead). Contrarily to what is observed in a standard policy decision setting, the optimal solution depends on the complex interplay between the economic and sanitary efficiency of a lockdown decision when plausible constraints such as hospital capacity constraints and availability of a vaccine are taken into account. This comes from the non-linear characteristics of the dynamics of a pandemic as formalized by the SIR model.

The policy instrument in this model is referred to as the “lockdown instrument”. Actually there exist many different instruments to tackle an expanding pandemic, in particular tracking, testing, appropriate individual equipements as masks and finally, isolation. A mix of measures is likely to be what defines an adequate policy toward the control of a pandemic. Any such measures are likely to have opposite health (positive) and economic (negative, if only because of direct costs) impacts and therefore meet our assumptions. Our policy instrument can thus be understood as a "composite" public health instrument (a combination of measures) for tackling a pandemic.

The model assumes that there is a single policy instrument and the population is homogeneous. It does not take into account the reaction of the population to the lockdown decision and assumes a simple framing of economic and sanitary losses. As it is, it proves an adequate basis for understanding the basic policy issues related to the control of a pandemic, in particular in relation with economic consequences of a lockdown policy when stylized laws of the dissemination of a pandemic are explicitly taken into accunt. The model can be complexified so as to take into consideration different assumptions. Relaxing these assumptions as well as analyzing lockdown policy in variants of the SIR model which have been offered in the epidemiology literature is left for further research.[20]

Finally, the model rules out uncertainty. Epidemiologists have developed a stochastic approach to capture the randomness in the matching process between infected and susceptible people. This is when the number of infected is very low and the law of large numbers does not apply (see Britton (2010)). We do not claim that the results obtained here (in particular the self-defeating nature of a zero-Covid policy) are transferable to a stochastic approach. The strength of a deterministic approach is to obtain analytical results without recouring to simulation techniques which clarify the impact of a lockdown policy. In particular it appears that these results are useful to understand the pitfalls of a lockdown policy, in particular the adverse consequence of a short-sighted lockdown policy.

[20] On the need to combine epidemiology and economics, Murray (2020, p.106) writes: "As an epidemiologist, I ask economists interested in Covid-19 to build on their expertise and ours. Indeed, the efforts of economists in tackling the economic sequelae of this pandemic are vitally needed, as are the development of tools for tracking, predicting, and preventing future pandemics based on understanding the flow of people, goods, and other economic activity around the globe."

A Appendix

A.1 Proof of Proposition 1

We know that:

$\frac{dr}{ds} = \frac{\gamma i}{-\beta_0 i s} = -\frac{1}{\mathcal{R}_0 s}$, i.e., $\frac{ds}{\mathcal{R}_0 s} = -dr$

$\forall t \in \mathbb{R}$, $\int_{s(-\infty)}^{s(t)} \frac{ds}{\mathcal{R}_0 s} = -\int_{r(-\infty)}^{r(t)} dr$, thus $\frac{1}{\mathcal{R}_0}\left(\ln s(t) - \ln s(-\infty)\right) = -\left(r(t) - r(-\infty)\right)$

i.e., $\forall t \in \mathbb{R}$, $\frac{1}{\mathcal{R}_0} \ln s(t) + r(t) = \frac{1}{\mathcal{R}_0} \ln s(-\infty) + r(-\infty)$.

The boundary conditions $s(-\infty) = 1$, $r(-\infty) = 0$ give $\frac{1}{\mathcal{R}_0} \ln s(t) + r(t) = 0$, $\forall t \in \mathbb{R}$, which gives (4).

Similarly, we know that

$\frac{di}{ds} = \frac{\beta_0 i s - \gamma i}{-\beta_0 i s} = -1 + \frac{1}{\mathcal{R}_0 s}$, i.e., $di = \left(-1 + \frac{1}{\mathcal{R}_0 s}\right) ds$

$\forall t \in \mathbb{R}$, $\int_{i(-\infty)}^{i(t)} di = \int_{s(-\infty)}^{s(t)} \left(-1 + \frac{1}{\mathcal{R}_0 s}\right) ds$, thus $i(t) - i(-\infty) = s(-\infty) - s(t) + \frac{1}{\mathcal{R}_0}\left(\ln s(t) - \ln s(-\infty)\right)$, i.e., $i(t) + s(t) - \frac{1}{\mathcal{R}_0} \ln s(t) = i(-\infty) + s(-\infty) - \frac{1}{\mathcal{R}_0} \ln s(-\infty)$.

The boundary conditions $s(-\infty) = 1$, $i(-\infty) = 0$ give $i(t) + s(t) - \frac{1}{\mathcal{R}_0} \ln s(t) = 1$, $\forall t \in \mathbb{R}$, which gives (5).

$\frac{ds}{dt} = -\beta_0 i(t) s(t) = -\beta_0 s(t)\left[1 - s(t) + \frac{1}{\mathcal{R}_0} \ln s(t)\right]$

thus $\frac{ds}{-\beta_0 s\left[1 - s + \frac{1}{\mathcal{R}_0} \ln s\right]} = dt$

i.e., $t = \int_{s(0)}^{s(t)} \frac{ds}{-\beta_0 s\left[1 - s + \frac{1}{\mathcal{R}_0} \ln s\right]} = \int_{s(t)}^{s(0)} \frac{ds}{\beta_0 s\left[1 - s + \frac{1}{\mathcal{R}_0} \ln s\right]}$.

$t \mapsto s(t)$ is a decreasing function since $\frac{ds}{dt} = -\beta_0 i s < 0$

$t \mapsto r(t)$ is an increasing function since $\frac{dr}{dt} = \gamma i > 0$, so we have proven (i) of Proposition 1.

From $i(t) = 1 - s(t) + \frac{1}{\mathcal{R}_0} \ln(s(t))$, we get:

$i'(t) = -s'(t) + \frac{s'(t)}{s(t)\mathcal{R}_0} = -s'(t)\left[1 - \frac{1}{s(t)\mathcal{R}_0}\right]$ which is of the sign of $1 - \frac{1}{s(t)\mathcal{R}_0}$ since $s'(t) < 0$. Thus $i(t)$ is maximal when $s(t) = \frac{1}{\mathcal{R}_0}$, which gives

$i_{\max} = 1 - \frac{1}{\mathcal{R}_0} + \frac{1}{\mathcal{R}_0} \ln(\frac{1}{\mathcal{R}_0}) = 1 - \frac{1}{\mathcal{R}_0}\left[1 + \ln(\mathcal{R}_0)\right]$, which proves (ii).

At the end of the epidemic, $i = 0$ thus $1 - s + \frac{1}{\mathcal{R}_0} \ln s = 0$, i.e., $\frac{\ln(s)}{\mathcal{R}_0} = s - 1$ then $\mathcal{R}_0 = \frac{\ln(s)}{s-1}$. □

A.2 Proof of Proposition 2

Proposition 1 gives the dynamics for $t < t_0$.

For $t \in [t_0, T]$,

$\frac{dr}{ds} = \frac{\gamma i}{-\beta i s} = -\frac{1}{\mathcal{R}s}$, i.e., $\frac{ds}{\mathcal{R}s} = -dr$

$\forall t \in [t_0, T]$, $\int_{s(t_0)}^{s(t)} \frac{ds}{\mathcal{R}s} = -\int_{r(t_0)}^{r(t)} dr$, i.e., $\frac{1}{\mathcal{R}}\left(\ln s(t) - \ln s(t_0)\right) = r(t_0) - r(t)$

$\frac{di}{ds} = \frac{\beta i s - \gamma i}{-\beta i s} = -1 + \frac{1}{\mathcal{R}s}$, i.e., $di = \left(-1 + \frac{1}{\mathcal{R}s}\right) ds$

$\forall t \in [t_0, T]$, $\int_{i(t_0)}^{i(t)} di = \int_{s(t_0)}^{s(t)} \left(-1 + \frac{1}{\mathcal{R}s}\right) ds$

thus $i(t) - i(t_0) = -\left(s(t) - s(t_0)\right) + \frac{1}{\mathcal{R}}\left(\ln(s(t) - \ln(s(t_0))\right)$.

$\frac{ds}{dt} = -\beta i(t) s(t) = -\beta s(t)\left[i(t_0) - \left(s(t) - s(t_0)\right) + \frac{1}{\mathcal{R}}\left(\ln s(t) - \ln s(t_0)\right)\right]$

thus $\frac{ds}{-\beta s\left[i(t_0) - (s - s(t_0)) + \frac{1}{\mathcal{R}}(\ln s - \ln s(t_0))\right]} = dt$

i.e., $\int_{s(t_0)}^{s(t)} \frac{ds}{-\beta s\left[i(t_0) - (s - s(t_0)) + \frac{1}{\mathcal{R}}(\ln s - \ln s(t_0))\right]} = t - t_0$

To sum up, for $t \in [t_0, T]$:

$r(t) = r(t_0) + \frac{1}{\mathcal{R}} \ln s(t_0) - \frac{1}{\mathcal{R}} \ln s(t)$

$i(t) = -s(t) + \frac{1}{\mathcal{R}} \ln s(t) + i(t_0) + s(t_0) - \frac{1}{\mathcal{R}} \ln s(t_0)$

$\int_{s(t)}^{s(t_0)} \frac{ds}{\beta s\left[i(t_0) + s(t_0) - \frac{1}{\mathcal{R}} \ln s(t_0) - s + \frac{1}{\mathcal{R}} \ln s\right]} = t - t_0$

which gives the second part of Proposition 2 (i).

It is assumed that $s(t_0) > \frac{1}{\mathcal{R}_0}$, thus the maximum of $i(t)$ for $t \in]-\infty; t_0]$ is attained at t_0, thus $\max_{t \in]-\infty; T]} i(t) = \max_{t \in [t_0, T]} i(t)$.

Moreover, on $t \in [t_0, T]$, $i'(t) = -s'(t) + \frac{s'(t)}{\mathcal{R}s(t)} = -s'(t)\left[1 - \frac{1}{\mathcal{R}s(t)}\right]$ with $s'(t) < 0$, thus the maximum of $i(t)$ for $t \in [t_0, T]$ is attained at t such that $s(t) = \frac{1}{\mathcal{R}}$ if $s(t_0) \geq \frac{1}{\mathcal{R}}$, and is attained at $t = t_0$ if $s(t_0) < \frac{1}{\mathcal{R}}$. This gives Proposition 2 (ii).

Now, we prove that if $\mathcal{R}' > \mathcal{R}$, the curve $\left(s_{\mathcal{R}'}(t), i_{\mathcal{R}'}(t)\right)_{t \geq t_0}$ is strictly above $\left(s_{\mathcal{R}}(t), i_{\mathcal{R}}(t)\right)_{t \geq t_0}$ in the plane (s, i), except a unique common point at $\left(s(t_0), i(t_0)\right)$.

We need to study i as a function of s. Since

$i_{\mathcal{R}}(t) = i(t_0) + s(t_0) - s(t) + \frac{1}{\mathcal{R}} \ln\left(s(t)/s(t_0)\right)$ according to Eq. (12).

We set:

$$\mathcal{I}_{\mathcal{R}}(s) = i(t_0) + s(t_0) - s + \frac{1}{\mathcal{R}} \ln\left(s/s(t_0)\right). \tag{29}$$

It is clear that $s = s_{\mathcal{R}}(t) \Rightarrow \mathcal{I}_{\mathcal{R}}(s) = i_{\mathcal{R}}(t)$. It means that in the plane (s,i), the curve $(s_{\mathcal{R}}(t), i_{\mathcal{R}}(t))_{t\geq t_0}$ is the curve $(s, \mathcal{I}_{\mathcal{R}}(s))_{s\leq s(t_0)}$.

We need to prove that $(s, \mathcal{I}_{\mathcal{R}'}(s))_{s\leq s(t_0)}$ is strictly above $(s, \mathcal{I}_{\mathcal{R}}(s))_{s\leq s(t_0)}$ if $\mathcal{R}' > \mathcal{R}$, except at $s = s(t_0)$.

$\mathcal{I}_{\mathcal{R}'}(s) - \mathcal{I}_{\mathcal{R}}(s) = [i(t_0) + s(t_0) - s + \frac{1}{\mathcal{R}'} \ln(s/s(t_0))] - [i(t_0) + s(t_0) - s + \frac{1}{\mathcal{R}} \ln(s/s(t_0))]$
$= \frac{1}{\mathcal{R}'} \ln(s/s(t_0)) - \frac{1}{\mathcal{R}} \ln(s/s(t_0)) = (\frac{1}{\mathcal{R}} - \frac{1}{\mathcal{R}'}) \ln(s(t_0)/s) > 0$ if $s < s(t_0)$, since $\frac{1}{\mathcal{R}} > \frac{1}{\mathcal{R}'}$.

Now we prove that $s_{\mathcal{R}'}(t) < s_{\mathcal{R}}(t)$ if $\mathcal{R}' > \mathcal{R}$.

According to Eq. (13), $\gamma(t-t_0) = \int_{s_{\mathcal{R}}(t)}^{s(t_0)} \frac{ds}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)}$ and $\gamma(t-t_0) = \int_{s_{\mathcal{R}'}(t)}^{s(t_0)} \frac{ds}{\mathcal{R}' s \mathcal{I}_{\mathcal{R}'}(s)}$

$\forall s$, $\mathcal{I}_{\mathcal{R}'}(s) \geq \mathcal{I}_{\mathcal{R}}(s)$, thus $\forall s$, $\mathcal{R}'\mathcal{I}_{\mathcal{R}'}(s) > \mathcal{R}\mathcal{I}_{\mathcal{R}}(s)$,

$\forall s$, $\frac{1}{\mathcal{R}'\mathcal{I}_{\mathcal{R}'}(s)} < \frac{1}{\mathcal{R}\mathcal{I}_{\mathcal{R}}(s)}$ with $\int_{s_{\mathcal{R}}(t)}^{s(t_0)} \frac{ds}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)} = \int_{s_{\mathcal{R}'}(t)}^{s(t_0)} \frac{ds}{\mathcal{R}' s \mathcal{I}_{\mathcal{R}'}(s)}$

thus the interval $[s_{\mathcal{R}'}(t), s(t_0)]$ must be larger than $[s_{\mathcal{R}}(t), s(t_0)]$, i.e. $s_{\mathcal{R}'}(t) < s_{\mathcal{R}}(t)$.

This gives Proposition 2 (iii). $\square$

A.3 Proof of Proposition 3

V_T is assumed to be concave, with $V_T(\mathcal{R}) = (T-t_0)y(\mathcal{R}) - \lambda N\delta(1 - s_{\mathcal{R}}(T))$ and $V_T'(\mathcal{R}) = (T-t_0)y'(\mathcal{R}) + \lambda N\delta \frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}$ (where $y' > 0$ and $\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} < 0$ according to Proposition 2 (iii)).

We can distinguish 3 cases:

- First corner solution: $\mathcal{R}^{opt} = 0$ if $V_T'(0) \leq 0$. But $V_T'(0) \leq 0$ is not possible since we have assumed that $y'(0) = +\infty$. Thus $V_T'(0) > 0$ in the sequel.

- Second corner solution: $\mathcal{R}^{opt} = \mathcal{R}_0$ if $V_T'(\mathcal{R}_0) \geq 0$, i.e., if $(T-t_0)y'(\mathcal{R}_0) \geq -\lambda\delta N \left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\mathcal{R}_0}$, i.e. if $\lambda \leq \lambda_0$, setting $\lambda_0 = \frac{(T-t_0)y'(\mathcal{R}_0)}{-\delta N \left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\mathcal{R}_0}} \geq 0$.

- Interior solution: $0 < \mathcal{R}^{opt} < \mathcal{R}_0$ if $V_T'(\mathcal{R}_0) < 0 < V_T'(0)$, which is true if $\lambda > \lambda_0$. Here $\mathcal{R}^{opt}$ satisfies $V_T'(\mathcal{R}^{opt}) = 0$.

In this last case, applying the implicit function theorem:
$\frac{d\mathcal{R}^{opt}}{d\lambda} = -\frac{\frac{\partial V_T'}{\partial \lambda}}{\frac{\partial V_T'}{\partial \mathcal{R}}} = -\frac{\delta N \frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}}{V_T"(\mathcal{R})} < 0$ since $V_T" < 0$ and $\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} < 0$. This gives Prop 3

(i), except $\lim_{\lambda\to+\infty}\mathcal{R}^{opt}=0$ proven below.

$\frac{d\mathcal{R}^{opt}}{dt_0}=-\frac{\frac{\partial V_T'}{\partial t_0}}{\frac{\partial V_T'}{\partial \mathcal{R}}}=\frac{-y'(\mathcal{R}^{opt})+\lambda\delta N\frac{\partial}{\partial t_0}\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\left(\mathcal{R}^{opt}\right)\right)}{-V_T"(\mathcal{R})}$ where $-V_T">0$. This gives Prop 3 (ii).

$\frac{d\mathcal{R}^{opt}}{dT}=-\frac{\frac{\partial V_T'}{\partial T}}{\frac{\partial V_T'}{\partial \mathcal{R}}}=\frac{y'\left(\mathcal{R}^{opt}\right)+\lambda\delta N\frac{\partial}{\partial T}\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\left(\mathcal{R}^{opt}\right)\right)}{-V_T"(\mathcal{R})}$. This gives Prop 3 (iii).

If $y(\mathcal{R})=A\mathcal{R}^{\alpha}$, then $y'(\mathcal{R})=\alpha A\mathcal{R}^{\alpha-1}$, and on $\lambda\geq\lambda_0$:

$\frac{d\mathcal{R}^{opt}}{dA}=\frac{\frac{\partial V_T'}{\partial A}}{-V_T"}=\frac{(T-t_0)\alpha\mathcal{R}^{\alpha-1}}{-V_T"(\mathcal{R})}>0$

$\frac{d\mathcal{R}^{opt}}{d\alpha}=\frac{\frac{\partial V_T'}{\partial \alpha}}{-V_T"}=\frac{(T-t_0)A\left[\mathcal{R}^{\alpha-1}+\alpha\ln(\mathcal{R})\mathcal{R}^{\alpha-1}\right]}{-V_T"(\mathcal{R})}$ which is positive if and only if $1+\alpha\ln(\mathcal{R})>0$.

Now let us prove that $\lim_{\lambda\to+\infty}\mathcal{R}^{opt}=0$.

According to Eq. (13), $\gamma(T-t_0)=\int_{s_{\mathcal{R}}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}s\mathcal{I}_{\mathcal{R}}(s)}=\int_{s_{\mathcal{R}'}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}'}(s)}$ if $\mathcal{R}<\mathcal{R}'$.

If $\mathcal{R}<\mathcal{R}'$, $s_{\mathcal{R}'}(T)<s_{\mathcal{R}}(T)$ and $\mathcal{I}_{\mathcal{R}'}(s)\geq\mathcal{I}_{\mathcal{R}}(s)$

$$\int_{s_{\mathcal{R}'}(T)}^{s_{\mathcal{R}}(T)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}}(s)}+\int_{s_{\mathcal{R}}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}}(s)}=\int_{s_{\mathcal{R}'}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}}(s)}\geq\int_{s_{\mathcal{R}'}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}'}(s)}=\gamma(T-t_0)$$

thus

$$\int_{s_{\mathcal{R}'}(T)}^{s_{\mathcal{R}}(T)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}}(s)}\geq\gamma(T-t_0)-\int_{s_{\mathcal{R}}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}'s\mathcal{I}_{\mathcal{R}}(s)}=\gamma(T-t_0)-\frac{\mathcal{R}}{\mathcal{R}'}\int_{s_{\mathcal{R}}(T)}^{s(t_0)}\frac{ds}{\mathcal{R}s\mathcal{I}_{\mathcal{R}}(s)}$$

$$=\gamma(T-t_0)-\frac{\mathcal{R}}{\mathcal{R}'}\gamma(T-t_0)$$

i.e.

$$\frac{1}{\mathcal{R}'}\int_{s_{\mathcal{R}'}(T)}^{s_{\mathcal{R}}(T)}\frac{ds}{s\mathcal{I}_{\mathcal{R}}(s)}\geq\gamma(T-t_0)\left[1-\frac{\mathcal{R}}{\mathcal{R}'}\right]$$

$$\int_{s_{\mathcal{R}'}(T)}^{s_{\mathcal{R}}(T)}\frac{ds}{s\mathcal{I}_{\mathcal{R}}(s)}\geq\gamma(T-t_0)(\mathcal{R}'-\mathcal{R})$$

$$\frac{1}{(\mathcal{R}'-\mathcal{R})}\int_{s_{\mathcal{R}'}(T)}^{s_{\mathcal{R}}(T)}\frac{ds}{s\mathcal{I}_{\mathcal{R}}(s)}\geq\gamma(T-t_0)$$

with $\mathcal{R}'\to\mathcal{R}$, we have

$$\lim_{\mathcal{R}'\to\mathcal{R}^+}\frac{1}{(\mathcal{R}'-\mathcal{R})}\int_{s_{\mathcal{R}'}(T)}^{s_{\mathcal{R}}(T)}\frac{ds}{s\mathcal{I}_{\mathcal{R}}(s)}=\frac{1}{s_{\mathcal{R}}(T)\mathcal{I}_{\mathcal{R}}(s_{\mathcal{R}}(T))}\times\lim_{\mathcal{R}'\to\mathcal{R}^+}\frac{s_{\mathcal{R}}(T)-s_{\mathcal{R}'}(T)}{(\mathcal{R}'-\mathcal{R})}$$

$$=\frac{1}{s_{\mathcal{R}}(T)i_{\mathcal{R}}(T)}\times\frac{-\partial s_{\mathcal{R}}(T)}{\partial\mathcal{R}}$$

thus

$$\frac{-\partial s_{\mathcal{R}}(T)}{\partial\mathcal{R}}\geq\gamma(T-t_0)s_{\mathcal{R}}(T)\,i_{\mathcal{R}}(T)\,. \tag{30}$$

Now we go back to the proof of $\lim_{\lambda\to+\infty}\mathcal{R}^{opt}=0$.

In the case of an interior solution $V_T'(\mathcal{R}^{opt})=0$, i.e.

$$(T-t_0)y'(\mathcal{R}^{opt})+\lambda N\delta\frac{\partial s_{\mathcal{R}}(T)}{\partial\mathcal{R}}=0$$

$(T-t_0)y'(\mathcal{R}^{opt})=-\lambda N\delta\frac{\partial s_{\mathcal{R}}(T)}{\partial\mathcal{R}}\geq\lambda N\delta\gamma(T-t_0)s_{\mathcal{R}}(T)\,i_{\mathcal{R}}(T)$ according to (30).

$y'(\mathcal{R}^{opt}) \geq \lambda N \delta \gamma s_{\mathcal{R}}(T)\, i_{\mathcal{R}}(T)$

Since $T < +\infty$, there exists $\eta > 0$ such that:

$s_{\mathcal{R}}(T)\, i_{\mathcal{R}}(T) \geq \eta > 0$ for all $\mathcal{R}$ (T given).

$y'(\mathcal{R}^{opt}) \geq \lambda N \delta \gamma \eta$

When $\lambda \to +\infty$, then $y'(\mathcal{R}^{opt}) \to +\infty$, i.e. $\mathcal{R}^{opt} \to 0$. $\square$

A.4 Proof of Proposition 4

$\frac{dr}{dt} = \gamma i(t)$ thus without hospital contraints: $m_\infty = \delta r_\infty = \delta \int_{-\infty}^{+\infty} dr = \delta \int_{-\infty}^{+\infty} \gamma i(t) dt$.

With hospital constraints, we get:

$m_\infty = \delta \int_{-\infty}^{+\infty} \gamma i(t) dt + (\delta' - \delta) \int_{i(t)>\bar{i}} \gamma \left(i(t) - \bar{i}\right) dt$.

The surmortality due to hospital constraints is $(\delta' - \delta) z(\mathcal{R}, \bar{i})$ where:

$z(\mathcal{R}, \bar{i}) = \int_{i(t)>\bar{i}} \gamma \left(i(t) - \bar{i}\right) dt = \int_{i(t)>\bar{i}} \gamma i(t) dt - \int_{i(t)>\bar{i}} \gamma \bar{i} dt$.

$[T_1, T_2]$ denoting the period during which $i(t) \geq \bar{i}$, with $t_0 < T_1 < T_2 < T$ by assumption, we have:

$z(\mathcal{R}, \bar{i}) = \int_{r_{T_1}}^{r_{T_2}} dr - \gamma \bar{i} \int_{T_1}^{T_2} dt = r(T_2) - r(T_1) - \gamma \bar{i}(T_2 - T_1)$.

- Let us show that $z(\mathcal{R}, \bar{i})$ is an increasing function of $\mathcal{R}$.

$i(T_1) + s(T_1) + r(T_1) = i(T_2) + s(T_2) + r(T_2) = 1$ where $i(T_1) = i(T_2) = \bar{i}$

thus $z(\mathcal{R}, \bar{i}) = s(T_1) - s(T_2) - \gamma \bar{i}(T_2 - T_1)$.

$\frac{1}{\bar{i}} z(\mathcal{R}, \bar{i}) = \frac{s(T_1) - s(T_2)}{\bar{i}} - \int_{s(T_2)}^{s(T_1)} \frac{ds}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)}$ according to Eq. (13), and setting $\mathcal{I}_{\mathcal{R}}(s) = i(t_0) + s(t_0) - s + \frac{1}{\mathcal{R}} \ln(s) - \frac{1}{\mathcal{R}} \ln(s(t_0))$ as in Eq. (29), we write

$\frac{1}{\bar{i}} z(\mathcal{R}, \bar{i}) = \int_{s(T_2)}^{s(T_1)} \left[\frac{1}{\bar{i}} - \frac{1}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)}\right] ds$. Here $s(T_1) = s_{\mathcal{R}}(T_1)$ is an increasing function of $\mathcal{R}$, $s(T_2) = s_{\mathcal{R}}(T_2)$ is a decreasing function of $\mathcal{R}$, and $\left[\frac{1}{\bar{i}} - \frac{1}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)}\right]$ is an increasing function of $\mathcal{R}$. Thus $z(\mathcal{R}, \bar{i})$ is an increasing function of $\mathcal{R}$, with $z(\overline{\mathcal{R}}, \bar{i}) = 0$ since $T_1 = T_2$ and $s(T_1) = s(T_2)$ if $\mathcal{R} = \overline{\mathcal{R}}$, and

$z(\mathcal{R}, \bar{i}) > 0$ if $\mathcal{R} > \overline{\mathcal{R}}$.

- Let us show that $z(\mathcal{R}, \bar{i})$ is a decreasing function of $\bar{i}$.

$z(\mathcal{R}, \bar{i}) = \int_{s(T_2)}^{s(T_1)} \left[1 - \frac{\bar{i}}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)}\right] ds$ where $\left[1 - \frac{\bar{i}}{\mathcal{R} s \mathcal{I}_{\mathcal{R}}(s)}\right]$ is a decreasing function of $\bar{i}$, $s(T_1)$ is a decreasing function of $\bar{i}$, $s(T_2)$ is an increasing function of $\bar{i}$. Thus $z(\mathcal{R}, \bar{i})$ is a decreasing function of $\bar{i}$. $\square$

A.5 Proof of Proposition 5

Let $V_T^* = \max_{0 \le \mathcal{R} \le \mathcal{R}_0} V_T(\mathcal{R}) = \max\left(V_1^*, V_2^*\right)$, where:

$V_1^* = \max_{0 \le \mathcal{R} \le \overline{\mathcal{R}}} V_T(\mathcal{R}) = \max_{0 \le \mathcal{R} \le \overline{\mathcal{R}}} [(T - t_0)y(\mathcal{R}) - \lambda\delta N(1 - s_{\mathcal{R}}(T))]$.

$V_2^* = \max_{\overline{\mathcal{R}} \le \mathcal{R} \le \mathcal{R}_0} V_T(\mathcal{R})$

$= \max_{\overline{\mathcal{R}} \le \mathcal{R} \le \mathcal{R}_0} [(T - t_0)y(\mathcal{R}) - \lambda\delta N(1 - s_{\mathcal{R}}(T)) - \lambda(\delta' - \delta)Nz(\mathcal{R})]$.

$V_T(\mathcal{R})$ is assumed to be concave on $0 \le \mathcal{R} \le \mathcal{R}_0$.

We denote by $V'_{T^-}(\overline{\mathcal{R}})$ the left-hand derivative of V_T at $\overline{\mathcal{R}}$, and by $V'_{T^+}(\overline{\mathcal{R}})$ the right-hand derivative of V_T at $\overline{\mathcal{R}}$.

$V'_{T^-}(\overline{\mathcal{R}}) = (T - t_0)y'(\overline{\mathcal{R}}) + \lambda\delta N \left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\overline{\mathcal{R}}}$

$V'_{T^+}(\overline{\mathcal{R}}) = (T - t_0)y'(\overline{\mathcal{R}}) + \lambda\delta N \left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\overline{\mathcal{R}}} - \lambda(\delta' - \delta)Nz'_+(\overline{\mathcal{R}})$.

Since $y'(0) = +\infty$ thus $V'_T(0) > 0$, two cases are possible on $0 \le \mathcal{R} \le \overline{\mathcal{R}}$:

- If $V'_{T^-}(\overline{\mathcal{R}}) \ge 0$ by concavity of V_T, we have $V_1^* = V_T(\overline{\mathcal{R}})$, where $V'_{T^-}(\overline{\mathcal{R}}) \ge 0$ is equivalent to $\lambda \le \frac{(T-t_0)y'(\overline{\mathcal{R}})}{-\delta N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\overline{\mathcal{R}}}} := \lambda_3$.

- If $V'_{T^-}(\overline{\mathcal{R}}) < 0$ then $V_1^* = V_T(\mathcal{R}^*)$ where $\mathcal{R}^* \in (0; \overline{\mathcal{R}})$, and $\mathcal{R}^*$ satisfies the first order condition $(T - t_0)y'(\mathcal{R}) + \lambda\delta N\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} = 0$.

Note that $V'_{T^-}(\overline{\mathcal{R}}) < 0$ is equivalent to $\lambda > \lambda_3$.

Three cases on $\overline{\mathcal{R}} \le \mathcal{R} \le \mathcal{R}_0$:

- If $V'_{T^+}(\overline{\mathcal{R}}) \le 0$ then $V_2^* = V_T(\overline{\mathcal{R}})$, where $V'_{T^+}(\overline{\mathcal{R}}) \le 0$ is equivalent to $\lambda \ge \frac{(T-t_0)y'(\overline{\mathcal{R}})}{-\delta N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\overline{\mathcal{R}}} + (\delta' - \delta)Nz'_+(\overline{\mathcal{R}})} := \lambda_2$.

- If $V'_T(\mathcal{R}_0) \ge 0$ then $V_2^* = V_T(\mathcal{R}_0)$.

But $V'_T(\mathcal{R}_0) \ge 0$ means that $\lambda \le \frac{(T-t_0)y'(\mathcal{R}_0)}{-\delta N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\mathcal{R}_0} + (\delta' - \delta)Nz'(\mathcal{R}_0)} := \lambda_1$.

- If $V'_T(\mathcal{R}_0) < 0 < V'_{T^+}(\overline{\mathcal{R}})$ then $V_2^* = V_T(\mathcal{R}^{**})$ where $\mathcal{R}^{**} \in (\overline{\mathcal{R}}; \mathcal{R}_0)$, and $\mathcal{R}^{**}$ satisfies the first order condition $(T - t_0)y'(\mathcal{R}) + \lambda\delta N\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} - \lambda(\delta' - \delta)Nz'(\mathcal{R}) = 0$. Note that $V'_T(\mathcal{R}_0) < 0 < V'_{T^+}(\overline{\mathcal{R}})$ is equivalent to $\lambda_1 < \lambda < \lambda_2$.

Since V_T is concave, we have $V'_{T^-}\left(\overline{\mathcal{R}}\right) > V'_{T^+}(\overline{\mathcal{R}}) > V'(\mathcal{R}_0)$, which implies that

$\lambda_3 > \lambda_2 > \lambda_1 > 0$.

Now we study V_T^*.

1. If $\lambda \leq \lambda_1$, then $V_1^* = V_T(\overline{\mathcal{R}})$ and $V_2^* = V_T(\mathcal{R}_0)$, thus $V_T^* = V_T(\mathcal{R}_0)$. Thus here $\mathcal{R}_h^{opt} = \mathcal{R}_0$, which gives Prop 5(i).

2. If $\lambda_1 < \lambda < \lambda_2$ then $V_1^* = V_T(\overline{\mathcal{R}})$, $V_2^* = V_T(\mathcal{R}^{**})$ and $V_T^* = V_T(\mathcal{R}^{**})$, where $\mathcal{R}^{**} \in (\overline{\mathcal{R}}; \mathcal{R}_0)$, and $\mathcal{R}^{**}$ satisfies the first order condition $V_T'(\mathcal{R}) = (T - t_0)y'(\mathcal{R}) + \lambda\delta N \frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} - \lambda(\delta' - \delta)Nz'(\mathcal{R}) = 0$.

Thus here $\mathcal{R}_h^{opt} = \mathcal{R}^{**}$, and applying the implicit function theorem, we get

$$\frac{d\mathcal{R}_h^{opt}}{d\lambda} = \frac{d\mathcal{R}^{**}}{d\lambda} = -\frac{\frac{d}{d\lambda}\left((T-t_0)y'(\mathcal{R})+\lambda\delta N\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}-\lambda(\delta'-\delta)Nz'(\mathcal{R})\right)}{V_T''(\mathcal{R})} = \frac{\delta N\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}-(\delta'-\delta)Nz'(\mathcal{R})}{-V_T\text{”}(\mathcal{R})} < 0,$$

since $V_T'' < 0$, $\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} < 0$ and $z' > 0$, which gives Prop 5(ii).

3. If $\lambda_2 \leq \lambda \leq \lambda_3$ then $V_1^* = V_T(\overline{\mathcal{R}})$, $V_2^* = V_T(\overline{\mathcal{R}})$ and $V_T^* = V_T(\overline{\mathcal{R}})$. Thus here $\mathcal{R}_h^{opt} = \overline{\mathcal{R}}$ which does not depend on λ.

$$i_{\max}(\overline{\mathcal{R}}) = i(t_0) + s(t_0) - \frac{1}{\overline{\mathcal{R}}} - \frac{\ln(\overline{\mathcal{R}})}{\overline{\mathcal{R}}} - \frac{\ln(s(t_0))}{\overline{\mathcal{R}}} \text{ with } i_{\max}(\overline{\mathcal{R}}) = \bar{i}$$

$$\frac{d\overline{\mathcal{R}}}{dt_0} = -\frac{\frac{\partial}{\partial t_0}i_{\max}(\overline{\mathcal{R}})}{\frac{\partial}{\partial \overline{\mathcal{R}}}i_{\max}(\overline{\mathcal{R}})} = -\frac{\left(\frac{di}{dt_0}(t_0)+\frac{ds}{dt_0}(t_0)-\frac{1}{\overline{\mathcal{R}}s(t_0)}\frac{ds}{dt_0}(t_0)\right)}{\left(\frac{1}{\overline{\mathcal{R}}^2}-\frac{(1-\ln(\overline{\mathcal{R}})}{\overline{\mathcal{R}}^2}+\frac{1}{\overline{\mathcal{R}}^2}\ln(s(t_0))\right)} = -\frac{\left(\beta_0 i(t_0)s(t_0)-\gamma i(t_0)-\beta_0 i(t_0)s(t_0)+\frac{\beta_0 i(t_0)}{\overline{\mathcal{R}}}\right)}{\left(\frac{\ln(\overline{\mathcal{R}})+\ln(s(t_0))}{\overline{\mathcal{R}}^2}\right)}$$

$$= -\frac{i(t_0)\left(\frac{\beta_0}{\overline{\mathcal{R}}}-\gamma\right)}{\left(\frac{\ln(\overline{\mathcal{R}}s(t_0))}{\overline{\mathcal{R}}^2}\right)} = -\frac{\gamma i(t_0)\left(\frac{\mathcal{R}_0}{\overline{\mathcal{R}}}-1\right)}{\left(\frac{\ln(\overline{\mathcal{R}}s(t_0))}{\overline{\mathcal{R}}^2}\right)} < 0, \text{ since } \mathcal{R}_0 > \overline{\mathcal{R}} \text{ and } s(t_0) > \frac{1}{\overline{\mathcal{R}}}.$$

Thus $\overline{\mathcal{R}}$ is a decreasing function of t_0 , which gives Prop 5(iii).

4. If $\lambda > \lambda_3$ then $V_1^* = V_T(\mathcal{R}^*)$, $V_2^* = V_T(\overline{\mathcal{R}})$ and $V_T^* = V_T(\mathcal{R}^*)$ where $\mathcal{R}^* \in (0; \overline{\mathcal{R}};)$, and $\mathcal{R}^*$ satisfies the first order condition $(T - t_0)y'(\mathcal{R}) + \lambda\delta N\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}} = 0$.

Thus here $\mathcal{R}_h^{opt} = \mathcal{R}^*$, and:

$$\frac{d\mathcal{R}_h^{opt}}{d\lambda} = \frac{d\mathcal{R}^*}{d\lambda} = -\frac{\frac{d}{d\lambda}\left((T-t_0)y'(\mathcal{R}^*)+\lambda\delta N\left(\frac{\partial s_{\mathcal{R}}(T)}{\partial \mathcal{R}}\right)_{\mathcal{R}=\mathcal{R}^*}\right)}{V_T''(\mathcal{R}^*)} = \frac{\delta N\left(\frac{\partial s_{\mathcal{R}}}{\partial \mathcal{R}}(T)\right)_{\mathcal{R}=\mathcal{R}^*}}{-V_T\text{”}(\mathcal{R}^*)} < 0$$

since $V_T'' < 0$ and $\frac{\partial s}{\partial \mathcal{R}} < 0$. This gives Prop 5(iv). $\square$

A.6 Proof of Proposition 6

(i) According to (21), for $t \geq T$, we have

$$i(t) = i(T) + s(T) - s(t) + \frac{1}{\mathcal{R}_0}\ln(s(t)) - \frac{1}{\mathcal{R}_0}\ln(s(T))$$

thus $i'(t) = -s'(t) + \frac{s'(t)}{\mathcal{R}_0 s(t)} = -s'(t)\left[1 - \frac{1}{\mathcal{R}_0 s(t)}\right]$, since $s'(t) < 0$, thus $i'(t)$ is of the sign of $1 - \frac{1}{\mathcal{R}_0 s(t)}$.

There is a rebound after T if and only if $i'(T) > 0$, i.e. if $1 > \frac{1}{\mathcal{R}_0 s(T)}$, which means that there is a rebound after T if and only if $s(T) > \frac{1}{\mathcal{R}_0}$.
If there is a rebound, $i(t)$ is maximal on $t \geq T$ when $i'(t) = 0$, i.e. for $s(t) = \frac{1}{\mathcal{R}_0}$.
If there is no rebound, $i(t)$ is a decreasing function on $t \geq T$, thus $i(t)$ is maximal on $t \geq T$ when $t = T$.

(ii) There is a rebound after $T \iff s(T) > \frac{1}{\mathcal{R}_0}$.
We must study under what conditions on $\mathcal{R}$ and T do we have $s(T) > \frac{1}{\mathcal{R}_0}$.
$s(T)$ is a decreasing function of T; we set $\widetilde{s}_\infty = \lim_{T \to +\infty} s(T)$, where $\widetilde{s}_\infty$ is the value of $s(t)$ such that $i(t) = 0$ in (12).
Thus $i(t_0) + s(t_0) - \widetilde{s}_\infty + \frac{1}{\mathcal{R}} \ln(\widetilde{s}_\infty) - \frac{1}{\mathcal{R}} \ln(s(t_0)) = 0$ and

$\mathcal{R} = \frac{\ln(s(t_0)) - \ln(\widetilde{s}_\infty)}{i(t_0) + s(t_0) - \widetilde{s}_\infty}$.

We claim that $\widetilde{s}_\infty$ is a decreasing function of $\mathcal{R}$.
Indeed, $\frac{d\mathcal{R}}{d\widetilde{s}_\infty} = \frac{-\frac{1}{\widetilde{s}_\infty}\left(i(t_0)+s(t_0)-\widetilde{s}_\infty\right)+\left(\ln(s(t_0))-\ln(\widetilde{s}_\infty)\right)}{\left(i(t_0)+s(t_0)-\widetilde{s}_\infty\right)^2} = \frac{\left(\mathcal{R}-\frac{1}{\widetilde{s}_\infty}\right)\left(i(t_0)+s(t_0)-\widetilde{s}_\infty\right)}{\left(i(t_0)+s(t_0)-\widetilde{s}_\infty\right)^2} < 0$, since $\widetilde{s}_\infty \leq s(T) < \frac{1}{\mathcal{R}}$.
We have: $s(T) > \frac{1}{\mathcal{R}_0}$ for all $T \iff \widetilde{s}_\infty \geq \frac{1}{\mathcal{R}_0}$,
i.e.

$s(T) > \frac{1}{\mathcal{R}_0}$ for all $T \iff \mathcal{R} \leq \widetilde{\mathcal{R}}_{t_0}$, setting $\widetilde{\mathcal{R}}_{t_0} = \frac{\ln(s(t_0)) - \ln\left(\frac{1}{\mathcal{R}_0}\right)}{i(t_0)+s(t_0)-\frac{1}{\mathcal{R}_0}}$,

$s(T) \leq \frac{1}{\mathcal{R}_0}$ for T sufficiently high $\iff \mathcal{R} > \widetilde{\mathcal{R}}_{t_0}$.

A.7 Proof of Proposition 7

(i) s_∞ is the value of $s(t)$ such that $i(t) = 0$ in (21).
Thus $i(T) + s(T) - s_\infty + \frac{1}{\mathcal{R}_0} \ln(s_\infty) - \frac{1}{\mathcal{R}_0} \ln(s(T)) = 0$
which gives $\mathcal{R}_0 = \frac{\ln(s(T)) - \ln(s_\infty)}{i(T) + s(T) - s_\infty}$.
- Let us show that $s_\infty(\mathcal{R}) < \frac{1}{\mathcal{R}_0}$.

If $s(T) \leq \frac{1}{\mathcal{R}_0}$, then $s_\infty(\mathcal{R}) < s(T) \leq \frac{1}{\mathcal{R}_0}$.

If $s(T) > \frac{1}{\mathcal{R}_0}$, there is a rebound after T with an epidemic peak when $s(t) = \frac{1}{\mathcal{R}_0}$, thus $s_\infty(\mathcal{R}) < s(t) = \frac{1}{\mathcal{R}_0}$.
In both cases we have $s_\infty(\mathcal{R}) < \frac{1}{\mathcal{R}_0}$.

- Let us show that $s_\infty(\mathcal{R})$ is an increasing function of T (for a given $\mathcal{R}_0$).

As $\mathcal{R}_0 = \frac{\ln(s(T))-\ln(s_\infty)}{i(T)+s(T)-s_\infty} = \frac{\ln(s(T))-\ln(s_\infty)}{1-r(T)-s_\infty}$, we get: $\mathcal{R}_0\left(1-r(T)-s_\infty\right) = \ln(s(T))-\ln(s_\infty)$.

Derivating with respect to T :

$\mathcal{R}_0\left(-\dot{r}(T)-\dot{s}_\infty\right) = \frac{\dot{s}(T)}{s(T)} - \frac{\dot{s}_\infty}{s_\infty}$.

As $\dot{r}(T) = \frac{dr}{dt}(T) = \gamma i(T)$ and $\dot{s}(T) = \frac{ds}{dt}(T) = -\beta i(T)s(T)$, we get:

$-\mathcal{R}_0\left(\gamma i(T)+\dot{s}_\infty\right) = -\beta i(T) - \frac{\dot{s}_\infty}{s_\infty}$.

As $\dot{s}_\infty\left(\frac{1}{s_\infty}-\mathcal{R}_0\right) = (\beta_0-\beta)i(T)$, we write

$\dot{s}_\infty = \frac{(\beta_0-\beta)i(T)}{\frac{1}{s_\infty}-\mathcal{R}_0} > 0$ which is positive since the numerator and the denominator are both positive.

- Let us show that $s_\infty(\mathcal{R})$ is a decreasing function of $\mathcal{R}_0$ (for a given T).

From (23) we get $\frac{d\mathcal{R}_0}{ds_\infty} = \frac{-\frac{1}{s_\infty}(i(T)+s(T)-s_\infty)+\ln(s(T))-\ln(s_\infty)}{(i(T)+s(T)-s_\infty)^2} = \frac{\left(\mathcal{R}_0-\frac{1}{s_\infty}\right)(i(T)+s(T)-s_\infty)}{(i(T)+s(T)-s_\infty)^2} < 0$

since $s_\infty < \frac{1}{\mathcal{R}_0}$. This proves Prop 7(i).

(ii) The representative curve of the function $s \mapsto \mathcal{I}_{\mathcal{R}_0}(s)$ is above the one representing $s \mapsto \mathcal{I}_{\mathcal{R}}(s)$, thus $s_\infty(\mathcal{R}) \geq s_\infty(\mathcal{R}_0)$.

- Let us show that $s_\infty(\mathcal{R})$ is a non-monotonic function of $\mathcal{R}$ if T is high.

$\mathcal{R}_0 = \frac{\ln(s(T))-\ln(s_\infty)}{i(T)+s(T)-s_\infty}$, thus $\mathcal{R}_0\left(i(T)+s(T)-s_\infty\right) = \ln(s(T))-\ln(s_\infty)$

i.e., $i(T)+s(T)-s_\infty = \frac{1}{\mathcal{R}_0}\ln(s(T)) - \frac{1}{\mathcal{R}_0}\ln(s_\infty)$

$-s_\infty + \frac{1}{\mathcal{R}_0}\ln(s_\infty) = -i(T)-s(T)+\frac{1}{\mathcal{R}_0}\ln(s(T))$

and since $i(T)+s(T) = 1-r(T)$:

$1-s_\infty+\frac{1}{\mathcal{R}_0}\ln(s_\infty) = r(T)+\frac{1}{\mathcal{R}_0}\ln(s(T))$.

If T is sufficiently high, then $s(T) \simeq \widetilde{s}_\infty$ (since $\widetilde{s}_\infty = \lim_{T\to\infty} s(T)$), and $r(T) \simeq 1-\widetilde{s}_\infty$. This leads to:

$1-s_\infty+\frac{1}{\mathcal{R}_0}\ln(s_\infty) \simeq 1-\widetilde{s}_\infty+\frac{1}{\mathcal{R}_0}\ln(\widetilde{s}_\infty)$

i.e.

$h_{\mathcal{R}_0}(s_\infty) \simeq h_{\mathcal{R}_0}(\widetilde{s}_\infty)$

where $h_{\mathcal{R}_0}$ is defined by $h_{\mathcal{R}_0}(s) = 1-s+\frac{1}{\mathcal{R}_0}\ln(s)$.

According to the proof of Prop 6 (ii), $\widetilde{s}_\infty$ is a decreasing function of $\mathcal{R}$ on $\mathcal{R} \in (0;\mathcal{R}_0]$, with $\widetilde{s}_\infty \geq \frac{1}{\mathcal{R}_0}$ if $\mathcal{R} \leq \widetilde{\mathcal{R}}_{t_0}$, and $\widetilde{s}_\infty \leq \frac{1}{\mathcal{R}_0}$ if $\mathcal{R} \geq \widetilde{\mathcal{R}}_{t_0}$.

For a given s_∞, the equation $h_{\mathcal{R}_0}(s_\infty) = h_{\mathcal{R}_0}(\widetilde{s}_\infty)$ of unknown $\widetilde{s}_\infty$ has two different roots: the first one is higher than $\frac{1}{\mathcal{R}_0}$, the second one is lower than $\frac{1}{\mathcal{R}_0}$.

For $\mathcal{R} \in [\widetilde{\mathcal{R}}_{t_0}, \mathcal{R}_0]$, we have $s_\infty = \widetilde{s}_\infty \leq \frac{1}{\mathcal{R}_0}$ and s_∞ is a decreasing function of $\mathcal{R}$.

For $\mathcal{R} \in (0, \widetilde{\mathcal{R}}_{t_0})$, we have $s_\infty \leq \frac{1}{\mathcal{R}_0} < \widetilde{s}_\infty$ and s_∞ is here an increasing function of $\mathcal{R}$.

Summing up, for T high enough, s_∞ is a non-monotonic function of $\mathcal{R}$, it is first increasing, then decreasing.

A.8 Proof of Proposition 8

(i) $\mathcal{R}_\infty^{opt}$ minimizes $L_\infty(\mathcal{R}) = (T - t_0)\,[y(\mathcal{R}_0) - y(\mathcal{R})] + \lambda M_\infty(\mathcal{R})$.

L_∞ is assumed to be a convex function of $\mathcal{R}$, since y is very concave (same argument as for the concavity of $V_T(\mathcal{R})$).

$L'_\infty(\mathcal{R}) = -(T - t_0)y'(\mathcal{R}) - \lambda N\delta s'_\infty(\mathcal{R})$ with $y' > 0$.

We can distinguish 3 cases:

- First corner solution:

$\mathcal{R}_\infty^{opt} = 0$ if $L'_\infty(0) \geq 0$. But $L'_\infty(0) \geq 0$ is not possible since $y'(0) = +\infty$.

- Second corner solution:

$\mathcal{R}_\infty^{opt} = \mathcal{R}_0$ if $L'_\infty(\mathcal{R}_0) \leq 0$, i.e., if $-(T - t_0)y'(\mathcal{R}_0) - \lambda\delta N s'_\infty(\mathcal{R}_0) \leq 0$, i.e. if $\lambda \leq \lambda'_0$, setting $\lambda'_0 = \frac{(T-t_0)y'(\mathcal{R}_0)}{-\delta N s'_\infty(\mathcal{R}_0)} \geq 0$.

- Interior solution: $0 < \mathcal{R}_\infty^{opt} < \mathcal{R}_0$ if $L'_\infty(\mathcal{R}_0) > 0 > L'_\infty(0)$, which is true if $\lambda > \lambda'_0$. Here $\mathcal{R}_\infty^{opt}$ satisfies $L'_\infty(\mathcal{R}_\infty^{opt}) = 0$.

$L'_\infty(\mathcal{R}_\infty^{opt}) = 0$ then $\lambda N\delta s'_\infty(\mathcal{R}_\infty^{opt}) = -(T - t_0)y'(\mathcal{R}_\infty^{opt}) < 0$. Thus $s'_\infty(\mathcal{R}_\infty^{opt}) < 0$ and then according to Proposition 7 (ii), we have $\mathcal{R}_\infty^{opt} > \widetilde{\mathcal{R}}_{t_0}$ if T is sufficiently high.

If $\lambda > \lambda'_0$, applying the implicit function theorem:

$\frac{d\mathcal{R}_\infty^{opt}}{d\lambda} = -\frac{\frac{\partial L'_\infty}{\partial \lambda}}{\frac{\partial L'_\infty}{\partial \mathcal{R}}} = \frac{\delta N s'_\infty}{L_\infty"} < 0$ since $L_\infty" > 0$ and $s'_\infty(\mathcal{R}_\infty^{opt}) < 0$.

This gives Prop 8 (i).

(ii) If $y(\mathcal{R}) = A\mathcal{R}^\alpha$, then $y'(\mathcal{R}) = \alpha A \mathcal{R}^{\alpha-1}$, and on $\lambda \geq \lambda'_0$:

$\frac{d\mathcal{R}^{opt}}{dA} = \frac{\frac{\partial L'_\infty}{\partial A}}{-L_\infty"} = \frac{(T-t_0)\alpha\mathcal{R}^{\alpha-1}}{L_\infty"} > 0$

$\frac{d\mathcal{R}^{opt}}{d\alpha} = \frac{\frac{\partial L'_\infty}{\partial \alpha}}{-L_\infty"} = \frac{(T-t_0)A\left[\mathcal{R}^{\alpha-1}+\alpha\ln(\mathcal{R})\mathcal{R}^{\alpha-1}\right]}{L_\infty"}$ which is positive if and only if $1+\alpha\ln(\mathcal{R}) > 0$.

(iii) For T high enough, we have $\mathcal{R}_\infty^{opt} \geq \widetilde{\mathcal{R}}_{t_0}$. But for λ high enough, $\mathcal{R}^{opt}$ is small, lower than $\widetilde{\mathcal{R}}_{t_0}$. Thus for T and λ high enough, $\mathcal{R}^{opt} < \widetilde{\mathcal{R}}_{t_0} \leq \mathcal{R}^{opt}_{\infty}$. $\square$

A.9 Proof of Proposition 9

First lockdown: $\mathcal{R} = \overline{\mathcal{R}}_1$, where $i_{\max}(\overline{\mathcal{R}}_1) = i(t_0) + s(t_0) - \frac{1}{\overline{\mathcal{R}}_1} - \frac{\ln(\overline{\mathcal{R}}_1)}{\overline{\mathcal{R}}_1} - \frac{\ln(s(t_0))}{\overline{\mathcal{R}}_1}$ and $i_{\max}(\overline{\mathcal{R}}_1) = \bar{i}$.

At the end of the lockdown according to (12), we have $i(t_1) = i(t_0) + s(t_0) - s(t_1) + \frac{\ln(s(t_1) - \ln(s(t_0))}{\overline{\mathcal{R}}_1}$, where t_1 is the date of end of the first lockdown.

We have $i(t_1) = i(t_0)$ by assumption, thus $-s(t_1) + \frac{\ln(s(t_1)}{\overline{\mathcal{R}}_1} = -s(t_0) + \frac{\ln(s(t_0)}{\overline{\mathcal{R}}_1}$, i.e. $h_{\overline{\mathcal{R}}_1}(s(t_1)) = h_{\overline{\mathcal{R}}_1}(s(t_0))$, where $h_{\overline{\mathcal{R}}_1}$ is defined by $h_{\overline{\mathcal{R}}_1}(s) = 1 - s + \frac{1}{\overline{\mathcal{R}}_1}\ln(s)$.
For a given $s(t_0)$, the equation $h_{\overline{\mathcal{R}}_1}(x) = h_{\overline{\mathcal{R}}_1}(s(t_0))$ of unknown x has two different roots: the first one is $s(t_0)$ and is higher than $\frac{1}{\overline{\mathcal{R}}_1}$, the second one is $s(t_1)$ and is lower than $\frac{1}{\overline{\mathcal{R}}_1}$. We know that $\overline{\mathcal{R}}_1$ is an increasing function of $\bar{i}$, so that if $\bar{i}$ is sufficiently high, $s(t_1) < \frac{1}{\mathcal{R}_0}$ and one lockdown is sufficient to achieve herd immunity.

If $\bar{i}$ is not sufficiently high, several lockdowns are neccesary to obtain herd immunity. The k^{th} lockdown is characterized by: $\mathcal{R} = \overline{\mathcal{R}}_k$, where $i_{\max}(\overline{\mathcal{R}}_k) = i(t_{k-1}) + s(t_{k-1}) - \frac{1}{\overline{\mathcal{R}}_k} - \frac{\ln(\overline{\mathcal{R}}_k)}{\overline{\mathcal{R}}_k} - \frac{\ln(s(t_{k-1}))}{\overline{\mathcal{R}}_k}$ and $i_{\max}(\overline{\mathcal{R}}_k) = \bar{i}$.

At the end of the lockdown, we have $i(t_k) = i(t_{k-1}) + s(t_{k-1}) - s(t_k) + \frac{\ln(s(t_k) - \ln(s(t_{k-1}))}{\overline{\mathcal{R}}_k}$, where t_k is the date of end of the k^{th} lockdown.

We have $i(t_k) = i(t_{k-1}) = i(t_0)$ by assumption, thus $-s(t_k) + \frac{\ln(s(t_k)}{\overline{\mathcal{R}}_k} = -s(t_{k-1}) + \frac{\ln(s(t_{k-1})}{\overline{\mathcal{R}}_k}$, i.e. $h_{\overline{\mathcal{R}}_k}(s(t_k)) = h_{\overline{\mathcal{R}}_k}(s(t_{k-1}))$, where $h_{\overline{\mathcal{R}}_k}$ is defined by $h_{\overline{\mathcal{R}}_k}(s) = 1 - s + \frac{1}{\overline{\mathcal{R}}_k}\ln(s)$ then similarly $s(t_k) < \frac{1}{\overline{\mathcal{R}}_k} < s(t_{k-1})$.

We obtain n successive lockdowns, with:

$s(t_0) > \frac{1}{\overline{\mathcal{R}}_1} > s(t_1) > \frac{1}{\overline{\mathcal{R}}_2} > s(t_2) > ... > s(t_{n-1}) > \frac{1}{\overline{\mathcal{R}}_n} > s(t_n)$.
The lockdowns will be less and less strict, since $\frac{1}{\overline{\mathcal{R}}_1} > \frac{1}{\overline{\mathcal{R}}_2} > ... > \frac{1}{\overline{\mathcal{R}}_n}$ implies $\overline{\mathcal{R}}_1 < \overline{\mathcal{R}}_2 < ... < \overline{\mathcal{R}}_n$.
The n^{th} lockdown is the last one if $s(t_n) < \frac{1}{\mathcal{R}_0} < s(t_{n-1})$. $\square$

A.10 Proof of Proposition 10

(i) Let us denote by $(\mathcal{C}_\varepsilon)$ the curve in the plane (s,i) representing the end of lockdowns $(s_{\mathcal{R}}(T), i_{\mathcal{R}}(T))$ which lead after release of lockdown to $s_\infty = s_\infty(\mathcal{R},T) = \frac{1}{\mathcal{R}_0} - \varepsilon$.

According to Eq. (21) with $t \to +\infty$, and $i = i(T)$, $s = s(T)$, since $\lim_{t\to\infty} i(t) = 0$ and $\lim_{t\to\infty} s(t) = s_\infty$, the equation of $(\mathcal{C}_\varepsilon)$ in the plane (s,i) is $0 = i + s - s_\infty + \frac{1}{\mathcal{R}_0}\ln(s_\infty) - \frac{1}{\mathcal{R}_0}\ln(s)$,

i.e.

$i = s_\infty - \frac{1}{\mathcal{R}_0}\ln(s_\infty) - s + \frac{1}{\mathcal{R}_0}\ln(s)$

Under the assumption $s_\infty = \frac{1}{\mathcal{R}_0} - \varepsilon$, with $\varepsilon > 0$, $(\mathcal{C}_\varepsilon)$ meets the x-axis at two points: $s = s_\infty$ and $s = \widetilde{s}_\infty$, with $s_\infty < \frac{1}{\mathcal{R}_0} < \widetilde{s}_\infty$,

thus $-s_\infty + \frac{1}{\mathcal{R}_0}\ln(s_\infty) = -\widetilde{s}_\infty + \frac{1}{\mathcal{R}_0}\ln(\widetilde{s}_\infty)$, i.e. $h_{\mathcal{R}_0}(s_\infty) = h_{\mathcal{R}_0}(\widetilde{s}_\infty)$.

Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be defined by $\lim_{T\to\infty} s_{\mathcal{R}_1}(T) = \widetilde{s}_\infty$ and $\lim_{T\to\infty} s_{\mathcal{R}_2}(T) = s_\infty$.

Then for $\mathcal{R} \le \mathcal{R}_1$ or $\mathcal{R} \ge \mathcal{R}_2$, we have $s_\infty(\mathcal{R}) < \frac{1}{\mathcal{R}_0} - \varepsilon$, i.e. $s_\infty(\mathcal{R}) = \frac{1}{\mathcal{R}_0} - \varepsilon$ is impossible.

If $\mathcal{R} \in (\mathcal{R}_1, \mathcal{R}_2)$, there exists a value of T (denoted by $T_\varepsilon(\mathcal{R})$), such that $s_\infty(\mathcal{R}) = \frac{1}{\mathcal{R}_0} - \varepsilon$.

(ii) For a given mortality $M_\infty(\mathcal{R},T) = \delta N s_\infty(\mathcal{R},T) = \delta N\left(\frac{1}{\mathcal{R}_0} - \varepsilon\right)$, the economic cost of the lockdown is $C_\varepsilon(\mathcal{R}) = (T_\varepsilon(\mathcal{R}) - t_0)\,(y(\mathcal{R}_0) - y(\mathcal{R}))$.

$\mathcal{R} \mapsto C_\varepsilon(\mathcal{R})$ is a continuous function on $(\mathcal{R}_1, \mathcal{R}_2)$, with $\lim_{\mathcal{R}\to\mathcal{R}_1} C_\varepsilon(\mathcal{R}) = \lim_{\mathcal{R}\to\mathcal{R}_2} C_\varepsilon(\mathcal{R}) = +\infty$.

$\mathcal{R} \mapsto C_\varepsilon(\mathcal{R})$ is continuous on any closed (compact) interval included in $(\mathcal{R}_1, \mathcal{R}_2)$, thus there exists a minimum in $(\mathcal{R}_1, \mathcal{R}_2)$.

The first order equation is $C'_\varepsilon(\mathcal{R}) = 0$, with $C'_\varepsilon(\mathcal{R}) = T'_\varepsilon(\mathcal{R})\,(y(\mathcal{R}_0) - y(\mathcal{R})) - (T_\varepsilon(\mathcal{R}) - t_0)\,y'(\mathcal{R})$

The minimal time is obtained with $\mathcal{R} = \mathcal{R}^\circ_\varepsilon$ satisfying $T'_\varepsilon(\mathcal{R}) = 0$.

The minimal economic cost is obtained with $\mathcal{R} = \widehat{\mathcal{R}}_\varepsilon$ satisfying $C'_\varepsilon(\mathcal{R}) = 0$.

Since $T'_\varepsilon(\mathcal{R}) = 0$ and $C'_\varepsilon(\mathcal{R}) = 0$ cannot be simultaneously obtained, thus $\mathcal{R}^\circ_\varepsilon \neq \widehat{\mathcal{R}}_\varepsilon$.

If $\mathcal{R}^\circ_\varepsilon > \widehat{\mathcal{R}}_\varepsilon$, then $\mathcal{R}^\circ_\varepsilon$ is not only the fastest, but also the less costly. It is impossible since the less costly is $\widehat{\mathcal{R}}_\varepsilon$. Thus $\mathcal{R}^\circ_\varepsilon < \widehat{\mathcal{R}}_\varepsilon$.

We have $\widehat{T}_\varepsilon > T^\circ_\varepsilon$ by definition. $\square$

A.11 Proof of Proposition 11

$L_{\infty}(\mathcal{R}, T) = (T - t_0)(y(\mathcal{R}_0) - y(\mathcal{R})) + \lambda M_{\infty}(\mathcal{R})$ is a continuous function on $(\mathcal{R}, T) \in [0; \mathcal{R}_0] \times [t_0; +\infty)$.

Let $a = \inf_{\substack{0 \leq \mathcal{R} \leq \mathcal{R}_0 \\ T \geq t_0}} L_{\infty}(\mathcal{R}, T)$. Is this infimum a minimum?

We can write $a = \inf_{0 \leq \varepsilon \leq \frac{1}{\mathcal{R}_0}} \left[\inf_{\left\{(\mathcal{R},T) \text{ with } s_{\infty} = \frac{1}{\mathcal{R}_0} - \varepsilon\right\}} L_{\infty}(\mathcal{R}, T) \right] = \inf_{0 \leq \varepsilon \leq \frac{1}{\mathcal{R}_0}} a(\varepsilon)$,
where $a(\varepsilon) = \inf_{(\mathcal{R},T) \text{ with } s_{\infty} = \frac{1}{\mathcal{R}_0} - \varepsilon} L_{\infty}(\mathcal{R}, T)$ has been studied in Proposition 10.

$a(0) = L_{\infty}\left(\widetilde{\mathcal{R}}_{t_0}, +\infty\right) = +\infty$, thus $a = \inf_{0 < \varepsilon \leq \frac{1}{\mathcal{R}_0}} a(\varepsilon)$ and for ε very close to 0, $a(\varepsilon)$ is arbitrarily high, thus if $\varepsilon_0 > 0$ is sufficiently small, we have $a = \inf_{\varepsilon_0 \leq \varepsilon \leq \frac{1}{\mathcal{R}_0}} a(\varepsilon)$.

The function $\varepsilon \mapsto a(\varepsilon)$ is continuous on the compact interval $[\varepsilon_0, \frac{1}{\mathcal{R}_0}]$, thus the infimum is a minimum according to the extreme value theorem, i.e. there exists a couple $(\mathcal{R}^*, T^*) \in [0; \mathcal{R}_0] \times [t_0; +\infty)$ such that $a = L_{\infty}(\mathcal{R}^*, T^*)$. $\square$

A.12 Proof of Proposition 12

According to Equations (25) and (27), we have:

$$\frac{di}{ds} = \frac{\beta_0 si - \gamma i}{-\beta_0 si - \rho s} = \frac{\beta_0 i(s - \frac{1}{\mathcal{R}_0})}{-\beta_0 s(i + \frac{\rho}{\beta_0})} = \frac{-(1 - \frac{1}{s\mathcal{R}_0})}{(1 + \frac{\rho}{\beta_0 i})} \tag{31}$$

i.e.,

$(1 + \frac{\rho}{\beta_0 i})di = -(1 - \frac{1}{s\mathcal{R}_0})ds$.

Integrating between T and t we obtain for $t \geq T$:

$i(t) - i(T) + \frac{\rho}{\beta_0}\ln(i(t)) - \frac{\rho}{\beta_0}\ln(i(T)) = -s(t) + s(T) + \frac{1}{\mathcal{R}_0}\ln(s(t)) - \frac{1}{\mathcal{R}_0}\ln(s(T))$.

Setting $H_{\rho}(s, i) = i + s + \frac{\rho}{\beta_0}\ln(i) - \frac{1}{\mathcal{R}_0}\ln(s)$,

we have $H_{\rho}(s(t), i(t)) = H_{\rho}(s(T), i(T))$ for $t \geq T$.

There is a rebound after T if $i'(T) > 0$. According to (31), we have:

$i'(t) = \frac{di}{dt} = \frac{di}{ds}\frac{ds}{dt} = -\frac{\left(1 - \frac{1}{\mathcal{R}_0 s(t)}\right)}{\left(1 + \frac{\rho}{\beta_0 i(t)}\right)} s'(t)$,

$i'(T) = -s'(T)\frac{\left(1 - \frac{1}{\mathcal{R}_0 s(T)}\right)}{\left(1 + \frac{\rho}{\beta_0 i(T)}\right)}$ which is of the sign of $\left(1 - \frac{1}{\mathcal{R}_0 s(T)}\right)$ since $s'(T) < 0$.

Thus $i'(T) > 0$ if $s(T) > \frac{1}{\mathcal{R}_0}$, i.e., there is a rebound after T if $s(T) > \frac{1}{\mathcal{R}_0}$.

The maximum of this rebound is attained at $s = \frac{1}{\mathcal{R}_0}$, with here $i_{\max}$ satisfying:

$H_\rho(\frac{1}{\mathcal{R}_0}, i_{\max}) = H_\rho(s(T), i(T))$.

We set

$A(\rho) = H_\rho(\frac{1}{\mathcal{R}_0}, i_{\max}) - H_\rho(s(T), i(T))$

$A(\rho) = \left[i_{\max} + \frac{1}{\mathcal{R}_0} + \frac{\rho}{\beta_0}\ln(i_{\max}) - \frac{1}{\mathcal{R}_0}\ln\left(\frac{1}{\mathcal{R}_0}\right)\right] - \left[i(T) + s(T) + \frac{\rho}{\beta_0}\ln(i(T)) - \frac{1}{\mathcal{R}_0}\ln(s(T))\right]$.

We have

$\frac{\partial i_{\max}}{\partial \rho} = -\frac{\frac{\partial A}{\partial \rho}}{\frac{\partial A}{\partial i_{\max}}} = -\frac{\left(\frac{1}{\beta_0}\ln(i_{\max}) - \frac{1}{\beta_0}\ln(i(T))\right)}{1+\frac{\rho}{\beta_0 i_{\max}}} < 0.$

Thus $i_{\max}$ is a decreasing function of ρ, which means that the higher is ρ, the lower the rebound.

Now, we study $(s_\infty, v_\infty, i_\infty, r_\infty) = \lim_{t\to+\infty}(s(t), v(t), i(t), r(t))$.

$H_\rho(s(t), i(t)) = H_\rho(s(T), i(T))$ for $t \geq T$, thus $\lim_{t\to+\infty} H_\rho(s(t), i(t)) = H_\rho(s(T), i(T))$, i.e,

$\lim_{t\to+\infty}\left[i(t) + s(t) + \frac{\rho}{\beta_0}\ln(i(t)) - \frac{1}{\mathcal{R}_0}\ln(s(t))\right] = H_\rho(s(T), i(T))$.

But we have of course $i_\infty = \lim_{t\to+\infty} i(t) = 0$, thus we must have $s_\infty = \lim_{t\to+\infty} s(t) = 0$

To find v_∞, we note first that

$$\frac{ds}{dv} = \frac{-\beta_0 is - \rho s}{\rho s} = -1 - \frac{\beta_0 i}{\rho}$$

$$dv = \frac{-ds}{1 + \frac{\beta_0}{\rho} i}.$$

We are here on $H_\rho(s, i) = H_\rho(s(T), i(T))$. Applying the implicit function theorem, there exists a function I_ρ such that: $H_\rho(s, i) = H_\rho(s(T), i(T)) \Leftrightarrow i = I_\rho(s)$. This leads to:

$$dv = \frac{-ds}{1 + \frac{\beta_0}{\rho} I_\rho(s)}.$$

Integrating from T to $+\infty$, we obtain:

$$v_\infty - v(T) = \int_{s(T)}^{s_\infty} \frac{-ds}{1 + \frac{\beta_0}{\rho} I_\rho(s)} = \int_{s_\infty}^{s(T)} \frac{ds}{1 + \frac{\beta_0}{\rho} I_\rho(s)}$$

with $v(T) = 0$ and $s_\infty = 0$ here, thus

$$v_\infty = \int_0^{s(T)} \frac{ds}{1 + \frac{\beta_0}{\rho} I_\rho(s)}.$$

The final proportion of vaccinated people v_∞ being an increasing function of the rate of vaccination ρ, then the end-of-time number of deaths $M_\infty = \delta N r_\infty = \delta N(1 - v_\infty)$ is a decreasing function of ρ. $\square$

References

Acemoglu, Daron, Chernozhukov, Victor, Werning, Iván, & Whinston, Michael D. 2021. Optimal targeted lockdowns in a multigroup SIR model. *American Economic Review: Insights*, **3**(4), 487–502.

Alvarez, Fernando E, Argente, David, & Lippi, Francesco. 2020. *A simple planning problem for covid-19 lockdown.* Tech. rept. National Bureau of Economic Research, w26981.

Atolia, Manoj, Papageorgiou, Chris, & Turnovsky, Stephen J. 2021. Re-opening after the lockdown: Long-run aggregate and distributional consequences of COVID-19. *Journal of Mathematical Economics*, **93**, 102481.

Avery, Christopher, Bossert, William, Clark, Adam, Ellison, Glenn, & Ellison, Sara Fisher. 2020. An economist's guide to epidemiology models of infectious disease. *Journal of Economic Perspectives*, **34**(4), 79–104.

Bailey, Norman TJ. 1975. The mathematical theory of infectious diseases and its applications. *Pages 413–413 of: The mathematical theory of infectious diseases and its applications.*

Bohner, Martin, Streipert, Sabrina, & Torres, Delfim FM. 2019. Exact solution to a dynamic SIR model. *Nonlinear Analysis: Hybrid Systems*, **32**, 228–238.

Bosi, Stefano, Camacho, Carmen, & Desmarchelier, David. 2021. Optimal lockdown in altruistic economies. *Journal of Mathematical Economics*, **93**, 102488.

Britton, Tom. 2010. Stochastic epidemic models: a survey. *Mathematical biosciences*, **225**(1), 24–35.

Brodeur, Abel, Gray, David M, Islam, Anik, & Bhuiyan, Suraiya. 2020. *A Literature Review of the Economics of COVID-19*. Tech. rept. IZA Discussion Paper.

Camera, Gabriele, & Gioffré, Alessandro. 2021. The economic impact of lockdowns: A theoretical assessment. *Journal of Mathematical Economics*, **97**, 102552.

Caulkins, Jonathan P, Grass, Dieter, Feichtinger, Gustav, Hartl, Richard F, Kort, Peter M, Prskawetz, Alexia, Seidl, Andrea, & Wrzaczek, Stefan. 2021. The optimal lockdown intensity for COVID-19. *Journal of Mathematical Economics*, **93**, 102489.

Eichenbaum, Martin S, Rebelo, Sergio, & Trabandt, Mathias. 2020a. *The macroeconomics of epidemics*. Tech. rept. National Bureau of Economic Research.

Eichenbaum, Martin S, Rebelo, Sergio, & Trabandt, Mathias. 2020b. *The macroeconomics of testing and quarantining*. Tech. rept. National Bureau of Economic Research.

Federico, Salvatore, & Ferrari, Giorgio. 2021. Taming the spread of an epidemic by lockdown policies. *Journal of mathematical economics*, **93**, 102453.

Gonzalez-Eiras, Martín, & Niepelt, Dirk. 2020a. Optimally controlling an epidemic. *Available at SSRN 3753941*.

Gonzalez-Eiras, Martín, & Niepelt, Dirk. 2020b. *On the optimal'lockdown'during an epidemic*. Tech. rept.

Harko, Tiberiu, Lobo, Francisco SN, & Mak, MK3197716. 2014. Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation*, **236**, 184–194.

Kermack, William Ogilvy, & McKendrick, Anderson 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), 700–721.

Kröger, Martin, & Schlickeiser, Reinhard. 2020. Analytical solution of the SIR-model for the temporal evolution of epidemics. Part A: time-independent reproduction factor. *Journal of Physics A: Mathematical and Theoretical*, **53**(50), 505601.

La Torre, Davide, Liuzzi, Danilo, & Marsiglio, Simone. 2021. Epidemics and macroeconomic outcomes: Social distancing intensity and duration. *Journal of Mathematical Economics*, **93**, 102473.

Loertscher, Simon, & Muir, Ellen V. 2021. Road to recovery: managing an epidemic. *Journal of Mathematical Economics*, 102482.

Miclo, Laurent, Spiro, Daniel, & Weibull, Jörgen. 2022. Optimal epidemic suppression under an ICU constraint: An analytical solution. *Journal of Mathematical Economics*, 102669.

Murray, Eleanor J. 2020. Epidemiology's time of need: COVID-19 calls for epidemic-related economics. *Journal of Economic Perspectives*, **34**(4), 105–20.

Murray, James D. 2007. *Mathematical biology: I. An introduction.* Vol. 17. Springer Science & Business Media.

Piguillem, Facundo, & Shi, Liyan. 2020. *Optimal COVID-19 quarantine and testing policies.* Tech. rept. CEPR Discussion Paper No. DP14613.

Rowthorn, Robert, & Maciejowski, Jan. 2020. A cost–benefit analysis of the COVID-19 disease. *Oxford Review of Economic Policy*, **36**(Supplement_1), S38–S55.

Schlickeiser, Reinhard, & Kröger, Martin. 2021. Analytical solution of the SIR-model for the temporal evolution of epidemics: part B. Semi-time case. *Journal of Physics A: Mathematical and Theoretical*, **54**(17), 175601.