SIAM J. CONTROL OPTIM.
Vol. 60, No. 2, pp. S221–S245

A DISTRIBUTED OPTIMAL CONTROL MODEL APPLIED TO COVID-19 PANDEMIC*

RAIMUND M. KOVACEVIC†, NIKOLAOS I. STILIANAKIS‡, AND VLADIMIR M. VELIOV§

Abstract. In this paper, a distributed optimal control epidemiological model is presented. The model describes the dynamics of an epidemic with social distancing as a control policy. The model belongs to the class of continuous-time models, usually involving ordinary/partial differential equations, but has a novel feature. The core model—a single integral equation—does not explicitly use transition rates between compartments. Instead, it is based on statistical information on the disease status of infected individuals, depending on the time since infection. The approach is especially relevant for the coronavirus disease 2019 (COVID-19) in which infected individuals are infectious before onset of symptoms during a relatively long incubation period. Based on the analysis of the proposed optimal control problem, including necessary optimality conditions, this paper outlines some efficient numerical approaches. Numerical solutions show some interesting features of the optimal policy for social distancing, depending on the weights attributed to the number of isolated individuals with symptoms and to economic losses due to the enforcement of the control policy. The general nature of the model allows for inclusion of additional epidemic features with minor adaptations in the basic equations. Therefore, the modeling approach may contribute to the analysis of combined intervention strategies and to the guidance of public health decisions.

Key words. epidemiology, optimal control, integral equation

AMS subject classifications. 92D30, 49K21

DOI. 10.1137/20M1373840

1. Introduction.

The COVID-19 pandemic, caused by the severe acute respiratory coronavirus 2 (SARS-CoV-2), shows characteristics that are challenging for public health care systems. High transmissibility of the causative virus, transmission of the virus by asymptomatic and presymptomatic infected individuals (subclinical infection), and lack of effective treatments and vaccines impede pandemic control. Therefore, nonpharmaceutical interventions such as population- and individual-based social distancing, testing, and contact tracing have so far been the principal public health measures.

In the case of COVID-19, infected individuals are initially asymptomatic for a period of approximately five to six days on average. Some of them may not develop symptoms and recover without complications. A fraction of the asymptomatic infected individuals progress to symptoms that, depending on the severity of the disease, may lead to death. In most cases symptomatic individuals are isolated.

On the other hand, infectiousness starts shortly after infection during the asymp-

*Received by the editors October 19, 2020; accepted for publication (in revised form) November 9, 2021; published electronically March 24, 2022.
https://doi.org/10.1137/20M1373840
Funding: This research was supported by the Austrian Science Foundation (FWF) under grant P31400-N32.

†Department for Economy and Health, Danube University Krems, 3500 Krems an der Donau, Austria, and Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, A-1040 Vienna, Austria (raimund.kovacevic@donau-uni.ac.at).

‡European Commission, Joint Research Centre (JRC), Ispra, 21027, Italy, and Department of Biometry and Epidemiology, University of Erlangen-Nuremberg, 91054 Erlangen, Germany (nikolaos.stilianakis@ec.europa.eu). Disclaimer: The views expressed are purely those of the writer (NIS) and may not in any circumstance be regarded as stating an official position of the European Commission.

§Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, A-1040 Vienna, Austria (vladimir.veliov@tuwien.ac.at).

tomatic period. The degree of infectiousness depends on the time since infection and on the stadium of the disease at which the infected individual is. It is estimated that infectiousness is at its highest shortly before or upon the onset of symptoms; see, e.g., He et al. (2020). Thus, epidemiological population groups such as asymptomatics, presymptomatics, and symptomatics are infectious, can transmit the virus, and are causes of major concern for the control of the epidemic; see He et al. (2020); Nishiura et al. (2020); Mizumoto et al. (2020); Park et al. (2020); and Zhang et al. (2020).

Mathematical models have been developed to assess the transmission dynamics of the virus, the severity of the disease, as well as the effectiveness of public health measures. Major social distancing interventions have been population-based (e.g., lockdown at local, regional, or national level) and individual-based social distancing; see, e.g., Peak et al. (2020); Hellewell et al. (2020); Tsay et al. (2020); and Kretzschmar et al. (2020). Predominant social distancing strategies at the individual level are isolation of asymptomatically or symptomatically infected individuals, contact tracing, and quarantine.

The vast majority of mathematical models are based on deterministic compartmental modeling, where the epidemiological subpopulations are classified into a susceptible category, several infected population groups as well as groups with recovered individuals or fatalities. Dynamic transitions from one class to the next are based on transition rates, usually constant and sometimes also variable over time. Interventions are built into the models in a similar way. Stochastic approaches have similar rationale allowing for uncertainty, and statistical approaches are usually used for the estimation of epidemiological parameters.

A major issue with ordinary or partial differential equation models with compartments is that transition rates are not directly observable and have to be identified from available observable data. A main problem for the latter is that transition rates are intrinsically variable—they may change in the course of the epidemic even in a stationary environment (see section 2).

The main advantage of the model proposed in this paper is that it does not involve transition rates between compartments. The only dynamic variable, $y(t)$, is the number of new infections as a function of the time, t. The dynamics of this variable is described by a single evolutionary integral equation. It requires observable statistical information on the medical status of infected individuals during the course of infection, that is, depending on the time since infection, θ. Examples for needed information are as follows: fraction of infected individuals showing symptoms θ days after infection; fraction of infected individuals who die θ days after infection; fraction of infected individuals who recover θ days after infection, etc. Knowledge of the function $y(\cdot)$ on a time interval $[0, t]$ and the needed statistical information allows one to evaluate the size of various subpopulations at time t (infected, asymptomatic, symptomatic, recovered, died, etc.) as well as to calculate several meaningful *effective reproduction numbers*.

Since the infection age plays a decisive role in the type of required data and for the model, we stress that models in which the infectiousness varies with the infection age are well known in the literature which goes back to the 1920s of the 20th century (Kermack and McKendrick (1927)). More about the history and the development of models with infection-age-dependent infectiousness can be found in Thieme and Castillo-Chavez (1993). Such models are especially relevant to epidemic diseases with a long incubation period and high variability of the infectiousness with the infection age, in particular to COVID-19, as argued above.

The basic model of the evolution of the epidemic is presented in section 2. As

mentioned above, it consists of a single integral equation, which is not standard; therefore existence of a (nonnegative) solution and some additional properties are presented. This section also includes formulas for the sizes of the subpopulations of interest, and for the effective reproduction numbers.

Optimal control theory has become a powerful tool designing of policies for prevention and medical treatment of infectious diseases. The work by Feichtinger et al. (2004) investigates an optimal control problem for a model involving infection-age-dependent infectiousness. In the present paper we only consider social distancing as control policy; therefore the main issue is to find a reasonable compromise between the positive effect of the social distancing for the attenuation of the epidemic, and the resulting economic losses due to the induced "lockdown."

In section 3, we present an optimal control model that combines the two optimization criteria mentioned above. The effect of social distancing on the course of epidemic is measured by the total number of isolated (or dead) individuals, while the economic component in the objective functional involves a simple estimation of the economic losses. Since the considered optimal control problem is not standard we derive a necessary optimality condition, together with some additional properties (subsection 3.1) which provide a basis for numerical approaches (subsection 3.2).

In order to demonstrate the usage of the proposed approach, section 4 analyzes a numerical case study. The parametrization of the implemented instance is based on distributions for the incubation period, the serial interval, and the time between onset of symptoms and death. Subsection 4.1 discusses related literature and the estimates used in the case study, and subsection 4.2 describes the construction of the implemented parameter functions, which depend on the time since infection. Finally, in subsection 4.3 a number of numerical results for the case study are presented, including the evolution of the epidemic subpopulations with or without policy measures, the optimal control policies for various optimization scenarios, the trade-off between the economic and the humanitarian objectives, and the effective reproduction numbers. An interesting observation is, for example, that for a larger weight attributed to the humanitarian objective, the contact restrictions begin earlier and are removed later, but the magnitude of the restrictions is milder. If more weight is attributed to the economy, severe restrictions are undertaken, but later and for a relatively short time.

In the concluding section, section 5, we indicate some possible extensions.

2. The basic model. An individual is infected after exposure to the virus, which allows the virus to enter, remain, and multiply in the host. One of the main features of COVID-19 is that infected individuals may be infectious during a long period of time, even until recovery, without having symptoms (asymptomatics). The infectiousness of infected individuals strongly depends on the time since infection. For this reason, the time since infection is essential in the model presented below. This time, denoted further by θ, will be shortly called *infection age* (not to be confused with the biological age of individuals). Thus an individual infected at time t will have infection age $\theta = t_1 - t$ at time $t_1 \geq t$.

Along with the infectiousness, the contact rates of individuals play a crucial role in epidemiological models. These depend on the stage of the individual in the course of the infection: asymptomatic/symptomatic, nonisolated/isolated, nonhospitalized/hospitalized, etc. The same applies to the other parameters (related to mortality, recovery, etc.) used in building an epidemiological model. In this paper, we focus on the following subpopulations (with corresponding notation), characterizing the status of an individual in the context of an epidemic:

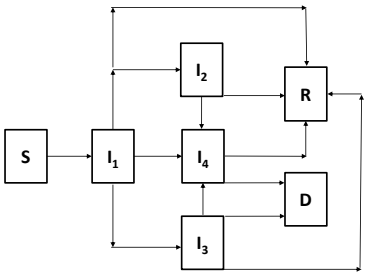

FIG. 2.1. *Possible paths for the development of the disease of an infected individual.*

S – susceptible;
I_1 – asymptomatic, nonisolated;
I_2 – asymptomatic, isolated;
I_3 – symptomatic, nonisolated;
I_4 – symptomatic, isolated;
R – recovered;
D – dead.
The possible stages that an infected individual may undergo are shown in Figure 2.1. For example, $S \longrightarrow I_1 \longrightarrow R$ and $S \longrightarrow I_1 \longrightarrow I_3 \longrightarrow I_4 \longrightarrow D$ are possible paths.

Typical ODE epidemiological models (such as SIR, SIRS, SEIR, etc.) and some PDE models describe the transitions between epidemiological population groups by *transition rates*. Such models have (among others) the following two drawbacks: (i) the transition rates cannot be directly extracted from observable data; (ii) even in a stationary environment without, e.g., seasonal fluctuations and pharmaceutical or nonpharmaceutical measures, the transition rates may be intrinsically time-dependent. For example, the transition rate from the group of infected to the group of recovered or dead individuals during the expansion phase of the epidemic is smaller than that in a phase of decreasing numbers of new infections. For this reason, the use of models involving transition rates for real predictions gives reasonable results only for relatively short time horizons if the dynamics of the epidemic are fast. Hence, the transition rates have to be permanently updated using current measurements. Various techniques for that have been published, especially after the emergence of COVID-19, e.g., the recent papers of Ma (2020); Kounchev et al. (2020); Margenov et al. (2020). The main advantage of the model proposed in this paper, compared with ODE models (see, e.g., Giordano et al. (2020) where almost the same compartments are considered), is that our model is mainly based on data about the course of infection. These data may be estimated from course of infection curves for, e.g., the incubation period, serial intervals, or the time span between onset of symptoms and death. Such information is the subject of epidemiological studies; see the discussion in subsection 4.1 below. Other relevant information about the course of infection can be observed by public health authorities, e.g., by monitoring flows from and to the isolated classes or using seroepidemiological studies of parts of the population. In contrast to the transition rates, this information can be considered as intrinsically stationary, although it may depend on external factors—seasonality effects, pharmaceutical interventions, etc. In particular, instead of transition rates, we use information about the probability that an infected individual belongs to one or another population group at a given infection age. Clearly, the availability and accuracy of these data increase during the

evolution of the epidemic. Section 4 discusses in more detail the concrete information used for a numerical case study.

Before formulating the model we make a few preliminary assumptions. First, the environment is stationary, except possibly global restrictions on the contact rates (social distancing). Second, the recovered individuals remain immune after recovery; see, e.g., Huang et al. (2020), Peng et al. (2020), and Kojima and Klausner (2022). Some more assumptions, concerning emergence of new cases, will be made below.

2.1. Formulation of the model. The model has a time-range $[0, T]$ (the time measured in days), while the disease may have begun earlier than time 0. The following data related to the progress of the disease along the infection age θ are required. Each of them represents the fraction of all individuals infected at the same time, having a given status $(I_1, \dots, I_4, R, D)$ at infection age θ:
$\alpha_1(\theta)$ – fraction of asymptomatic nonisolated individuals (I_1);
$\alpha_2(\theta)$ – fraction of asymptomatic isolated individuals (I_2);
$\alpha_3(\theta)$ – fraction of symptomatic nonisolated individuals (I_3);
$\alpha_4(\theta)$ – fraction of symptomatic isolated individuals (I_4);
$\rho(\theta)$ – fraction of recovered individuals (R);
$\mu(\theta)$ – fraction of dead individuals (D).
These fractions have to be known on the interval $[0, \Theta]$, where Θ is such that the infected individuals of infection age Θ can be assumed to be either recovered or dead. Consistent with the meaning of Θ, we extend $\alpha_k(\theta) = 0$, $k = 1, \dots, 4$, and $\rho(\theta) = \rho(\Theta)$, $\mu(\theta) = \mu(\Theta)$ for $\theta > \Theta$. All of these fractions can be considered as time-dependent without substantial changes in the model, but in this paper they are assumed to be stationary.

Clearly, it must hold that for all $\theta \geq 0$

$$\sum_{k=1}^{4} \alpha_k(\theta) + \rho(\theta) + \mu(\theta) = 1.$$

Figure 2.2 represents the above functions for the case study presented in section 4.

The above data allow one to express the sizes of each group with a given status (S, I_1, …, D) at any time $t \geq 0$ by means of a single function of time, $y(\cdot)$, which gives the amount of new infections at any given time $t \in [-\Theta, T]$. Somewhat overloading the notation, $I_k(t)$ will be the size of the group I_k at time $t \in [0, T]$ $(k = 1, \dots, 4)$ and similarly for $S(t)$, $R(t)$, and $D(t)$. Then, obviously,

$$I_k(t) = \int_0^{\Theta} \alpha_k(\theta) y(t - \theta)\, \mathrm{d}\theta, \quad k = 1, \dots, 4, \tag{2.1}$$

$$S(t) = S(-\Theta) - \int_0^{\Theta + t} y(t - \theta)\, \mathrm{d}\theta, \tag{2.2}$$

$$R(t) = R_0 + \int_0^{\Theta + t} \rho(\theta) y(t - \theta)\, \mathrm{d}\theta, \tag{2.3}$$

$$D(t) = D_0 + \int_0^{\Theta + t} \mu(\theta) y(t - \theta)\, \mathrm{d}\theta, \tag{2.4}$$

where $S(-\Theta)$ is the size of the susceptible population at time $-\Theta$, R_0 is the size of the group of infected individuals before $t = -\Theta$ who have recovered until time $t = 0$, and D_0 is similar but for the dead individuals.

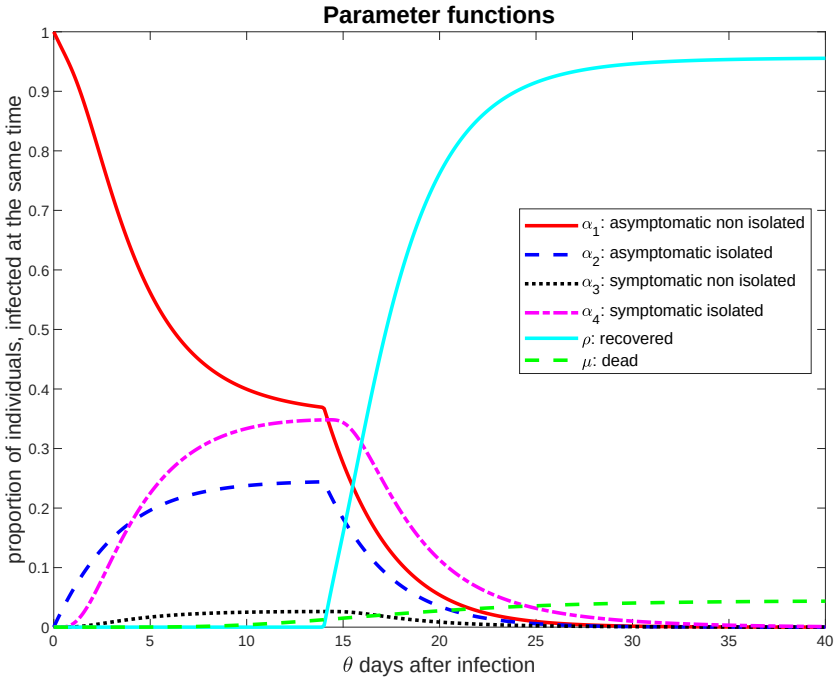

FIG. 2.2. *Parameter functions α_k, ρ, and μ.*

Thus, the only unknown variable in the model will be $y(t)$—the quantity of individuals who get infected (for the first time) at time t. All other variables of interest are expressed in terms of $y(\cdot)$ by (2.1)–(2.4). According to the definition of Θ, all infected individuals either recover or die at most Θ days after infection.

More data is needed for modeling the dynamics of $y(t)$, some of which will be time-dependent in order to later incorporate the distancing policy, namely

$c(t)$ – contact rate of susceptible and recovered individuals,
$c_k(t)$ – contact rate of individuals of infection status I_k, $k = 1, \dots, 4$,
$i_k(\theta)$ – infectiousness of individuals of infection age θ and infection status I_k, $k = 1, \dots, 4$.

Here it is assumed that the infectiousness is independent of time, which may not be the case if medication treatment is applied. These parameters are more difficult to be evaluated. One may consider $c(t)$—the normal contact rate of susceptibles—as a reference value, and assume that

(2.5)
$$c_1(t) = c(t), \quad c_2(t) = c_4(t) = \varepsilon c(t) \ \text{ with a small } \varepsilon, \quad c_3(t) = \delta c(t) \ \text{ with } \delta \in [\varepsilon, 1].$$

Moreover, it is reasonable to assume that

$$i_1(\theta) = i_2(\theta) =: i_A(\theta), \qquad i_3(\theta) = i_4(\theta) =: i_S(\theta), \tag{2.6}$$

since the infectiousness does not depend on whether a person is isolated or not, but may depend (in addition to the infection age) on presence of symptoms (a symptomatic individual may shed more virus loaded droplets than an asymptomatic one).

In the derivation of the equation for y below, natural deaths and births are ignored, which is plausible if the duration of the epidemic is not too long or the natural

demographic change is slow. The foundation of this equation follows.

First, the restriction of $y(\cdot)$ on $[-\Theta, 0)$, denoted further by $y_0(\cdot)$, is considered as known from records for the new infections before time $t = 0$.

It is assumed that the amount of newly infected individuals at time t is proportional to the number of contacting susceptibles, $S(t)$, regarding their contact rate, $c(t)$, and to the infectiousness of the environment in which the contacts take place:

$$y(t) = \mathcal{I}(t)c(t)S(t).$$

In the representation of $\mathcal{I}(t)$ it is assumed, roughly speaking, that an infectious individual with a certain level of infectiousness $i(\theta)$, and having a certain number of risky contacts per unit of time, is equally infectious as one who is twice as infectious but has half of the risky contacts. In addition, it is assumed that the individuals participating in contacts are homogeneously mixed. Then

$$y(t) = \frac{V(t)}{W(t)}c(t)S(t), \tag{2.7}$$

where $V(t)$ is the participation of infected individuals, weighed by their infectiousness and their contact rates, and $W(t)$ is the total population participating in risky contacts, weighed by their contact rates. Thus,

$$V(t) = \int_0^{\Theta} y(t-\theta) \sum_{k=1}^{4} i_k(\theta)c_k(t)\alpha_k(\theta)\,\mathrm{d}\theta.$$

Similarly,

$$W(t) = \int_0^{\Theta} y(t-\theta) \sum_{k=1}^{4} c_k(t)\alpha_k(\theta)\,\mathrm{d}\theta + c(t)(S(t) + R(t)).$$

Using (2.5) and (2.6), and having in mind that all $\alpha_k(\theta) = 0$ for $\theta > \Theta$, we represent

$$V(t) = c(t) \int_0^{\Theta+t} \big[i_A(\alpha_1 + \varepsilon\alpha_2) + i_S(\delta\alpha_3 + \varepsilon\alpha_4)\big](\theta)\, y(t-\theta)\,\mathrm{d}\theta,$$

$$W(t) = c(t)\bigg[S(-\Theta) + R_0 + \int_0^{\Theta+t} \big[\alpha_1 + \varepsilon\alpha_2 + \delta\alpha_3 + \varepsilon\alpha_4 + \rho - 1\big](\theta)\, y(t-\theta)\,\mathrm{d}\theta\bigg].$$

Substituting these expressions in (2.7) we obtain the following equation describing the dynamics of the new cases, $y(\cdot)$:

$$y(t) = c(t)\frac{\int_0^{\Theta+t} \tilde{q}(\theta)y(t-\theta)\,\mathrm{d}\theta \left(S(-\Theta) - \int_0^{\Theta+t} y(t-\theta)\,\mathrm{d}\theta\right)}{S(-\Theta) + R_0 + \int_0^{\Theta+t}[q(\theta) + \rho(\theta) - 1]y(t-\theta)\,\mathrm{d}\theta}, \tag{2.8}$$

where

$$\tilde{q}(\theta) = [i_A(\alpha_1 + \varepsilon\alpha_2) + i_S(\delta\alpha_3 + \varepsilon\alpha_4)](\theta), \quad q(\theta) = [\alpha_1 + \varepsilon\alpha_2 + \delta\alpha_3 + \varepsilon\alpha_4](\theta). \tag{2.9}$$

Equation (2.8) is an evolutionary integral equation for $y(\cdot)$, supplemented with the initial condition

$$y(t) = y_0(t) \quad \text{for} \quad t \in [-\Theta, 0). \tag{2.10}$$

This is the basic equation of the model. Knowing its solution, one may calculate the quantities of interest in (2.1)–(2.4).

Having in mind (2.2)–(2.3) with $t = 0$, we can equivalently reformulate equation (2.8) in the following way, explicitly including y_0 in the equation:

$$y(t) = c(t)\frac{\left(d_1(t) + \int_0^t \tilde{q}(\theta)y(t-\theta)\,\mathrm{d}\theta\right)\left(d_2 - \int_0^t y(t-\theta)\,\mathrm{d}\theta\right)}{d_3(t) + \int_0^t [q(\theta)+\rho(\theta)-1]y(t-\theta)\,\mathrm{d}\theta}, \tag{2.11}$$

where

$$d_1(t) := \int_{-\Theta}^0 \tilde{q}(t-\tau)y_0(\tau)\,\mathrm{d}\theta, \quad d_2 := S(0), \tag{2.12}$$

$$d_3(t) := S(0) + R_0 + \int_{-\Theta}^0 [q(t-\tau)+\rho(t-\tau)]y_0(\tau)\,\mathrm{d}\theta.$$

We mention that the model can be further extended, say by including hospitalized individuals, hospitalized with intensive care, etc., provided that data similar to α_k are available for these groups. Such extensions, however, do not change the structural form of the transmission dynamics (2.8) or (2.11). More about extensions is given in section 5.

2.2. Preliminary analysis of the basic integral equation. Concerning the data involved in (2.8) we make the following assumptions:

(i) the functions α_k $(k = 1, \dots, 4)$, ρ, μ, i_A, i_S, c, y_0 are measurable, bounded, and nonnegative, $c(t) \le \bar{c}$, $\varepsilon, \delta \in (0, 1]$;

(ii) $S(-\Theta) - \int_{-\Theta}^0 y_0(\theta) \ge 0$;

(iii) $\eta := R_0 + \int_0^\Theta \rho(\theta)y_0(-\theta)\,\mathrm{d}\theta > 0$.

Here $\bar{c}$ represents the average contact rate of individuals in absence of the epidemic—a positive number that may become lower, $c(t) \le \bar{c}$, due to restrictions on the contacts during the epidemic. Assumption (ii) formally means that the susceptible population, $S(0)$, at time $t = 0$ is nonnegative (see (2.2) with $t = 0$). Assumption (iii) requires that at time $t = 0$ there are recovered individuals. This assumption can be relaxed, because it is only needed to ensure that the denominator in (2.8) does not vanish. For that it is enough to assume that the population does not go extinct before $t = T$, but such an assumption is not a priori checkable.

Below $\|\cdot\|_i$ denotes the norm in the usual space $L^i[0, \tau]$ on a scalar interval $[0, \tau]$ (which is sometimes skipped in the notation).

PROPOSITION 2.1. *For every nonnegative function $c \in L^2[0, T]$, equation* (2.11) *has a unique solution $y \in L^2[0, T]$. This solution is nonnegative, and the denominator in* (2.11) *is a.e. strictly larger than η along the solution y. Moreover, $y(t) \le c(t)\max_{\theta\in[0,\Theta]}\{i_A(\theta), i_S(\theta)\}S(0)$.*

The proof is based on the contraction mapping theorem. The specific form of (2.11), where the denominator can vanish for some $y \in L^2$, makes the proof not straightforward. Therefore, it is given in the appendix.

2.3. Reproduction numbers. The notion of an *effective reproduction number* may have various meanings depending on the group of infected individuals to which

it applies. Roughly, this number should give the average number of secondary cases per primary cases from the considered group at time t. The model presented in this paper allows one to define in a meaningful way several types of reproduction numbers (further on we drop the adjective "effective"). Let $I(t)$ be the size of the infected population at time t (the latter merely called "group $I(t)$"). We consider the following:

- $\mathcal{R}(t)$ – *aggregated reproduction number* at time t—the average number of individuals who are directly infected by an individual from the group $I(t)$.
- $\mathcal{R}_{\mathrm{coh}}(t)$ – *reproduction number of newly infected individuals* at time t—the average number of individuals who are directly infected by a member of the cohort of newly infected individuals emerging at time t.

Given that the basic equation (2.11) (or (2.8)–(2.10)) is solved until time $t+\Theta$ (with any $t \in [0, T-\Theta]$), one can represent the above reproduction numbers as follows:

$$\mathcal{R}(t) = \frac{1}{I(t)} \int_t^{t+\Theta} c(\tau) \frac{S(\tau) \int_{\tau-t}^{\Theta} \tilde{q}(\theta) y(\tau-\theta)\,\mathrm{d}\theta}{S(\tau) + R(\tau) + \int_0^{\Theta} q(\theta) y(\tau-\theta)\,\mathrm{d}\theta}\,\mathrm{d}\tau, \tag{2.13}$$

$$\mathcal{R}_{\mathrm{coh}}(t) = \int_t^{t+\Theta} c(\tau) \frac{\tilde{q}(\tau-t) S(\tau)}{S(\tau) + R(\tau) + \int_0^{\Theta} q(\theta) y(\tau-\theta)\,\mathrm{d}\theta}\,\mathrm{d}\tau. \tag{2.14}$$

Here $S(\tau)$, $R(\tau)$, and $I(\tau) := \sum_{k=1}^{4} I_k(\tau)$ have to be calculated from (2.1)–(2.3). To obtain the above expressions one may just follow the derivation of the basic model (2.8), tracing the role of the respective group (either $I(t)$ or $y(t)$) in the evolution of the size of the infected population in $[t, t+\Theta]$, not counting infections caused by secondary cases.

Essentially, $\mathcal{R}(t)$ indicates what is the averaged per capita "production" of new cases by the individuals infected prior to time t, while $\mathcal{R}_{\mathrm{coh}}(t)$ is the averaged per capita "production" of newly infected individuals at time t.

Notice the number $\mathcal{R}(t)$ only includes the per capita secondary cases produced by individuals who are infectious at time t. This undervalues the contribution of these individuals to the epidemic, because they may have already produced secondary cases before time t. If also these secondary infections are counted, then we have a more relevant effective reproduction number of the currently infected individuals:

$$\begin{aligned}\mathcal{R}^{\mathrm{all}}(t) = {} & \frac{1}{I(t)} \int_t^{t+\Theta} c(\tau) \frac{S(\tau) \int_{\tau-t}^{\Theta} \tilde{q}(\theta) y(\tau-\theta)\,\mathrm{d}\theta}{S(\tau) + R(\tau) + \int_0^{\Theta} q(\theta) y(\tau-\theta)\,\mathrm{d}\theta}\,\mathrm{d}\tau \\ & + \frac{1}{I(t)} \int_{t-\Theta}^{t} c(\tau) \frac{S(\tau) \int_0^{\tau-t+\Theta} \tilde{q}(\theta) y(\tau-\theta)\,\mathrm{d}\theta}{S(\tau) + R(\tau) + \int_0^{\Theta} q(\theta) y(\tau-\theta)\,\mathrm{d}\theta}\,\mathrm{d}\tau.\end{aligned}$$

In addition, one can distinguish the contribution of members of the cohort of new cases at time t during various phases of their infection period. For example, a newly infected individual emerging at time t will create an average of $\mathcal{R}^1_{\mathrm{coh}}(t)$ secondary cases during her asymptomatic period and before potential isolation, where

$$\mathcal{R}^1_{\mathrm{coh}}(t) = \int_t^{t+\Theta} c(\tau) \frac{i_A(\tau-t)\alpha_1(\tau-t) S(\tau)}{S(\tau) + R(\tau) + \int_0^{\Theta} q(\theta) y(\tau-\theta)\,\mathrm{d}\theta}\,\mathrm{d}\tau.$$

Similarly, one can calculate the contribution of the members of the cohort infected at time t during the other phases of their infection period. This may give valuable information for efficiently focusing potential policies for attenuation of the epidemic.

Of course, it is of particular interest to evaluate the *basic reproduction number* of the epidemic. The effect of a single infected individual put in a completely susceptible population depends on the specific infection age at the time of emergence. If (in the worst case) the infected individual has infection age $\theta = 0$, the number of secondary cases will be

$$\mathcal{R}_0 = \int_0^\Theta c_1(\theta) i(\theta)\, \mathrm{d}\theta.$$

This expression is straightforward, but it could also be obtained from (2.3) for $t = 0$ and passing to the limit with $I(0) \to 0$.

The calculation of the effective reproduction numbers at time t requires knowledge of the evolution of the basic variable $y(\cdot)$ on $[t, t + \Theta]$; thus their usefulness may look questionable. However, the calculation of the reproduction numbers at time $t - \Theta$ may be based (using the same formulas) on the real measurements on $y(\tau)$ in the interval $[t - \Theta, t]$. Moreover, having in mind that secondary cases dominantly appear in a much shorter period than Θ, one may replace Θ in the formulas for $\mathcal{R}(t)$ and $\mathcal{R}_{\mathrm{coh}}(t)$ with a much smaller number (e.g., 7 days instead of $\Theta = 40$ days). The resulting approximations of the reproduction numbers based only on measured data can be helpful for assessment of the current potential of the epidemic.

3. Optimal "social/physical distancing" policy. In this section we consider the contact rate $c(t)$ as a policy variable and introduce an objective functional to be minimized, including in a simple way human/medical concerns, but also regarding the economic losses due to the epidemic and due to restrictions on the contact rates ("physical distancing"). For this, the contact rate c in (2.8) will take the form $c(t) = \bar{c}u(t)$, where $u(t) \in [u_0, 1]$ is a control function to be chosen (we remind the reader that $\bar{c}$ is the "normal" average contact rate without epidemic).

We denote the total human damage of the epidemic in $[0, T]$ (assuming that the chosen time horizon is long enough, so that the epidemic is close to extinction at T) by

$$\int_0^T I_4(t)\, \mathrm{d}t,$$

which is the total number of isolated symptomatic individuals in the period $[0, T]$; see (2.1). Keeping in mind that the number of individuals who need hospitalization or intensive care is a statistically estimated proportion of the isolated symptomatic ones, the same expression, weighted appropriately, can represent the total number of each of these groups, or the total number of deaths.[1]

Modeling the economic damages of an epidemic (with or without "lockdown") is a complicated issue (see, e.g., the review by Bloom et al. (2020) and Acemoglu et al. (2020) for comprehensive analysis of the issue). Here we keep the model simple by making a reasonable shortcut. The economic part of the objective functional is a simple version of the "cost-of-illness" approach (see section 3 in Bloom et al. (2020)), where medical costs and the (temporarily or ultimately) lost labor are counted. The instantaneous gross domestic product (GDP) without epidemic is assumed to be proportional to $K^{1-\sigma} L^{\sigma}$ (the Cobb–Douglas production function), where K is the capital stock at a given time (which will be assumed constant in the horizon $[0, T]$), L is the

[1]Here we ignore possible susceptible individuals in quarantine, and the infected asymptomatic individuals, because quarantine/isolation of such individuals can only be achieved by additional policy measures (testing and contact tracing), which will be the subject of a further upgrade of the model.

available labor, and $\sigma \in (0,1)$. With the restriction $u \in [u_0, 1]$, and due to the fact that isolated symptomatic individuals normally do not work, the economic losses at time t can be assumed to be proportional to

$$K^{1-\sigma}L^{\sigma} - K^{1-\sigma}(u(t))^{\sigma}\Big(L - \frac{L}{N}I_4(t)\Big)^{\sigma} = K^{1-\sigma}L^{\sigma}\Big(1 - (u(t))^{\sigma}\Big(1 - \frac{I_4(t)}{N}\Big)^{\sigma}\Big).$$

Here N and L are the total population and labor before the epidemic; thus L/N is the fraction of labor in the total population. The dead individuals are disregarded in the above expression, because the fraction of individuals in working ages who die due to the COVID-19 epidemic is relatively low. Summarizing, we consider the following objective functional to be minimized:

$$\int_0^T \Big[\pi I_4(t) + p\left(1 - (u(t))^{\sigma}\left(1 - \beta I_4(t)\right)^{\sigma}\right)\Big]\,\mathrm{d}t, \tag{3.1}$$

where $\pi \geq 1$ is a weight (formally redundant), which allows one to take into account the costs per capita for medical treatment $(\pi - 1)$. The parameter β may take values between zero and L/N, and p is a weight which incorporates (in a multiplicative way) the following three components: the pre-epidemic GDP (proportional to $K^{1-\sigma}L^{\sigma}$), the discounted postepidemic economic losses (which may be considered as proportional to the economic losses during the epidemic), and the weight needed to put on a common scale the human losses and the economic losses. The latter is a matter of political decision. Using the parameter p one can analyze the trade-off between humanitarian and economic objectives. Summarizing and disregarding the constant summand p in (3.1), we obtain the following optimal control problem:

$$\min \int_0^T \Big(\pi I_4(t) - p\,(u(t))^{\sigma}\,[1 - \beta I_4(t)]^{\sigma}\Big)\,\mathrm{d}t \tag{3.2}$$

subject to the equations

$$y(t) = u(t)\,\bar{c}\,\frac{\Big(d_1(t) + \int_0^t \tilde{q}(\theta)y(t-\theta)\,\mathrm{d}\theta\Big)\Big(d_2 - \int_0^t y(t-\theta)\,\mathrm{d}\theta\Big)}{d_3(t) + \int_0^t [q(\theta) + \rho(\theta) - 1]y(t-\theta)\,\mathrm{d}\theta}, \tag{3.3}$$

$$I_4(t) = d_4(t) + \int_0^t \alpha_4(\theta)y(t-\theta)\,\mathrm{d}\theta, \tag{3.4}$$

and the control constraint $u(t) \in [u_0, 1]$, where

$$d_4(t) = \int_{-\Theta}^0 \alpha_4(t-s)y_0(s)\,\mathrm{d}s. \tag{3.5}$$

The parameter β can be assumed to belong to $[0,1]$, provided that the population size N before the beginning of the epidemic is normalized to 1 as we further assume.

3.1. Analysis of the optimization problem. Problem (3.2)–(3.4) can be written in the following more general form with states $y \in \mathbb{R}$ and $z \in \mathbb{R}^n$ and control values $u \in U \subset \mathbb{R}^m$:

$$\min \int_0^T g(z(t), u(t))\,\mathrm{d}t \tag{3.6}$$

subject to

$$y(t) = f(z(t), u(t)) \quad \text{for} \quad t \in [0, T], \tag{3.7}$$

$$z_k(t) = d_k(t) + \int_0^t \varphi_k(\theta) y(t-\theta)\, \mathrm{d}\theta, \quad k = 1, \dots, n, \tag{3.8}$$

$$u(t) \in U. \tag{3.9}$$

Indeed, with $z = (z_1, z_2, z_3, z_4)$ and $U = [u_0, 1]$, one can set

$$\varphi_1(\theta) = \tilde{q}(\theta), \quad \varphi_2(\theta) = -1, \quad \varphi_3(\theta) = q(\theta) + \rho(\theta) - 1, \quad \varphi_4(\theta) = \alpha_4(\theta), \tag{3.10}$$

$$f(u, z) := u\, \bar{c}\, \frac{z_1\, z_2}{z_3}, \qquad g(z, u) = \pi z_4 - p\, u^\sigma\, (1 - \beta z_4)^\sigma,$$

where d_1, d_2, d_3 are defined in (2.12), d_4 is defined in (3.5), and q and $\tilde{q}$ are introduced in (2.9). According to the last statement of Proposition 2.1, $z_3 > \eta$ for every z_3 in the domain of interest in the particular model (3.2)–(3.4). Since the analysis below is local, the function f is smooth in the domain of interest. Thus, also taking into account Proposition 2.1, the assumptions made in the next paragraph for the more general problem (3.6)–(3.9) can be considered as fulfilled for the particular problem (3.2)–(3.4).

Assumptions for problem (3.6)–(3.9). The functions $f, g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are differentiable, $\varphi, d_k \in L^\infty[0, T]$, and U is convex and closed. Moreover, g, and all first derivatives of f and g, are (globally) Lipschitz continuous in $u \in \mathbb{R}^m$, uniformly with respect to z in any compact set, and locally Lipschitz continuous in z, uniformly with respect to u in any compact set. In addition, there exists a constant M such that for every $u \in L^2[0, T]$ the system (3.7)–(3.8) has a unique solution (y, z) in L^2, and $\|z\|_\infty \le M \|u\|_2$.

Remark 3.1. More general problems than (3.6)–(3.9) can be investigated in the same way as below, but we do not seek generality in this paper (the same applies to the assumptions above). We consider (3.6)–(3.9) mainly in order to simplify the notation involved in (3.2)–(3.4).

In the particular problem of interest, (3.2)–(3.4), the function f is linear in u, and I_4 linearly depends on y. Moreover, the multiplier $[1 - \beta I_4(t)]^\sigma$ is nonnegative, since $I_4(t) < N = 1$, hence the integrand in (3.2) is convex in u. Then existence of a solution of this problem is implied by the classical Tonelli argument. In the more general problem (3.6)–(3.9) we just assume existence of a local solution, further denoted by $(\hat{u}, \hat{y}, \hat{z})$. To be more precise, denote by $J(u)$ the value of the functional (3.6), where z results from (3.7)–(3.8). Also denote $\mathcal{U} := \{u \in L^\infty[0, T] : u(t) \in U \text{ for a.e. } t \in [0, T]\}$—the set of admissible controls. Then "local solution" means that $J(u) \ge J(\hat{u})$ for all $u \in \mathcal{U}$ belonging to an L^∞-neighborhood of $\hat{u}$.

Further on, we indicate by subscripts the partial derivatives of functions. For example, $g_u(z, u)$ denotes the derivative of g with respect to u (a $(1 \times m)$-dimensional matrix), $f_z(z, u)$ is a $(1 \times n)$-dimensional, etc.

PROPOSITION 3.2. *Under the assumptions made above, the functional J is Fréchet differentiable in the set $\mathcal{U}$ with respect to the L^2-norm, and its derivative has the following representation (belonging to $L^\infty[0, T]$): for every $u \in \mathcal{U}$*

$$J'(u)(t) = \lambda(t) f_u(z(t), u(t)) + g_u(z(t), u(t)), \tag{3.11}$$

where z is the solution of (3.7)–(3.8) *corresponding to u, and λ is the unique solution in $L^\infty[0,T]$ of*

$$\lambda(t) = \int_t^T \varphi(s-t)[\lambda(s) f_z(z(s),u(s)) + g_z(z(s),u(s))]\,\mathrm{d}s. \tag{3.12}$$

Notice that (3.12) is a Volterra integral equation of the second kind (in inverse time), and therefore it has a unique solution in $L^\infty[0,T]$; see, e.g., Gripenberg (1990) (we remind that any solution z of (3.7)–(3.8) corresponding to $u \in \mathcal{U}$ belongs to $L^\infty[0,T]$).

The considered optimal control problem is not standard, but the proof of Proposition 3.2 is more or less routine for specialists in optimal control theory. We present a sketch of the proof in the appendix.

Corollary 3.3. *If $(\hat u, \hat y, \hat z)$ is a locally optimal solution of problem* (3.6)–(3.9)*, then there exists a unique solution $\hat\lambda \in L^\infty[0,T]$ of* (3.12)*, with (z,u) replaced with $(\hat z, \hat u)$, such that*

$$\hat\lambda(t) f_u(\hat z(t),\hat u(t)) + g_u(\hat z(t),\hat u(t)) \in N_U(\hat u(t)) \quad \text{for a.e. } t \in [0,T],$$

where $N_U(u)$ is the usual normal cone to the convex set U at the point u.

To obtain the above necessary optimality condition it is not necessary to have differentiability of J in the space L^2 (the straightforward differentiability in L^∞ is enough). However, the differentiability in L^2 is important for investigation of various approximation approaches, such as gradient projection or Newton-type methods (see the next subsection).

Further in this section, we focus on the particular case of problem (3.2)–(3.4). Formula (3.11) takes the form

$$J'(u) = \lambda(t)\,\bar c\,\frac{z_1(t)\,z_2(t)}{z_3(t)} - p\sigma(u(t))^{\sigma-1}\big(1-\beta z_4(t)\big)^{\sigma}, \tag{3.13}$$

where λ is the solution of the Volterra equation

$$\begin{aligned}\lambda(t) = {} & \int_t^T \lambda(s)u(s)\bar c\left[\frac{z_2(s)}{z_3(s)}\varphi_1(s-t) - \frac{z_1(s)}{z_3(s)} - \frac{z_1(s)\,z_2(s)}{(z_3(s))^2}\varphi_3(s-t)\right]\mathrm{d}s \\ & + \int_t^T \left[\pi + p(u(s))^\sigma \sigma\beta\big(1-\beta z_4(s)\big)^{\sigma-1}\varphi_4(s-t)\right]\mathrm{d}s.\end{aligned} \tag{3.14}$$

Let us assume that the functions α_k, ρ, i_A, i_S are Lipschitz continuous. This is a natural assumption, since there is no reason to expect an abrupt change of these functions with the infection age, as also the statistically based estimation in Figure 2.2 suggests. Hence, the functions φ_k in (3.10) are Lipschitz continuous. From (3.14) we obtain that λ is also Lipschitz continuous. The solution with respect to u of the variational inequality in Corollary 3.3 is obvious:

$$\hat u(t) = \mathrm{proj}_{[u_0,1]}\, u^{\#}(\lambda(t), z(t)),$$

where $\mathrm{proj}_{[u_0,1]}(v)$ is the projection of the number v on the interval $[u_0,1]$ and

$$u^{\#}(\lambda,z) := \Big(\frac{\bar c \lambda z_1 z_2}{p\sigma z_3(1-\beta z_4)^\sigma}\Big)^{\frac{1}{\sigma-1}}.$$

In particular, we obtain the following result.

PROPOSITION 3.4. *Any optimal control $\hat{u}$ in problem* (3.2)–(3.4) *is Lipschitz continuous.*

Indeed, from (3.8) we obtain (after a change of the variable of integration) that z_i are Lipschitz continuous. Since $\sigma \in (0,1)$ and the assumptions made imply that $z_1(t)$ and $z_2(t)$ are separated from zero, we have that $u^{\#}(\lambda(t), z(t))$ is Lipschitz continuous, hence also $\hat{u}(t)$.

3.2. Numerical approach. At the abstract level, one can make use of Proposition 3.2 and implement any gradient projection method for minimization of the functional J subject to the constraint $u \in \mathcal{U}$ in the Hilbert space $L^2[0,T]$. Of course, in practice one has to pass to a discretized version of the problem.

To be specific, we focus on the particular problem (3.2)–(3.4), although the approach briefly discussed below is also applicable to problem (3.6)–(3.9), in principle. With a mesh of step size h in $[0,T]$, and for a given piecewise linear admissible control $u \in \mathcal{U}$, one can solve numerically the integral equation (3.3) by similar discretization techniques used for Volterra integral equations of the second kind. Namely, one can implement the rectangular rule for integration (which will result in a first order accuracy with respect to h) or the trapezoidal formula (which becomes implicit, but easily tractable, and may provide a second order accuracy). Then, one can solve the adjoint equation (3.14) for λ and obtain a discrete approximation of the gradient of J at u using (3.13). (Notice that the Volterra equation (3.14) has a separable kernel; therefore it does not provide a heavy numerical burden.) This enables implementation of any gradient procedure for mathematical programming problems. The error analysis of this solution procedure is not simple and is not a subject of this paper.

4. Parameter identification and case study. In the following, we demonstrate our proposed modeling and optimization approaches by applying them to a numerical case study. It is explained how the parametrization can be based on available epidemiological data. While we present a reasonable parametrization, additional empirical research is necessary in order to develop more realistic, fully data driven instances.

We first discuss some important waiting time distributions, estimated in the literature to describe the course of disease. Based on this, we describe the construction of the parameter functions α_k, μ, ρ. Finally, we apply the proposed optimization approach to the resulting illustrative model.

4.1. Basic data: The course of disease. Many publications on COVID-19 describe aspects of its infection course, i.e., the dependency of relevant parameters on the time since infection, or since onset of symptoms. Usually the distribution of the random time until some relevant event, e.g., onset of symptoms or death, is characterized by an estimated probability density function (PDF). This information can be restated using the related cumulative distribution function (CDF), which can be interpreted as the fraction of infected individuals for which the event already happened up to some time. Using such information fits perfectly with the proposed mathematical modeling approach.

The incubation period is the time interval between infection and the onset of symptoms. For COVID-19, several probability densities for the length of this time span have been estimated in the literature. In the subsequent numerical example we use Li et al. (2020), who estimated the incubation period density for COVID-19 as the density of a Log-normal distribution with parameters $\mu = 1.434$, $\sigma = 0.661$. Alternative estimates based on the Weibull distribution can be found, e.g., in Zhang

et al. (2020) and Guan et al. (2020).

If one analyzes chains of infection, serial intervals are important observable quantities, measuring the time between the respective onset of symptoms for a pair of an infecting and an infected individual. In the present paper we base our findings on an early estimation of the serial interval for COVID-19 from He et al. (2020). The authors analyzed the (publicly available) data of 77 transmission pairs from mainland China and estimated a $\Gamma[2.116, 2.307]$ as the serial interval distribution.

We use these early estimations, because there is evidence that the changed behavior of individuals and the measures taken by the states had an impact on the effective serial interval.

The third basic time information used in the following is the time between symptom onset and death. We use here Verity et al. (2020), who analyzed the early deaths in mainland China and estimated a Log-normal distribution with parameters $\alpha = 2.81,\ \beta = 0.370$. We mention that, based on a much larger sample, Wu and McGoogan (2020) communicated very similar estimates.

While these three distributions contain the basic information on the course of disease for the subsequent case study, further information, e.g., the time between onset of symptoms and reporting of a case (see, e.g., MIDAS Network (2020)), may be used for more detailed modeling.

4.2. The benchmark parameters. The available information on the course of infection can be used to construct the parameter functions α_k, μ, ρ, and c. Without going into the details, we give a sketch of the considerations that lead to the parameter functions depicted in Figure 2.2, which are used subsequently for the optimization case study.

The generation interval is the difference between the infection time of an infected individual and the infection time of the infecting person. An estimated PDF of the generation interval is based on pairs of infectors and infected individuals, and must be interpreted as the conditional PDF that an infection happened at time θ after the infector was infected, given that the secondary infection takes place. The infectiousness function i_S is obtained by multiplying the generation time PDF with the probability that a contact between an infected and a susceptible individual leads to an infection. For the case study we assume that asymptomatic and symptomatic cases have the same infectiousness, i.e., $i_S(\theta) = i_A(\theta)$.

The generation time PDF can be derived from the incubation time density and the serial interval density: let $T_I^1,\ T_I^2$ denote i.i.d. versions of the incubation time (i.e., the incubation time for a first and a second infection). Moreover, let T_S be the serial interval for these two infections. Assuming independence, the time between the first and the second infection is then given by

$$T_{II} = T_I^1 + T_S - T_I^2. \tag{4.1}$$

Because we know that the first individual infects the second, it is necessary to apply the condition $\tilde{T}_{II} > 0$. Consequently, the PDF for the time between consecutive infections is

$$f_{II}(\theta) = \frac{f_{T_I} * f_{T_S} * f_{-T_I}(\theta)}{\int_0^\infty f_{T_I} * f_{T_S} * f_{-T_I}(s)\ \mathrm{d}s} \tag{4.2}$$

for $\theta \geq 0$ where for any random variable X the related probability density is denoted by f_X and $*$ denotes the convolution of densities. For the case study, we applied this approach to the estimated densities in Li et al. (2020) and He et al. (2020).

TABLE 1
Estimated parameter values.

Overall probability of infection per contact	0.1199, estimated
Basic contact rate	15.37, estimated
Reduction of contact rate for isolated individuals	0.057, estimated
Reduction of contact rate for nonisolated symptomatic individuals	0.98, estimated
Surviving of symptomatically infected	0.89, JHCRC (2020)
Infected individuals without symptoms	0.425, estimated
Percentage of isolated asymptomatic individuals	0.41, estimated
Percentage of isolated symptomatic individuals	0.93, estimated
Individuals infected before time zero	0.00001, chosen

The class of infected individuals consists of individuals who show symptoms over some time and of *completely asymptomatic* infected individuals who recover without having shown any symptoms. Asymptomatic cases at time t in the sense of the proposed model (I_1 and I_2) are infected individuals who do not have symptoms at time t, whether or not they show symptoms later on.

Only individuals with symptoms may die, and we use the mortality function estimated by Verity et al. (2020), as discussed above. The function μ can then be constructed by multiplying the mortality function with the proportion of eventually symptomatic individuals within the class of infected and with the proportion of fatally ill individuals within the class of eventually symptomatics.

The joint recovery function ρ for symptomatic individuals is constructed based on information about symptomatic cases such that individuals recover on an average of 14 days after the onset of symptoms. We assume that recovery of fully asymptomatic cases follows the same timing.

The way into isolation works in different ways for symptomatic and completely asymptomatic individuals. We assume that a (high) proportion of eventually symptomatic cases becomes isolated when symptoms occur. Within this group, the fraction of isolated cases at time θ is given by the CDF of the incubation time, discussed above. On the other hand, a (lower) proportion of completely asymptomatic cases is isolated.

The way out of isolation is modeled as follows: individuals that show symptoms begin to leave isolation 14 days after the onset of symptoms, a small proportion still remains infected and leaves isolation later, according to decay of the infectiousness function. All infected individuals without any symptoms leave isolation after 14 days.

At several points, proportionality factors and further basic parameters (like starting values for the new cases y) are needed to fully calibrate the model. Most of these values were estimated by fitting the model to official data (Jan. 31–May 1, 2020) from Great Britain, downloaded from CDCP (2020). Table 1 states the used values.

In order to estimate the parameter vector σ, containing the parameters in Table 1, we use the number of daily (day i) observed new infections y_i and the number of daily observed newly dead individuals D_i for fitting the model. Recall that according to our assumptions we observe infected only when they get symptoms. Therefore, $y(t)$—the number of newly infected—in our model cannot be directly compared with the observed number of infected and it is important to consider the incubation time and the fact that not all infected show symptoms.

We now write $y^\sigma(t)$ instead of $y(t)$ to emphasize the dependence of the central model variable on the parameter vector. Then, using our model based on a parameter

vector σ, an estimate $\hat{y}_i^\sigma$ of the observed y_i is obtained by

$$\hat{y}_i^\sigma = z^\sigma(i) - z^\sigma(i-1), \tag{4.3}$$

where

$$z^\sigma(t) = \pi_{SQ} \int_0^t \int_0^\Theta y^\sigma(t-s) f_{T_I}(s)\, ds \tag{4.4}$$

and π_{SQ} is the product of the fraction of infected with symptoms and the fraction of isolated asymptomatic cases (both in Table 1).

In the same manner, given a parameter vector σ, the dead individuals resulting from the model, $\hat{D}_i^\sigma$ can be estimated by

$$\hat{D}_i^\sigma = D^\sigma(i) - D^\sigma(i-1), \tag{4.5}$$

where $D^{\sigma(t)}$ denotes the dead individuals up to time t based on the model when parameter vector σ is used. This estimate is compared with the observed dead at day i, D_i.

Based on these preparations, we use nonlinear regression to get an estimate $\hat{\sigma}$ for the parameter vector σ:

$$\hat{\sigma} = \arg\min_\sigma \left\{ \sum_{i=1}^N |\hat{y}_i^\sigma - y_i| + \omega \sum_{i=1}^N |\hat{D}_i^\sigma - D_i| \right\}, \tag{4.6}$$

where ω is a weight, in our numerical example chosen such that the weighted numbers for dead individuals and newly infected have the same order of magnitude. The absolute deviation is used here instead of the squared-deviation to increase robustness of the estimated parameters.

4.3. Numerical results. Below we present some numerical results showing the evolution of the epidemic in the benchmark scenario without and with implementation of optimal control policies. Three optimal control scenarios are considered: case 1, where the weight $p = 0.0375$; case 2, with $p = 0.0350$, and case 3, with $p = 0.0310$. Thus, in case 1 more weight is attributed to the economic losses, in case 2 the humanitarian component is more important, and even more in case 3.

Optimal policies of global social distancing are computed for the benchmark scenario in cases 1–3 with $\pi = 1$ and control restrictions $u(t) \in [0.5, 1]$. The left plot in Figure 4.1 shows the corresponding optimal controls. A remarkable observation is that the more weight is attributed to the economic losses (case 1), the later emerge and the sooner are removed the social contact restrictions, while the strength of the restrictions is higher. In case 3, for example, the control never reaches the lower bound $u = 0.5$, but the period in which restrictions are in effect is more than 50 weeks, while for case 1 it is only about 18 weeks. The reason is that when the prevalence is small, the contact restrictions bring a small reduction of the new cases (in absolute numbers), but the economic losses from the restrictions are only slightly dependent on the prevalence. Thus contact restrictions are not efficient in times of low prevalence, unless the parameter p is small enough.

Another remarkable observation is that in all cases the epidemic becomes almost extinct, both in the uncontrolled and the optimally controlled cases. For the uncontrolled case this is expected due to the obtained "herd immunity," but for the optimally controlled scenarios the number of recovered individuals is too low for that.

In fact, in the controlled cases a second wave actually appears (beyond the plotted time horizon), but it is due to the finiteness of the time interval on which the optimal control problem is solved ($T = 450$ days). We mention that the above observations are valid in numerous additional experiments not presented here.

The right plot in Figure 4.1 and the two plots in Figure 4.2 show the evolution of the main compartment sizes without and with implementation of optimal control policies in cases 1 and 2. The shaded areas become rather tiny in case 3 and are not plotted. Applying no policy measures leads to 1.53% dead individuals during the considered time horizon of one year, and a maximum of 5.29% of the population is infected simultaneously. These numbers decrease to $(0.5\%, 2.72\%)$ in case 1, $(0.2\%, 1.0\%)$ in case 2, and $(0.05\%, 0.9\%)$ in case 3.

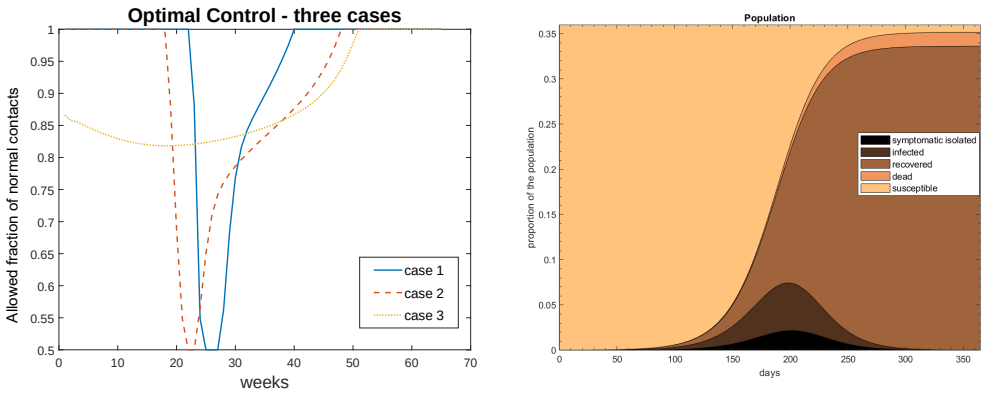

FIG. 4.1. *Left plot shows three optimal control functions: for case* 1 *($p = 0.0375$), case* 2 *($p = 0.0350$), and case* 3 *($p = 0.0310$). Right plot: evolution of the population compartments without control.*

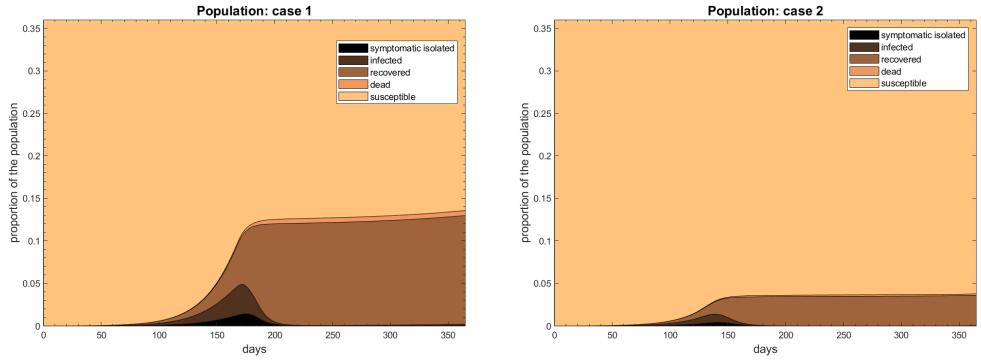

FIG. 4.2. *Optimal evolution of the population compartments in case* 1 *($p = 0.0375$) and case* 2 *($p = 0.0370$).*

Figure 4.3 shows the optimal trade-off between the proportion of dead individuals at the end of the planning horizon versus the reduction of GDP as a percentage of the pre-epidemic GDP. The figure can therefore be interpreted as showing the efficient frontier for the two parts of the objective (3.2). It is computed by solving the optimal control problem for a large number of weights p. The points on the efficiency frontier

corresponding the uncontrolled case and to cases 1 and 2 are indicated on the figure. The reduction of the GDP is about $0.5\,‰$ in the uncontrolled case, while it increases to about $41\,‰$ in case 1 and to $61\,‰$ in case 2 (we remind that only direct economic costs due to lost working hours are counted).

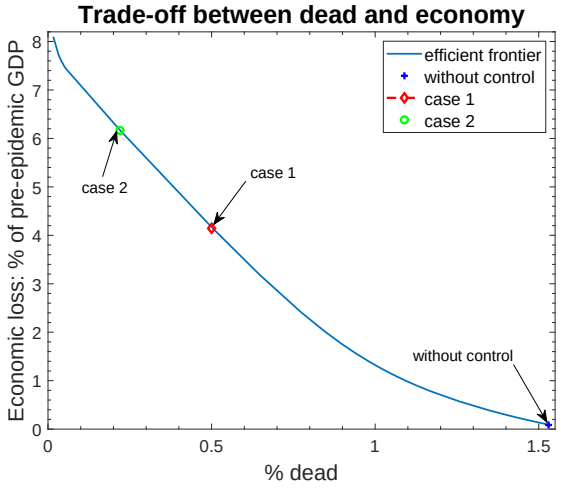

FIG. 4.3. *Efficient frontier: Trade-off between the proportion of dead individuals and the economic loss. The case without social distancing and the two main cases are placed on the curve.*

Finally, Figure 4.4 depicts the reproduction number $\mathcal{R}_{\mathrm{coh}}(t)$ together with the number of new cases $y(t)$ (appropriately rescaled to fit to the same figure) for the uncontrolled case and in case 1. Clearly, the value $\mathcal{R}_{\mathrm{coh}}(t)$ gives an indication for the change in the number of new cases in the near future. In particular 4–7 days after $\mathcal{R}_{\mathrm{coh}}(t)$ crosses the line $\mathcal{R} = 1$, the trend of the new cases changes qualitatively (increase versus decrease).

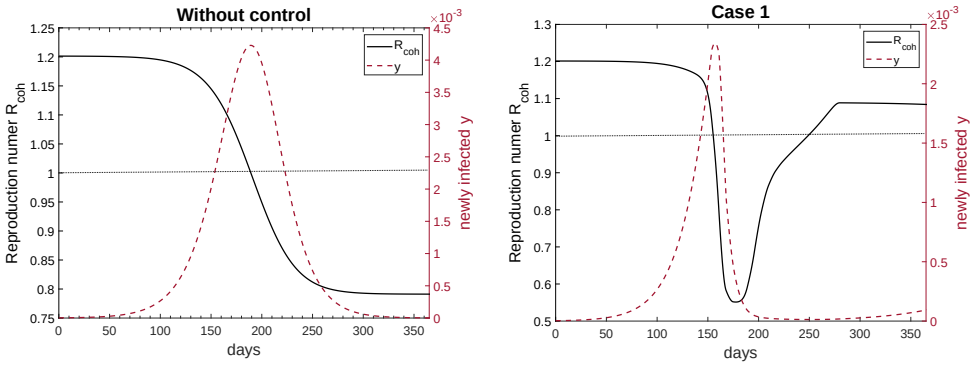

FIG. 4.4. *Reproduction number $\mathcal{R}_{\mathrm{coh}}$ and the number of new cases $y(t)$ (multiplied by 10^{-3}. Left plot: no control is applied. Right plot: the optimal control as in case 1 is applied.*

5. Discussion. The aim of this work is to develop a new epidemiological modeling approach that describes the dynamics of an epidemic also under interventional conditions that differs from the typical compartmental models with transition rates

between the involved epidemic subpopulations. The model is based on a minimal number of dynamic variables with high generalization potential. Specifically, the model is built up using the number of new infections as a single time-dependent variable. Optimal control theory approaches can be embedded in the model including substantial expansions of the specific features of an epidemic. The model is applied to the COVID-19 pandemic; however, it can be employed and adapted for other types of epidemics with different transmission modes and a variety of intervention measures.

The main feature of the presented model is that it is based on the probability of an infected individual being in a particular stage of the disease (asymptomatic, symptomatic, isolated, dead, etc.) and of the infectiousness, depending on the time since infection.

As already mentioned, more possible stages an infected individual might undergo (hospitalization, intensive care) can be easily incorporated in the basic integral equation without any structural change. Changes of immunological status after recovery can also be modeled within the same framework.

More challenging are other widely used policy measures, in particular quarantine as a result of testing and/or contact tracing. The model will need an upgrade which is a subject of current work. Another issue is the optimization of vaccination or treatment strategies that may become available during the course of the epidemic. For instance, the population groups to be prioritized to receive a vaccine as well as which vaccination strategy should be implemented. The intuitive strategy for vaccination might not necessarily be the prioritization for health care workers and vulnerable populations but rather those with high risk of transmitting the pathogen due to their working, living, or other conditions. An upgraded version of the model in this paper, which includes vaccination as a control policy, is a subject of a separate paper.

Several optimal control problems can be meaningful in the context of pandemics the size of COVID-19. For example, the one of minimization of the economic losses subject to the constraint that the number of infected individuals does not exceed a certain bound. This is an optimal control problem of the same form as the one considered in this paper, but with a state constraint, which makes it more demanding.

The economic component of the optimal control problem is introduced mainly for illustrative purposes. More detailed economic modeling should include at least a dynamic economic component (cf. Acemoglu et al. (2020) and Bloom et al. (2020)).

6. Appendix.

Proof of Proposition 2.1. For any nonnegative function $y \in L^2[0,T]$ one can define $S[y](t)$ and $R[y](t)$ as in (2.2)–(2.3) (after extending y as y_0 on $[0,-\Theta]$). Apparently, both $S[y]$ and $R[y]$ belong to $L^2[0,T]$, $S[y]$ is continuous, $S[y](0) = S(0)$ and $R[y](0) = R(0)$ are independent of y, and $S[y](t) \le S(0)$ for every $t \in [0,T]$. Equation (2.11) can be rewritten as

$$y(t) = c(t)\frac{d_1(t) + \int_0^t \tilde{q}(\theta)y(t-\theta)\,\mathrm{d}\theta}{S[y](t) + R[y](t) + \int_{-\Theta}^0 q(t-\tau)y_0(\tau)\,\mathrm{d}\tau + \int_0^t q(\theta)y(t-\theta)\,\mathrm{d}\theta}S[y](t). \tag{6.1}$$

Denote $i := \max_{\theta\in[0,\Theta]}\{i_A(\theta), i_S(\theta)\}$, and let $\sigma > 0$ be such that $i\int_0^\sigma c(t)\,\mathrm{d}t \le 1$. Define the set

$$\mathcal{Y}_\sigma := \{y \in L^\infty[0,\sigma] : y(t) \ge 0,\ S[y](t) \in [0,\ ic(t)S(0)] \text{ for a.e. } t \in [0,\sigma]\}.$$

Let the mapping $\mathcal{F} : \mathcal{Y}_\sigma \to L^2[0,\sigma]$ be defined so that $\mathcal{F}(y)(t)$ is the right-hand side of (6.1) for $t \in [0,\sigma]$. Since $S[y](t) \ge 0$ for $y \in \mathcal{Y}_\sigma$, assumption (iii) implies that the

denominator in (6.1) is not smaller than η; thus $\mathcal{F}(y)$ is well defined on $\mathcal{Y}_\sigma$. Next, we show that $\mathcal{Y}_\sigma$ is invariant with respect to $\mathcal{F}$. From (6.1) it is obvious that $\mathcal{F}(y)(t) \geq 0$. Moreover, due to assumption (i) and the inequality $\tilde{q}(\theta) \leq iq(\theta)$, we have (now using (2.8)) that for $y \in \mathcal{Y}_\sigma$

(6.2)
$$\mathcal{F}(y)(t) \leq \frac{c(t) i \int_0^{\Theta+t} q(\theta) y(t-\theta)\,\mathrm{d}\theta}{S[y](t) + R[y](t) + \int_0^{\Theta+t} q(\theta) y(t-\theta)\,\mathrm{d}\theta} S[y](t) \leq ic(t) S[y](t) \leq ic(t) S(0).$$

In particular, $\mathcal{F}(y) \in L^2[0,\sigma]$. Finally, for $t \in [0,\sigma]$,

$$\begin{aligned} S[\mathcal{F}(y)](t) = S[y](0) - \int_0^t \mathcal{F}(y)(\tau)\,\mathrm{d}\tau &\geq S(0) - i\int_0^t c(\tau)\,\mathrm{d}\tau S(0) \\ &\geq S(0)\left(1 - i\int_0^\sigma c(\tau)\,\mathrm{d}\tau\right) \geq 0. \end{aligned}$$

Thus $\mathcal{F}(\mathcal{Y}_\sigma) \subset \mathcal{Y}_\sigma$. We define in $\mathcal{Y}_\sigma$ the metric induced by the weighted norm

$$\|y\|_{2,\nu} := \left(\int_0^\sigma e^{-\nu t} |y(t)|\,\mathrm{d}t\right)^{\frac{1}{2}}, \quad \nu > 0,$$

in the space $L^2[0,\sigma]$. With this metric, $\mathcal{Y}_\sigma$ is a complete metric space. In order to prove that $\mathcal{F}$ is contractive, we mention (skipping the details) that there exists a constant L (independent of $c \in L^\infty[0,T]$) such that

$$|\mathcal{F}(y_1)(t) - \mathcal{F}(y_2)(t)| \leq Lc(t) \int_0^t |y_1(s) - y_2(s)|\,\mathrm{d}s \tag{6.3}$$

for every $y_1, y_2 \in \mathcal{Y}_\sigma$ and a.e. $t \in [0,\sigma]$. This is due to assumption (iii), which implies that the denominator in (6.1) is not smaller than $\eta > 0$. Then, denoting $\Delta y(t) := y_1(t) - y_2(t)$, we have

$$\begin{aligned} \|\mathcal{F}(y_1) - \mathcal{F}(y_2)\|_{2,\nu}^2 &\leq L^2 \int_0^\sigma e^{-\nu t} (c(t))^2 \left(\int_0^t |\Delta y(s)|\,\mathrm{d}s\right)^2 \mathrm{d}t \\ &\leq L^2 \|c\|_2^2 \sup_{t\in[0,t]} e^{-\nu t} \left(\int_0^t |\Delta y(s)|\,\mathrm{d}s\right)^2 \\ &= L^2 \|c\|_2^2 \sup_{t\in[0,t]} e^{-\nu t} \left(\int_0^t e^{\frac{\nu s}{2}} \left(e^{-\frac{\nu s}{2}} |\Delta y(s)|\right) \mathrm{d}s\right)^2 \\ &\leq L^2 \|c\|_2^2 \sup_{t\in[0,t]} e^{-\nu t} \int_0^t e^{\nu s}\,\mathrm{d}s \int_0^t e^{-\nu s} |\Delta y(s)|^2\,\mathrm{d}s \\ &\leq L^2 \|c\|_2^2 \sup_{t\in[0,t]} e^{-\nu t} \frac{1}{\nu}\left(e^{\nu t} - 1\right) \|\Delta y\|_{2,\nu} \leq \frac{L^2 \|c\|_2^2}{\nu} \|\Delta y\|_{2,\nu}. \end{aligned}$$

We can choose $\nu > L^2\|c\|_2^2$, so that $\mathcal{F}$ is contractive. Thus $y = \mathcal{F}(y)$ has a unique solution $\hat{y} \in \mathcal{Y}_\sigma$.

Now, we have to extend it (if $\sigma < T$) to a solution of (6.1) on $[0,T]$. For this, using (6.1) we estimate $\hat{S} := S[\hat{y}]$ for $t \in [0,\sigma]$ as follows:

$$\hat{S}(t) = S(0) - \int_0^t \hat{y}(s)\,\mathrm{d}s = S(0) - \int_0^t l(s)\hat{S}(s),$$

where

$$l(s) := c(t)\frac{\int_0^{\Theta+s} \tilde{q}(\theta)\hat{y}(s-\theta)\,\mathrm{d}\theta}{S[\hat{y}](s) + R[\hat{y}](s) + \int_0^{\Theta+s} q(\theta)\hat{y}(s-\theta)\,\mathrm{d}\theta} \le ic(t)$$

as already argued above. Hence,

$$\hat{S}(\sigma) \ge e^{-i\int_0^\sigma c(s)\,\mathrm{d}s} S(0) > 0. \tag{6.4}$$

Thus assumption (ii) is also fulfilled for the same equation (2.11), but starting at time σ. Moreover, $R(\sigma) \ge R(0) \ge \eta$, thus (iii) is also satisfied. One can repeat the above argument to show that a solution exists on $[\sigma, 2\sigma]$ (or $[0, T]$) with $y(t) = \hat{y}(t)$ on $[\sigma - \Theta, \sigma]$, and so on until time T is reached. Thus a unique solution of (6.1) exists in the set $\mathcal{Y}_T$.

Now, consider an arbitrary solution y of (6.1) in $L^2[0, T]$. Assume that $S := S[y](t)$ is not nonnegative. Since S is continuous and $S(0) > 0$ according to assumption (ii), there exists a first time $\tau > 0$ such that $S(\tau) = 0$. Since $S(t) > 0$ on $[0, \tau)$, one can easily argue from (6.1) that $y(t) \ge 0$ for a.e. $t \in [0, \tau]$. But then we can estimate as in (6.4) that $S(\theta) \ge e^{-i\int_0^\theta c(s)\,\mathrm{d}s} S(0) > 0$, which contradicts the assumption that $S(\tau) = 0$. Thus S is nonnegative, hence also y. The inequality $S(t) \le ic(t)S(0)$ is fulfilled by the same argument as in (6.2). Then $y \in \mathcal{Y}_T$ and it must coincide with $\hat{y}$.

The other claims of the proposition have already been proved. ∎

Derivation of the reproduction rates (2.13)–(2.14). The cohort of newly infected individuals at time $t \in [0, T - \Theta]$ may infect other individuals on $[t, t + \Theta]$. According to the random mixing assumptions, the number of new cases at time $\tau \in [t, t + \Theta]$ caused by the cohort $y(t)$ is

$$\frac{\tilde{q}(\tau - t)y(t)}{S(\tau) + R(\tau) + \int_0^\Theta q(\theta)y(\tau - \theta)\,\mathrm{d}\theta}\, S(\tau).$$

Integrating on $[t, t + \Theta]$ we obtain the total numbers of secondary cases caused by the cohort $y(t)$. The per capita secondary cases are given by (2.14).

Now, we shall derive (2.13). Observe that the new cases at time $\tau \in [t, t + \Theta]$ caused by all infectious, $I(t)$, at time t may result from any cohort $y(s)$ with $s \in [t - \Theta, t]$. Thus the number of new cases at time τ caused by $I(t)$ is

$$\frac{\int_{\tau-\Theta}^t \tilde{q}(\tau - s)y(s)\,\mathrm{d}s}{S(\tau) + R(\tau) + \int_0^\Theta q(\theta)y(\tau - \theta)\,\mathrm{d}\theta}\, S(\tau) = \frac{\int_{\tau-t}^\Theta \tilde{q}(\theta)y(\tau - \theta)\,\mathrm{d}\theta}{S(\tau) + R(\tau) + \int_0^\Theta q(\theta)y(\tau - \theta)\,\mathrm{d}\theta}\, S(\tau).$$

Integrating on $[t, t + \Theta]$ to obtain the total number of secondary cases caused by $I(t)$ and dividing by $I(t)$ we obtain the expression (2.13).

Proof of Proposition 3.2. Let us fix an arbitrary $u \in \mathcal{U}$ and consider an increment $\Delta u \in L^2[0, T]$. In the calculation of the derivative of J it is enough to consider Δu with $\|u\|_2 \le 1$. The corresponding solutions (y, z) and $(y + \Delta y, z + \Delta z)$ are assumed to exist. Moreover, z and $z - \Delta z$ are bounded (uniformly in Δu $\|u\|_2 \le 1$) due to the last assumption in section 3.1. Then $J(u)$ and $J(u + \Delta u)$ are finite because of the assumed Lipschitz properties of g.

Below we use first order Taylor expansions, where r_g and r_f denote the corresponding residuals, which will be analyzed at the end. Moreover, we shorten $f_z(z(t),$

$u(t)) =: f_z(t)$, similarly for f_u, g_z, g_u. We have

$$J(u+\Delta u) - J(u) = \int_0^T [g_z(t)\Delta z(t) + g_u(t)\Delta u(t) + r_g(t)]\,\mathrm{d}t, \tag{6.5}$$

$$\Delta z(t) = \int_0^t \varphi(\theta)\Delta y(t-\theta)\,\mathrm{d}\theta,$$

$$0 = -\Delta y(t) + f_z(t)\int_0^t \varphi(\theta)\Delta y(t-\theta)\,\mathrm{d}\theta + f_u(t)\Delta u(t) + r_f(t). \tag{6.6}$$

We multiply (6.6) with the solution λ of (3.12) (which exists, as explained after the formulation of the proposition), integrate on $[0,T]$, and add to (6.5) to obtain

$$\begin{aligned} J(u+\Delta u) - J(u) &= \int_0^T \Big[g_z(t)\int_0^t \varphi(\theta)\Delta y(t-\theta)\,\mathrm{d}\theta + g_u(t)\Delta u(t) + r_g(t) \\ &\quad - \lambda(t)\Delta y(t) + \lambda(t) f_z(t)\int_0^t \varphi(\theta)\Delta y(t-\theta)\,\mathrm{d}\theta + \lambda(t) f_u(t)\Delta u(t) + \lambda(t) r_f(t)\Big]\,\mathrm{d}t \\ &= \int_0^T \Big[g_u(t)\Delta u(t) + \lambda(t) f_u(t)\Delta u(t) + r_g(t) + \lambda(t) r_f(t)\Big]\,\mathrm{d}t + \Lambda, \end{aligned}$$

where

$$\begin{aligned} &\Lambda \\ &= \int_0^T \Big[-\lambda(t)\Delta y(t) + g_z(t)\int_0^t \varphi(t-s)\Delta y(s)\,\mathrm{d}s + \lambda(t) f_z(t)\int_0^t \varphi(t-s)\Delta y(s)\,\mathrm{d}s\Big]\,\mathrm{d}t \\ &= \int_0^T \Big[-\lambda(s)\Delta y(s) + \int_s^T \big(g_z(t)\varphi(t-s)\Delta y(s) + \lambda(t) f_z(t)\varphi(t-s)\Delta y(s)\big)\,\mathrm{d}t\Big]\,\mathrm{d}s = 0 \end{aligned}$$

due to the choice of λ in (3.12).

To estimate the terms r_f and r_g is a matter of elementary analysis. We give only a sketch. Knowing that $z+\Delta z$ is bounded in L^∞, uniformly in Δu with $\|\Delta u\|_2 \le 1$, and using the Grönwall inequality we obtain the estimate $\|\Delta z\|_\infty \le c_1\|\Delta u\|_1 \le c_2\|\Delta u\|_2$ with appropriate constants c_1 and c_2. Then we can use the inequality

$$\begin{aligned} |r_f(t)| \le \sup_{s\in[0,1]} \big[&|f_z(z(t)+s\Delta z(t), u(t)+s\Delta u(t) - g_z(t))|\,|\Delta z(t)| \\ &+ |f_u(z(t)+s\Delta z(t), u(t)+s\Delta u(t)) - g_u(t)|\,|\Delta u(t)|\big] \end{aligned}$$

and the similar one for r_g to estimate $\|r_f\|_2 \le c_3\|\Delta u\|_2^2$ and complete the proof. ∎

REFERENCES

D. Acemoglu, V. Chernozhukov, I. Werning, and M. D. Whinston (2020), *Optimal Targeted Lockdowns in a Multi-Group SIR Model*, Working Paper 27102, National Bureau of Economic Research.

D. E. Bloom, M. Kuhn, and K. Prettner (2020), *Modern Infectious Diseases: Macroeconomic Impacts and Policy Responses*, IZA Discussion Papers 13625.

Centers for Disease Control and Prevention (CDCP) (2020), *COVID-19 Coronavirus Data – European Union Open Data Portal*, https://data.europa.eu/euodp/en/data/dataset/covid-19-coronavirus-data.

A. L. Dontchev and V. M. Veliov (2009), *Metric regularity under approximations*, Control Cybernet., 38, pp. 1283–1303.

G. Feichtinger, V. M. Veliov, and T. Tsachev (2004), *Maximum principle for age and duration structured systems: A tool for optimal prevention and treatment of HIV*, Math. Popul. Stud., 11, pp. 3–28.

G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri (2020), *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy*, Nat. Med., 26, pp. 855–860.

G. Gripenberg, S. O. Londen, and O. Staffans (1990), *Volterra Integral and Functional Equations*, Encyclopedia Math. Appl. 34, Cambridge University Press, Cambridge.

W. Guan, Z. Ni, Yu Hu, W. Liang, C. Ou, J. He, L. Liu, H. Shan, C. Lei, D. S. C. Hui, B. Du, L. Li, G. Zeng, K.-Y. Yuen, R. Chen, C. Tang, T. Wang, P. Chen, J. Xiang, S. Li, J.-L. Wang, Z. Liang, Y. Peng, L. Wei, Y. Liu, Y.-H. Hu, P. Peng, J.-M. Wang, J. Liu, Z. Chen, G. Li, Z. Zheng, S. Qiu, J. Luo, C. Ye, S. Zhu, and N. Zhong for the China Medical Treatment Expert Group for Covid19 (2020), *Clinical characteristics of Coronavirus disease* 2019 *in China*, N. Engl. J. Med., 382, pp. 1708–1720.

X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li , and G. M. Leung (2020), *Temporal dynamics in viral shedding and transmissibility of COVID-19*, Nat. Med., 26, pp. 672–675.

J. Hellewell, S. Abbott, A. Gimma, N. I. Bosse, C. I. Jarvis, H. Gibbs, K. van Zandvoort, S. Funk, and R. M. Eggo (2020), *Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts*, Lancet Glob. Health, 8, pp. e488–e496.

A. T. Huang, B. Garcia-Carreras, M. D. T. Hitchings, B. Yang, L. C. Katzelnick, S. M. Rattigan, B. A. Borgert, C. A. Moreno, B. D. Solomon, I. Rodriguez-Barraquer, J. Lessler, H. Salje, D. Burke, A. Wesolowski, and D. A. T. Cummings (2020), *A systematic review of antibody mediated immunity to coronaviruses: Kinetics, correlates of protection, and association with severity*, Nat. Commun., 11, 4704.

John Hopkins Coronavirus Resource Center (JHCRC) (2020), *Mortality Analyses, status: Sept.* 2020, https://coronavirus.jhu.edu/data/mortality.

W. O. Kermack and A. G. McKendrick (1927), *A contribution to the mathematical theory of epidemics*, Proc. Roy. Soc. London Ser. A, 115, pp. 700–721.

M. Kretzschmar, G. Rozhnova, M. C. J. Bootsma, M. van Boven, J. H. H. M. van den Wijgert, and M. J. M. Borten (2020), *Impact of delays on effectiveness of contact tracing strategies for COVID-19: A modelling study*, Lancet Public Health, 5, pp. 452–459.

N. Kojima, J. D. Klausner (2022), *Protective immunity after recovery to SARS-CoV-2 infection*, Lancet Infect. Dis., 22, pp. 12–14, https://doi.org/10.1016/S1473-3099(21)00676-9.

O. Kounchev, G. Simeonov, and Z. Kuncheva (2020), *The TVBG-SEIR Spline Model for Analysis of COVID-19 Spread and a Tool for Prediction Scenarios*, preprint, https://arxiv.org/abs/2004.11338, 2020.

Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, and Z. Feng (2020), *Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia*, N. Engl. J. Med., 382, pp. 1199–1207.

J. Ma (2020), *Estimating epidemic exponential growth rate and basic reproduction number*, Infect. Dis. Model., 5, pp. 129–141.

S. Margenov, N. Popivanov, I. Ugrinova, S. Harizanov, and T. Hristov (2020), *Mathematical and Computer Modeling of COVID-19 Transmission Dynamics in Bulgaria by Time-depended Inverse SEIR Model*, preprint, https://arxiv.org/abs/2008.10360, 2020.

MIDAS Network, *midas-network/COVID-19*, GitHub, https://github.com/midas-network/COVID-19.

K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell (2020), *Estimating the asymptomatic proportion of coronavirus disease* 2019 *(COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan,* 2020, Euro. Surveill., 25, 2000180.

H. Nishiura, T. Kobayashi, T. Miyama, A. Suzuki, S.-M. Jung, K. Hayashi, R. Kinoshita, Y. Yang, B. Yuan, A. R. Akhmetzhanov, and N. M. Linton (2020), *Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19)*, Int. J. Infect. Dis., 94, pp. 154–155.

S. W. Park, D. M. Cornforth, J. Dushoff, and J. S. Weitz (2020), *The time scale of asymp-*

tomatic transmission affects estimates of epidemic potential in the COVID-19 outbreak, Epidemics, 31, 100392.

C. M. Peak, R. Kahn, Y. H. Grad, L. M. Childs, R. Li, M. Lipsitch, and C. O. Buckee (2020), *Individual quarantine versus active monitoring of contacts for the mitigation of COVID-19: A modelling study*, Lancet Infect. Dis., 20, pp. 1025–1033.

Y. Peng, A. J. Mentzer, G. Liu, X. Yao, Z. Yin, D. Dong, W. Dejnirattisai, T. Rostron, P. Supasa, C. Liu, C. López-Camacho, J. Slon-Campos, Y. Zhao, D. I. Stuart, G. C. Paesen, J. M. Grimes, A. A. Antson, O. W. Bayfield, D. E. D. P. Hawkins, D.-S. Ker, B. Wang, L. Turtle, K. Subramaniam, P. Thomson, P. Zhang, C. Dold, J. Ratcliff, P. Simmonds, T. de Silva, P. Sopp, D. Wellington, U. Rajapaksa, Y.-L. Chen, M. Salio, G. Napolitani, W. Paes, P. Borrow, B. M Kessler, J. W. Fry, N. F. Schwabe, M. G. Semple, J. K. Baillie, S. C. Moore, P. J. M. Openshaw, M. Azim Ansari, S. Dunachie, E. Barnes, J. Frater, G. Kerr, P. Goulder, T. Lockett, R. Levin, Y. Zhang, R. Jing, L.-P. Ho, Oxford Immunology Network Covid-19 Response T cell Consortium; ISARIC4C Investigators; R. J. Cornall, C. P. Conlon, P. Klenerman, G. R. Screaton, J. Mongkolsapaya, A. McMichael, J. C. Knight, G. Ogg, and T. Dong (2020), *Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19*, Nat. Immunol., 21, pp. 1336–1345.

H. Thieme and C. Castillo-Chavez (1993), *How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS?*, SIAM J. Appl. Math., 53, pp. 1447–1479, https://doi.org/10.1137/0153068.

C. Tsay, F. Lejarza, M. A. Stadtherr, and M. Baldea (2020), *Modeling, state estimation, and optimal control for the US COVID-19 outbreak*, Sci. Rep., 10, 10711.

A. Pan, L. Liu, C. Wang, H. Guo, X. Hao, Q. Wang, J. Huang, N. He, H. Yu, X. Lin, S. Wei, and T. Wu (2020), *Associations of public health interventions with the epidemiology of COVID-19 outbreak in Wuhan, China*, JAMA, 323, pp. 1915–1923.

R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, J. T. Griffin, M. Baguelin, S. Bhatia, A. Boonyasiri, A. Cori, Z. Cucunubá, R. FitzJohn, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y. Wang, X. Xi, C. A. Donnelly, A. C. Ghani, and N. M. Ferguson (2020), *Estimates of the severity of coronavirus disease 2019: A model-based analysis*, Lancet Infect. Dis., 20, pp. 669–677.

Z. Wu and J. M. McGoogan (2020), *Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese center for disease control and prevention*, JAMA, 323, pp. 1239–1242.

J. Zhang, M. Litvinova, W. Wang, Y. Wang, X. Deng, X. Chen, M. Li, W. Zheng, L. Yi, X. Chen, Q. Wu, Y. Liang, X. Wang, J. Yang, K. Sun, I. M. Longini, Jr., M. E. Halloran, P. Wu, B. J. Cowling, S. Merler, C. Viboud, A. Vespignani, M. Ajelli, and H. Yu (2020), *Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: A descriptive and modelling study*, Lancet Infect. Dis., 20, pp. 793–802.