SIAM J. CONTROL OPTIM.
Vol. 60, No. 2, pp. S145–S169

HOW MUCH TESTING AND SOCIAL DISTANCING IS REQUIRED TO CONTROL COVID-19? SOME INSIGHT BASED ON AN AGE-DIFFERENTIATED COMPARTMENTAL MODEL*

SARA GRUNDEL†, STEFAN HEYDER‡, THOMAS HOTZ‡, TOBIAS K. S. RITSCHEL†, PHILIPP SAUERTEIG‡, AND KARL WORTHMANN‡

Abstract. In this paper, we provide insights on how much testing and social distancing is required to control COVID-19. To this end, we develop a compartmental model that accounts for key aspects of the disease: incubation time, age-dependent symptom severity, and testing and hospitalization delays; the model's parameters are chosen based on medical evidence, and, for concreteness, adapted to the German situation. Then, optimal mass-testing and age-dependent social distancing policies are determined by solving optimal control problems both in open loop and within a model predictive control framework. We aim to minimize testing and/or social distancing until herd immunity sets in under a constraint on the number of available intensive care units. We find that an early and short lockdown is inevitable but can be slowly relaxed over the following months.

Key words. compartmental modeling, SARS-CoV-2, decision support, model predictive control, optimal control

AMS subject classifications. 93C10, 93C15, 93C35, 93A30

DOI. 10.1137/20M1377783

1. Introduction.

The *severe acute respiratory syndrome coronavirus* 2 (SARS-CoV-2) is a strain of coronavirus which causes the respiratory illness *coronavirus disease* 2019 (COVID-19). On March 11, 2020, the World Health Organization (WHO) declared the outbreak of SARS-CoV-2 a pandemic [58]. Due to the novelty of the virus, there was (at the time of submitting this manuscript) significant uncertainty concerning the severity and mortality of COVID-19. Furthermore, as of October 2020, no vaccine has completed the trials necessary for approving widespread use [36]. Therefore, many countries are enforcing nonpharmaceutical countermeasures [28, 29], e.g., social distancing, increased public hygiene, travel restrictions, self-isolation (quarantine), and population-wide mass testing. However, enforcing these countermeasures for long periods of time has severe economic and social consequences, both at the national and global scale [42]. Therefore, there is a need for identifying economic strategies for simultaneously relaxing the countermeasures and containing the pandemic. Model-based decision support systems can be used for exactly this purpose.

Predictive models are used to assist decision makers in identifying and evaluating candidate strategies; see, e.g., [2]. In particular, given a dynamical model of the spread of SARS-CoV-2, economically optimal (open-loop) mitigation strategies can be identified by solving optimal control problems (OCPs) over several months or even years.

*Received by the editors November 2, 2020; accepted for publication (in revised form) September 14, 2021; published electronically January 13, 2022.
https://doi.org/10.1137/20M1377783
Funding: This work was funded by the Federal Ministry of Education and Research (BMBF, grants 05M18EVA and 05M18SIA). The sixth author gratefully acknowledges funding by the German Research Foundation (DFG, grant WO 2056/6-1, project 406141926).

†Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg 39108, Germany (grundel@mpi-magdeburg.mpg.de, ritschel@mpi-magdeburg.mpg.de).
‡Institute for Mathematics, Technische Universität Ilmenau, Ilmenau 98693, Germany (stefan.heyder@tu-ilmenau.de, thomas.hotz@tu-ilmenau.de, philipp.sauerteig@tu-ilmenau.de, karl.worthmann@tu-ilmenau.de).

A key advantage of this approach is that it can directly account for constraints, e.g., related to the capacity of public healthcare systems. However, given the uncertainty surrounding SARS-CoV-2 and COVID-19, it is advisable to implement the optimal mitigation strategies in closed-loop, i.e., to repeatedly update the strategies when new data becomes available. This is referred to as model predictive control (MPC) [46], also known as receding-horizon control, and is an established method for advanced process control [45]. The predictive capabilities of the underlying model are crucial for the efficacy of the resulting mitigation strategy, and a common challenge is to identify suitable model parameters.

Epidemics are often modelled using deterministic *compartmental models* [27], e.g., the classical susceptible, infected, and recovered (SIR) model, where individuals are either susceptible, infectious, or removed, or the susceptible-exposed-infectious-removed (SEIR) model which, additionally, takes the incubation time into account. Optimal control of compartmental models was already an active research topic before the SARS-CoV-2 pandemic; see [50] for a review. In particular, optimal control of SIR models has been considered, e.g., for arbitrary social interaction models [3] and to identify time-optimal mitigation strategies [5, 25]. Optimal control of more complex models has also been considered. For instance, Fischer, Chudej, and Pesch [19] consider optimal control of a model with two species, Bussell et al. [8] demonstrate the importance of closed-loop mitigation strategies (i.e., of incorporating feedback), and Watkins, Nowzari, and Pappas [56] considered MPC of stochastic compartmental models. In [18] the authors determined control strategies to maintain hard infection caps in a disease-vector model based on the theory of barriers. This approach, however, exploits the low dimensionality of the model. Application of these techniques to complex compartmental models is not straightforward. Furthermore, using reduced-order models instead makes it challenging to actually satisfy hard constraints.

In response to the SARS-CoV-2 pandemic, many researchers have presented optimal control strategies, for instance based on Pontryagin's maximum principle (e.g. [31, 44, 59]). These strategies typically involve extended SIR or SEIR models, nonpharmaceutical countermeasures (often social distancing), and minimization of the number of infected as well as the economic cost of the countermeasures (and often other quantities as well, e.g., the number of deaths). However, hard constraints, for instance related to health care or testing capacities, are hardly taken into account. In the following, we highlight some of the key developments in decision support for SARS-CoV-2 mitigation based on optimal control. Gondim and Machado [21] use a model with three age groups to compute optimal quarantine strategies (for susceptible individuals), which minimize the number of infected and the cost of quarantining. Bonnans and Gianatti [6] compute social distancing strategies based on a model with a continuous age structure. Here, the strategies minimize a combination of the number of deaths, the peak number of hospitalized, and the cost of social distancing. Similarly, Richard et al. [47] present optimal social distancing strategies based on a model with a continuous age and infection duration structure, which minimize the number of deaths and the cost of social distancing. Morato et al. [38] compute on-off (also called bang-bang) social distancing strategies which minimize the number of symptomatic infectious people and the duration of the social distancing policies subject to constraints on intensive care unit (ICU) occupancy. They use extended SIR models. In [10, 48] the authors use MPC to compute social distancing and travel restriction strategies for an extended multi-region SIR model, minimizing the cost of the countermeasures and preventing an overload on the hospitals. Köhler et al. [34] use MPC to minimize the number of fatalities caused by COVID-19, subject to constraints on

the economic cost of social distancing. They take a modified SIDARTHE model [20] as basis and use interval arithmetic in the MPC to propagate model uncertainties. Finally, Tsay et al. [53] use MPC to minimize the cost of social distancing and testing, subject to an upper bound on the peak number of infectious people who have tested positive. They use the unscented Kalman filter to estimate the noisy state variables of an extended SEIR model.

In this work, we address some of the key questions that decision makers involved in the mitigation of the SARS-CoV-2 pandemic are facing:

1. Is mass testing alone sufficient to avoid overloading of ICUs?
2. If not, how much social distancing is then required?
3. How much can social distancing measures be reduced by targeting specific age groups?
4. How do strategies obtained by short- and long-term planning differ?
5. What are the benefits of increasing the daily testing or ICU capacity?

In this paper, the limited ICU capacity is considered as an example for constraints imposed by the health care system or political considerations. Of course, different constraints such as limited personnel for contact tracing could be incorporated as well.

We address the above questions by proposing a novel compartmental model and using optimal control as well as MPC to compute open- and closed-loop social distancing and testing strategies. The model contains three age groups, and it accounts for several of the key characteristics of COVID-19, i.e., (1) the incubation time, (2) different levels of symptom severity depending on age, (3) delay of testing results (and the following self-isolation), and (4) delay of hospitalization.

Furthermore, we choose values of the epidemiological model parameters based on the best of our knowledge to ensure that our numerical results match reality. For concreteness, we use the COVID-19 outbreak in Germany to determine parameters depending on demographics and the health care system. However, we expect our conclusions to carry over to outbreaks in other developed countries as well.

The remainder of this paper is structured as follows. In section 2, we describe the novel compartmental model of the SARS-CoV-2 outbreak in Germany, and in section 3, we motivate our choice of model parameters. In section 4, we demonstrate that optimal control can be used as a decision support tool based on the proposed model, and we conclude this paper in section 5.

2. Modelling pandemics. In this section, we propose a dynamical model tailored to COVID-19. The aim is to be able to evaluate the effect of population-wide mass testing (in combination with quarantine) and social distancing measures on the development of the pandemic. To this end, we extend the well-known SIR model.

2.1. Interpretation of deterministic compartmental models. We start with an illustration of the connection between (1) infectious disease models based on randomly acting individual agents and (2) their approximation by ordinary differential equation compartmental models. This exposition will highlight the interpretation and conversion of parameters when moving from a random to a deterministic model. Our final model will be based on parameter values from the epidemiological literature and not be obtained by parameter identification. As such parameters will often describe the behavior of individuals, we deem a brief explanation on how to convert these to the ODE setting to be useful to the reader. The actual parameter values used in this model are described in section 3. For simplicity, we consider the classical SIR model in this subsection. However, the connection, especially the interpretation of parameters,

is similar for more complex models such as the one described in section 2.2. For a more in-depth analysis of these models we refer the reader to [14, Chapter 3] which contains a derivation similar to ours.

Consider a population of n_{pop} individuals or agents each being either *susceptible*, *infectious*, or *removed*. At time $t \in [0, \infty)$ denote the (random) set of susceptibles by $\mathcal{S}_t$, the set of infectious by $\mathcal{I}_t$, and the set of removed by $\mathcal{R}_t$. Time is modelled continuously and measured in days.

We assume a homogeneous population with contacts between agents a and b following a Poisson process with intensity λ which does not depend on the agents considered. Infections occur randomly upon contact with a fixed probability α if one of the agents is susceptible and the other infectious. Thus, potentially infectious contacts also follow a Poisson process with respective intensity $\alpha\lambda$. Similarly, we model other events, in this simple model only recoveries, to occur according to a Poisson process. This implies that the time an agent spends in the infectious compartment is exponentially distributed with rate η, say, which we also assume to be the same for each agent (see [41] for models where these quantities follow other distributions).

We denote by $S(t) = \mathbf{E}\frac{|\mathcal{S}_t|}{n_{\text{pop}}}, I(t) = \mathbf{E}\frac{|\mathcal{I}_t|}{n_{\text{pop}}}$, and $R(t) = \mathbf{E}\frac{|\mathcal{R}_t|}{n_{\text{pop}}}$ the expected share of the population which are susceptible, infectious, and removed, respectively. Since for large n_{pop} the change of $\frac{|\mathcal{S}_t|}{n_{\text{pop}}}$ over a short time interval can, due to the law of large numbers, be well approximated by its expectation, $S(t)$ will provide a sufficient approximation of $\frac{|\mathcal{S}_t|}{n_{\text{pop}}}$ over the finite time horizon considered for a country the size of Germany. By the same argument, $I(t)$ and $R(t)$ approximate $\frac{|\mathcal{I}_t|}{n_{\text{pop}}}$ and $\frac{|\mathcal{R}_t|}{n_{\text{pop}}}$, respectively, sufficiently well.

If a is susceptible, he will transition to the infectious compartment upon having an infectious contact. At a fixed time t with $a \in \mathcal{S}_t$, there are two possible sources of infection for a: either some $b \in \mathcal{I}_t$ which is already infectious or some $c \in \mathcal{S}_t$ which will become infectious himself at some later time.

To determine the probability that b infects a in the time frame $(t, t + \Delta t]$, we analyze two competing events: The first is an infectious contact between a and b, and the second is b's recovery from the infectious state. Both events happen independently of one another with exponentially distributed time of occurrence, the first with rate $\alpha\lambda$ and the second with rate η. Thus the first time of occurrence of one of these is again exponentially distributed with rate $\alpha\lambda + \eta$ and the probability that the first occurrence is an infectious contact is $\frac{\alpha\lambda}{\alpha\lambda+\eta}$. In total

$$\begin{aligned}\mathbf{P}\big(b \text{ infects } a \text{ in } (t, t+\Delta t] \;\big|\; a \in \mathcal{S}_t, b \in \mathcal{I}_t\big) &= \big(1 - \exp(-(\alpha\lambda + \eta)\Delta t)\big)\frac{\alpha\lambda}{\alpha\lambda + \eta} \\ &= \alpha\lambda\Delta t + o(\Delta t).\end{aligned}$$

For $c \in \mathcal{S}_t$ to infect a in $(t, t + \Delta t]$, c has to become infectious himself before he in turn can infect a. This happens only with probability $o((\Delta t)^2)$ and can, thus, be neglected in the following calculations. In total a is moved out of the susceptible compartment with probability

$$\begin{aligned}\mathbf{P}\Big(a \notin \mathcal{S}_{t+\Delta t} \;\Big|\; \mathcal{S}_t, \mathcal{I}_t\Big) &= 1 - \prod_{b\in\mathcal{I}_t}\Big(1 - \alpha\lambda\Delta t - o(\Delta t)\Big)\prod_{c\in\mathcal{S}_t}\Big(1 - o\big((\Delta t)^2\big)\Big) \\ &= -\alpha\lambda\Delta t|\mathcal{I}_t| + o(\Delta t).\end{aligned}$$

Approximating $\frac{|\mathcal{S}_t|}{n_{\text{pop}}}$ and $\frac{|\mathcal{I}_t|}{n_{\text{pop}}}$ by $S(t)$ and $I(t)$ using the law of total expectation

yields

$$\frac{S(t+\Delta t)-S(t)}{\Delta t} \approx \frac{1}{n_{\text{pop}}\Delta t}\sum_{b\in\mathcal{S}_t} -\alpha\lambda\Delta t|\mathcal{I}_t| + o(\Delta t) = -\frac{\alpha\lambda}{n_{\text{pop}}}|\mathcal{S}_t||\mathcal{I}_t| + o(1)$$
$$\approx -\alpha\lambda n_{\text{pop}} S(t)I(t) + o(1).$$

As we assume the time from infection to removal to be exponentially distributed with rate η, a similar but more straightforward calculation reveals

$$\frac{R(t+\Delta t)-R(t)}{\Delta t} = \eta I(t) + o(1),$$

where η^{-1} is the mean stay of a single agent in the infectious compartment. We now set $\beta = n_{\text{pop}}\alpha\lambda$, which can be interpreted in this model as the daily amount of (potentially) infectious contacts a single agent has.

Since $S(t) + I(t) + R(t) = 1$ for all t, we obtain the classical SIR compartmental model:

$$\begin{aligned} \dot{S}(t) &= -\beta S(t)I(t), \\ \dot{I}(t) &= \beta S(t)I(t) - \eta I(t), \\ \dot{R}(t) &= \eta I(t). \end{aligned} \tag{2.1}$$

To determine suitable parameter values for β and η in this model, we reiterate that these are best thought of in the probabilistic setting. For the coefficients of the linear terms on the right-hand side, the interpretation is straightforward: it is the rate of the exponential distribution underlying the time until an agent leaves the compartment. Its inverse is the mean stay in this compartment.

For coefficients of interaction (product) terms, here β, the interpretation is the rate at which an agent in the first compartment causes other agents to leave the second compartment. In our setting, this is the daily amount of infections one infectious agent causes which can readily be seen from the definition of β. See section 3 for a more detailed discussion of the parameter values we use in our model. As mentioned above, these interpretations for the parameters carry over in a straightforward manner to more sophisticated models such as the one considered in the following.

2.2. A compartmental model for COVID-19. The SIR model provides a good starting point to study the dynamics of pandemics. However, due to its simple structure it is not suited to model the COVID-19 pandemic adequately. In particular it does not include hospitalization, age-specific disease progressions, and interventions. Therefore, we extend the SIR model in three ways.

1. We introduce eight additional compartments. In detail, we take into account that people can be infected, but not yet infectious. We call them *exposed* (or latent) and denote the compartment by E; see also [27]. Moreover, we split the infectious compartment into three depending on how the course of the infection will be. We distinguish between *severe* cases I^S that are going to need intensive care, i.e., they will move to H^{ICU} at some point in time; *mild* cases I^M that are going to visit a physician and hence be quarantined, i.e., removed; and *asymptomatic* cases I^A that might recover without being detected. Furthermore, we incorporate the possibility of being *tested* but not yet detected by introducing the compartments T^S (severe) and T^O (other). We assume that the patients with severe cases will visit a physician at some

point before being sent to an ICU. To this end, we introduce P as a pre-ICU compartment which comprises isolated patients at home or on a regular hospital ward. Moreover, we split the compartment of removed people into *known* and *unknown* cases $R = R^U + R^K$.

2. Each compartment is further divided into n_{g} groups, $n_{\mathrm{g}} \in \mathbb{N}$, depending on the age of a subject in order to study how these groups affect each other.
3. Social distancing and hygiene measures affect the contact rate as well as the transmission probability. Therefore, β can be used as a time-dependent control input $\beta(t)$.

The resulting susceptible-exposed-infectious-tested-prehospitalized-hospitalized-removed (SEITPHR) model reads as

$$\dot{S}_i(t) = -\sum_{j=1}^{n_{\mathrm{g}}} \beta_{ij}(t) S_i(t) \left[I_j^S(t) + I_j^M(t) + I_j^A(t) + T_j^S(t) + T_j^O(t) \right], \tag{2.2a}$$

$$\dot{E}_i(t) = \sum_{j=1}^{n_{\mathrm{g}}} \beta_{ij}(t) S_i(t) \left[I_j^S(t) + I_j^M(t) + I_j^A(t) + T_j^S(t) + T_j^O(t) \right] - \gamma E_i(t), \tag{2.2b}$$

$$\dot{I}_i^S(t) = \pi_i^S \gamma E_i(t) - (\eta^S + \theta_i(t)) I_i^S(t), \tag{2.2c}$$

$$\dot{I}_i^M(t) = \pi_i^M \gamma E_i(t) - (\eta^M + \theta_i(t)) I_i^M(t), \tag{2.2d}$$

$$\dot{I}_i^A(t) = \pi_i^A \gamma E_i(t) - (\eta^A + \theta_i(t)) I_i^A(t), \tag{2.2e}$$

$$\dot{T}_i^S(t) = \theta_i(t) I_i^S(t) - \tau^S T_i^S(t), \tag{2.2f}$$

$$\dot{T}_i^O(t) = \theta_i(t) \left[I_i^M(t) + I_i^A(t) \right] - \tau^O T_i^O(t), \tag{2.2g}$$

$$\dot{P}_i(t) = \eta^S I_i^S(t) + \tau^S T_i^S(t) - \rho P_i(t), \tag{2.2h}$$

$$\dot{H}_i^{\mathrm{ICU}}(t) = \rho P_i(t) - \sigma H_i^{\mathrm{ICU}}(t), \tag{2.2i}$$

$$\dot{R}_i^K(t) = \eta^M I_i^M(t) + \tau^O T_i^O(t) + \sigma H_i^{\mathrm{ICU}}(t), \tag{2.2j}$$

$$\dot{R}_i^U(t) = \eta^A I_i^A(t), \tag{2.2k}$$

where the subscript $i \in \{1, 2, \ldots, n_{\mathrm{g}}\}$ denotes the age group in ascending order. We enforce $\sum_{i=1}^{n_{\mathrm{g}}} N_i = 1$, where N_i denotes the relative size of age group i. We assume a mean incubation time γ^{-1} independent of both the course of infection and the age of the patient. However, depending on the age, there are different probabilities π_i^S, π_i^M, and π_i^A for the three courses of infection, where $\pi_i^S + \pi_i^M + \pi_i^A = 1$ for all i. Similar to the SIR model (2.1), the parameters denoted by η correspond to people being removed from the system, i.e., η^S and η^M denote those who visit a physician and, therefore, are put into quarantine immediately, while η^A represents unreported recovery. We denote the total number of susceptibles and unreported cases by $U_i = S_i + E_i + I_i^S + I_i^M + I_i^A + R_i^U$. The control input $\theta_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ describes the rate of those being tested per day, where tests are distributed uniformly at random among all individuals in U_i. In addition, symptomatic cases who visit physicians are assumed to be tested as well. Therefore, the total number of tests at time $t \geq 0$ is given by

$$T^{\mathrm{tot}}(t) = n_{\mathrm{pop}} \cdot \left(\sum_{i=1}^{n_{\mathrm{g}}} \theta_i(t) U_i(t) + \eta^S I_i^S(t) + \eta^M I_i^M(t) \right).$$

Note that testing does not affect the state of noninfectious subjects. Parameters τ^S and τ^O denote the rate from being tested positive to being detected, and hence being

put into quarantine. Furthermore, ρ is the rate from prehospital quarantine to hospitalization and σ from hospitalization to being reportedly removed, i.e., σ incorporates both mortality and recovery rate of hospitalized patients. The basic structure of the SEITPHR model (2.2) is depicted in Figure 1.

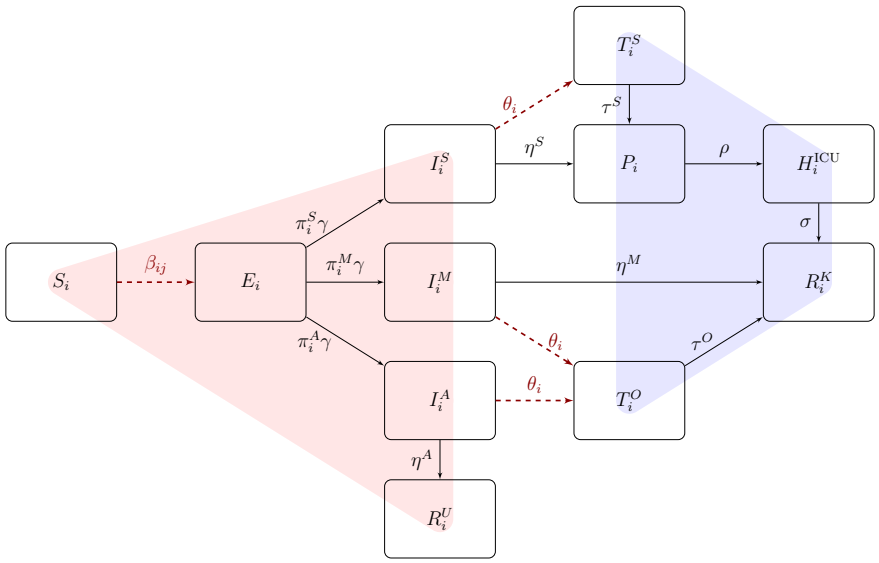

FIG. 1. *Flow of the SEITPHR model for one age group. The controls are indicated with dashed red edges. Unreported compartments are highlighted by the left red triangle, while tested and detected compartments are highlighted by the right blue trapezium. (Color is available online only.)*

The purpose of our model is to make qualitative statements on how countermeasures affect the spread of COVID-19 and how these countermeasures should be coordinated. Therefore, we make some simplifying assumptions about the nonpharmaceutical countermeasures and the underlying sociodynamics. Specifically, we do not attempt to quantify the effects of individual nonpharmaceutical countermeasures, such as masks and travel restrictions, on the overall transmission rate; see [30] for a discussion on how such effects can be quantified. Furthermore, we assume that political measures such as contact restrictions become active immediately after they are announced, i.e., we disregard delays; see [38] for a discussion of such delays. The latter assumption reflects the situation in Germany well since the population adapts quickly to changes in the contact restrictions. For instance, an emergency-break policy (*Bundesnotbremse*) was established in April 2021, where restrictions were enforced/lifted districtwise based on the 7-day-incidence value [1]. See [37, 12, 10] for work involving spatially differentiated contact restrictions.

For a concise notation we stack the state vectors into $x = (x_1, \dots, x_{n_g})$ and the controls into $u = (\beta, \theta)$, where $x_i = (S_i, E_i, I_i^S, I_i^M, I_i^A, T_i^S, T_i^O, P_i, H_i^{\mathrm{ICU}}, R_i^U, R_i^K)$, $\beta = (\beta_{ij})_{i,j=1}^{n_g}$ with $\beta_{ij} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ and $\theta = (\theta_1, \dots, \theta_{n_g})$. Similarly, we denote $\pi \in \mathbb{R}^{3n_g}$, $\tau \in \mathbb{R}^2$, and $\eta \in \mathbb{R}^3$. Thus, we write system (2.2) as

$$\dot{x}(t) = f(x(t), u(t), p), \tag{2.3}$$

where $p = (\pi, \eta, \tau, \rho, \sigma, \gamma) \in \mathbb{R}^{3n_g+8}$ collects all parameters. Furthermore, we intro-

TABLE 1
Overview of the parameter values used and their sources.

Parameter	Description	$i=1$	$i=2$	$i=3$	References
	Age-dependent parameters				
N_i	fraction of population	0.14	0.58	0.28	[51]
β^0_{i1}	base contact rate age group 1	0.46	0.48	0.12	[39, 43]
β^0_{i2}	base contact rate age group 2	0.48	0.63	0.29	[39, 43]
β^0_{i3}	base contact rate age group 3	0.12	0.29	0.18	[39, 43]
π_i^S	fraction severe cases (in %)	0.53	0.31	3.02	[49]
π_i^M	fraction mild cases (in %)	12.11	22.01	25.12	[49]
π_i^A	fraction asymptomatic cases (in %)	87.37	77.68	71.86	[49]
	Age-independent parameters				
γ	incubation time		0.19		[35]
η^S	removal rate for severe case		0.25		[26, 57]
η^M	removal rate for mild case		0.25		[26, 57]
η^A	removal rate for asymptomatic case		0.17		[26, 57]
τ^S	removal rate for tested severe cases		0.75		—
τ^0	removal rate for tested nonsevere case		0.92		—
ρ	ICU admittance rate		0.09		[17]
σ	ICU removal rate		0.1		[17]
$H^{\mathrm{ICU}}_{\max}$	max. ICU capacity		$10,000$		[16]
$T^{\max}$	max. daily test capacity		$\frac{1,200,000}{7}$		—

duce the initial condition $x(0) = x^0$ for some $x^0 \in \Omega$, where

$$\Omega = \left\{ x \in \mathbb{R}^{11\cdot n_{\mathrm{g}}}_{\geq 0} \;\middle|\; \sum_{j=1}^{11\cdot n_{\mathrm{g}}} x_j = 1 \right\}$$

denotes the set of possible states. Note that Ω is forward invariant under (2.3), i.e., if $x^0 \in \Omega$, then $x(t) \in \Omega$ for all $t \geq 0$.

3. Parameters. We selected the parameter values based on a literature research on the epidemiology of SARS-CoV-2. For parameter identification, which is outside the scope of this paper, we refer the reader to [11]. Before we present our choices for the parameter values of model (2.2), an overview of which can be found in Table 1, let us reiterate that some of the parameters of our model depend on age. We indicate this dependence by an appropriate index which we drop if the parameter is constant across age groups. For example, π_i^S is the age-dependent probability of having a severe course of disease while we assume η^A, the rate with which asymptomatic cases recover, to be age-independent.

$\boldsymbol{N_i}$. We use data on the population size of Germany at the end of 2019 from the GENESIS-Online Database of the Federal Statistical Office (DESTATIS) [51]. The first age group consists of individuals younger than 15 years, the second of those older than 15 but younger than 60 years, and the last comprises all individuals older than 60 years. These groupings result in proportions $N_1 = 0.14, N_2 = 0.58$, and $N_3 = 0.28$.

$\boldsymbol{\beta^0_{ij}}$. The rate at which an infected agent in compartment I_j infects susceptibles in compartment S_i depends on the contact structure of a population as well as the probability that a contact between a susceptible and infectious agent leads to a transmission of the virus. We base our contact process on data from the POLYMOD study

on daily contacts in several European countries [39]. From this data we calculate a contact matrix $C = (c_{ij})$ whose (i,j)th entry is the mean amount of contacts an individual in age group i has with age group j; here we only consider those contacts labeled as *physical*, since those are more likely to lead to viral transmission.

Let us denote by β_{ij}^0 the rate at which a single infectious agent from age group j infects susceptible agents from age group i if no countermeasures, such as social distancing, are in place. We model β_{ij}^0 to be proportional to c_{ij}, and let α be the corresponding proportionality constant. Consider a single infectious agent who is introduced into the otherwise completely susceptible population. Such a population is often called a *virgin* population, for which we also assume that no countermeasures are in place. The mean amount of secondary cases this agent causes is the basic reproduction number $\mathcal{R}_0$, which we approximate by

$$\mathcal{R}_0 \approx \sum_{i,j=1}^{n_g} N_i N_j \beta_{ij}^0 \frac{1}{\eta} = \sum_{i,j=1}^{n_g} N_i N_j c_{ij} \frac{\alpha}{\eta}. \tag{3.1}$$

Recall from the end of section 2.1 that in a simple SIR model a single infectious agent would cause on average β many new infections per time unit and stays infectious for a mean duration of η^{-1} time units, yielding a mean reproduction number of $\beta\eta^{-1}$. With this in mind we note that (3.1) is a straightforward extension of this to our age-stratified model: an infectious agent belonging to age group i would cause $N_j\beta_{ij}^0$ many cases in age group j per unit of time and the probability of this agent being in age group i is N_i.

There is a wide variety of estimates for $\mathcal{R}_0$ in the literature [43], with most estimates in the interval $[2, 3.5]$. We choose a value of $\mathcal{R}_0 = 2.5$ as early, higher estimates might be biased upwards due to imported and undetected cases. Fixing $\eta^{-1} = 6$ (see the discussion on η_A below) we calculate $\alpha = 5.79\%$ and in turn $\beta_{ij}^0 = \alpha c_{ij}$ from (3.1):

$$(\beta_{ij}^0)_{1\le i,j\le n_g} = \begin{pmatrix} 0.46 & 0.48 & 0.12 \\ 0.48 & 0.63 & 0.29 \\ 0.12 & 0.29 & 0.18 \end{pmatrix}. \tag{3.2}$$

$\boldsymbol{\gamma}$. The rate at which latent cases become infectious is the inverse of the mean incubation time. This parameter is modeled age-independent and chosen to be 0.19, which corresponds to a mean incubation time of 5.2 days [35].

$\boldsymbol{\pi_i^S, \pi_i^M, \pi_i^A}$. These parameters denote the proportion of individuals in age group i that have severe, mild, or asymptomatic course of disease. For Germany, the Robert Koch Institut (RKI) has published data on severity of disease progression for $12,178$ cases by age groups [49]. For our purposes we define a severe case to be a case that will eventually be admitted to intensive care, a mild case being one developing influenza-like symptoms, pneumonia, or being admitted to hospital for other reasons. All other cases we classify as asymptomatic.

Based on the data in [49] we choose $\pi_i = (\pi_i^S, \pi_i^M, \pi_i^A)$, the proportion of severe, mild, and asymptomatic cases in age group i, respectively, as

$$(\pi_1, \pi_2, \pi_3) = \frac{1}{100} \begin{pmatrix} 0.53 & 0.31 & 3.02 \\ 12.11 & 22.01 & 25.12 \\ 87.37 & 77.68 & 71.86 \end{pmatrix}.$$

Observe that the oldest age group is at highest risk with 3.02% of infected individuals admitted to ICU. Also the proportion of severe cases in the youngest age group is higher than in the middle age group. This might be explained by the fact that cases in the youngest age group are detected less frequently due to them being tested less, leading to overreporting of severe cases.

$\boldsymbol{\eta^S, \eta^M, \eta^A}$. These are the rates at which infectious individuals are removed from the infection process, if no mass-testing is implemented, i.e., if $\theta_i = 0$. For individuals with severe or mild course of disease we assume this to occur when they develop symptoms leading to self-isolation, quarantine prescribed by a physician, or to direct hospitalization.

One characteristic of COVID-19 is that even presymptomatic cases transmit SARS-CoV2 [26]. We assume the time from being infectious to symptom onset to be two days after which we add two more days which it takes before the infectee visits a physician. Thus we choose $\eta^S = 0.25$.

For mild progressions we assume the same mean duration from being infectious to symptomatic, though in this case individuals self-isolate, visit a physician, or receive a positive test result after a mean waiting time of again two more days; consequently, we also set $\eta^M = 0.25$.

For asymptomatic cases in I_i^A the only way to be removed from the infection process is by recovery from the infection. In [57] positive virus samples were found in patients' throats for up to eight days after symptom onset. Assuming a lower viral load for asymptomatic cases with only four days of potential infectiousness and adding the two days of presymptomatic transmission we chose $\eta^A = 0.17$, corresponding to a mean time of six days to recovery for asymptomatic cases.

$\boldsymbol{\tau^O, \tau^S}$. As we assume the testing related to the controls θ_i to be of a random nature, tested individuals are not yet removed from the infection process. Instead we assume positive test results to become available after a mean delay of two days. However, severe cases may visit a physician and thus go into immediate quarantine before receiving their test result. The latter transition occurs with rate η^S, and hence the faster transition occurs with rate $\tau^S = \frac{1}{2} + \eta^S = 0.75$.

Nonsevere cases that are tested, T_i^O, are removed if they recover naturally (with rate η^A), or receive a positive test result, or visit a physician, which leads to $\tau^O = \eta^M + \eta^A + \frac{1}{2} = 0.92$ for each age group.

$\boldsymbol{\rho}$. This parameter is the rate at which severe cases move from being in the pre-ICU state to the ICU. This includes time spent in quarantine at home as well as time spent in the hospital in normal care while being isolated.

In [17] the median time from symptom onset to being in intensive care for 50 patients was 9 days. As the median of an exponential distribution $\mathrm{Exp}(\zeta)$ is $\frac{\log 2}{\zeta}$ we choose a mean stay of $\frac{9}{\log 2} - 2 \approx 10.98 = \rho^{-1}$ days, accounting for the mean two days from symptom onset to the transition into the pre-ICU compartment.

$\boldsymbol{\sigma}$. This is the inverse of the mean time spent on intensive care until discharge or death. According to [17] patients with acute respiratory distress symptoms (ARDS) spent a median amount of 13 days in intensive care and patients without ARDS spent a median amount of two days in intensive care. Of the 50 patients considered in this study, 24 were afflicted with ARDS. Converting again between median and mean for the assumed exponential distribution yields a mean time of $\sigma^{-1} = \frac{13}{\log 2}\frac{24}{50} + \frac{2}{\log 2}\frac{26}{50} = 10.5$ days spent in intensive care.

$\boldsymbol{H^{\mathrm{ICU}}_{\mathrm{max}}}$. The DIVI-Intensivregister offers daily information on the amount of free intensive care hospital beds in Germany. On October 20, 2020 they reported a capacity

of 8,872 free beds with 879 actively treated COVID-19 patients [16]. We therefore round the maximal ICU-beds available for COVID-19 patients to 10,000.

$\boldsymbol{T^{\max}}$. In late August until the beginning of October the RKI conducted between 1 and 1.2 million weekly SARS-CoV-2 tests in Germany. This motivates our upper bound $T^{\max} = \frac{1{,}200{,}000}{7}$ of daily tests.

$\boldsymbol{x^0}$. We initialize our model at time $t = 0$ with entries of x^0 set to 0 except for those related to the susceptible, latent, and infectious compartments. Our choice of initial values is informed by the number of active cases reported by the RKI in late march assuming the proportion of underreporting to be 50%. We hence set the total number of infectious agents at $t = 0$ to 524 and the number of latent agents to 1672 distributed among the age groups according to N_i. As we explain in Remark 4.1 the solutions of the optimal control problems we consider in section 4 are robust against misspecification of the initial values.

Figure 2 demonstrates the simulation capabilities of our model. Here, the course of the pandemic is visualized if no countermeasures are implemented, i.e., no social distancing ($\beta(t) = \beta^0$) and no mass-testing ($\theta(t) = 0$). As expected, the pandemic

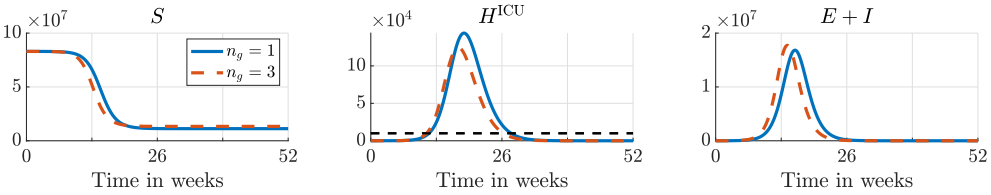

FIG. 2. *Evolution of the pandemic without countermeasures for one and three age groups over one year; the dashed black horizontal line in the middle figure marks* $H^{\mathrm{ICU}}_{\max}$.

evolves too fast to satisfy any reasonable cap on the number of required ICUs. In particular, the number H^{ICU} of required ICUs exceeds 100,000 whereas we noted above that in Germany only about 10,000 ICUs are available to treat COVID-19 patients. Therefore, countermeasures are indispensable to avoid an overload on the hospitals. Note that if we distinguish different age groups, the pandemic evolves faster, but less ICUs are required, as the pandemic spreads mostly in the less vulnerable, younger age groups. Similar observations, viz. herd immunity being achieved faster in heterogeneous populations in comparison to homogeneous ones, have already been made by [7].

4. Optimal testing and social distancing. In this section, we provide information on how to keep the epidemic manageable. To this end, we formulate suitable optimal control problems (OCPs) and solve them numerically. Since we neither take vaccines nor reinfections into account, we consider the epidemic to be over once herd immunity is achieved, i.e., a state where the introduction of new infectious agents does not lead to an outbreak. Therefore, our main goal is to reach herd immunity with as little social distancing as possible while maintaining strict limits on the ICU occupancy to avoid a breakdown of the healthcare system. We call a control $u = (\beta, \theta)$ of the system (2.3) *feasible* if $\beta_{ij}(t) \in [0, \beta^0_{ij}]$, $\theta_i(t) \geq 0$, $i, j \in \{1, \dots, n_{\mathrm{g}}\}$, and

$$n_{\mathrm{pop}} \cdot \sum_{i=1}^{n_{\mathrm{g}}} H_i^{\mathrm{ICU}}(t) \;\leq\; H^{\mathrm{ICU}}_{\max}$$

is satisfied for all t.

A natural stopping point for simulations is when the share of susceptibles has decreased enough to ensure herd immunity even when all countermeasures are lifted completely. The time-dependent effective reproduction number $\mathcal{R}(t)$, that is, the mean number of secondary cases a primary case will cause at time t, can be used to determine whether herd immunity has been reached: this will be the case if $\mathcal{R}(t)$ is less than 1. If there is only one infected compartment, as in a simple SIR model, the latter condition is equivalent to $\dot{I}(t) < 0$. If there is more than one infected compartment, as in model (2.2), [15, 54] have suggested computing $\mathcal{R}^{\mathrm{NGM}}$, the largest eigenvalue of the so-called next-generation matrix evaluated at the disease-free equilibrium, i.e., $S_i = N_i$ for all $i = 1, \dots, n_g$ and all other compartments equal to 0. To account for the depletion of susceptibles over time we define $\mathcal{R}^{\mathrm{NGM}}(t)$ to be the largest eigenvalue of the next-generation matrix evaluated at another disease-free equilibrium $S_i = S_i(t)$ and $R_i = R_i(t) = N_i - S_i(t)$ for $i = 1, \dots, n_g$. Here, $R_i(t)$ accounts for all removed people in age group i at time t; the ratio between detected and undetected does not effect the next-generation matrix. Then, $\mathcal{R}^{\mathrm{NGM}}(t)$ exhibits the same threshold property as $\mathcal{R}(t)$, that is, $\mathcal{R}^{\mathrm{NGM}}(t) < 1$ implies herd immunity for an epidemic in a population with a fraction of $S(t)$ susceptibles. Thus we use $\mathcal{R}^{\mathrm{NGM}}(t)$ to check whether our simulations have reached herd immunity. A time horizon of two years (104 weeks) turned out to be sufficient for all our simulations.

This section is structured as follows. In section 4.1, we verify the existence of a feasible testing strategy, i.e., without enforcing social distancing. Due to delays in testing, the existence of a solution is not trivial and depends on, e.g., the initial condition x^0 and the upper bound $H^{\mathrm{ICU}}_{\max}$. Next, we enforce an upper bound on the daily number of tests and investigate to what extent social distancing is required to ensure feasibility. We consider both age-homogeneous and age-differentiated social distancing in sections 4.2 and 4.3, respectively. In section 4.4, we present closed-loop optimal testing and social distancing strategies which only involve short-term predictions, and in section 4.5, we investigate the sensitivity of these strategies to the amount of available tests and the acceptable ICU occupancy. Finally, in section 4.6, we comment on the numerical solution of the OCPs.

Throughout our simulations, we assume the length of each control interval to be one week. This reflects the practical constraint that the government cannot change policies arbitrarily often but more realistically on a weekly basis.

4.1. Optimal testing strategy. Here, our goal is to maintain a hard cap on the number of required ICUs with as few tests as possible without enforcing social distancing, i.e., $\beta \equiv \beta^0$. To this end, we solve the OCP

$$\min_{\theta} \quad J_1(\theta) = \int_{t_0}^{t_f} T^{\mathrm{tot}}(t)\,\mathrm{d}t \tag{4.1a}$$

$$\text{s.t.} \quad n_{\mathrm{pop}} \cdot \sum_{i=1}^{n_{\mathrm{g}}} H_i^{\mathrm{ICU}}(t) \le H^{\mathrm{ICU}}_{\max} \quad \forall\, t \in [t_0, t_f], \tag{4.1b}$$

$$\dot{x}(t) = f(x(t), u(t), p) \quad \forall\, t \in [t_0, t_f], \quad x(0) = x^0, \tag{4.1c}$$

$$\beta(t) = \beta^0 \quad \forall\, t \in [t_0, t_f], \tag{4.1d}$$

$$\theta_i(t) \ge 0 \quad \forall\, t \in [t_0, t_f],\ i \in \{1, 2, \dots, n_{\mathrm{g}}\}. \tag{4.1e}$$

The objective function penalizes the total number of tests over the entire time horizon $[t_0, t_f]$ with $t_f > t_0 \ge 0$. The equality constraint (4.1c) captures the system dynamics while the one-sided box constraints (4.1e) ensure that the testing rates

cannot be negative.

Figure 3 depicts the optimal controls as well as the total number of tests and the number of detected cases per day while Figure 4 shows the impact on the evolution of the epidemic. Here, we computed the effective reproduction number $\mathcal{R}^{\mathrm{NGM}}(t)$ at each time step, demonstrating that we reached herd immunity.

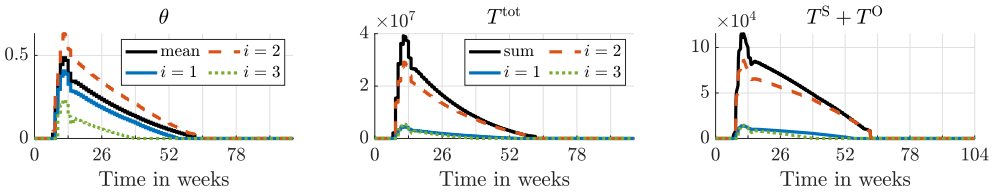

FIG. 3. *Optimal testing strategy for three age groups over two years. The middle-age group (most contacts) is prioritized, then children, and, finally, the elderly.*

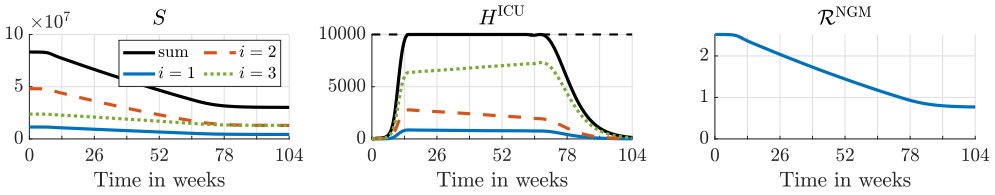

FIG. 4. *Evolution of the epidemic corresponding to the optimal testing strategy visualized in Figure 3 and reproduction number based on the next-generation matrix. The dashed black line in the middle plot depicts the maximal capacity of ICUs.*

We observe that there exists a testing strategy that ensures feasibility, which was not obvious from the outset because of the assumed delays. In particular, the bound (4.1b) is active once it's reached, i.e., $H^{\mathrm{ICU}} \equiv H^{\mathrm{ICU}}_{\max}$, and becomes inactive when the number of susceptible people falls below a certain threshold and $\mathcal{R}^{\mathrm{NGM}}(t) < 1$ indicating the onset of herd immunity.

Remark 4.1. The steady-state like behavior with respect to H^{ICU} suggests that problem (4.1) satisfies the so-called turnpike property [13, 32, 23]. Typically, turnpikes indicate the optimal operating state of a system. These are steady states at which the running costs are minimized. In our example, since we do not penalize the number of required ICUs, the best strategy is to stay at the upper bound while saving tests. Once the objective function value is zero the system leaves the state eventually. In particular, regardless of the initial value, the system is steered towards this optimal operating point. As a consequence, a rough estimation suffices as the initial guess for our simulations. A rigorous analysis of these turnpikes, however, is left for future research.

However, these results are only of theoretical interest, since this optimal testing strategy would be prohibitively expensive and might not even be implementable at all. For instance, regarding Figure 3, one observes that the mean testing rate reaches about 0.5, which corresponds to being tested every two days on average. Moreover, the total number of tests per day required for this approach is more than 12,000,000 on average (over 65 weeks), compared to $T^{\max} \approx 170,000$ daily tests, which are currently conducted in Germany. Note that, even with this enormous testing effort, the number

of detected cases, $T^{\mathrm{S}} + T^{\mathrm{O}}$, is rather small since the number of infectious individuals is small compared to the total population.

In conclusion, mass-testing alone currently does not suffice to maintain hospitalization caps in reality. These arguments support the government's decision to introduce additional measures like social distancing and hygiene concepts. However, cheap rapid test kits might change the situation favorably as they could be made widely available and self-administered while giving immediate test results.

In the following subsections, we enforce $T^{\max}$ as an upper bound on the amount of daily tests. Under this additional constraint we then determine the minimal amount of social distancing required to reach herd immunity. The success of such measures depends on the acceptance and thus compliance by the general population.

4.2. Optimal homogeneous social distancing. In a first step, we determine an optimal social distancing strategy by penalizing the deviation of β from β^0 equally over all age groups. This might increase acceptance in the general population due to the (perceived) fairness of such measures: everyone is treated equally and contacts are reduced by the same proportion for everyone. In reality such strategies may be hard to conceive as different measures affect the age groups differently, i.e., closing schools and nurseries affects those in the lowest age group the most while leaving the oldest age group unaffected. Nevertheless, a mixture of many different nonpharmaceutical measures may be able to achieve such a reduction in contacts.

We introduce a time-varying factor $\delta = \delta(t)$ describing the amount of social distancing that is implemented. Moreover, we choose to penalize the ℓ^2 deviation of this control input from $\delta = 1$ in the objective function in order to smooth the optimal control. For instance, penalizing the ℓ^1 deviation yields bang-bang controls, i.e., the optimal solution jumps back and forth between the two extremal options: no contact restrictions and lockdown (simulations not shown). See, e.g., [55] for the relation between ℓ^1 minimization and bang-bang optimal control. Therefore, we determine an optimal homogeneous social distancing policy by solving

$$\min_{\theta,\delta} \quad J_2(\theta,\delta) = \int_{t_0}^{t_f} \big(1-\delta(t)\big)^2 + \kappa \sum_{i=1}^{n_{\mathrm{g}}} \theta_i(t)\,\mathrm{d}t \tag{4.2a}$$

$$\text{s.t.} \quad n_{\mathrm{pop}} \cdot \sum_{i=1}^{n_{\mathrm{g}}} H_i^{\mathrm{ICU}}(t) \le H_{\max}^{\mathrm{ICU}} \quad \forall\, t \in [t_0, t_f], \tag{4.2b}$$

$$\beta_{ij}(t) = \delta(t)\beta_{ij}^0 \quad \forall\, i,j \in \{1,\dots,n_{\mathrm{g}}\}, \tag{4.2c}$$

$$\dot{x}(t) = f(x(t),u(t),p) \quad \forall\, t \in [t_0,t_f], \quad x(0) = x^0, \tag{4.2d}$$

$$T^{\mathrm{tot}}(t) \le T^{\max} \quad \forall\, t \in [t_0,t_f], \tag{4.2e}$$

$$\delta(t) \in [0,1] \quad \forall\, t \in [t_0,t_f], \tag{4.2f}$$

$$\theta_i(t) \ge 0 \quad \forall\, t \in [t_0,t_f],\ i \in \{1,2,\dots,n_{\mathrm{g}}\}. \tag{4.2g}$$

Note that we allow one to distribute the tests among the age groups by not fixing θ_i, but enforcing (4.2e) and (4.2g). Furthermore, we introduce a regularization term with weight $\kappa = 10^{-5}$; see section 4.6 for a detailed discussion. In contrast to OCP (4.1), we always find a feasible solution of OCP (4.2) if the epidemic has not yet evolved too far. More precisely, by choosing $\delta = 0$, which corresponds to a complete lockdown, we are (theoretically) able to stop the spread. Therefore, if the initial number of people with a severe course of infection is sufficiently low, the upper bound on the number of ICUs will not be violated.

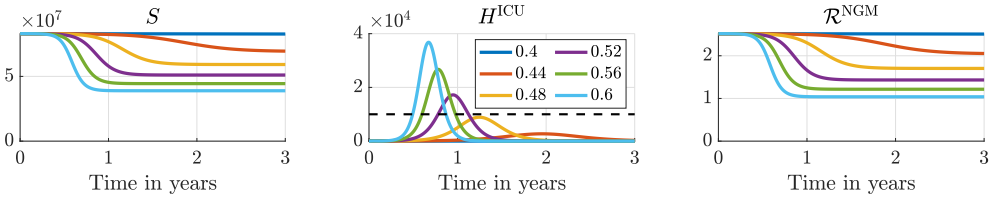

FIG. 5. *The impact of constant contact reduction rate* $\delta \in [0.4, 0.6]$ *on the speed of evolution of the epidemic and on the number of required ICUs.*

A highly fluctuating social distancing strategy may lead to low acceptance in the population, because people have to adapt to new rules every few weeks. Thus before we solve OCP (4.2) let us have a look at what happens if we consider a constant value for δ over time, i.e., a social distancing strategy without fluctuations. Figure 5 (left) shows that fewer contacts result in a longer time for the epidemic to abate on the one hand, but a lower number of total infections within the considered time horizon on the other hand. Moreover, Figure 5 (middle) visualizes that quite strict social distancing is needed in order to meet the ICU capacities. The maximal value of δ to stay feasible is 0.487, i.e., contacts needed to be more than halved over three years. Furthermore, once we lift the restrictions (see Figure 6) there might be another outbreak. In particular, the stronger the restrictions were in the beginning, the stronger the second outbreak will be. Therefore, it is essential to establish herd immunity before lifting all restrictions, and to adapt the policy over time.

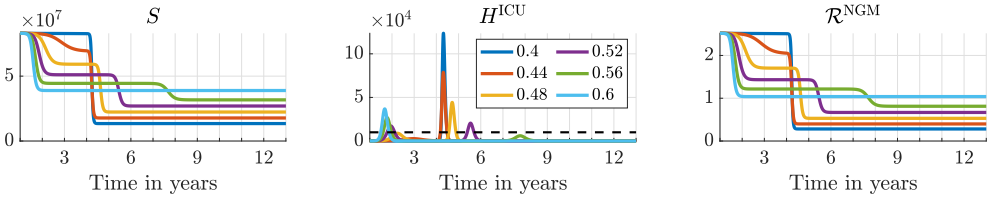

FIG. 6. *Solution for fixed* δ *for three years and complete lift of restrictions afterwards.*

A visualization of the optimal solution of OCP (4.2) can be found in Figures 7 and 8. As mentioned above, the bound on H^{ICU} is not violated. Since the weight κ is chosen sufficiently small, the upper bound on the total number of tests per day is active as long as the upper bound on δ is not. However, note that not all age groups are tested equally. More precisely, only the middle-aged group is tested at all. The reason is that this group is the largest ($N_2 > N_1 + N_3$) and has the highest contact rates (cf. (3.2)) and therefore contributes more to the spread of the epidemic than the other groups. Furthermore, we observe that the social distancing policy has to be quite strict in the beginning. In particular, $\min_t \delta(t) \approx 0.3$ which corresponds to a reduction of average contacts per person by 70%. However, this can be qualitatively compared to the measures taken in Germany starting in mid March 2020 when contacts were reduced by school and restaurant closures as well as other contact restrictions.

In conclusion, social distancing is an effective tool to keep the epidemic manageable. Comparing the results of (4.2) to the simulations with constant δ we see that a (partial) lockdown appears inevitable. However, our simulations suggest to let the epidemic evolve for a few weeks and then enforce a contact reduction down to approximately 30% for 2–4 weeks before slowly lifting the restrictions over the next

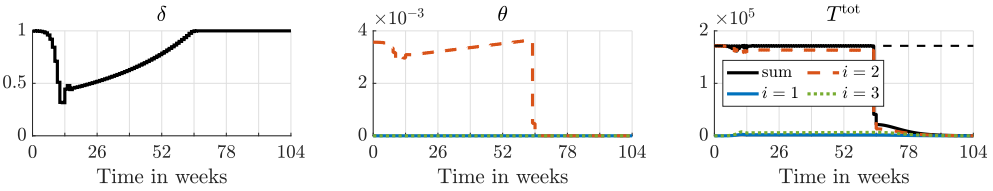

FIG. 7. *Optimal combination of testing and (homogeneous, time-varying) social distancing for three age groups over two years. The dashed black line in the second subplot depicts the upper bound on the total number of tests per day.*

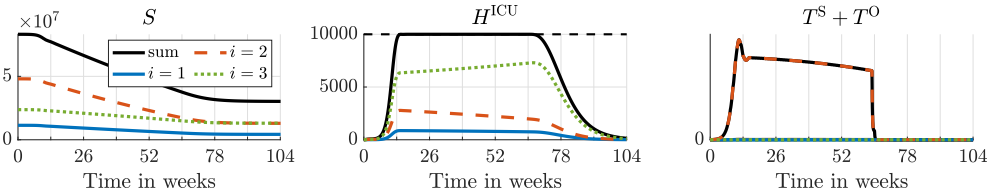

FIG. 8. *Evolution of the epidemic based on the controls visualized in Figure 7. The dashed black line in the plot of* H^{ICU} *depicts the upper bound on the number of available ICUs.*

12 months until herd immunity is achieved.

4.3. Age-dependent social distancing. The constraint that contacts are reduced by the same proportion for each age group is restrictive and it is plausible that more efficient solutions exist when contact reductions are distributed differently across age groups. One reason to consider such a strategy is that it may be more efficient at stopping the spread of the epidemic; as mentioned above the middle-age group is the driver of the epidemic while the oldest age group consists of the most vulnerable individuals. In any case, such an age-differentiated social distancing strategy needs to be accepted by the whole population to be successful.

Hence, we improve the social distancing policy computed above by allowing it to depend on age. Given the solution $(\theta^\star, \delta^\star)$ of (4.2), we solve the OCP

$$\text{(4.3a)} \qquad \min_{\theta,\beta} \quad J_3(\theta,\beta) = \int_{t_0}^{t_f} \sum_{i,j=1}^{n_{\text{g}}} N_i N_j \big(\beta_{ij}(t) - \beta_{ij}^0\big)^2 + \kappa\theta_i(t)\,\mathrm{d}t$$

$$\text{(4.3b)} \qquad \text{s.t.} \quad n_{\text{pop}} \cdot \sum_{i=1}^{n_{\text{g}}} H_i^{\text{ICU}}(t) \le H_{\max}^{\text{ICU}} \quad \forall\, t \in [t_0, t_f],$$

$$\text{(4.3c)} \qquad \dot{x}(t) = f\big(x(t), u(t), p\big) \quad \forall\, t \in [t_0, t_f], \quad x(0) = x^0,$$

$$\text{(4.3d)} \qquad T^{\text{tot}}(t) \le T^{\max} \quad \forall\, t \in [t_0, t_f],$$

$$\text{(4.3e)} \qquad \theta_i(t) \ge 0 \quad \forall\, t \in [t_0, t_f], \quad i \in \{1, 2, \ldots, n_{\text{g}}\},$$

$$\text{(4.3f)} \qquad \beta_{ij}(t) \in [\beta_{ij}^{\min}, \beta_{ij}^0] \quad \forall\, t \in [t_0, t_f], \quad i, j \in \{1, 2, \ldots, n_{\text{g}}\}.$$

Here, we use $\delta^\star$ to define $\beta_{ij}^{\min} = \min_t \delta^\star(t)\beta_{ij}^{\text{nom}}$, i.e., the lower bound on β in (4.3) is the worst case of (4.2). Therefore, no one is treated worse than when applying homogeneous social distancing. Note that $(\theta^\star, \beta)$ with $\beta = \delta^\star\beta^0$ is feasible for OCP (4.3). As in (4.2) we penalize testing as soon as $\beta(t) = \beta^0$ holds.

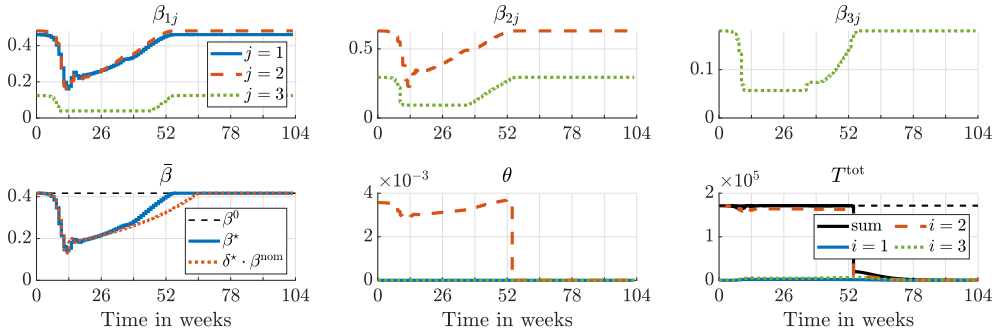

FIG. 9. *Optimal age-dependent social-distancing strategy for three age groups over two years.*

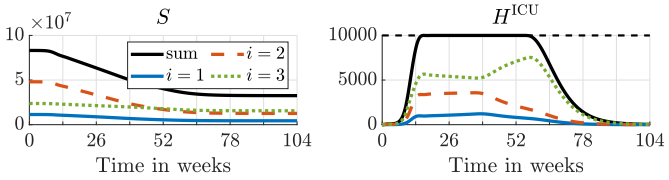

FIG. 10. *Evolution of the compartments associated with controls depicted in Figure* 9.

Results for OCP (4.3) can be found in Figures 9 and 10, where

$$\bar{\beta}(t) \;=\; \sum_{i,j=1}^{n_{\text{g}}} N_i N_j \beta_{ij}(t)$$

describes the average number of contacts per person and day in a heterogeneous population. The corresponding value for β^0 is $\bar{\beta}^0 = 0.4167$. We used again $\kappa = 10^{-5}$. This allows one to compare the solution $\beta_{ij}(t)$ with $\bar{\beta}^0 \delta^\star(t)$ obtained from (4.2). Similar to the solution of (4.2), the upper bound on testing is active most of the time, while essentially only the middle-aged group is tested. The social distancing measures are less restrictive than for (4.2) which makes compliance with the measures more likely. However, the measures could be perceived as unfair, since the contacts of the oldest age group are restricted most. Moreover, the contacts of the middle-age group are least restricted. Therefore, the working class would be allowed to go to work, which is beneficial from an economic point of view. The reason for this imbalance lies within the problem formulation. Since we try to minimize the total amount of social distancing, the age-differentiated strategy allows for a complete lift of restrictions almost three months earlier; see Figure 9 (bottom, left). On the other hand, the main constraint is to maintain the ICU capacity. Therefore, it is essential to reduce the contacts of the most vulnerable group. Note that people who are allowed to have more contacts need to get tested more frequently.

In conclusion, social distancing is crucial to avoid an overload on the hospitals. In addition, testing middle-aged people helps to reduce the required amount of social distancing. Furthermore, all presented strategies support a lockdown for a few weeks into the epidemic, which is followed by lifting the restrictions step by step until herd immunity sets in. Age-differentiated social distancing might be hard to argue for, but it helps to end the epidemic several months earlier and, therefore, support the economy.

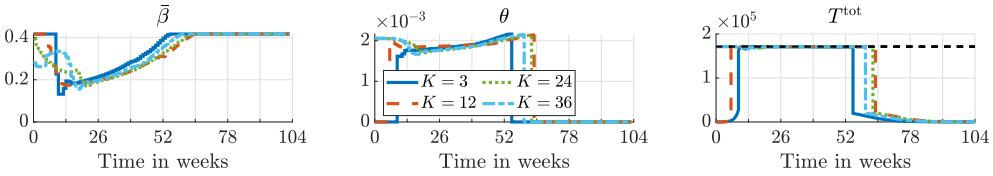

FIG. 11. *Optimal control for solving* (4.3) *in closed loop for varying prediction horizon length. For the sake of readability, we depicted average values of* θ *and sums of* T^{tot} *over the age groups.*

4.4. Short-term decision making. The control strategies derived in the previous subsections provide rough guidelines to contain the epidemic. However, from a decision maker's perspective, it will be hard to argue for policies taking effect in the far future. In particular, there are many uncertainties that might affect the performance of the control strategy over the time span of two years, and hence the control strategy needs to be adjusted over time. One major source of uncertainty lies in the model's parameters, which change over time; see, e.g., [33] for the duration of acute infection with the B.1.1.7 mutant. Therefore, solving any of the above mentioned OCPs over two years based on the data available in March 2020 would result in completely different and (most likely) not at all optimal solutions. Throughout the pandemic huge amounts of data are gathered, based on which the model parameters can be updated continuously and sometimes it can be beneficial to start countermeasures only after one has reduced the uncertainties in the parameters [52]. Model predictive control (MPC) provides a state-of-the-art methodology to make use of newly gathered information. The basic idea of MPC is to consecutively solve a series of OCPs over a smaller horizon of K control intervals rather than solving a single OCP over the whole horizon. Then, only the first part of the optimal control derived by solving such an auxiliary OCP is implemented. Next, the time window is shifted, and the procedure is repeated based on updated measurements. For a detailed introduction to MPC we refer the reader to [46]. Here, we tackle (4.3) via MPC since it is the most elaborated one; the earlier problems can be treated analogously. The MPC scheme for (4.3) is summarized in Algorithm 4.1.

Algorithm 4.1 MPC scheme for solving OCP (4.3)

Input: Prediction horizon length K, length of control interval Δt. Set time $t = t_0$.
Repeat:

1. Obtain current states $\hat{x} = x(t)$.
2. Determine optimal solution $u^\star : [t, t + K\Delta t) \to \mathbb{R}^9$ of (4.3) on $[t, t + K\Delta t)$ with $x(t_0) = \hat{x}$.
3. Implement $u^\star|_{[t,t+\Delta t)}$. Increment time $t \leftarrow t + \Delta t$.

Results based on varying prediction horizon lengths are shown in Figures 11 and 12. The basic structure of both the optimal control and the associated states is comparable to the open-loop solution presented in the previous subsection. Therefore, we stopped the simulations after one and a half years. The length of the prediction horizon affects mainly the optimal social distancing policy. In particular, the larger the prediction horizon, the less social distancing is needed in total. More precisely, for bigger K, we implement a slightly stricter lockdown but can start it later and relax it earlier. Furthermore, the larger K gets the closer the optimal solution is to the open-loop solution. In particular, the MPC solutions qualitatively resemble the

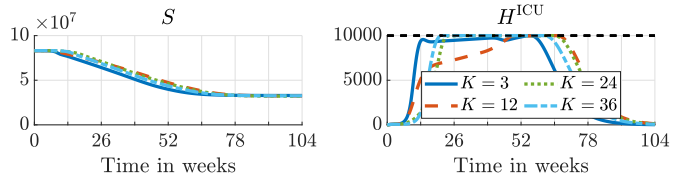

FIG. 12. *Evolution of the epidemic based on the controls depicted in Figure* 11. *For the sake of readability, only the sum over the age groups is visualized.*

open-loop solution: after an early lockdown, social distancing is slowly lifted.

For $K = 3$, the ICU capacity reaches its upper limit earlier due to the laissez-faire policy in the beginning. However, this constraint also becomes inactive earlier. For even shorter prediction horizons the ICU cap might be violated, i.e., recursive feasibility cannot be guaranteed (simulations not shown).

Moreover, in [22] we extend the model in order to incorporate vaccination. Here, the choice of the prediction horizon length crucially affects the optimal control. In particular, for short prediction horizon lengths K it is optimal to vaccinate the elderly first (immediate impact on ICU compartment), whereas with increasing K the people with highest contact rates should be prioritized.

4.5. Impact of upper bounds on number of tests and ICUs. So far, we assumed both the upper bounds on the number of tests per day and the number of ICUs to be fixed at our chosen values. In practice, these conditions may change: free ICU capacity might exhibit seasonal patterns and the number of possible tests per day depends on infrastructure and available personnel. In addition, varying the upper bounds is useful to illustrate the benefits of increased testing and higher ICU capacities. In this subsection, we investigate the impact of these parameters on the optimal social distancing policy numerically.

First, we study the effect $T^{\max}$ has on the social distancing by solving (4.2) via MPC; see Figure 13 (left). As pointed out in the previous subsection, the prediction horizon length affects the start and end time of measures as well as its peak (simulations not shown). In addition, increasing $T^{\max}$ by some factor $T^{\max}_{\text{fac}} \geq 1$ shifts the whole δ curve upwards, i.e., as expected, the more tests are available, the less social distancing is required. Furthermore, Figure 13 (left) visualizes the impact of $T^{\max}$ on the objective function value of (4.2).

Second, we investigate the impact of the number of ICUs on the optimal solution of (4.3). Results can be found in Figures 13 and 14. For the simulations in Figure 13 (middle and right) we used MPC with prediction horizon $K = 12$ weeks. Figure 13 (middle) clearly shows that the number of available ICUs directly affects the cost function value, while for a small value of $H^{\text{ICU}}_{\max}$, every additional ICU contributes, for large values, a saturation seems to take place. In particular, doubling the current number of available ICUs does help, but the benefit becomes negligible when increasing it further. These phenomena are almost unaffected by doubling or halving $T^{\max}$. However, when there are not enough ICUs, then the upper bound on T^{tot} is always active (see Figure 13 (right)), where T^{tot} is at its maximum value all the time. Moreover, an increase in the number of ICUs clearly leads to a reduction in the social distancing measures, as can be seen in Figure 14.

In summary, increasing test capacities and/or ICU capacities helps to reduce measures like social distancing. However, the impact of the number of available ICUs appears to be much stronger. Nonetheless the qualitative shape of the solutions over

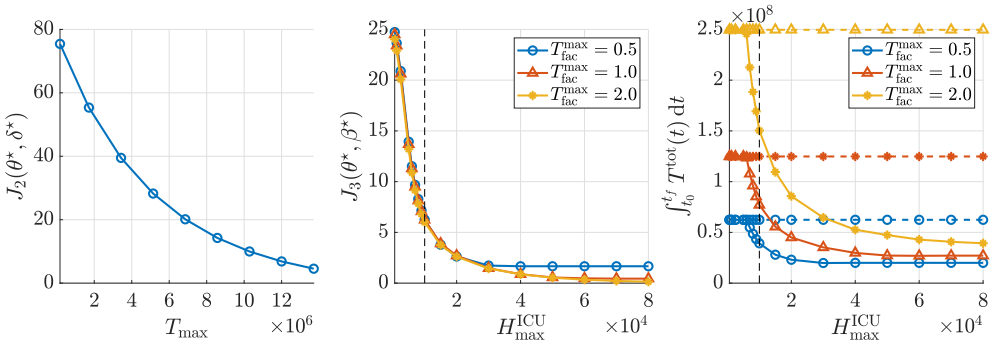

FIG. 13. *Impact of* $T^{\max}$ *on social distancing costs (left) and of* $H^{\mathrm{ICU}}_{\max}$ *on both social distancing costs (middle) and testing (right). In the last two subfigures the currently available number of ICUs in Germany is highlighted by a vertical dashed line. The dashed horizontal lines in the rightmost figure indicate the total testing capacities over the entire simulation horizon. Factor of modification of* $T^{\max}$ *denoted by* $T^{\max}_{\mathrm{fac}}$.

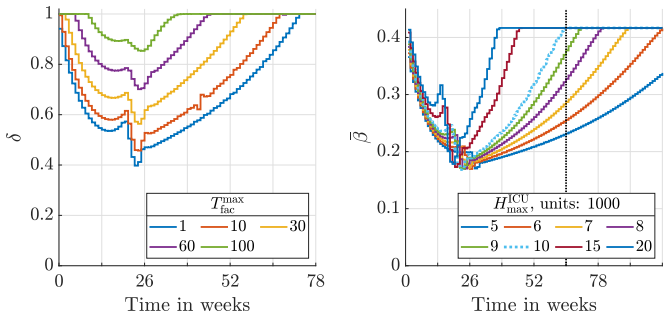

FIG. 14. *Impact of the available number* $H^{\mathrm{ICU}}_{\max}$ *of ICUs and the prediction horizon* K *on the average social distancing. The dotted cyan line refers to the number of currently available ICUs in Germany. The vertical dotted black line marks the end of social distancing measures for that setting. (Color is available online only.)*

time is not affected by varying these constraints.

4.6. Numerical solution of the optimal control problems. We use a direct single-shooting approach [4] to transcribe the OCPs (4.1)–(4.3) to nonlinear programs (NLPs), and as mentioned previously, we parametrize the control variables u using a zero-order-hold (ZOH) parametrization [46], i.e., we assume that they are piecewise constant. Furthermore, we enforce the nonlinear inequality constraints on the number of tests and the ICU occupancy, (4.1b), (4.2b), (4.2e), (4.3b), and (4.3d), pointwise on an equidistant time grid with intervals of 1 day. We solve the NLPs using the sequential quadratic programming (SQP) method [40] implemented in `fmincon` of MATLAB. The SQP method uses the first-order derivatives of the left-hand sides of the nonlinear inequality constraints with respect to the control variables. We compute these derivatives using a continuous forward method, i.e., by solving the sensitivity equations [24] associated with the model equations (2.2). We use `ode45` of MATLAB routine for approximating the solutions to both the model equations and the sensitivity equations.

As mentioned earlier in this section, we introduce a regularization term in the objective functions, J_2 and J_3, in the OCPs (4.2) and (4.3), where κ is the regular-

TABLE 2

The minimum, average, and maximum computation time for solving the OCP (4.3) *(with age-differentiated social distancing) in closed loop using different prediction horizons, K. The computation times correspond to the results presented in section* 4.4.

Prediction horizon, K	3 weeks	12 weeks	24 weeks	36 weeks
Maximum	2.2 s	12.7 s	40.1 s	60.8 s
Average	1.0 s	7.2 s	17.7 s	29.2 s
Minimum	0.5 s	2.2 s	5.1 s	8.9 s

ization parameter. The reason is that the COVID-19 outbreak can no longer sustain itself towards the end of the considered time interval. Therefore, in this final period of time, social distancing and mass testing are not required in order to satisfy the upper bound on the ICU occupancy. Consequently, without a regularization term, θ is not uniquely defined for this final period of time, and its value will be arbitrarily determined by the iterations of the SQP method used to solve the NLPs. The regularization term ensures that θ is zero during this final time period, and we choose κ such that the regularization term is negligible compared to the first terms in the objective functions, i.e., it is several orders of magnitude smaller.

On a standard desktop computer the computation times required to obtain the numerical solutions to the OCPs (4.1)–(4.3) presented in sections 4.1–4.3 are 70 minutes, 49 minutes, and 12 minutes, respectively. We use the solution to (4.2) to warmstart the solution of (4.3) which significantly reduces the computation time. Table 2 shows the computation time of solving (4.3) in closed loop for different prediction horizons, K. We warmstart the solution of the OCPs (except the first one) using the solution to the previous one. As is evident, the computation time is negligible compared to the prediction horizons and the size of the control intervals (one week). The computation time can possibly be improved by implementing the upper bounds on the number of tests and the ICU occupancy as soft constraints, e.g., using a barrier method or a penalty method [40]. In that case, using an adjoint method [9] to compute the gradient of the objective function is likely faster than using a forward method. However, these extensions are outside the scope of the present work.

5. Conclusions and outlook. In this paper, we demonstrated how mitigation of the COVID-19 epidemic can be achieved by a combination of age-stratified testing and social distancing measures while avoiding a breakdown of the healthcare system. We believe that our model with the chosen parameters reflects reality sufficiently well to provide qualitatively valid insight on how testing and social distancing can control the spread of SARS-CoV-2. To summarize our findings, we answer the five key questions formulated in the introduction.

(1) Is mass testing alone sufficient to avoid overloading of the ICUs? Assuming realistic testing capacities, mass testing alone does not suffice to avoid a breakdown of the healthcare system in Germany.

(2) If mass testing alone does not suffice, how much social distancing is then required? To this end, we designed optimal social distancing strategies with a focus on applicability and acceptance in the general population, i.e., strategies with slowly changing contact reductions. Our results show that one has to implement strict reductions early on and slowly relax them over the course of several months. In particular, the social distancing measures imitate the measures actually taken in Germany, but are lifted at a much slower pace.

(3) How much can social distancing measures be reduced by targeting specific age groups? Age-differentiated contact restrictions improve upon the aforementioned results as they yield qualitatively similar social distancing strategies and prioritize relaxing restrictions for the work-force and children at the cost of reducing contacts of the more vulnerable older population.

(4) How do strategies obtained by short- and long-term planning differ? We used MPC to model the process of real-life policy making where the planned countermeasures are reevaluated and updated at regular time intervals. Specifically, the plan is updated weekly by solving the optimal control problem over a shifted horizon. Our analysis shows that longer prediction horizons allow for faster lifting of the restrictions although long-term predictions may be infeasible in practice. While short-term planning of measures is unable to control the exponential growth of COVID-19 cases, medium-term planning leads to strategies that, qualitatively, do not differ from the optimal ones.

(5) What are the benefits of increasing the daily testing capacity or the ICU capacity? As expected, the number of available ICUs dictates how fast herd immunity can be reached and how much total social distancing is necessary. Increasing the testing capacity may yield a similar effect, but the required increase seems unrealistic.

However, we caution the reader against interpreting these results in a quantitative way, as our model has not been devised to produce precise predictions. Similarly, we want to stress that we do not provide concrete policies to implement, as the impact of particular countermeasures on β is not easily quantified.

In a follow-up work, which has been published in the meantime, we extended the model to account for vaccination [22].

Another possible extension concerns reinfections, which could be included in our model without difficulties if parameters are available to model them. As our model is based on ODEs, interaction effects such as contact tracing cannot be included. Agent-based (stochastic) models are able to handle these critical effects and could be seen as a natural extension of our (deterministic) compartmental model. To solve the resulting stochastic optimal control problems would then require more sophisticated techniques, however.

Acknowledgments. We thank Kurt Chudej (University of Bayreuth) for insights on modelling pandemics and Manuel Schaller (TU Ilmenau) for fruitful discussions on optimal control and the turnpike property.

REFERENCES

[1] *Bundesweite Notbremse*, https://www.bundesregierung.de/breg-de/suche/bundesweite-notbremse-1888982, 2021 (accessed 2021-05-08).

[2] M. S. Aronna, R. Guglielmi, and L. M. Moschen, *A model for COVID-19 with isolation, quarantine and testing as control measures*, Epidemics, 34 (2021), 100437, https://doi.org/10.1016/j.epidem.2021.100437.

[3] H. Behncke, *Optimal control of deterministic epidemics*, Optim. Contr. Appl. Met., 21 (2000), pp. 269–285, https://doi.org/10.1002/oca.678.

[4] T. Binder, L. Blank, H. G. Bock, R. Bulirsch, W. Dahmen, M. Diehl, T. Kronseder, W. Marquardt, J. P. Schlöder, and O. von Stryk, *Introduction to model based optimization of chemical processes on moving horizons*, in Online Optimization of Large Scale Systems, Springer, Berlin, Heidelberg, 2001, pp. 295–339, https://doi.org/10.1007/978-3-662-04331-8_18.

[5] L. Bolzoni, E. Bonacini, C. Soresina, and M. Groppi, *Time-optimal control strategies in SIR epidemic models*, Math. Biosci., 292 (2017), pp. 86–96, https://doi.org/10.1016/j.mbs.2017.07.011.

[6] J. F. Bonnans and J. Gianatti, *Optimal control techniques based on infection age for the study of the COVID-19 epidemic*, Math. Model. Nat. Phenom., 15 (2020), 48, https://doi.org/10.1051/mmnp/2020035.

[7] T. Britton, F. Ball, and P. Trapman, *A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2*, Science, 369 (2020), pp. 846–849, https://doi.org/10.1126/science.abc6810.

[8] E. H. Bussell, C. E. Dangerfield, C. A. Gilligan, and N. J. Cunniffe, *Applying optimal control theory to complex epidemiological models to inform real-world disease management*, Philos. Trans. R. Soc., 374 (2019), 20180284, https://doi.org/10.1098/rstb.2018.0284.

[9] Y. Cao, S. Li, L. Petzold, and R. Serban, *Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution*, SIAM J. Sci. Comput., 24 (2003), pp. 1076–1089, https://doi.org/10.1137/S1064827501380630.

[10] R. Carli, G. Cavone, N. Epicoco, P. Scarabaggio, and M. Dotoli, *Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario*, Annu. Rev. Control, 50 (2020), pp. 373–393, https://doi.org/10.1016/j.arcontrol.2020.09.005.

[11] D. W. Clarke, *Generalized least-squares estimation of the parameters of a dynamic model*, in Proceedings of IFAC Symposium on Identification in Automatic Control Systems, Prague, 1967, International Federation of Automatic Control (IFAC), 1967.

[12] M. Coraggio, S. Xie, F. D. Lellis, G. Russo, and M. di Bernardo, *Intermittent Non-pharmaceutical Strategies to Mitigate the COVID-19 Epidemic in a Network Model of Italy via Constrained Optimization*, preprint, https://arxiv.org/abs/2103.16502, 2021.

[13] T. Damm, L. Grüne, M. Stieler, and K. Worthmann, *An exponential turnpike theorem for dissipative discrete time optimal control problems*, SIAM J. Control Optim., 52 (2014), pp. 1935–1957, https://doi.org/10.1137/120888934.

[14] O. Diekmann, H. Heesterbeek, and T. Britton, *Mathematical Tools for Understanding Infectious Diseases Dynamics*, Princeton Ser. Theor. Comput. Biol., Princeton University Press, Princeton, NJ, 2013.

[15] O. Diekmann, J. Heesterbeek, and J. Metz, *On the definition and the computation of the basic reproduction ratio R0 in models for infectious-diseases in heterogeneous populations*, J. Math. Biol., 28 (1990), pp. 365–382, https://doi.org/10.1007/BF00178324.

[16] Intensivregister-Team am RKI, *Tagesreport aus dem DIVI-Intensivregister*, 2020, http://doi.org/10.25646/7417 (accessed 2020-10-20).

[17] M. Dreher, A. Kersten, J. Bickenbach, P. Balfanz, B. Hartmann, C. Cornelissen, A. Daher, R. Stöhr, M. Kleines, and S. Lemmen, *Charakteristik von 50 hospitalisierten COVID-19-Patienten mit und ohne ARDS*, Dtsch. Arztebl. Int., 117 (2020), pp. 271–278, https://doi.org/10.3238/arztebl.2020.0271.

[18] W. Esterhuizen, T. Aschenbruck, J. Lévine, and S. Streif, *Maintaining hard infection caps in epidemics via the theory of barriers*, IFAC-PapersOnLine, 53 (2020), pp. 16100–16105, https://doi.org/10.1016/j.ifacol.2020.12.429.

[19] A. Fischer, K. Chudej, and H. J. Pesch, *Optimal vaccination and control strategies against dengue*, Math. Methods Appl. Sci., 42 (2019), pp. 3496–3507, https://doi.org/10.1002/mma.5594.

[20] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri, *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy*, Nat. Med., 26 (2020), pp. 855–860, https://doi.org/10.1038/s41591-020-0883-7.

[21] J. A. M. Gondim and L. Machado, *Optimal quarantine strategies for the COVID-19 pandemic in a population with a discrete age structure*, Chaos Solit. Fract., 140 (2020), 110166, https://doi.org/10.1016/j.chaos.2020.110166.

[22] S. Grundel, S. Heyder, T. Hotz, T. K. S. Ritschel, P. Sauerteig, and K. Worthmann, *How to coordinate vaccination and social distancing to mitigate SARS-CoV-2 outbreaks*, SIAM J. Appl. Dyn. Syst., 20 (2021), pp. 1135–1157, https://doi.org/10.1137/20M1387687.

[23] L. Grüne, M. Schaller, and A. Schiela, *Exponential sensitivity and turnpike analysis for linear quadratic optimal control of general evolution equations*, J. Differential Equations, 268 (2020), pp. 7311–7341, https://doi.org/10.1016/j.jde.2019.11.064.

[24] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I: *Non-stiff Problems*, 2nd ed., Springer Ser. Comput. Math. 8, Springer-Verlag, Berlin, 1993, https://doi.org/10.1007/978-3-540-78862-1.

[25] E. Hansen and T. Day, *Optimal control of epidemics with limited resources*, J. Math. Biol., 62 (2011), pp. 423–451, https://doi.org/10.1007/s00285-010-0341-0.

[26] 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, *Temporal dynamics in viral shedding and transmissibility of COVID*-19, Nat. Med., 26, pp. 672–675, https://doi.org/10.1038/s41591-020-0869-5.

[27] H. W. Hethcote, *The mathematics of infectious diseases*, SIAM Rev., 42 (2000), pp. 599–653, https://doi.org/10.1137/S0036144500371907.

[28] International Monetary Fund, *Policy Responses to COVID*-19, https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19, 2020 (accessed 2020-10-26).

[29] International Monetary Fund, *World Economic Outlook: A Long and Difficult Ascent*, https://www.imf.org/en/Publications/WEO/Issues/2020/04/14/weo-april-2020, 2020 (accessed 2020-10-26).

[30] D. C. P. Jorge, M. S. Rodrigues, M. S. Silva, L. L. Cardim, N. B. da Silva, I. H. Silveira, V. A. F. Silva, F. A. C. Pereira, A. R. de Azevedo, A. A. S. Amad, S. T. R. Pinho, R. F. S. Andrade, P. I. P. Ramos, and J. F. Oliveira, *Assessing the nationwide impact of COVID*-19 *mitigation policies on the transmission rate of SARS-CoV*-2 *in Brazil*, Epidemics, 35 (2021), 100465, https://doi.org/10.1016/j.epidem.2021.100465.

[31] M. Kantner and T. Koprucki, *Beyond just flattening the curve: Optimal control of epidemics with purely non-pharmaceutical interventions*, J. Math. Ind., 10 (2020), 23, https://doi.org/10.1186/s13362-020-00091-3.

[32] C. M. Kellett, S. R. Weller, T. Faulwasser, L. Grüne, and W. Semmler, *Feedback, dynamics, and optimal control in climate economics*, Annu. Rev. Control, 47 (2019), pp. 7–20, https://doi.org/10.1016/j.arcontrol.2019.04.003.

[33] S. Kissler, J. R. Fauver, C. Mack, C. G. Tai, M. I. Breban, A. E. Watkins, R. M. Samant, D. J. Anderson, D. D. Ho, N. D. Grubaugh, and Y. Grad, *Densely Sampled Viral Trajectories Suggest Longer Duration of Acute Infection with B*.1.1.7 *Variant Relative to Non-B*.1.1.7 *SARS-CoV*-2, https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37366884.

[34] J. Köhler, L. Schwenkel, A. Koch, J. Berberich, P. Pauli, and F. Allgöwer, *Robust and optimal predictive control of the COVID*-19 *outbreak*, 51 (2021), pp. 525–539, https://doi.org/10.1016/j.arcontrol.2020.11.002.

[35] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler, *The incubation period of coronavirus disease* 2019 *(COVID*-19*) from publicly reported confirmed cases: Estimation and application*, Ann. Intern. Med., 172 (2020), pp. 577–582, https://doi.org/10.7326/M20-0504.

[36] T. T. Le, J. P. Cramer, R. Chen, and S. Mayhew, *Evolution of the COVID*-19 *vaccine development landscape*, Nat. Rev. Drug Discov., 19 (2020), pp. 667–668, https://doi.org/10.1038/d41573-020-00151-8.

[37] J. G. V. Miranda, M. S. Silva, J. G. Bertolino, R. N. Vasconcelos, E. C. B. Cambui, M. L. V. Araújo, H. Saba, D. P. Costa, S. G. Duverger, M. T. de Oliveira, H. S. de Araújo Neto, W. de Jesus Sant'anna Franca-Rocha, D. C. P. Jorge, J. F. de Oliveira, R. F. S. Andrade, and R. S. do Rosário, *Scaling effect in COVID*-19 *spreading: The role of heterogeneity in a hybrid ODE-network model with restrictions on the inter-cities flow*, Phys. D, 415 (2021), 132792, https://doi.org/10.1016/j.physd.2020.132792.

[38] M. M. Morato, S. B. Bastos, D. O. Cajueiro, and J. E. Normey-Rico, *An optimal predictive control strategy for COVID*-19 *(SARS-CoV*-2*) social distancing policies in Brazil*, Annu. Rev. Control, 50 (2020), pp. 417–431, https://doi.org/10.1016/j.arcontrol.2020.07.001.

[39] J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and W. J. Edmunds, *Social contacts and mixing patterns relevant to the spread of infectious diseases*, PLoS Med., 5 (2008), pp. 381–391, https://doi.org/10.1371/journal.pmed.0050074.

[40] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed., Springer, 2006, https://doi.org/10.1007/978-0-387-40065-5.

[41] M. Ogura and V. M. Preciado, *Stability of SIS spreading processes in networks with non-Markovian transmission and recovery*, IEEE Trans. Control Netw. Syst., 7 (2020), pp. 349–359, https://doi.org/10.1109/TCNS.2019.2905131.

[42] Organisation for Economic Co-operation and Development, *The Territorial Impact of COVID*-19*: Managing the Crisis Across Levels of Government*, http://www.oecd.org/coronavirus/policy-responses/the-territorial-impact-of-covid-19-managing-the-crisis-across-levels-of-government-d3e314e1, 2020 (accessed 2020-10-26).

[43] M. Park, A. R. Cook, J. T. Lim, Y. Sun, and B. L. Dickens, *A systematic review of COVID*-19 *epidemiology based on current evidence*, J. Clin. Med., 9 (2020), 967, https:

//doi.org/10.3390/jcm9040967.

[44] T. A. Perkins and G. España, *Optimal control of the COVID-19 pandemic with non-pharmaceutical interventions*, Bull. Math. Biol., 82 (2020), 118, https://doi.org/10.1007/s11538-020-00795-y.

[45] S. J. Qin and T. A. Badgwell, *A survey of industrial model predictive control technology*, Control Eng. Pract., 11 (2003), pp. 733–764, https://doi.org/10.1016/S0967-0661(02)00186-7.

[46] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, *Model Predictive Control: Theory, Computation, and Design*, 2nd ed., Nob Hill Publishing, Madison, WI, 2019.

[47] Q. Richard, S. Alizon, M. Choisy, M. T. Sofonea, and R. Djidjou-Demasse, *Age-structured non-pharmaceutical interventions for optimal control of COVID-19 epidemic*, PLoS Comput. Biol., 17 (2021), e1008776 https://doi.org/10.1371/journal.pcbi.1008776.

[48] P. Scarabaggio, R. Carli, G. Cavone, N. Epicoco, and M. Dotoli, *Stochastic Non-Pharmaceutical Optimal Control Strategies to Mitigate COVID-19*, preprint, https://doi.org/10.36227/techrxiv.14413259.v1, 2021.

[49] J. Schilling, M. Diercke, D. Altmann, W. Haas, and S. Buda, *Vorläufige Bewertung der Krankheitsschwere von COVID-19 in Deutschland basierend auf übermittelten Fällen gemäß Infektionsschutzgesetz*, Epid. Bull., 2020 (2020), pp. 3–9, https://doi.org/http://dx.doi.org/10.25646/6670.2.

[50] O. Sharomi and T. Malik, *Optimal control in epidemiology*, Ann. Oper. Res., 251 (2017), pp. 55–71, https://doi.org/10.1007/s10479-015-1834-4.

[51] Statistisches Bundesamt, *Bevölkerung: Deutschland, Stichtag, Altersjahre*, 2019, https://www-genesis.destatis.de/genesis/online/ (accessed 2020-10-16).

[52] R. N. Thompson, C. A. Gilligan, and N. J. Cunniffe, *Control fast or control smart: When should invading pathogens be controlled?*, PLoS Comput. Biol., 14 (2018), e1006014, https://doi.org/10.1371/journal.pcbi.1006014.

[53] C. Tsay, F. Lejarza, M. A. Stadtherr, and M. Baldea, *Modeling, state estimation, and optimal control for the US COVID-19 outbreak*, Sci. Rep., 10 (2020), 10711, https://doi.org/10.1038/s41598-020-67459-8.

[54] P. van den Driessche and J. Watmough, *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission*, Math. Biosci., 180 (2002), pp. 29–48, https://doi.org/10.1016/S0025-5564(02)00108-6.

[55] G. Vossen and H. Maurer, *On L^1-minimization in optimal control and applications to robotics*, Optim. Control Appl. Meth., 27 (2006), pp. 301–321, https://doi.org/10.1002/oca.781.

[56] N. J. Watkins, C. Nowzari, and G. J. Pappas, *Robust economic model predictive control of continuous-time epidemic processes*, IEEE Trans. Automat. Control, 65 (2020), pp. 1116–1131, https://doi.org/10.1109/TAC.2019.2919136.

[57] R. Wölfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Bruenink, J. Schneider, R. Ehmann, K. Zwirglmaier, C. Drosten, and C. Wendtner, *Virological assessment of hospitalized patients with COVID-2019*, Nature, 581 (2020), pp. 465–469, https://doi.org/10.1101/2020.03.05.20030502.

[58] World Health Organization, *Coronavirus Disease 2019 (COVID-19) Situation Report 51*, 2020, https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10 (accessed 2020-08-20).

[59] M. Zamir, Z. Shah, F. Nadeem, A. Memood, H. Alrabaiah, and P. Kumam, *Non pharmaceutical interventions for optimal control of COVID-19*, Comput. Methods Programs Biomed., 196 (2020), 105642, https://doi.org/10.1016/j.cmpb.2020.105642.