Networks and Heterogeneous Media

https://www.aimspress.com/journal/nhm

NHM, 19(2): 500–525.
DOI: 10.3934/nhm.2024022
Received: 15 January 2024
Revised: 25 April 2024
Accepted: 06 May 2024
Published: 14 May 2024

Research article

Optimal control of pandemics via a sociodemographic model of non-pharmaceutical interventions

Ryan Weightman, Temitope Akinode and Benedetto Piccoli *

Center for Computational and Integrative Biology, Rutgers Camden, Camden NJ USA

* **Correspondence:** Email: piccoli@camden.rutgers.edu.

Abstract: The COVID-19 pandemic highlighted the need to quickly respond, via public policy, to the onset of an infectious disease breakout. Deciding the type and level of interventions a population must consider to mitigate risk and keep the disease under control could mean saving thousands of lives. Many models were quickly introduced highlighting lockdowns, testing, contact tracing, travel policies, later on vaccination, and other intervention strategies along with costs of implementation. Here, we provided a framework for capturing population heterogeneity whose consideration may be crucial when developing a mitigation strategy based on non-pharmaceutical interventions. Precisely, we used age-stratified data to segment our population into groups with unique interactions that policy can affect such as school children or the oldest of the population, and formulated a corresponding optimal control problem considering the economic cost of lockdowns and deaths. We applied our model and numerical methods to census data for the state of New Jersey and determined the most important factors contributing to the cost and the optimal strategies to contained the pandemic impact.

Keywords: epidemiological compartmental model; COVID-19; sociodemographic variables; optimal control

1. Introduction

The human population is no stranger to a modern epidemic, with resurgences of Dengue fever throughout hundreds of years of history [47], repeated Ebola outbreaks since 1976 [21], the AIDS epidemic beginning in 1981 [19], and more recently the COVID-19 pandemic, which took center stage in the year 2020 [42] and has been a unifying motivation of researchers across all scientific fields. With each surge of virus comes a desire to understand both its biology and transmission. Different viruses can be transmitted differently; via direct physical contact, airborne droplets produced from coughing, sneezing, etc [32], and in some cases via a nonhuman vector population [44]. Understanding the transmission route is imperative to developing a mitigation strategy. For example, in the case of the

West Nile Virus, it is more important to control the human interactions with the vector population than to control human-to-human contact [14]. However, in the case of a virus like COVID-19, it is very important to understand human-to-human contact to make forecasts and ultimately policy decisions that are best for both a population and its economy [1, 18].

In this paper, we introduce a general model for studying interactions between people of different sociodemographic categories in the case of a virus where transmission occurs during a human-to-human interaction. We start by developing the model which uses an interaction matrix to govern interaction levels based on a sociodemographic characteristic. Next, an optimal control problem is considered with regards to optimizing demographic-targeted lockdowns, followed by simulations to test the model and, lastly, a conclusion with future steps.

1.1. Epidemiological modeling of populations and optimal control

Epidemiological models representing a population in terms of interacting susceptible/infected subpopulations date back to Kermack and McKendrick who subdivided populations into compartments and provided differential equations driven by infection and recovery rates [27]. The model type that was born from this work is often called a "SIR" model or "SEIR" model standing for Susceptible, Exposed, Infected, and Recovered which are the labels representing the compartments of the model. The equations of such a model read $\dot{S} = -\frac{\beta SI}{N}$, $\dot{E} = \frac{\beta SI}{N} - \delta E, \dot{I} = \delta E - \gamma I, \dot{R} = \gamma I$, where β and γ are the infection rate, and recovery rate respectively, with δ being the latent period of the virus, or the time between exposure and infectivity. The key parameter $R_0 = \frac{\beta}{\gamma}$ [15] represents the "basic reproduction number", that is, the number of secondary cases in which one case would produce in a completely susceptible population.

This basic model has been adapted in many ways by increasing the number and complexity of compartments to: capture disease progression [20], consider undetectable infections or asymptomatic infections [11, 36], consider social parameters such as age-structure and spatial distribution of populations [8, 26, 53], control vaccination and other mitigation strategies [13, 35], consider in-host dynamics [4, 43], and more.

SIR models are used as a basic logic for many modeling efforts, including those that do not necessarily use differential equations to govern movement. With the increasing computing power available to researchers, Agent-Based Modeling (ABM), otherwise known as individual-based modeling, has gained quite a bit of interest [5]. In [22], for example, an ABM's ability to easily model spatial parameters is employed to study indoor non-pharmaceutical interventions of contagious respiratory diseases. These types of models also allow for easy parameterization of a unique population and the characteristics identified as important to viral transmission.

Model results, including forecasts, are improved by more descriptive parameters. In the case of a SEIR model, once such population-specific information is provided, and the researcher can answer questions and make forecasts based on model results. In [52], optimal control is implemented to measure the impact that awareness has on a population being plagued by AIDS. In [6] Pontryagin's Maximum Principle (PMP) is applied to a control problem for disease in livestock where the controls involve vaccination, isolation, culling, and reduction of transmission. The optimal lockdown schedule was shown in [49] to depend strongly on the mutation of the virus itself over a time period. In fact, when viral mutation is considered, waves of virus can match very closely to the waves seen during the COVID-19 pandemic.

1.2. Paper's contributions

In this paper, we propose a general SEIR model that uses a sociodemographic interaction matrix to capture the interaction levels between subpopulations. The additional parameterization of subpopulation allows us to track more specifically the disease progression and allows the description of targeted non-pharmaceutical measures. The availability of specific data is usually limited, thus we focus on age-brackets to subdivide the population in groups which match well the US census data. The mathematical approach is easily adaptable to other sociodemographic variables.

Then, we formulate an optimal control problem, where the control variables are the level of severity of a finite set of possible lockdown interventions. We choose three specific interventions: general lockdown, school closures, and restriction to elderly interactions. In practice, the latter can be achieved via food delivery services, home care, and retirement homes policies. With a virus pandemic in mind, as the recent COVID-19 one, the selection of the three interventions allows to capture the importance of protecting vulnerable subpopulations (elderly had a much higher death rate) as well as understanding the role played by intense social interactions (as those happening in schools). The cost function is designed to capture the economic cost of interventions. More precisely, a first component accounts for the cost of interventions, due to production loss and increased costs, while a second component accounts for cost of deaths.

Once the model is detailed, we perform standard analysis using mathematical control theory. The existence of optimal control follows from classic results and we point out that the PMP suggests that optimal control is likely bang-bang, i.e., for most of the time, the intervention will be either not performed or performed to its maximum extent. More conclusive results are impeded by the high-dimensionality and nonlinearity of the dynamics, thus we resort to numerical methods.

Extensive simulations allows us to study both the importance of various factors on the efficiency of interventions, as well as the characteristics of optimal policies. We first compare the cost of no intervention with the optimal one for R_0=1.7 (similar to COVID-19 in some of the pandemic phases), showing an improvement of 90% of the cost. Even for more aggressive viruses (R_0=2.5) the cost reduction is of 80%. Interestingly, the optimal lockdown policies include an initial general lockdown, restrict the elderly interactions for most of the time horizon, and include alternating closure and opening of schools. Subsequently, we show how the timing of interventions is more critical than their intensity. As expected, if limited interventions are allowed, then the optimal policy tends to extend them over time.

The cost per person of each intervention is estimated from literature, thus we include an investigation of the impact of daily costs over optimal policies. The total cost here is the sum of the costs of lockdown with the cost of death. This represents a realized economic cost of dealing with the pandemic. The daily cost per person represents the cost, each day, of locking down that person. This estimate includes things like not going to work, not using public resources, and not shopping/eating at public establishments. An estimated cost of future deaths has been included in order to deal with poor behavior of the optimizer in the final few days since the simulation ends before the final infected can pass to recovered and subsequently be included in deaths calculations. The general conclusion is that the optimal policies are very consistent for different daily costs, with optimal policies always consisting of initial general lockdowns limited in time, lengthy restrictions to elderly, and intermittent school closures. We finally look at the different interventions separately. School closures are the least effective when used as only measures, while restrictions for elderly and general lockdowns provide

significant cost reductions. We also observe that general lockdowns may happen to be intermittent.

Our main conclusion is that models including sociodemographic information can be of paramount importance in highlighting key factors in pandemic management. We compute optimal policies for non-pharmaceutical interventions, and estimate the economic and social impact on the population. Future development will include the exploration of different sociodemographic variables and a wider set of interventions.

It is worth noting that when attempting to model the lockdown of a population via specific sociodemographic characteristics using equation-based modeling, things can get more complicated when overlap exists between two subpopulations. For example, if rather than age we chose to parameterize based on "beach goers" and "restaurant goers", a nonzero lockdown in both could have a double counting effect on a person that eats at restaurants and goes to the beach. Naive attempts at addressing such cases may result in unrealistic dependencies such as a model that can only limit attendance of either restaurants or beach visits, but not both at the same time. For this reason, the decisions of which variables to consider when labeling populations as well as how to incorporate such labels into the modeling effort are extremely important.

2. A sociodemographic SEIR model

We incorporate sociodemographic variables into a classical SEIR model in such a way that a more complex mitigation strategy can be found. We augment the basic SEIR model to differentiate between subgroups of the population who may have unique interactions, as for instance, those due to school activities, and focus on age groups. However, our method may be adapted to other types of sociodemographic variables, such as geographic location, income, health access, and others, which plays an important role in pandemic progression. Our model will be fit to census data for the state of New Jersey (but could be easily adapted to other states and countries), thus we consider 7 different groups corresponding to age brackets of census data denoted by a subscript, e.g., S_j, $j = 1, \dots, 7$. See Section 5 for the specific age brackets.

The interactions among different age groups will be encoded by a matrix $l = (l_{k,j})$, where $l_{k,j}$ quantifies the interactions that group k is having with group j. In particular, $l_{k,j}$ will be used to weight the infections resulting from susceptibles in group j, i.e., S_j, interacting with the infected in group k, i.e., I_k. For the sake of generality, we will state our problem with a number $h \in \mathbb{N}$ of groups, while simulations will be performed using census data, thus with $h = 7$. Let us denote by $N_j = S_j + E_j + I_j + R_j$ the total population in the j-th age group (which is constant), and set $\widetilde{N}_j = \sum_{k=1}^{h} l_{k,j} N_k$, which represents the total number of individuals the j-th susceptibles are interacting with. Then, we set:

$$\begin{cases} \dot{S}_j = -u \dfrac{\beta S_j}{\widetilde{N}_j} \displaystyle\sum_{k=1}^{h} l_{k,j} I_k \\ \dot{E}_j = u \dfrac{\beta S_j}{\widetilde{N}_j} \displaystyle\sum_{k=1}^{h} l_{k,j} I_k - \delta E_j \\ \dot{I}_j = \delta E_j - \gamma I_j \\ \dot{R}_j = \gamma I_j \end{cases} \tag{2.1}$$

for $j = 1, \ldots, h$, where $1 - u$ is the lockdown rate, β the infection rate, l the interaction matrix, δ the latent period, and γ the recovery rate. Often, as in the case of the simulations here, interaction information is available for the general population but not exactly available for the interactions between susceptible and infected people. A more severe symptomatic virus would have fewer interactions due to the realization that the infectious person is infected but a higher infection rate per interaction, while a less severe virus might have more interactions but less infectivity. One method that could be employed to capture such nuances is separating asymptomatic versus symptomatic infection into two separate compartments [33]. Here, we introduce the model in its most basic state, so the assumption is that the parameter choice of our infection rate and interaction matrix capture these nuances.

To study how the control of interactions amongst different groups could affect the system, we subdivide the interaction matrix $l = (l_{k,j})$ into a well-known socially driven subset of uniquely interacting groups. We define a set of l^i to be n interaction matrices where $l^i_{k,j}$ encodes some portion of the interaction between groups k and j due to the i-th subset of interactions. For example, if there are four sociodemographic variables that are known to drive the interactions between age groups, to capture the different interactions, $n = 4$. Notice that while l^i is the same size as l, each element is either a portion of the corresponding element of l or 0. It holds that

$$\sum_{i=1}^{n} l^i_{k,j} = l_{k,j} \tag{2.2}$$

for every k, j and, thus, $l = \sum_{i=1}^{n} l^i$. Each new interaction matrix brings the ability to encode the regulation of that subgroup of interactions, and we will denote by u_i the "amount of allowed interactions" such that the severity of lockdown for each group of interactions is $1 - u_i$. Therefore, our final model reads

$$\begin{cases} \dot{S}_j = -\beta \dfrac{S_j}{\widetilde{N}_j} \left(u_1 \displaystyle\sum_{k=1}^{h} l^1_{k,j} I_k + u_2 \sum_{k=1}^{h} l^2_{k,j} I_k + \ldots + u_n \sum_{k=1}^{h} l^n_{k,j} I_k \right) \\ \dot{E}_j = \beta \dfrac{S_j}{\widetilde{N}_j} \left(u_1 \displaystyle\sum_{k=1}^{h} l^1_{k,j} I_k + u_2 \sum_{k=1}^{h} l^2_{k,j} I_k + \ldots + u_n \sum_{k=1}^{h} l^n_{k,j} I_k \right) - \delta E_j \\ \dot{I}_j = \delta E_j - \gamma I_j \\ \dot{R}_j = \gamma I_j \end{cases} \tag{2.3}$$

with $j = 1, \ldots, h$ and other parameters set as in Eq (2.1).

2.1. Optimization of lockdown policies

We consider an optimal control problem for the system (2.3) representing the interactions between the various age groups. For this, we define a cost functional with two main goals; minimizing the economic cost of locking down a population or subpopulation, and minimizing the deaths that the whole population will incur over a given time horizon. Because this model has no reinfection considered, the recovered compartment is monotonically increasing and accounts for all people who have experienced the full tenure of the virus. We then can compute deaths simply by multiplying death rates with the total recovered at the end of a simulation. This total deaths is represented in the

following by the variable "D".

As is customary in optimal control, we will consider a running cost $\mathcal{L}$ to be integrated over the time horizon and a final cost ψ depending on the state at the end of the time horizon. Specifically, we set

$$\mathcal{L} = \sum_{i=1}^{n} c_i(1-u_i)P_i + c_{n+1}D, \quad \psi = c_{n+2}I_{\mathbf{end}} \tag{2.4}$$

with P_i as the total population isolated due to the i-th control action (lockdown), c_i is the cost of the $i-th$ lockdown per person, usually ranging between USD 10 and USD 100, and c_{n+1} the cost per death, usually set to USD $1,500,000$. Additionally, $I_{\mathbf{end}}$ represents the total infected at the end of the simulation. This part of the cost function takes into account the cost of deaths that will happen beyond the simulation horizon. The weight c_{n+2} needs to take into account future deaths due to current and future infected and will be calibrated in simulations.

The values of c_i are typically estimated in terms of the economic impact of the lockdown measure for lost production output or increased costs. For instance, isolating workers will likely decrease the production rate, while school closure will increase childcare costs. A more complex phenomenon may be measured as the loss in educational outcomes due to online school activities compared to in-class ones; see [16]. Many studies have shown that disparities in education between vulnerable and non-vulnerable populations grow substantially during events like school closure events [17]. Therefore, similar to assigning a cost of mortality, the true cost of school closure must incorporate not only a financial value but also a social value (which in the end may result in financial deficits as well).

3. Model analysis

The proposed optimal control problem can be stated as follows:

$$\dot{x} = f(x) + \sum_{i=1}^{n} u_i f_i(x), \quad \min_{u(\cdot)\in\mathcal{U}} J(u), \quad J(u) = \int_0^T \mathcal{L}(x,u) + \psi(T, x(T)), \tag{3.1}$$

where f is the drift term, f_i the controlled fields, $u = (u_1, \dots, u_n)$ the control vector, $\mathcal{U}$ the set of admissible controls, T the time horizon, $J(u)$ the cost functional, $\mathcal{L}$ the Lagrangian or running cost, and ψ the final cost.

Referring to model (2.3), the system can be written in the general form Eq (3.1) with the following choices: $x = (S_j, E_j, I_j, R_j)_{1\le j\le 7}$, f and f_i are defined by Eq (2.3), $\mathcal{U}$ is the set of measurable functions $u : [0,T] \to ([0,1])^n$, and $\mathcal{L}$ and ψ are given in Eq (2.4).

Due to the specific properties of Eqs (2.3) and (2.4), we get the following:

Theorem 1. *Consider the optimal control problem* (3.1) *with dynamics given by* (2.3) *and cost function by* (2.4). *There exists an optimal control for every time horizon $T > 0$ and initial conditions $x(0) = x_0$.*

Proof. For every $i = 1, \dots, N$, we have $u_i \in [0,1]$, thus the control set $U = ([0,1])^N$ is compact and convex. Moreover, the dynamics (2.3) is linear in the control, thus the associated multifunction $F(x) = \{f(x) + \sum_{i=1}^n u_i f_i(x) : u_i \in [0,1]\}$ has compact and convex values.

The dynamics (2.3) is smooth in all variables and admits the natural invariant region $\{x = (S_j, E_j, I_j, R_j)_{1\le j\le 7} : S_j \ge 0, E_j \ge 0, I_j \ge 0, R_j \ge 0, S_j + E_j + I_j + R_j = N_j\}$ where $N_j > 0$; thus, for every initial condition, the support of trajectories on $[0, T]$ is uniformly bounded. Moreover, the cost function (2.4) is smooth in all variables.

We can apply Theorem 5.2.1 of [7], obtaining the existence of an optimal control. □

Being that the optimal control problem is linear in the controls, one may attempt to use necessary conditions, such as the PMP, (see Theorem 6.5.1 of [7]) to find optimal controls. However, the number of variables (7 age groups with 4 populations each) and the nonlinearity w.r.t. state variables (e.g., terms SI to describe infections) render the problem too complex to obtain detailed information. Therefore, to compute optimal controls we will resort to numerical methods described in the next section. However, we provide some information derived from PMP for the general case and some results for a special case. Let us start recalling that, from the PMP, there exists a covector (λ_0, λ) so that the optimal control minimizes the Hamiltonian:

$$H(\lambda_0, \lambda, x, u) = \lambda \cdot (f(x) + \sum_{i=1}^{n} u_i f_i(x)) + \lambda_0 \cdot \mathcal{L}(x, u). \tag{3.2}$$

Since both the dynamics (2.3) and the cost function (2.4) are linear in u, this condition implies $u_i \in \{0, 1\}$ unless the corresponding switching function vanishes:

$$\psi_i(\lambda_0, \lambda, x, u) = \lambda \cdot f_i(x) + \lambda_0 \cdot \mathcal{L}_i(x, u) = 0$$

where we use the fact that we can write $\mathcal{L} = \mathcal{L}_0 + \sum_i u_i \mathcal{L}_i$ being the running cost linear in the control. This is indeed confirmed by simulations of Section 4, where we observe $u_i \in \{0, 1\}$ at all times. However, the number of switchings, i.e., changes from the value $u(t) = 1$ to the value $u(t) = 0$ and vice versa, is difficult to bound as observed in Figure 2. The resulting number of oscillations could be too frequent to be considered as actual policy changes over a time horizon.

Singular controls, i.e., in $]0, 1[$, may appear if the functions ψ_i vanish identically on an interval. While a general analysis is out of scope, we provide a result for the simple case of one control and one age group, thus we consider the dynamics:

$$\dot{S} = -u\frac{\beta S I}{N}, \dot{E} = u\frac{\beta S I}{N} - \delta E, \dot{I} = \delta E - \gamma I \tag{3.3}$$

where we omit the dynamics of R (not influencing the system). The system can be written compactly as $\dot{x} = f(x) + u_1 f_1(x)$ with $x = (S, E, I)$ and f, f_1 defined by (3.3). Singular controls may appear only if trajectories run on singular points, i.e., points $x = (S, E, I)$ such that the vectors $f(x)$, $f_1(x)$ and $[f, f_1](x)$ do not span the whole space [31,34]. Here, as usual, $[f, f_1]$ indicates the Lie bracket of f and f_1; see [7] for details. We have the following:

Theorem 2. *Consider the optimal control problem* (3.1) *with dynamics given by* (3.3)*. Then, the only singular points in the first orthant satisfy either* $S = 0$ *or* $I = 0$*.*

Proof. We can easily compute:

$$f = \begin{pmatrix} 0 \\ -\delta E \\ \delta E - \gamma I \end{pmatrix}, \quad f_1 = \begin{pmatrix} -\frac{\beta S I}{N} \\ \frac{\beta S I}{N} \\ 0 \end{pmatrix}, \quad [f, f_1] = \begin{pmatrix} -\frac{\beta S}{N}(\delta E - \gamma I) \\ \frac{\beta \delta S I}{N} + \frac{\beta S}{N}(\delta E - \gamma I) \\ -\frac{\beta \delta S I}{N} \end{pmatrix}. \tag{3.4}$$

If the vectors are linearly dependent, then the matrix formed with columns f, f_1, and $[f, f_1]$ must have determinant equal to zero. A simple computation shows that the determinant is given by a constant times S^2I^3, thus we conclude. □

The consequence of Theorem 2 is that trajectories corresponding to singular controls, also called singular trajectories, may only run on the coordinate axes $S = 0$ or $I = 0$. Now, if $S \equiv 0$, then we are in the case of whole population exposed, infected, or recovered, thus the control policy is not much relevant. If $I \equiv 0$ on a time interval, then, necessarily, $E \equiv 0$, thus we are in the case of no disease. Concluding in all relevant cases, singular trajectories do not appear.

4. Simulation of optimal control policies

In this Section, we provide simulation results obtained using the numerical algorithms described in Section 4.3. The values of parameters are reported in the Appendix A.

4.1. Description of lockdown policies

We now look at the case where $n = 3$. Our focus is on the three possible lockdown strategies:
(1) general lockdown with home mandate for the whole population;
(2) school closure affecting school-age population;
(3) restriction of elderly interaction (for instance, by retirement homes policies, food delivery, home care, and others.).

The second strategy corresponds to control u_2 and an interaction matrix l^2 defined as the 7×7 matrix with nonzero elements in the upper left 3×3 square representing interactions between groups 1–3. This group represents "school interactions".

The third strategy corresponds to control u_3 and the matrix l^3 is defined as the 7×7 matrix with nonzero elements in the final column and row, representing all interactions the elderly have with each other and the other age groups.

It was clear during the COVID-19 pandemic that curfews and social distancing measures specifically targeted at the elderly communities were crucial in slowing the virus spread and mortality [3]. Interestingly, the elderly represent the most vulnerable population while the school-age children represent the population with the strongest interaction coefficients. We assume that not all interactions are affected by the second and third lockdown measure, thus will use parameters τ_i, $i = 2, 3$, to represent the amount of interactions cancelled by the lockdown. We set $l^2_{k,j} = \tau_2 l_{k,j}$ for the indices $k, j = 1, \ldots, 3$ (and zero elsewhere), and $l^3_{k,j} = \tau_3 l_{k,j}$ for the indices k, j such that $k = 7$ or $j = 7$ (and zero elsewhere). Finally, because of Eq (2.2), we have for every k and j,

$$l^1_{k,j} = l_{k,j} - l^2_{k,j} - l^3_{k,j}. \tag{4.1}$$

Thus, l^1 is the matrix representing the remaining interactions not affected by the second and third strategy. We choose $\tau_2 = .66$ in our simulations. In the case of a lockdown, children will likely still see certain friends and family in their age ranges as well as their siblings. We choose $\tau_3 = .8$. This is because a lockdown of the elderly will result in continued interactions with the elderly person's family or the staff at any institution providing care.

4.2. Description of analyzed scenarios

In Section 4.4 we study how the optimal lockdown changes depending on the delayed response to the virus. We start with immediate response and then take a look at how the optimal changes if it takes some months to identify the virus and pass legislation to begin a lockdown. We also explore how a more severe R_0 or a less intense maximum lockdown affect the optimal lockdown's plan.

In Section 4.5, we then study the impact that the cost per person has on the policy surrounding each of our populations being locked down. There are varying opinions about how much a lockdown at different levels costs a society, so we explore how these shifting numbers might change optimal strategies.

Lastly, in Section 4.6, we perform a cost analysis of optimal lockdown strategies for each of our three populations asking the questions "If only one type of lockdown is employed, how is the cost affected ?" and "Which lockdown on its own impacts the population in the most positive way? "

All simulations are parametrized and assumed to be starting on December 20 2019. All simulation time horizons are discretized on the order of days.

4.3. Numerical methods for optimization

To compute the optimal controls for this problem, we use CasADi, an open-source tool for nonlinear optimization and algorithmic differentiation with a large degree of flexibility [2] . The dynamics of (2.3) are implemented with controls using an implicit Runge-Kutta 2nd order method and the cost functional with the trapezoidal rule. All controls are initialized with a "best guess" of the maximum control over the time horizon where each time-step represents one day. We then use `IpOpt` [46] (the interior-point optimization routine) to solve the optimization problem. The discretization of the controls is chosen to be half of the number of days, meaning over 180 days, and the control is updated 90 times. This is more than sufficient for our purposes (no government will update policy more than every other day) and reduces run-time from a model that would choose optimal controls daily.

4.4. Optimal lockdown varying pandemic stages, infectivity, intervention delays, and lockdown intensity

We first study the difference in costs, lockdown dynamics, and virus trajectories when allowing the lockdown to begin at different initial times. For new types of viruses, the initial time is affected by recognizing the new infection and devise screening methods.

To start the analysis, we report in Figure 1 a simulation of the age-based SEIR model (2.3) with no lockdown intervention, a replication rate, R_0, of 1.7 and initial S, E, I, and R values taken directly from COVID-19 data for the state of New Jersey. We take the total population of New Jersey from the 2021 census, and develop our exposed and infected as approximations of 14 days of COVID-19 data. Specifically, we begin simulations with a total population of 9.267 million people with initial infected across all age groups totaling 13736. We see that essentially all people in age groups one and three are infected over the time horizon of the virus. These dynamics mimic closely to that of a classic SIR model with a wave of cases emerging from few initial infections and then the virus dying out. Under the curtain, this model is enriched with the interactions between the various subpopulations at each time step, driven by our interaction matrix. In this case, the corresponding total cost consists only of

the cost of deaths, being the lockdown controls set to zero. This cost represents an upper limit for any control policy representing non-pharmaceutical intervention and is around USD 183 billions.

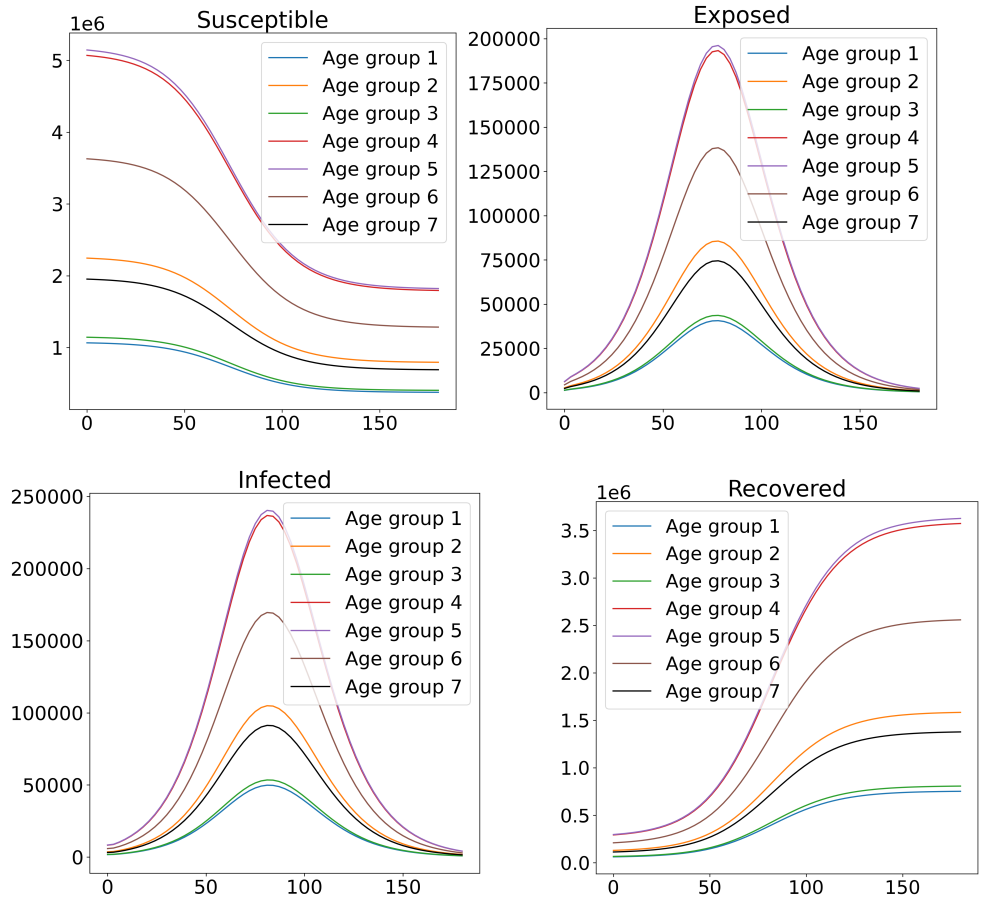

Figure 1. Dynamics of S, E, I, and R for a model with no lockdown. Here, $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0.25$, $\tau_2 = .66$, $\tau_3 = .8$ and the cost of death is 1.83×10^{11}.

In each of the following simulations, the lockdown cost is \$70 per person for a general lockdown, \$40 per person to lockdown the elderly, and \$25 per student to lockdown schools [9, 12, 28, 38], unless otherwise stated. There is a wide range of suggested costs in the literature where a more severe lockdown costs more per person than a less severe. We consider a lockdown of only the elderly to be a first step and, thus, a less severe mitigation strategy than locking down the entire population. Therefore the cost per person should be on the lower side of the estimates. Quantifying the "cost of death" in an economic sense can be tricky. This value is assumed in the literature to be anywhere from hundreds of thousands of dollars to the assumed VSL (value per statistical life) of \$10 million dollars [24, 29]. Because most COVID-19 deaths occur in the elderly, the VSL may be an inappropriately high measure of the average cost of a COVID-19 death. Therefore, while we explore the parameter in this section, we choose to follow the estimated value of \$1.5 million dollars to be the average cost of a COVID-19 death in the United States [10]. We incorporate a "cost of future deaths" to dampen end behavior effects of a lockdown fully dismissing with an influx if infected who do not have time (on the chosen time horizon) to be moved to recovered. This cost is exactly

$C_{n+2} = 100C_{n+1}$. For brevity, some figures of dynamics are omitted for each simulation. For example, because the exposed mimics so closely to the infected, in future figures, the exposed compartment will be omitted.

In Figure 2, we report the best case scenario: A situation where every type of lockdown is available at a 100% level at any instant of time. Notice that this is an idealized situation for various reasons: (1) A 100% level lockdown may be not feasible; (2) usually lockdowns need to be limited in time; (3) different types of lockdowns may need coordination; and (4) change in lockdown policies cannot occur too frequently. Interestingly, if we can start our lockdown policies early, the optimal case is to have intermittent lockdowns for school activities, keeping the exponential growth of a full wave of virus under control from the beginning. This, in a sense, was adopted by many states and school districts who chose to creatively reduce school interactions by splitting students into groups and having certain groups attend on certain days, taking certain days off to do remote instruction, etc. We also see that the optimal solution reduces interactions with the elderly as much as possible. We assume the ability to lockdown 100% of each population, which may be unrealistically effective, to create a lower bound for the realized cost to the population, while we consider differing levels of lockdown in future simulations. The total cost for New Jersey is now reduced to USD 31.4 billion, with a reduction of more than 80% of the cost without intervention.

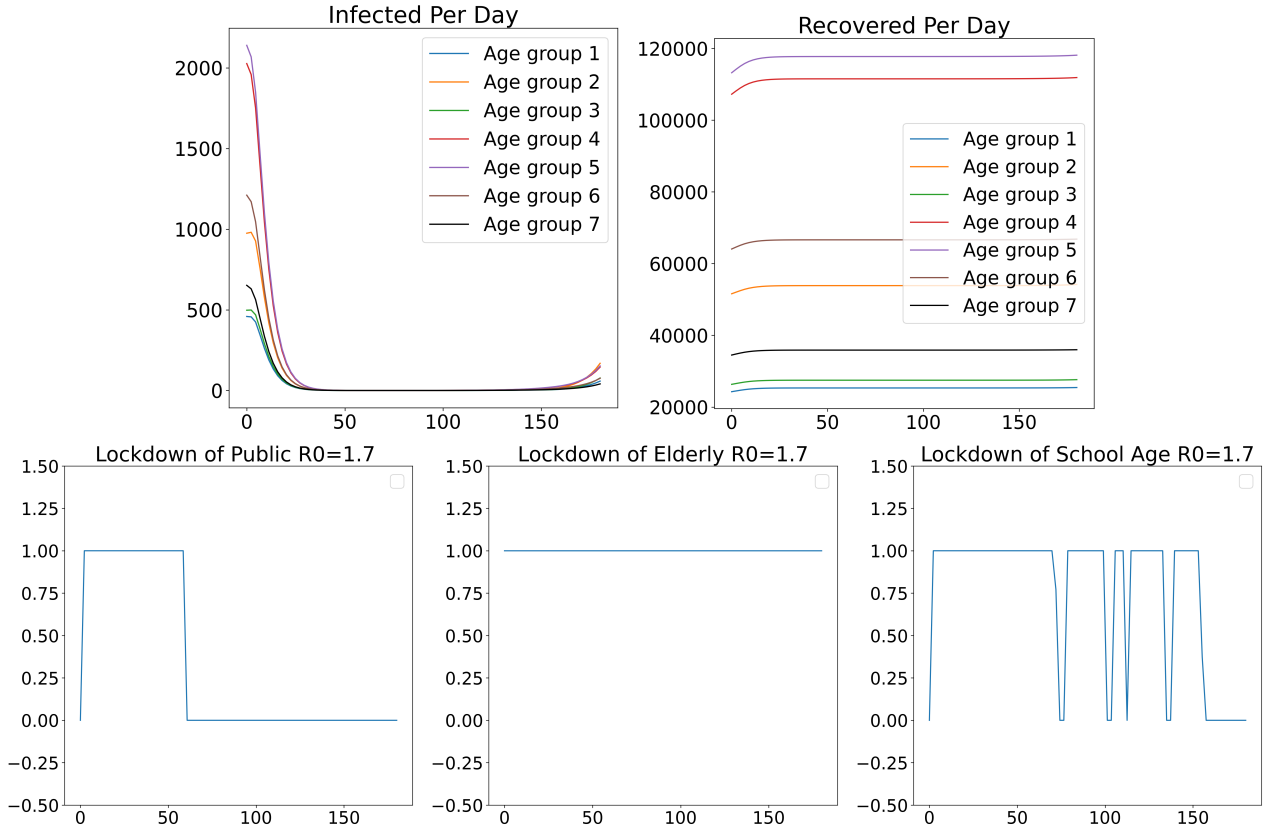

Figure 2. Dynamics of I, R (top) and optimal lockdowns policies (bottom). Here $R_0 = 1.7$, $\gamma = 0.2, \beta = 0.34, \delta = 0.25, \tau_2 = .66, \tau_3 = .8$. The total cost is 3.14×10^{10}.

In Figure 3, we see that when the virus is more aggressive ($R_0 = 2.5$), the optimal lockdown of

schools and the elderly remains fairly similar while the general public must get more severe restrictions with recurring lockdowns. Even so, there is still almost an 80% reduction in cost compared to a less aggressive virus with no mitigation strategy employed. This highlights the important point that a lockdown is crucial as a first mitigation; even a much more aggressive virus has a fraction of the impact of a less aggressive virus if a lockdown is employed for the former.

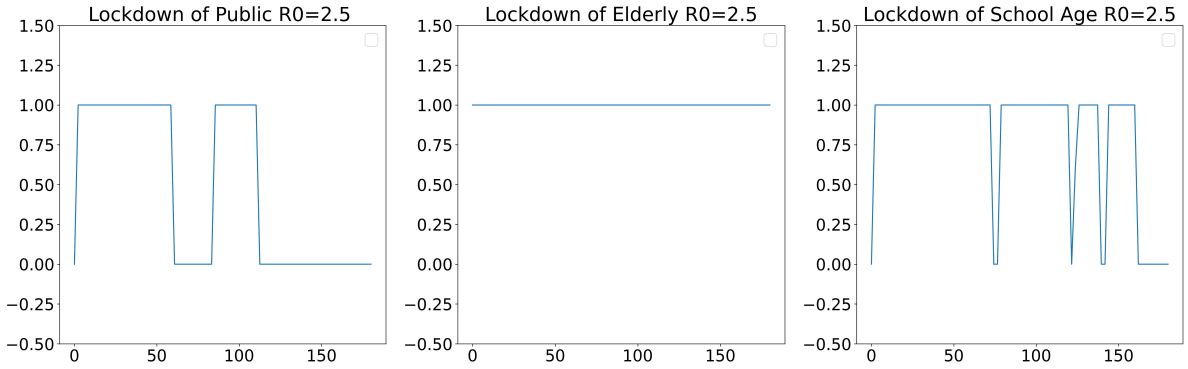

Figure 3. Optimal lockdowns with $R_0 = 2.5$, $\gamma = 0.2$, $\beta = 0.50$, $\delta = 0.25$, $\tau_2 = .66$, $\tau_3 = .8$. The total cost is 3.90×10^{10}.

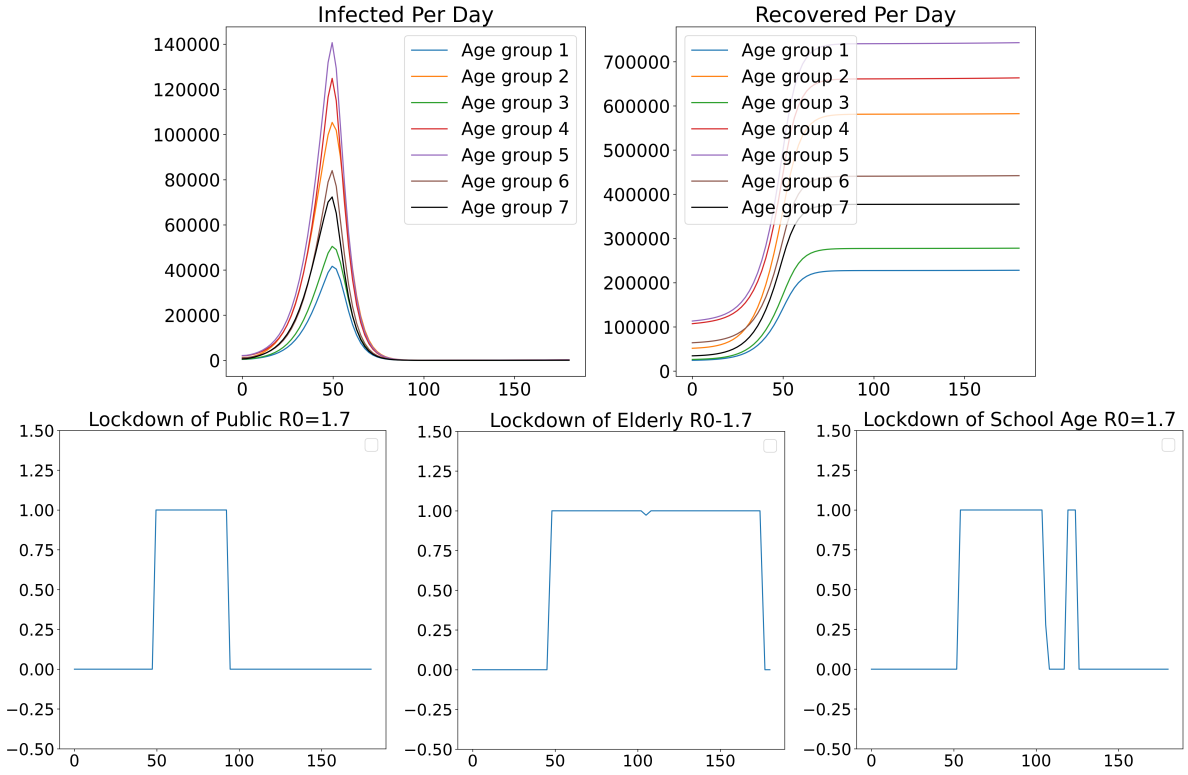

Figure 4. Dynamics of I, R, (top) and optimal lockdowns (bottom) with a 50 day delay before the policy is implemented. $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0.25$, $\tau_2 = .66$, $\tau_3 = .8$. The total cost here is 8.98×10^{10}.

To further investigate real scenarios, we consider the case where a lockdown is a response to an already propagating virus. It is known that a delay in the initial lockdown most likely results in a higher number of cases [25]. In Figure 4, we consider a 50 days delay in interventions with an R_0 of 1.7, and, as expected, the lockdowns start immediately after the delay. The COVID-19 pandemic started in the United States with the first reported case in January, while the first lockdown of the state of New Jersey was not until March, so a 50 day delay in response is a reasonable choice [50]. Additionally, the World Health Organization estimated COVID-19's initial R_0 to be between 1.4–2.4, so a choice of 1.7 is in line with a real virus [40]. Moreover, there is a greater strain due to the virus's ability to propagate, thus the locking down of the most vulnerable population (the elderly) and school activities are prolonged for almost the entire optimization horizon. We also see that the cost is much larger when mitigation strategies are not employed early. The total cost is USD 89.8 billion. The main message is that delays in policy activation are much more impactful than the virus infectivity.

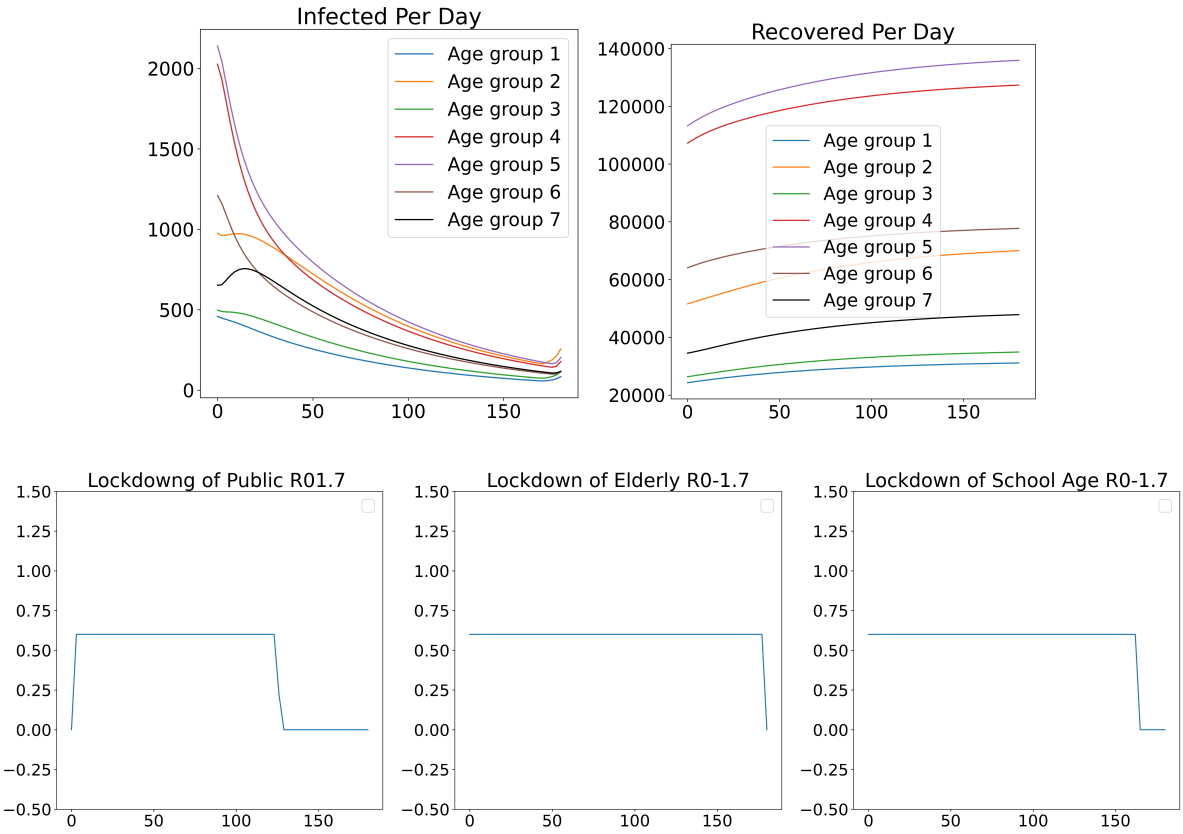

Figure 5. Dynamics of I, R, (top) and optimal lockdowns (bottom) with a cap at 60% maximum lockdowns at any given time. $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0.25$, $\tau_2 = .66$, $\tau_3 = .8$. The total cost here is 4.78×10^{10}.

Additionally, for logistical, economic, or legal reasons, most populations may be unable to perform a 100% lockdown. When we reduce the severity of possible lockdown from 100% to 60%, the optimal policies correspond to lockdowns extended over time. In Figure 5, we see that even a reduced lockdown rate can greatly curb the cost of a virus to a population with a total cost of USD 47.8 billion for New Jersey. The optimal solution here is to lock down the elderly and school activities as much as

possible for as long as possible with the rest of the public able to go back to no lockdown earlier. However, with less severe lockdowns, the trade-off is that there must be infrastructure in place to accommodate a longer lockdown for certain populations. In the case of school children, it is known that a prolonged loss of in-person education can be detrimental, so a short but strong lockdown is likely more beneficial [23]. However, it may be easier to accommodate a long-term but less severe lockdown for the elderly. A key takeaway here is that early intervention may outweigh more intense intervention in cost reduction. In fact, with the 50 day delay in Figure 4, there are 16, 322 deaths, whereas with our 60% lockdown allowed but with no delay, there are only 8, 298 deaths, as seen in Table 1.

Table 1. Data for Figures 1–5. Legend: OL = "Optimal Lockdown".

Fig	Case	R_0	Total Cost	Deaths
1	No Lockdown	1.7	1.83×10^{11}	122, 463
2	OL	1.7	2.2×10^{10}	3, 921
3	OL	2.5	3.75×10^{10}	4, 593
4	50-day delay OL	1.7	6.17×10^{10}	16, 322
5	60% OL	1.7	4.78×10^{10}	8, 298

Interestingly, by the end of 2020, the state of New Jersey had seen approximately 19, 200 COVID-19 deaths, with over 16, 000 after the first 6 months [51]. This suggests that a well-tuned delay in response is important in parameterizing an epidemiological model. Earlier intervention almost assuredly would have reduced deaths in the first year of the pandemic in a significant way. This is reflected in the fact that when mimicking New Jersey parameters, the total deaths is very close to the realized COVID-19 deaths, while if we force earlier intervention in the model, the deaths are much smaller.

4.5. Impact of cost of lockdown

We now study how varying the cost of lockdown per person, differentiating among different type of lockdowns, will affect the optimal control policy. In each simulation, our initial conditions are employed to mimic a pandemic that, if left with no intervention, is on its ascent to a peak in cases. Notice that without reinfection or more complex ways to choose parameters, a standard SIR model will generally have one wave of cases where the endemic equilibrium has no infected (all of the population has either stayed in susceptible, or moved through the system of equations to recovered) [30]. Figure 6 reports the optimal policy with equal daily cost of USD 10 per person for all the three types of lockdown. We notice that, despite different death rates and population sizes, the optimal lockdowns are very similar. In fact, all three run until exactly day 74, with a resurgence of only locking down the elderly. This suggests that a general lockdown would make the most sense for this situation.

We then keep the daily cost constant across lockdown types, but increase it. This causes a a differentiation in optimal lockdown policies, as seen in Figure 7. The higher daily cost, now USD 70, causes a more subtle lockdown than in Figure 6. However, while there is a reduction in lockdown for school age children and the general public, the lockdown in fact increases in time for the elderly who have a much higher death rate than the younger populations.

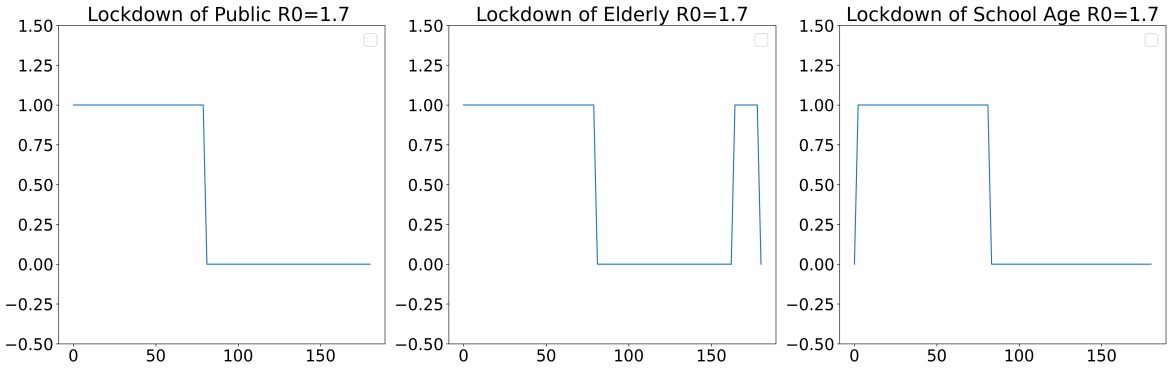

Figure 6. Optimal Lockdown with USD 10 daily cost per person. $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0.25$, $\tau_2 = .66$, $\tau_3 = .8$. The total cost is 1.39×10^{10}.

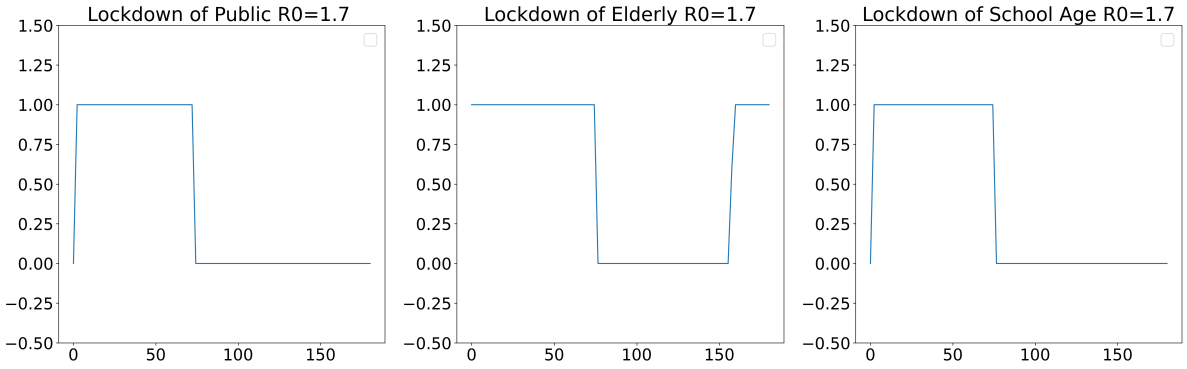

Figure 7. Optimal Lockdown with equal daily cost of USD 70 for all types. $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0.25$, $\tau_{2,3} = 1$. The total cost is 3.72×10^{10}.

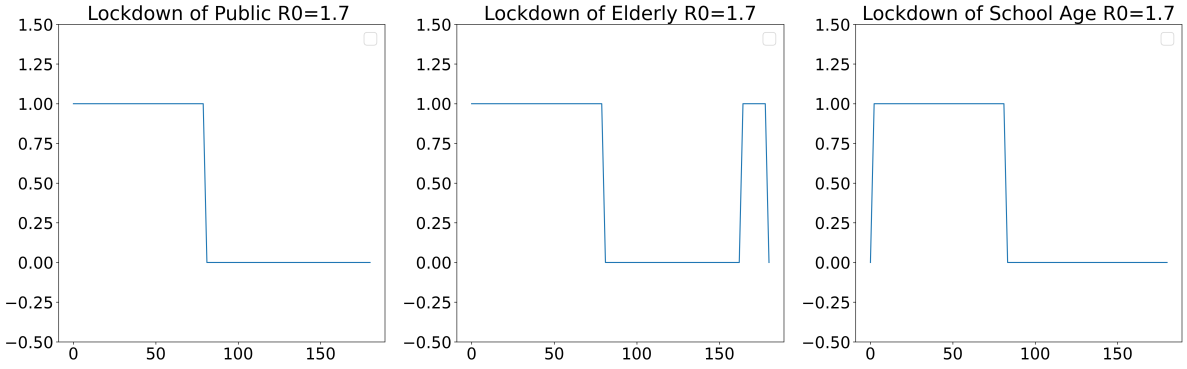

Figure 8. Optimal Lockdown with equal daily cost of USD 70 for all types, but with a cost of death of \$10 million dollars per person. $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0.25$, $\tau_2 = .66$, $\tau_3 = .8$. The total cost is 9.38×10^{10}.

Now, a lockdown of the general public comes with a much higher cost than one of a subgroup of the population, in part because there are more people affected, but also because when everybody is affected, the economic strain begins to take hold on small businesses, public transportation, and more.

For this reason, we see a tendency like in Figure 2 to minimize the lockdown of the public compared to our subpopulations.

In Figure 8, we see an extreme case where the parameters are the same as Figure 7, but the cost of a death is \$10 million dollars instead of 1.5 million. Interestingly, while the total cost has a significant increase, the optimal policy is not significantly different; a prolonged public lockdown and a slightly lessened resurgence of lockdown for the elderly. This could indicate that the optimal policy is fairly robust to a difference in cost of death.

4.6. Cost analysis by lockdown

A major goal of this work is to exemplify the power of stratifying the population into subgroups with specific sociodemographic characteristics and using these characteristics to develop targeted mitigation strategies when addressing an epidemic. In the following simulations, we will look at a situation where it takes a month to start the first mitigation strategies, the maximum lockdown is 60%, and like in the previous section, we will use a replication rate of 1.7. Now, we study the net benefit of a stratify-able population in the face of lockdown measures through the study of the cases of no lockdown, single targeted lockdowns, and when one can use the three lockdowns independently. To set the stage, we refer back to Figure 1, which provides a baseline for how the virus would progress over the given time horizon if there were no lockdown policies enacted at all. We see immediately the recognizable peak and descent of a virus wave in the infected compartment.

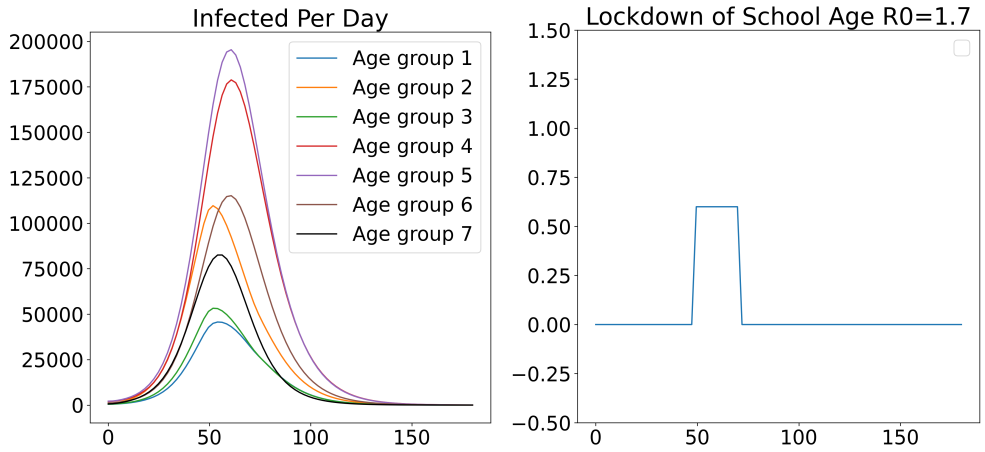

Figure 9. Dynamics of I (left) and optimal school closure (right).$R_0 = 1.7, \gamma = 0.2, \beta = 0.34$, $\delta = 0.25, \tau_2 = .66$. The maximum lockdown is 60% at any given time. The total cost between lockdown and deaths is 1.54725×10^{11}.

Researchers using a variety of models have reported conflicting results on the impact on the viruses spread of school closures. In the case of COVID-19, many estimates show that a full school closure results in less than 5% of prevented deaths [45]. In comparison, the costs of no lockdown outweigh the costs in Figure 9 of a school lockdown by only about 15%, meaning that optimizing school closure will at most reduce the direct societal cost by 15% (this does not take into account the loss of education long-term). It is also worth noting that even when school closure is the only method used to mitigate virus risk, the optimal strategy is not to fully lock down the school students for long-term.

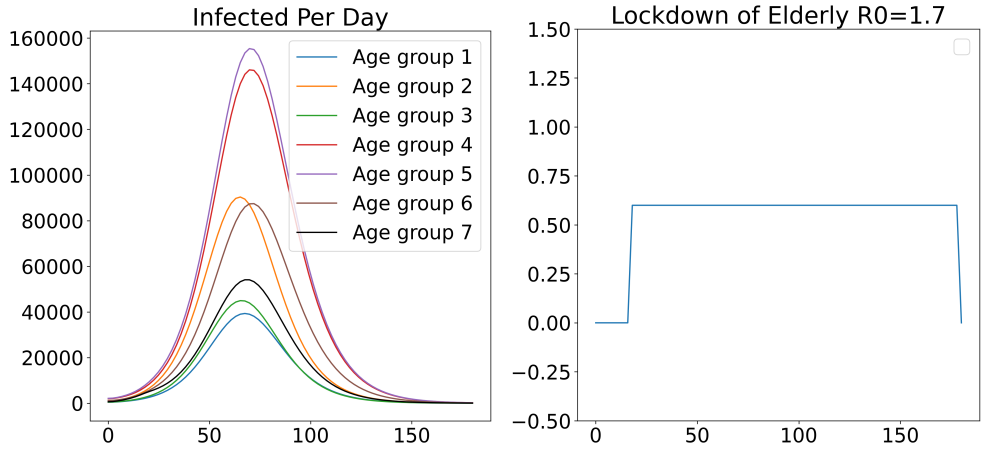

Figure 10. Dynamics of I (left) and optimal lockdown of the elderly (right). $R_0 = 1.7$, $\gamma = 0.2, \beta = 0.34, \delta = 0.25, \tau_3 = .8$. The maximum lockdown is 60% at any given time. The total cost between lockdown and deaths is 1.07×10^{11}.

Comparing the results from the school-only lockdown in Figure 9 with those of an elderly-only lockdown in Figure 10, we see an immensely more important role played by the elderly. While school closure reduces cost by 15%, a full lockdown of the elderly over the time horizon (the optimal decision) results in a 42% reduction in cost. Notice the steep reduction of slope in infected at the start of the lockdown period.

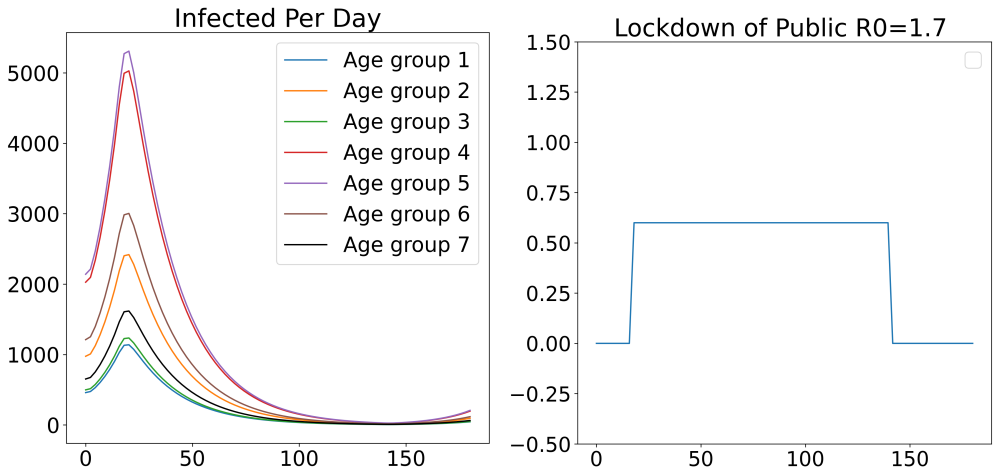

Figure 11. Dynamics of I (left) and optimal lockdown of a general lockdown (right). $R_0 = 1.7, \gamma = 0.2, \beta = 0.34, \delta = 0, \tau_3 = 0$. The maximum lockdown is 60% at any given time. The total cost between lockdown and deaths is 6.46×10^{10}.

Next, we look at a situation where there is only the opportunity to lock down the entire population to varying levels with no consideration of subpopulation targeting. We see from Figure 11 that a general lockdown provides even stronger protection. In reality, locking down the public results in

many unforeseen costs such as strain on small business and loss of education for children missing school. These results also suggest that, if given the opportunity, a focus should be on maximizing the possible lockdown of the elderly specifically, with this lockdown being almost the entire time horizon.

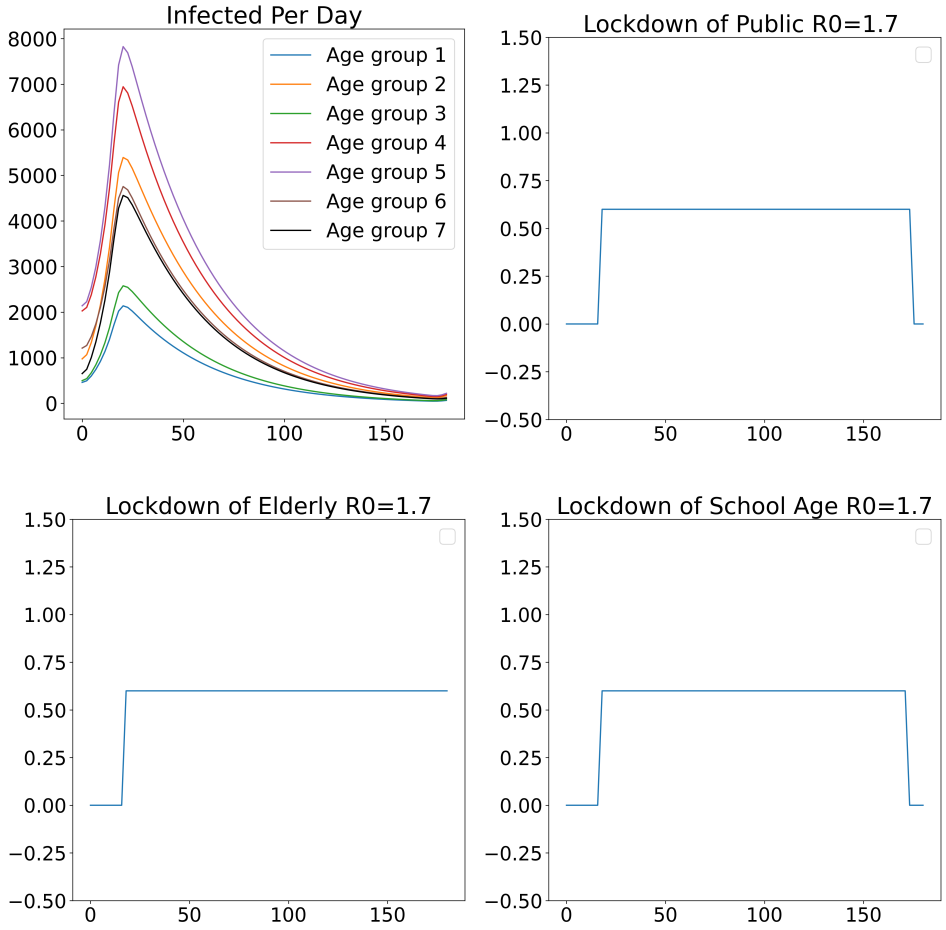

Figure 12. Dynamics of I (left) and optimal lockdown of all subpopulations (other panels). $R_0 = 1.7, \gamma = 0.2, \beta = 0.34, \delta = 0, \tau_2 = .66, \tau_3 = .8$. The maximum lockdown is 60% at any given time. The total cost between lockdown and deaths is 5.24×10^{10}.

Lastly, we see from Figure 12 that when given the ability to lock down various parts of the population separately, the cost is the lowest, resulting in a reduced cost of about 70%. While this is expected to be the best-case scenario, it highlights not only the power of having a flexible population lockdown but also the sheer importance of considering heterogeneity within a population.

Holding a population to a 60% lockdown across the board may be considered unrealistic, as, one could consider the ability to close schools to be 100%, general population maybe only around 60%, and the elderly somewhere in between since they can be supported in many ways by caretakers, say 80%. In Figure 13 we see that our model is able to handle such heterogeneity in order to better fit the population being modeled and the sociodemographic variables being considered.

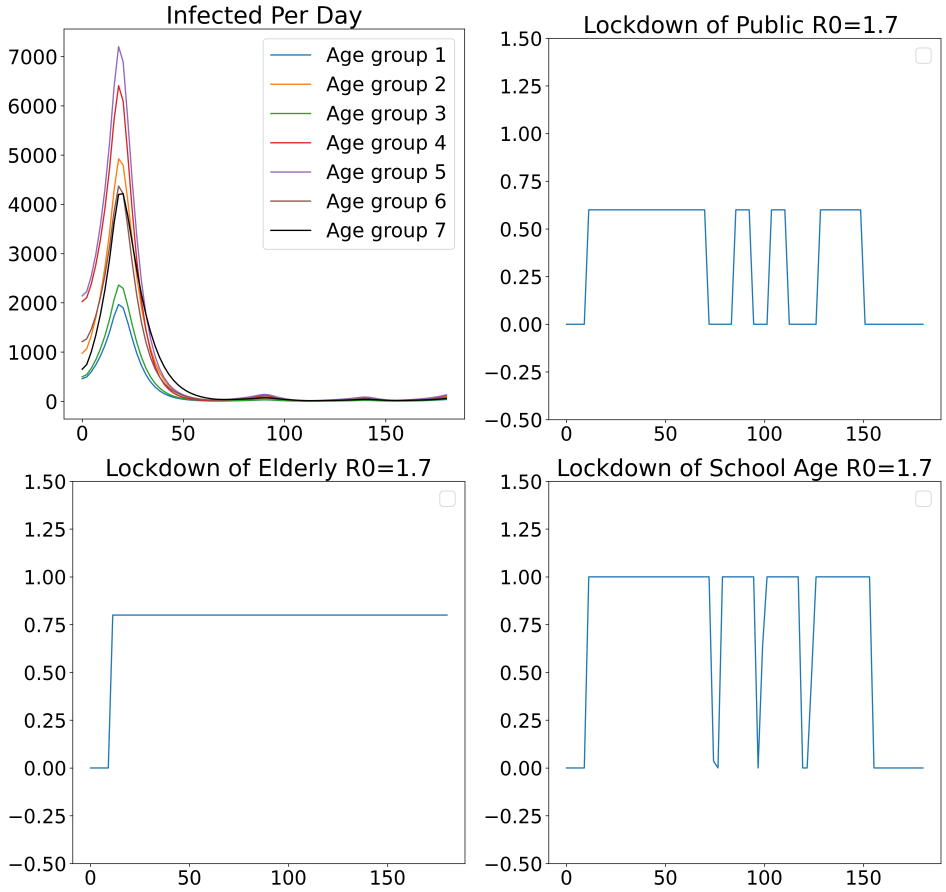

Figure 13. Dynamics of I (left)and optimal lockdown of all subpopulations (other panels). $R_0 = 1.7$, $\gamma = 0.2$, $\beta = 0.34$, $\delta = 0$, $\tau_2 = .66$, $\tau_3 = .8$. The maximum lockdown is 60% for general public, 80% for the elderly, and 100% for school children at any given time. The total cost between lockdown and deaths is 4.57×10^{10}.

5. Conclusion

In this paper, we introduce a method of combining classical epidemiological models with data about sociodemographic characteristics of a population to develop an information tool that maybe crucial to public policymakers. We choose data related specifically to age and interactions between age groups, and tune the model to census data of New Jersey. We then chose subsets of these interactions to be governed by different levels of lockdown policies, namely general interactions with the elderly, interactions most likely to occur in a school setting, and the remaining interactions after those two are considered. Costs are developed and used to guide an optimal lockdown schedule for these three strategies. We saw from preliminary COVID-19 data, that our model is strong at predicting deaths over time. We found overwhelmingly that when lockdowns must be employed, there is a heightened effect to lockdown not the most active interactions (school-age children), but the most vulnerable population (the elderly). Also, it seems that early intervention is key in curbing a virus from spreading and keeping the total mortality low. In fact, even a reduced lockdown (here 60%) greatly reduces mortality

if employed early enough. Lastly, a lockdown of the general public is crucial in curbing virus spread, but if given the opportunity to fit lockdown measures to specific sociodemographic groups, the cost to the population is further lessened.

Future work will focus on an advanced model able to handle multiple sociodemographic variables as well as multiple lockdown policies. Different lockdowns may have overlapping effects, thus variables and controls must be accurately defined in the development of the model itself. Other extensions will focus on including virus mutation dynamics as well as vaccination.

All script files are written in the Python programming language and all code produced here is placed on Github and made publicly available.

Author contributions

Weightman, Piccoli, and Akinode collaborated across various stages of the paper. In the conceptualization phase, Weightman and Piccoli contributed their expertise. Methodologically, Weightman, Akinode, and Piccoli worked together to develop research methodologies. During writing, Weightman and Piccoli drafted the original manuscript, while also engaging in review and editing alongside Akinode. Piccoli provided supervision throughout the project, and played a crucial role in securing funding.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors acknowledge the support of the Joseph and Loretta Lopez Chair endowment, the support of the NSF CMMI project # 2033580 “Managing pandemic by managing mobility” in collaboration with Cornell University and Vanderbilt University, Rutgers Global grant “Innovative models for pandemics with mutating viruses: mitigation strategies for vulnerable populations”, and the Institute for Advanced Study in Princeton for hosting last author.

Conflict of interest

Benedetto Piccoli is an editor-in-chief for [Networks and Heterogeneous Media] and was not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

References

1. K Aabed, M M.A. Lashin, An analytical study of the factors that influence Covid-19 spread, *Saudi J Biol Sci*, **28** (2021), 1177–1195. https://doi.org/10.1016/j.sjbs.2020.11.067
2. J A E Andersson, J Gillis, G Horn, J B Rawlings, M Diehl, Casadi–A software framework for nonlinear optimization and optimal control, *Math Program Comput*, **11** (2019), 1–36. https://doi.org/10.1007/s12532-018-0139-4

3. C Baunez, M Degoulet, S Luchini, P A Pintus, M Teschl, *An early assessment of curfew and second Covid-19 lock-down on virus propagation in france*, medRxiv [Preprint], (2020), [cited 2024 May 14]. Available from: https://doi.org/10.1101/2020.11.11.20230243
4. N Bellomo, R Bingham, M A J Chaplain, G Dosi, G Forni, D A Knopoff, et al., A multiscale model of virus pandemic: heterogeneous interactive entities in a globally connected world, *Math. Mod. Meth. Appl. S*, **30** (2020), 1591–1651. https://doi.org/10.1142/S0218202520500323
5. M Bicher, N Popper, Agent-based derivation of the sir-differential equations, *2013 8th EUROSIM Congress on Modelling and Simulation*, 2013, 306–311.
6. L Bolzoni, E Bonacini, C Soresina, M Groppi, Time-optimal control strategies in sir epidemic models, *Math Biosci*, **292** (2017), 86–96. https://doi.org/10.1016/j.mbs.2017.07.011
7. A Bressan, B Piccoli, *Introduction to the mathematical theory of control*, Springfield: American institute of mathematical sciences, 2007.
8. T Britton, F Ball, P Trapman, A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2, *Science*, **369** (2020), 846–849. https://doi.org/10.1126/science.abc6810
9. US Census Bureau, Survey of income and program participation (SIPP), 2022. https://www.census.gov/programs-surveys/sipp.html.
10. C Conover, How economists calculate the costs and benefits of Covid-19 lockdowns, Forbes, 2020. Available from: https://www.forbes.com/sites/theapothecary/2020/03/27/how-economists-calculate-the-costs-and-benefits-of-covid-19-lockdowns/?sh=44a20e846f63.
11. Y Chen, P Lu, C Chang, T Liu, A time-dependent SIR model for Covid-19 with undetectable infected persons, *IEEE Trans. Network Sci. Eng.*, **7** (2020), 3279–3294.
12. G Cho, Y J Kim, S Seo, G Jang, H Lee, Cost-effectiveness analysis of Covid-19 variants effects in an age-structured model, *Sci. Rep.*, **13** (2023), 15844. https://doi.org/10.1038/s41598-023-41876-x
13. M Chyba, T Klotz, Y Mileyko, C Shanbrom, A look at endemic equilibria of compartmental epidemiological models and model control via vaccination and mitigation, *Math. Control Signals Syst.*, (2023), 1–31. https://doi.org/10.1007/s00498-023-00365-2
14. A T Ciota, L D Kramer, Vector-virus interactions and transmission dynamics of west nile virus, *Viruses*, **5** (2013), 3021–3047. https://doi.org/10.3390/v5123021
15. K Dietz, The estimation of the basic reproduction number for infectious diseases, *Stat Methods Med Res*, **2** (1993), 23–41.
16. R Donnelly, H A Patrinos, Learning loss during Covid-19: An early systematic review, *Prospects*, **51** (2021), 601–609. https://doi.org/10.1007/s11125-021-09582-6
17. E Dorn, B Hancock, J Sarakatsannis, E Viruleg, Covid-19 and learning loss—disparities grow and students need help. McKinsey & Company, 2020. Available from: https://www.mckinsey.com/industries/public-sector/our-insights/covid-19-and-learning-loss-disparities-grow-and-students-need-help.
18. D Fanelli, F Piazza, Analysis and forecast of Covid-19 spreading in china, Italy and france, *Chaos Solitons Fractals*, **134** (2020), 109761. https://doi.org/10.1016/j.chaos.2020.109761

19. A S Fauci, The aids epidemic—considerations for the 21st century, *New Engl J Med.*, **341** (1999), 1046–1050. https://doi.org/10.1056/NEJM199909303411406

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

21. L. O Gostin, D Lucey, A Phelan, The ebola epidemic: a global health emergency, *JAMA*, **312** (2014), 1095–1096.

22. C. Gunaratne, R. Reyes, E. Hemberg, UM O'Reilly, Evaluating efficacy of indoor non-pharmaceutical interventions against Covid-19 outbreaks with a coupled spatial-SIR agent-based simulation framework, *Sci. Rep.*, **12** (2022), 6202. https://doi.org/10.1038/s41598-022-09942-y

23. S Hammerstein, C König, T Dreisörner, A Frey, Effects of Covid-19-related school closures on student achievement-a systematic review, *Front. Psychol.*, **12** (2021), 746289. https://doi.org/10.3389/fpsyg.2021.746289

24. J K Hammitt, Valuing mortality risk in the time of Covid-19, *J Risk Uncertain*, **61** (2020),129–154. https://doi.org/10.1007/s11166-020-09338-1

25. X Huang, X Shao, L Xing, Y Hu, D D. Sin, X Zhang, The impact of lockdown timing on Covid-19 transmission across us counties, *EClinicalMedicine*, **38** (2021), 101035. https://doi.org/10.1016/j.eclinm.2021.101035

26. V Kala, K Guo, E Swantek, A Tong, M. Chyba, Y Mileyko, et al., Pandemics in hawaii: 1918 influenza and Covid-19, *The Ninth International Conference on Global Health Challenges*, 2020.

27. W. O. Kermack, A. G. McKendrick, Contributions to the mathematical theory of epidemics. II.—The problem of endemicity, *Proc. R. Soc. London, Ser. A*, **138** (1932), 55–83. https://doi.org/10.1098/rspa.1932.0171

28. M. L. King, M. Tertilt, IPUMS-CPS: An integrated version of the march current population survey, 1962–2002, *Hist. Methods*, **36** (2003), 35–40. https://doi.org/10.1080/01615440309601213

29. J. M. Kirigia, R. N. D. K. Muthuri, The fiscal value of human lives lost from coronavirus disease (Covid-19) in China, *BMC Res. Notes*, **13** (2020), 198. https://doi.org/10.1186/s13104-020-05044-y

30. N. Kousar, R. Mahmood, M. Ghalib, A numerical study of SIR epidemic model, *Int. J. Sci.: Basic Appl. Res.*, **25** (2016), 354–363.

31. A. J. Krener, H. Schättler, The structure of small-time reachable sets in low dimensions, *SIAM J. Control Optim.*, **27** (1989), 120–147. https://doi.org/10.1137/0327008

32. J. S. Kutter, M. I. Spronken, P. L. Fraaij, R. A. M. Fouchier, S. Herfst, Transmission routes of respiratory viruses among humans, *Curr. Opin. Virol.*, **28** (2018), 142–151. https://doi.org/10.1016/j.coviro.2018.01.001

33. K. Y. Leung, P. Trapman, T. Britton, Who is the infector? Epidemic models with symptomatic and asymptomatic cases, *Math. Biosci.*, **301** (2018), 190–198. https://doi.org/10.1016/j.mbs.2018.04.002

34. C. Lobry, Contro labilite des syste mes non lineaires, *SIAM J. Control Optim.*, **8** (1970), 573. https://doi.org/10.1137/0308042

35. Q. Luo, R. Weightman, S. T. McQuade, M. DÍaz, E. Trélat, W. Barbour, et al., Optimization of vaccination for Covid-19 in the midst of a pandemic, *Netw. Heterog. Media*, **17** (2022), 443–466. https://doi.org/10.3934/nhm.2022016

36. S. T. McQuade, R. Weightman, N. J. Merrill, A. Yadav, E. Trélat, S. R. Allred, et al., Control of Covid-19 outbreak using an extended seir model, *Math. Mod. Meth. Appl. S.*, **31** (2021), 2399–2424. https://doi.org/10.1142/S0218202521500512

37. M. Miller, 2019 novel coronavirus Covid-19 (2019-nCoV) data repository: Johns hopkins university center for systems science and engineering, *Assoc. Can. Map Lib. Arch. Bull.*, **164** (2020), 47–51. https://doi.org/10.15353/acmla.n164.1730

38. B. W. Mol, J. Karnon, Strict lockdown versus flexible social distance strategy for Covid-19 disease: A cost-effectiveness analysis, *Arch. Clin. Biomed. Res.*, **7** (2023), 58–63. https://doi.org/10.26502/acbr.50170319

39. Centers for Disease Control and Prevention, Weekly updates by select demographic and geographic characteristics, National Center for Health Statistics, 2021. Available from: file:///C:/Users/mabin/Desktop/cdc_105341_DS1.pdf.

40. World Health Organization, Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV), 2020. Available from: https://www.who.int/zh/news/item/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).

41. K. Prem, A. R. Cook, M. Jit, Projecting social contact matrices in 152 countries using contact surveys and demographic data, *PLoS Comput. Biol.*, **13** (2017), e1005697. https://doi.org/10.1371/journal.pcbi.1005697

42. H. E. Rodriguez, Collecting Covid-19: Documenting the CDC response, *Collections*, **17** (2021), 102–111. https://doi.org/10.1177/1550190620980411

43. M. Roy, R. D. Holt, Effects of predation on host–pathogen dynamics in SIR models, *Theor. Popul. Biol.*, **73** (2008), 319–331. https://doi.org/10.1016/j.tpb.2007.12.008

44. C. A. Stafford, G. P. Walker, D. E. Ullman, Hitching a ride: Vector feeding and virus transmission, *Commun. Integr. Biol.*, **5** (2012), 43–49. https://doi.org/10.4161/cib.18640

45. R. M. Viner, S. J. Russell, H. Croker, J. Packer, J. Ward, C. Stansfield, et al., School closure and management practices during coronavirus outbreaks including Covid-19: A rapid systematic review, *Lancet Child Adolesc.*, **4** (2020), 397–404. https://doi.org/10.1016/S2352-4642(20)30095-X

46. A. Wächter, L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Program.*, **106** (2006), 25–57. https://doi.org/10.1007/s10107-004-0559-y

47. S. H. Waterman, D. J. Gubler, Dengue fever, *Clin. Dermatol.*, **7** (1989), 117–122. https://doi.org/10.1016/0738-081x(89)90034-5

48. R. Weightman, Age based control github repository, PLabCOVIDModeling, 2023. Available from:https://t.ly/Ms8C_.

49. R. Weightman, B. Piccoli, Optimization of non-pharmaceutical interventions for a mutating virus, *2023 American Control Conference (ACC)*, San Diego: IEEE, (2023), 307–312. https://doi.org/10.23919/ACC55779.2023.10156293

50. Q. Wen, J. Yang, T. Luo, First case of Covid-19 in the united states, *N. Engl. J. Med.*, **382** (2020), e53. https://doi.org/10.1056/NEJMc2004794

51. Worldometer, Cases and deaths from Covid-19 virus pandemic, 2020. Available from:https://www.worldometers.info/coronavirus/country/us/.

52. O. Zakary, M. Rachik, I. Elmouki, On the impact of awareness programs in HIV/AIDS prevention: An SIR model with optimal control, *Int. J. Comput. Appl.*, **133** (2016), 1–6. https://doi.org/10.5120/ijca2016908030

53. J. Zhang, M. Litvinova, W. Wang, Y. Wang, X. Deng, X. Chen, et al., Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: A descriptive and modelling study, *Lancet Infect. Dis.*, **20** (2020), 793–802. https://doi.org/10.1016/S1473-3099(20)30230-9

Appendix A: Data, Parameters, and Contact Matrix

Each simulation reported in Section 4 uses the SEIR parameter set in Table A1 based on data collected during the Covid-19 pandemic.

Table A1. Parameters for the SEIR model (2.3).

Name	Description	Estimate	Units
β	rate of infection	0.2–0.5	-
δ	latent period	0.25	days
γ	infectious period	0.2	days
τ_i	isolated interactions	0–100	percent

To calibrate our model, we choose an age structure to drive the interactions and a lockdown as our mitigation strategy. We begin by partitioning our population into the seven age groups in Table A2.

Table A2. Age brackets defining age groups.

Name	Description
Group 1	Age 0–4 population
Group 2	Age 5–14 population
Group 3	Age 15–19 population
Group 4	Age 20–39 population
Group 5	Age 40–59 population
Group 6	Age 60–69 population
Group 7	Age 70+ population

We partition the total population into these age groups based on percentages provided by the United States census [9]. For each piece of information captured in a SEIR model, initial conditions and parameters must be estimated. The Johns Hopkins University (JHU) repository of Covid-19 data collected case and death numbers for each state in the US carefully from the start of the pandemic until early 2023 [37]. We then choose a starting date and use the JHU data to get the number of infected. From this, we build initial conditions for our compartments.

The number of exposed is set equal to a certain percent of the infected from the first 5 days of the model (December 16–20 2019). Then, the number of infected is taken to be the number of cases from the first 5 days of the model minus exposure assigned proportionally based on the number of people in each age group. We consider the number of recovered to be the sum of cases reported from the start of reporting until two weeks before the chosen start date. Lastly, we define the number of susceptibles to be the sum of the population in each group minus the sum of the other compartments.

We take the general SEIR model and couple it with an age-based interaction matrix. Our interaction matrix l is developed starting from data from [41], which takes data from population-based contact diaries in eight countries and projects to various other countries using a Bayesian model. The contact matrix is actually inferred data for the United States. The age groups in the original contact matrix provided by the authors contains 16 age groups at a higher resolution than we include here, so averages are taken where two of the original age groups make up our age group. For example, the second element of l is 0.8003 which makes up two of the age groups in the original dataset; the 5–9 element being 1.1013 and the 9–14 element being 0.4994, the average of which is 0.8003. Notice also that this matrix is not symmetric. The data is used to develop vectors for each age group which represent their interactions at home, interactions in the workplace, interactions at school, etc. Therefore, a phenomenon occurs where, for example, the 70+ age group, in their daily interactions, are very likely to interact with someone of the 40–59 age group, as these are often caretakers, children, or employees at consumer establishments. So, their element of interacting with 40–59 is high. However, relative to other interactions, 40–59 age group members are unlikely to interact with the 70+ age group the most, resulting in their relative interactions with 70+ being much lower. It is worth noting that while we used only data from a more recent source, the contact matrices found in other papers are very similar in structure to the ones we use [8].

$$l = \begin{bmatrix} 2.5982 & 0.8003 & 0.3160 & 0.7934 & 0.3557 & 0.1548 & 0.0564 \\ 0.6473 & 4.1960 & 0.6603 & 0.5901 & 0.4665 & 0.1238 & 0.0515 \\ 0.1737 & 1.7500 & 11.1061 & 0.9782 & 0.7263 & 0.0815 & 0.0273 \\ 0.5504 & 0.5906 & 1.2004 & 1.8813 & 0.9165 & 0.1370 & 0.0397 \\ 0.3894 & 0.7848 & 1.3139 & 1.1414 & 1.3347 & 0.2260 & 0.0692 \\ 0.3610 & 0.3918 & 0.3738 & 0.5248 & 0.5140 & 0.7072 & 0.1469 \\ 0.1588 & 0.3367 & 0.3406 & 0.2286 & 0.3673 & 0.3392 & 0.3868 \end{bmatrix}$$

We then use matrix l coupled with general information about the population of the United States of America to separate our subpopulations l^i.

Death rates are computed from CDC Covid-19 data, but follow the general pattern of many viruses where as age increases, as does death rate [39]. They are as follows in order of our age groups: 0.00016, 0.00016, 0.00006, 0.00007, 0.00229, 0.01915, 0.13527.

In the end, we would like to use optimal control to answer the question of which sub-groups lockdown would be most impactful in curbing the effect a virus as on a population. As such, the cost function must be developed carefully to reflect the population in question and the age categories being locked down. To achieve this, we define the cost per child per day of a school lockdown by averaging estimates provided by the Survey of Income and Program Participation and the Current Population Survey [9, 28]. This number fell in the range of as low as \$25 per day to as high as \$60 per day. It is important to note that this does not include the extremely high potential reduction of education. For the cost of locking down the elderly populations, we consider the same cost as locking down another adult person, which we find to be \$40–\$350. This number has a large range, with some reporting a cost of as much as \$350 per day for a strict lockdown and much less for a less severe, more flexible lockdown [12, 38].

Lastly, choosing parameters such as costs for a cost function requires quantifying the "cost of death" in an economic sense. We follow the estimates of one analyst who claims \$1.5 million dollars to be the average cost of a Covid-19 death in the United States [10]. All codes used to generate simulations are available at [48].

AIMS Press