EPIDEMIOLOGY

Balancing economic and epidemiological interventions in the early stages of pathogen emergence

Andy Dobson[1,2,3]*, Cristiano Ricci[4], Raouf Boucekkine[5], Fausto Gozzi[6], Giorgio Fabbri[7], Ted Loch-Temzelides[8], Mercedes Pascual[2,9]

The global pandemic of COVID-19 has underlined the need for more coordinated responses to emergent pathogens. These responses need to balance epidemic control in ways that concomitantly minimize hospitalizations and economic damages. We develop a hybrid economic-epidemiological modeling framework that allows us to examine the interaction between economic and health impacts over the first period of pathogen emergence when lockdown, testing, and isolation are the only means of containing the epidemic. This operational mathematical setting allows us to determine the optimal policy interventions under a variety of scenarios that might prevail in the first period of a large-scale epidemic outbreak. Combining testing with isolation emerges as a more effective policy than lockdowns, substantially reducing deaths and the number of infected hosts, at lower economic cost. If a lockdown is put in place early in the course of the epidemic, it always dominates the "laissez-faire" policy of doing nothing.

INTRODUCTION

Coronavirus disease 2019 (COVID-19) has now infected nearly half a billion people globally and has led to more than 6 million fatalities, although the true total may be closer to 20 million [see (*1*)]. Part of the economic damages associated with the pandemic stem directly from people's inability to work when sick. A substantial proportion is also driven by "lockdown" policies put in place to minimize disease transmission through contact between infected hosts and potential susceptible hosts. A template for initial response to future pandemics is needed that will avoid the mistakes and ambiguities that occurred in the initial response to COVID-19. This is particularly important, as at least 20 pathogens have emerged as threats to human populations over the past 50 years [see (*2*)]. In the 1- to 2-year period before vaccines become widely available to control a novel pathogen, the only public health measures to control outbreaks are isolation of susceptible hosts and testing for infection once tests have been developed and distributed. Both of these measures can generate nontrivial financial costs associated with a reduction in employment and income. Policy-makers have the task of determining the best way to offset these costs against those caused by sickness and possibly death of infected hosts.

There are increasing calls for better foresight of how to respond to future pandemics (*3*, *4*). In a recent article, Persad and Pandya (*5*) argue that combining epidemiological studies with economic cost-benefit analysis is essential for an effective policy response. Policy-makers must be equipped with tools to rigorously compare various interventions, both when evaluating individual policies and when determining which policies to include in a regulatory package. They argue that COVID-19–related policy decisions require considering not only trade-offs between health outcomes and the direct costs of interventions but also additional dimensions related to economic activity, social justice, and individual liberty. Decision-making always uses some type of mental model to weigh the pros and cons of different policy options. Rigorous economic evaluation formalizes this process. Value judgments will still be present, but economic evaluation can make the decision-making process more systematic, comprehensive, and transparent. A particularly relevant quote in (*5*) states: "Critics of this type of approach might argue that in the midst of a pandemic that is still killing thousands of people globally every day, we don't have time to engage in economic evaluation — that we should do the best we can, without fully weighing the costs and benefits of the options under consideration. In contrast, we believe the severity of the pandemic makes the need for evaluation all the more urgent. Choosing optimal interventions is associated with a bigger payoff when risks are higher." Economic evaluations of COVID-19–related policies must consider nonlinear effects, as policies might have different results in combination than they do independently.

We have developed a model that combines existing, widely used economic models and an expansion of the SEIR (susceptible-exposed-infected-recovered) epidemiological framework. The resultant hybrid epi-econ model introduces novel features and allows us to develop optimization methods to concomitantly minimize deaths and damages to the economy. Our model is innovative in two ways: First, the expanded SEIR model accounts for a larger set of possible states for the hosts in an attempt to consider the dynamics of isolating and testing contacts of infected hosts, which may themselves develop infections. Second, we take into account that realistic epidemic control policies are subject to inefficiencies resulting from "economic frictions" inherent in the implementation of such policies. A partial list of these frictions includes incomplete information, transactions costs associated with initiating multiple rounds of lockdowns in rapid succession, incomplete enforcement, and costs associated with transitions in and out of lockdown. By

[1]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [2]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA. [3]Smithsonian Tropical Research Institute, Panama City, Panama. [4]Department of Economics and Management, Università di Pisa, Pisa, Italy. [5]Centre for Unframed Thinking, Rennes School of Business, 35000 Rennes, France. [6]Department of Economics and Finance, Luiss University, Roma, Italy. [7]Université Grenoble Alpes, CNRS, INRAE, Grenoble INP, GAEL, 38000 Grenoble, France. [8]Department of Economics and Baker Institute for Public Policy, Rice University, 6100 Main St., Houston, TX 77005, USA. [9]Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA.
*Corresponding author. Email: dobson@princeton.edu

improving the realism of both epidemic dynamics and policy modeling, we can better understand the structure of the mechanisms through which public health and economic factors interact.

Before the current health crisis, economic research on epidemics has been limited and has mostly attempted to connect epidemics with economic development in the global South [see, for example, (*6*, *7*)]. One of the first substantial contributions connecting economics and epidemic compartmental modeling (hereafter, epi-econ modeling) traces back to Gersovitz and Hammer in 2004 (*8*). The current COVID-19 crisis has stimulated additional research at this intersection. Most studies in the area use standard compartmental epidemic models ranging from SIS to SEIRAD ("A" stands for asymptomatic) and typically investigate appropriate controls to target specific aspects of the COVID-19 crisis. Examples include (*9*–*13*). The policies examined in these papers tend to focus on lockdowns as a control strategy, in line with (*9*). Testing and/or social distancing are less frequently addressed [see (*10*, *14*)]. These interventions are often modeled in a stylized way [for example, lockdown controls are assumed to be continuous-time control variables; Aspri *et al.* (*15*) provide a notable exception]. They suggest that targeted isolation avoids the sharp economic decline of lockdown by creating incentives for infectious individuals to isolate while allowing unexposed individuals to continue to consume and work. Ash *et al.* (*16*) calibrate a dynamic economic model to COVID-19–related epidemiological data and use it to evaluate the effects of different scenarios, including voluntary isolation, targeted isolation, and blanket lockdowns.

Unlike the majority of the recent epi-econ literature, we use a finite-horizon model to concentrate on short-term outcomes. More precisely, we build a framework that provides insights on how policy-makers should respond in the first 2 years of a novel emergent pathogen for which (i) there is very limited epidemiological information, (ii) there are no available specific drugs or vaccines, and (iii) tests for infectivity are in the early stages of development. Given these constraints, a lockdown is one of the main tools available to policy-makers. We specifically depart here from the common assumption that lockdown policies can be adjusted in continuous time and consider that they take place in a finite number of phases, the lockdown parameters (intensity of the lockdown and duration) being optimally chosen at each phase. In addition, we incorporate technological implementation delays (e.g., in efficient testing) and capacity constraints (e.g., in test/mask production). This, in turn, allows us to compare optimal control policies for countries at different levels of development or governments with different levels of concern for the welfare of their citizens/workforce.

RESULTS

Our model has two main components: an epidemiological submodel and an economic one. They are coupled together by the population of hosts who can either work or are restricted from working to full capacity by exposure to infected hosts or by being unable to work when ill and incapacitated. In both cases, we will assume an underlying well-mixed "mean-field" structure, a framework that has been widely used in both epidemiology and economics. The flow diagram and state variables of the epidemiological model are described in Fig. 1. The corresponding equations are given in Methods.

The economic analysis is based on a standard intertemporal production model (Methods). We postulate that a policy-maker solves an "optimal control problem." They maximize a function representing the society's well-being or "welfare" over time, subject to the set of differential equations describing the evolution of the epidemiological variables. The welfare function is a linear combination of economic and epidemiological objectives.

The policy-maker's objective is to optimally balance between two components. The first involves maximizing economic well-being associated with the flow consumption of goods and services produced in the economy. The second is to minimize the society's direct costs associated with deaths resulting from the pathogen. In addition to reducing output as a result of declining labor supply, deaths impose several other costs to society. Our modeling of these costs is meant to capture the intrinsic value of lives lost, as well as the resulting social and psychological effects on families and affected communities more broadly. As is standard in economic models, we assume that utility from consumption increases at a decreasing rate as consumption increases. The power function specification is also standard; it features constant elasticity of substitution in consumption over time, which captures the willingness of the policy-maker to switch consumption/production over time. In addition, we assume that disutility from deaths increases at an increasing rate with the number of deaths. Last, we use the parameter θ to weight the importance of deaths relative to consumption of goods and services in the evaluation of the policy-maker. At one extreme, $\theta = 0$ corresponds to a purely economic model, where costs from lives lost are not directly taken into consideration (they reduce economic well-being indirectly but only as a result of lost production). As θ increases, the epidemiological objective (lives saved) increases in importance. Different values of θ can thus trace a Pareto frontier between economic performance and lives lost.

We use the model to examine the concomitant response of the economy and the pathogen to three different nonpharmaceutical interventions (NPIs) that are likely to be available before a vaccine or antiviral drugs can be developed. Until that time, we assume that the only interventions available are (A) a lockdown of a fraction of the economy, (B) isolation of contacts of people known to be infected, and (C) wearing of surgical masks. In what follows, we initially ignore mask-wearing and assume that this activity is a voluntary adjunct to A and B. We discuss this further in the conclusions (and the Supplementary Materials). We assume that tests will be developed quite quickly that allow identification of infected people before symptoms appear. This is particularly important for COVID-19, when substantial levels of transmission are undertaken by hosts who do not yet show symptoms. The efficacy of these tests will also improve as they are more widely used. We can modify our model to include improving levels of specificity and sensitivity in the accuracy of tests used to identify recently infected hosts. Here, we simplify by assuming average recorded values for specificity and sensitivity for the tests used for COVID-19 (*17*, *18*).

Transient dynamics under different epidemic control policies

In this section, we study the dynamics generated by our model under four possible scenarios: The first benchmark that we consider is to do nothing to control the epidemic (case 0 below). This "laissez-faire" policy is always available to policy-makers, particularly if the costs of intervention are high. This essentially represents the

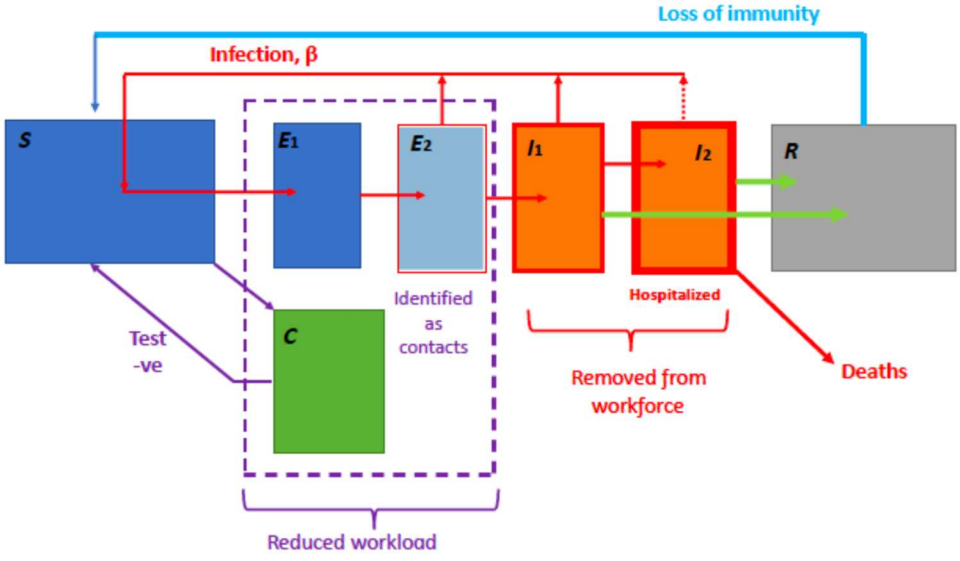

Fig. 1. Flow diagram of the epidemiological components of the model described in Results. We have modified the basic SEIR formulation by dividing the exposed (E) and infectious (I) classes into two sequential classes, E_1 and E_2 and I_1 and I_2. Exposed hosts, who are not yet infectious are classified as E_1, while asymptomatic, contagious hosts are classified as E_2. We assume that E_1 individuals transform to E_2 at an exponential rate determined by ϕ_1. The presymptomatic hosts, E_2, transform to symptomatic infected hosts, I_1, at a rate ϕ_2. Both E_2 and I_1 are infectious. This rate largely determines the duration of time during which exposed hosts are able to transmit infection before they show symptoms of infection. If ϕ_2 is large (~365; around 1 day), then exposed hosts quickly exhibit signs of symptoms and can be identified as infectious (as occurred with SARS). In contrast, if ϕ_2 is slower (~365/7; a week), then asymptomatic hosts may transmit the disease for up to a week before showing symptoms, as in the case of COVID-19 (or many years in the case of HIV or tuberculosis, when ϕ_2 may range from 0.1 to 0.5). In a similar way, infected hosts, I_1, may become sick and get hospitalized, I_2. These hosts have a higher mortality rate but are assumed to be in relative isolation and are thus unable to transmit the pathogen, except to unprotected health care workers. The majority of the pathogen-induced mortality occurs in the I_2 class. We also include an additional class, C, into our model structure; these are contacts of infectious hosts who do not develop infection. Contact tracing identifies $C + E_1 + E_2$ as contacts of infected hosts; testing is used to differentiate uninfected contacts, C, from exposed hosts (E_1 and E_2); the former can return to work, and the latter remain in isolation and go on to develop infection.

policies pursued against COVID-19 in Sweden and Tanzania. The second case (case 1) is one in which the policy-maker responds to the presence of the pathogen only with a lockdown policy (sensu China and North Korea). We characterize the optimal duration and intensity of the lockdown in this context and discuss its impact on economic dynamics. Symmetrically, the following sections (cases 2 and 3) are devoted to the case where the policy-maker copes with the epidemic only by testing individuals; this can be done at random or by focusing on those who have been in contact with hosts who have developed symptoms that progress to illness. Last, we investigate the case where, after some initial delay, the policy-maker may use both tools at their disposal (case 4). In all scenarios, we characterize the optimal sequential combination of lockdown followed by testing.

We first consider the case where no formal attempt is made to initiate a lockdown of the economy, nor to restrict interactions between infected and susceptible hosts, or to isolate those who have contacted the pathogen and those who have developed symptoms. For this scenario, we also assume that no tests are available. Such a laissez-faire scenario could, for example, be favored by those entities whose economic interests would be most severely affected by reductions in economic activity under lockdowns. We use this case as a benchmark for our initial examination of the model's dynamics, as it provides a comparison with the different forms of intervention that we introduce later. This case also allows us to identify the conditions under which minimal policy intervention might be desirable from the perspective of maximizing social welfare.

One view that has been put forward in support of the laissez-faire case is that the pathogen will generate levels of "herd immunity" among survivors of infection that will slow the further spread of infections while allowing uninfected hosts to continue to operate in the economy. These arguments assume relatively low rates of mortality and prolonged periods of immunity (*19*–*22*). Crucially, loss of immunity over time always leads to a resurgence of the epidemic. We explore this possibility by allowing immunity to wane at four different average rates: 6 months and 1, 2, and 3 years. As we are only focusing on the first 3 years of the epidemic, the slowest rate of loss of immunity is essentially equivalent to the classical studies of measles, which assume life-long immunity and form the basis of most SEIR modeling approaches.

Case 0: The laissez-faire

In addition to providing a benchmark, this scenario also quantifies the economic damages that result from reduction of the workforce due to illness and through loss of life (Fig. 2). In all cases, we emphasize that the pathogen does not die off but continues to generate more deaths and a continuous reduction in economic activity. It is worth pointing out that our laissez-faire scenario is restrictive in the sense that it abstracts from people's endogenous decisions to change their behavior, for example, by isolating or wearing facemasks. There is increasing evidence from epidemiological models that such changes in behavior do have a substantial impact on reducing rates of pathogen transmission, particularly when these are driven by reduced levels of aggregation in the host population (*27*–*30*). The large differences observed between different countries in the early

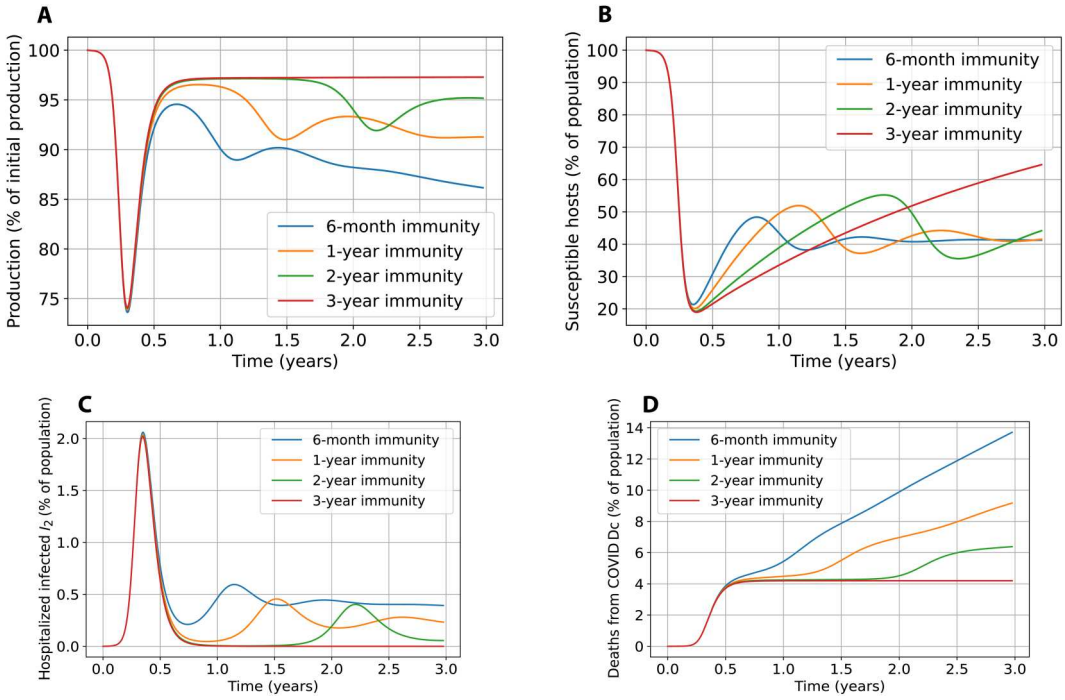

Fig. 2. The laissez-faire case: Impact of duration of immunity on epidemic dynamics and the economy in the absence of intervention. (**A**) Impact of the pathogen on the economy for different levels of duration of immunity. (**B**) Reduction in the numbers of susceptible hosts in the population. The lines are colored to reflect different rates of loss of immunity (blue, 6 months; orange, 1 year; green, 2 year; red, 3 years). (**C**) The numbers of hospitalized patients. (**D**) Time course of utility loss due to deaths during the course of the epidemic under four different strategies. Values of the economic and epidemiological parameters are described in Table 1.

months of the COVID-19 pandemic reflect differences in both government-mandated and individual choice behaviors (*28*, *31*).

The rate at which immunity is lost has profound effects on both the dynamics of the epidemic and its influence on the economy. The upper two graphs illustrate the impact on the economy and on the number of susceptible hosts as sequential waves of infection pass through the population (Fig. 2). If immunity lasts for 3 years, then the epidemic is experienced as a short outbreak that peaks when herd immunity is attained. As the duration of immunity shortens, it becomes progressively easier for the pathogen to reassert itself, causing repeated waves of infection resulting in corresponding sequential impacts on the economy. Deaths initially peak at around 4% of the population. They continue to rise after the short duration of immunity provided by transient levels of protection begins to diminish and those who have recovered from infection become susceptible and potentially reinfected. The deaths are matched by a prolonged 4% reduction in economic output, as births have not yet had time to enter the workforce. Notice that loss of economic productivity continues in repeated waves, as each cohort loses their immunity, become sick, and transiently leave the workplace. When immunity lasts for less than a year, economic productivity experiences a steady long-term decline. We have assumed that those who receive second or third bouts of infection have no residual immunity to infection after their immunity has waned. This is pessimistic in that it is likely that recovered hosts do receive some residual immunological benefits from prior infection, but incorporating these would require additional stages to our model that acknowledge a second and third class of resistant host.

Case 1: Lockdowns

Next, we consider the case where the policy-maker responds to the presence of the pathogen by initiating a lockdown that closes a proportion of the economy for a sequence of time intervals, eventually leading to the eradication of the pathogen from the local population. Our optimization algorithm minimizes economic losses by setting the level and duration of the lockdown, in principle for up to "*n*" distinct time intervals. In each case, we characterize the combined optimal duration and depth required to maximize the total social welfare functional.

Policy-makers have to make crucial decisions at the beginning of an outbreak. These decisions trade off the political expediency of being seen to act promptly against the cost of slowing down economic activity, which might expose them to claims of overreaction, particularly if reports of an epidemic run the risk of being false alarms.

We assume that 2 weeks after the epidemic emergence is the fastest time when lockdowns can be put in place. This allows for identification of sufficient initial cases before it is concluded that something needs to be done, and a lockdown is the only possible response, as no treatments or vaccines are available in this scenario. We examine the economic and human costs resulting from more extended delays (Fig. 3). We do this by starting the lockdown process at sequentially later dates: two, 4, 8, and 16 weeks after the initial detection of transmission.

The first main message emerging from this analysis is quite intuitive: The longer the delay in initiating lockdown, the deeper its impact on the economy. However, the lockdown does not necessarily have to last longer, particularly when the delay extends toward

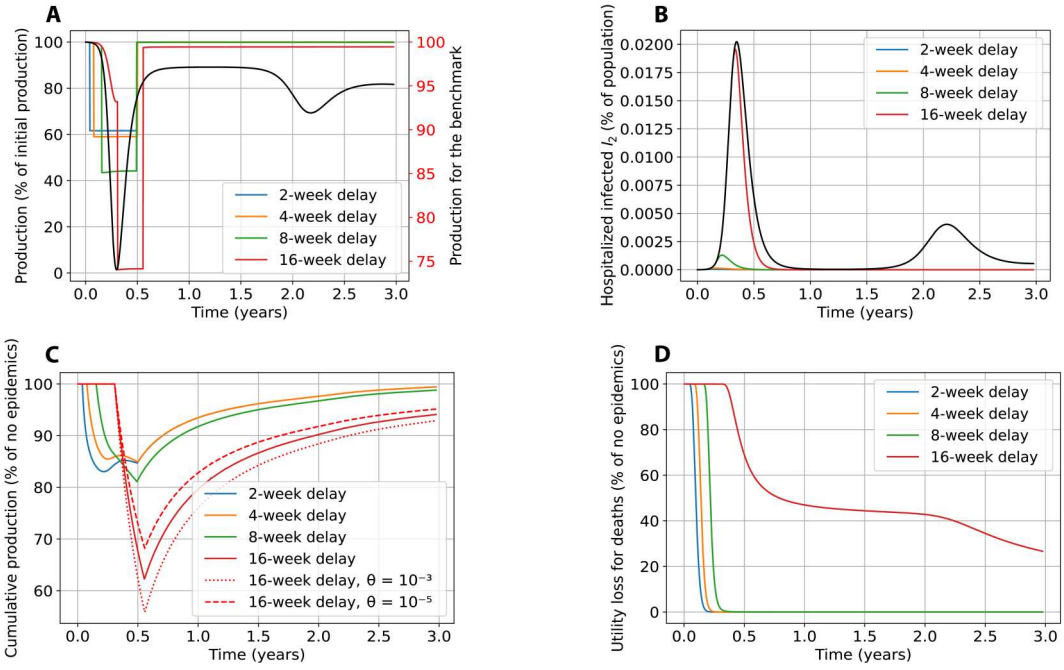

Fig. 3. Lockdown without testing. Depth and duration of optimal lockdown and the related impact on the economy and on the epidemic of varying the delay in the policy response: two weeks (blue), 4 weeks (orange), 8 weeks (green), and 16 weeks (red). In black, we illustrate the benchmark laissez-faire case (calibrated on the right-hand axis). The figures show (**A**) the optimal duration and intensity of the lockdown, (**B**) the related flow of hospitalizations, (**C**) the cumulative production (from time 0 to each period) in terms of the cumulative production in an epidemic-free dynamics, and (**D**) the utility loss (from time 0 to each period) due to epidemic-related deaths weighted by the parameter θ. The dashed and dotted lines around the 16-week delay illustrate the sensitivity of lockdown to θ, the index that parameterizes the value of human life. The values of the economic and epidemiological parameters are described in Table 1.

the natural underlying first peak of the epidemic. In practice, it is not easy to know a priori when an epidemic wave of infections will peak. A second important finding is that lockdown levels do not have to match the levels of herd immunity. Eradication can be achieved by a 33 to 36% reduction in economic activity for a period of up to 10 months. The relative insensitivity of the duration of the lockdown is not matched by differences in initiating a lockdown: Eight times as many people are hospitalized, and five times as many people die if the lockdown is delayed from 2 to 4 weeks. Delay always leads to more deaths up until the time when a lockdown corresponds to the time when the epidemic has peaked. With delays of this duration, transient levels of herd immunity in the recovered section of the population reduce the number of future deaths but not the economic cost of achieving eradication. Perhaps a more relevant way to measure the relative efficiency of a lockdown policy with respect to laissez-faire is to compare cumulative costs over the whole (3-year-long) period. Figure 3 (C and D) points to some important related factors. First, while the lockdown policy is outperformed in terms of production by the laissez-faire after 1 year (from slightly more than 5% below in the 2-week delay scenario to 20% in the 16-week delay case), this gap tends to vanish at the end of the 3-year-long period (with the notable exception of the 16-week delay scenario). Provided that the lockdown starts early enough, the economic rebound offsets the initial drop in production relative to laissez-faire after 3 years. This is largely due to the lives saved and the infections prevented as a result of the lockdown policy. Second, our model asserts that a lockdown policy is more effective in reducing cumulative deaths than the laissez-faire. Even in the worse lockdown case, the 16-week delay scenario, the cumulative death rate over the whole period is half the one generated by the laissez-faire. It is worth pointing out that the balance between the economic cost and the welfare losses due to aversion to deaths in our model depends on the value of the parameter θ. As reflected in Fig. 3C, the cumulative economic losses (relative to laissez-faire) also depend on this parameter: As θ rises, more human lives are saved under a longer optimal lockdown, which, in turn, increases the cumulative production losses due to the lockdown.

Our results are sensitive to the two key parameters that determine the pathogen's transmission efficiency: (i) the average rate of transmission per contact, β, and (ii) the duration of time for which an asymptomatic host is infectious before symptoms appear and the host is isolated. Figure 4 (A to D) illustrates the effect of variation in transmission. When transmission is relatively low, the pathogen can be eliminated by a short, deep lockdown. As transmission efficiency increases, the depth of lockdown increases, but the duration is relatively constant. Contrasting results emerge if the duration of infectiousness changes. Initially, the pathogen can be controlled by deepening the lockdown level. However, if the duration of infectivity is substantially prolonged, then the lockdown has to last longer to eliminate infections. This implies that the optimal depth of a lockdown is partly driven by the pathogen transmission rate and the lockdown's duration is dependent on the duration of infectivity.

Case 2: Random testing and isolation

The development of tests to identify people who are infectious but not yet showing symptoms is potentially a powerful tool to contain an emerging epidemic (*32*–*34*). The speed with which these tests

Table 1. Parameters used in the simulations. These are based on those used in published studies (*23*–*26*).

Variable	Value	
SEIR		
μ	1/70	Birth and death rate
$\mu/1$	0.02	Death rate of I_1
μ_R	1/70	Death rate of R
μ_E	0.012	Death rate of E_2
p_M	1/3	Probability of death of an individual during severe infection
γ	0.05	Proportional contribution to infection of E_2 and I
ϕ	1/2	Transition rate from R to S
ϕ_1	365/5	Transition rate from E_1 to E_2
ϕ_2	365/5	Transition rate from E_2 to I
δ	365/18	Transition rate from C to S
δ_1	365/18	Duration^{-1} of I_1
δ_2	365/13	Duration^{-1} of I_2
β	150.00	Per capita transmission rate
η	0.15	Hospitalization fraction during infection period
Testing		
c	Variable > 0	Number of contacts per infection generated (E_2 and I_1)
s_p	0.90	Specificity of test
s_e	0.90	Sensitivity of test
τ	365/12	Default testing rate
τmax	365/2	Maximum testing capacity for a developed country
r	0 or 1	Contact or random testing respectively
testCost	50,00	Cost for one test in dollars
yearly US GDP	21.5 T	Yearly U.S. total GDP in 2019 in dollars
Y_0	$A0N^{\alpha}$	Yearly model total GDP with no disease spread
ρ_0	$Y_0 \frac{\text{testCost}}{\text{yearly US GDP}}$	Cost for one test in model units
ρ_1	0.1	Testing cost parameter
N	1×10^5	Total number of individuals
Production		
A_0	1×10^5	(Capital-adjusted) Total factor productivity
α	2/3	Curvature of production function
Δ	1/5	Short-run elasticity of capital utilization
ϵC	0.8	Work efficiency of C-class (fraction of contacts that work and do so under isolation)
Utility parameters		
σ	1/2	Curvature of utility from consumption
ω	3	Curvature of disutility from deaths
θ	1×10^{-4}	Weight of utility from consumption versus disutility from deaths

can be developed is dependent on obtaining viral material from infected hosts. It may then take time for tests to be manufactured and be made widely available. There are a variety of methods that have been used to identify contacts of infected hosts. These methods vary in accuracy: Some may simply be based on self-recognition of contacts; others may be based on cell phone–based associations that identify when a potential contact has been within the vicinity of a cell phone–bearing exposed host [see (*35*)]. We attempt to capture this range of efficiencies using a parameter, c. When $c = 0$, there is "perfect knowledge" and only hosts known to be exposed and infected, E_1 and E_2, are identified. As c increases ($c > 0$), a larger pool of potential contacts are identified, a smaller proportion of whom are actually infected. This creates a pool of individuals, ($C + E_1 + E_2$), who are placed in partial isolation from both susceptible

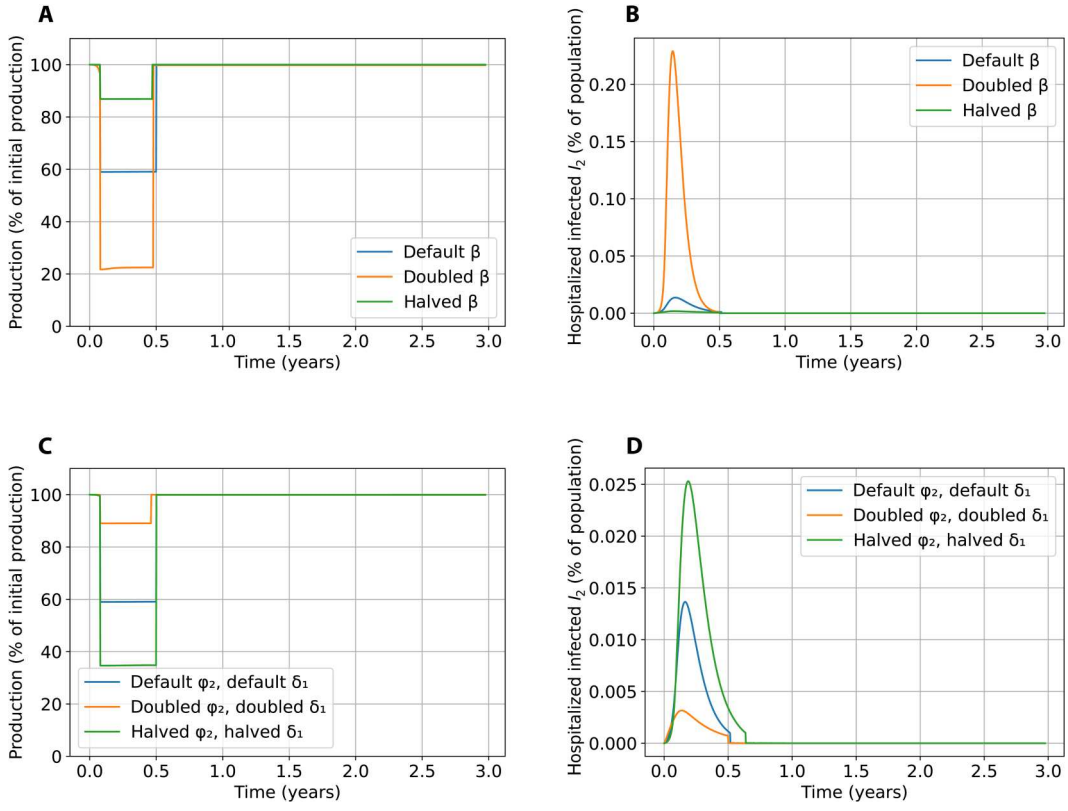

Fig. 4. Lockdown without testing: Effects of variations in infectious rate and in the transition rates. Impact of varying the transmission rate and duration of infectivity rates (from E_2 to I_1 and from I_1 to R) on the size of the epidemic and the optimal lockdown duration. (**A**) Changes in production (relative to initial pre-epidemic production). (**B**) The number of hospitalized individuals for different values of the transmission rate β. (**C** and **D**) The same variables for different durations of infectivity ϕ_2 (transition from E_2 to I_1) and δ_1 (transition from I_1 to R). Values of the rest of the economic and epidemiological parameters are described in Table 1.

and infected hosts. Concomitantly, this reduces their contributions to the economy by a factor ϵ_C. It is important to notice that we lump C, E_1, and E_2 together for the purposes of our economic calculations, as they have all been exposed to infection, but we cannot differentiate between their infectious statuses without testing. Once they have been tested, they can either progress to the E_2 classes of infection, if positive, or be returned to the susceptible class, S, if they have a negative test. Larger values of c reflect a larger degree of caution, which is synonymous with a larger impact on the economy.

We initially assume two types of testing: (i) random testing of a proportion of uninfected hosts and (ii) contact testing of people who have been placed in the C, E_1, and E_2 classes. We note that the rate at which tests identify infected hosts is a function of the type of tests used. In the case of COVID-19, polymerase chain reaction tests would identify all four categories of hosts as infected (E_1, E_2, I_1, and I_2), although it may take several days for the laboratory to return the results of the tests. In contrast, contacts can run laminar flow tests for themselves and have results in 15 min. These tests will likely only identify E_2, I_1, and some I_2 hosts as positive, but these are the hosts that are most likely to be transmitting the virus (particularly E_2). Plainly, tests that can be self-administered and that produce rapid results will minimize absences from work and, hence, the impact of the pathogen on the economy.

Case 3: Contact identification, testing, and isolation

It is natural to assume that tests are likely to be imperfect, at least initially. They will give rise to both false positives and false negatives. We have included two parameters for sensitivity and specificity that reflect measured value of false positive and negatives [see (*34*)]. Contacts who have tested negative are returned to the susceptible class, S, where they continue to work and mix at the same rate as other susceptible hosts, S.

Our results consistently suggest that testing and isolation of infected hosts are considerably more effective than lockdowns in controlling the epidemic outbreaks (Fig. 5). In the case of purely random testing, we again see that the longer the policy is delayed, the larger numbers of people are identified as asymptomatically infected and isolated from work until tested for infection status. Isolation always reduces the number of people hospitalized and dying. However, these increase at a more rapid rate than they do when lockdown is delayed. Crucially, the economic costs of testing and isolation are always lower than when using a lockdown as a control measure, particularly when testing is started early. In addition, there are substantially fewer deaths when testing is initiated early.

If contact tracing is perfect and all contacts of infected hosts are identified and isolated, then testing is unsurprisingly highly efficient at both containing the epidemic and in minimizing economic costs (Fig. 6). However, as mentioned earlier, it is unlikely that contact tracing is perfect, so we allow our model to test different

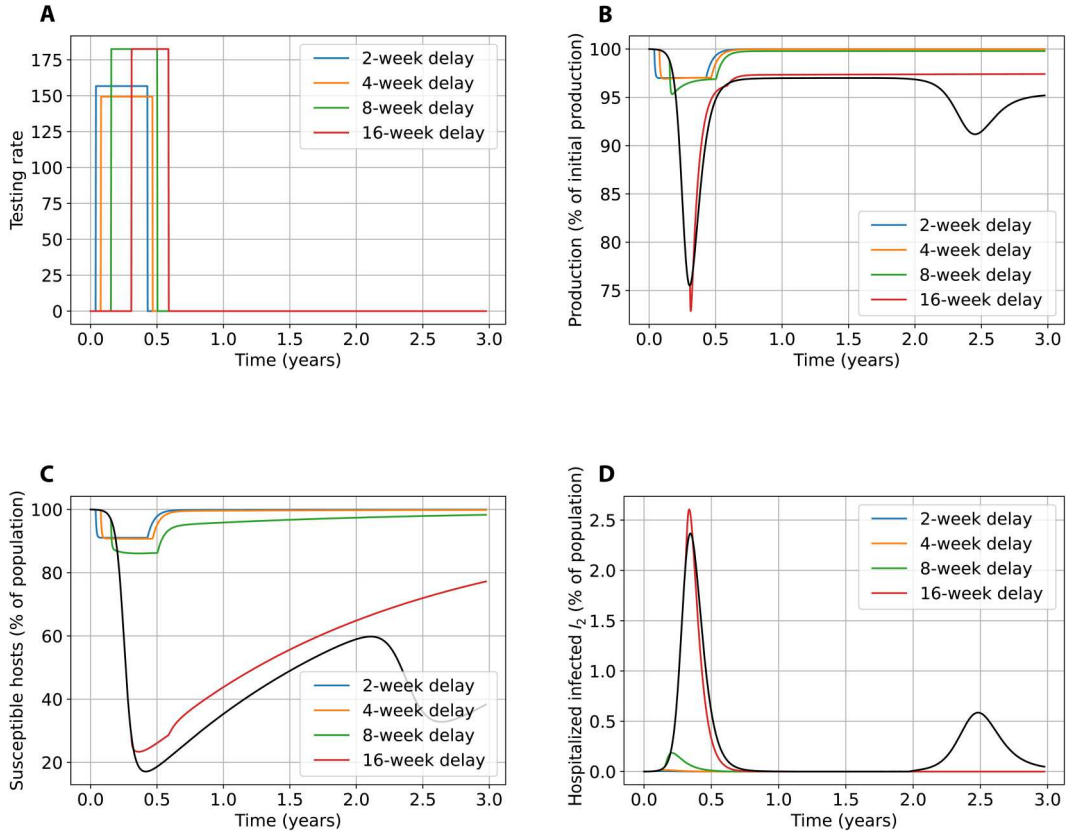

Fig. 5. Random testing without lockdown. Strength and duration of optimal random testing and its impact on the epidemic. (**A**) The intensity and duration of optimal random testing. (**B**) Resultant production compared to initial pre-epidemic production. (**C**) The relative size of the susceptible population. (**D**) The numbers of hospitalized individuals. In (B) to (D), the black line illustrates the laissez-faire case of zero testing. In each figure, we vary the initiation of testing: two weeks (blue), 4 weeks (yellow), 8 weeks (green), and 16 weeks (red). Values of the economic and epidemiological parameters are described in Table 1 ($r = 1$, $c = 0$).

numbers of potential contacts when an ever-widening net of contacts is offset by a reduction in efficiency in identifying infected contacts (we essentially assume that this is linear as the time taken to test and identify all infected costs increases linearly with the number of people tested). Delays in starting contact testing again increase the levels of mortality and economic damages, as more people have to be isolated as contacts when the epidemic has progressed and more hosts are infectious (Fig. 6). Testing reduces the time that healthy contacts, *C*, are removed from the workforce but enhances the background force of infection as rapid testing returns them to the pool of susceptible hosts, *S*. Testing not only reduces the mortality and hospitalizations associated with the epidemic, it also minimizes the epidemic's overall size. Even delays of up to a month result in only 5% reductions in economic productivity, for periods of up to 2 years. In contrast, starting testing within 2 weeks generates a 3% reduction in economic growth for a period of around a year.

Case 4: Lockdown followed by testing

We now consider joint policies where lockdown and testing/tracing are jointly optimally determined. We concentrate of the realistic scenario where, while lockdown can be started at any point in time after the epidemic begins, testing at full capacity and maximal efficiency requires additional time. Accordingly, we focus on the case where full-capacity testing can only be implemented with some delay. In the results below, we have assumed that a lockdown starts after 4 weeks and testing after 16 weeks. Our optimization algorithm then determines the optimal length and intensity of the lockdown and of testing under these constraints. An interesting question concerns whether optimal lockdown and testing policies will overlap for certain time intervals or not. We shall see that this depends not only on the maximal testing capacity but also on the efficiency of testing and tracing.

The first panel (Fig. 7) presents the results for a developed country with a maximal testing capacity $\tau_{max} = 365/2$, which is our reference value (see Table 1); we then consider the case of a developing country with more limited access to testing. In addition to the benchmark laissez-faire case (black curves), we have considered four different types of testing/tracing ranging from the most efficient combination of both ($r = 0$, $c = 0$, blue curves) to less efficient combinations (e.g., $r = 0$, $c = 2$, green curves). Figure 7 (A and B) displays the optimal lockdown and testing policies for these four types of testing/tracing. In all cases, we find that the lockdown is optimally stopped as soon as testing starts, regardless of the type of testing (random or targeted) and for all the values of tracing efficiency considered. Even when contact tracing casts a broad net, $c = 2$, which generates a higher level of an initial lockdown, it is still stopped as soon as testing becomes available. This strongly illustrates the superiority of testing/tracing over lockdown when testing is efficient and capacity is high. Note also that, although the optimal duration of testing may differ, optimal testing always

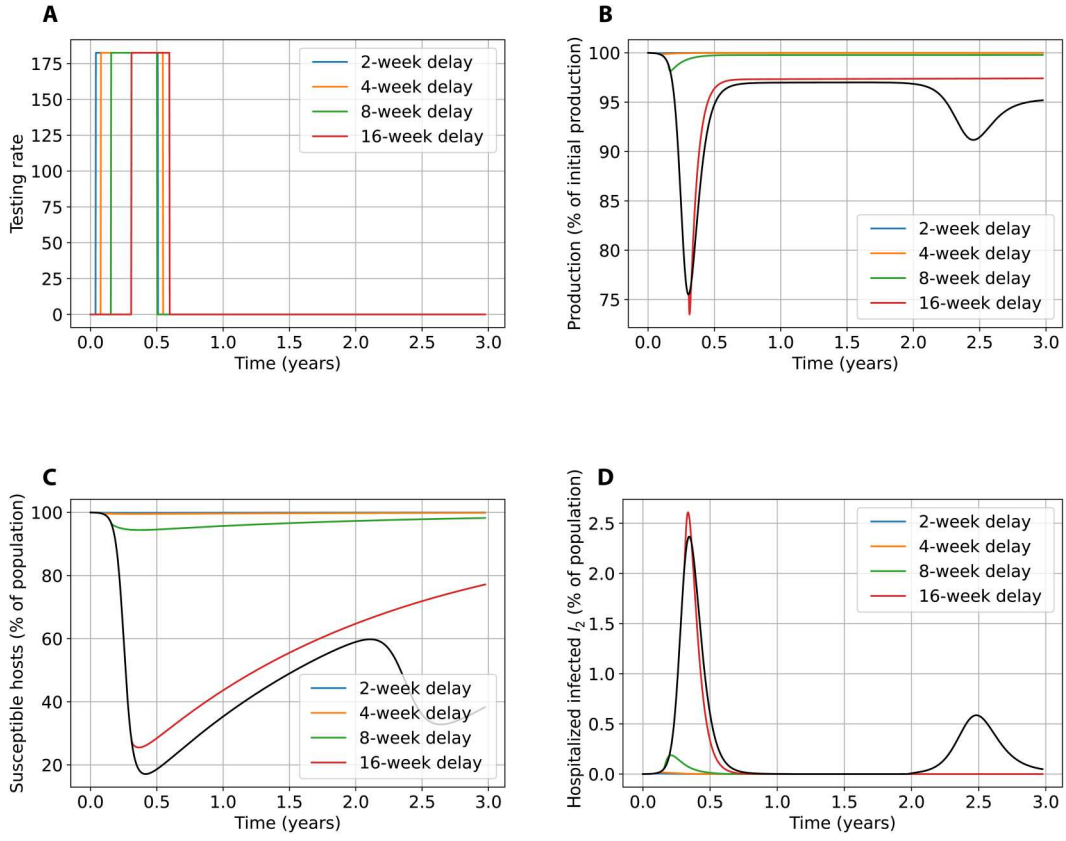

Fig. 6. Targeted ("perfect") testing without lockdown. Strength and duration of the optimal random testing and its impact on the epidemic in case of target testing. (**A**) The intensity and duration of the targeted testing. (**B**) Resultant production with respect to initial pre-epidemic production. (**C**) The relative size of the susceptible population. (**D**) The number of hospitalized individuals. In each case, we again vary the initiation of testing: two weeks (blue), 4 weeks (yellow), 8 weeks (green), and 16 weeks (red). Values of the economic and epidemiological parameters are described in Table 1 ($r = 0$, $c = 0$).

reaches full capacity under all parameterizations that we consider. We find that the policy leading to the highest social welfare outcome is obtained when $r = c = 0$. We also find that (optimal) targeted testing is not always superior to random testing from the total social welfare point of view, for example, if tracing efficiency is sufficiently low, random testing ($r = 1$, $c = 0$) dominates contact tracing ($r = 0$, $c = 2$) from a social welfare perspective.

As mentioned earlier, different countries have different amounts of resources to deal with an epidemic and this includes their testing capacity. To mimic the case of developing countries, with a smaller capacity to produce or import effective tests, we consider (Fig. 8) the case where the testing capacity parameter is reduced by a factor of one-third (which implies setting $\tau_{max} = 365/6$). This single modification leads to substantially altered results. Under this scenario, we no longer observe that lockdown is outperformed by testing as soon as the latter is available. As testing costs are now more binding in relative terms, a lockdown remains in place for an initial period after testing is introduced. The only exception is when tracing is highly efficient at identifying contacts ($r = c = 0$). When tracing is less efficient, the lockdown is optimally extended to about 6 months, which is more than 2 months after testing becomes available. This finding serves as an important reminder for policymakers. Optimal epidemic policies might be different in developed versus developing countries, and what works well in one country might not necessarily work well in a different context.

DISCUSSION

In his classic volume on pandemic prevention, Chris Dye quotes, "Everything we do before a pandemic will seem alarmist. Everything we do after a pandemic will seem inadequate" (*36*). The primary motivation of this paper was to develop a quantitative framework that provides guidance for policy-makers when faced with this dilemma. Different countries pursued different strategies in their attempts to control and minimize the damage caused by the recent COVID-19 pandemic. At the earliest stages, the only control tool available was either a lockdown or the laissez-faire response of doing nothing. The primary purpose of lockdowns is to "flatten the epidemic curve" and reduce the peak levels of hospitalizations (*37*). Our results suggest that, when we optimize taking into account economic and public health/mortality objectives, both economic damages and pathogen-induced mortality can be substantially reduced relative to the laissez-faire case if a lockdown is put in place as swiftly as possible. This is true even if we take into account the substantial adverse economic consequences resulting from a lockdown. Ash *et al.* (*16*) calibrate a dynamic economic model to COVID-19–related epidemiological data to evaluate the effects of different scenarios, including voluntary isolation, targeted isolation, and blanket lockdowns. Their numerical findings assert that voluntary isolation or blanket lockdowns suppress the epidemic nearly as effectively as targeted isolation but impose higher

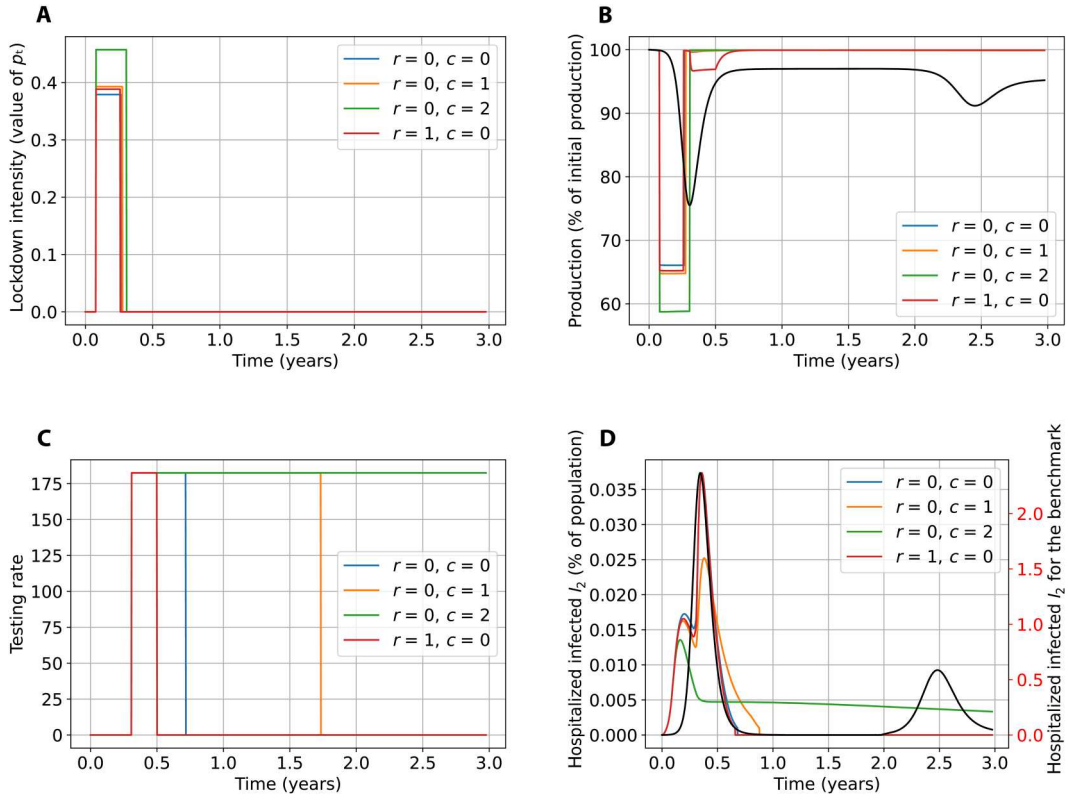

Fig. 7. Optimal combination of lockdown and testing in the high testing capacity, *s*. Strength and duration of the optimal lockdown, strength and duration of the optimal testing, and their impact on the epidemic varying *r* and *c*: $r = 0$, $c = 0$ (blue), $r = 0$, $c = 1$ (yellow), $r = 0$, $c = 2$ (green), and $r = 1$, $c = 0$ (red) when the capacity of the lockdown is high. (**A**) The strength and duration of the optimal random lockdown. (**B**) The production level relative to initial production. (**C**) The strength and duration of the optimal random testing. (**D**) The share of hospitalized individuals. The black line again provides a comparison with the laissez-faire case. The values of the economic and epidemiological parameters are described in Table 1.

economic costs. Similar results have been found in earlier models for influenza pandemics (*27*). Our results also suggest that lockdown policies could be replaced by contact tracing and testing as soon as viable tests become available. Delays in initiating this transition in policy will always lead to higher economic damages and enhanced mortality. Looking into how to best prepare for the future, this suggests that developing a genetic library of potential pathogens that may cross over to humans in the future would provide an important safety net for minimizing the potential economic and human costs of epidemics, particularly if they can be used to rapidly develop effective tests. Investment in developing these tests from material gathered in broad-scale surveys of potential novel viruses in wild reservoir hosts is likely to prove economically valuable (*2*, *38*).

The potential magnitude of voluntary changes in behavior can be partially gauged by considering Sweden, when our laissez-faire case would predict around 400,000 deaths in a population of around 10 million people. In contrast, Sweden only experienced around 20,000 deaths, strongly suggesting that people's personal response to the epidemic may have had a pronounced impact on buffering transmission. In this paper, we have also chosen to study in depth the interaction of an extensive epidemiological model and a comprehensive (frictional) epidemic control policy menu without endogenizing individual responses. In the recent epi-econ literature, several authors have tried to account for these endogenous responses. For example, Eichenbaum *et al.* (*28*), among others, calibrate their model to the U.S. economy and assume that agents endogenously choose to reduce their consumption and labor supply by taking the probability of infection into consideration. However, as documented in (*39*), crucial aspects of individual behavior derive from heterogeneous characteristics, as well as strategic reactions to epidemic control policies and local considerations. Modeling strategic behavior by both heterogeneous individuals in the economy and the policy-maker(s) is beyond the scope of our analysis. Rather than incorporating simple ad hoc specifications of individual behavior, we focused our present analysis on the implications of the added epidemiological factors that we consider and on a more realistic set of epidemic control policies.

Another important open question for future research concerns the connections between different pathogen characteristics and the optimal duration/depth of the lockdown policy. The desirable lockdown depth is clearly a function of the pathogen's transmissibility. Our analysis pointed to the conclusion that more transmissible pathogens require deeper lockdowns (Fig. 4). We also suspect that the duration of the lockdown would increase with the duration of infectivity. Establishing such connections between pathogen characteristics and policy response is an important topic for future research.

We have not included age, or sex, structure in our model; both of these are potentially important extensions, as mortality from

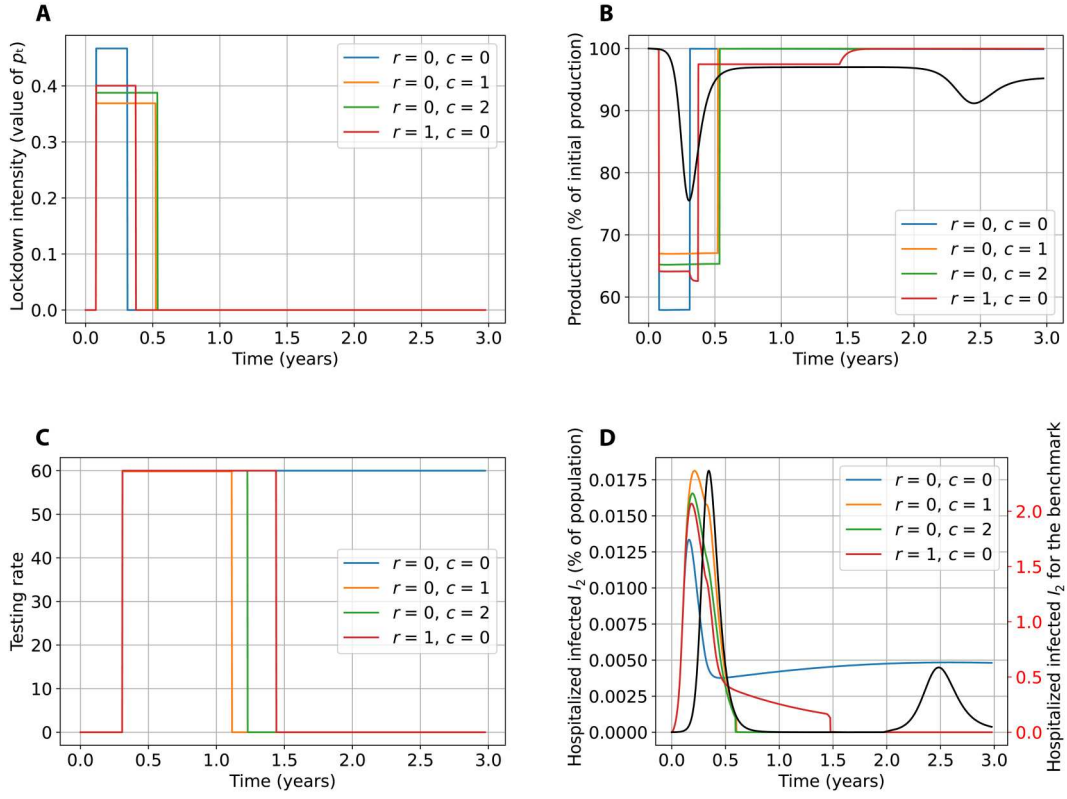

Fig. 8. Optimal combination of lockdown and testing under reduced testing capacity. Strength and duration of the optimal lockdown, strength and duration of the optimal testing, and their impact on the epidemic varying the parameters r and c: $r = 0$, $c = 0$ (blue), $r = 0$, $c = 1$ (yellow), $r = 0$, $c = 2$ (green), and $r = 1$, $c = 0$ (red), when the capacity of the lockdown is reduced. (**A**) The strength and duration of the optimal random lockdown. (**B**) The production relative to initial production level. (**C**) The strength and duration of the optimal random testing. (**D**) The share of hospitalized individuals. The values of the economic and epidemiological parameters are described in Table 1.

COVID-19 increases substantially with age. Men also seem to suffer higher mortality rates than women (*40*, *41*). We are currently developing an extended version of our model that incorporates these aspects, but our primary focus in this paper was to examine epidemiological and economic interactions between a generic "emerging pathogen" and a homogeneous workforce. As the sharpest increase in COVID-19 mortality occurs in age classes that are usually post-retirement and no longer part of the active workforce, we do not think the addition of these important aspects of population structure will have a major impact on our main conclusions. We do explicitly acknowledge that older retired people will place substantial additional pressure on health care services in the early stage of a pandemic, and this may lead to increased mortality within the workforce when the capacity of the health care system is reduced by large numbers of elderly sick people (*27*, *31*).

Our model assumes that asymptomatic individuals eventually exhibit symptoms of infection. This is not always the case for COVID-19 where asymptomatic patients may play a substantial role in transmission. Modification of the basic structure of the epidemiological model can incorporate this assumption [e.g., (*25*)]. Because our model was parameterized to obtain levels of transmission and hospitalization consistent with those reported for COVID-19, such an extension would effectively shift some of the contribution to the force of infection from presymptomatics to asymptomatics. Thus, we do not expect the major results to be modified except for further emphasizing the importance of testing efforts.

Arguably, the most economical control strategy during the COVID-19 pandemic has been the wearing of surgical masks at work and in public places (*42*). We have not incorporated the use of masks in our analysis, partly because we see it as an important "safety play" that is unlikely to stop the pandemic. Masks may provide some protection to those that wear them while also substantially reducing transmission from asymptomatic infected hosts. An important limitation here is that we do not have good estimates of the efficacy of masks (neither for transmission blocking nor for protection). The best available methods suggest an average efficiency of around 45% (averaged across both transmission blocking and susceptible protection) (*43*). Similarly, the proportion of people wearing masks varies widely and may follow current levels of infection and perceived risk in the population. A simple static analysis can compare the efficacy of face masks with that of a vaccine (Supplementary Materials). A high proportion of people have to wear very efficient masks if they are to be effective in reducing R_0 below unity. They will still serve the useful function of slowing the epidemic and reducing the pressure on health care services. Ultimately, we see them as a useful adjunct to the other forms of NPI discussed in more detail below. If more people wear masks, then

lockdown can be shorter and affect a smaller proportion of the economy.

We have also not considered "long COVID-19" in our model framework, although there is increasing evidence that this is a substantial problem. Current estimates suggest that between 1 and 5% of people who acquired COVID-19 in the early stages of the epidemic continue to feel debilitated and unable to work. These symptoms last between 6 months and the 3 years for which we have data. They could be included in our model as an additional equation that diverts a proportion of people from the I_2 class into a long COVID-19 class rather than the recovered and immune class. Hosts would remain in this class for a number of years and would make a reduced contribution to the economy during this period. All of the additional terms required to include this into our current model structure are linear; whence, the additional cost to the economy is essentially the product of the proportion of people who develop long COVID-19 and their reduced ability to work. All of which makes our calculations optimistic. The specter of long COVID-19 always increases the economic value of interventions that minimize the number of hosts ever infected (as the number of people with long COVID-19 will always scale with the number of people ever infected).

The development of vaccines for COVID-19 has had a substantial impact on the dynamics of the pandemic. Vaccines have increased levels of personal protection and permitted the beginnings of an economic recovery from the substantial initial collapse in the spring and summer of 2020. Unfortunately, vaccines have not provided us with a perfect solution, as vaccinated people seem still able to transmit the pathogen while exhibiting considerably reduced levels of morbidity. Immunity also seems disconcertingly short lived (6 to 8 months). Everything we have described above assumes that we do not yet have a vaccine for the hypothetical emergent pathogen. In a second paper, we will expand our framework to consider the role of vaccines and their efficacy and resistance to vaccine adoption within the host population.

Our framework is quite flexible and allows for various additional extensions. A few illustrative examples are worth mentioning. Adding hospitalization costs and a capacity constraint on the health system, requiring that the amount of hospitalized people I_2 remain under a given congestion threshold. This would provide an additional realistic constraint on optimal policies. We could also consider additional policy variables, for example, regarding the modality/intensity of contact tracing. We could further investigate the robustness of our results to the use of different utility/disutility functions. For example, we could use linear U and V as in (*10*). We believe that the qualitative aspects of our findings related to the relative effectiveness of lockdowns versus testing would remain intact. It is also possible to use our framework to study the optimal number of lockdown phases, given certain adjustment costs associated with introducing additional phases. In our simulations, comparing the optimal strategies and the welfare function under different scenarios, we observed that in most cases, one lockdown phase is sufficient, even if the policy-maker could, in principle, implement additional lockdowns. Last, we could pursue a more detailed sensitivity analysis, for example, in connection to the choice of the parameter θ. This parameter establishes the relative weights between the utility from production and the disutility from deaths during the epidemic. An extension of the model could vary θ to characterize the Pareto frontier, i.e., the policies that cannot lead to improvements in both economic and public health objectives.

Concerns about future pandemics motivated this work (*2*, *38*). We have therefore included sufficient flexibility within the model structure to allow us to adjust parameters to consider novel pathogens with different characteristics such as severe acute respiratory syndrome (SARS) or another influenza strain, where symptoms do not appear until a host is infectious, or pathogens with long asymptomatic phases (such as HIV). Our framework could also be expanded to consider vector-borne infections such as those caused by the dengue or Zika virus.

COVID-19 was not the first pandemic, and it will not be the last. A huge proportion of the impact of COVID-19 has been driven by different governments responding to the pandemic in an "ad hoc" fashion as political pressures to maintain economic activity clash with epidemiological advice. The contrasting responses of different nations reflect the relative magnitude of economic and epidemiological forces and national levels of expertise in these areas. We believe that combining economic modeling, which highlights incentive constraints, with epidemiological modeling, which focuses on public health considerations, will be increasingly relevant in designing policy interventions that are both effective and attainable during a future public health crisis. The structure described here provides crucial initial steps in this direction.

METHODS

Epidemiological model

Our epidemiological model is based on the standard *SEIR* framework (*44*, *45*). We have modified the basic framework by dividing the exposed (E) and infectious (I) classes into two sequential classes, E_1 and E_2 and I_1 and I_2 (Fig. 1). Exposed hosts who are not yet infectious are classified as E_1, while asymptomatic, contagious hosts are classified as E_2. We assume that E_1 individuals transform to E_2 at an exponential rate determined by ϕ_1. The presymptomatic hosts, E_2, transform to symptomatic infected hosts, I_1, at a rate ϕ_2. Both E_2 and I_1 are infectious. This rate largely determines the duration of time during which exposed hosts are able to transmit infection before they show symptoms of infection. If ϕ_2 is large ($\sim$365; around 1 day), then exposed hosts quickly exhibit signs of symptoms and can be identified as infectious (as occurred with SARS). In contrast, if ϕ_2 is slower ($\sim$365/7; a week), then asymptomatic hosts may transmit the disease for up to a week before showing symptoms, as in the case of COVID-19 (or many years in the case of HIV or tuberculosis). In a similar way, infected hosts, I_1, may become sick and get hospitalized, I_2. These hosts have a higher mortality rate but are assumed to be in relative isolation and are thus unable to transmit the pathogen, except to unprotected health care workers. The majority of the pathogen-induced mortality occurs in the I_2 class.

We also include an additional class, C, into our model structure, these are contacts of infectious hosts who do not develop infection. Contact tracing identifies $C + E_1 + E_2$ as contacts of infected hosts, testing is used to differentiate uninfected contacts, C, from exposed hosts (E_1 and E_2); the former can return to work, the latter remain in isolation and go on to develop infection.

The main equations of our *SCEEIIR* models are the following

$$\dot{S} = \mu N - \mu S - (1-p)^2\,\beta\frac{S[E_2\,(1-\gamma)+\gamma I_1](1+c)}{N} + \phi R + \tau\left(r+\frac{1-r}{1+c}\right)Cs_p + \delta C - \tau r S(1-s_p) \quad (1)$$

$$\dot{C} = (1-p)^2\,\beta[cS-(1-\in_c)C]\,\frac{E_2(1-\gamma)+\gamma I_1}{N} + \tau r S(1-s_p) - \mu C - \tau\left(r+\frac{1-r}{1+c}\right)Cs_p - \delta C \quad (2)$$

$$\dot{E}_1 = (1-p)^2\,\beta[S+(1-\in_C)C]\frac{E_2(1-\gamma)+\gamma I_1}{N} - (\mu+\varphi_1)E_1 - \tau\left(r+\frac{1-r}{1+c}\right)s_e E_1 \quad (3)$$

$$\dot{E}_2 = \varphi_1 E_1 - (\mu+\varphi_2)\,E_2 - \tau\left(r+\frac{1-r}{1+c}\right)s_e E_2 \quad (4)$$

$$\dot{I}_1 = \varphi_2 E_2 + \tau s_e\left(r+\frac{1-r}{1+c}\right)(E_1+E_2) - [\mu+\mu_{I_1}+(1-\eta)\delta_1+\eta\delta_1]I_1 \quad (5)$$

$$\dot{I}_2 = \eta\delta_1 I_1 - [\mu+p_M\delta_2+(1-p_M)\delta_2]I_2 \;(\text{where } \mu_{I_2}=p_M\delta_2) \quad (6)$$

$$\dot{R} = (1-\eta)\delta_1 I_1 + (1-p_M)\delta_2 I_2 - (\mu+\mu_R+\phi)\,R + \nu V_w \quad (7)$$

Here, N denotes the size of the host population, all of whom contribute to economic output when healthy. Once the pathogen has established, N is the sum of all possible types of host

$$N = S + C + E_1 + E_2 + I_1 + I_2 + R \quad (8)$$

As the death rates of the epidemics is low, we will approximate N with N_0 [the initial size of the population; see, e.g., (*44*, *45*)]. In the absence of the pathogen, the total number of deaths at any time is given by $\dot{D} = \mu N$; when the pathogen is present, the additional deaths due to the pathogen are given by $D_C = \mu_{I1}I_1 + p_M\delta_2 I_2 + \mu_E E_2 + \mu_R R$, where severe infections I_2 last an average duration $\frac{1}{\delta_2}$ with a fraction p_M of individuals leaving this class resulting in death, whereas a fraction $(1-p_M)$ are able to recover. Similarly, only a fraction $(1-\eta)$ of individuals leaving class I_1 recovers, whereas η proceeds to severe infection and hospitalization.

Lockdown intensity at time t is parameterized by $p(t)$, the probability that an individual, susceptible or infected, is protected from contact and therefore excluded from transmission events. Per capita transmission rate is given by β, and for each infection, a number c of uninfected contacts is generated. Contacts are restricted in their ability to work, with only a fraction ϵ able to do so and under isolated conditions, for example from home, which also reduces their exposure to infection.

Testing occurs at rate τ. The indicator variable r takes values of 1 or 0 to control the respective implementation of random testing versus contact tracing. The specificity and sensitivity of testing are given respectively by s_p and s_e.

Structure of economic model

Flow output, Y, is produced through a production function using labor, L, as the only production factor. Abstracting from capital accumulation seems reasonable because our analysis concentrates on a short time horizon. The short horizon of the epidemic also prevents a large-scale substitutability between capital and labor. Infections affect the economy by reducing labor supply and by making the labor used less productive. The latter could result from disruptions to supply chains or related shortages in necessary intermediate inputs. To model these effects, we assume that infections reduce both the labor used in the production process and the total factor productivity parameter, A. We assume that the costs, Φ, associated with the number of people tested, x, are subtracted from total output. The specific form imposed on the function Φ ensures that when x is very small, the costs of testing increase linearly with x. We assume that it is forbiddingly costly to test the entire population. Thus, the maximum rate of testing is constrained.

The proportion of people $p(t)$ who are in lockdown are unable to contribute to the economy. We define the production function as

$$Y(t) = \underbrace{A(t)\,\{[1-p(t)]L(t)\}^{\alpha}}_{\text{production function}} - \underbrace{\Phi[x(t)]}_{\text{testing cost}} \quad (9)$$

where

$$A(t) = A_0[1-p(t)]^{\Delta}$$
$$L(t) = S(t) + \in_C[C(t) + E_1\,(t) + E_2(t)] + R(t)$$
$$\Phi(x) = \rho_0 x + \exp\left(\frac{\rho_1}{N-x}\right) - \exp\left(\frac{\rho 1}{N}\right)$$
$$x(t) = \tau[C(t) + E_1(t) + E_2 + rS(t)]$$

In the above expression, $A(t)$ stands for the (capital-adjusted) total factor productivity. This parameter determines the effectiveness of the labor input in producing the consumption goods. The flow labor supply consists of individuals who are susceptible or recovered from the pathogen, as well as those who are exposed or potentially exposed (contacts). The variable x indicates the flow of people tested. The form of the testing cost function, $\Phi(x)$, is meant to capture the property that, at low levels of testing, the costs increase linearly in the number of tests, while at the same time, it is prohibitively expensive to test the entire population. More precisely, when ρ_1 and x are small, the marginal cost of testing, i.e., the cost of administering one additional test, is approximately ρ_0. To calibrate ρ_0, we convert the dollar value of a test to units of daily U.S. per capita GDP (gross domestic product). As a benchmark, we use the per capita U.S. GDP value of $63,416 in 2020 and assume an average cost of testing of $174. Following (*14*), we set $\rho_1 = 0.1$, implying that the cost of testing increases relatively slowly as large numbers of people are tested.

The objective of the policy-maker is to maximize the total social welfare (TSW) function, this is given by

$$\text{TSW}(T) = \int_0^T U[Y(t)] - \theta V[D_c(t)]dt \quad (10)$$

with

$$U(Y) = \frac{Y^{1-\sigma}}{1-\sigma}$$
$$V(D_c) = \frac{D_c^{\omega}}{\omega}$$

In the above expressions, $U(Y)$ stand for the satisfaction (utility) from consuming goods and services, while $V(D_c)$ stands for the direct utility loss of lives lost.

Calibration

We calibrate the model using the parameters described in the Supplementary Materials and Table 1; these are based on those used in published studies (*23*–*26*).

Optimization

The optimization problem that we study is nontrivial and cannot be treated with the standard tools used in the current literature. In the Supplementary Materials, we describe in detail our approach. Here, we provide a brief summary. First, our problem falls into the class of deterministic optimal control with exit time. A description of this class of problems is provided, e.g., in chapter 8 of (*46*). The main ingredients of this class of problems are (i) the time horizon of the problem, (ii) the state/control variables and the space where they belong (the state/control space), (iii) the state equation, which provides the dynamic behavior of the state variables as a function of their initial data and of the choice of the control strategies, (iv) the set of admissible control strategies, (v) the objective function to optimize over all admissible control strategies, and (vi) the target set O where the epidemic ends.

The main differences with respect to the papers in the existing literature are the following: (i) the "exit time feature," i.e., the fact that the epidemic stops when all associated variables (namely, E_1, E_2, I_1, and I_2) fall below 1 and (ii) the "discrete control strategies," i.e., the fact that the control strategies are piecewise constant with a given finite number of switching times.

Both features are crucial in making the model more realistic, the first in connection to the behavior of the epidemic and the second to take account of the constraints faced by policy-makers. In this context, the existence of optimal strategies can be demonstrated using standard arguments; however, the uniqueness of the optimal solution is not guaranteed. Concerning the numerical approximations, to compute the objective function, we rely on classical numerical methods for ordinary differential equations. However, because of the lack of regularity of the controls, we need to use a numerical method outside the class of those used for "stiff problems," such as implicit Runge-Kutta methods of high order. Instead, we approximate the continuous-time integration using a Gauss-Kronrod quadrature rule.

The process of numerical optimization is challenging because of the lack of convexity in the objective function. In the absence of convexity, there is no guarantee that a local optimum will also be a global one. Therefore, during the numerical optimization, we have to rely on a global optimization algorithm, which is numerically more demanding than a local numerical optimizer. Global optimization suffers from the curse of dimensionality, that is, the number of function evaluations required for a thorough search in the state space grows exponentially with the dimension of the problem. For the optimization procedure, we used both the DIRECT and DIRECT-L algorithm proposed in (*47*) and (*48*), respectively. Both algorithms are deterministic procedures based on a subdivision of the domain in iteratively smaller rectangles until convergence is reached. To ensure the correctness of the results, we also performed additional tests based on a combination of a brute-force approach on a very fine grid and a local refinement based on a local optimizer. An additional confirmation of the accuracy of the optimization procedure is that the maximum value of the objective function shifts in the expected direction when changing some of the parameters (e.g., when increasing the delay in the control policies, the objective function decreases) (*49*).

Supplementary Materials

This PDF file includes:

Supplementary Materials
Figs. S1 and S2
References

REFERENCES AND NOTES

1. D. Adam, The pandemic's true death toll: Millions more than official counts. *Nature* **601**, 312–315 (2022).
2. A. S. Bernstein, A. W. Ando, T. Loch-Temzelides, M. M. Vale, B. V. Li, H. Li, J. Busch, C. A. Chapman, M. Kinnaird, K. Nowak, M. C. Castro, C. Zambrana-Torrelio, J. A. Ahumada, L. Xiao, P. Roehrdanz, L. Kaufman, L. Hannah, P. Daszak, S. L. Pimm, A. P. Dobson, The costs and benefits of primary prevention of zoonotic pandemics. *Sci. Adv.* **8**, eabl4183 (2022).
3. S. Jones, How to eradicate the next pandemic disease. *Nature* **610**, S48–S49 (2022).
4. D. Sridhar, Five ways to prepare for the next pandemic. *Nature* **610**, S50 (2022).
5. G. Persad, A. Pandya, A comprehensive covid-19 response—The need for economic evaluation. *N. Engl. J. Med.* **386**, 2449–2451 (2022).
6. J. T. Cuddington, J. D. Hancock, Assessing the impact of aids on the growth path of the malawian economy. *J. Dev. Econ.* **43**, 363–368 (1994).
7. A. Young, The gift of the dying: The tragedy of AIDS and the welfare of future african generations. *Q. J. Econ.* **120**, 423–466 (2005).
8. M. Gersovitz, J. S. Hammer, The economical control of infectious diseases. *Econ. J.* **114**, 1–27 (2004).
9. D. Acemoglu, V. Chernozhukov, I. Werning, M. D. Whinston, Optimal targeted lockdowns in a multigroup sir model. *Am. Econ. Rev. Insights* **3**, 487–502 (2021).
10. F. Alvarez, D. Argente, F. Lippi, A simple planning problem for covid-19 lock-down, testing, and tracing. *Am. Econ. Rev. Insights* **3**, 367–382 (2021).
11. C. Gollier, Cost–benefit analysis of age-specific deconfinement strategies. *J. Public Econ. Theory* **22**, 1746–1771 (2020).
12. G. Fabbri, F. Gozzi, G. Zanco, Verification results for agestructured models of economic–Epidemics dynamics. *J. Math Econ.* **93**, 102455 (2021).
13. N. Hritonenko, O. Yatsenko, Y. Yatsenko, Model with transmission delays for covid-19 control: Theory and empirical assessment. *J. Public Econ. Theory* **24**, 1218–1244 (2022).
14. F. Piguillem, L. Shi, Optimal covid-19 quarantine and testing policies. *Econ. J.* **132**, 2534–2562 (2022).
15. A. Aspri, E. Beretta, A. Gandolfi, E. Wasmer, Mortality containment vs. economics opening: Optimal policies in a seiard model. *J. Math Econ.* **93**, 102490 (2021).
16. T. Ash, A. Bento, D. Kaffine, A. Rao, A. Bento, Diseaseeconomy trade-offs under alternative pandemic control strategies. *medRxiv*, 2021.02.12.21251599 (2021).
17. B. Böger, M. M. Fachi, R. O. Vilhena, A. F. Cobre, F. S. Tonin, R. Pontarolo, Systematic review with meta-analysis of the accuracy of diagnostic tests for covid-19. *Am. J. Infect. Control* **49**, 21–29 (2021).
18. M. L. Bastos, G. Tavaziva, S. K. Abidi, J. R. Campbell, L.-P. Haraoui, J. C. Johnston, Z. Lan, S. Law, E. M. Lean, A. Trajman, D. Menzies, A. Benedetti, F. A. Khan, Diagnostic accuracy of serological tests for covid-19: Systematic review and meta-analysis. *BMJ* **370**, m2516 (2020).
19. P. Fine, K. Eames, D. L. Heymann, "Herd immunity": A rough guide. *Clin. Infect. Dis.* **52**, 911–916 (2011).
20. P. E. Fine, Herd immunity: History, theory, practice. *Epidemiol. Rev.* **15**, 265–302 (1993).
21. H. E. Randolph, L. B. Barreiro, Herd Immunity: Understanding COVID-19. *Immunity* **52**, 737–741 (2020).
22. A. Fontanet, S. Cauchemez, Covid-19 herd immunity: Where are we? *Nat. Rev. Immunol.* **20**, 583–584 (2020).

23. N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. C. Okell, S. van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, Patrick GT Walker, C. Walters, P. Winskill, C. Whittaker, C. A. Donnelly, S. Riley, A. C. Ghani, Impact of non-pharmaceutical interventions (NPIs) to reduce covid-19 mortality and healthcare demand. (2020).
24. S. Flaxman, S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod; Imperial College COVID-19 Response Team, P. N. Perez-Guzman, N. Schmit, L. Cilloni, K. E. C. Ainslie, M. Baguelin, A. Boonyasiri, O. Boyd, L. Cattarino, L. V. Cooper, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, B. Djaafara, I. Dorigatti, S. L. van Elsland, R. G. FitzJohn, K. A. M. Gaythorpe, L. Geidelberg, N. C. Grassly, W. D. Green, T. Hallett, A. Hamlet, W. Hinsley, B. Jeffrey, E. Knock, D. J. Laydon, G. Nedjati-Gilani, P. Nouvellet, K. V. Parag, I. Siveroni, H. A. Thompson, R. Verity, E. Volz, C. E. Walters, H. Wang, Y. Wang, O. J. Watson, P. Winskill, X. Xi, P. G. T. Walker, A. C. Ghani, C. A. Donnelly, S. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, S. Bhatt, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
25. R. Subramanian, Q. He, M. Pascual, Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2019716118 (2021).
26. J. S. Lavine, O. N. Bjornstad, R. Antia, Immunological characteristics govern the transition of COVID-19 to endemicity. *Science* **371**, 741–745 (2021).
27. T. Déirdre Hollingsworth, D. Klinkenberg, H. Heesterbeek, R. M. Anderson, Mitigation strategies for pandemic influenza A: Balancing conflicting policy objectives. *PLoS Comput. Biol.* **7**, e1001076 (2011).
28. M. S. Eichenbaum, S. Rebelo, M. Trabandt, The macroeconomics of epidemics. *Rev. Financ. Stud.* **34**, 5149–5187 (2021).
29. E. P. Fenichel, C. Castillo-Chavez, M. G. Ceddia, G. Chowell, P. A. G. Parra, G. J. Hickling, G. Holloway, R. Horan, B. Morin, C. Perrings, M. Springborn, L. Velazquez, C. Villalobos, Adaptive human behavior in epidemiological models. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6306–6311 (2011).
30. Y. Yan, A. A. Malik, J. Bayham, E. P. Fenichel, C. Couzens, S. B. Omer, Measuring voluntary and policy-induced social distancing behavior during the COVID-19 pandemic. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2008814118 (2021).
31. R. M. Anderson, H. Heesterbeek, D. Klinkenberg, T. D. Hollingsworth, How will country-based mitigation measures influence the course of the COVID-19 epidemic? *Lancet* **395**, 931–934 (2020).
32. M. J. Mina, R. Parker, D. B. Larremore, Rethinking covid-19 test sensitivity—A strategy for containment. *N. Engl. J. Med.* **383**, e120 (2020).
33. M. J. Mina, K. G. Andersen, Covid-19 testing: One size does not fit all. *Science* **371**, 126–127 (2021).
34. D. B. Larremore, B. Wilder, E. Lester, S. Shehata, J. M. Burke, J. A. Hay, M. Tambe, M. J. Mina, R. Parker, Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Sci. Adv.* **7**, eabd5393 (2021).
35. L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).
36. Christopher Dye, *The Great Health Dilemma: Is Prevention Better Than Cure?* (Oxford Univ. Press, 2021).
37. K. Koelle, M. A. Martin, R. Antia, B. Lopman, N. E. Dean, The changing epidemiology of sars-cov-2. *Science* **375**, 1116–1121 (2022).
38. A. P. Dobson, S. L. Pimm, L. Hannah, L. Kaufman, J. A. Ahumada, A. W. Ando, A. Bernstein, J. Busch, P. Daszak, J. Engelmann, M. F. Kinnaird, B. V. Li, T. Loch-Temzelides, T. Lovejoy, K. Nowak, P. R. Roehrdanz, M. M. Vale, Ecology and economics for pandemic prevention. *Science* **369**, 379–381 (2020).
39. R. Chetty, J. N. Friedman, N. Hendren, M. Stepner, *How did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker based on Private Sector Data* (National Bureau of Economic Research, 2020).
40. N. G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, R. M. Eggo, Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).
41. D. E. L. Promislow, A geroscience perspective on COVID-19 mortality. *J. Gerontol. A Biol. Sci. Med. Sci.* **75**, e30–e33 (2020).
42. N. G. Davies, A. J. Kucharski, R. M. Eggo, A. Gimma, W. J. Edmunds; Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: A modelling study. *Lancet Public Health* **5**, e375–e385 (2020).
43. S. E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, A. B. Gumel, To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic. *Infect Dis. Model* **5**, 293–308 (2020).
44. R. M. Anderson, R. M. May, in *Infectious Diseases of Humans: Dynamics and Control* (Oxford Univ. Press, 1992), pp. 757.
45. M. J. Keeling, P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton Univ. Press, 2011).
46. P. Cannarsa, C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control* (Springer Science & Business Media, 2004), vol. 58.
47. D. R. Jones, C. D. Perttunen, B. E. Stuckman, Lipschitzian optimization without the Lipschitz constant. *J. Optim Theory Appl.* **79**, 157–181 (1993).
48. J. M. Gablonsky, C. T. Kelley, A locally-biased form of the DIRECT algorithm. *J. Global Optim.* **21**, 27–37 (2001).
49. P. Cannarsa, C. Pignotti, C. Sinestrari, Semiconcavity for optimal control problems with exit time. *Discrete Contin. Dyn. Syst. Ser. A* **6**, 975–997 (2000).
50. K. Dietz, J. A. P. Heesterbeek, Daniel Bernoulli's epidemiological model revisited. *Math. Biosci.* **180**, 1–21 (2002).
51. M. Bardi, I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman equations. Systems & Control: Foundations & Applications* (Birkhäuser Boston, 1997).

Acknowledgments

Funding: R.B. was supported by funds from French Government Subsidy, SCSP no. 210 292 814; these were used to hire three interns: A. Ndiaye, V. T. M. Ky, and M. E. Znaidi. We thank them for insightful participation. F.G. and C.R. are supported by the Italian Ministry of University and Research (MIUR), in the framework of PRIN project 2017FKHBA8 001 (The Time-Space Evolution of Economic Activities: Mathematical Models and Empirical Applications). The work of G.F. is partially supported by the French National Research Agency in the framework of the "Investissements d'Avenir" program (ANR-15-IDEX-02) and of the center of excellence LABEX MME-DII (ANR11-LABX-0023-01). A.D. and T.L.-T. have no funding sources to declare. M.P. thanks the support of the University of Chicago. **Author contributions:** All the authors contributed to the development of the model structure. Computation and creation of the figures were undertaken by C.R., F.G. and G.F. The manuscript was drafted by A.D. and Supplementary Materials by F.G. R.B., T.-L.T., G.F., C.R., and M.P. all contributed to the editing. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. All the code for numerical optimization and model simulation is available at https://doi.org/10.5281/zenodo.7498153.

Submitted 27 August 2022
Accepted 19 April 2023
Published 24 May 2023
10.1126/sciadv.ade6169

ScienceAdvances

Balancing economic and epidemiological interventions in the early stages of pathogen emergence

Andy Dobson, Cristiano Ricci, Raouf Boucekkine, Fausto Gozzi, Giorgio Fabbri, Ted Loch-Temzelides, and Mercedes Pascual

Sci. Adv. **9** (21), eade6169. DOI: 10.1126/sciadv.ade6169

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.