Optimally Controlling an Epidemic

Martín Gonzalez-Eiras and Dirk Niepelt

Working Paper 20.06

Optimally Controlling an Epidemic*

Martín Gonzalez-Eiras† Dirk Niepelt‡

December 11, 2020

Abstract

We propose a flexible model of infectious dynamics with a single endogenous state variable and economic choices. We characterize equilibrium, optimal outcomes, static and dynamic externalities, and prove the following: (i) A lockdown generically is followed by policies to stimulate activity. (ii) Re-infection risk lowers the activity level chosen by the government early on and, for small static externalities, implies too cautious equilibrium steady-state activity. (iii) When a cure arrives deterministically, optimal policy is dis-continous, featuring a light/strict lockdown when the arrival date exceeds/falls short of a specific value. Calibrated to the ongoing COVID-19 pandemic the baseline model and a battery of robustness checks and extensions imply (iv) lockdowns for 3–4 months, with activity reductions by 25–40 percent, and (v) substantial welfare gains from optimal policy unless the government lacks instruments to stimulate activity after a lockdown.

JEL codes: I18

Keywords: Epidemic, lockdown, forced opening, SIR model, SIS model, SI model, logistic model, COVID-19.

*For comments and discussions we thank seminar participants at ANCE, CEMFI, DTMC COVID-19 Workshop, and the University of Copenhagen.

†University of Copenhagen. Oster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: mge@alum.mit.edu. Web: alum.mit.edu/www/mge.

‡Study Center Gerzensee; University of Bern; CEPR; CESifo. P.O. Box 21, CH-3115 Gerzensee, Switzerland. E-mail: dirk.niepelt@szgerzensee.ch. Web: www.niepelt.ch.

1 Introduction

An epidemic generates trade-offs. Since economic activity involves personal interaction, higher activity accelerates infectious dynamics, with detrimental effects on public health, mortality, and future well-being. The trade-offs are particularly stark at the outset as the ongoing COVID-19 drama has shown.[1] Since it took time in many countries to adopt efficient and targeted strategies to contain the infections, governments relied on non-targeted forced reductions in economic activity—lockdowns—as a first response.

In this paper, we propose a generic framework to analyze optimal lockdown policies and we derive a series of results that seem to have gone mostly unnoticed in the recently burgeoning literature focusing on the intersection of epidemiology and economics.[2] Almost all models in that recent literature build on the classical epidemiological "SIR"-model developed by Kermack and McKendrick (1927). But this model is just one among several "compartmental" epidemiological workhorse models, and certainly not the most relevant one when re-infection risk or other factors are of primary concern.[3] In addition, SIR-models of various flavours feature two endogenous epidemiological state variables; this makes it difficult to embed economic choices in those frameworks without sacrificing analytical tractability, transparency, and generality.

The generic epidemiological framework at the basis of our economic analysis generalizes the "simple epidemiological model" (Bailey, 1975) and captures the essence of infectious dynamics, which mainly depend on the population shares of two groups, those that have contracted the disease and those that have not yet but are susceptible. New infections are driven by complementarities between the two groups, and cumulative infections therefore approximately follow a logistic law of motion. We show that a framework with a single endogenous state variable, the share of the population that has contracted the disease at some time in the past, constitutes a tractable and accurate approximation of infectious dynamics in models with additional state variables. Moreover, it offers a flexible environment to analyze how economic choices and epidemiological dynamics interact and how optimal policy shapes them.

The economic layer that we superimpose on the epidemiological framework incorporates households and a government. We assume that households derive utility from their choices of economic activity, both positive because activity generates consumption and negative because it requires effort. In addition, higher activity increases the risk of getting infected which is privately costly due to mortality risk or health care expenses. Households are fully aware of the aggregate infection dynamics and the risks they face and they behave individually rationally.[4]

[1] The fact that some countries do better than others in terms of both economic and public health outcomes does not invalidate the presence of trade-offs but indicates differences in the quality of government responses.

[2] Early papers in that recent literature include Atkeson (2020) and Eichenbaum et al. (2020). Within a couple of weeks many papers by other authors have followed, see for instance CEPR's *Covid Economics: Vetted and Real-Time Papers* series.

[3] Hethcote (1989; 2000) reviews workhorse epidemiological models.

[4] Goolsbee and Syverson (2020) document voluntary behavioral changes in the U.S. before states introduced lockdown measures.

Nevertheless, the privately optimal activity choices fail to fully internalize the social consequences, and this gives rise to "static" and "dynamic" externalities. The static externality reflects differences between the immediate private and social marginal costs of activity. Such differences may arise for two reasons: First, because households only bear a part of the direct marginal costs of infection, for example due to public health insurance; and second, because they perceive their actions to have a linear effect on health outcomes while symmetric activity choices in equilibrium may give rise to nonlinearities and stronger complementarities. The dynamic externality arises because households do not internalize the consequences of their choices for the aggregate state variable.

While the private sector always ends up shouldering the full social costs of infection the static and dynamic externalities imply that the private activity choices are distorted. In equilibrium, activity evolves inversely with the flow of new infections as households trade off the economic net benefits and the immediate private health costs. At the outset of an epidemic, households therefore take only minor precautions while they reduce activity much more drastically when infections peak.

In contrast, the activity level implemented by a benevolent social planner reflects the full immediate social costs as well as the dynamic welfare consequences of higher infection numbers. In most, but not all, parts of the analysis we assume that the government has sufficient instruments at its disposal to implement the social planner outcome. When it does, the government imposes a lockdown in the early phase of the epidemic and curtails activity below the privately optimal level. This slows down infection dynamics and shifts peak infections into the future.

That infection dynamics give rise to externalities and call for corrective government action is well understood (Gersovitz and Hammer, 2004). More surprisingly, we find that this corrective action can go both ways. One of our results characterizes the optimal timing not only of lockdowns but also of "inverse lockdowns," namely interventions that aim at stimulating private sector activity. Arguably, some measures resembling such inverse lockdowns have been imposed or promoted in the context of the COVID-19 pandemic, for instance in the form of monetary easing, temporary sales tax reductions, employment subsidies, or "return-to-work bonuses."[5]

We find that inverse lockdowns are imposed even if individual households fail to fully internalize the immediate social costs of activity, that is, even when the static externality is negative. Intuitively, as the infection spreads, the dynamic externality eventually becomes positive and as long as the static externality remains limited in size the total externality turns positive as well. At a later stage of the epidemic the optimal government intervention therefore amounts to stimulating activity.

Underlying the dynamic externality and the fact that it eventually becomes positive are capital losses and gains. Early in the epidemic, the value function of society is decreasing in the aggregate state, i.e., the stock of previously infected households. Intuitively, peak infection rates and the associated costs of infection and of reduced economic activity are approaching and this pushes the present value of future suffering up. But after a point,

[5]In the U.S. the director of the White House National Economic Council urged lawmakers in June 2020 to replace a USD 600-a-week lump-sum transfer with a "return-to-work bonus" (*The Wall Street Journal*, June 15, 2020).

the slope of the value function turns positive because higher cumulative infection numbers imply fewer future infections to suffer from, and the advent of a return to normality. This positive implication of infection flows is not internalized by the private sector and provides the rationale for the government to stimulate activity.

When infection and subsequent recovery does not confer permanent immunity the calculus of optimal government intervention changes.[6] We model this scenario under the assumption that a share of the population that contracted the disease in the past becomes susceptible again, opening the path for recurrent infections. With such re-infection risk, the economy converges to a—policy dependent—steady state with strictly positive rather than zero infection flows; that is, the disease becomes endemic. As a consequence, not only the infection dynamics are generically inefficient if the government does not intervene, but the steady state is as well.

We prove that, absent a static externality, the government's choice of activity level necessarily exceeds the activity level in equilibrium. When the static externality is sufficiently large or society cares little about the future, in contrast, the opposite result holds true. Independently of these steady-state properties the government's activity level early on in an epidemic, when a lockdown is in place, decreases with re-infection risk.

Another key result of our analysis concerns the optimal policy when a cure like a vaccine or effective treatment is anticipated to arrive deterministically such that time constitutes another state variable. Such a scenario might reflect, for example, that several vaccine candidates, of which at least one will work with near certainty, undergo final trials or require final regulatory approval (as is the case in Fall 2020 in the context of COVID-19), or that a poor country waits for supplies of vaccine doses that are provided by donor institutions.

We prove that in this kind of situation the exact date at which the cure becomes available is critical in the sense that the strictness of an optimal lockdown is dis-continuous. When the arrival date exceeds a specific value then the lockdown is very light or the government even imposes an inverse lockdown. But when the arrival date falls short of the value, even just, then the optimal lockdown is very strict. Intuitively, as the arrival date increases, the optimal policy first calls for lower and lower activity in order to keep total infections in check until the cure arrives. Eventually, the costs of curtailing activity by more and more, for longer and longer become prohibitive and it is optimal to "give up" on infections.

Our framework with a single endogenous state variable can easily be solved and simulated numerically. Calibrated to match COVID-19 infection data in the U.S. the model's baseline version suggests that starting from mid March 2020 a lockdown should optimally have lasted for three months with economic activity reduced by twenty five percent. Compared with laissez faire the optimal policy would have increased welfare on the order of six percent of lifetime consumption. Moreover, exploiting the model's tractability and flexibility we run a battery of robustness checks. We change key parameter values (for instance, the intertemporal elasticity of substitution); allow for constant or increasing returns of activity on infections; and analyze different specifications of the cost function

[6] In the context of the COVID-19 pandemic scientists currently cannot rule out that individuals with antibodies may re-contract the disease. See, for example, Gudbjartsson et al. (2020) and To et al. (2020).

associated with infections—allowing for congestion effects and "flattening-the-curve" motives as well as learning effects. We allow for virus mutations, enhancements in therapy or test-and-trace technology, or multiple infection waves by introducing stochastic change of epidemiological parameters. And we study the effects of restricting the government's instruments such that inverse lockdowns become infeasible.

Across almost all of these alternative specifications, as well as the extensions discussed previously, the basic message remains one of optimal lockdowns for three to four months, with activity reductions by twenty five to forty percent. Only when activity enters the law of motion quadratically or in the presence of re-infection risk does the optimal lockdown extend over a much longer period, around six to seven months. Only when the intertemporal elasticity of substitution equals one half rather than unity is the reduction of activity during the lockdown much smaller, roughly twelve percent; and only with re-infection risk is it much larger, close to fifty percent. The welfare gains from the optimal government intervention are on the order of three to nine percent of lifetime consumption. Only when the government can impose a lockdown but lacks instruments to subsequently stimulate activity are the welfare gains nonsignificant.

Through most of the paper we assume that infection status is unobserved and as a consequence, that all households choose the same activity level. While this is a plausible approximation in the context of some epidemics, including the COVID-19 pandemic, it is less so for others. In a final extension we therefore consider the consequences of either publicly or privately observed infection status. We find that when ten percent of infected households become aware of their status this has only minor effects on the strictness of the optimal lockdown and the associated welfare effects, increasing the optimal lockdown duration by two weeks.

Related Literature Workhorse epidemiological models due to Kermack and McKendrick (1927) and Bailey (1975) are reviewed, e.g., in Hethcote (1989) and Hethcote (2000). To derive our generic epidemiological framework and to calibrate the model we rely on this literature; the mapping between different frameworks laid out in Gonzalez-Eiras and Niepelt (2020b); as well as on Atkeson (2020), Ferguson et al. (2020), Greenstone and Nigam (2020), Hall et al. (2020), and Russell et al. (2020).

Since mid March 2020 there has been an explosion of papers focusing on the intersection of epidemiological dynamics and economic cost-benefit analysis. Early contributions include Atkeson (2020) and Eichenbaum et al. (2020). Alvarez et al. (2020) compute the optimal lockdown policy when there is a rationale to flatten the infection curve in order to relax health care system capacity constraints. Farboodi et al. (2020) argue that in equilibrium and under the optimal policy the effective reproduction number always remains close to unity. Gersovitz and Hammer (2004), Bethune and Korinek (2020) and Jones et al. (2020) assess externalities. Kaplan et al. (2020), Acemoglu et al. (2020), and Ellison (2020) analyze the implications of heterogeneity, including differential costs of reduced activity, asymmetric infection dynamics due to "super spreaders," age-dependent fatality rates, or welfare losses due to nontargeted measures. Çenesiz and Guimaraẽs (2020) and Giannitsarou et al. (2020) analyze immunity loss and demographic dynamics. Taking spatial aspects into account Bisin and Moro (2020) show how local interactions

give rise to matching frictions and local herd immunity effects.

Most of this work focuses almost exclusively on numerical analyses, with Toxvaerd (2020), Gonzalez-Eiras and Niepelt (2020a), Abel and Panageas (2020), and Miclo et al. (2020) constituting some notable exceptions.[7] Our paper combines analytical results with numerical simulations. Because the model captures the essence of epidemiological dynamics and economic responses to it, our analytical results on capital gains, inverse lockdowns, steady-state inefficiencies, and dis-continous policy functions should apply much more broadly.[8] In fact, in parallel work by Garibaldi et al. (2020) who study a discrete-time setup with multiple states a dynamic externality similar to the one in our model is present; we highlight that such dynamic externalities necessarily start negative and eventually become positive. Moreover, our paper presents a battery of extensions and robustness checks and contrasts various epidemiological and economic environments: Laissez faire vs. optimal policy; observable vs. unobservable infection status; lockdowns vs. forced openings; stochastic vs. deterministic arrival of a cure; permanent vs. temporary immunity, generating either a disease-free steady state or an endemic equilibrium; or, stationary epidemiological environments vs. environments with changing characteristics. Our simulations offer insights into how these different features, some of which have individually been studied before, affect the optimal policy.

Outline The remainder of the paper is organized as follows. We lay out the model in section 2 and present the conditions characterizing equilibrium and first best in section 3. The analysis of the baseline model is contained in section 4. Sections 5–8 consider various modifications and extensions. In particular, section 5 offers a host of robustness checks. We consider alternative specifications of the law of motion, the costs of infection, or of parameter values as well as stochastic change in the infection rate, a stochastic arrival of a new type of disease, or restrictions on government instruments. Section 6 analyzes the consequences of lack of immunity. Section 7 derives the implications of a deterministic rather than stochastic arrival of a cure. Section 8 considers the implications of observable infection status and thus heterogeneity. Section 9 concludes. Appendix A discusses the calibration against the background of a comparison of epidemiological models. All proofs are relegated to appendix B.

2 The Model

We consider an infinite horizon economy with households and a government. Time is continuous and indexed by $t \geq 0$.

[7] Toxvaerd (2020) characterizes privately optimal social distancing; Gonzalez-Eiras and Niepelt (2020a) characterize optimal lockdown policies; and Abel and Panageas (2020) characterize the optimal steady state in a SIR model with vital dynamics. Miclo et al. (2020) derive the optimal policy under a capacity constraint.

[8] Farboodi et al. (2020) briefly consider the case with a deterministic vaccine arrival but they do not find our result that optimal policy may be dis-continuous.

2.1 Epidemiology

We adopt an epidemiological framework that is closely related to several canonical models in the epidemiological literature: The SIR model due to Kermack and McKendrick (1927), a modified SIR model and the simple epidemic model, the SI model, due to Bailey (1975), and the SIS model derived from it.[9] Our framework incorporates one endogenous state variable (rather than the two in the typical SIR model), possibly time as a second state variable (unlike SIR and SIS models), as well as economic activity (unlike SIR and SIS models).

2.1.1 Dynamics

At time t the population consists of $x(t)$ "pre-infection" (for short: "pre") households; $1-\bar{y}$ "neutral" households; and $y(t) = \bar{y} - x(t)$ "post-infection" ("post") households. Members of the post group have been infected in the past, members of the pre group might be infected in the future, and members of the neutral group cannot be infected, for instance because they are immune. We allow for members of the post group to return to the pre group; this captures the fact that households might repeatedly be infected because they do not permanently develop immunity. Since we normalize the mass of the total population to unity, the mass of each population group also represents the group's population share,

$$0 \leq y(t) \leq \bar{y} \leq 1.$$

The initial population shares of the pre and post groups are given by $x(0) = \bar{y} - y(0) > 0$ and $y(0) > 0$.

While the infection status of neutral households never changes post households transmit the disease to members of the pre group according to a logistic law of motion. Moreover, as mentioned before, some post households may return to the pool of pre households. As a consequence, the share of post households evolves according to the law of motion

$$\dot{y}(t) = g(a(t))\,\beta\,y(t)\,(\bar{y} - y(t)) - \gamma y(t), \tag{1}$$

where a dot (as in $\dot{y}(t)$) denotes the time derivative. By definition, $\dot{x}(t) = -\dot{y}(t)$.

Variable $a(t)$ in the law of motion denotes an index of economic activity. We assume that higher activity fosters infections, that is function g is increasing and smooth. Parameters $\beta > 0$ and $\gamma \geq 0$ capture the biological characteristics of the disease. According to equation (1), the number of pre households that get infected depends on their number, $x(t) = \bar{y} - y(t)$; the infection rate, $g(a(t))\beta$; and the number of post households, $y(t)$. The total increase of post households equals the number of newly infected post households net of the share γ of post households who return to the pre group.

When $\gamma = 0$ post households cannot be re-infected. The law of motion then implies that for any $g(a(t))\beta > 0$ and as long as $y(0) > 0$, as we assume, the share of the post group is strictly increasing over time and converges to $\bar{y}$; conversely, the share of the pre

[9] The "S," "I," and "R" in SIR, SI, and SIS stands for "susceptible," "infectious," and "removed," respectively. See Hethcote (1989) and Hethcote (2000) for an overview over epidemiological models of infectious diseases.

group is strictly decreasing and converges to 0 in this case. When $\gamma > 0$, in contrast, dynamics need not be monotone unless $a(t)$ is constant.

When $a(t) = a$, equation (1) has the solution[10]

$$y(t) = \frac{\bar{y} - \frac{\gamma}{g(a)\beta}}{1 + \mathrm{e}^{-(g(a)\beta\bar{y}-\gamma)t}\left(\frac{\bar{y}}{y(0)} - 1 - \frac{\gamma}{g(a)\beta y(0)}\right)} \tag{2}$$

and $y(t)$ converges to zero or to $y_\infty(a) \equiv \bar{y} - \gamma/(g(a)\beta)$, depending on whether $g(a)\beta\bar{y}$ is smaller or larger than γ. Throughout, we consider the latter case; that is, we focus on the case where $g(a)\beta\bar{y} > \gamma$ such that $y_\infty(a) > 0$. For future reference we note that solving equation (2) for t yields

$$t(y_0, y) = \ln\left(\frac{y(y_\infty(a) - y_0)}{y_0(y_\infty(a) - y)}\right)/(g(a)\beta\bar{y} - \gamma), \quad y_0 < y < y_\infty(a). \tag{3}$$

We assume that with Poisson arrival rate ν the disease and its consequences (described below) disappear and β drops to zero (a "cure arrives"). For example, this might be due to medical progress or the development of a vaccine. We also allow for the possibility that the disease deterministically disappears in finite time, at date T.

2.1.2 Costs of Infection

Infections impose costs, for example because the health care system requires resources, output is lost, utility foregone, or persons die. We represent such costs as a function of the intensity of transitions from the pre to the post group—new infections—that is, as a function of $\dot{y}(t) + \gamma y(t)$. We assume that the direct and indirect social costs associated with these transitions equal

$$\psi\, g(a(t))\, \beta\, y(t)\, (\bar{y} - y(t)), \tag{4}$$

where $\psi > 0$ denotes the unit costs of infection. In section 5 we allow ψ to depend on the state, $\psi(y(t))$. This allows us to analyze learning or congestion effects. In the former case, the unit costs decrease with cumulative case numbers and in the latter, the unit costs increase with infection flows, for instance due to capacity constraints in the health care sector that lead to a deterioration of care quality and increased fatality rates, generating a motive to "flatten the curve."[11] In the baseline analysis, we assume that ψ is a constant.

To save on notation, we sometimes use the definition

$$\dot{y}^g(t) \equiv g(a(t))\, \beta\, y(t)\, (\bar{y} - y(t)) = \dot{y}(t) + \gamma y(t)$$

for the gross flow (as opposed to the net flow $\dot{y}(t)$) from the pre to the post group.

[10] We assume that $g(a)\beta\bar{y} \neq \gamma$. See for example Hethcote (1989) for the case of $g(a) = \bar{y} = 1$.

[11] For an analysis of optimal policy with a specific focus on capacity constraints, see for example Miclo et al. (2020).

2.1.3 Relation to SIR and SIS Models

The law of motion (1) nests a range of well known epidemiological models, augmented by an effect of economic activity on the infection rate. For example, under the restrictions $\gamma = 0$ and $\bar{y} = 1$ it corresponds to the SI model (Bailey, 1975) in which the number of transitions from the pre to the post group first increases and then decreases as $y(t)$ moves from near zero towards unity.

The law of motion subject to $\gamma = 0$ also corresponds to a special case of the canonical SIR model (Kermack and McKendrick, 1927) and the modified SIR model (Bailey, 1975). In their general form, both SIR models characterize the evolution of three population groups—susceptible, currently infected, and removed households—and therefore contain two endogenous state variables (the shares of two of the three groups). Susceptible households are infected by currently infected households, they remain infectious for a random time span, and eventually they join the pool of removed (recovered or deceased) households. These dynamics reduce to the law of motion (1) with $\gamma = 0$ when the distinction between currently infected and removed households are blurred and the two are combined into a single group of post households.[12]

Importantly, blurring the distinction between currently infected and removed households does not undermine the model's capacity to represent societal costs of infection or death, for representing these costs does not require to explicitly account for the *stock* of currently infected households or the deceased population. It suffices to account for the *flow* of infections, that is the flow from the pre to the post infection state, and to associate costs with this flow.[13] In Gonzalez-Eiras and Niepelt (2020b) we offer a detailed discussion of the connection between SIR models and the law of motion (1). We establish theoretically and based on numerical examples that the law of motion (1) allows to flexibly capture epidemiological dynamics very much in line with those in conventional SIR models.

While the canonical SIR model (Kermack and McKendrick, 1927) and the modified SIR model (Bailey, 1975) yield similar predictions for transition dynamics they differ with respect to their implications for the long-run share of the population that never gets infected. In the canonical SIR model this long-run share is endogenous while in the modified SIR model it always equals zero. A hybrid model augments the modified SIR model with an additional parameter that allows to regulate the long-run population shares (Gonzalez-Eiras and Niepelt, 2020b). In the law of motion (1) the parameter $\bar{y}$ plays a similar role. A lower $\bar{y}$ implies a larger share of the population that never gets infected. While the exogeneity of $\bar{y}$ is restrictive in comparison with the endogenous herd immunity level in some SIR models the law of motion (1) still allows to analyze long-run cumulative infection levels. Specifically, since we allow for the random arrival of a cure, slowing down infections constitutes a powerful strategy to increase the expected population share that

[12] That is, when the transition rate from currently infected to removed in the SIR model equals zero.

[13] Since the law of motion (1) does not explicitly account for deaths it does not account for changes in the population size and in implied population shares due to death. These effects are negligible when death rates are small as we assume. Blurring the distinction between currently infected and removed households also amounts to assuming that the groups are indistinguishable in terms of their economic characteristics (e.g., productivity). We view the consequences of this assumption as minor as well.

is never infected.

The law of motion (1) also is closely related to the SIS model in which households, once infected, randomly recover and return to the susceptible pool (rather than the removed pool as in SIR models) because infection does not confer immunity (Hethcote, 1989).[14] Our framework differs from the SIS model insofar as we represent infections in terms of *flows*, $\dot{y}^g(t)$, while in the SIS model $y(t)$ corresponds to the *stock* of currently infected persons. In the SIS model a higher value of γ decreases the steady-state stock y while in our framework the relationship between γ and the steady-state infection flow $\dot{y}^g_\infty = \gamma y_\infty$ is inverse-U-shaped. In an extended SIR model with loss of immunity (a SIR-S model) the effect of γ on y_∞ would be positive, contrary to the SIS model. We view such a positive effect of γ on (the costs of) infections as more plausible and in our specification based on the law of motion (1), we therefore restrict attention to small values of γ which guarantee that γy_∞ increases in γ. Specifically, we require that $g(a_\infty)\beta\bar{y} > 2\gamma$.[15]

In conclusion, the law of motion (1) nests standard epidemiological models with one endogenous state variable and constitutes a flexible and tractable alternative to other standard models with two endogenous state variables (see also the discussion in appendix A). Relative to epidemiological models it introduces an effect of economic activity on infection dynamics.

We summarize our assumptions regarding the epidemiological part of the model as follows:

Assumption 1. Epidemiological dynamics are described by the law of motion (1) with $g(a)\beta\bar{y} > 2\gamma$ for any a along the equilibrium path. Function g is smooth, strictly positive, and strictly increasing. The social costs of infection are given by (4).

2.2 Economics

2.2.1 Households

A household i chooses the activity level $a_i(t)$ over time in order to maximize an intertemporal objective which accounts for the immediate economic effects of activity and for the costs of infection that the household bears. Households take aggregates as well as the law of motion (1) as given and discount the future at the rate ρ.

We represent the immediate economic effects by an indirect utility function, u, that depends on the individual choice, $a_i(t)$, and satisfies the following assumptions:

Assumption 2. Function u is smooth, twice differentiable, and strictly concave. It satisfies $\lim_{a\downarrow 0} u'(a) = \infty$ and reaches a maximum at $a^\star \in (0, \infty)$, $u'(a^\star) = 0$.

We represent the household's costs of infection as the product of two terms: The social costs given in (4) and a factor ξ that depends on the household's and the aggregate

[14] The SIS model has $\bar{y} = 1$.

[15] Both in the SIS model and our framework the steady-state values satisfy $g(a_\infty)\beta(\bar{y} - y_\infty) = \gamma$. In the SIS model we therefore have $dy_\infty/d\gamma = -1/(g(a_\infty)\beta)$ while in our framework, $d\dot{y}^g_\infty/d\gamma = d(\gamma y_\infty)/d\gamma = y_\infty - \gamma/(g(a_\infty)\beta) = \bar{y} - 2\gamma/(g(a_\infty)\beta)$. In a SIR-S model the steady-state values satisfy $dy_\infty/d\gamma > 0$ (because $\beta x_\infty y_\infty = c y_\infty = \gamma z_\infty$, see appendix A).

activity level. Accordingly, the household's costs equal

$$\xi(a_i(t), a(t))\, \psi \dot{y}^g(t).$$

Let ξ_{a_i} denote the partial derivative of ξ with respect to $a_i(t)$. We make the following assumption regarding ξ:

Assumption 3. Function ξ is smooth, twice differentiable, and convex in $a_i(t)$. Moreover, $\xi(a(t), a(t)) \equiv 1$ and $\xi_{a_i}(a_i(t), a(t))$ is homogeneous.

Convexity implies that the household's marginal costs of infection are nondecreasing in the household's activity level. The condition $\xi(a(t), a(t)) \equiv 1$ imposes the natural restriction that the total (but not necessarily marginal) private costs sum to the social costs. The homogeneity assumption is made for tractability reasons.

Assumption 3 allows for a variety of cost specifications and our formal results do not require any assumptions regarding ξ beyond those introduced before.[16] Our preferred specification is given by

$$\xi(a_i(t), a(t)) \equiv \zeta \frac{a_i(t)}{a(t)} + (1 - \zeta), \quad \zeta \in [0, 1],$$

where we interpret ζ as the share of costs that a household perceives to depend on its own actions. Specifically, suppose that a pre household transitions to the post pool with a probability that is proportional to aggregate gross infection flows as well as its relative activity level, $a_i(t)/a(t)$. A share σ of the newly infected develops symptoms generating health care costs h_0 per case which are covered by taxpayers, as well as costs h_1 which are actuarially fairly insured by an insurer whose premia are conditioned on $a_i(t)/a(t)$. In addition, a share $\delta \leq 1$ of the newly infected dies, with associated costs d.[17] This maps into our formal structure as follows:

$$\begin{aligned} \psi &= \sigma(h_0 + h_1 + \delta d), \\ \xi(a_i(t), a(t)) &= \frac{h_1 + \delta d}{h_0 + h_1 + \delta d} \frac{a_i(t)}{a(t)} + \frac{h_0}{h_0 + h_1 + \delta d} = \zeta \frac{a_i(t)}{a(t)} + (1 - \zeta). \end{aligned}$$

Note that $\xi(a_i(t), a(t)) = 1$ in equilibrium.

To preserve a specification with $y(t)$ as the single endogenous state variable we need to assure that a household's infection status is not a state in that household's program. To that effect we assume that households do not know their infection status (whether they are asymptomatic or develop symptoms), for example because the cause of sickness

[16] For example, strict convexity of ξ would follow with $\xi(a_i(t), a(t)) \equiv \zeta \left(\frac{a_i(t)}{a(t)}\right)^{\omega} + (1 - \zeta),\ \omega > 1$.

[17] In the law of motion (1) we treat $\sigma\delta$ as negligible (and assume symmetric activity choices of all households, for reasons discussed shortly). If we explicitly incorporated it then the population size at date t would equal $1 - \sigma\delta y(t)$; the law of motion would be $\dot{y}(t) = g(a(t))\beta y(t)(1 - \sigma\delta)(\bar{y} - y(t)) - \gamma y(t)(1 - \sigma\delta)$; and the payoff from activity would equal $u(a(t))(1 - \sigma\delta y(t))$. The programs that we analyze in section 3 therefore would remain unchanged except that β and γ would be multiplied by $1 - \sigma\delta \approx 1$ and u and $U^\star$ by $1 - \sigma\delta y(t) \approx 1$.

cannot be identified from symptoms or because all survivors remain asymptomatic (i.e., $\delta = 1$).[18] As a consequence every household bears a share $\xi(a_i(t), a(t))$ of the social costs, independently of its infection status, and all pre, post, and neutral households choose the same activity level.

We view the assumption that households are unaware of their infection status—and therefore behave symmetrically—as a plausible approximation in the context of many epidemics, including the COVID-19 pandemic.[19] However, it is clearly less plausible in the context of other epidemics. In an extension discussed in section 8 we therefore relax the assumption and analyze whether privately or publicly observable infection status, for instance related to the development of symptoms, changes the results.

2.2.2 Government

We assume that policy makers have instruments to control economic activity along the activity-infection margin, for instance by imposing social distancing measures, closure of non-essential businesses, or other lockdown measures. If so desired, the government may also stimulate activity. Using these instruments the government faces the same program as a social planner.

In section 5 we analyze the consequences of instrument restrictions which imply that the government faces a more limited choice set than a social planner.

2.3 Functional Form Assumptions and Calibration

To sharpen analytical results we sometimes impose the preference assumption

$$u(a) = \ln(a) - a + 1,$$

implying $a^\star = 1$ and $u(a^\star) = 0$.[20] We use this functional form because it is flexible and convenient and yields a tractable first-order condition.[21] Also, we sometimes impose the assumption

$$g(a) = a^n, n = 1, 2,$$

for the effect of activity in the law of motion (1). This specification allows for both constant and increasing returns to scale as far as the effect of activity on infectious dynamics is concerned.[22]

[18] An alternative assumption could be that excess continuation values relative to the continuation value of the average household are fully taxed away.

[19] Many individuals that have been infected with COVID-19 have remained asymptomatic or shown only mild symptoms, see also the discussion in Farboodi et al. (2020, p. 39). As long as COVID-19 tests are only selectively applied, as is still the case in many places at the time of this writing, most infected individuals necessarily behave like individuals who have not been infected.

[20] Scaling the utility function by a constant would not affect the results.

[21] Several other authors adopt the same preference specification, see for instance Çenesiz and Guimaraẽs (2020) or Farboodi et al. (2020).

[22] While in the canonical SIR model a doubling of population shares implies a quadrupling of infections this is not the case in a modified SIR model (Gonzalez-Eiras and Niepelt, 2020b). The epidemiological evidence on constant versus increasing returns to scale ("frequency dependence" versus "density depen-

Throughout the paper we use numerical simulations to illustrate our results. The simulations make use of the functional form assumptions described above and are based on parameter values that we calibrate to match properties of the COVID-19 pandemic. Our unit of time is a day and $t = 0$ corresponds to mid March 2020. Accordingly, we set $\rho = -\ln(0.95)/365$ (five percent annual discount rate) and $\nu = 1/(365 * 1.5)$ (one-and-a-half years expected duration until discovery of a vaccine).[23]

In appendix A we describe in detail how we calibrate the remaining parameters, using information about parameter values in the canonical SIR model and theoretical results connecting SIR models and the logistic model.[24] This yields $y(0) = 0.1893 \cdot 10^{-3}$, $\beta = 0.9660 \cdot 10^{-1}$ (corresponding to an infection rate in the SIR model (at normal activity level) of 0.1333), and $\bar{y} = 0.75$. We set $\gamma = 0$ except in section 6 where we analyze the consequences of lack of immunity.

To calibrate the derivative $\xi_{a_i}(a, a)$ and the social cost parameter ψ we use estimates of expected health care and mortality costs as well as households' willingness to pay to eliminate COVID-19 induced mortality risk.[25] We assume that households fully internalize mortality risk but not the social marginal costs of health care implying a social cost parameter $\psi = 227.1$ and an internalization rate $\zeta = 0.8266$. Table 1 summarizes the baseline calibration.

Parameter	Value
ρ	$0.1405 \cdot 10^{-3}$
ν	$0.1826 \cdot 10^{-2}$
$y(0)$	$0.1893 \cdot 10^{-3}$
β	$0.9660 \cdot 10^{-1}$
$\bar{y}$	0.7500
ψ	$0.2271 \cdot 10^{3}$
ζ	0.8266

Table 1: Baseline calibration. See the text and appendix A for explanations.

dence") is mixed (Hethcote, 1989). Acemoglu et al. (2020) allow for $n \in [1, 2]$. Farboodi et al. (2020) opt in their baseline for the quadratic formulation, taking as a reference the distinction made by Diamond and Maskin (1979).

[23] See, e.g., Alvarez et al. (2020). The probability of discovery until time t equals $1 - \mathrm{e}^{-\nu t}$; the expected duration until discovery thus equals $\int_{t=0}^{\infty} t\nu \mathrm{e}^{-\nu t} dt = \nu^{-1}$.

[24] We rely on parameter estimates by Atkeson (2020), Ferguson et al. (2020), Greenstone and Nigam (2020), Hall et al. (2020), and Russell et al. (2020).

[25] We rely on parameter estimates by Bartsch et al. (2020), Hall et al. (2020), and Menachemi et al. (2020).

3 First Best and Equilibrium

Let $U^\star \equiv u(a^\star)/\rho$ denote the value of the representative household when first-best activity is chosen permanently and no costs of infection occur. The value $U^\star$ is attained once a vaccine is developed or all households have gained immunity because $y(t) \approx \bar{y}$ and $\gamma = 0$.

We represent the government's optimality conditions as well as the equilibrium conditions recursively. The state in the program of the government or an individual household is given by (y, t). When $T = \infty$ such that there is no deterministic terminal date of the disease then the state only includes y.

3.1 Government Program

As discussed earlier, the government effectively solves the program of a social planner (except when it lacks sufficient instruments as in the setting analyzed in section 5). Letting V denote the value function of the government, its Hamilton-Jacobi-Bellman (HJB) equation reads

$$(\rho + \nu)V(y,t) = \max_a u(a) - \psi \dot{y}^g(y,a) + V_t(y,t) + \dot{y}(y,a)V_y(y,t) + \nu U^\star$$

subject to (1) where value function subscripts denote partial derivatives.[26] The left-hand side of the HJB equation represents the risk-adjusted required return on the value and the right-hand side represents the dividend and capital gains components of the return. The dividend component contains the immediate economic net benefit of activity net of the costs of infection. The capital gains component reflects the change in the value due to variations in the state or a sudden arrival of a cure; in the latter case all households immediately adopt the activity level $a^\star$ and they (as well as the government) attain $U^\star$.

Recall from (1) that the derivative of $\dot{y}^g(y,a)$ and $\dot{y}(y,a)$ with respect to a equals $g'(a)\beta y(\bar{y} - y)$. The activity level chosen by the government thus solves the first-order condition

$$u'(a(y,t)) = g'(a(y,t))\beta y(\bar{y} - y)\left(\psi - V_y(y,t)\right). \tag{5}$$

3.2 Decentralized Equilibrium

The value function of the representative household, U, solves the HJB equation

$$(\rho+\nu)U(y,t) = \max_{a_i} u(a_i) - \xi(a_i, a(y,t))\,\psi \dot{y}^g(y, a(y,t)) + U_t(y,t) + \dot{y}(y, a(y,t))U_y(y,t) + \nu U^\star$$

subject to (1). As in the case of the government's HJB equation the right-hand side of the household's HJB equation represents dividend and capital gains components. Unlike the government, however, an individual household does not perceive its activity choice, a_i, to influence the law of motion of the aggregate state although in equilibrium, individual and aggregate activity levels coincide ($a_i = a(y,t)$). The programs of the government and

[26] Since we switch to recursive notation $\dot{y}$ is now a function of y and a.

an individual household therefore only differ insofar as the government appreciates the symmetry of activity levels across households.

The household's first-order condition with respect to a_i yields

$$u'(a_i) = \xi_{a_i}(a_i, a(y,t))\, \psi \dot{y}^g(y, a(y,t)) \text{ s.t. } (1).$$

Since individual and aggregate choices coincide in equilibrium the activity level in the decentralized equilibrium satisfies

$$u'(a(y,t)) = \xi_{a_i}(a(y,t), a(y,t))\, \psi g(a(y,t)) \beta y(\bar{y} - y), \tag{6}$$

where we substitute equation (1).

4 Baseline Analysis

We now turn to the analysis of equilibrium and optimal policy in the baseline model.

4.1 Optimal Allocation

If the initial value of y equals zero the law of motion (1) implies that there will be no infections. Accordingly, the government's optimal choice is to engage in the ideal level of economic activity, $a^\star$, such that $V(0,t) = U^\star$. The same holds true for $y = \bar{y}$ but only if $\gamma = 0$ (no re-infections): if $\gamma = 0$ then $V(\bar{y}, t) = U^\star$.

For any other value of y the law of motion in combination with $g' > 0$ (assumption 1) and the equilibrium requirement $a(y,t) > 0$ (assumption 2) implies that the government faces current and future costs of infection as well as, possibly, losses due to reduced economic activity. Since the discount rate is finite this implies that $V(y,t) < U^\star$ for all $y \in (0, \bar{y})$, and also $V(\bar{y}, t) < U^\star$ if $\gamma > 0$. Since V is continuous, as established in the following lemma, $V(y,t)$ is decreasing in a neighborhood of $y = 0$.

Lemma 1. Under assumptions 1 and 2, $V(0,t) = U^\star$ and $V(y,t) < U^\star$ for all $y \in (0, \bar{y})$. Moreover, if $\gamma = 0$, then $V(\bar{y}, t) = U^\star$; if $\gamma > 0$, then $V(\bar{y}, t) < U^\star$; and V is continuous.

Figure 1 illustrates additional properties of the solution of the government's problem under the assumption that there is no re-infection risk ($\gamma = 0$) and focusing on the time autonomous case ($T = \infty$). When $\gamma = 0$ the appropriate terminal condition to solve the government's HJB equation is $\lim_{y \to \bar{y}} V(y) = U^\star$ (and parallel for the decentralized equilibrium discussed below). The top left panel of the figure displays the government's value function, the top right panel the government's choice of activity level, and the left panel in the bottom row infections, all as functions of y.[27] We discuss the other panel later.

[27]The figure is drawn using the baseline calibration introduced earlier. Unless otherwise noted, we use that calibration throughout when presenting numerical examples. When we plot functions of the state or of time it is always understood that these relationships apply before a cure arrives. (Afterwards $a(t)$ would equal unity even if $y < \bar{y}$, etc.) Also, for simplicity we omit qualifications of this type when stating our formal results.

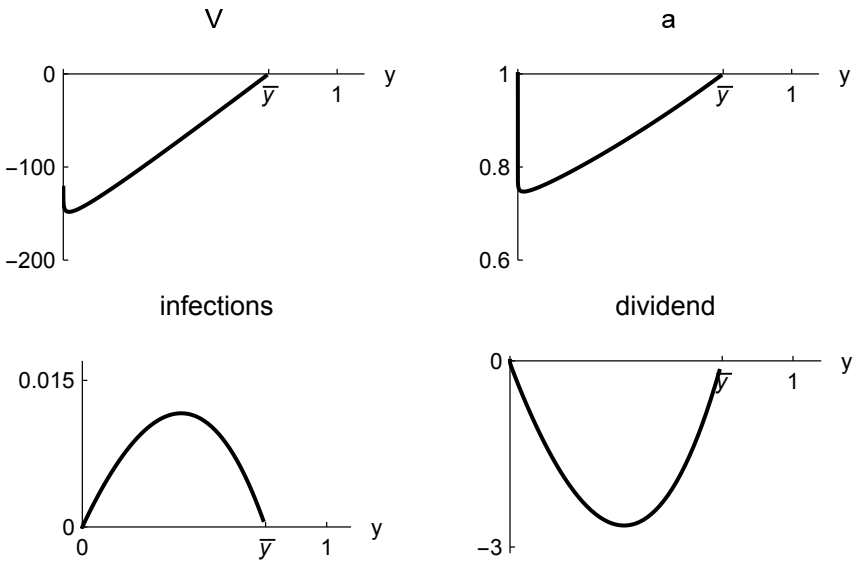

Figure 1: Value function, activity level, infections, and dividend in the government's program.

We note several general features. First, there exists a $y^{\min} \in (0, \bar{y})$ such that V attains its global minimum at $y^{\min}$ with $V(y^{\min}) < U^\star$. This follows directly from lemma 1. Clearly, higher costs of infection (ψ) lower the government's value function. The following lemma further characterizes $y^{\min}$:

Lemma 2. Under assumptions 1 and 2 and if $T = \infty$, V has a unique minimum at $y^{\min} \leq \bar{y}/2$. Moreover, if $\gamma = 0$, parameter changes that imply a higher (lower) $y^{\min}$ also imply less (more) pronounced convexity of V around $y^{\min}$.

Second, the optimal choice of activity is not symmetric although the function $\beta y(\bar{y}-y)$ is symmetric around the point $y = \bar{y}/2$. The asymmetry follows from the fact that the government's first-order condition accounts for the effect of activity on both the costs of infection (this alone would yield a symmetric policy function) and the continuation value. For example, in our benchmark specification with $u(a) = \ln(a) - a + 1$ and $g(a) = a$, the government's first-order condition (5) reads $1/a(y) - 1 = \beta y(\bar{y} - y)(\psi - V'(y))$, which reduces to

$$a(y) = \frac{1}{1 + \beta y(\bar{y} - y)(\psi - V'(y))}. \tag{7}$$

Note that this implies $(\rho + \nu)V(y) = \ln(a(y))$.[28]

Compare this to the decentralized equilibrium condition, equation (6). As we show in the proof of proposition 1 below, $\xi_{a_i}(a, a) = \xi_{a_i}(1, 1)/a = \zeta/a$ for any ξ function that satisfies assumption 3, not only for our preferred specification $\xi(a_i, a) = \zeta a_i/a + (1 - \zeta)$. Accordingly, the decentralized equilibrium condition reduces to

$$a(y) = \frac{1}{1 + \beta y(\bar{y} - y)\zeta\psi}. \tag{8}$$

[28]Using equation (7), we have $(\rho + \nu)V(y) = \ln(a(y)) + 1 - a(y)\{1 + \beta y(\bar{y} - y)(\psi - V'(y))\} + \nu U^\star = \ln(a(y)) + 1 - 1 + \nu U^\star = \ln(a(y))$.

Note that the equilibrium choice of activity is symmetric around $\bar{y}/2$.

Third, to gain intuition for the shape of V consider first the path of infections displayed on the left in the bottom row of figure 1. Infections are hump shaped because the logistic function is S shaped; the fact that $a(y)$ varies with the state does not fundamentally alter that result. As discussed earlier the hump shaped path of infections is consistent with the predictions of many epidemiological frameworks. Since the costs of infection are increasing in the number of infections they are hump shaped as well.

Consider next the dividend component of the government's return on the right-hand side of the government's HJB equation, $u(a(y)) - \psi \dot{y}^g(y, a(y))$. This dividend component is displayed on the right-hand side in the bottom row of figure 1. It reflects the costs of infection as well as the losses from reduced activity.[29] The HJB equation implies that the capital gains component of the return, $\dot{y}(y, a(y))V'(y) + \nu U^\star$, equals the required return on the government's value, $(\rho+\nu)V(y)$, net of the dividend component. Equivalently, the dividend yield and the capital gains yield add to $\rho + \nu$.

Over most of the state space the slope of V is relatively constant. Since infections are hump shaped, however, capital gains are hump shaped as well. When many infections occur, costs of infection and losses from low activity quickly materialize. Accordingly, the value function, which discounts the future costs and losses, swiftly increases; the capital gains are large and dividends low. In contrast, capital gains are negative for small values of y. Infections are still low in this part of the state space but they accelerate and as a consequence, dividends are not yet strongly depressed but the period when they will be is approaching quickly. The value function which discounts the future costs and losses reflects this.

4.2 Externalities

Individual households do not perceive their own choice of activity to affect the aggregate activity level. As a consequence they do not fully internalize the welfare consequences of higher activity. This gives rise to two externalities: a "static" one related to the contemporaneous costs of infection, and a "dynamic" one related to the effect of activity on the state and the continuation value. We analyze these externalities in turn.

A comparison of the first-order conditions (5) and (6) reveals two differences between the marginal costs of activity perceived by an individual household and the government. Subtracting the right-hand side of (5) from the right-hand side of (6) and evaluating terms at a common activity level yields $\mathcal{S}(a, y) + \mathcal{D}(a, y, t)$ with

$$\begin{aligned} \mathcal{S}(a, y) &\equiv \psi \dot{y}^g(y, a) \left(\xi_{a_i}(a, a) - \frac{g'(a)}{g(a)} \right), \\ \mathcal{D}(a, y, t) &\equiv \dot{y}^g(y, a) \frac{g'(a)}{g(a)} V_y(y, t). \end{aligned}$$

$\mathcal{S}(a, y)$ represents the static externality which arises because households do not fully internalize the effects of their choice of activity on the costs of infection. $\mathcal{D}(a, y, t)$ represents

[29] It is negative because $u(a^\star)$ is normalized to zero.

the dynamic externality which arises because households do not internalize that higher activity increases the infection rate and thus affects the continuation value.

The static externality in turn may arise for two conceptually different reasons. To see this consider the specification $\xi(a_i, a) = \zeta a_i/a + (1 - \zeta)$. The first type of static externality is present if $\zeta < 1$, that is, if households view the costs they bear to only partially depend on their own choice of activity. For example, if $g(a) = a$, then the static externality is proportional to $\zeta - 1 < 0$. The second type is present even if $\zeta = 1$ when $g(a)$ is convex, $g(a) = a^2$. In this case the static externality is proportional to $-a < 0$ because the individual household perceives a linear cost function ($\xi(a_i, a) = a_i/a$) while at the aggregate level the cost function is convex. Of course, the two types of externality may compound each other.

The following proposition characterizes the externalities for general ξ functions (recall that $\zeta \equiv \xi_{a_i}(1, 1)$):

Proposition 1. Under assumptions 1 and 3 and if $g(a) = a^n$, the total externality equals

$$\frac{\dot{y}^g(y, a)}{a} n \left(\psi(\zeta/n - 1) + V_y(y, t)\right). \tag{9}$$

The static externality is negative when $\zeta \leq n$.

The total externality is proportional to infections, $\dot{y}^g(y, a)$, because infections drive the costs of infection and change the state variable, which in turn affects the continuation value. The factor of proportionality contains two terms, $\psi(\zeta/n - 1)$ and $V_y(y, t)$. The former reflects the fact that with $\zeta < n$, households do not fully internalize the negative consequences of their actions for the costs of infection; that is, there is a negative static externality (of either type). The latter factor represents the effect of economic activity on the continuation value, due to a higher infection rate. Individual households do not internalize this effect at all, giving rise to the dynamic externality.

Figure 2 illustrates the consequences of the externalities. As in figure 1, we focus on the time autonomous case without re-infections and we let $n = 1$; as discussed previously, we calibrate $\zeta = 0.8266$, that is, households internalize roughly eighty three percent of the social costs of infection. The solid lines in the figure represent the outcomes implemented by the government and correspond to the schedules in figure 2; the dashed lines represent the equilibrium outcomes.

The externalities lower the value in equilibrium relative to the situation with government intervention, pushing the dashed line in the left panel below the solid one. More interestingly, the activity levels displayed in the right panel differ; early on, the government chooses a lower activity level than in equilibrium, later on the opposite holds true. The driving force behind the reversal is the capital gains component which lies at the source of the dynamic externality. For $y \geq y^{\min}$ the capital gains component is positive but only the government internalizes the capital gains when choosing the activity level. When this effect is sufficiently strong to compensate for the negative static externality (due to $\zeta \leq n$) then the equilibrium activity level falls short of the level chosen by the government. In the figure, the capital gains component (and dynamic externality) equals zero at $y = y^{\min} \approx 0.0212$. At $y \approx 0.0257$ the total externality equals zero and the

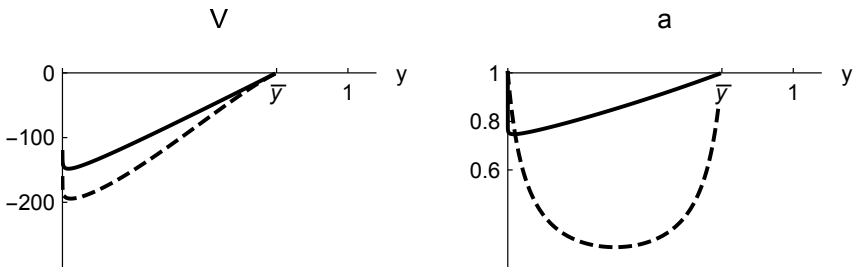

Figure 2: Value function and activity level in the government's program (solid) and in equilibrium (dashed).

activity levels chosen by the government and in equilibrium coincide. For higher values of the state the total externality is positive, that is, equilibrium activity is too *low*.

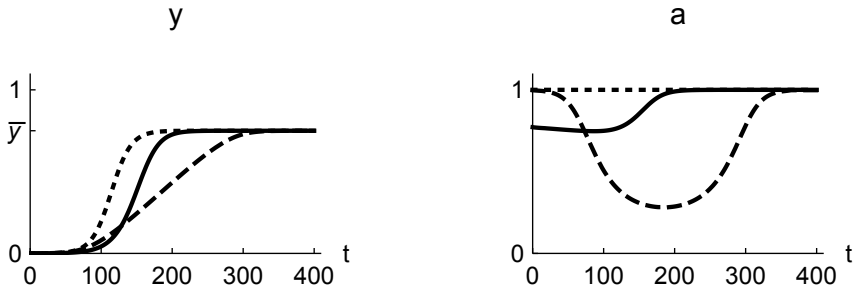

Figure 3: Infections and activity level in the government's program (solid), in equilibrium (dashed), and with no intervention (dotted).

Figure 3 illustrates how the corrective government interventions shape infections and activity over time. In contrast to figures 1 and 2 the horizontal axis now depicts time (in days), not the level of the endogenous state. The figure shows that early government intervention delays infections relative to a scenario without any adjustment of activity, and also—but much more modestly—relative to a scenario with only voluntary reductions in activity.

After around four months the government increases activity and infections rise as quickly as they would have risen a month or so earlier in the scenario without adjustment of activity. In equilibrium, in contrast, activity now falls strongly because the product $y(\bar{y}-y)$ approaches its maximum and households do not internalize the positive dynamic externality. Evaluated at $y(0)$, the share of the U.S. population infected in mid March 2020, we find $U(y(0)) \approx -173.8$ and $V(y_0) \approx -131.9$. Solving

$$\begin{aligned} \frac{1}{\rho+\nu}(\ln(a^\star(1-\phi^u)) - a^\star + 1) &= -173.8, \\ \frac{1}{\rho+\nu}(\ln(a^\star(1-\phi^v)) - a^\star + 1) &= -131.9 \end{aligned}$$

yields $\phi^u = 0.2896$ and $\phi^v = 0.2286$. Compared with the equilibrium outcome, the welfare gains due to optimal government intervention therefore amount to the equivalent of roughly six percent of lifetime consumption.

4.3 Lockdowns and Inverse Lockdowns

We refer to a "lockdown" as a situation where the government wishes to depress economic activity below the level chosen in equilibrium. Conversely, we refer to an "inverse lockdown" as a situation where the government wishes to stimulate economic activity beyond the level chosen in equilibrium. Instruments to implement a lockdown include for instance stay-at-home-orders, social distancing rules, business closures, or school closures. Instruments to implement an inverse lockdown may take the form of stimulation measures like monetary easing, temporary sales tax reductions, employment subsidies, or a "return-to-work bonus."

Our preceding analysis of externalities has direct implications for lockdowns and inverse lockdowns. We have the following result:

Proposition 2. Under assumptions 1 and 3 and if $g(a) = a^n$, the economy is in lockdown when

$$\psi(\zeta/n - 1) + V_y(y, t) < 0,$$

and in inverse lockdown when the reverse inequality holds.

Intuitively, starting from $a = a^\star$, there is a second-order loss of reducing activity but a first-order gain from slowing down infections. Unless there are zero infection dynamics both households and the government choose activity levels below $a^\star$. If the static externality (represented by the term $\psi(\zeta/n - 1)$) and the dynamic externality (represented by the term $V_y(y, t)$) combined are negative then the government perceives a larger first-order gain from lowering activity. In this case the government imposes a lockdown in order to correct the distorted individual choices. Note that the condition in proposition 2 is not directly affected by the parameters γ or ν which determine the rate of re-infections or the arrival rate of a cure. These parameters matter only indirectly because they affect the value function.

Turning to the timing of lockdowns and inverse lockdowns, recall that infections increase the government's continuation value—infections induce capital gains—once y reaches the value $y^{\min}$. Let y^c denote the value of y (if it exists) at which the total externality equals zero, $\psi(\zeta/n - 1) + V_y(y^c, t) = 0$; and let ${V_y}^{\max}$ denote the maximum value of $V_y(y, t)$, both along the path implemented by the government. We have the following result:

Proposition 3. Under assumptions 1 and 3 and if $g(a) = a^n$ and $\zeta < n$, lockdowns occur as follows:

i. Starting from small y, the government immediately imposes a lockdown;

ii. if $\gamma = 0$ and ${V_y}^{\max} > \psi(1 - \zeta/n)$ then the government also imposes an inverse lockdown;

iii. if V is locally convex at $y = y^c$ then an inverse lockdown immediately follows the lockdown.

The first part of proposition 3 is consistent with the fact that during the ongoing COVID-19 pandemic many governments imposed lockdown measures early on. The last part, in contrast, which concerns the reversal from a lockdown to an inverse lockdown, appears more surprising. It might partly explain stimulus measures such as temporary sales tax reductions or employment subsidies; and it suggests quite different policy interventions at later stages of an epidemic.[30] In section 5 we analyze the consequences of constraints that make it impossible for the government to impose an inverse lockdown.

Note that the basic intuition underlying the reversal result is very general: Since an epidemic generates costs the value function during the transition is lower than after the transition; that is, at some point, society experiences capital gains. These capital gains arise due to the change of an aggregate state variable (or many such state variables) which an household takes as given; that is, the capital gains are not internalized by households. As long as the capital gains are sufficiently large to outweigh negative static externalities the reversal result thus follows.

Under our baseline calibration for the time autonomous case the optimal lockdown lasts about ninety days. During that period the average activity level lies at roughly seventy six percent; without government intervention it would have averaged more than ninety three percent.

5 Robustness

In this section, we check the robustness of the quantitative results that we have derived so far. Because the model is so easy to solve we consider a host of different scenarios. They are distinguished by the specification of the law of motion, the costs of infection, or the parameter values we impose. We also consider scenarios in which the regime stochastically changes or in which the government faces constraints on policy instruments. Subsequent sections are devoted to the analysis of more wide ranging extensions of the model, including model variants with re-infections, a deterministic arrival of a cure, and observable infection status.

5.1 Quadratic Effect of Activity on Infections

As discussed previously the epidemiological evidence on returns to scale in infections is mixed (see, e.g., Hethcote, 1989). In a first robustness check we therefore allow function g to exhibit increasing returns to scale.

When $n = 2$ such that $g(a) = a^2$ rather than a, reductions in activity suppress infections more strongly and this alters the trade-offs perceived both by the government and individual decision makers. Equations (7) and (8) no longer apply. Instead, the optimal and equilibrium activity levels are the (positive) solutions of quadratic equations, and the value functions U and V change accordingly.

[30] Note that the proposition allows for repeated reversals between lockdowns and inverse lockdowns (depending on the shape of V), a feature that we do not find in any of our simulations.

Scenario	y^c	$t^c(days)$	a^c	$\phi^u - \phi^v$
Baseline	0.0257	90	0.7570	0.0610
Quadratic effect of activity	0.0745	203	0.6423	0.0342
Stronger curvature of u	0.0239	77	0.8766	0.0235
Higher costs of infection	0.0276	105	0.6616	0.0926
Endogenous costs: Congestion	0.0624	108	0.7529	0.0869
Endogenous costs: Learning	0.0194	85	0.7559	0.0622
Higher arrival rate of a cure	0.0385	108	0.6881	0.0648
Regime change: Reduction in β	0.0361	108	0.6750	0.0370
Regime change: Multiple waves	0.0650	136	0.5994	0.0670
Constraints on policy instruments	0.0359	105	0.6956	0.0048

Table 2: Key statistics for different scenarios. y^c denotes the value of the state at which the lockdown ends and, absent constraints on policy instruments, the total externality equals zero. With such constraints the total externality equals zero slightly after y^c is reached. t^c denotes the duration of the lockdown and a^c the average activity level during the lockdown. $\phi^u - \phi^v$ measures the welfare gain of the government intervention expressed in terms of the equivalent life-time consumption variation.

Figure 9 in appendix C compares the optimal and equilibrium activity levels when $n = 2$ with the corresponding outcomes in the baseline model. The panels on the left-hand side represent the baseline model and the panels on the right-hand side the case with $n = 2$. The two panels in the top row represent the results in y-space and the two panels in the bottom row illustrate dynamics over time. In addition, table 2 summarizes key statistics: The value y^c at which the total externality equals zero; the time span t^c until y^c is reached; and the average activity level a^c under the lockdown. The table collects these statistics for all scenarios we consider in this section and compares them to the outcomes in the baseline model and the case with re-infections considered in section 6.

With $n = 2$ both $y^{\min}$ and y^c increase (to approximately 0.0301 and 0.0745 respectively). The government reduces activity substantially more strongly than when $n = 1$, in contrast to what households choose in equilibrium. Accordingly, the transition dynamics in the two cases look more similar than with $n = 1$ and the welfare gains from intervention decrease to roughly three and a half percent. Nevertheless, due to the significantly slowed down infection dynamics, the optimal lockdown extends over a much longer period, 203 days.

5.2 Stronger Curvature of u

Next we consider a modified specification of preferences: We assume that rather than $u(a) = \ln(a) - a + 1$ the net direct utility from activity is given by

$$u(a) = \frac{a^{-1}}{-1} - a + 2;$$

that is, the intertemporal elasticity of substitution for consumption equals one half rather than one as posited in the baseline model. Figure 10 in appendix C and table 2 summarize the results.

With the lower elasticity of substitution activity is reduced by less, both in equilibrium and under the optimal policy. The lockdown also is shorter and the welfare gains due to the optimal government intervention are smaller.

5.3 Higher Costs of Infection

Next we consider the consequences of higher costs of infection. Mulltiplying the cost parameter ψ by 1.5 implies, not surprisingly, that both the government and households reduce activity more strongly than in the baseline case. As a consequence, transition dynamics slow down, see figure 11 in appendix C and table 2. The optimal lockdown is stricter, lasts half a month longer than in the baseline, and yields higher benefits from government intervention.

5.4 Endogenous Costs of Infection: Congestion and Learning Effects

Next we relax the assumption that the unit costs of infection—the factor ψ entering the costs of infection (4)—are an exogenous constant. We consider two alternative specifications. In the first specification, we replace ψ in (4) by $\psi^f y(\bar{y} - y)$ where $\psi^f > 0$. That is, we assume that the costs of infection are quadratic in $y(\bar{y} - y)$ rather than linear as we have assumed so far. This specification captures congestion effects, for instance due to the fact that capacity constraints in the health care sector lead to a deterioration of care quality and increased fatality rates as infection flows rise. In COVID-19 related policy discussions congestion effects are an important motivation for calls to "flatten the curve."

In the second specification, we replace ψ in (4) by $\psi^l(2 - y/\bar{y})$ where $\psi^l > 0$. That is, we assume that the costs of infection per unit of infection flow are decreasing in the share of the population that has been infected in the past. More specifically, we assume that the unit costs halve during an epidemic. This specification captures learning-by-doing effects, for instance due to the fact that medical personnel learns by experience how to diagnose, triage, or treat patients while health authorities improve administration and logistics.

Both specifications require slight modifications in our calibration strategy. As we explain in appendix A we set $\psi^f = 6\psi/\bar{y}^2$ and $\psi^l = \psi 2/3$.

We find that congestion effects lengthen the optimal lockdown but have no discernible effect on the optimal activity level; the welfare gains due to the optimal lockdown policy

rise. Learning shortens the lockdown duration slightly but again has little effect on the optimal activity level, see figures 12 and 13 in appendix C as well as table 2.

5.5 Higher Arrival Rate of a Cure

Next we consider the consequences of a higher arrival rate of a cure, reflected in a multiplication of the parameter ν by 1.5. The government responds aggressively to the prospect of a more likely exit from the epidemic, exploiting the possibility of intertemporal substitution by more strongly reducing activity. In contrast, households do not respond because their activity choice is unaffected by the arrival rate. Accordingly, the optimal lockdown lasts longer and is stricter but the welfare gains of that policy are of a similar magnitude as in the baseline case, see figure 14 in appendix C and table 2.

5.6 Stochastic Regime Change: Reduction in β and Multiple Waves

Next, we ask how expected changes in the epidemiological environment affect optimal policy. Scenarios where such a question might arise include settings where policy makers anticipate virus mutations; new medical treatments; improved test, trace, and quarantine strategies; improved implementability of lockdown restrictions; or, in contrast, increased political resistance against such restrictions.

We consider two specific scenarios. In the first one we allow for a permanent reduction in β by fifty percent (reflecting, e.g., improved test, trace, and quarantine strategies) that materializes stochastically, with arrival rate μ. We assume that $\mu = 1/90$, reflecting an expected duration of three months until the regime changes.[31]

The left panel of figure 4 depicts the government's value function when β assumes the baseline value (solid line); when β permanently assumes the lower value (dashed line); or when β assumes the baseline value but may randomly fall (dotted line). The right panel depicts the corresponding activity levels chosen by the government in the three cases.

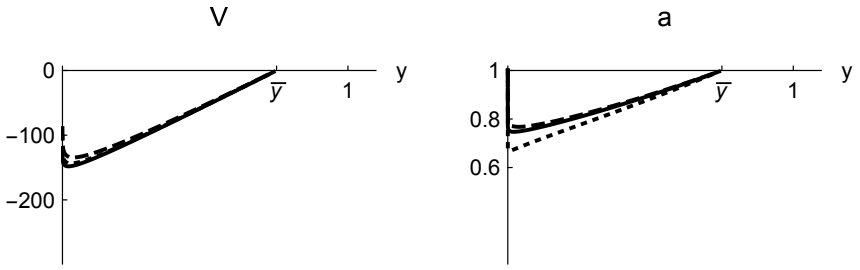

Figure 4: Value function and activity level in the government's program when β assumes the baseline value (solid), permanently assumes the lower value (dashed), or assumes the baseline value and may randomly fall (dotted).

[31]For example, Fetzer and Graeber (2020) report that the U.K. implemented a test-and-trace regime by late May; exploiting a data processing error they conclude that the new regime significantly lowered infection rates.

Not surprisingly, the value function when regime change is possible is (slightly) higher than in the baseline case; this reflects the upside risk. More importantly, the prospect of this upside risk leads the government to wait longer before relaxing the lockdown, and to substantially reduce activity levels until β falls. Relative to the baseline case, y^c and the duration of the lockdown as well as its strictness increase. Similarly to the case with a higher ν the government intertemporally substitutes, anticipating a better trade-off between lives and livelihoods in the future. Nevertheless, the welfare gains from optimal government intervention relative to no intervention fall compared with the baseline case, see figure 15 in appendix C and table 2.

In the second scenario we allow for multiple waves of infection. More specifically, we assume that with arrival rate μ the number of persons in the post group reverts back to y_0.[32] As a consequence the HJB equation features an additional capital gains term, $\mu V(y_0)$. To capture the fact that seasonal factors might affect the start of a new wave (as was the case during the Spanish influenza) we let $\mu = 1/365$, that is, a new wave arrives on average after a year. We assume that there is no upper bound to the maximum number of possible waves.

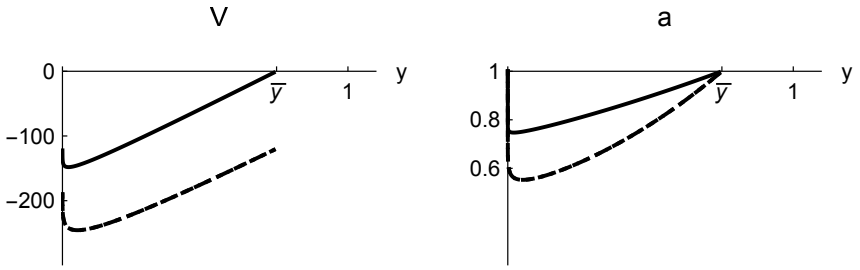

Figure 5: Value function and activity level in the government's program without (solid) and with (dashed) the possibility of multiple waves.

We find that the prospect of multiple waves induces the government to behave even more cautiously. The optimal lockdown duration increases to more than four and a half months, with an average reduction of the activity level to less than sixty percent. The welfare gains due to optimal government intervention increase to nearly seven percent, see figure 16 in appendix C and table 2.

5.7 Constraints on Policy Instruments

Finally, we analyze how important "inverse lockdown" measures are for the welfare gains due to optimal government intervention. More specifically, we assume that government can curtail economic activity (impose a lockdown) but lacks the powers to correct households' unwillingness to reengage after a lockdown (impose an inverse lockdown). The relationship between the value functions of the social planner, the government, and households in equilibrium differs in this case from the situation considered so far: In the range where the social planner imposes an inverse lockdown (high y values) the government's

[32] This could reflect a new strain of virus that is associated with a massive loss of immunity.

value function coincides with the *equilibrium* value function; and in the range where the social planner imposes a lockdown (low y values) the government's value function lies between the value function of the social planner and the equilibrium value function.

To solve for the government's value function subject to the constraints on its instruments we need to impose a modified boundary condition. In the baseline scenario this boundary condition stipulates that the government's and the household's equilibrium value functions coincide at $y = \bar{y}$ because the epidemic has ceased to exist at that point. Now, we instead impose the condition that the value functions coincide starting from some endogenous value of the state, $\hat{y}$ say, at which the government ends the lockdown (and would like to switch to an inverse lockdown but cannot do that). Formally, we first solve for U and then find V and $\hat{y}$ by means of the value matching and smooth pasting conditions

$$V(\hat{y}) = U(\hat{y}) \quad \text{and} \quad V'(\hat{y}) = U'(\hat{y}).$$

We find that the constraints imply a longer and stricter lockdown than in the baseline scenario. The activity trend changes abruptly at $y \approx 0.0359$ when the lockdown ends. Thereafter, the dynamics under the optimal government policy resemble those in equilibrium, with a time lag. By definition, the welfare gains due to government intervention are smaller than when the government is unconstrained. In fact, the welfare gains are very small—the difference between ϕ^u and ϕ^v falls to less than one half percent, see figure 17 in appendix C and table 2.

5.8 Summary

The robustness checks suggest that the optimal lockdown lasts for three to four months, with activity reductions by twenty five to forty percent and welfare gains from the optimal government intervention on the order of three to nine percent of lifetime consumption. There are three outlier cases, however. First, when activity enters the law of motion quadratically the optimal lockdown extends over a much longer period, nearly seven months. Second, when the intertemporal elasticity of substitution equals one half the reduction of activity during the lockdown only amounts to roughly twelve percent. Third, when the government can impose a lockdown but lacks instruments to subsequently stimulate activity then the welfare gains are nonsignificant.

6 Lack of Immunity

In the baseline specification households who undergo infection and recovery are protected from re-infection such that herd immunity is eventually attained (i.e., y reaches $\bar{y}$). But in some epidemics re-infection risk is nonnegligible. For example, at the time of this writing, research on COVID-19 suggests that immunity wanes and individuals with antibodies may re-contract the disease (see the discussion in the introduction).

To study the implications of lack of immunity we consider the case with $\gamma > 0$, that is, we assume that persons who have undergone infection in the past may join the pre-group

again and subsequently re-contract the disease. Accordingly, the steady-state value of y is given by $y_\infty(a) \equiv \bar{y} - \gamma/(g(a)\beta)$; due to $\gamma > 0$ it lies strictly below $\bar{y}$.

In decentralized equilibrium, the steady state (a_∞, y_∞) is characterized by the law of motion and the first-order condition. Letting for simplicity $g(a) = a$, equations (1) and (8) yield

$$\begin{aligned} \gamma &= a_\infty \beta(\bar{y} - y_\infty), \\ a_\infty &= \frac{1}{1 + \beta y_\infty(\bar{y} - y_\infty)\zeta\psi}. \end{aligned}$$

This system has a unique positive solution,

$$\begin{aligned} a_\infty &= \frac{1 - \gamma\psi\bar{y}\zeta}{2} + \frac{\varphi}{2\beta}, \\ y_\infty &= \frac{1 + \gamma\psi\bar{y}\zeta}{2\gamma\psi\zeta} + \frac{\varphi}{2\beta\gamma\psi\zeta} \end{aligned}$$

with $\varphi \equiv \sqrt{\beta}\sqrt{4\gamma^2\psi\zeta + \beta(1 - \gamma\psi\bar{y}\zeta)^2}$. Differentiating this solution implies that in a neighborhood of $\gamma = 0$ (and $a_\infty = 1$), a_∞ is U-shaped; y_∞ decreases; and steady-state infections $\dot{y}^g = \gamma y_\infty$ are inverse-U-shaped in γ. For the reasons discussed in subsection 2.1.3 only the increasing segment of this inverse-U-shaped relationship is relevant.

The steady state in the government's program solves the conditions

$$\begin{aligned} \gamma &= a_\infty \beta(\bar{y} - y_\infty), \\ a_\infty &= \frac{1}{1 + \beta y_\infty(\bar{y} - y_\infty)(\psi - V'(y_\infty))}, \\ (\rho + \nu)V'(y_\infty) &= -a_\infty \beta(\bar{y} - 2y_\infty)(\psi - V'(y_\infty)) - \gamma V'(y_\infty), \end{aligned}$$

where the first equation follows from equation (1); the second from the first-order condition (7); and the last from the envelope condition (using the steady-state property $\dot{y}(y) = 0$). Eliminating y_∞ and $V'(y_\infty)$ yields the equation

$$(\rho + \nu + \gamma)\gamma\psi = a_\infty(1 - a_\infty)\beta\left(\frac{\rho + \nu}{a_\infty\beta\bar{y} - \gamma} + 1\right), \tag{10}$$

which characterizes the government's choice of a_∞. The ratio in parentheses is positive (because $y_\infty(a_\infty) > 0$) and $a_\infty \in (0, 1)$ such that both the left- and the right-hand side of equation (10) are strictly increasing in γ. Moreover, the right-hand side is non-monotone in a_∞. Equation (10) thus generally admits multiple solutions. Of course, only one of them constitutes the outcome implemented by the government. In fact, for small γ (such that the model is well specified) the solution of equation (10) necessarily yields a government choice for a_∞ which is decreasing in γ.

The following proposition compares the equilibrium steady state and the steady state in the government's program. Moreover, it indicates that an increase in γ reduces activity in the government's program around $y^{\min}$ where lockdowns are in place.

Proposition 4. Under assumptions 1–3 and if $\gamma > 0$, $T = \infty$, $u(a) = \ln(a) - a + 1$, and $g(a) = a$, there exists a unique steady state in decentralized equilibrium. The government's choices of a_∞ and y_∞ exceed the corresponding equilibrium values if $(\rho + \nu)(1 - \zeta) < \zeta\beta\bar{y}$. In a neighborhood of $y = y^{\min}$ the government's choice of activity satisfies $da(y)/d\gamma < 0$ when $2\beta\bar{y} > (4 + \beta\bar{y}^2\psi)\gamma$.

This result implies that absent a static externality ($\zeta = 1$) the government's choice of a_∞ necessarily exceeds the equilibrium level of a_∞, and more concern about the future (low ρ and/or ν) as well as a high infection rate (β) and a high $\bar{y}$ render the same result more likely. When the static negative externality is maximal ($\zeta = 0$), in contrast, then the steady-state equilibrium activity level is suboptimally high. Intuitively, this reflects the contrasting effects of static and dynamic externalities. On the one hand, the equilibrium level of a_∞ tends to exceed the optimal level when households do not fully internalize the costs of infection ($\zeta < 1$). On the other hand, the parameters β, $\bar{y}$, ρ, and ν increase the capital gains in the government's problem for given (a_∞, y_∞) and this reduces the net costs of infection perceived by the government.

Next, we turn to the dynamic analysis. Since with $\gamma > 0$, $\bar{y}$ is no longer a rest point the boundary conditions that we had imposed so far ($\lim_{y\to\bar{y}} U(y) = U^\star$ and $\lim_{y\to\bar{y}} V(y) = U^\star$) no longer apply. Instead, we impose the boundary condition that the value function evaluated in steady state (characterized above) equals the capitalized steady-state utility flows net of the steady-state flow costs of infection. That is, for example, we solve the government's HJB equation subject to the boundary condition

$$(\rho + \nu)V(y_\infty) = u(a_\infty) - \psi a_\infty \beta y_\infty(\bar{y} - y_\infty),$$

where a_∞ is optimal.

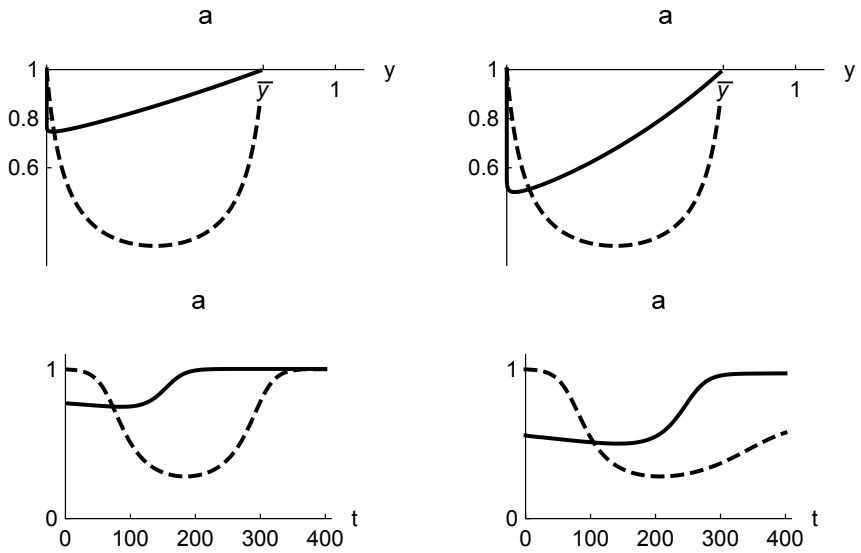

Figure 6: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with lack of immunity (right panels).

In the quantitative analysis, we let $\gamma = 1/365$ (i.e., immunity is lost on average after a year).[33] We find that the steady-state activity level drops to $a_\infty = 0.9696$ and $y_\infty = 0.7208$, while in equilibrium, $a_\infty = 0.6372$ and $y_\infty = 0.7055$. Moreover, the strictly positive value of γ substantially shifts the value functions U and V downward. While in the case with $\gamma = 0$, U and V converged as $y \to \bar{y}$ this is no longer the case when there is lack of immunity.

Proposition 4 implies that γ reduces the government's choice of activity when a lockdown is in place. We find that, indeed, the optimal lockdown lasts much longer than in the baseline case—approximately half a year—and is very strict as the activity level is reduced to just above fifty percent on average. The welfare gains from the optimal government intervention exceed seven percent, see figure 6 and table 2.

7 Deterministic Arrival of a Cure

In the baseline specification a cure arrives stochastically with arrival rate ν. But in some settings it appears more plausible to view the arrival of a vaccine or other form of cure as a deterministic rather than stochastic process. For example, several promising candidates for an effective vaccine, of which at least one would work with near certainty, might undergo final trials or require final regulatory approval (as was the case in Fall 2020 in the context of COVID-19). Or a poor country which is dependent on international financial and logistic assistance might have to wait for a pre-specified period until doses are supplied by donor institutions and distributed domestically.

To represent such a setting with a deterministic arrival of a cure, we let $T < \infty$ and $\nu = \gamma = 0$. Time therefore becomes a second state variable and the government's problem may be formulated more conveniently using the Hamiltonian

$$\mathcal{H}(a(t), y(t), t) = u(a(t)) - \psi g(a(t))\beta y(t)(\bar{y} - y(t)) + \mu(t) g(a(t))\beta y(t)(\bar{y} - y(t)),$$

where $\mu(t)$ denotes the (present value) multiplier attached to the law of motion. An optimal plan satisfies the boundary condition $\mu(T) = 0$ as well as the conditions $\mathcal{H}_a(a(t), y(t), t) = 0$ and $\mathcal{H}_y(a(t), y(t), t) = -\dot{\mu}(t) + \rho\mu(t)$.[34] This implies

$$\begin{aligned} u'(a(t)) &= g'(a(t))\beta y(t)(\bar{y} - y(t))\left(\psi - \mu(t)\right), \\ \dot{\mu}(t) &= g(a(t))\beta(\bar{y} - 2y(t))\left(\psi - \mu(t)\right) + \rho\mu(t) \end{aligned}$$

as well as the law of motion, $\dot{y}(t) = g(a(t))\beta y(t)(\bar{y} - y(t))$.

As we will show next the government's optimal policy may not be continuous in T. We establish this result for the special case with $\rho = 0$ which allows for a sharper characterization:

[33] In the absence of reliable estimates about the duration of immunity against COVID-19 we set this number to make the results comparable with the multiple waves scenario analyzed in section 5.

[34] The multiplier $\mu(t)$ corresponds to the partial derivative $V_y(y(t), t)$ such that $\dot{\mu}(t) = V_{yt}(y(t), t) + V_{yy}(y(t), t) g(a(t))\beta(\bar{y} - 2y(t))$.

Proposition 5. Under assumption 1 and $\gamma = \rho = \nu = 0$ and $T < \infty$, the government's optimal choice satisfies $a(t) = a$. If, moreover, $u(a) = \ln(a) - a + 1$, $g(a) = a$, and y_0 is sufficiently small (as defined in the proof) then the optimal a as a function of T is dis-continuous.

The first part of proposition 5 implies that the government's program reduces to the choice of a *constant* activity level. Starting from a low value of T, this optimal choice responds to an increase in T by reducing activity and allowing for a higher stock of post households, $y(T)$, when the cure arrives.

The second part of the proposition states that as T increases further, the costs of curtailing activity eventually become so large that it is no longer optimal to impose a severe lockdown in order to keep $y(T)$ in check. Instead it becomes optimal to choose a much higher activity level and to "give up," i.e., to accept higher infection numbers.[35] To illustrate this important result figure 7 displays the contour lines of the government's objective function under our baseline calibration. The figure indicates that, as T increases (along the vertical axis), the optimal activity level first falls before eventually (when $T \approx 200$) dis-continuously jumping to a value close to unity.

The result of a nonmonotone (in T) optimal activity reduction can potentially rationalize large differences in public health policies across time and space in an epidemic. Applied to our motivating examples, it implies that when vaccine trials in rich countries or the time to distribution of vaccines in poor countries are expected to last very long, strict lockdowns are unlikely to be optimal.

8 Observable Infection Status

In the baseline specification infection status is unobserved and as a consequence all households in the pre, post, and neutral groups choose the same activity level. As we discussed in section 2 this is a plausible approximation in the context of some epidemics, including the COVID-19 pandemic, but less so for others. We now relax the assumption and consider the consequences of observable infection status.

Specifically, we assume that infected households who develop symptoms can trace these symptoms to the infection and thus become aware of their infection status. Upon recovery they gain immunity ($\gamma = 0$).[36] Since σ denotes the symptomatic share in the post pool a share σy of the population knows that it is immune. Households in this group choose the activity level $a^\star$ since they do not face any risk of future infection.

When the infection status or activity choice of a household is public information households in the pre group will avoid contact with known members of the post group in order to minimize privately costly infection risk. Accordingly, switching back to recursive notation, the law of motion changes to

$$\dot{y}(y, a) = g(a)\beta(1 - \sigma)y(\bar{y} - y), \tag{11}$$

[35] The government's program does not satisfy Arrow's second order conditions because the maximized Hamiltonian fails to be concave in y for all $t \in [0, t]$. Accordingly, the first-order conditions are not sufficient for a global maximum.

[36] We maintain the assumption that the death rate δ is negligible.

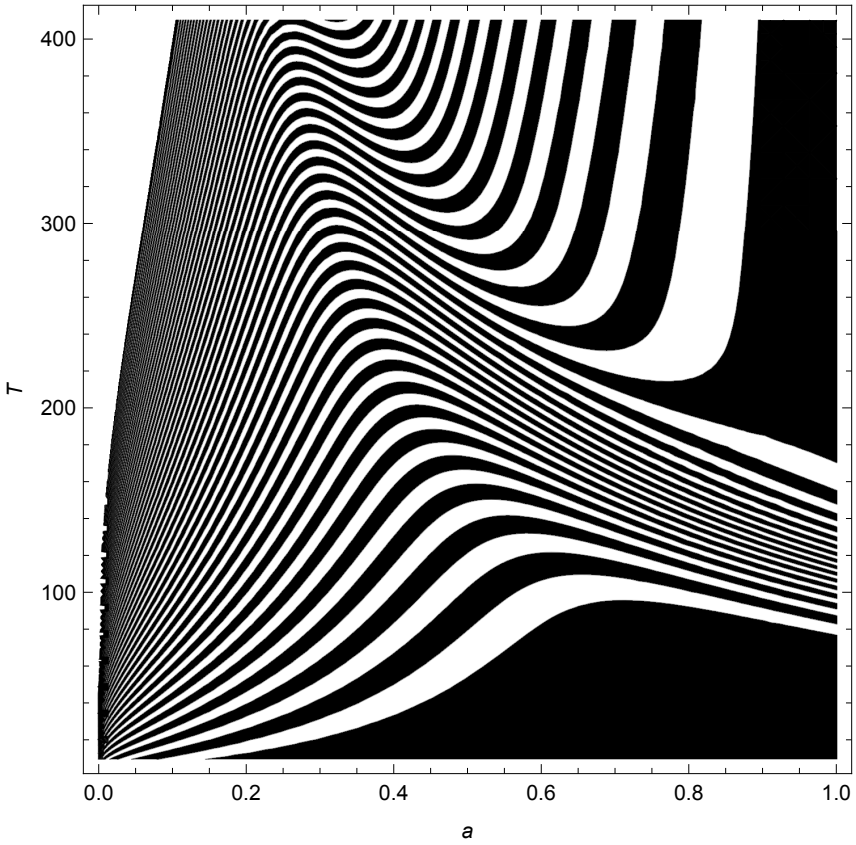

Figure 7: Contours of the government's objective function when T is finite.

as households in the pre group (with mass $\bar{y}-y$) only interact with asymptomatic members of the post group (with mass $(1-\sigma)y$).

Observability implies that the value functions of households differ depending on whether they have or have not developed symptoms. The value function of a household that experienced symptoms in the past is given by $U^\star$ because that household chooses $a^\star$ and is not exposed to any further infection risk. The value function (in the time autonomous case and re-introducing stochastic arrival of a cure) of a household that did not experience symptoms so far, say $\tilde{U}(y)$, resembles the value function in the baseline model, $U(y)$, with two differences. First, infection risk in that group reflects the fact that aggregate infection flows are fully concentrated in the group. That is, households without symptoms face the infection flow $\dot{y}(y,a)/(1-\sigma y)$. Second, infections do not only generate costs, ψ, but also release the affected households with probability σ into the symptomatic pool with continuation value $U^\star$.

Accordingly, the HJB equation reads

$$\begin{aligned}(\rho+\nu)\tilde{U}(y) &= \max_{a_i} u(a_i) + \frac{a_i}{a(y)}\frac{\dot{y}(y,a(y))}{1-\sigma y}\left(-\zeta\psi+\sigma\left(U^\star-\tilde{U}(y)\right)\right)\\ &\quad -\frac{\dot{y}(y,a(y))}{1-\sigma y}(1-\zeta)\psi+\dot{y}(y,a(y))\tilde{U}'(y)+\nu U^\star\end{aligned}$$

subject to the law of motion (11). The first-order condition is given by

$$(1-\sigma y)u'(a_i) = \frac{\zeta\psi - \sigma(U^\star - \tilde{U}(y))}{a(y)} g(a(y))\beta(1-\sigma)y(\bar{y}-y).$$

The government's program, which is represented by the HJB equation

$$(\rho+\nu)V(y) = \max_a \sigma y u(a^\star) + (1-\sigma y)u(a) - \psi\dot{y}(y,a) + \dot{y}(y,a)V'(y) + \nu U^\star,$$

yields the first-order condition

$$(1-\sigma y)u'(a(y)) = (\psi - V'(y))g'(a(y))\beta(1-\sigma)y(\bar{y}-y).$$

Unlike the household's program, which is now type dependent, the government's program is nearly unchanged and the first-order condition differs from the baseline case only insofar as β is replaced by $\beta(1-\sigma)/(1-\sigma y) < \beta$.

When we impose the standard functional form assumptions $g(a) = a$ and $u(a) = 1 + \ln(a) - a$, the first-order conditions imply that the activity choice of asymptomatic households in equilibrium, say $\tilde{a}(y)$, and by the government, respectively, satisfy[37]

$$\begin{aligned}
\tilde{a}(y) &= \frac{1}{1 + \Big(\zeta\psi - \sigma(U^\star - \tilde{U}(y))\Big)\beta(1-\sigma)y(\bar{y}-y)/(1-\sigma y)}, \\
a(y) &= \frac{1}{1 + (\psi - V'(y))\beta(1-\sigma)y(\bar{y}-y)/(1-\sigma y)}.
\end{aligned}$$

Note that the three differences between $\tilde{a}(y)$ and $a(y)$ reflect the static externality (the ζ term in the household's first-order condition), the dynamic externality (only the government's first-order condition contains the derivative of the value function), and the probability of a type change (the $U^\star - \tilde{U}(y)$ term in the household's first-order condition). In equilibrium, the latter effect introduces an incentive for asymptomatic households to more strongly engage in activity. Substituting the activity choices back into the respective HJB equations yields differential equations that can numerically be solved.

Under our baseline calibration and assuming that $\sigma = 0.1$ we find similar results as in the baseline scenario.[38] The critical value y^c equals roughly 0.0315, the lockdown duration t^c amounts to 104 days, the activity level of asymptomatic households during the lockdown averages 0.7604, and the welfare gains from optimal government intervention are roughly five percent of lifetime consumption.

Next, we turn to the case with private information about one's infection status and activity choice.[39] The law of motion now reads

$$\dot{y}(y) = g(\bar{a}(y))\beta y(\bar{y}-y), \tag{12}$$

[37] Average activity in equilibrium is given by $\sigma y a^\star + (1-\sigma y)\tilde{a}(y)$; under the optimal policy, it is given by the same expression with $\tilde{a}(y)$ replaced by $a(y)$.

[38] For evidence on the share of symptomatic post households see footnote 48.

[39] In the context of COVID-19 private information may seem more relevant than public information as some symptoms, such as loss of smell and taste, are unobserved by third parties.

where $\bar{a}(y) \equiv (\sigma y a^\star + (\bar{y} - \sigma y)\tilde{a}(y))/\bar{y}$ denotes the average activity of pre and post households. There are two differences between the laws of motion (11) and (12). First, the $1-\sigma$ term in the former equation is replaced by unity, as in the baseline law of motion (1), reflecting the fact that all post households interact with pre households. Second, related, the groups whose activity affects infection dynamics are the pre households, the post households unaware of their infection status, and the aware post households. While the former two groups choose the activity level $\tilde{a}(y)$ the latter chooses $a^\star$.

We assume that symptomatic households drop out of the insurance scheme such that their value function is given by $U^\star$, as in the case with public information.[40] The value function of an asymptomatic household, $\tilde{U}(y)$, corresponds to the function in the public information case except for the modified law of motion. The activity level $a(y)$ entering the HJB equation continues to represent the activity level chosen by all insured households. The first-order condition,

$$(1 - \sigma y)u'(a_i) = \frac{\zeta\psi - \sigma(U^\star - \tilde{U}(y))}{a(y)} g(\bar{a}(y))\beta y(\bar{y} - y),$$

implies the following equilibrium activity choice of asymptomatic households once we let $g(a) = a$ and $u(a) = 1 + \ln(a) - a$:

$$\tilde{a}(y) = \frac{1 - \sigma(y/\bar{y})a^\star \left(\zeta\psi - \sigma(U^\star - \tilde{U}(y))\right) \beta y(\bar{y} - y)/(1 - \sigma y)}{1 + (1 - \sigma(y/\bar{y})) \left(\zeta\psi - \sigma(U^\star - \tilde{U}(y))\right) \beta y(\bar{y} - y)/(1 - \sigma y)}.$$

As expected, $d\tilde{a}(y)/d\sigma < 0$ (given $\tilde{U}$), because higher average activity and thus, infection rates, lead the households unaware of their infection status to shield more. The government's program is the same as in the baseline model (where $\sigma = 0$) because the government chooses the same activity level for everybody.[41]

Imposing the same parameter values as in the public information case we find similar results. The optimal lockdown lasts 103 days, the activity level during the lockdown averages 0.7567, and the welfare gains from the optimal government intervention are approximately six percent. The transition dynamics in the economies with public or private information about the infection status broadly resemble those in the baseline model as far as asymptomatic households are concerned, see figures 18 and 19 in appendix C.

[40] This is consistent with the private information assumption when insurers must not disclose their clients, perhaps for contractual or legal reasons.

[41] The government cannot gain from implementing two different activity levels. To see this, suppose that incentive constraints would not bind and the asymptomatic and symptomatic households had activity levels a_1 and $a_2 > a_1$, respectively, such that $\bar{a}(y) = (\sigma y a_2 + (\bar{y} - \sigma y)a_1)/\bar{y}$. The direct utility gain relative to the situation with a common activity level a_1 would equal $\sigma y(u(a_2) - u(a_1)) \approx u'(a_1)(a_2 - a_1)\sigma y$, while the cost would be equal to $(\psi - V'(y))(g(\bar{a}) - g(a_1))\beta y(\bar{y} - y) \approx (\psi - V'(y))g'(a_1)(a_2 - a_1)\beta y(\bar{y} - y)\sigma y/\bar{y}$. The first-order condition for a common activity level implies $u'(a_1) = (\psi - V'(y))g'(a_1)\beta y(\bar{y} - y)$. Since $\bar{y} < 1$ it follows that the losses from a discriminatory policy would exceed the gains.

9 Conclusion

We have developed a flexible model of infectious dynamics with a single endogenous state variable and economic choices by households and a government. Government intervention yields welfare gains relative to laissez faire because households do not internalize static and dynamic externalities although they do adjust their behavior for fear of infection.

We find several novel theoretical results. First, a lockdown is generally followed by its opposite—policies to stimulate activity beyond the privately optimal level. Economic activity generates negative static externalities and initially, these induce the government to impose restrictions. But eventually, activity and the associated infections also generate positive dynamic externalities—capital gains for society—and once the latter are sufficiently strong they outweigh the static externalities and call for higher activity than what households would privately choose. This mechanism extends to many richer models independently of the number of state variables.

Second, re-infection risk implies more cautious activity choices in steady state than optimal if static externalities are small. Intuitively, even if infections are endemic there are capital gains that the government internalizes but the private sector does not. Moreover, when a lockdown is in place the government's choice of activity level is decreasing in re-infection risk.

Third, when a cure arrives deterministically, optimal policy is dis-continous. The lockdown is light when the arrival date exceeds a specific value, but strict otherwise. Intuitively, a later arrival date requires increasingly large reductions in activity, for a longer time in order to keep infections in check. At some point the government is no longer willing to bear these increasingly high costs of "lost livelihoods" and optimally accepts higher infection levels earlier on, turning a blind eye on the epidemic.

Calibrated to the ongoing COVID-19 pandemic our baseline specification suggests that starting from mid March 2020, a lockdown should optimally have lasted for three months with economic activity reduced by twenty five percent. Exploiting the flexibility of our framework we run a battery of robustness checks, with a focus on different parameter values or infection cost functions, potential regime change, or restrictions on government instruments, and we analyze various extensions. Across almost all of these alternative specifications the basic message remains one of optimal lockdowns for three to four months with activity reductions by twenty five to forty percent and significant welfare gains from optimal government intervention.

Only when activity enters the law of motion quadratically or in the presence of re-infection risk does the optimal lockdown extend over a much longer period, between six and seven months, and only when the intertemporal elasticity of substitution equals one half rather than unity is the reduction of activity during the lockdown much smaller, roughly twelve percent. With re-infection risk, in contrast, the optimal lockdown is much stricter, with activity reduced by almost fifty percent. Finally, only when the government can impose a lockdown but lacks instruments to subsequently stimulate activity are the welfare gains nonsignificant.

Our framework can serve as a workhorse model and allows for many extensions. One interesting avenue for further research would be to introduce additional dimensions of

heterogeneity relative to those present in section 8. Another, related one would be to introduce conflicts of interest in order to analyze political economy considerations for governments that fight an epidemic.

A Calibration Strategy

A.1 Law of Motion (1)

In this appendix, we describe how we use information about parameter values in the canonical SIR model to deduce parameter values for the law of motion (1).[42]

A.1.1 Canonical SIR Model

The canonical SIR model due to Kermack and McKendrick (1927) specifies laws of motion for the population shares of three groups: the "susceptible," the "infected" or "infectives," and the "removed." Their respective population shares at time $t \geq 0$ are denoted by $x(t)$, $y(t)$, and $z(t)$, respectively, where $x(t) + y(t) + z(t) = 1$.[43] We normalize the mass of the total population at time $t = 0$ to unity.

At time $t = 0$ the population consists of $x(0)$ susceptible persons and a few infected persons, $y(0)$. There are no removed persons at this time, $z(0) = 0$. In each instant after time $t = 0$, infected persons transmit the disease to members of the susceptible group and a share of the infected either dies or recovers and develops immunity. Formally,

$$\dot{x}(t) = -b(t)x(t)y(t), \tag{13}$$
$$\dot{y}(t) = -\dot{x}(t) - (c^d + c^r)y(t), \tag{14}$$
$$\dot{z}(t) = (c^d + c^r)y(t). \tag{15}$$

Here, $b(t)$ denotes a possibly time-varying infection rate. The extent to which susceptible persons are infected depends on their number, $x(t)$; the infection rate, $b(t)$; and the population share of infected persons. The number of infected persons increases one-to-one with the susceptible persons that get infected, while a share $c \equiv c^d + c^r$ of the infected population dies or recovers; the coefficients c^d and c^r parameterize the flow into death and recovery, respectively.

Consider the case where $b(t)$ is constant at value b. Inspection of equations (13) and (14) reveals that for $bx(0) > c$ the share of infected persons increases until it reaches a maximum when $x(t) = c/b$; thereafter, the share declines. Intuitively, when $x(0)$ falls short of c/b (the "herd immunity level") then there are fewer new infections of susceptible persons than outflows from the infected pool due to recoveries and death. As is well known (e.g., Theorem 2.1 in Hethcote, 2000), $x(\infty)$ falls short of the herd immunity level unless $x(0) = c/b = x(\infty)$ and $y(0) = 0$.[44]

In the SIR-S model a share γ of the removed population loses immunity and moves

[42] See Gonzalez-Eiras and Niepelt (2020b) for a broader and more detailed discussion, also with respect to other SIR models.

[43] We follow the notation introduced by Kermack and McKendrick (1927).

[44] Note also, from equation (14), that at the beginning of an epidemic with $x(t) \approx 1$ and $z(t) \approx 0$, b approximately equals the growth rate of the number of persons who are or were infected, $\frac{\dot{y}(t)+\dot{z}(t)}{y(t)+z(t)} = b\frac{x(t)y(t)}{y(t)+z(t)} \approx b$.

back to the susceptible pool. Accordingly, the dynamic system is given by

$$\begin{aligned}\dot{x}(t) &= -b(t)x(t)y(t) + \gamma z(t),\\ \dot{y}(t) &= b(t)x(t)y(t) - cy(t),\\ \dot{z}(t) &= cy(t) - \gamma z(t).\end{aligned}$$

In steady state this reduces to

$$\gamma z = bxy = cy.$$

Calibration We measure time in days and use information about the spread of COVID-19 in the United States to calibrate the model. We associate time $t = 0$ with mid March 2020, the date around which public health authorities considered to impose restrictions. We assume that at this time, $z(0)$ equalled practically nil.

Following Atkeson (2020) and the sources cited therein we assume that the flow rate from the infected to the removed population equals $c = 1/18$, corresponding to an exponentially distributed infection duration that averages 18 days.[45]

From Russell et al. (2020), Greenstone and Nigam (2020), and the sources cited therein we infer that the inverse of the infection fatality rate, c/c^d, lies in the range $[100, 200]$.

To calibrate $y(0)$ we use data on COVID-19 deaths in mid March 2020 as well as information about c^d and c. The number of deaths on March 16 equalled 23.[46] Based on equation (15) we infer that the initial share of the infected population in mid March, $y(0)$, equalled $1.8933 \cdot 10^{-4}$.[47] This compares to a reported case count of 4507, corresponding to a population share of $1.3745 \cdot 10^{-5}$.[48]

Finally, to calibrate b we rely on information in Ferguson et al. (2020) who argue that the "basic reproduction number" $\mathcal{R}_0 = b/c$ for COVID-19 equals approximately 2.4 which implies $b = 0.1333$. When we simulate the canonical SIR model subject to these parameter values we find that infections peak after roughly 114.34 days, on 7 July 2020. We use this date below.

A.1.2 Logistic Model

The law of motion (1) (subject to $\gamma = 0$) follows from the canonical SIR model (and other related models, see the discussion in Gonzalez-Eiras and Niepelt, 2020b) by letting

[45] Note that $\int_0^\infty c\mathrm{e}^{-ct}t\ dt = 1/c$.

[46] See `https://github.com/nytimes/covid-19-data/blob/master/us.csv`. Regressing the full set of March data on an exponential trend generates a similar point estimate for March 16.

[47] We have $y(0) \cdot (\text{US population}) = (\text{new deaths})/c^d = (\text{new deaths})/c \cdot c/c^d$. We use US population = 328 million, new deaths = 23, and $c/c^d = 150$.

[48] See `https://github.com/nytimes/covid-19-data/blob/master/us.csv`. The reported number corresponds to the cumulative case count but there are very few removed cases at the time. Common estimates of the extent of underreporting suggest a factor of ten, in line with our results; see, e.g., `https://www.medrxiv.org/content/10.1101/2020.03.14.20036178v2.full.pdf+html` or `https://fondazionecerm.it/wp-content/uploads/2020/03/Using-a-delay-adjusted-case-fatality-ratio-to-estimate-under-reporting-_-CMMID-Repository.pdf` or `https://www.npr.org/sections/coronavirus-live-updates/2020/06/25/883520249/cdc-at-least-20-million-americans-have-had-coronavirus-heres-who-s-at-highest-ri`.

$c^d = c^r = 0$ such that $z(t) = 0$, and by letting $b(t) = g(a(t))\beta$. Variable $y(t)$ now has the interpretation of the stock of persons who underwent infection in the past, not of the number of currently infected persons. For $g(a(t)) = 1$ the law of motion (1) (subject to $\gamma = 0$) implies a logistic path for $y(t)$ that converges to $\bar{y}$,

$$y(t) = \frac{\bar{y}}{1 + \mathrm{e}^{-\beta\bar{y}t}(\bar{y}/y(0) - 1)}.$$

Here, $\bar{y}$ has the interpretation of $1 - x(\infty)$ where $x(\infty)$ denotes the long-run share of susceptible households in the canonical SIR model who do not contract the disease.

We have the following standard result:

Proposition 6. Consider the law of motion (1) with $\gamma = 0$, $g(a(t)) = 1$, and $y(0) < \bar{y} \leq 1$. Then, $\dot{y}(t)$ reaches a maximum at

$$t = \ln\left(\frac{\bar{y} - y(0)}{y(0)}\right)/(\beta\bar{y}).$$

Proof. Solving $\ddot{y}(t) = 0$ (or $y(t) = \bar{y} - y(t)$) for t yields the result. □

Calibration Following Hall et al. (2020) who in turn rely on Ferguson et al. (2020) we assume that 75 percent of the population would contract the disease eventually in the absence of any mitigation measures, $\bar{y} = 0.75$. Moreover, we use proposition 6 to infer the value of β that corresponds to the b value in the canonical SIR model such that both models predict peak infections at the same date. That is, we choose β such that the t value in proposition 6 corresponds to 7 July 2020. Using $y(0) = 1.8933 \cdot 10^{-4}$ and $\bar{y} = 0.75$, this yields $\beta = 0.9660 \cdot 10^{-1}$.

Figure 8 illustrates the close connection between the predictions of the canonical SIR model (in blue) and the logistic model (in black). The figure is plotted under the assumption that the parameter values are given by those described above. The long-run share of households that got infected at some point (represented by the limiting values of the dashed curves) differs between the canonical SIR model and the logistic model because we assume that $\bar{y} = 0.75$, in line with the literature described above, while the parameter value for β stipulated by Ferguson et al. (2020) and fed into the SIR model implies a higher limiting value.[49]

A.2 Costs of Infection

In this appendix, we discuss the calibration of the parameters representing private and social costs of infection. Assume that $\xi(a_i, a)$ has the functional form

$$\xi(a_i, a) = \zeta\frac{a_i}{a} + (1 - \zeta),$$

[49] We could trivially make the two limiting values coincide by choosing $\bar{y}$ appropriately. Subject to the modified $\bar{y}$ value, matching the date of peak infections would require to slightly adjust the value of β.

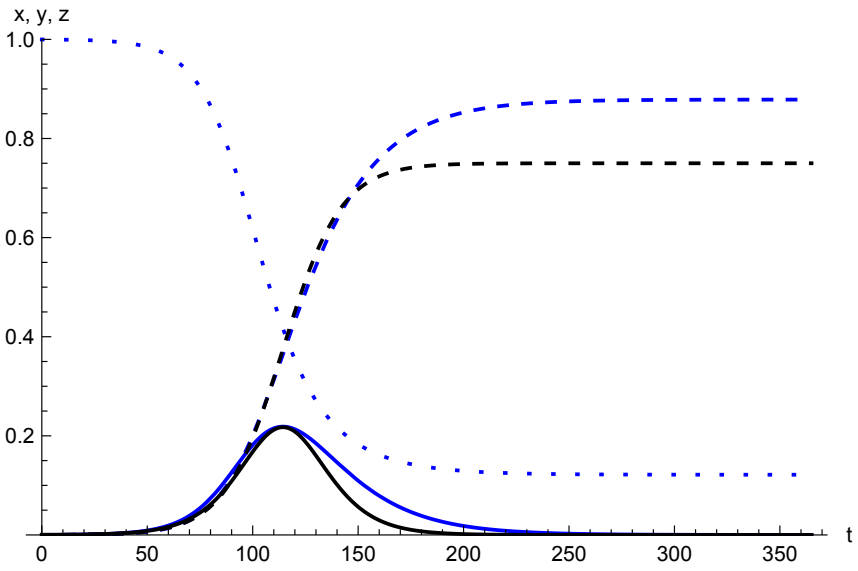

Figure 8: Dynamics in the canonical SIR model (blue) and in the logistic model (black). SIR model: $x(t)$ (dotted), $y(t)$ (solid), and $z(t)$ (dashed). Logistic model: $x(t)$ (dotted, including $\bar{x} = 1 - \bar{y}$), $\dot{y}(t)$ (solid, scaled), and $y(t)$ (dashed).

which satisfies assumption 3 (namely $\xi(a, a) = 1$ and $\xi_{a_i}(a_i, a)$ homogeneous) and therefore, in accordance with proposition 1, $\xi_{a_i}(a, a) = \xi_{a_i}(1, 1)/a = \zeta/a$. In words, households perceive that they bear a share $1-\zeta$ of the social costs of infection independently of their behavior, and the remaining share ζ proportionally to their choice of activity relative to the aggregate activity level.[50]

We calibrate ζ based on U.S. estimates of hospitalization costs and the value of life by Bartsch et al. (2020) and Hall et al. (2020), respectively. Bartsch et al. (2020) estimate direct medical costs including follow up expenses (over a year) of $1.25 trillion under the assumption that eighty percent of the U.S. population are infected. This translates into conditional per-capita costs of about $4,764 (eighty percent of 328 million persons). Hall et al. (2020) assess the value of life at $270,000 per year. With an average remaining life expectancy of 14.5 years every life lost to COVID-19 thus costs $3,915,000. Menachemi et al. (2020) estimate an infection fatality rate of 0.58 percent implying a conditional expected cost of dying from COVID-19 of $22,707 per infected individual. Under the assumption that individuals fully internalize mortality risk but not marginal social medical costs we conclude that $\zeta = 22,707/(22,707 + 4,764) \approx 0.8266$.

To calibrate ψ based on the dollar amount $22,707 + $4,764 we use Hall et al.'s (2020) estimate according to which households would be willing to sacrifice 32 percent of consumption to eliminate all COVID-19 related mortality risk over one year (neglecting other costs).[51] Let $1 - \phi = 0.32$ denote this share. In the model the utility cost of

[50] When $g(a) = a$ (but not when $g(a) = a^2$), $\zeta - 1$ is proportional to the static externality. To see this recall from proposition 1 that the private marginal costs equal $\xi_{a_i}(a, a)\psi\dot{y}^g(y) = \zeta\psi\dot{y}^g(y)/a$ while the social marginal costs equal $\psi d\dot{y}^g(y)/da = \psi n\dot{y}^g(y)/a$. The difference between private and social marginal costs thus is proportional to $\zeta - n$.

[51] Hall et al. (2020) stipulate CRRA preferences with a coefficient of relative risk aversion of two and find

sacrificing the share $1-\phi$ of consumption during N days equals[52]

$$N \cdot \{(1+\ln(a^\star)-a^\star)-(1+\ln(a^\star\phi)-a^\star)\} = -N\ln(\phi).$$

Suppose that (almost) all infections occur during N days such that $\int_0^N \dot{y}(t)dt \approx \bar{y}$ and the total (undiscounted) mortality costs amount to $\hat{\psi}\bar{y}$. We conclude that

$$\hat{\psi} \approx -N\ln(\phi)/\bar{y}.$$

With $N = 365$ and $\bar{y} = 0.75$ this implies social costs due to mortality risk of $\hat{\psi} \approx 187.7$. Adding medical costs we arrive at an estimate for ψ of $\psi = \hat{\psi}/\zeta \approx 227.1$.

Endogenous Costs To analyze the two cases in which the ψ factor entering the costs of infection (4) is a function of the state (rather than an exogenous constant) we need to modify the calibration.

In the first case, we replace ψ in (4) by $\psi^f y(t)(\bar{y}-y(t))$ where $\psi^f > 0$. As we discuss in the text this modification captures congestion effects. Using

$$\int_0^\infty \dot{y}(t)y(t)(\bar{y}-y(t))dt = \int_0^{\bar{y}} y(\bar{y}-y)dy = \frac{\bar{y}^3}{6}$$

and under the maintained assumption that (almost) all infections occur during N days we conclude that $\hat{\psi}^f = \hat{\psi}6/\bar{y}^2$. Accordingly, we multiply ψ by a factor of $6/\bar{y}^2 \approx 10.6667$ in the modified calibration.

In the second case, we replace ψ in (4) by $\psi^l(\ell - y(t)/\bar{y})$ where $\psi^l > 0$. This modification captures learning-by-doing effects. Consistent with recent evidence[53] we assume that the learning effects let unit costs drop by fifty percent over the course of the epidemic: $\ell = 2$. Using

$$\int_0^\infty \dot{y}(t)(2-y(t)/\bar{y})dt = \int_0^{\bar{y}} (2-y/\bar{y})dy = 1.5\bar{y}$$

and under the maintained assumption that (almost) all infections occur during N days we conclude that $\hat{\psi}^l = \hat{\psi}/1.5$. Accordingly, we multiply ψ by a factor of $2/3$ in the modified calibration.

B Proofs

B.1 Proof of Lemma 1

Proof. To prove continuity at $y = 0$, we consider without loss of generality the case of $\nu = 0$. Consider y_0 and y_1 with $0 < y_0 < y_1 \leq y_\infty(a^\star)$ where $y_\infty(a^\star)$ denotes the steady-state level of y when activity equals $a^\star$ (see equation (2)). Let $\mathcal{C}(a^\star, y)$ denote the costs

that the willingness to sacrifice consumption equals 28 percent. With logarithmic utility this translates into 32 percent.

[52]We neglect time discounting as do Hall et al. (2020). Note that only the benefit of economic activity ("consumption") not the cost associated with it ("labor supply") is reduced by the fraction $1-\phi$.

[53]See, for example, Armstrong et al. (2020), Dennis et al. (2020), or *The Financial Times* (`ig.ft.com/coronavirus-global-data/`).

of infection when activity equals $a^\star$ and the state is given by (y,t). Define

$$\tilde{V}(y_0,t_0)=\int_0^{t(y_0,y_1)}\mathrm{e}^{-\rho\,\tau}u(a^\star)d\tau-\int_{y_0}^{y_1}\mathrm{e}^{-\rho\,t(y_0,y)}\mathcal{C}(a^\star,y)dy+\mathrm{e}^{-\rho\,t(y_0,y_1)}V(y_1,t_0+t(y_0,y_1)).$$

Recall that $t(y_0,y)$ denotes the time span over which the state variable moves from y_0 to y. $\tilde{V}(y_0,t_0)$ represents the value at time t_0 conditional on $y=y_0$ when activity is fixed at level $a^\star$ until $y=y_1$ is reached, from which point on activity is chosen optimally. By construction, $\tilde{V}(y_0,t_0)$ constitutes a lower bound for $V(y_0,t_0)$ such that $\tilde{V}(y_0,t_0)\leq V(y_0,t_0)\leq U^\star$. Note that $\mathcal{C}(a^\star,y)$ is bounded for $y\in[0,\bar{y}]$. Also, from equation (1), $\lim_{y\downarrow 0}\mathcal{C}(a^\star,y)=0$. Finally, from equation (3), $\lim_{y_0\downarrow 0}t(y_0,y)=\infty$. It follows that $\lim_{y_0\downarrow 0}\tilde{V}(y_0,t_0)=U^\star$ and therefore $\lim_{y_0\downarrow 0}V(y_0,t_0)=U^\star$. □

B.2 Proof of Lemma 2

Proof. From the government's HJB equation, the envelope condition reads

$$(\rho+\nu)V'(y)=-g(a(y))\beta\left[(\bar{y}-2y)(\psi-V'(y))-y(\bar{y}-y)V''(y)\right]-\gamma\left(V'(y)+yV''(y)\right).$$

Let $\hat{y}$ denote a point where V reaches a local minimum or maximum. Evaluated at $\hat{y}$ the envelope condition reduces to

$$g(a(\hat{y}))\beta(\bar{y}-2\hat{y})\psi=\left[g(a(\hat{y}))\beta(\bar{y}-\hat{y})-\gamma\right]\hat{y}V''(\hat{y}). \tag{16}$$

To see that V has a unique minimum suppose to the contrary that there exist multiple local minima. Consider two neighboring minima at, say, y^a and y^c with $y^a<y^c$. Then there must exist a local maximum at some y^b with $y^a<y^b<y^c$. Note that $V''(y^a)>0$, $V''(y^b)<0$, and $V''(y^c)>0$. This requires at least two inflection points between y^a, y^b, and y^c at which $V''(y)$ and thus the right-hand side of equation (16) equals zero. Since the left-hand side of the equation has at most one zero we have arrived at a contradiction which proves that V has a unique minimum, $y^{\min}$.

To see that $y^{\min}\leq\bar{y}/2$ suppose to the contrary that $y^{\min}>\bar{y}/2$. Since the minimum is unique we have $V'(\bar{y}/2)<0$. From the envelope condition,

$$(\rho+\nu+\gamma)V'(\bar{y}/2)=\left(g(a(\bar{y}/2))\beta(\bar{y}/2)^2-\gamma(\bar{y}/2)\right)V''(\bar{y}/2)$$

and thus, since $g(a)\beta\bar{y}>2\gamma$ (from assumption 1), $V''(\bar{y}/2)<0$. Since the minimum lies to the right of $\bar{y}/2$ there must exist an inflection point, say y^i, with $\bar{y}/2<y^i<y^{\min}$, $V'(y^i)<0$, and $V''(y^i)=0$. But evaluated at y^i the envelope condition implies

$$\underbrace{(\rho+\nu+\gamma)V'(y^i)}_{<0}=-\underbrace{g(a(y^i))\beta}_{>0}\underbrace{(\bar{y}-2y^i)}_{<0}\underbrace{(\psi-V'(y^i))}_{>0},$$

which yields a contradiction. We conclude that $y^{\min}\leq\bar{y}/2$.

From equation (16), when $\gamma=0$, $V''(y^{\min})=(\bar{y}-2y^{\min})\psi/y^{\min}/(\bar{y}-y^{\min})$ which is strictly decreasing in $y^{\min}$. This proves the last claim. □

B.3 Proof of Proposition 1

Proof. From assumption 3, $\xi_{a_i}(a_i, a)$ is homogenous of degree m say. Accordingly, $\xi(a_i, a)$ is homogeneous of degree $m + 1$. Since $\xi(\lambda a, \lambda a) = 1$ for all $\lambda \neq 0$, m equals -1: $\xi_{a_i}(a, a) = \lambda \xi_{a_i}(\lambda a, \lambda a)$. Letting $\lambda = 1/a$ implies $\xi_{a_i}(a, a) = \xi_{a_i}(1, 1)/a \equiv \zeta/a$ and yields the first result. The derivations also imply that $\zeta - n$ is proportional to the static externality, establishing the second result. □

B.4 Proof of Proposition 2

Proof. The result follows directly from equation (9) in proposition 1. □

B.5 Proof of Proposition 3

Proof. Part i. follows because the value function is decreasing in a neighborhood of $y = 0$ (from lemma 1), implying that the static and dynamic externalities both are negative such that the government imposes a lockdown. Part ii. follows from the fact that $V_y(y, t)$ eventually is positive. For "small" $1-\zeta/n$ the total externality therefore eventually turns positive. As a consequence, the government imposes an inverse lockdown if and when the economy reaches the relevant part of the state space (requiring the condition on γ). If V is locally convex then the total externality switches signs at y^c, establishing part iii. □

B.6 Proof of Proposition 4

Proof. Uniqueness of equilibrium follows by construction. Differentiating the equilibrium solution with respect to γ implies that in a neighborhood of $\gamma = 0$ (and $a_\infty = 1$), $da_\infty/d\gamma|_{\gamma=0} = -\psi\bar{y}\zeta$. For the government's solution, equation (10) implies $da_\infty/d\gamma|_{\gamma=0} = -\psi(\rho + \nu)\bar{y}/(\rho + \nu + \beta\bar{y})$. With $\gamma = 0$ the two steady-state values are identical. We conclude that for small values of γ the government chooses a higher steady-state activity level (and, from equation (1), a higher y_∞) than in decentralized equilibrium if

$$(\rho + \nu)(1 - \zeta) < \zeta\beta\bar{y}.$$

Consider next the derivative $da/d\gamma$ under the optimal policy. Note that for any y we have

$$\begin{aligned}(\rho + \nu)V(y) &= \ln(a(y)) + 1 - a(y)\{1 + \beta y(\bar{y} - y)(\psi - V'(y))\} - \gamma y V'(y) + \nu U^\star \\ &= \ln(a(y)) - \gamma y V'(y) + \nu U^\star = \ln(a(y)) - \gamma y V'(y).\end{aligned}$$

Differentiating the equation with respect to γ yields

$$(\rho + \nu)\frac{dV(y)}{d\gamma} = \frac{1}{a(y)}\frac{da(y)}{d\gamma} - yV'(y) - \gamma y\frac{dV'(y)}{d\gamma} < 0,$$

where the inequality follows from the fact that higher re-infection risk must reduce the government's value for any $y > 0$. Since $V'(y^{\min}) = 0$ we have

$$\frac{1}{a(y^{\min})}\frac{da(y^{\min})}{d\gamma} - \gamma y^{\min}\frac{dV'(y^{\min})}{d\gamma} < 0.$$

Using the expression for the government's optimal choice, equation (7), this implies

$$a(y^{\min})\beta y^{\min}(\bar{y}-y^{\min})\frac{dV'(y^{\min})}{d\gamma}-\gamma y^{\min}\frac{dV'(y^{\min})}{d\gamma}<0.$$

We conclude that $dV'(y^{\min})/d\gamma<0$ and, since $a(y^{\min})$ depends positively on $V'(y^{\min})$ (see equation (7)), also $da(y^{\min})/d\gamma<0$ as long as $\gamma<a(y^{\min})\beta(\bar{y}-y^{\min})$.

To establish that the latter inequality is indeed satisfied note that assumption 1 and $g(a)=a$ imply $a(y^{\min})\beta\bar{y}>2\gamma$, or, using the expression for the government's choice,

$$\frac{\beta\bar{y}}{1+\beta y^{\min}(\bar{y}-y^{\min})(\psi-V'(y^{\min}))}=\frac{\beta\bar{y}}{1+\beta y^{\min}(\bar{y}-y^{\min})\psi}>2\gamma.$$

From lemma 2, $y^{\min}\leq\bar{y}/2$ such that the left-hand side of the preceding inequality is bounded below by

$$\frac{\beta\bar{y}}{1+\beta(\bar{y}/2)(\bar{y}-(\bar{y}/2))\psi}=\frac{\beta\bar{y}}{1+\beta\bar{y}^2/4\psi}.$$

The result therefore follows under the additional, imposed, parameter restriction

$$\frac{\beta\bar{y}}{1+\beta\bar{y}^2/4\psi}>2\gamma\quad\text{or}\quad\frac{2\beta\bar{y}}{4+\beta\bar{y}^2\psi}>\gamma.$$

□

B.7 Proof of Proposition 5

Proof. Differentiating the first-order condition, $\mathcal{H}_a(a(t),y(t),t)=0$, with respect to time implies

$$\dot{a}(t)\frac{d}{d\,a(t)}\frac{u'(a(t))}{g'(a(t))}=\beta(\bar{y}-2y(t))\dot{y}(t)\,(\psi-\mu(t))-\beta y(t)(\bar{y}-y(t))\dot{\mu}(t).$$

The right-hand side of this condition collapses to zero when we substitute for $\dot{y}(t)$ from the law of motion and for $\dot{\mu}(t)$ from the condition $\mathcal{H}_y(a(t),y(t),t)=-\dot{\mu}(t)$. We conclude that $\dot{a}(t)=0$.

Using the functional form assumptions, an optimal allocation thus is characterized by the following two equations in two unknowns, a and $y_T\equiv y(T)$:

$$a = \frac{1}{1+\beta y_T(\bar{y}-y_T)\psi}, \tag{17}$$

$$y_T = \frac{\bar{y}}{1+e^{-a\beta\bar{y}T}\left(\frac{\bar{y}}{y(0)}-1\right)}. \tag{18}$$

Equation (17) represents the government's first-order condition at time T; it uses the fact that $\mu(T)=0$. Equation (18) follows from the law of motion, see equation (2). Both

conditions can be represented as functions in (y_T, a) space. Specifically, let $a(y_T)$ denote the right-hand side of equation (17) and let

$$A(y_T) \equiv \frac{1}{\beta \bar{y} T} \ln \left(\frac{\frac{\bar{y}}{y_0} - 1}{\frac{\bar{y}}{y_T} - 1} \right)$$

denote the value of a that solves equation (18). A candidate optimal allocation $(a^\diamond, y_T^\diamond)$ satisfies $a^\diamond = a(y_T^\diamond) = A(y_T^\diamond)$.

Note that in the relevant range, function $a(y_T)$ is U-shaped; strictly positive; symmetric around $y_t = \bar{y}/2$; and satisfies $a(0) = 1 = a(\bar{y})$. Function $A(y_T)$ is strictly increasing; satisfies $A(y_0) = 0$; $\lim_{y_T \to \bar{y}} A(y_T) = \infty$; and an increase in T proportionally reduces $A(y_T)$ for any given y_T. This implies the following: (i) For small T, there exists a unique candidate optimal allocation (a, y_T) with $a(y_T) = A(y_T)$ and $y_T \approx y_0$. (ii) For $T \to \infty$, there exists a unique candidate optimal allocation (a, y_T) with $a(y_T) = A(y_T)$ and $y_T \approx \bar{y}$.

Let

$$\hat{a} \equiv \frac{1}{1 + \beta \bar{y}^2 \psi / 4} < 1, \quad \hat{T} \equiv \frac{\ln(\bar{y}/y(0) - 1)}{\hat{a} \beta \bar{y}} < \infty,$$

and suppose that y_0 is sufficiently small, namely y_0 satisfies both $y_0 < \bar{y}/2$ and

$$\frac{1 + \left(\frac{\bar{y}}{y(0)} - 1 \right)^{\frac{\hat{a}-1}{\hat{a}}}}{1 - \left(\frac{\bar{y}}{y(0)} - 1 \right)^{\frac{\hat{a}-1}{\hat{a}}}} \ln \left(\frac{\bar{y}}{y(0)} - 1 \right) > - \frac{\hat{a} \beta \psi \bar{y}^2 / 2}{1 + \ln(\hat{a}) - \hat{a}}. \tag{19}$$

Starting from $T \approx 0$ and thus, $y_T \approx y_0$, consider a continuous increase in T. This scales the $A(y_T)$ schedule down such that the intersection of the $a(y_T)$ and $A(y_T)$ schedules shifts to the right and down in (y_T, a) space. When T reaches $\hat{T}$, the intersection reaches the point $(y_T, a) = (\bar{y}/2, \hat{a})$. The welfare generated under the policy $a = \hat{a}$ given $T = \hat{T}$ equals $1 + \ln(\hat{a}) - \hat{a} - \psi \bar{y}/2$. In contrast, welfare under the policy $a = a^\star$ given $T = \hat{T}$ equals

$$(1 + \ln(a^\star) - a^\star) \hat{T} - \psi \frac{\bar{y}}{1 + e^{-\beta \bar{y} \hat{T}} \left(\frac{\bar{y}}{y(0)} - 1 \right)} = -\psi \frac{\bar{y}}{1 + \left(\frac{\bar{y}}{y(0)} - 1 \right)^{1 - \frac{1}{\hat{a}}}}.$$

Under condition (19), this last term strictly exceeds $(1 + \ln(\hat{a}) - \hat{a}) \hat{T} - \psi \bar{y}/2$. Accordingly, $a = \hat{a}$ is suboptimal when $T = \hat{T}$; in other words, when $T = \hat{T}$ there exists an activity level $a \neq \hat{a}$ which also solves the equation system (17)–(18) and dominates the choice $a = \hat{a}$.

We have established that the system (17)–(18) has a unique solution for $T \approx 0$, the lowest solution of the system (17)–(18). We have also established that the system (17)–(18) has multiple solutions for $T = \hat{T}$ and that the optimal solution is not that lowest solution. It follows that there exists some $T^\diamond < \hat{T}$ at which the optimal solution ceases to be the lowest solution. Since the lowest solution and the optimal solution are different, a is dis-continuous at $T = T^\diamond$. □

C Additional Figures

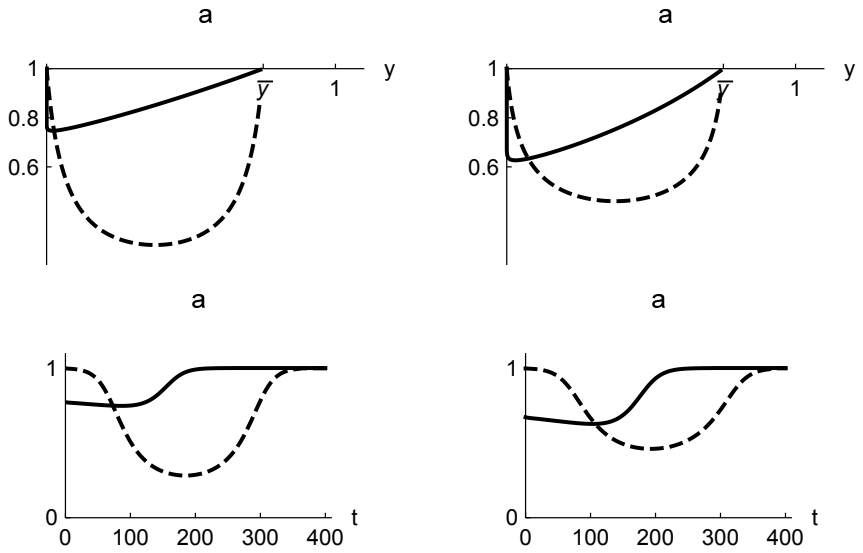

Figure 9: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with quadratic effect of activity on infections (right panels).

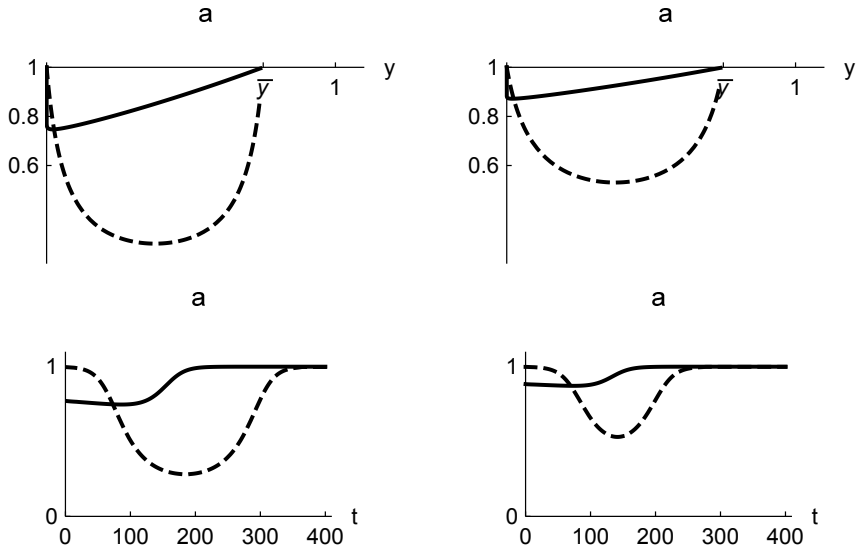

Figure 10: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with stronger curvature of u (right panels).

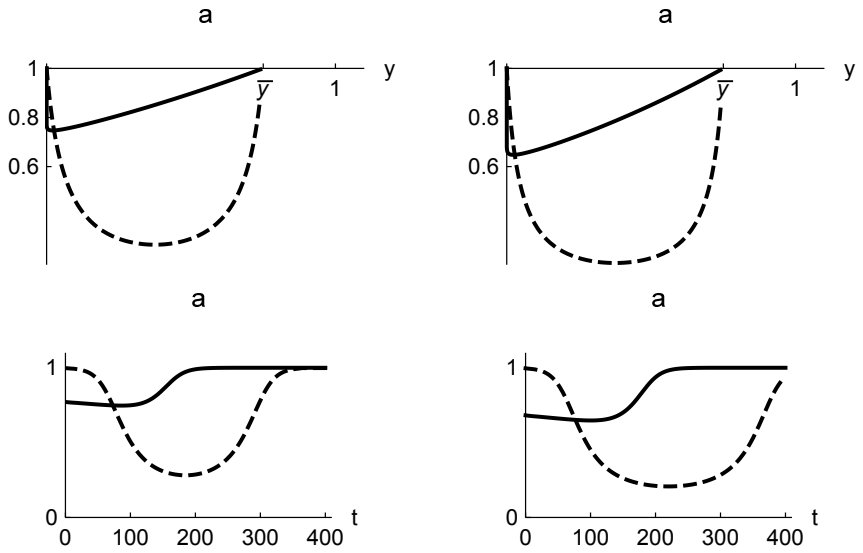

Figure 11: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with higher costs of infection (right panels).

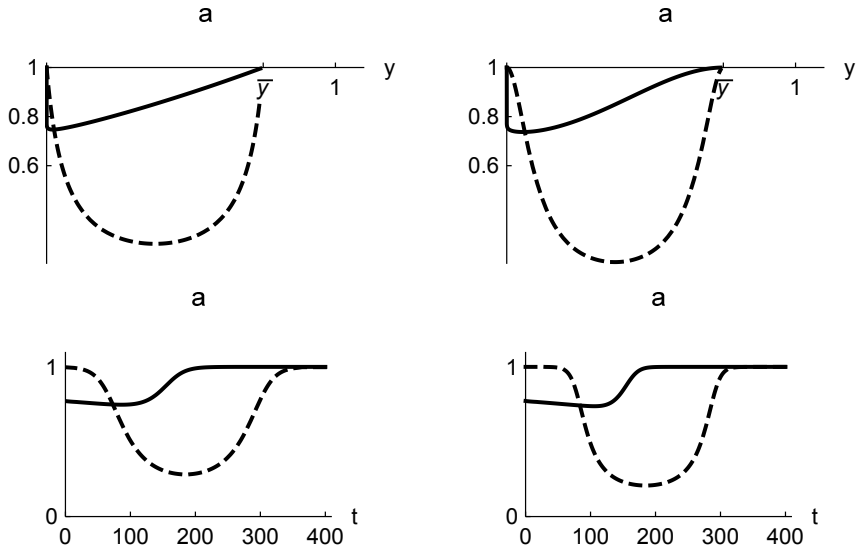

Figure 12: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with congestion effects (right panels).

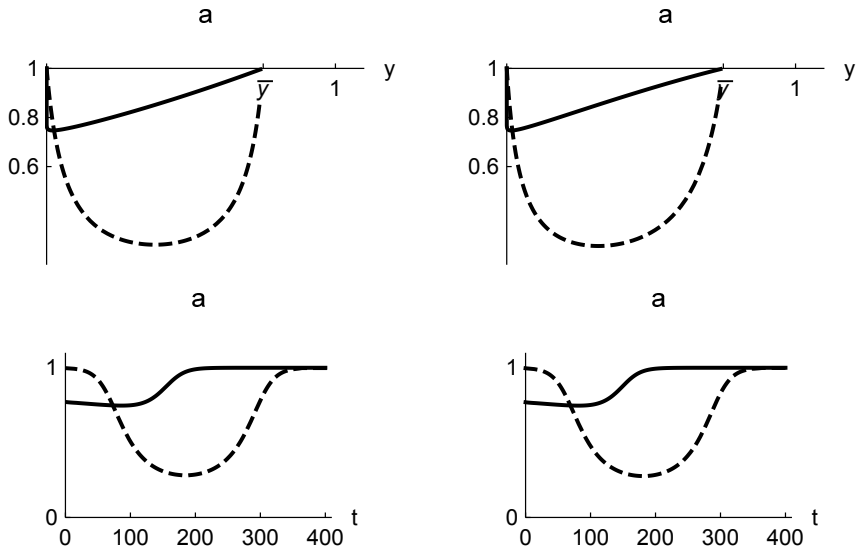

Figure 13: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with learning effects (right panels).

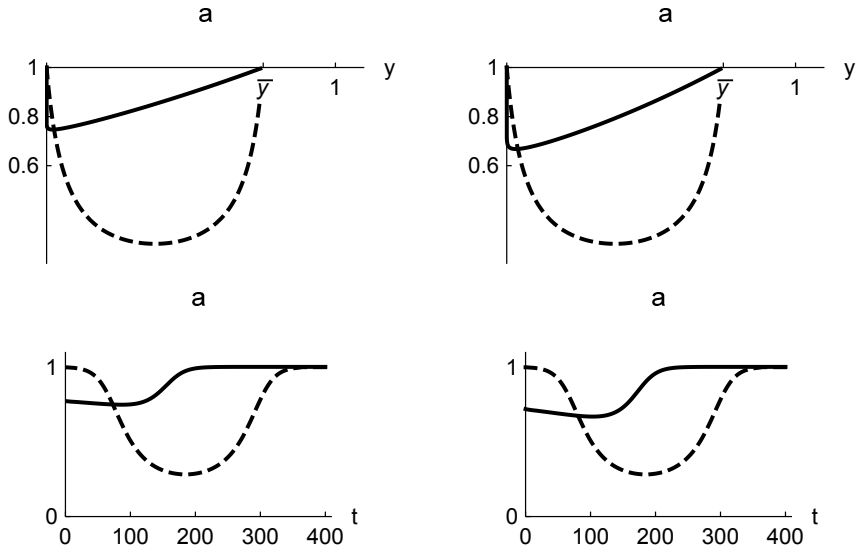

Figure 14: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with a higher arrival rate of a cure (right panels).

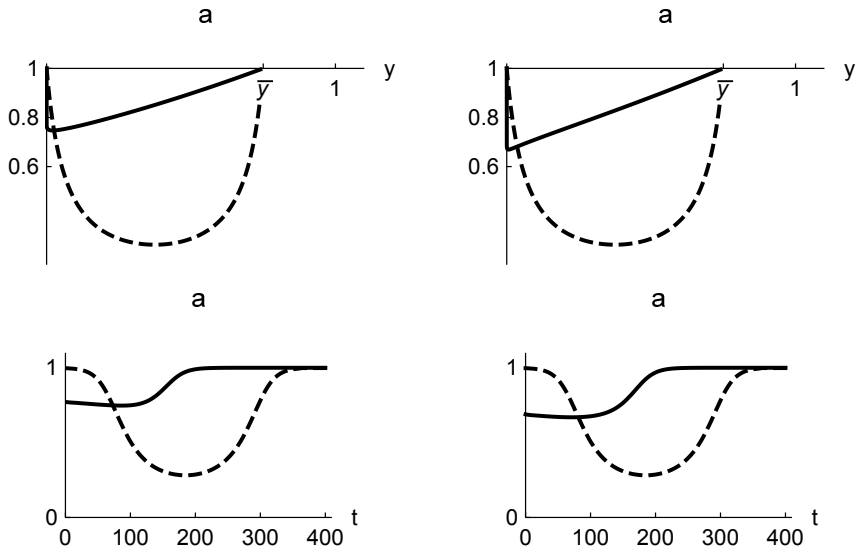

Figure 15: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with stochastic reduction in β (right panels).

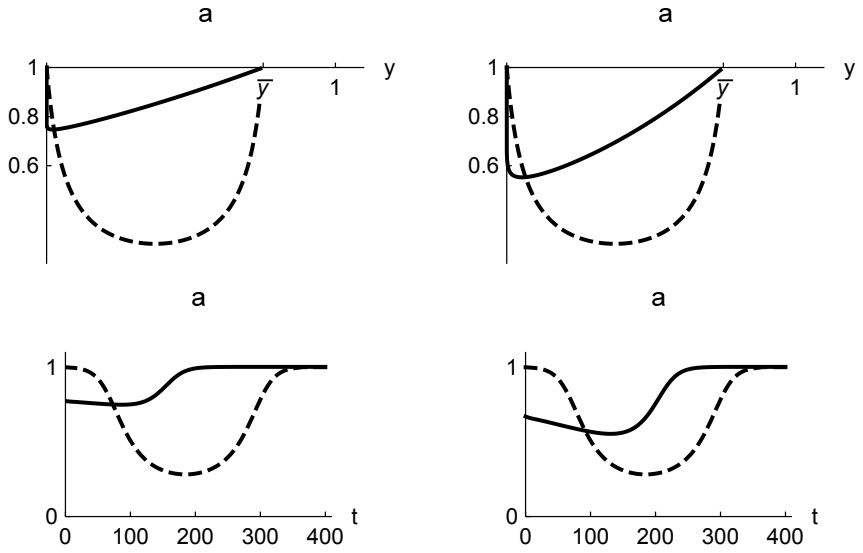

Figure 16: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with multiple waves (right panels).

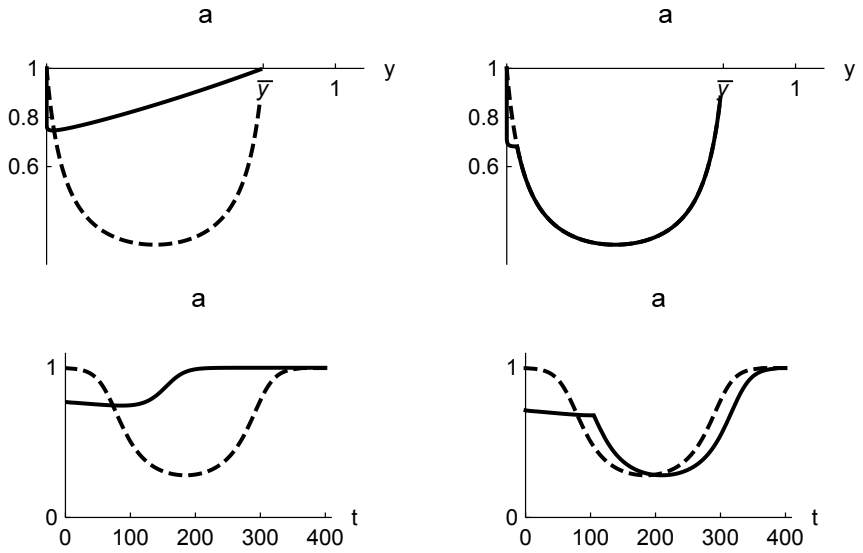

Figure 17: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with constraints on policy instruments (right panels).

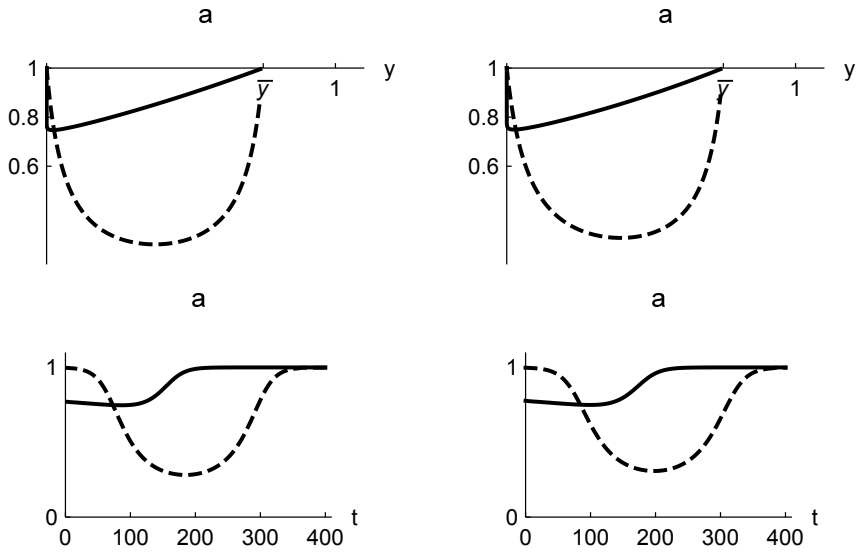

Figure 18: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with observable infection status (public information, right panels).

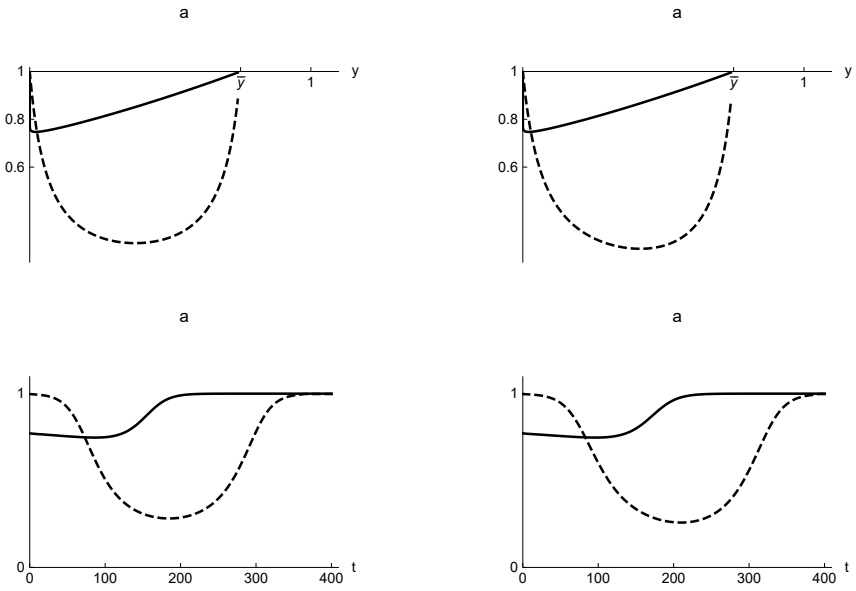

Figure 19: Activity level in the government's program (solid) and in equilibrium (dashed): Baseline model (left panels) and model with observable infection status (private information, right panels).

References

Abel, A. B. and Panageas, S. (2020). Optimal management of a pandemic in the short run and the long run, *Working Paper 27742*, NBER, Cambridge, Massachusetts.

Acemoglu, D., Chernozhukov, V., Werning, I. and Whinston, M. D. (2020). Optimal targeted lockdowns in a multi-group SIR model, *Working Paper 27102*, NBER, Cambridge, Massachusetts.

Alvarez, F., Argente, D. and Lippi, F. (2020). A simple planning problem for COVID-19 lockdown, *Working Paper 26981*, NBER, Cambridge, Massachusetts.

Armstrong, R. A., Kane, A. D. and Cook, T. M. (2020). Decreasing mortality rates in ICU during the COVID-19 pandemic, *Anaesthesia* .

Atkeson, A. (2020). What will be the economic impact of COVID-19 in the US? rough estimates of disease scenarios, *Working Paper 26867*, NBER, Cambridge, Massachusetts.

Bailey, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and its Applications*, 2 edn, Hafner Press, New York.

Bartsch, S. M., Ferguson, M. C., McKinnell, J. A., O'Shea, K. J., Wedlock, P. T., Siegmund, S. S. and Lee, B. Y. (2020). The potential health care costs and resource use associated with COVID-19 in the United States, *Health Affairs* **forthcoming**.

Bethune, Z. A. and Korinek, A. (2020). Covid-19 infection externalities: Trading off lives vs. livelihoods, *Working Paper 27009*, NBER, Cambridge, Massachusetts.

Bisin, A. and Moro, A. (2020). Learning epidemiology by doing: The empirical implications of a spatial-SIR model with behavioral responses, *Working Paper 27590*, NBER, Cambridge, Massachusetts.

Çenesiz, A. and Guimaraẽs, L. (2020). COVID-19: What if immunity wanes? arXiv https://arxiv.org/abs/2008.03283.

Dennis, J. M., McGovern, A. P., Vollmer, S. J. and Mateen, B. A. (2020). Improving COVID-19 critical care mortality over time in England: A national cohort study, March to June 2020. medRxiv https://doi.org/10.1101/2020.07.30.20165134.

Diamond, P. A. and Maskin, E. (1979). An equilibrium analysis of search and breach of contract, i: Steady states, *The Bell Journal of Economics* **10**(1): 282–316.

Eichenbaum, M. S., Rebelo, S. and Trabandt, M. (2020). The macroeconomics of epidemics, *Working Paper 26882*, NBER, Cambridge, Massachusetts.

Ellison, G. (2020). Implications of heterogeneous SIR models for analyses of COVID-19, *Covid Economics* **53**: 1–32.

Farboodi, M., Jarosch, G. and Shimer, R. (2020). Internal and external effects of social distancing in a pandemic, *Working Paper 27059*, NBER, Cambridge, Massachusetts.

Ferguson, N. M., Laydon, D., Nedjati-Gilani, G. and collaborators (2020). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, *Report 9*, Imperial College, London.

Fetzer, T. and Graeber, T. (2020). Does contact tracing work? quasi-experimental evidence from an excel error in England, *Working Paper 521*, Centre for Competitive Advantage in the Global Economy, Coventry.

Garibaldi, P., Pissarides, C. and Moen, E. R. (2020). Static and dynamic inefficiencies in an optimizing model of epidemics, *Discussion Paper 15439*, CEPR, London, UK.

Gersovitz, M. and Hammer, J. S. (2004). The economical control of infectious diseases, *Economic Journal* **114**: 1–27.

Giannitsarou, C., Kissler, S. and Toxvaerd, F. (2020). Waning immunity and the second wave: Some projections for SARS-CoV-2, *Discussion Paper 14852*, CEPR, London, UK.

Gonzalez-Eiras, M. and Niepelt, D. (2020a). On the optimal 'lockdown' during an epidemic, *Covid Economics* **7**: 68–87.

Gonzalez-Eiras, M. and Niepelt, D. (2020b). Tractable epidemiological models for economic analysis, *Discussion Paper 14791*, CEPR.

Goolsbee, A. and Syverson, C. (2020). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020, *Working Paper 27432*, NBER, Cambridge, Massachusetts.

Greenstone, M. and Nigam, V. (2020). Does social distancing matter?, *Covid Economics* **7**: 1–23.

Gudbjartsson, D. F., Norddahl, G. L., Melsted, P., Gunnarsdottir, K., Holm, H., Eythorsson, E., Arnthorsson, A. O., Helgason, D., Bjarnadottir, K., Ingvarsson, R. F., Thorsteinsdottir, B., Kristjansdottir, S., Birgisdottir, K., Kristinsdottir, A. M., Sigurdsson, M. I., Arnadottir, G. A., Ivarsdottir, E. V., Andresdottir, M., Jonsson, F., Agustsdottir, A. B., Berglund, J., Eiriksdottir, B., Fridriksdottir, R., Gardarsdottir, E. E., Gottfredsson, M., Gretarsdottir, O. S., Gudmundsdottir, S., Gudmundsson, K. R., Gunnarsdottir, T. R., Gylfason, A., Helgason, A., Jensson, B. O., Jonasdottir, A., Jonsson, H., Kristjansson, T., Kristinsson, K. G., Magnusdottir, D. N., Magnusson, O. T., Olafsdottir, L. B., Rognvaldsson, S., le Roux, L., Sigmundsdottir, G., Sigurdsson, A., Sveinbjornsson, G., Sveinsdottir, K. E., Sveinsdottir, M., Thorarensen, E. A., Thorbjornsson, B., Thordardottir, M., Saemundsdottir, J., Kristjansson, S. H., Josefsdottir, K. S., Masson, G., Georgsson, G., Kristjansson, M., Moller, A., Palsson, R., Gudnason, T., Thorsteinsdottir, U., Jonsdottir, I., Sulem, P. and Stefansson, K. (2020). Humoral

immune response to SARS-CoV-2 in Iceland, *The New England Journal of Medicine* **forthcoming**.

Hall, R. E., Jones, C. I. and Klenow, P. J. (2020). Trading off consumption and covid-19 deaths, *Working Paper 27340*, NBER, Cambridge, Massachusetts.

Hethcote, H. W. (1989). Three basic epidemiological models, *in* L. Gross, T. G. Hallam and S. A. Levin (eds), *Applied Mathematical Ecology*, Springer, Berlin, pp. 119–144.

Hethcote, H. W. (2000). The mathematics of infectious diseases, *SIAM Review* **42**(4): 599–653.

Jones, C. J., Philippon, T. and Venkateswaran, V. (2020). Optimal mitigation policies in a pandemic: Social distancing and working from home, *Working Paper 26984*, NBER, Cambridge, Massachusetts.

Kaplan, G., Moll, B. and Violante, G. L. (2020). The great lockdown and the big stimulus: Tracing the pandemic possibility frontier for the U.S., *Working Paper 27794*, NBER, Cambridge, Massachusetts.

Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society, Series A* **115**(772): 700–721.

Menachemi, N., Yiannoutsos, C. T., Dixon, B. E., Duszynski, T. J., Fadel, W. F., Wools-Kaloustian, K. K., Unruh Needleman, N., Box, K., Caine, V., Norwood, C., Weaver, L. and Halverson, P. K. (2020). Population point prevalence of SARS-CoV-2 infection based on a statewide random sample—Indiana, *Morbidity and Mortality Weekly Report, April 25–29* **69**: 960–964.

Miclo, L., Spiro, D. and Weibull, J. W. (2020). Optimal epidemic suppression under an ICU constraint. Unpublished, Stockholm School of Economics.

Russell, T. W., Hellewell, J., Jarvis, C. I., van Zandvoort, K., Abbott, S., Ratnayake, R., CMMID COVID-19 working group, Flasche, S., Eggo, R. M. and Kucharski, A. J. (2020). Estimating the infection and case fatality ratio for COVID-19 using age-adjusted data from the outbreak on the Diamond Princess cruise ship, *Eurosurveillance* **25**(12).

To, K. K.-W., Hung, I. F.-N., Ip, J. D., Chu, A. W.-H., Chan, W.-M., Tam, A. R., Fong, C. H.-Y., Yuan, S., Tsoi, H.-W., Ng, A. C.-K., Lee, L. L.-Y., Wan, P., Tso, E., To, W.-K., Tsang, D., Chan, K.-H., Huang, J.-D., Kok, K.-H., Cheng, V. C.-C. and Yuen, K.-Y. (2020). COVID-19 re-infection by a phylogenetically distinct SARS-coronavirus-2 strain confirmed by whole genome sequencing, *Clinical Infectious Diseases* **forthcoming**.

Toxvaerd, F. (2020). Equilibrium social distancing, *Working Paper 2020/08*, Cambridge Institute for New Economic Thinking, Cambridge.