OPTIMAL CONTROL OF AN EPIDEMIC THROUGH SOCIAL DISTANCING

By

Thomas Kruse and Philipp Strack

April 2020

Revised July 2020

COWLES FOUNDATION DISCUSSION PAPER NO. 2229R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Optimal Control of an Epidemic through Social Distancing*

Thomas Kruse† Philipp Strack‡

July 28, 2020

Abstract

We analyze how to optimally engage in social distancing in order to minimize the spread of an infectious disease. We identify conditions under which any optimal policy is single-peaked, i.e., first engages in increasingly more social distancing and subsequently decreases its intensity. We show that an optimal policy might delay measures that decrease the transmission rate substantially to create herd-immunity and that engaging in social distancing sub-optimally early can increase the number of fatalities. Finally, we find that optimal social distancing can be an effective measure and can substantially reduce the death rate of a disease.

Keywords— Social Distancing, SIR model, Time-Optimal Control of an Epidemic

1 Introduction

This paper theoretically analyzes how to optimally engage in measures to contain the spread of an infectious disease. We formalize this question in the context of a standard model from epidemiology, the Susceptible-Infected-Recovered (SIR) model (Kermack and McKendrick, 1927). This model divides the population into the three groups susceptible, infected and recovered. People transition from one group into another at given exogenously specified rates depending on the size of each sub-population. We extend this model by introducing an additional parameter controlled by the planner that affects the rate at which the disease is transmitted. We think of this parameter as capturing

*We want to thank Stefan Ankirchner, Alexis Toda, and Aleh Tsyvinski for helpful discussions. We plan on updating this paper frequently, the current version of the paper can be found at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3581295. The first version of this paper was circulated on April 20, 2020.

†Giessen University, thomas.kruse@math.uni-giessen.de

‡Yale University and CEPR research fellow, philipp.strack@yale.edu

political measures such as social distancing, the lockdown of businesses, schools, universities and other institutions. While such measures reduce the spread of the disease, they often come at a substantial economic and social cost. We model this trade-off by considering a planner who faces convex cost in the number of infected (capturing the number of people whose death is caused by the disease) and the reduction in transmission rate (capturing the cost of shutting down society). Convexity of the cost allows for the possibility that the probability of dying from the disease increases in the fraction of the population which is infected, for example due to capacity constraints of the healthcare system.

Our analysis identifies several features of any optimal policy: First, whenever the probability of dying from the disease does not increase to quickly in the number of infected, the optimal policy is single peaked in the sense that first the measures to reduce the transmission rate are escalated until some point in time, and after this point measures are reduced. Second, if furthermore the cost of reducing the transmission rate is linear, meaning that closing half of society for two days is equally costly as closing all of society for one day, only the most extreme policies are used at any point in time. Either, the planner imposes the maximal possible lockdown or no restrictions at all. Intuitively, the planner can achieve a greater effect by imposing a more extreme policy for a shorter time and thus does not find it optimal to use intermediate policies. These results imply that for linear cost the optimal policy has a simple structure and consists of maximally three phases: first it imposes no restrictions then it imposes as many restrictions as possible, and finally in the third phase imposes no restrictions at all. This result drastically simplifies the search for an optimal policy as the planner has to only optimize over the start and end time of the social distancing period. We furthermore show that in this case the number of infected peaks at most twice under any optimal policy.

We then illustrate some further insights using parameters that are commonly chosen to model the Covid-19 epidemic in 2020. We first characterize the optimal timing of the social distancing period given that the planner has access to a certain budget of days of social distancing. We find that the optimal social distancing is often substantially delayed. For example, if the planer has a budget of 100 days of social distancing in the next 360 days after 0.1% of the population are infected it is optimal to delay social distancing by 50 days. This initial period of letting the disease spread uncontrolled is useful as it creates "herd immunity" and thereby reduces the overall severity of the epidemic. We show by an example that the benefit of herd immunity is so strong that sometimes more social distancing can *increase* the number of people that die from the epidemic. We show that in this example more people die when social distancing is imposed from day 0-100 compared to day 50-100. As this example suggests, benefit of optimally timing social distancing measures is often large and we illustrate this by comparing social distancing in the first t days after 0.1 percent are infected to t days of optimally timed social distancing. Finally, we quantify the optimal amount of social distancing.

Related Literature Our theoretical results extend the literature on the optimal control of an infectious disease (for an overview see chapter 5 in Wickwire, 1977). In general, there are three policy tools commonly used to control an infectious disease: (i) immunization, (ii) testing and isolation of infected individuals and (iii) lockdown measures that lead to a reduction in the contact rate for the whole population. Most of the preceding literature has focused on immunization and selective isolation measures. Abakuks (1973) considers how to optimally isolate infectious population if infectious population can be instantaneously isolated. Abakuks (1972, 1974) determine the optimal vaccination strategy in the same framework. Morton and Wickwire (1974) and Wickwire (1975) extend the previous work on vaccination and isolation by considering flow controls. Behncke (2000) considers more general functional forms and Hansen and Day (2011) allows for hard bounds on the control, while considering vaccination and isolation policies simultaneously. The general insight from this literature is that for linear cost of vaccination/testing the optimal policy switches from vaccinating/testing the population at the maximal feasible intensity until some point in time to vaccinating/testing no one after that point in time.

While the previously discussed literature has analyzed vaccination and isolation policies, little is known about optimal lockdown policies. In general, the analysis of optimal lockdown policies in the standard SIR model is a challenging problem, due to its non-linear structure and the three dimensional state space.[1] To the best of our knowledge the first article that discusses the optimal social distancing or lockdown policies in the SIR model is chapter 4 in Behncke (2000). In a model without terminal cost, this paper observes that the optimal policy depends only on the shadow price difference between infected and susceptible.[2] Our paper contributes to this literature by theoretically deriving properties of the optimal lockdown policy about which little has been known before. Most closely related to our work, Miclo et al. (2020) derive the optimal policy that keeps the number of infected below some exogenous threshold when the cost of reducing the transmission rate is linear. The two studies nicely complement each other as we focus on the case where the cost of infected is not too convex[3], while their model intuitively corresponds to the case where the cost is infinitely convex at the threshold that captures the capacity of the healthcare system.

Finally, our paper relates to the recent, quickly growing literature in economics that numerically studies optimal policies for the epidemic of Covid-19 in the context of SIR models (see for example Alvarez et al., 2020; Kissler et al., 2020; Toda, 2020; Acemoglu et al., 2020; Akbarpour et al.,

[1]Parts of the literature focused on variants of the model that allow for a reduction in the dimension of the state space. Rowthorn and Toxvaerd (2012) solved the model where infected population always recovers (does not die) and can become susceptible again. Gonzalez-Eiras and Niepelt (2020) adopt a model from Bohner et al. (2019) that assumes that there is no contact between infectious population and population that was previously infected (see the discussion in Bohner et al. (2019) on page 2) and are able to obtain closed form solutions for the optimal policy.

[2]This insight generalizes to our analysis (see Proposition 1).

[3]See Assumption 1 below.

2020; Garibaldi et al., 2020; Kantner, 2020). While it is not a goal of this paper to make any recommendations for the current Covid-19 epidemic we hope that the formal analysis and insights into the structure of optimal policies will be useful in the rapidly evolving discussion of how to optimally react to the Covid-19 epidemic.[4]

2 The Evolution of an Epidemic

The SIR Model To model the spread of an infectious disease we rely on a basic model from epidemiology, the *Susceptible Infected Recovered* (SIR) model introduced in Kermack and McKendrick (1927). We divide society into three groups: susceptible s, infected i, and the rest which is either immune to the disease as they recovered from it or died. We denote by $s(t)$ the fraction of the population that is healthy, but susceptible to disease at time t, and by $i(t)$ the fraction of the population that is infected. The SIR model assumes the number of people that gets infected by a single infected person is deterministic proportional to the fraction of society $s(t)$ that is still susceptible to the disease. Intuitively, if only a small fraction of society is susceptible to the disease it is unlikely that an infected person meets a susceptible person. The mass of healthy people that become infected during dt thus equals

$$\beta(t)i(t)s(t)\,,$$

where the transmission rate $\beta(t) \in [\underline{b}, \bar{b}]$ captures both the infectivity of the disease, as well as measures society has taken to influence the speed at which the disease spreads (like social distancing). At rate $\gamma > 0$ infected become non-infected, by either recovering from the disease, or dying of it, such that during a short time span dt, the fraction of infected is reduced by $\gamma i(t)$. The susceptible and infected populations thus for every $t \in [0, \infty)$ evolve according to the following dynamics

$$\begin{aligned} s'(t) &= -\beta(t)i(t)s(t), & s(0) &= s_0, \\ i'(t) &= \beta(t)i(t)s(t) - \gamma i(t), & i(0) &= i_0, \end{aligned} \tag{1}$$

where $s_0, i_0 \in (0, 1)$ are given initial values satisfying $s_0 + i_0 \leq 1$.

Reproduction Rate The *basic reproduction rate* R_0 of the disease equals the number of people a single infected person infects when he is the only person carrying the disease and the transmission rate is maximal

$$R_0 = \frac{\bar{b}}{\gamma}\,.$$

[4]See for example Atkeson (2020); Barro et al. (2020); Dewatripont et al. (2020); Piguillem et al. (2020); Stock (2020); de Walque et al. (2020); Guerrieri et al. (2020).

The *effective reproduction* rate $R(t) = \frac{1}{\gamma}\beta(t)s(t)$ at time t equals the number of people infected by an infected person at time t and depends on the current transmission rate.

Control of the Transmission Rate The time-dependent transmission rate $\beta\colon [0,\infty) \to B$ takes values in an compact interval $B = [\underline{b}, \overline{b}] \subset (0,\infty)$. We denote by $\overline{b}$ the maximal transmission rate and by $\underline{b}$ the minimal transmission rate that can be achieved through policy measures, like social distancing, the closure of schools, wearing masks, etc. The set of admissible controls $\mathcal{B}$ consists of all measurable functions $\beta\colon [0,\infty) \to B$.

We introduce three cost functions $v, \bar{v}\colon [0,1] \to [0,\infty)$ and $c\colon B \to [0,\infty)$, where $v, \bar{v}$ capture the economic and health cost of infected population and c captures the cost of reducing the transmission rate. The planner trades-off the consequences of a higher number of infected with the economic and social cost of reducing the transmission rate and thus aims at minimizing the cost functional

$$J(\beta) = \int_0^T v(i(t)) + c(\beta(t))\, dt + \bar{v}(i(T))\,. \tag{2}$$

A policy β^* is *optimal* if it minimizes J over $\mathcal{B}$

$$\beta^* \in \underset{\beta\in\mathcal{B}}{\arg\min}\; J(\beta)\,. \tag{3}$$

The cost $v(i)$ measures the number of people that die per unit of time if a share $i \in [0,1]$ of the population is infected. We suppose that v is convex, continuously differentiable, and strictly increasing and without loss set $v(0) = 0$. Convexity of v captures that the probability of dying from the disease might be higher if a large share of the population is infected and the hospital system is overwhelmed.[5]

We assume that a vaccine arrives at time $T > 0$ and every susceptible person is vaccinated immediately, but no cure is available at that point in time.[6] As, after the comprehensive vaccination of the population, there are no new infections, the share of infected population evolves according to $i'(t) = -\gamma i(t)$ and is given by $i(t) = i(T)e^{-\gamma(t-T)}$ for $t \geq T$. The share of the population that dies after the arrival of the vaccine thus equals

$$\bar{v}(i(T)) = \int_T^\infty v\left(i(T)e^{-\gamma(t-T)}\right) dt = \int_0^{i(T)} \frac{v(z)}{z\gamma} dz\,. \tag{4}$$

The cost function c captures the economic and social cost of measures taken to reduce the transmission rate. For example if social distancing measures are imposed which require the closure

[5]The cost v can not only capture the people who die of the disease directly, but also those who die because other medical conditions remain untreated as an indirect consequence of the disease.

[6]If a cure also becomes available at time T, the terminal cost equals zero $\bar{v}(i) = 0$ as all infected get cured immediately at time T. We considered this case in the working paper version of this paper.

of most businesses this comes at a substantial economic cost. We only make minimal assumption on c and assume that it is non-increasing, convex, and continuous. Without loss we normalize the cost associated with the highest transmission rate to zero, $c(\bar{b}) = 0 > c(b)$ for all $b \in [\underline{b}, \bar{b})$. The convexity of c captures that some businesses are less costly to close down than others and the planner could start by closing down these "least essential" businesses or impose other low cost measures such as making masks mandatory. The convexity of c creates a incentive for the planner to smooth out lockdown measures over time and it is often instructive to focus on the linear case where such an incentive is absent.

3 Optimal Policies

The next result shows existence of an optimal policy and provides necessary conditions that any solution of the optimal control problem (3) must satisfy.

Proposition 1. *An optimal policy exists. Let $\beta^* \in \mathcal{B}$ be an optimal policy and denote by $s^*, i^* : [0,T] \to [0,1]$ the associated number of susceptible and infected satisfying (1). There exists a function $\eta^* : [0,T] \to \mathbb{R}$ with $\eta^*(T) = \frac{v(i^*(T))}{\gamma i^*(T)}$ such that for almost all $t \in [0,T]$ it holds*

$$(\eta^*)'(t) = \eta^*(t)\beta^*(t)i^*(t) - v'(i^*(t)) + \frac{v(i^*(t)) + c(\beta^*(t)) - \min_{b\in B}\left[\frac{1}{\gamma}v(i^*(T))s^*(T)b + c(b)\right]}{i^*(t)}, \quad (5)$$

$$(\beta^*)(t) \in \arg\min_{b\in B}\left[\eta^*(t)i^*(t)s^*(t)b + c(b)\right]. \quad (6)$$

Moreover, we have $\eta^*(t) > 0$ for all $t \in [0,T]$.

The proof of Proposition 1 relies on a sequence of auxiliary results we establish in the appendix using standard arguments from control theory that can, e.g., be found in Clarke (2013). The existence of an optimal policy follows as the convexity of c and B ensures compactness of the policy space which leads to the existence of an optimal policy. Pontryagin's optimality principle then yields that there exist two Lagrange multipliers λ_1^* - the marginal cost of susceptible population – and λ_2^* – the marginal cost of infected population – such that the optimal control is only a function of these multipliers. As the transmission rate controls how fast susceptible population becomes infected, the optimal control can be determined from the difference in marginal cost $\eta^* = \lambda_2^* - \lambda_1^* > 0$ according to (6).[7] This Lagrange multiplier $\eta^*(t)$ has a clear interpretation as the marginal cost of having an additional susceptible person infected. Proposition 1 goes beyond the Pontryagin maximum principle as it shows that the two Lagrange multipliers λ_1^*, λ_2^* can be summarized in a single Lagrange multiplier η^* whose dynamics can be expressed as a forward-backward differential

[7] A similar observation is made in Behncke (2000).

equation (independent of λ_1^*, λ_2^*) thereby effectively reducing the dimension of the problem by 1. This, simplification of the problem allows us to explicitly characterize features of the optimal policy later. Furthermore, Proposition 1 shows that $\eta^*(t) > 0$ which establishes that the planner always benefits from having fewer infected. While this result seems intuitive it is not at all obvious as we later show that the planner sometimes benefits from a *higher* transmission rate (see Figure 1).

We throughout restrict attention to optimal policies where (6) is satisfied for *all* $t \in [0, T]$. This assumption is inconsequential in the sense that changing the policy β^* on a set of measure zero does not affect the share of susceptible and infected population (s^*, i^*).[8]

3.1 Gain from Reducing the Transmission Rate

An important quantity is the *gain from reducing the transmission rate* at time t along the optimal path which we define as

$$g^*(t) = \eta^*(t) i^*(t) s^*(t) \,.$$

Proposition 1 implies that $g^* : [0, T] \to (0, \infty)$ is strictly greater zero and completely determines the optimal control through (6). To simplify notation we define for all $g \in \mathbb{R}$

$$M(g) = \min_{b \in B} \left[g\, b + c(b) \right].$$

Our next result characterizes how the gain from reducing the transmission rate evolves over time.

Proposition 2 (Properties of the Gain from Reducing the Transmission Rate)**.**

(i) The gain from reducing the transmission rate $g^* : [0, T] \to (0, \infty)$ *evolves according to*

$$(g^*)'(t) = -\gamma g^*(t) + s^*(t) \left(M(g^*(t)) - M(g^*(T)) - [v'(i^*(t)) i^*(t) - v(i^*(t))] \right) \,, \tag{7}$$

with terminal condition $g^*(T) = \frac{1}{\gamma}[v(i^*(T)) s^*(T)]$.

(ii) For all $t \in [0, T)$ *the gain* g^* *satisfies the following bounds*

$$0 < g^*(T) < g^*(t) < \frac{1}{\gamma} \left[v(1) + v'(1) \left(e^{(T-t)\gamma} - 1 \right) \right]. \tag{8}$$

(iii) Any optimal control β^* *is non-increasing in* g^*

$$g^*(t) > g^*(t') \Rightarrow \beta^*(t) \leq \beta^*(t') \,.$$

Equation (7) characterizes how the incentive to reduce the transmission rate evolves over time. The first term $-\gamma g^*(t)$ shows that the incentive is exponentially decaying at the rate at which the

[8] It is a priori not clear that an optimal β^* can be modified such that (6) is satisfied for all $t \in [0, T]$ and such that β^* remains measurable. We proof this in Lemma 13 in the appendix.

infected die or become cured. Intuitively, when there are fewer infected the benefits from reducing the transmission rate is smaller.

The term $M(g^*(t)) - M(g^*(T))$ quantifies how much the objective function of the planner increases as a function of $g^*(t)$. This term is strictly positive, and increasing in $g^*(t)$. Finally, the third term $v'(i^*(t))i^*(t) - v(i^*(t))$ is non-negative, increasing in i^* and a measure of the increase in cost due to the convexity of v.[9] The term implies that the gain from reducing the transmission rate falls more quickly when more people are currently infected.

Part (ii) of the proposition states that the incentive to reduce the transmission rate is minimal at the final time T. Part (iii) establishes that the optimal control is non-increasing in the transmission rate. Together part (ii) and (iii) imply that the optimal control is maximal at the terminal time T.

Corollary 3. *Any optimal control $\beta^* \in \mathcal{B}$ satisfies $\beta^*(t) \leq \beta^*(T)$ for all $t \in [0, T]$.*

Intuitively, as the benefit of having fewer infected is smallest at time T as people which get infected at time T do not cause further infections due to the availability of a vaccine at time T.

3.2 Herd-Immunity

We denote by τ the first time society achieves *herd-immunity*, that is the first time every infected person infects fewer than one person on average without intervention of the planner[10]

$$\tau = \min \left\{ t \in [0, T] \colon \frac{\bar{b}}{\gamma} s(t) \leq 1 \right\} .$$

It is immediate from the definition of τ that the share of infected population is strictly decreasing after τ, independent of the policy chosen after time τ. Our next result shows that the optimal policy always implies a non-decreasing transmission rate after τ, i.e. that lockdown measures get less stringent after herd-immunity has been achieved in society.

Proposition 4. *Fix an optimal policy β^*. The gain from reducing the transmission rate g^* is strictly decreasing on $[\tau, T]$ and the optimal transmission rate β^* is non-decreasing on $[\tau, T]$.*

Intuitively, after herd-immunity has been achieved the incentives to reduce the transmission rate are decreasing as each infected has a smaller chance of infecting someone in the future as over time one gets closer to the arrival of a vaccine, and the share of susceptible population decreases.

For any disease with basic reproduction rate R_0 below 1 herd immunity is achieved at time zero and an immediate corollary of Proposition 4 is that any optimal policy reduces the lockdown

[9] For example suppose that the percentage infected who dies grows linearly in the fraction of the population that is infected $v(i) = ki^2$. In this case $v'(i^*(t))i^*(t) - v(i^*(t)) = ki^2$.

[10] Throughout the article we use the conventions that $\inf \emptyset = \min \emptyset = \infty$ and $\sup \emptyset = \max \emptyset = -\infty$

measures over time. Especially, in the case of linear costs, i.e. there exists $\delta > 0$ such that $c(\beta) = \delta(\bar{b} - \beta)$, this implies that the planner optimally engages in the maximal lockdown until some time t^* and afterwards uses no lockdown measures at all.

Corollary 5. *Assume that $R_0 \leq 1$ and let β^* be an optimal policy. Then, β^* is non-decreasing on $[0, T]$ and if c is linear then there exists $t^* \in [0, T]$ such that*[11]

$$\beta^*(t) = \begin{cases} \underline{b} & \text{for } t \in [0, t^*) \\ \bar{b} & \text{for } t \in (t^*, T] \end{cases} . \tag{9}$$

Less Social Distancing Leads to Faster Herd-Immunity As we argue next, the fact that for $R_0 \leq 1$ the optimal policy engages in a lockdown as quickly as possible does not generalize to diseases with $R_0 > 1$. Intuitively, for such diseases there is an additional motive at play for the planner before herd immunity is reached. By increasing the transmission rate the planner generates more infections, which lead to more people who are immune to the disease, and faster herd-immunity.

Consider two policies β and $\hat{\beta}$ and their associated paths of susceptible and infected (s^β, i^β) and $(s^{\hat{\beta}}, i^{\hat{\beta}})$. We say that β induces *less social distancing* than $\hat{\beta}$ if for all $t, t' \in [0, T]$ with $s^\beta(t) = s^{\hat{\beta}}(t')$ we have $\beta(t) \geq \hat{\beta}(t')$. Our definition of less social distancing requires that for any given share of susceptible population the transmission rate is higher if the planner engages in less social distancing.

Proposition 6. *If β induces less social distancing than $\hat{\beta}$ then there will be fewer susceptible $s^\beta(t) \leq s^{\hat{\beta}}(t)$ at every point in time $t \in [0, T]$ and herd-immunity will be reached faster under the policy β.*

Proposition 6 shows that less social distancing can lead to faster herd-immunity. A natural conjecture following this observation is that too much social distancing can *increase* the number of people who die from a disease. Figure 1 provides such an example. The example considers a disease with a basic reproduction rate of 2.5, an average infection length of 12 days, where social distancing reduces the transmission rate by 65%, and 1% of infected die. These parameters are chosen in line with estimates for the Covid-19 pandemic in 2020.[12] A vaccine is assumed to arrive after 360 days. Starting with 0.1% of the population infected at time zero it compares the number of infected if social distancing is enacted from day $0 - 100$ to social distancing from day $50 - 100$.

[11] Here and in the sequel we use the convention that $[x, x) = (x, x] = (x, x) = \emptyset$ for all $x \in \mathbb{R}$. In particular, any of the intervals in (9) might be empty.

[12] Wu et al. (2020) estimate R_0 for Covid-19 to be between 2.47 and 2.86 or between 2.32 and 2.71 depending on the modelling assumptions. For an average infection length of 10 days Toda (2020) finds a median reproduction rate of 2.9 and documents significant variation across countries. Kantner (2020) calibrate an SIR model to Germany and find a basic reproduction rate of $R_0 = 2.7$.

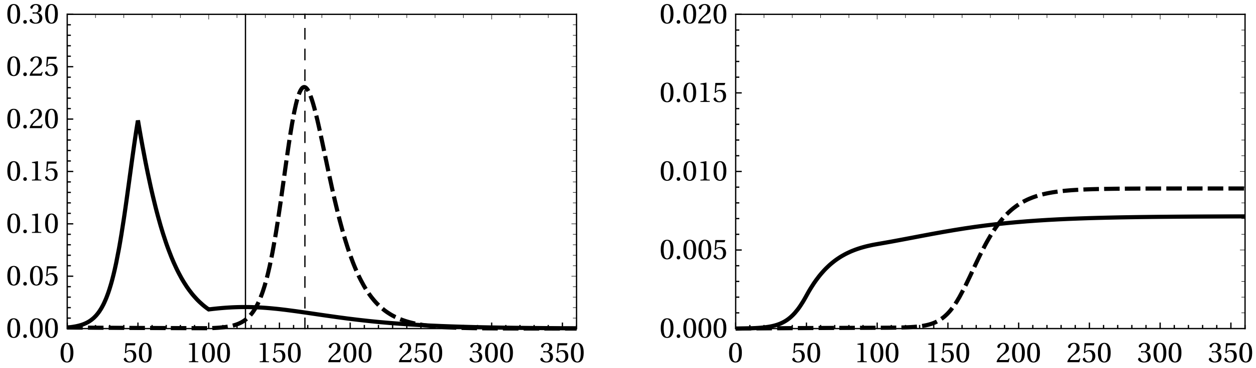

Figure 1: On the left there is the fraction of infected over time for maximal social distancing from day 50-100 (solid line) and from day 0-100 (dashed line). The vertical lines mark the time of herd-immunity. On the right the fraction of dead population over time for maximal social distancing from day 50-100 (solid line) and from day 0-100 (dashed line). The parameters are $R_0 = 2.5, \gamma = 1/12, \bar{b} = R_0\,\gamma, \underline{b} = 0.35\bar{b}, v(i) = 0.01\gamma i, T = 360$.

On the left one can see how the number of infected evolves over time under either policy. If there is no social distancing initially the number of infected peaks early and then decreases until day 100, after which it mildly increases until herd immunity is reached at day 126, after which the number of infected decrease. If social distancing is enacted from day 0-100 the number of infected peaks much later and higher at day 168 at which herd-immunity is reached. The right-hand-side plot illustrates the accumulated share of people who died. While with social distancing more people die early from the disease, fewer people die later on. Overall, the share of the population that dies from the disease is 0.177% less when social distancing is enacted from day 50-100 in this example. Notably, this example does not assume convex v, which corresponds to limited hospital capacity, and would amplify the difference in dead due to the different height of the peak number of infected.

3.3 Limited Capacity Effects

We next introduce an assumption that restricts to environments where the cost is not too convex in the share of infected population.[13]

Assumption 1 (Limited Capacity Effects)**.** *The cost v is twice continuously differentiable and satisfies* $\frac{v''(i)}{\gamma} < \frac{27}{4}\,\frac{\underline{b}}{\bar{b}-\underline{b}}\left|c'(\bar{b}-)\right|$ *for all* $i \in [0,1]$.

Assumption 1 admits an intuitive interpretation: As $\frac{1}{\gamma}v'(i)$ is the probability with which an additional infected dies when a share i of the population is infected. The condition thus requires that the probability of dying from the disease can not be too sensitive to the share of infected

[13] We denote by $c'(\bar{b}-) = \lim_{b \nearrow \bar{b}} \frac{c(\bar{b})-c(b)}{\bar{b}-b} \leq 0$ the left-derivative of c at $\bar{b}$ (which exist due to convexity of c).

population, compared to the marginal cost of reducing the transmission rate. It is always satisfied if the probability of dying from the disease is independent of which fraction of the population is infected, which corresponds to the case where v is linear. The main reason for non-linear v is that the death rate increases for larger number of infected due to limited hospital capacity.

This assumption thus rules out too large capacity effects that arise from the overload of the medical system. It is thus a reasonable approximation if the number of infected is kept within levels that do not overburden the healthcare system, or if the death rate does not increase too quickly when the number of infected exceeds the hospital capacity.

Our next result shows that in this case the optimal transmission rate is quasi-convex.

Proposition 7. *Suppose that Assumption 1 is satisfied and let $\beta^* \in \mathcal{B}$ be an optimal control. Then there exists $t^* \in [0, T]$ such that*

(i) β^* is non-increasing on $[0, t^*]$ and non-decreasing on $[t^*, T]$, i.e., β^* is quasi-convex, and

(ii) the fraction of infected population i^* is strictly log-concave on $[0, t^*]$ and thus admits at most one local maximum on $[0, t^*]$.

Proposition 7 establishes that any optimal policy is single peaked, in the sense that the measures to decrease the transmission rate are first escalated until some point in time and then reduced over time. Any policy where a reduction in measures is followed by an increase is suboptimal. Furthermore, in the initial period where SD measures are escalated and the transmission rate falls the fraction of infected is single peaked. The proof of Proposition 7 is quite involved. The main idea of the proof is to differentiate the dynamics of the gain from reducing the transmission rate derived in (7) and then bounding the second derivative of g^* using Assumption 1 to establish that g^* is quasi-concave.

We next show that if Assumption 1 is satisfied and c is linear then the optimal policy involves only the two most extreme controls $\beta^*(t) \in \{\underline{b}, \bar{b}\}$. The assumption that the cost c of measures that reduce the transmission rate is linear has a simple interpretation in the context of social distancing: Shutting down half of the economy for two days is equally costly as shutting down the whole economy for a single day.[14] While there is no normative reason for this assumption we think of it as a natural baseline for the analysis, and potentially a good approximation of the trade-offs once "cheap measures" to reduce the transmission rate (such as wearing masks) are exhausted.[15]

Proposition 8. *Suppose that Assumption 1 is satisfied and c is linear and let $\beta^* \in \mathcal{B}$ be an optimal control. Then there exist $0 \le t_1^* \le t_2^* \le T$ such that*

[14] This implicitly assumes that the transmission rate β depends linearly on the shut down of the economy.

[15] We note here that the theoretical literature on optimal policies to manage an infectious disease has almost exclusively focused on the case of linear cost.

(i) for a.e. $t \in [0,T]$ we have

$$\beta^*(t) = \begin{cases} \bar{b} & \text{for } t \in [0, t_1^*) \\ \underline{b} & \text{for } t \in (t_1^*, t_2^*) \\ \bar{b} & \text{for } t \in (t_2^*, T] \end{cases} \tag{10}$$

(ii) and the fraction of infected i^* under the policy β^* is strictly increasing on $[0, t_1^*]$, and has at most one local maximum in $[0, t_2^*]$, and at most one local maximum in $[t_2^*, T]$.

Proposition 8 drastically simplifies the search for an optimal policy as it implies that any optimal policy is characterized by the two points in time (t_1^*, t_2^*). Note that the proposition does not rule out that any of the intervals is empty. In particular, reducing the transmission rate by the maximal amount at every point in time as well as taking no measures at all to reduce the transmission rate can be optimal.

The first economic insight of Proposition 8 is that when the cost of reducing the transmission rate is linear and the hospital capacity does not have a too strong effect on the death rate it is never optimal to use intermediate measures for a longer time (i.e. closing only parts of the economy) as doing so is dominated by implementing maximal measures for a shorter time. The second part of the proposition establishes that under the optimal policy the number of infected peaks at most twice. The second peak always happens after the planner has stopped using lockdown measures. Thus, whenever one observes the number of infected peak more than twice one can conclude that either the planner acted suboptimally or one of the assumptions of the model must be violated.

3.4 The Optimal Timing of Social Distancing

We next illustrate the optimal timing of social distancing. For this illustration we again consider a disease with a basic reproduction rate of 2.5, an average infection length of 12 days, where social distancing reduces the transmission rate by 65%, and 1% of infected die.[16] A vaccine is assumed to arrive after 360 days. We consider linear cost of reducing the transmission rate for which Proposition 8 implies that the optimal policy consist of three subsequent periods, first no social distancing, followed by a period of social distancing, and a period of no social distancing.

Without loss we suppose that the planner has a fixed budget of days of social distancing and plot during which time period he optimally engages in social distancing. This allows us to display the optimal policy for various costs. For example, consider the case where the planner has a budget of 100 days of social distancing. In this case the optimal policy is to start social distancing on day 50 and end it on day 150. As one can see in the left graph of Figure 2 this leads to a flatter curve of infected over time than social distancing in the first 100 days or no social distancing. Interestingly,

[16] These parameter values are chosen in line with the Covid-19 pandemic in 2020 c.f. Footnote 12.

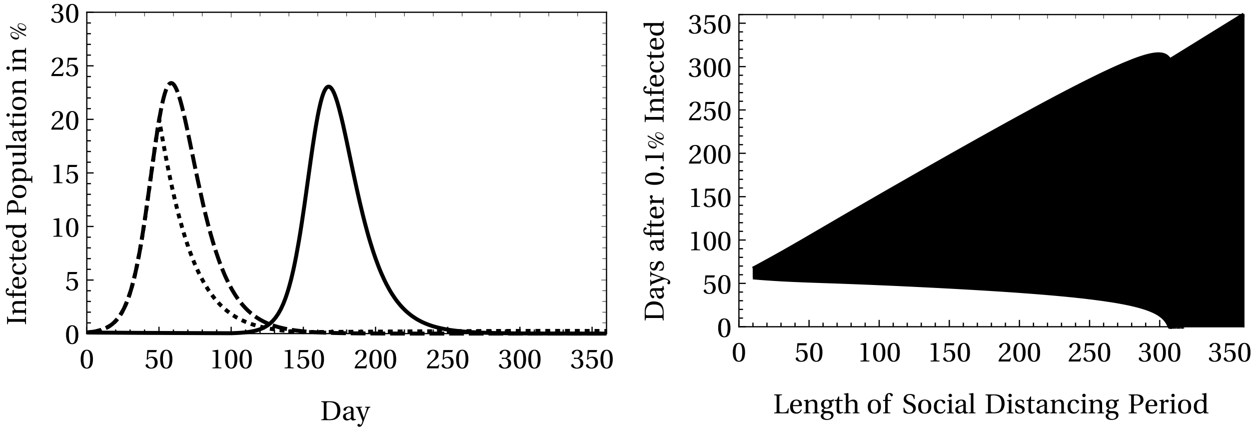

Figure 2: On the left: The number of infected over time without SD (dashed line), with SD from day 0-100 (solid line), and for the optimal SD period from day 50 to day 150 (dotted line). On the right: the optimal period of SD as a function of the length of SD.

the effect of suboptimal social distancing is marginal in the sense that while it initially reduces the number of infected substantially, it essentially only delays the peak of infected, but does not substantially flatten it. Optimal social distancing leads to a substantial reduction in the implied death rate within a year: 0.607% under optimal social distancing, 0.892% with social distancing in the first hundred days, and 0.893% without social distancing.

The right graph of Figure 2 depicts the optimal timing of social distancing as a function of the length of social distancing. As one can see it is often optimal to delay social distancing beyond the date where 0.1% of the population is infected. For example even if it is optimal for the planner to engage in 150 days of social distancing within the next year it is only optimal to start social distancing after 50 days. This observation might be surprising as it implies that if it is not optimal to maintain permanent social distancing (until the arrival of a vaccine/cure), then it is optimal to substantially delay the period of social distancing.

The Value of Optimal Social Distancing We next illustrate the importance of the optimal timing of social distancing. The left graph of Figure 3 plots the death rate as a function of the number of days of social distancing.[17] The solid line depicts the death rate when the social distancing measures are optimally timed and the dashed line depicts the death rate if the planner engages in social distancing immediately. As one can see in the figure social distancing can be an effective measure to prevent the death of population. For example, 50 days of optimally timed social distancing (from day 50 to day 100) reduce the death rate by roughly 0.2%. The figure however shows that without optimal timing social distancing is much less effective and to achieve an equal reduction in the example one needs 270 days of social distancing.

[17]We used the same parameters as in Figure 1 and 2.

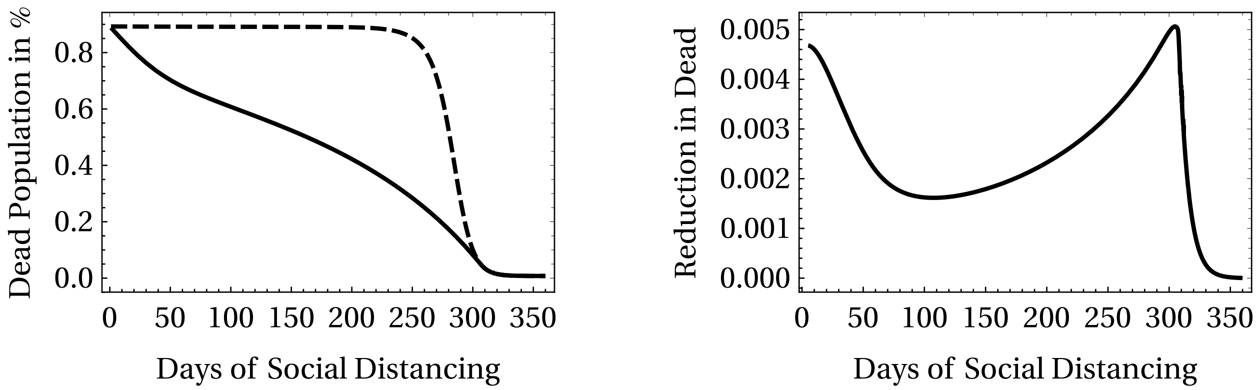

Figure 3: On the Left: Fraction of the population that dies within 360 days of 0.1% infected as a function of the length of social distancing for the optimal timing of social distancing (solid line) and social distancing starting at day 0 (dashed line). On the right: the benefit of an additional day of social distancing.

The right graph of Figure 3 plots the reduction in cost from engaging in an additional day of optimally timed social distancing. As one can see in the picture the benefit is non-concave. It is initially high and the first days of social distancing safe around 0.005% of the population. It then decreases and peaks again around 300 days of social distancing. Intuitively, this non-concavity results from the fact that engaging in social distancing over almost the whole time until a vaccine arrives can prevent the disease from ever substantially spreading. In contrast engaging in social distancing over a shorter time horizon can not avoid a substantial spread of the disease before the arrival of a vaccine.

4 Conclusion

In this short note we derived the optimal policy for social distancing during an epidemic. Our analysis revealed several features of the optimal policy. If the death rate is not too sensitive to the number of infected, the optimal policy consists of two phases: a first phase where the planner engages in more social distancing over time and a second phase where the planner engages in less social distancing over time. Furthermore, if the cost of reducing the transmission rate is linear, the optimal policy is always extreme. At any point in time either social distancing is carried out to the maximal extend possible or not at all. The intuitive reason for this result is that more extreme measures over a shorter time horizon are more effective than less extreme measures over a longer horizon. We illustrated through an example that the effectiveness of social distancing depends crucially on its optimal timing. Within the context of this example optimal social distancing is often substantially delayed in order to generate herd immunity. Engaging in more, but too early social distancing can increase the peak number of infected and the fatalities caused by the disease.

A Appendix

We first prove a sequence of auxiliary results which together imply Proposition 1. Recall that the function $M\colon \mathbb{R} \to \mathbb{R}$ is defined by $M(y) = \min_{b\in B}[yb + c(b)]$, $y \in \mathbb{R}$, and note that M is non-decreasing.

Lemma 9. *An optimal policy β^* exists that solves (3). Let $s^*, i^*\colon [0,T] \to [0,1]$ be the state processes associated with an optimal control satisfying (1). Then there exist absolutely continuous functions $\lambda_1^*, \lambda_2^*\colon [0,T] \to \mathbb{R}$ which satisfy for almost all $t \in [0,T]$ the dynamics*

$$\begin{aligned} (\lambda_1^*)'(t) &= (\lambda_1^*(t) - \lambda_2^*(t))\beta^*(t)i^*(t), & \lambda_1^*(T) &= 0, \\ (\lambda_2^*)'(t) &= (\lambda_1^*(t) - \lambda_2^*(t))\beta^*(t)s^*(t) + \gamma\lambda_2^*(t) - v'(i^*(t)), & \lambda_2^*(T) &= \frac{v(i^*(T))}{\gamma i^*(T)}, \end{aligned} \tag{11}$$

and the optimality condition

$$\beta^*(t) \in \operatorname*{arg\,min}_{b\in B} \left[(\lambda_2^*(t) - \lambda_1^*(t))i^*(t)s^*(t)b + c(b)\right]. \tag{12}$$

Moreover, for all $t \in [0,T]$ we have

$$M\big(\lambda_2^*(t) - \lambda_1^*(t))i^*(t)s^*(t)\big) - \gamma\lambda_2^*(t)i^*(t) + v(i^*(t)) = M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right). \tag{13}$$

Proof of Lemma 9. Suppose an optimal policy β^* exists. The existence of functions λ_1^*, λ_2^* : $[0,T] \to \mathbb{R}$ that satisfy (11), (12) and (13) follows from the Pontryagin principle (see, e.g., Clarke, 2013, Theorem 22.2 and Corollary 22.3). We show the existence of an optimal policy by verifying the conditions of Theorem 23.11 in Clarke (2013).

(a) $g(t,(s,i)) = \begin{pmatrix} -is \\ +is \end{pmatrix}$ which implies that $|g(t,(s,i))| \le 2|is| < 2$.

(b) $B = [\underline{b}, \overline{b}]$ is closed and convex by definition.

(c) The sets $E = \{(s_0, i_0)\} \times \mathbb{R}_+$ and $Q = [0,T] \times [0,1]^2$ are closed and $\ell(s_0, i_0, s_T, i_T) = \bar{v}(i_T)$ is lower semicontinuous.

(d) The running cost $\beta \mapsto v(i) + c(\beta)$ is convex as c is convex. Furthermore, $v(i) + c(\beta) \ge 0$.

(e) The projection set is given by $\{(s_0, i_0)\}$ and thus bounded.

(f) As $\beta \in B$ it follows that $|\beta| \le \overline{b}$. This verifies (f) (ii).

Moreover, the constant control $\beta(t) = \overline{b}$ has finite costs. We have hence verified that there exists an optimal policy. □

In the following we suppose that $\beta^* \in \mathcal{B}$ is an optimal control and denote by $s^*, i^*\colon [0,T] \to [0,1]$ the associated state processes satisfying (1). Moreover, we denote by $\lambda_1^*, \lambda_2^*\colon [0,T] \to \mathbb{R}$

the Lagrange variables from Lemma 9. Compactness of B and continuity of c ensure that for all $t \in [0,T]$ the function $b \mapsto [\lambda_2^*(t) - \lambda_1^*(t)]i^*(t)s^*(t)b + c(b)$ attains its minimum on B.

We introduce the new Lagrange variable

$$\eta^*(t) = \lambda_2^*(t) - \lambda_1^*(t)\,. \tag{14}$$

The variable $\eta^*(t)$ has a clear interpretation: it measures the marginal change in the cost with respect to infecting susceptible population. Intuitively speaking, $\eta^*(t)$ measures the additional cost if one additional person is infected at time t given the optimal policy is used. Note that by (12) at each time t the optimal control $\beta^*(t)$ depends on $\lambda_1^*(t)$ and $\lambda_2^*(t)$ only through their difference $\eta^*(t) = \lambda_2^*(t) - \lambda_1^*(t)$.

Lemma 10. *Let $\beta^* \in \mathcal{B}$ be an optimal control and suppose that the optimality condition (12) holds for all $t \in [0,T]$. Suppose that $t_0 \in [0,T]$ satisfies $\eta^*(t_0) \leq 0$. Then it holds that $\lim_{t\to t_0} \beta^*(t) = \beta^*(t_0) = \bar{b}$.*

Proof of Lemma 10. First note that the assumption $\eta^*(t_0) \leq 0$ ensures that the function $b \mapsto \eta^*(t)i^*(t)s^*(t)b + c(b)$ attains its global minimum on B at $\bar{b}$. Hence (12) implies that $\beta^*(t_0) = \bar{b}$. Next let (t_n) be a sequence such that $t_n \to t_0$ as $n \to \infty$. Suppose by contradiction that there exists a subsequence such that $\lim_{n\to\infty} \beta^*(t_n) =: b_0 < \bar{b}$. Next note that (12) ensures for all $n \in \mathbb{N}$ that (recall that $c(\bar{b}) = 0$)

$$\eta^*(t_n)i^*(t_n)s^*(t_n)\beta^*(t_n) + c(\beta^*(t_n)) \leq \eta^*(t_n)i^*(t_n)s^*(t_n)\bar{b}. \tag{15}$$

This implies that

$$\eta^*(t_n)i^*(t_n)s^*(t_n) \geq \frac{c(\beta^*(t_n))}{\bar{b} - \beta^*(t_n)}. \tag{16}$$

Taking the limit $n \to \infty$ yields the contradiction

$$0 \geq \lim_{n\to\infty} \eta^*(t_n)i^*(t_n)s^*(t_n) \geq \lim_{n\to\infty} \frac{c(\beta^*(t_n))}{\bar{b} - \beta^*(t_n)} = \frac{c(b_0)}{\bar{b} - b_0} > 0. \tag{17}$$

Therefore, we have $\lim_{t\to t_0} \beta^*(t) = \bar{b} = \beta^*(t_0)$. $\square$

The next result shows that the cost of additional infected η^* is characterized by a differential equation that does not depend on λ_1^* and λ_2^*. Moreover, we show that both λ_1^* and λ_2^* can be recovered from η^*.

Lemma 11. *The variable η^* solves*

$$(\eta^*)'(t) = \eta^*(t)\beta^*(t)i^*(t) - v'(i^*(t)) + \frac{v(i^*(t)) + c(\beta^*(t)) - M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right)}{i^*(t)} \tag{18}$$

with terminal condition $\eta^*(T) = \frac{v(i^*(T))}{\gamma i^*(T)}$. *Conversely, suppose that* $i, s, \beta, \eta\colon [0,T] \to \mathbb{R}$ *satisfy*

$$\begin{aligned}
&s'(t) = -\beta(t)i(t)s(t), \quad s(0) = s_0,\\
&i'(t) = \beta(t)i(t)s(t) - \gamma i(t), \quad i(0) = i_0,\\
&\eta'(t) = \eta(t)\beta(t)i(t) - v'(i(t)) + \frac{v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)}{i(t)}, \quad \eta(T) = \frac{v(i(T))}{\gamma i(T)},\\
&\beta(t) \in \operatorname*{arg\,min}_{b\in B}\left[\eta(t)i(t)s(t)b + c(b)\right],
\end{aligned} \tag{19}$$

then

$$\begin{aligned}
\lambda_1(t) &= \frac{1}{\gamma}\left[\eta(t)\beta(t)s(t) + \frac{v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)}{i(t)}\right] - \eta(t),\\
\lambda_2(t) &= \frac{1}{\gamma}\left[\eta(t)\beta(t)s(t) + \frac{v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)}{i(t)}\right]
\end{aligned} \tag{20}$$

solves (11).

Proof of Lemma 11. First note that it follows from (13) and (12) that

$$\begin{aligned}
-\eta^*(t)\beta^*(t)i^*(t)s^*(t) + \gamma\lambda_2^*(t)i^*(t) &= -M(\eta^*(t)i^*(t)s^*(t)) + \gamma\lambda_2^*(t)i^*(t) + c(\beta^*(t))\\
&= v(i^*(t)) + c(\beta^*(t)) - M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right).
\end{aligned} \tag{21}$$

Then (11) implies that

$$\begin{aligned}
(\eta^*)'(t) = (\lambda_2^*)'(t) - (\lambda_1^*)'(t) &= -\eta^*(t)\beta^*(t)s^*(t) + \gamma\lambda_2^*(t) - v'(i^*(t)) + \eta^*(t)\beta^*(t)i^*(t)\\
&= \eta^*(t)\beta^*(t)i^*(t) - v'(i^*(t)) + \frac{v(i^*(t)) + c(\beta^*(t)) - M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right)}{i^*(t)}.
\end{aligned} \tag{22}$$

Next suppose that s, i and η solve (19) and that λ_1 and λ_2 are given by (20). Observe that it holds that $\lambda_1(T) = 0$ and $\lambda_2(T) = \frac{v(i(T))}{\gamma i(T)}$. Next note that the envelope theorem ensures that

$$\frac{\partial}{\partial t}\left[\eta(t)i(t)s(t)\beta(t) + c(\beta(t))\right] = \frac{\partial}{\partial t}\min_{b\in B}\left[\eta(t)i(t)s(t)b + c(b)\right] = \beta(t)\frac{\partial}{\partial t}\left[\eta(t)i(t)s(t)\right]. \tag{23}$$

Then it holds that

$$
\begin{aligned}
\lambda_2'(t) &= \frac{\partial}{\partial t}\left[\frac{\eta(t)\beta(t)i(t)s(t) + v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)}{\gamma i(t)}\right] \\
&= \frac{\beta(t)\frac{\partial}{\partial t}\left[\eta(t)i(t)s(t)\right] + v'(i(t))i'(t)}{\gamma i(t)} \\
&\quad - \frac{\left(\eta(t)\beta(t)i(t)s(t) + v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)\right)i'(t)}{\gamma (i(t))^2} \\
&= \frac{1}{\gamma}(\beta(t)\eta'(t)s(t) + \beta(t)\eta(t)s'(t)) \\
&\quad + \frac{i'(t)}{\gamma i(t)}\left(v'(i(t)) - \frac{v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)}{i(t)}\right) \\
&= \frac{1}{\gamma}(\beta(t)\eta'(t)s(t) + \beta(t)\eta(t)s'(t)) + \frac{i'(t)}{\gamma i(t)}\left(\eta(t)\beta(t)i(t) - \eta'(t)\right) \\
&= \frac{\beta(t)\eta(t)}{\gamma}(s'(t) + i'(t)) + \frac{\eta'(t)}{\gamma}\left(\beta(t)s(t) - \frac{i'(t)}{i(t)}\right) \\
&= -\beta(t)\eta(t)i(t) + \eta'(t).
\end{aligned}
$$

Therefore we obtain that

$$
\begin{aligned}
\lambda_2'(t) &= \frac{v(i(t)) + c(\beta(t)) - M\left(\frac{v(i(T))s(T)}{\gamma}\right)}{i(t)} - v'(i(t)) = \gamma\lambda_2(t) - \eta(t)\beta(t)s(t) - v'(i(t)) \\
&= \left(\lambda_1(t) - \lambda_2(t)\right)\beta(t)s(t) + \gamma\lambda_2(t) - v'(i(t))\,.
\end{aligned}
$$

Similarly, λ_1 satisfies

$$
\lambda_1'(t) = \lambda_2'(t) - \eta'(t) = [\eta'(t) - \eta(t)\beta(t)i(t)] - \eta'(t) = (\lambda_1(t) - \lambda_2(t))\beta(t)i(t)\,. \qquad \square
$$

Lemma 12 (More Infected are Costly)**.** *The function η^* satisfies $\eta^*(t) > 0$ for all $t \in [0, T)$.*

Proof of Lemma 12. Suppose that there exists $t \in [0, T)$ such that $\eta^*(t) \le 0$. Then Lemma 10 shows that $\beta^*(t) = \bar{b}$. Moreover, Lemma 10 ensures that β^* is continuous at t and hence η^* is differentiable at t. Then (18) shows (recall that $c(\bar{b}) = 0$)

$$
(\eta^*)'(t) = \eta^*(t)\beta^*(t)i^*(t) + \frac{v(i^*(t)) - M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right) - v'(i^*(t))i^*(t)}{i^*(t)}. \tag{24}
$$

Since v is convex and since $M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right) > 0$ we thus obtain that

$$(\eta^*)'(t) < \eta^*(t)\beta^*(t)i^*(t) \leq 0 \tag{25}$$

We conclude from the terminal condition $\eta^*(T) = \frac{v(i^*(T))}{\gamma i^*(T)} > 0$ that $\eta^*(t) > 0$ for all $t \in [0, T)$. $\square$

Since $(\lambda_1^*)'(t) = -\eta^*(t)\beta^*(t)i^*(t)$ we obtain from Lemma 12 that λ_1^* is decreasing in time and, in particular, that λ_1^* is non-negative. This means that a marginal increase of the susceptible population (while keeping the infected population constant) marginally increases the costs. This marginal effect decreases over time and vanishes at time T.

Combining the results of Lemma 9, Lemma 10, Lemma 11 and Lemma 12 proves Proposition 1.

Proof of Proposition 2. Recall that $g^*(t) = \eta^*(t)i^*(t)s^*(t)$, $t \in [0, T]$. We compute $(g^*)'(t)$

$$\begin{aligned}
(g^*)'(t) &= (\eta^*)'(t)i^*(t)s^*(t) + (\eta^*(t))(i^*)'(t)s^*(t) + \eta^*(t)i^*(t)(s^*)'(t) \\
&= [(\lambda_1^*(t) - \lambda_2^*(t))\beta^*(t)s^*(t) + \lambda_2^*(t)\gamma - v'(i^*(t)) - (\lambda_1^*(t) - \lambda_2^*(t))\beta^*(t)i^*(t)]i^*(t)s^*(t) \\
&\quad + (\lambda_2^*(t) - \lambda_1^*(t))s^*(t)[\beta^*(t)i^*(t)s^*(t) - \gamma i^*(t)] \\
&\quad - (\lambda_2^*(t) - \lambda_1^*(t))i^*(t)\beta^*(t)i^*(t)s^*(t) \\
&= (\lambda_1^*(t) - \lambda_2^*(t))\beta^*(t)[-(s^*(t))^2 i^*(t) + (i^*(t))^2 s^*(t) + (s^*(t))^2 i^*(t) - (i^*(t))^2 s^*(t)] \\
&\quad + [\gamma\lambda_2^*(t) - v'(i^*(t))]i^*(t)s^*(t) - \gamma(\lambda_2^*(t) - \lambda_1^*(t))s^*(t)i^*(t) \\
&= [\gamma\lambda_1^*(t) - v'(i^*(t))]i^*(t)s^*(t).
\end{aligned} \tag{26}$$

This and the fact that $g^*(t) = (\lambda_2^*(t) - \lambda_1^*(t))i^*(t)s^*(t)$ imply

$$(g^*)'(t) = \gamma\lambda_1^*(t)i^*(t)s^*(t) - v'(i^*(t))i^*(t)s^*(t) = -\gamma g^*(t) + \gamma\lambda_2^*(t)i^*(t)s^*(t) - v'(i^*(t))i^*(t)s^*(t).$$

Next (13) implies that

$$(g^*)'(t) = -\gamma g^*(t) + s^*(t)\left[M(g^*(t)) - M\left(\frac{v(i^*(T))s^*(T)}{\gamma}\right) + v(i^*(t)) - v'(i^*(t))i^*(t)\right].$$

Using that $g^*(T) = \eta^*(T)i^*(T)s^*(T) = \frac{v(i^*(T))s^*(T)}{\gamma}$ we obtain

$$(g^*)'(t) = -\gamma g^*(t) + s^*(t)\left[M(g^*(t)) - M(g^*(T)) - [v'(i^*(t))i^*(t) - v(i^*(t))]\right].$$

We next argue that we have that $g^*(t) > g^*(T) > 0$ for all $t \in [0, T)$. Let $t \in [0, T)$ and suppose that $g^*(t) \leq g^*(T)$. Since M is non-decreasing we have $M(g^*(t)) \leq M(g^*(T))$. Convexity of v ensures that $v'(i^*(t))i^*(t) - v(i^*(t)) \geq 0$. Moreover, by Lemma 12 we have $g^*(t) > 0$ and consequently (7) implies that $(g^*)'(t) < 0$. This contradicts continuity of g^* and therefore proves $g^*(t) > g^*(T) > 0$

for all $t \in [0, T)$.

Next, we verify the upper bound in (8). To this end let $Q\colon [0,1] \to \mathbb{R}$ satisfy $Q(i) = v'(i)i - v(i)$. Note that convexity of v implies that Q is non-decreasing. Note that (7), monotonicity of M and the fact that $g^*(t) > g^*(T)$ ensure that

$$\begin{aligned}(g^*)'(t) &= -\gamma g^*(t) + s^*(t)\left(M(g^*(t)) - M(g^*(T)) - [v'(i^*(t))i^*(t) - v(i^*(t))]\right) \\ &> -\gamma g^*(t) - s^*(t)Q(i^*(t)) \geq -\gamma g^*(t) - Q(1)\,.\end{aligned}$$

Gronwall's lemma shows that

$$g^*(t) < e^{(T-t)\gamma} g^*(T) + \frac{Q(1)}{\gamma}\left(e^{(T-t)\gamma} - 1\right)\,.$$

Using that $g^*(T) = \frac{v(i^*(T))s^*(T)}{\gamma} \leq \frac{v(1)}{\gamma}$ and $Q(1) = v'(1)1 - v(1)$ we get

$$g(t) < \frac{1}{\gamma}\left[e^{(T-t)\gamma} v(i^*(T))s^*(T) + (v'(1) - v(1))\left(e^{(T-t)\gamma} - 1\right)\right] \leq \frac{1}{\gamma}\left[v(1) + v'(1)\left(e^{(T-t)\gamma} - 1\right)\right]\,. \tag{27}$$

Finally, take two points in time $t, t' \in [0, T]$. Then (12) shows that

$$g^*(t)\beta^*(t) + c(\beta^*(t)) \leq g^*(t)\beta^*(t') + c(\beta^*(t')) \text{ and } g^*(t')\beta^*(t') + c(\beta^*(t')) \leq g^*(t')\beta^*(t) + c(\beta^*(t)). \tag{28}$$

Adding these two inequalities yields that

$$(g^*(t) - g^*(t'))(\beta^*(t) - \beta^*(t')) \leq 0 \tag{29}$$

Thus, $g^*(t) > g^*(t')$ implies $\beta^*(t) \leq \beta^*(t')$. □

By (12) we have that $\beta^*(t) \in \arg\min_{b \in B} g^*(t)b - c(b)$ for almost all $t \in [0, T]$. By potentially changing β^* on a set of measure zero we suppose in the sequel that $\beta^*(t)$ attains the minimum for all $t \in [0, T]$ (i.e., (12) holds for all $t \in [0, T]$). Note that this change does not affect the trajectories of s^*, i^*, λ_1^* and λ_2^*. In Lemma 13 below we show that this modification preserves the measurability of β^*.

Lemma 13. *There exists an optimal policy such that* $\beta^*(t) \in \arg\min_{b \in B}[g^*(t)b + c(b)]$ *for all* $t \in [0, T]$.

Proof. Let β^* be an arbitrary optimal policy and g^* the associated gain from reducing the transmission rate. Let $S\colon \mathbb{R} \to [\underline{b}, \overline{b}]$ be a function that satisfies for all $y \in \mathbb{R}$ that $M(y) = yS(y) + c(S(y))$. Such a function S exists by the axiom of choice and since $b \mapsto yb + c(b)$ is continuous and hence attains a minimum on $[\underline{b},\bar{b}]$. Then we can show as in Proposition 2 that S is non-increasing. Hence any such S is measurable. Now let β^* satisfy $\beta^*(t) \in \arg\min_{b\in B}[g^*(t)b + c(b)]$ for all $t \in \mathbb{T}$ where $\mathbb{T}$ is a measurable set of full Lebesgue measure. Then define $\hat{\beta}$ for all $t \in [0,T]$ by

$$\hat{\beta}(t) = S(g^*(t))1_{\mathbb{T}^c}(t) + \beta^*(t)1_{\mathbb{T}}(t).$$

Then $\hat{\beta}(t) \in \arg\min_{b\in B}[g^*(t)b + c(b)]$ for all $t \in [0,T]$, $\hat{\beta}$ is measurable as S is measurable and g^* is continuous, and $\hat{\beta}(t) = \beta^*(t)$ for almost all $t \in [0,T]$. □

Proof of Corollary 3. The statement follows directly from parts (i) and (ii) of Proposition 2. □

Proof of Proposition 4. Let $Q\colon [0,1] \to \mathbb{R}$ satisfy $Q(i) = v'(i)i - v(i)$. Note that for all $t > \tau$ we have that $\bar{b}s^*(t) - \gamma \leq 0$. This implies that

$$\begin{aligned}
(g^*)'(t) &= -\gamma g^*(t) + s^*(t)\left[M(g^*(t)) - M(g^*(T)) - Q(i^*(t))\right] \\
&\leq -\gamma g^*(t) + s^*(t)\left[M(g^*(t)) - M(g^*(T))\right] \\
&< -\gamma g^*(t) + s^*(t)\beta^*(T)(g^*(t) - g^*(T)) \\
&< -\gamma g^*(t) + s^*(t)\bar{b}(g^*(t) - g^*(T)) \\
&< g^*(t)[s^*(t)\bar{b} - \gamma] \leq 0\,.
\end{aligned}$$

This establishes that g^* is strictly decreasing on $[\tau,T]$ and as β^* is non-increasing in g^* the result follows. □

Proof of Corollary 5. As $1 \geq R_0 = \frac{\bar{b}}{\gamma}$ we have that $\tau = \min\{t \in [0,T]\colon \frac{\bar{b}}{\gamma}s(t) \leq 1\} = 0$. Thus, Proposition 4 implies that g^* is strictly decreasing on $[0,T]$ and β^* is non-decreasing on $[0,T]$. Let $t^* = \inf\{t \in [0,T]\colon g(t) = \delta\} \wedge T$. We note that $g^*(t) < \delta$ for $t < t^*$ and $g^*(t) > \delta$ for $t > t^*$. Since $c(b) = \delta(\bar{b} - b)$ it follows immediately from $\beta^*(t) \in \arg\max_{b\in B} g^*(t)b + c(b) = \arg\max_{b\in B}(g^*(t) - \delta)b + \delta\bar{b}$ that $\beta^*(t) = \underline{b}$ for $t < t^*$ and $\beta^*(t) = \bar{b}$ for $t > t^*$. □

Proof of Proposition 6. Denote by $\tau^\beta(q) = \inf\{t\colon s^\beta(t) \leq q\}$[18] and let $\tilde{i}^\beta(q) = i(\tau^\beta(q))$ for all $q \geq s^\beta(T)$ and define $\tilde{i}^{\hat{\beta}}$ analogously. The number of infected $\tilde{i}^\beta$ as a function of the number of susceptible solves

$$\tilde{i}^\beta(s) = i_0 + s_0 - s - \gamma\int_s^{s_0} \frac{1}{\beta(\tau^\beta(z))z}dz$$

for $s \in [s^\beta(T), s_0]$. As β induces less social distancing than $\hat{\beta}$ we have that $\beta(\tau^\beta(z)) \geq \hat{\beta}(\tau^{\hat{\beta}}(z))$ for all $z \geq \max\{s^\beta(T), s^{\hat{\beta}}(T)\}$. It thus follows that $\tilde{i}^\beta(z) \geq \tilde{i}^{\hat{\beta}}(z)$ for all $z \geq \max\{s^\beta(T), s^{\hat{\beta}}(T)\}$. This

[18]By convention $\inf\emptyset = \infty$.

implies that any given share of susceptible population is reached faster with less social distancing, i.e. for all $s \in [\max\{s^{\beta}(T), s^{\hat{\beta}}(T)\}, s_0]$

$$\tau^{\beta}(s) = \int_s^{s_0} \frac{1}{\beta(\tau^{\beta}(z)) z \tilde{i}^{\beta}(z)} dz \leq \int_s^{s_0} \frac{1}{\hat{\beta}(\tau^{\hat{\beta}}(z)) z \tilde{i}^{\hat{\beta}}(z)} dz = \tau^{\hat{\beta}}(s)\,.$$

This implies that $s^{\beta}(t) \leq s^{\hat{\beta}}(t)$ for all $t \in [0, T]$. This implies that herd immunity is achieved faster under a higher transmission rate $\tau^{\beta}(\gamma/\bar{b}) < \tau^{\hat{\beta}}(\gamma/\bar{b})$.

$\square$

Proof of Proposition 7. Let $Q\colon [0,1] \to \mathbb{R}$ satisfy $Q(i) = v'(i)i - v(i)$. As g solves

$$(g^*)'(t) = -\gamma g^*(t) + s^*(t)\left[M(g^*(t)) - M(g^*(T)) - Q(i^*(t))\right] \tag{30}$$

and M and Q are continuous it follows that $(g^*)'$ is continuous. Moreover, we obtain from (30) that

$$\begin{aligned}(g^*)'(t) &= -\gamma g^*(t) + s^*(t)\left[g^*(t)\beta^*(t) + c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))\right] \\ &= \frac{(i^*)'(t)}{i^*(t)} g^*(t) + s^*(t)\left[c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))\right].\end{aligned} \tag{31}$$

By the envelope theorem we get that M is absolutely continuous and for a.e. $t \in [0, T]$ we have $M'(g^*(t)) = \beta^*(t)$. We thus get by (30) that for a.e. $t \in [0, T]$

$$\begin{aligned}(g^*)''(t) = -\gamma(g^*)'(t) + (s^*)'(t)\left[M(g^*(t)) - M(g^*(T)) - Q(i^*(t))\right] \\ + s^*(t)\left[\beta^*(t)(g^*)'(t) - Q'(i^*(t))(i^*)'(t)\right].\end{aligned}$$

Using (30) once again we obtain for a.e. $t \in [0, T]$

$$\begin{aligned}(g^*)''(t) &= -\gamma(g^*)'(t) + \frac{(s^*)'(t)}{s^*(t)}\left[(g^*)'(t) + \gamma g^*(t)\right] \\ &\qquad + s^*(t)\left[\beta^*(t)(g^*)'(t) - Q'(i^*(t))(i^*)'(t)\right] \\ &= \left[-\gamma + \frac{(s^*)'(t)}{s^*(t)} + s^*(t)\beta^*(t)\right](g^*)'(t) + \frac{\gamma g^*(t)(s^*)'(t)}{s^*(t)} - s^*(t)Q'(i^*(t))(i^*)'(t) \\ &= \left[-\gamma + \beta^*(t)(s^*(t) - i^*(t))\right](g^*)'(t) - \gamma g^*(t)\beta^*(t)i^*(t) - s^*(t)Q'(i^*(t))(i^*)'(t).\end{aligned}$$

Next it follows from (31) that for a.e. $t \in [0,T]$

$$\begin{aligned}(g^*)''(t) &= \left[-\gamma + \beta^*(t)(s^*(t) - i^*(t)) - \frac{Q'(i^*(t))i^*(t)s^*(t)}{g^*(t)}\right](g^*)'(t) - \gamma g^*(t)\beta^*(t)i^*(t) \\ &\qquad + \frac{(s^*(t))^2 i^*(t)Q'(i^*(t))\left[c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))\right]}{g^*(t)} \\ &= \left[-\gamma + \beta^*(t)(s^*(t) - i^*(t)) - \frac{Q'(i^*(t))i^*(t)s^*(t)}{g^*(t)}\right](g^*)'(t) \\ &\qquad + g^*(t)\beta^*(t)i^*(t)\left[\frac{(s^*(t))^2 Q'(i^*(t))\left[c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))\right]}{\beta^*(t)(g^*(t))^2} - \gamma\right].\end{aligned}$$

It follows from the fact that $Q'(i) = v''(i)i$ that for a.e. $t \in [0,T]$

$$\begin{aligned}(g^*)''(t) &= \left[-\gamma + \beta^*(t)(s^*(t) - i^*(t)) - \frac{v''(i^*(t))(i^*(t))^2 s^*(t)}{g^*(t)}\right](g^*)'(t) \\ &\qquad + g^*(t)\beta^*(t)i^*(t)\left[\frac{(s^*(t))^2 i^*(t)v''(i^*(t))\left[c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))\right]}{\beta^*(t)(g^*(t))^2} - \gamma\right]\end{aligned} \tag{32}$$

Next note that the facts that $s^*(t) + i^*(t) \leq 1$ and for all $s \in [0,1] : s^2(1-s) \leq \frac{4}{27}$ imply that

$$(s^*(t))^2 i^*(t) \leq (s^*(t))^2(1 - s^*(t)) \leq \frac{4}{27}. \tag{33}$$

Moreover, note that convexity of v and concavity of M ensure that

$$\begin{aligned}c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t)) &\leq c(\beta^*(t)) - M(g^*(T)) \\ &= M(g^*(t)) - M(g^*(T)) - g^*(t)\beta^*(t) \\ &\leq \beta^*(T)(g^*(t) - g^*(T)) - \beta^*(t)g^*(t) \leq (\beta^*(T) - \beta^*(t))g^*(t).\end{aligned} \tag{34}$$

We next consider the two cases (i) $g^*(t) < -c'(\bar{b}-)$ and (ii) $g^*(t) \geq -c'(\bar{b}-)$.

Case (i): Note that convexity of c ensures that for all $b \in [\underline{b}, \bar{b})$ we have

$$\frac{c(\bar{b}) - c(b)}{\bar{b} - b} \leq c'(\bar{b}-).$$

In the case $g^*(t) < -c'(\bar{b}-)$ we thus obtain for all $b \in [\underline{b}, \bar{b})$ that

$$g^*(t)\bar{b} + c(\bar{b}) < g^*(t)b + c(b).$$

This implies that $\beta^*(t) = \bar{b}$. By Proposition 2 we have $g^*(t) > g^*(T)$ and hence we obtain similarly that $\beta^*(T) = \bar{b}$. It follows from (34) that in the case $g^*(t) < -c'(\bar{b}-)$ we have

$$\frac{c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))}{\beta^*(t)(g^*(t))^2} \leq 0. \tag{35}$$

Case (ii): In the case $g^*(t) \geq -c'(\bar{b}-)$ we have with (34) that[19]

$$\frac{c(\beta^*(t)) - M(g^*(T)) - Q(i^*(t))}{\beta^*(t)(g^*(t))^2} \leq \frac{\beta^*(T) - \beta^*(t)}{\beta^*(t)g^*(t)} \leq -\frac{1}{c'(\bar{b}-)}\left(\frac{\bar{b}}{\underline{b}} - 1\right). \tag{36}$$

Combining (33), (35), (36) and (32) implies that in both cases (i) and (ii) we have that for a.e. $t \in [0, T]$

$$\begin{aligned}(g^*)''(t) \leq &\left[-\gamma + \beta^*(t)(s^*(t) - i^*(t)) - \frac{v''(i^*(t))(i^*(t))^2 s^*(t)}{g^*(t)}\right](g^*)'(t) \\ &- g^*(t)\beta^*(t)i^*(t)\left[\frac{4v''(i^*(t))}{27c'(\bar{b}-)}\left(\frac{\bar{b}}{\underline{b}} - 1\right) + \gamma\right].\end{aligned} \tag{37}$$

To show that g^* is strictly quasi concave consider a point $t^* \in (0, T)$ such that $(g^*)'(t^*) = 0$. It follows from boundedness of β^*, i^* and s^*, the fact that $g^* > 0$ and continuity of $(g^*)'$ that

$$\lim_{t \to t^*}\left[-\gamma + \beta^*(t)(s^*(t) - i^*(t)) - \frac{v''(i^*(t))(i^*(t))^2 s^*(t)}{g^*(t)}\right](g^*)'(t) = 0.$$

Moreover, $\gamma\beta^* g^* i^*$ is bounded away from 0 on $[0, T]$. Therefore Assumption 1 that $\frac{v''(i)}{\gamma} < \frac{27\underline{b}}{4(\bar{b}-\underline{b})}(-c'(\bar{b}-))$ and (37) ensure that there exist $\epsilon, \delta > 0$ such that $(g^*)''(t) \leq -\delta$ for a.e. $t \in (t^* - \epsilon, t^* + \epsilon)$. This implies for all $t \in (t^* - \epsilon, t^*)$ that

$$(g^*)'(t) = -((g^*)'(t^*) - (g^*)'(t)) = -\int_t^{t^*}(g^*)''(r)dr \geq \delta(t^* - t) > 0.$$

Similarly, we get that $(g^*)'(t) < 0$ for all $t \in (t^*, t^* + \epsilon)$. It follows that t^* is a strict local maximum. Thus every local extremum of g^* in $(0, T)$ is a strict local maximum and hence it follows that g^* is strictly quasi concave. In particular, g^* is single peaked. For the sequel let $t^* \in [0, T]$ be such that g^* is strictly increasing on $[0, t^*]$ and strictly decreasing on $[t^*, T]$. It follows from Proposition 2 that β^* is quasi-convex.

We next show that i^* is strictly log-concave on $[0, t^*]$. Since β^* is non-increasing on $[0, t^*]$ and s^* is strictly decreasing it follows that $\frac{\partial}{\partial t}\log(i^*(t)) = \frac{(i^*)'(t)}{i^*(t)} = \beta^*(t)s^*(t) - \gamma$ is strictly decreasing on $[0, t^*]$. Consequently, i^* is strictly log-concave on $[0, t^*]$ and thus has at most one local maximum

[19] Note that Assumption 1 implies that $-c'(\bar{b}-) > 0$ as $-c'(\bar{b}-) = 0$ otherwise would imply that $v''(i) < 0$ which contradicts the convexity of v.

on $[0, t^*]$. $\square$

Proof of Proposition 8. Let $c(\beta) = \delta(\bar{b} - \beta)$. It follows from (12) that any optimal control satisfies

$$\beta^*(t) = \begin{cases} \bar{b} & \text{if } g^*(t) < \delta \\ \underline{b} & \text{if } g^*(t) > \delta \,. \end{cases} \tag{38}$$

As argued in the proof of Proposition 7, g^* is strictly quasi-concave. If $g^*(t) < \delta$ for all $t \in [0, T]$ we set $t_1^* = t_2^* = 0$. If there exists $\bar{t} \in [0, T]$ with $g(\bar{t}) \geq \delta$ we set $t_1^* = \inf\{t \in [0, T] \colon g^*(t) \geq \delta\}$ and $t_2^* = \sup\{t \in [0, T] \colon g^*(t) \geq \delta\}$. This together with (38) implies (10).

Next note that β^* is almost everywhere equal to a non-increasing function on the interval $[0, t_2^*)$. This implies that $\frac{\partial}{\partial t} \log(i^*(t)) = s^*(t) b^*(t) - \gamma$ is strictly decreasing and thus that i^* is strictly log-concave and admits at most one local maximum on $[0, t_2^*]$. It follows similarly that as β^* is almost everywhere equal to a constant function on $(t_2^*, T]$ that i^* has at most one local maximum on $[t_2^*, T]$. In the case $t_2^* > 0$ there can never be a local maximum at t_2^* as the left derivative is strictly less than the right derivative.

We finally show that i^* is strictly increasing on $[0, t_1^*]$. First assume that $g^*(T) \geq \delta$. Then we have $g^*(t) > g^*(T) \geq \delta$ for all $t \in [0, T)$ and hence $t_1^* = 0$ and there is nothing to show. For the remainder of the proof assume that $g^*(T) < \delta$. Then we have $\beta^*(T) = \bar{b}$ and hence for all $t \in [0, t_1^*)$

$$\begin{aligned} \frac{(g^*)'(t)}{(g^*)(t)} &= -\gamma + \frac{s^*(t)}{g^*(t)} \left[M(g^*(t)) - M(g^*(T)) - Q(i^*(t))\right] \\ &\leq -\gamma + \frac{s^*(t)}{g^*(t)} \left[M(g^*(t)) - M(g^*(T))\right] \\ &= -\gamma + \frac{s^*(t)}{g^*(t)} \left[g^*(t) - g^*(T)\right] \bar{b} \\ &= -\gamma + s^*(t) \left[1 - \frac{g^*(T)}{g^*(t)}\right] \bar{b} \leq \bar{b} s^*(t) - \gamma = \frac{(i^*)'(t)}{i^*(t)} \,. \end{aligned}$$

Observe, that as g^* is strictly increasing on $[0, t_1^*]$[20] it follows that i^* is strictly increasing on $[0, t_1^*]$. $\square$

[20]Note that if we have $g^*(t) < \delta$ for all $t \in [0, T]$ then by construction we have $t_1^* = 0$ and this claim clearly holds true.

References

Abakuks, Andris, "Some optimal isolation and immunisation policies for epidemics." PhD dissertation, University of Sussex 1972.

_ , "An optimal isolation policy for an epidemic," *Journal of Applied Probability*, 1973, *10* (2), 247–262.

_ , "Optimal immunisation policies for epidemics," *Advances in Applied Probability*, 1974, *6* (3), 494–511.

Acemoglu, Daron, Victor Chernozhukov, Iván Werning, and Michael D Whinston, "A Multi-Risk SIR Model with Optimally Targeted Lockdown," Technical Report, National Bureau of Economic Research 2020.

Akbarpour, Mohammad, Cody Cook, Aude Marzuoli, Simon Mongey, Abhishek Nagaraj, Matteo Saccarola, Pietro Tebaldi, Shoshana Vasserman, and Hanbin Yang, "Socioeconomic Network Heterogeneity and Pandemic Policy Response," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2020, (2020-75).

Alvarez, Fernando E, David Argente, and Francesco Lippi, "A simple planning problem for covid-19 lockdown," Technical Report, National Bureau of Economic Research 2020.

Atkeson, Andrew, "What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios," Technical Report, National Bureau of Economic Research 2020.

Barro, Robert J, Jose F Ursua, and Joanna Weng, "The Coronavirus and the Great Influenza Epidemic-Lessons from the," 2020.

Behncke, Horst, "Optimal control of deterministic epidemics," *Optimal control applications and methods*, 2000, *21* (6), 269–285.

Bohner, Martin, Sabrina Streipert, and Delfim FM Torres, "Exact solution to a dynamic SIR model," *Nonlinear Analysis: Hybrid Systems*, 2019, *32*, 228–238.

Clarke, Francis, *Functional analysis, calculus of variations and optimal control*, Vol. 264, Springer Science & Business Media, 2013.

de Walque, Damien, Jed Friedman, Roberta Gatti, and Aaditya Mattoo, "How Two Tests Can Help Contain COVID-19 and Revive the Economy," 2020.

Dewatripont, Mathias, Michel Goldman, Eric Muraille, and Jean-Philippe Platteau, "Rapidly identifying workers who are immune to COVID-19 and virus-free is a priority for restarting the economy," *VOX CEPR Policy Portal*, 2020, *23.*

Garibaldi, Pietro, Espen R Moen, and Christopher A Pissarides, "Modelling contacts and transitions in the SIR epidemics model," *Covid Economics Vetted and Real-Time Papers, CEPR*, 2020.

Gonzalez-Eiras, Martin and Dirk Niepelt, "On the Optimal "Lockdown" During an Epidemic," Technical Report 2020.

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Iván Werning, "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?," Technical Report, National Bureau of Economic Research 2020.

Hansen, Elsa and Troy Day, "Optimal control of epidemics with limited resources," *Journal of mathematical biology*, 2011, *62* (3), 423–451.

Kantner, Markus, "Beyond just" flattening the curve": Optimal control of epidemics with purely non-pharmaceutical interventions," *arXiv preprint arXiv:2004.09471*, 2020.

Kermack, William Ogilvy and Anderson G McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 1927, *115* (772), 700–721.

Kissler, Stephen M, Christine Tedijanto, Marc Lipsitch, and Yonatan Grad, "Social distancing strategies for curbing the COVID-19 epidemic," *medRxiv*, 2020.

Miclo, Laurent, Daniel Spiro, and Jörgen Weibull, "Optimal epidemic suppression under an ICU constraint," *arXiv preprint arXiv:2005.01327*, 2020.

Morton, Richard and Kenneth H Wickwire, "On the optimal control of a deterministic epidemic," *Advances in Applied Probability*, 1974, *6* (4), 622–635.

Piguillem, Facundo, Liyan Shi et al., "The optimal covid-19 quarantine and testing policies," Technical Report, Einaudi Institute for Economics and Finance (EIEF) 2020.

Rowthorn, Bob RE and Flavio Toxvaerd, "The optimal control of infectious diseases via prevention and treatment," 2012.

Stock, James H, "Data gaps and the policy response to the novel coronavirus," Technical Report, National Bureau of Economic Research 2020.

Toda, Alexis Akira, "Susceptible-Infected-Recovered (SIR) Dynamics of COVID-19 and Economic Impact," *arXiv preprint arXiv:2003.11221*, 2020.

Wickwire, K, "A note on the optimal control of carrier-borne epidemics," *Journal of Applied probability*, 1975, *12* (3), 565–568.

— , "Mathematical models for the control of pests and infectious diseases: a survey," *Theoretical population biology*, 1977, *11* (2), 182–238.

Wu, Joseph T, Kathy Leung, and Gabriel M Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study," *The Lancet*, 2020, *395* (10225), 689–697.