Optimal Case Detection and Social Distancing Policies to Suppress COVID-19*

Stefan Pollinger †

January 17, 2021 ‡

Latest version available here.

Abstract

This paper finds that the combination of case detection and social distancing is crucial for the efficient eradication of a new infectious disease. Theoretically, I characterize the optimal suppression policy as a simple function of observables, which eases its implementation. Together with the number of infected, optimal social distancing decreases over time. The fundamental trade-off is between its intensity and its duration. Quantitatively, I calibrate the model to the COVID-19 pandemic in Italy. Given the observed prevalence and detection efficiency on May 10th, suppression costs 11 % of annual GDP. Efficient digital contact tracing reduces this cost to 0.4 %.

JEL-Code: E1, I18, D6, H84.

Keywords: Covid-19, suppression policy, contact tracing, social distancing.

*I am especially grateful to Christian Hellwig, Nicolas Werquin, and Christian Gollier for their supervision and guidance. I thank Tiziana Assenza, Fabrice Collard, Martial Dupaigne, Patrick Féve, James Hammit, Alix de Loustal, Arnaud Reynaud, Paul Seabright, François Salanié, Nicolas Treich, and the participants of the Covid and the Macro workshop at the Toulouse School of Economics for their comments and suggestions. I acknowledge support from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). All errors are mine.

†Stefan Pollinger, Ph.D. candidate at the Toulouse School of Economics; email address: stefan.pollinger@tse-fr.eu; website: www.stefanpollinger.com

‡First public version, May 15th. Published in Covid Economics, Issue 23, May 28, 2020. I plan to update the paper frequently. Please find the latest version on www.stefanpollinger/research

1 Introduction

What is the optimal response to a rapidly spreading and deadly infectious disease, when no vaccine or efficient medication is available? Due to the COVID-19 pandemic, it has become very urgent to answer this policy question. In a broad sense, two different policy-approaches are possible: mitigation and suppression. On the one hand, mitigation controls the spread of the virus until contagions stop because the population achieves herd immunity.[1] However, to acquire immunity, a considerable number of individuals need to contract the disease and recover. Inevitably, some will die in the process. The consequent number of casualties is the main drawback to this policy-approach. On top of that, the strategy bears the risk that immunity vanishes over time, or that the virus mutates. In both cases, the virus becomes endemic, and the policy fails. On the other hand, suppression policies push the viral growth rate below zero, such that the disease dies-out over time. The policy avoids the infection of a large part of the population; however, controlling viral growth is costly.

This paper shows how to optimally suppress a virus when the policymaker has two tools: social distancing and case detection. Social distancing reduces the growth rate of the virus by reducing the rate of social contacts between all individuals in the population. Case detection, for instance, with the help of tests and contact tracing, reduces growth by actively finding infectious individuals and isolating them from the susceptible population. The optimal policy minimizes total economic and health costs. While the main focus of this paper is on suppression policies, I compare them with mitigation policies in the conclusion.

The first contribution of the paper is theoretical. I characterize the properties of the optimal policy to derive concrete policy implications. I find that optimal social distancing decreases when viral prevalence decreases. This simple property offers important guidance for policy-design. Suppose a policymaker discovers a first outbreak of a virus. In the optimum, she immediately implements social distancing measures to reverse viral growth. As a consequence, the number of infectious reduces and converges to zero. Importantly, as the number of infectious decreases, the policymaker gradually relaxes the social distancing measures. The optimal response is instantaneous and the largest at the onset. In particular, it is not optimal to "smooth in" social distancing or to wait before imposing measures, a

[1]A population reaches herd immunity when a large enough fraction is immune to infection, i.e., not susceptible. The fraction is large enough when, on average, one infected individual meets and transmits the disease to less than one susceptible individual.

mistake made by many countries at the beginning of the COVID-19 pandemic. Any initial hesitation or delay in implementing sufficiently strong social distancing measures worsens economic and health outcomes. The reason is that weak measures allow the virus to grow further, which increases the number of casualties and the time and intensity of an afterward necessary measures. Moreover, a too fast exit easing of social distancing is not optimal. On the optimal path, the number of infections does not rise. In case a policymaker discovers signs of rebounding case numbers - such as an increase in the flow of confirmed cases, symptomatic patients, hospitalizations, or death - a swift increase in social distancing is the optimal response.

Two simple and observable sufficient statistics characterize the optimal policy at each point in time: first, the instantaneous growth rate of the virus, and second, the instant flow of costs from suppression measures and health outcomes.[2] A policy at a certain point in time is optimal if the elasticity of the current flow cost to the current growth rate is equal to one. Note that this property is somewhat surprising. In principle, the optimal policy at a certain point in time depends on the past and the future. However, the two sufficient statistics contain all relevant dynamic information. The condition gives specific and straightforward guidance on how to stay on the optimal suppression path over time, and, in particular, on how fast to relax social distancing measures. To decide upon relaxing a particular measure, the policymaker only needs to evaluate the relative impact on the current flow of costs and the viral growth rate . If the percentage reduction in cost is larger than the percentage increase in growth, a measure should be relaxed.

Methodologically, I exploit the fact that when the number of infectious is low compared to the number of susceptible, a simple exponential process approximates the dynamic behavior of infections. Note that this is the relevant case for studying suppression because the number of infectious goes to zero. The approximation simplifies the heavy SIR machinery currently used in the literature. I solve the model with pen and paper, which allows me to study the dependence of the optimal policy and welfare on the unknown functions and parameters. These unknowns are the economic cost and viral growth impact of social distancing policies, the flow of death per infection and its social cost, the uncontrolled growth rate of the virus, and the speed of detection as a function of the overall stock of infected. Although possible to estimate, typically, very little is know about these key determinants

[2]For a real-time estimation of the economic costs of the pandemic, see Chetty, Friedman, Hendren, Stepner et al. (2020).

for optimal policy when dealing with a new infectious disease. Therefore, it is crucial to study the properties of optimal policy without making restrictive assumptions on the unknowns.

The exact characterization of the optimal policy at each point in time follows from simple intuition. The critical dynamic trade-off when suppressing the virus is between the intensity of social distancing and the time it needs to stay in place. Too extreme measures rapidly reduce the number of infections in the population; however, they have very high instantaneous costs, because even the most fundamental economic activities are on hold. Too weak measures have low instant costs; however, to eradicate the virus, they need to stay in place for a very long time. The optimal policy trades off these two margins at every instant. Consider a particular current level of infectious, and consider the cost of reducing it by one unit. This unit cost is the current flow cost from health outcomes and social distancing measures, multiplied by the time it takes to reduce infections by one unit. Both factors depend on the intensity of social distancing. The stricter social distancing, the higher the cost, and the lower the time it takes to suppress one unit. The unit cost is at its minimum when the percentage change in cost is equal to the percentage change in time - a property of interior extrema of products of functions. As time is inversely proportional to the growth rate, the same property is valid for the growth rate instead of time. What follows is the optimality condition as a relation of two simple sufficient statistics: the current flow cost and the current growth rate. A policy is optimal if, at each point in time, its relative impact on the flow cost is equal to its relative impact on the viral growth rate. The optimal total cost is simply the integral over the optimal unit costs.

In the long-run, the optimal degree of social distancing depends crucially on the detection-technology. A key property is the detection rate, i.e., the number of daily detected cases relative to the overall number of currently infectious. Due to decreasing returns to scale, the rate of detection is decreasing in prevalence. The optimal long-run policy depends on the detection rate at zero prevalence. On the one hand, if the rate of detection at zero is larger than the uncontrolled growth rate of the virus, optimal social distancing measures are entirely removed in the long run. I call this case efficient detection. It means that society is going back to normal, along with the decreasing number of overall infectious in the population. Intuitively, the smaller the number of infectious, the more significant is the relative amount of control coming from case detection. In the long run, case detection is efficient enough to control the virus completely. On the other hand, if the rate of detection at zero is lower than the uncontrolled growth rate of the virus, optimal long-run social distancing is constant and positive. Detection alone cannot control the disease and needs to be complemented by social distancing until the disease dies out.

The efficiency of case detection, and therefore, the long-run behavior of the optimal policy, has stark consequences for the total cost of suppression. The more efficient is the tracing technology, the lower is the necessary amount of long-run social distancing, and therefore, the cost of suppression. On the one hand, if case detection is inefficient, the total cost of suppression depends on the extinction threshold of the virus. The extinction threshold is the level of prevalence where the virus dies out. In the limit, when the extinction threshold goes to zero, the total cost of suppression goes to infinity. Typically, the extinction threshold is very small. The result shows that the total cost of suppression is substantial under inefficient case detection. The reason is that prevalence follows an exponential decay process. In the long run, the reduction in infections becomes infinitely slow. Besides, some degree of social distancing needs to stay in place until the virus becomes extinct. As a consequence, the total cost is unbounded in the limit. On the other hand, in stark contrast, if case detection is efficient, the total cost of suppression is bounded for any extinction threshold. As a result, the total cost of suppression is relatively low. These results suggest that efficient case detection, at least at low infection levels, has enormous benefits. Note, however, I do not take welfare losses from an eventual loss in privacy into account.

The results of the last paragraph show that to limit the overall cost of suppression, the combination of efficient case detection and social distancing is crucial. They become particularly intuitive when considering the case where only social distancing is used to suppress the virus. Relying on social distancing alone poses two problems. First, social distancing affects all individuals in a population, making it very costly. Second, because the decrease in viral prevalence follows an exponential decay process, social distancing becomes very inefficient when prevalence is low. It takes the same effort to reduce the number of infected from 20,000 to 10,000 as it does to reduce it from from 20 to 10. Towards the end of a pandemic, it is necessary to impose costly measures on the whole population just to avoid one last transmission of the virus.

The theoretical results give concrete policy advice on how to respond to a future pandemic. If a sufficiently efficient case detection procedure is in place when a new infectious disease appears, social distancing is unnecessary. As a consequence, high economic costs are avoidable. Otherwise, if detection is not efficient enough, it is necessary to complement case detection with social distancing immediately. In combination, the two measures need to be strong enough to stop the spread of the disease. If the policymaker fails to respond, infections will spread. Bringing them back under control necessitates more intense, and consequently, more costly social distancing later, just to bring the epidemic back to the starting point. Many countries painfully learned this lesson during the first outbreak of COVID-19.

The second contribution of the paper is quantitative. To compare how the total cost and time of suppression depends on realistic detection technologies, I calibrate the unknown functions and parameter values using data from Italy and South Korea. I use 0.07% prevalence - the estimate for Italy on May 10th - as an initial condition. I consider three different detection scenarios. In the first scenario, Italy uses fast and efficient digital contact tracing like South Korea. In the second scenario, Italy uses slower and less efficient manual tracing. In the third scenario, Italy continues to detect cases at the observed low rate.[3]

I find that the total cost of suppressing COVID-19, using digital tracing, is only 0.4% of annual GDP. The strategy allows for a fast and continuing reduction of social distancing. After 39 days already, optimal social distancing is at such a low level that its flow-cost is only 1% of daily GDP. Afterward, the daily cost continues to converge to zero. The virus is entirely under control, and social activity is back to a normal level well before a vaccine arrives. Additionally, the strategy is robust to a certain degree of imported cases. The number of additional casualties under this scenario would be 3,300. Under the second scenario, when using manual contact tracing, I find that the total cost of suppression is 1.7% of annual GDP. The flow-cost of social distancing drops below 1% of daily GDP after 3 months. Manual tracing is not efficient enough to allow for a total return to normality. In the long run, some degree of social distancing needs to stay in place; however, its flow-cost is only

[3] Digital contact tracing uses mobile phone data to identify and inform the past contacts of a confirmed infectious individual. It is particularly fast and efficient. Its maximal detection rate is 35% per day (Ferretti, Wymant, Kendall, Zhao, Nurtay, Abeler-Dörner, Parker, Bonsall, and Fraser, 2020). Manual contact tracing relies on teams of tracing personal who question confirmed infectious and find their contacts manually. Its maximal detection rate is 10% per day. See Ferretti et al. (2020) for an extensive discussion. Currently, Italy detects 3% of cases per day.

0.1% of daily GDP. The virus dies out after 15 months. The number of additional casualties under this scenario would be 4,200. In stark contrast, the total cost of suppression in the no tracing scenario is 11% of annual GDP. The reason is that, in this case, optimal social distancing is very close to constant. Its flow-cost is 19% of daily GDP. The cost accrues until the virus becomes extinct after 8 months. Additionally, if new cases are imported after extinction, the pandemic restarts. Therefore, meticulous border controls need to stay in place until a vaccine arrives. The number above does not take into account their costs. The number of additional casualties would be 2,900. I compare the optimal suppression policy with optimal mitigation policies in the conclusion.

Relevant Literature This paper contributes to the economic literature on optimal disease control. A large and recent literature studies mitigation policies, using variants of the SIR model augmented with economic interactions. As discussed above and in the conclusion, mitigation policies are very distinct from a suppression policies, which I study in this paper. The mitigation literature mostly uses numerical methods to solve for the optimal policy, or to simulate the impact of certain policies of interest. See Acemoglu, Chernozhukov, Werning, and Whinston (2020), Alvarez, Argente, and Lippi (2020), Atkeson (2020), Berger, Herkenhoff, and Mongey (2020), Bethune and Korinek (2020), Chari, Kirpalani, and Phelan (2020), Eichenbaum, Rebelo, and Trabandt (2020), Farboodi, Jarosch, and Shimer (2020), Favero, Ichino, and Rustichini (2020), Gollier (2020), Gonzalez-Eiras and Niepelt (2020), Hornstein (2020), Jones, Philippon, and Venkateswaran (2020), Miclo, Spiro, and Weibull (2020),Obiols-Homs (2020), and Piguillem and Shi (2020). The list is far from exhaustive. Assenza, Collard, Dupaigne, Fève, Hellwig, Kankanamge, Werquin et al. (2020), Garibaldi, Moen, and Pissarides (2020), and Rachel (2020) characterize the theoretical properties of optimal mitigation policies.

A smaller part of the literature studies suppression. Gollier (2020), Scherbina (2020), and Ugarov (2020) simulate the impact of a uniform lock-down. Dorn, Khailaie, Stoeckli, Binder, Lange, Lautenbacher, Peichl, Vanella, Wollmershaeuser, Fuest et al. (2020) simulate the effects of various control scenarios using a detailed economic and epidemiological model. Wang (2020) simulates the effect of mass testing and shows that it can lead to suppression before herd immunity. I contribute to this literature by explicitly characterizing the optimal time-variable suppression policy.

Closest to my paper are Gerlagh (2020), Piguillem and Shi (2020), Bethune and Korinek (2020), and Alvarez et al. (2020). Gerlagh (2020) solves for the optimal suppression policy when case-detection is constant and ineffective, and assumes extinction is not possible. I explicitly model case-detection and extinction. Piguillem and Shi (2020) solve for the optimal suppression policy in a SIR model with social distancing and random testing. Bethune and Korinek (2020) solve for the optimal suppression policy under two polar assumptions on the information set of the planner: the planner exactly knows who is infected, and the planner has no information at all on who is infected. Alvarez et al. (2020) solve for the optimal policy with tracing. They explicitly assume a functional form for tracing. The last three contributions are quantitative. I contribute to this literature by characterizing the optimal suppression policy as the solution of two simple sufficient statistics. I derive its properties under general functional forms and parameter values. It is important because very little is know about key parameters and relevant functions influencing optimal policy. In particular, I show that the optimal suppression policy and its cost depend crucially on well-defined properties of the detection technology. The tracing function used by Alvarez et al. (2020) is infinitely efficient in the limit. This property is at odds with the epidemiological literature on case detection; see Ferretti et al. (2020). The property leads to overly optimistic estimates for the efficiency of suppression. Quantitatively, I contribute to this literature by comparing the optimal suppression policy under realistic detection technologies. I use epidemiological estimates (Ferretti et al., 2020) to calibrate the detection technologies.

Pueyo (2020) gives an extensive informal discussion of possible policy responses.

2 The Model

Assume there is an initial mass I_0 of infectious individuals in a susceptible population. The virus transmits from infectious to susceptible. Infected individuals die or recover from the disease after a certain time. Assume that, when uncontrolled, the mass of infectious individuals I_t at time t follows:

$$\dot{I}_t = r^0 I_t. \tag{1}$$

The variable $\dot{I}_t$ denotes the time derivative, and r^0 is the uncontrolled viral growth rate. Assume that $r^0 > 0$; the virus is spreading. The equation describes an exponential growth process with a growth rate of r^0.

When the number of infectious is small compared to the number of susceptible, the above process approximates the standard SIR model. The SIR model is widely used in the economics literature on optimal disease control. When suppressing the disease, the number of infected needs to converge to zero, and therefore, at some point, the number of infectious is inevitably small compared to the number of susceptible. In the limit, when $I_t = 0$, the approximation is exact. The general SIR model can only be solved numerically. Using the above approximation has the advantage of simplifying the analysis considerably. It allows me to solve for the optimal suppression policy analytically. For a discussion of the SIR model, as-well-as its relation to the above equation, see appendix A.1. For the quantitative results I calibrate the model to Italy, using the situation on May 10th as an initial condition. This date is 3 months after the onset of the pandemic in Italy. The country was hardly hit by the virus, and it went through a period of strict lock-down. The mass of infectious on that date is equal to 0.07%, and the mass of susceptible is equal to 96%. The quantitative exercise in Section 4 shows that the change in the mass of susceptible on the optimal suppression path is smaller than 0.7%, and therefore negligible.

The policymaker can alter the spread of the virus by using two tools: case detection and social distancing.

2.1 Case Detection

Case detection allows for quarantining a mass X of infectious individuals at each instant of time. I assume infectious individuals in quarantine do not infect susceptible individuals. $X(I)$ is the flow of detected cases into quarantine as a function of the mass of infectious I. Intuitively, it is the speed of detection. When there is case detection the mass of infectious follows

$$\dot{I}_t = r^0 I_t - X(I_t). \tag{2}$$

For a derivation from the SIR model, see appendix A.1.2. Assume $X(0) = 0$; if there are no infectious, none can be detected. $X'(I) > 0$; the speed of detection increases with the number of infectious. When there are more infectious, it is easier to detect them. Assume that the overall resources for case detection in a country are fixed and constant over time. Consequently, assume that $X''(I) < 0$; the increase in speed is decreasing in the number of infectious. The detection-technology becomes overwhelmed if there are too many infected, i.e., there are decreasing returns to scale. In the limit, if I goes to infinity, $X'(I)$ goes to zero. A discussion of this assumption is in appendix A.1.1. Note that for many countries, it has been difficult to increase the detection capacity in the short term. For this reason, I focus on the simple case where resources for detection are constant. An extension to the case where resources for detection are variable is in appendix A.3.2.

Define the detection rate as

$$\frac{X(I)}{I}. \tag{3}$$

Intuitively, it gives the percentage of overall cases that are detected daily. Under the above assumptions, the detection rate is decreasing in I. The proof of this statement is in Lemma 2 in the appendix. It is the largest at zero. The detection rate at zero is a key parameter for the analysis. Denote it as ξ_0:

$$\xi_0 = \lim_{I \to 0} \frac{X(I)}{I} = X'(0). \tag{4}$$

There are two distinct cases:

Lemma 1. .

1. *If $\xi_0 \leq r^0$, the rate of case detection is never larger than the uncontrolled viral growth rate. Therefore, detection alone cannot suppress the virus.*
2. *If $\xi_0 > r^0$, there exists a level of infectious $I^* > 0$, such that for all $I < I^*$, it holds that $\dot{I}(I) < 0$. Therefore, if $I_t < I^*$, detection alone suppresses the virus. I^* is the point where $r^0 I^* = X(I^*)$.*

The proof is in appendix A.1.2.

2.2 Social Distancing

Assume social distancing policies are indexed by $p \in [0, 1]$. Each policy p has mass zero. As an example, think of policy p as closing a sector p of the economy. Each p reduces the growth rate of the virus by $dr(p)$ and has a social cost $dc(p)$. Assume policies are indexed such that the cost benefit ratio $\frac{dc(p)}{dr(p)}$ is increasing. Also, assume that $\frac{dc(0)}{dr(0)} = 0$ and $\frac{dc(1)}{dr(1)} = \infty$. Applying policies 0 to p has a growth impact of

$$r(p) = \int_0^p r'(\tilde{p})d\tilde{p}. \tag{5}$$

Following the sectoral interpretation, applying policies 0 to p means closing a fraction of p of the economy. The order of sectoral closure is such that the social cost relative to the growth benefit is increasing. Assume that there are enough policies available such that $r(1) >> r^0$. Strict enough measures allow pushing the growth rate of the virus below zero, i.e., exponential decay. Denote by p_t the fraction of policies applied by the policymaker at time t. The spread of the virus follows the process:

$$\dot{I}_t = (r^0 - r(p_t))I_t. \tag{6}$$

For a derivation from the SIR model, see appendix A.1.2. If $r(p_t) > r^0$ the process follows an exponential decay. Physically, for any initial level of infections I_0, the suppression of the virus is possible by keeping $r(p_t) > r^0$. However, $\dot{I}_t$ goes to zero as I_t goes to zero. The smaller I_t, the slower the suppression is advancing. In the limit, the process becomes infinitely slow. This property has important consequences for the cost of suppressing the virus, which I discuss in Section 3.4.

The flow cost of applying policies 0 to p is

$$c(p) = \int_0^p c'(\tilde{p})d\tilde{p}. \tag{7}$$

To summarize, the functions $r(p)$ and $c(p)$ have the following properties: they are increasing and zero at zero, $\frac{c'(p)}{r'(p)}$ is increasing and zero at zero, $r(1) >> r^0$, and $c(1) = \infty$.

For some derivations, it is more convenient to express the flow-cost as a function of r instead of p:

$$c(r) = c(p(r)). \tag{8}$$

It follows that $c(r)$ is increasing and convex: $c'(.) > 0$ and $c''(.) > 0$. The cost, as well as the marginal cost, is zero in the origin: $c(0) = 0$ and $c'(0) = 0$. Note that this is an abuse of notation. I use the same letter for two different functions. Which function is meant will be clear from the context.

For parts of the results, I assume a very small but positive extinction threshold $I_\epsilon > 0$ exits, i.e., as soon as the number of infectious falls below I_ϵ, the virus dies out. This assumption is a convenient short-cut for describing the extinction behavior when the number of infectious is low. It is common in the economics literature (see Gollier, 2020, and Piguillem and Shi, 2020; the assumption is implicit in quantitative models). As common in the deterministic SIR models, I assume the number of infected is a continuous mass. This assumption makes the model tractable. However, it implies that at a low mass of I, a fraction of an individual is infected, which is not possible in reality. For low numbers of infected, the true infection-process becomes discrete. At each instant of time, a discrete number of individuals becomes infected. New infections and the exit from the state of infectiousness are probabilistic. In particular, if I is sufficiently low, there is a positive probability that all infected recover without infecting a new susceptible individual. In this case, the virus dies out. Note that in the deterministic model, the virus never dies out. Even if there is constant and negative growth, in the limit, an infinitesimal fraction of one individual is infected. The extinction threshold I_ϵ is a convenient way to make the deterministic model more realistic and introduce extinction. I leave a generalization to probabilistic extinction for future research. Note that all results in this paper hold for arbitrarily small extinction thresholds. The results for the case when the rate of detection ξ_0 is larger than the uncontrolled growth rate r^0 also hold when the extinction threshold I_ϵ is equal to zero.

3 The Optimal Policy

Assume that each new infection has a health cost of v. Note that infectious exit the state of infectiousness at rate θ. When they exit they recover or die. $(r^0 - r(p_t))I_t$ is the net flow of infections. Therefore, the flow of new infections is $(r^0 - r(p_t) + \theta)I_t$. For the detailed derivation see appendix A.1.2. Denote the probability that the outcome of an infection is death by δ. Denote the value of a statistical life by vsl. The flow cost per infection v is equal to $v = \delta\, vsl$.[4] The problem of the policymaker is:

$$\min_{p_t} C(p_t) = \int_0^\infty c(p_t) + vI_t(r^0 - r(p_t) + \theta)dt \tag{9}$$

such that

$$\dot{I}_t = (r^0 - r(p_t))I_t - X(I_t). \tag{10}$$

Note that in problem (9), I neglect time discounting. This assumption simplifies the problem considerably. Time discounting is not very important for the problem of the optimal suppression policy. The time frame is days, and daily interest rates are very low. The probabilistic arrival of an effective cure or mass vaccine for a disease has the same effect as time discounting. However, the daily arrival-probability of an effective treatment or mass vaccine for COVID-19 is very low as-well. I solve for the general case with discounting and a random vaccine or treatment arrival in appendix A.3.1. I find that, as long as the discount rate and arrival-probability are low enough, the qualitative results in this chapter do not change.

The solution to the problem is an optimal control function p_t. Assume that at the optimum $\lim_{t\to\infty} I_t = 0$ and $\dot{I}_t < 0$ for all t. These assumptions can be verified ex-post. It follows that I_t is strictly decreasing in time and therefore invertible. Use the invertibility of I to eliminate time in the minimization problem (9):

$$\min_{p(.)} C(p(.)) = \int_{I_0}^{I_\epsilon} \frac{c(p(I)) + vI(r^0 - r(p(I)) + \theta)}{\dot{I}(I)} dI, \tag{11}$$

where

$$\dot{I}(I) = (r^0 - r(p(I)))I - X(I). \tag{12}$$

The solution to the problem is a control function $p(.)$. It is the solution to a simple pointwise minimization of the above integral. Its solution gives the main result of the paper:

Proposition 1. .

[4] v may be interpreted more broadly as containing all other costs caused by an infection, such as the dis-utility of being sick and chronic damages caused by the virus. The cost can be generalized to a nonlinear cost in I, accounting for congestion effects in the health care sector.

For each amount of currently infectious I, the optimal policy $p(I)$ solves:

$$\frac{c'(p) - vIr'(p)}{c(p) + vI(r^0 - r(p) + \theta)} = \frac{r'(p)}{r(p) + \frac{X(I)}{I} - r^0}. \tag{13}$$

An optimal solution $p(.)$ exists and is unique.

The proof is in appendix A.2.1. Note that the denominator on the left of Condition (13), $c(p) + vI(r^0 - r(p) + \theta)$, is the instantaneous flow cost from suppression measures and health outcomes, while the denominator on the right, $r(p) + \frac{X(I)}{I} - r^0$, is the instantaneous negative viral growth rate. The respective enumerators are the marginal impacts of a policy change on these two variables. This fact gives rise to the following corollary:

Corollary 1. .
A policy $p(I)$ is optimal if at each point in time, its relative effect on the flow cost is equal to its relative effect on the viral growth rate:

$$\frac{d\ \log\left(c(p) + vI(r^0 - r(p) + \theta)\right)}{d\,p} = \frac{d\ \log\left(r(p) + \frac{X(I)}{I} - r^0\right)}{d\,p}. \tag{14}$$

Corollary 1 has concrete policy implications, which I discuss in Section 3.2. Alternatively, one can express Condition (13) as an elasticity:

Corollary 2. .
A policy is optimal if at each point in time, the elasticity of the flow cost to the negative viral growth rate is equal to one:

$$\frac{d\ \log\left(c\left(-g + r^0 - \frac{X(I)}{I}\right) + vI\left(g + \frac{X(I)}{I} + \theta\right)\right)}{d\ \log(-g)} = 1. \tag{15}$$

Note that Corollary 2 follows from the fact that, for each level of infectious I, there is a one to one mapping from policy p to the instantaneous negative growth rate g:

$$-g(p, I) = r(p) + \frac{X(I)}{I} - r^0. \tag{16}$$

Therefore, one can change variable in (13) from p to $-g$.

3.1 The Intuition Behind Proposition 1

To better understand the intuition behind the optimality condition (13), it is useful to recall each mathematical step in the derivation intuitively:

The first step is the change in variable from t to I. Integrating over time means summing the flow costs at each point in time. Integrating over I means summing the flow cost for each reduction in I. The policymaker would like to reduce new infections from I_0 to 0. As a consequence of the extinction threshold, the policymaker only needs to reduce infections to $I_\epsilon > 0$. It is useful to think about the reduction as of a distance to cover. In particular, partition the distance into many small and constant intervals ΔI. The minimization problem consists in minimizing the cost for each of these intervals. The cost to reduce new infections at I to $I - \Delta I$ depends on the flow cost and the time it takes to cross the interval:

$$\underbrace{\big(c(p) + vI(r^0 - r(p) + \theta)\big)}_{flow\ cost} \times \underbrace{\Delta t(I,p)}_{crossing\ time} . \tag{17}$$

Note that the crossing time is a function of p and I:

$$\Delta t = \frac{\Delta I}{(r(p) - r^0)I + X(I)}. \tag{18}$$

The flow cost is increasing in the intensity of applied policies p while the crossing time is decreasing p. Note that for small I, vI is only marginally relevant. The following corollary summarizes this finding:

Corollary 3. .
The key trade-off for suppressing the disease is between the cost of social distancing, $c(p)$, and the time it needs to stay in place, $\Delta t(p)$. The optimal policy trades off these two margins at every instant of time.

To find the optimal policy $p(I)$, take the logarithm of the above expression and perturb the current policy p by a small amount Δp to derive a change in cost ΔC:

$$\Delta C = \left(\frac{c'(p) - vIr'(p)}{c(p) + vI(r^0 - r(p) + \theta)} + \frac{\frac{\partial \Delta t(I,p)}{\partial p}}{\Delta t(I,p)} \right) \Delta p. \tag{19}$$

A policy is optimal if there exits no policy perturbation that reduces the cost. It is the case when the expression in brackets is equal to zero.

Instead of using Δt in the condition above, it is possible to express the same condition as a function of the viral growth rate. Define the growth rate g as $\frac{\dot{I}}{I}$. The crossing time Δt is inversely proportional to the growth rate:

$$\Delta t(I, p) = \frac{\Delta I}{I} \frac{1}{-g(I, p)}. \tag{20}$$

Therefore, the change in cost ΔC as a function of the growth rate is:

$$\Delta C = \left(\frac{c'(p) - vIr'(p)}{c(p) + vI(r^0 - r(p) + \theta)} - \frac{\frac{\partial -g(I,p)}{\partial p}}{-g(I, p)} \right) \Delta p. \tag{21}$$

Using the definition of the growth rate gives the expression in Proposition 1:

$$\frac{c'(p) - vIr'(p)}{c(p) + vI(r^0 - r(p) + \theta)} = \frac{r'(p)}{r(p) + \frac{X(I)}{I} - r^0}. \tag{22}$$

3.2 The Policy Implications of Corollary 1

Two simple and observable sufficient statistics characterize the optimal policy: the current flow-cost and the current viral growth rate. The policymaker only needs to consider the relative change of these two observables to a change in policy to evaluate the current policy's optimality. Optimality solely depends on current variables, which is somewhat surprising. The problem is a dynamic optimization problem, and, to be optimal, a decision at a certain point in time needs to account for its effects on the whole future. However, the two observables summarize all relevant dynamic information.

The optimality condition gives specific guidance to organize a de-confinement after an extended lock-down. For relaxing a certain confinement measure, the policymaker only needs to evaluate its relative impact on the current social cost and viral growth rate. If the relative reduction in cost is larger than the relative increase in growth, a measure should be relaxed. For instance, a policymaker may want to evaluate reopening a particular sector of the economy, such as construction. To make an optimal decision, the policymaker only needs information on how many percentage points such a measure would ease the current cost of the confinement and how many percentage points it would increase the virus's current growth rate.

Note that how to reopen, which is which policy to reverse first, is determined by the ratio $\frac{dc(p)}{dr(p)}$. While the question of which policy to reverse first is an empirical problem, theoretically, it is not very difficult to answer. Policies with a high ratio should be relaxed first. The harder theoretical question is how fast to reopen, which is determined by the above optimality condition. The optimality condition is robust to complementaries between policies, both in cost and in growth impact. The optimal decision only depends on the marginal impact of the most efficient policy at a certain point in time.

3.3 The Properties of the Optimal Policy

To derive the optimal policy's properties, it is simpler to use r as a control variable instead of p. Note that such a change in the variable is without loss of generality. The optimal policy is characterized by a function $r(I)$.

Proposition 2. .

1. *In the optimum, social distancing measures are always positive, and increasing in the number of infectious. Social distancing is the largest at the beginning when $I = I_0$. It is gradually released, when the number of infectious decreases:*

$$r(I) > 0, \textit{ for all } I > 0, \textit{ and } r'(I) > 0. \tag{23}$$

2. *In the limit, at I_ϵ, the optimal policy $r(I_\epsilon)$ has the following properties:*
   - *If $\xi_0 \geq r^0$, social distancing goes to zero:* $\lim_{I_\epsilon \to 0} r(I_\epsilon) = 0$;
   - *If $\xi_0 < r^0$, social distancing goes to a constant:* $\lim_{I_\epsilon \to 0} r(I_\epsilon) = 2(r^0 - \xi_0) > 0$.

3. *Under the optimal policy, the growth rate of I is negative: $g(I) < 0$. In the limit it is equal to* $\lim_{I_\epsilon \to 0} g(I_\epsilon) = -|\xi_0 - r^0|$. *In particular, the growth rate goes to $-\infty$ if $\xi_0 = \infty$.*

The proof is in the appendix. Note that for the case $\xi_0 < r^0$, I assume a quadratic cost. The proposition underlines the crucial role of ξ_0, i.e., the rate of detection at zero. It governs the amount of time it takes to suppress the virus and the optimal policy in the limit. If detection is efficient enough, social contacts go gradually back to the pre-crisis level. The same is not true, when case detection is inefficient. Note that the efficiency of detection is characterized by the derivative of the flow of detections in zero. With inefficient detection, some amount of social distancing needs to stay in place until the extinction limit I_ϵ is reached. If this limit is infinitely small, suppression takes infinitely long. However, the efficiency of detection still matters in this case. It determines the level of necessary social distancing in the limit. The necessary level may consist of mild measures such as washing hands, wearing masks, and forbidding mass events. In the next step, I study the cost of suppression at the optimum.

3.4 The Cost of the Optimal Policy

In the optimum, the total cost of suppressing a mass I_0 of infectious is

$$C = \int_{I_\epsilon}^{I_0} \frac{c(r(I)) + vI(r^0 - r(I) + \theta)}{\left(r(I) + \frac{X(I)}{I} - r^0\right) I} dI, \tag{24}$$

where $r(I)$ denotes the optimal policy. The unit cost of suppression, intuitively, the cost to suppress one more infectious, is equal to

$$\frac{dC}{dI} = \frac{c(r(I))}{\left(r(I) + \frac{X(I)}{I} - r^0\right) I} + \frac{v(r^0 - r(I) + \theta)}{\left(r(I) + \frac{X(I)}{I} - r^0\right)}. \tag{25}$$

It consists of two parts: an economic unit cost, the first summand, which comes from the costs of the taken suppression policies, and a social unit cost, the second summand, which comes from the health costs.

Proposition 3. .
Case 1, $\xi_0 > r^0$:

- *As I converges to zero, the economic unit cost of suppression converges to zero, and the social unit cost of suppression converges to a constant. In particular, if $\xi_0 = \infty$, also the social unit cost converges to zero.*
- *The total cost of suppression is bounded, even if $I_\epsilon = 0$.*

Case 2, $\xi_0 < r^0$:

- *As I converges to zero, the economic unit cost of suppression converges to infinity, and the social unit cost of suppression converges to a constant.*

- *If the extinction limit I_ϵ goes to zero, the total cost of suppression goes to infinity.*

The case $\xi_0 < r^0$ assumes a quadratic cost. The proposition underlines the importance of the properties of case detection when I goes to zero. If the detection rate is high enough, social contacts gradually go back to normal, which bounds the total cost. If the rate is not high enough, the total cost goes to infinity if I_ϵ goes to zero. When $\xi_0 < r^0$, optimal social distancing does not go to zero in the limit; therefore, its cost does not go to zero. On top of that, the time to suppress the virus goes to infinity if I_ϵ goes to zero. It follows that the total cost of suppression goes to infinity if I_ϵ goes to zero. However, this does not mean suppression is not a good idea. The necessary long-run measures may be very mild, and therefore worth enduring. Even if $\xi_0 < r^0$, its size still matters, because it determines the amount of social distancing necessary in the long run. It may still be cheaper to suppress the virus than to use another solution, such as herd immunity. Especially, suppression avoids the risk that the virus mutates and becomes endemic. I further discuss these issues in the conclusion.

3.5 The Problem of Using Social Distancing Alone

Assume the policymaker would like to minimize only the total economic cost of suppressing the virus. Additionally, assume the only available tool to do so is social distancing:

$$C = \min_{r(.)} \int_{I_0}^{I_\epsilon} \frac{c(r(I))}{\dot{I}(I)} dI, \tag{26}$$

$$\text{where } \dot{I} = (r^0 - r(I))I. \tag{27}$$

Corollary 4. .

1. *When only using social distancing, the optimal cost-minimizing policy is constant over time.*

2. *The optimal policy r is equal to*

$$\frac{c'(r)}{c(r)} = \frac{1}{r - r^0}. \tag{28}$$

3. *Assume $c(r)$ is iso elastic. It follows that the optimal effect of social distancing r is*

equal to

$$r = r^0 \frac{\zeta_1}{\zeta_1 - 1}, \tag{29}$$

where $\zeta_1 > 1$ is the cost-elasticity.

4. *The optimal economic unit cost of reducing an infection goes to infinity as I goes to zero.*

Points one to three follow as special cases from Proposition 1. Point 4 follows as a special case from Proposition 3. In particular, point 4 shows the problem of using social distancing alone to suppress the virus. Even in the optimum, the cost-efficiency of social distancing measures decreases as I decreases, because the reduction in infectious becomes infinitely slow when I goes to zero. This result is quite intuitive. Given a certain intensity of social distancing, it takes the same time to reduce infections from 10 million to 1 million as reducing them from 10 to 1. Suppressing the virus by social distancing is possible, but very costly. One may be tempted to think that point 4 of Corollary 4 is not relevant in practice. The last unit to reduce is at $I = I_\epsilon$ and not at $I = 0$. However, I_ϵ is very small. Therefore, point 4 shows that the costs of reducing the last units close to I_ϵ are very high. Note that the discussed policy is the cost-minimizing policy in an economic sense. When maximizing social welfare, as discussed above, the result becomes even more extreme.

4 Quantitative Results

The results discussed so far are theoretical and robust to the huge parameter uncertainty related to Covid-19. However, they are unable to answer two crucial questions. What are the economic and health costs of the optimal suppression policy under different realistic detection technologies? And, how do these costs compare to the costs of optimal mitigation policies? To calculate the total cost of suppression, it is necessary to know the functions $c(p)$, $r(p)$ and $X(I)$, as well as the parameters r^0 and v. Therefore, due to the high parameter uncertainty, a precise answer to these questions is impossible. Following the literature, I use a calibration exercise to obtain approximate answers.

4.1 Data Sources

I use frequently updated epidemiological data from the Institute for Health Metrics and Evaluation (IHME) at the University of Washington.[5] They provide a time series of confirmed cases and estimates for the real number of daily infections for many countries. Their estimates are based on Murray et al. (2020). I use data from Italy and South Korea.

4.2 Calibration

4.2.1 Parameters Literature

I use the following parameters from the literature as a starting point for my calibration:

Parameter	Symbol	Value	Source
Mortality rate	δ	0.01	Alvarez et al. (2020)
Time of contagiousness	$\frac{1}{\theta}$	5 days	Fernández-Villaverde and Jones (2020)
Value of statistical life	vsl	$20 \frac{\text{GDP}}{\text{capita}}$	Alvarez et al. (2020)
Uncontrolled growth rate	r^0	0.14	Ferretti et al. (2020)
Max. rel. speed digital tracing	ξ_0^d	0.35	Ferretti et al. (2020)
Max. rel. speed manual tracing	ξ_0^m	0.1	Ferretti et al. (2020)
GDP loss strict lock-down	c_{LD}	0.5	Gollier (2020)

Table 1: Parameters Literature

4.2.2 The Cost Function

I use a direct relation between the cost of social distancing, measured as lost GDP, and the reduction in the viral viral growth rate r. I assume the function is iso-elastic:

$$c(r) = \zeta_0 r^{\zeta_1}, \tag{30}$$

where $\zeta_0 > 0$ and $\zeta_1 > 1$. Neglecting the value of lives lost as well as tracing, by Corollary 4, it holds that the optimal r solves

$$\zeta_1 = \frac{r}{r - r^0}. \tag{31}$$

[5] I downloaded the data from http://www.healthdata.org on May 12th, 2020.

From March 10th until April 26th, the Italian government imposed a nationwide lock-down. A lock-down is a strict form of social distancing. Any no-essential social contacts are forbidden. A large part of the population is forced to stay at home. Going outside is permitted only if essential. To calibrate ζ_1, I assume the strict lock-down in Italy was close to optimal. Note that Italy did not use much tracing during the time of the lock-down. The value of lives lost is small compared to the lost GDP. The assumption of optimality is certainly strong. However, in many countries such as France, Spain, the UK, and Germany, we have seen very similar lock-downs intensities. This is consistent with Equation (31). Note that the optimal intensity of r does not depend on the level of infections. It only depends on ζ_1, which parametrizes the convexity of the cost. Once a country discovers an outbreak, it should hit hard to reduce new infections. If tracing is infeasible in the short term, and the number of death is relatively small, Equation (31) is a good approximation for the optimal policy. The intensity of r does only depend on the convexity of the cost ζ_1. Note that the convexity should be similar between countries. The more convex the cost $c(.)$, the more it costs to implement a strict lock-down. The similarly intense lock-downs between countries suggest that governments approximately followed the optimal lock-down strategy. A different interpretation of the optimality assumption is that it makes the simulated costs consistent with the strict lock-down. Assume the government acted optimally, given its best estimate of the cost function. The implied simulated policies and costs are consistent with the estimate.

Under the assumption that the intensity of the lock-down was optimal, it is informative about the cost function's convexity. Using the epidemiological data from Murray et al. (2020), I estimate the growth rate under the Italian lock-down at $g_{LD} = -0.032$. I use the estimated number of new infections form the peak on March 11th until the most recent estimates on May 10th. Together with an uncontrolled growth rate of $r^0 = 0.14$ (see Ferretti et al. (2020)), I estimated growth reduction during the lock-down is $r_{LD} = r^0 - g_{LD} = 0.172$. Using Equation (31), it implies an elasticity of $\zeta_1 \approx 5$. Following Gollier (2020), I assume a strict lock-down implies a daily GDP loss of around $c_{LD} = 50\%$.[6] The implied $\zeta_0 \approx 3300$.

[6] The number in Gollier (2020) is based on GDP estimates from the "National Institute of Statistics and Economic Studies" in France (https://www.insee.fr/fr/statistiques/4485632).

4.2.3 The Case Detection Function

I use the following case detection function:

$$X'(I) = \left(\xi_0^{-\frac{1}{\alpha}} + \xi_1 I\right)^{-\alpha}, \tag{32}$$

and $X(0) = 0$. The function fulfills the necessary properties of a detection function specified above. $\xi_0 > 0$ is the value of the function for $I = 0$. Note that it is equal to $\lim_{I\to 0} \frac{X(I)}{I}$, i.e., the relative speed of tracing at zero. The parameter $\xi_1 > 0$ controls the function's behavior for large values of I. $\alpha \geq 0$ controls how fast $X'(I)$ goes from ξ_0 to its behavior for large I. This function is quite general and contains some intuitive tracing functions as special cases. For $\alpha = 0$ it reduces to a constant returns to scale tracing function: $X(I) = \xi_0 I$. In particular, if ξ_0 is equal to the daily flow of tests, it is equal to tracing under random testing. When ξ_0 goes to infinity, the function reduces to a power function as used in Alvarez et al. (2020). The disadvantage of a power function is that $X'(0) = \infty$ by assumption. This assumption is unrealistic. It makes tracing overly efficient at the end of suppression.

To calibrate the parameters, I distinguish two cases: digital tracing and manual tracing. I use micro estimates to calibrate the function for both cases. I use results from Ferretti et al. (2020). This epidemiological paper estimates how much optimal contact tracing can reduce daily new infections. They compare digital contact tracing with manual contact tracing. Ferretti et al. (2020) estimate that, under optimal conditions, digital contact tracing can find infectious individuals at a rate of $\xi_0^d = 35\%$ per day. It means that the stock of currently infectious can be reduced by 35% in one day. Manual contact tracing is much slower. The authors argue that optimal manual contact tracing achieves a rate of $\xi_0^m = 10\%$ per day because of unavoidable delays. I use these estimates as values for ξ_0 in the two cases. I assume that, if a country uses its full resources to find the last cases, tracing achieves its optimal rate. However, as soon as the caseload grows, the system becomes overwhelmed, and the efficiency of tracing decreases.

I calibrate ξ_1 such that at a prevalence level of 10%, i.e., 10% of the population is infected at the same time, $X'(I) = \xi_0/10000$, which is close to zero. It means that at a prevalence level of 10%, the system is so overloaded that any further increase in the number of infected will not lead to more traced cases. To calibrate α, I use estimates of the fraction of traced cases from Italy and South Korea. I use data from Murray et al. (2020), and I use the date with the maximal detection rate in both countries. I estimate that Korea, using digital tracing, at a prevalence of 56 infected per million, found 21 % of total cases daily. I assume the number of confirmed cases is equal to the number of traced cases. Note that the estimated rate is not too far from the theoretical limit of 35%. It implies that, for digital tracing, $\alpha_d = 1.2$. For manual tracing, I use the same procedure using data from Italy. On May 10th, at an estimated prevalence of 700 per million, Italy manages to confirm 3% of the total cases daily. It implies that, for manual tracing, $\alpha_m = 1.5$.

4.2.4 Remaining Parameter Values

As already mentioned, I use $r^0 = 0.14$ as in Ferretti et al. (2020).

To estimate the current prevalence in Italy, I use the data from Murray et al. (2020). I use May 10th as a starting date because it is the last date of available observations. Murray et al. (2020) estimate daily new infections. I use new infections to calculate the current stock of infectious by summing the infections over the 5 preceding days. I assume an infected stays infectious for $1/\theta = 5$ days, following Fernández-Villaverde and Jones (2020). On May 10th, I find a level of prevalence of $I_0 = 0.0007$ and a level of susceptible of $S = 0.96$. It shows that the number of infectious is indeed small compared to the number of susceptible. Therefore, The assumption that the number of susceptible is constant is approximately true. I will confirm the assumption ex-post in Section 4.3.1.

To estimate the health cost, I assume that an infectious dies with probability $\delta = 0.01$, following Alvarez et al. (2020). Following Alvarez et al. (2020), I use a value of statistical life of 20 times the annual output per capita. It follows that $v = vsl * 365 * \delta = 73$. Note that my results are very insensitive to the assumptions related to mortality because of the low prevalence level.

4.2.5 Summary Relevant Parameters

Parameter	Symbol	Value	Matched Moment or Source
Factor cost function	ζ_0	3300	GDP loss lock-down
Cost-elasticity	ζ_1	5	Lock-down intensity
Max. rel. speed digital tracing	ξ_0^d	0.35	Ferretti et al. (2020)
Max. rel. speed manual tracing	ξ_0^m	0.10	Ferretti et al. (2020)
Scalability digital tracing	α_d	1.2	Confirmed cases Korea
Scalability manual tracing	α_m	1.5	Confirmed cases Italy
Initial Prevalence	I_0	0.07%	Estimate for Italy May 11th
Flow value of casualties	v	73	Alvarez et al. (2020)
Uncontrolled growth rate	r^0	0.14	Ferretti et al. (2020)

Table 2: Relevant Parameters

4.3 Results

I take the current prevalence in Italy as given and analyze the optimal suppression policy for three different tracing scenarios:

1. Italy continues to isolate infectious at the current rate of 3% per day. I refer to this case as no tracing.
2. Italy adopts an optimal manual contact tracing strategy.
3. Italy adopts an optimal digital contact tracing strategy like South Korea.

To compare the three scenarios, I examine the intensity of the optimal social distancing measures, their implied flow costs, the time it takes to reach certain thresholds in daily cost, and the total cost.

Cost or intensity measure	No tracing	Manual tracing	Digital tracing
Optimal intensity r on May 10th	0.14	0.14	0.12
Implied daily cost [daily GDP]	19 %	19 %	8.7%
Time until cost drops to 1% of daily GDP	never	3.4 months	39 days
Daily cost in the limit	16%	0.1%	0
Time to reach extinction threshold 1 ppm	7.8 months	15 months	3.0 months
Total cost until extinction [annual GDP]	11 %	1.7%	0.44%
Total additional death until extinction	2,900	4,200	3,300

Table 3: Comparison Scenarios

The optimal reduction in social distancing intensity is such that r goes from 0.17 under the lock-down, to 0.14 (no tracing/manual tracing), and 0.12 (digital tracing). Some degree of easing is optimal. The reason is that identifying a fraction of the contagious takes over some of the burdens to keep viral growth at an optimal level. This modest reduction in social distancing already has an essential impact on economic cost. It reduces from 50% of daily GDP under the lock-down, to 19 % under no tracing/manual tracing, and 8.7% under digital tracing. Said differently, on May 10th, it is possible to ease the lock-down by around half, measured in lost daily GDP. An immediate switch to the Korean strategy would allow for an easing of a factor of almost 6.

The cost of social distancing drops over time because the number of infectious reduces and social distancing is gradually relaxed in the optimum. Under digital tracing, the cost drops below 1% of daily GDP after 39 days already. Under manual tracing, it takes 3.4 months to reach this point. However, under no tracing, this point is never reached. The optimal intensity is almost constant and stays close to 0.14. A cost close to 19% of daily GDP needs to be paid until the virus disappears. In the long run, the daily cost reduces to 0.1% for manual tracing, and 0% for digital tracing. The numerical results confirm the theoretical results. Only efficient tracing with $\xi_0 > r^0$ allows the society to go back to a normal activity level. If tracing is inefficient, i.e., $\xi_0 << r^0$ relatively strong and costly social distancing measures need to stay in place until the virus disappears. Mild efficiency implies that measures have to stay in place in the long run; however, they are mildly intense and not very costly. For instance, this may correspond to the case where society only imposes restrictions on mass events and general hygiene measures such as mask-wearing.

Next, to compare the different strategies' total cost, following Piguillem and Shi (2020), I assume the virus disappears when prevalence falls below an extinction threshold of 1 infectious per million inhabitants. Piguillem and Shi (2020) use a threshold of 10 per million. I use a more conservative threshold because South Korea already reached a prevalence of 6 per million, and the virus did not get extinct.

The differences in cost between the different strategies are enormous. No tracing takes 7.8 months and costs 11% of annual GDP. Note that this cost is in addition to the already incurred cost due to the strict lock-down. Also, under this policy, the virus causes around 2,900 additional victims. The suppression with the help of manual tracing takes 15 months.

Manual tracing is slower but much less costly. The reason is that social distancing is gradually relaxed in the optimum. It reaches a limit where its cost is only 0.1% of daily GDP. The total cost is 1.7% of annual GDP, and the number of additional casualties is around 4,200. Note that the number of casualties is higher in the manual tracing scenario than in the no tracing scenario. The health cost has only a very small influence on the optimal policy. Because manual tracing is cheaper, the optimal policy takes longer. Additionally, more control is exerted with the help of tracing. Tracing does not avoid new infections, because individuals are traced after they got infected. Both factors contribute to a higher number of casualties. The total cost under manual tracing is still substantial. The cheapest option is digital tracing. The virus disappears in 3 months. Social distancing is relaxed quickly and substantially, well before that date. The total cost is only 0.44 % of annual GDP. The number of additional casualties is around 3,300. For all strategies, after extinction, social distancing is completely relaxed. In case there is an import of new cases, the pandemic restarts under no tracing and manual tracing, but not under digital tracing. Digital tracing alone can keep small new outbreaks under control. I do not count the cost to avoid or control new outbreaks when digital tracing is not used. In case a country decides to suppress the virus using no tracing or optimal manual tracing, the policy needs to be complemented by meticulous border controls until a vaccine arrives. I leave the quantification of these additional costs for future research.

Note that the total cost of suppression with digital tracing is by an order of magnitude smaller than estimates for the total cost under optimal mitigation strategies. Acemoglu et al. (2020) and Gollier (2020) evaluate mitigation strategies with age-depended social distancing measures. Age-depended policies give the most optimistic estimates for total costs and casualties of mitigation policies. They find a total cost of mitigation in the range of 7 to 13 % of annual GDP. The death toll of optimal mitigation strategies in the most optimistic scenarios is 0.2% of the population. In Italy's case, these are around 120,000 casualties, which stand in stark contrast to the 3,300 additional casualties of an optimal suppression policy.

4.3.1 The Evolution of the Number of Susceptible

I find that the change in the mass of susceptible ΔS is small under all three scenarios:

No tracing	Manual tracing	Digital tracing
0.49%	0.69%	0.48%

Table 4: Change in the Number of Susceptible ΔS

The initial mass of susceptible is $S_0 = 96\%$, and therefore, it is indeed very close to constant.

5 Conclusion

This paper characterizes the optimal policy to suppress COVID-19. I find that complete and efficient eradication of COVID-19 is possible at a reasonable economic cost of 0.4% of annual GDP. A simple function of observables, the optimal policy is easily implementable. However, some crucial questions are still unanswered. In particular, is it more efficient to use mitigation or suppression?

Remember, mitigation controls the spread of the virus until contagions stop because the population achieves herd immunity. If the current number of infectious individuals is sufficiently low, and case detection is efficient enough, the answer to this question is undoubtedly suppression. The same is true if the value of lives lost is large enough. However, for all other cases, it becomes much harder to make an optimal decision. Moreover, a decision needs to be made. The two policy responses dictate a very different optimal time path of infections. A mitigation policy lets infections grow because the virus needs to reach a large enough part of the population. Optimal suppression never allows infections to grow. The policymaker stands at a crossroads and needs to decide which path to take. The total cost of either of them is still very uncertain. It depends crucially on: the cost and viral growth impact of social distancing policies, the speed of tracing, especially at low infection levels, the statistical value of life, and the capacity of the health care system and its impact on mortality rates. All of these variables are highly uncertain. Only the precise estimates of the mentioned unknowns can give a definite answer to the question.

However, the calibration exercise in this paper can give rough guidance on how to answer the question. I find that the total cost of suppression is 0.44% of annual GDP when using digital contact tracing and 1.7 % of GDP when using manual tracing. In comparison, the cost-estimates of an optimal mitigation strategy range from 7% (Gollier, 2020), to 7-

14% (Acemoglu et al., 2020), to around 30% (Alvarez et al., 2020). Acemoglu et al. (2020) estimate the additional number of casualties in the most optimistic scenario to around 0.2% of the population. In the case of Italy, these are around 120,000 additional casualties. The estimates in Acemoglu et al. (2020) and Gollier (2020) rely on the optimistic assumption that it is possible to shelter the most vulnerable part of the population. Even in that case, the number of casualties of an optimal mitigation policy stands in stark contrast to the 3,300 additional casualties of an optimal suppression policy. None of these estimates take the risk that the virus could become endemic into account. The comparison suggests that suppression is the most cost-efficient strategy. It is certainly the strategy that reduces the number of casualties.

Curiously, it is easier to find the exact optimal amount of social distancing at each point when following a suppression policy, than to decide on the optimal broad direction of policy. The policymaker can turn to the econometrician and the epidemiologist - they can estimate the local impact of a policy change on the flow cost and the viral growth rate - and apply the condition derived in this paper.

References

ACEMOGLU, D., V. CHERNOZHUKOV, I. WERNING, AND M. D. WHINSTON (2020): "A Multi-Risk SIR Model with Optimally Targeted Lockdown," Tech. rep., National Bureau of Economic Research.

ALVAREZ, F. E., D. ARGENTE, AND F. LIPPI (2020): "A simple planning problem for covid-19 lockdown," *Covid Economics*, 14.

ASSENZA, T., F. COLLARD, M. DUPAIGNE, P. FÈVE, C. HELLWIG, S. KANKANAMGE, N. WERQUIN, ET AL. (2020): "The Hammer and the Dance: Equilibrium and Optimal Policy during a Pandemic Crisis," Tech. rep., Toulouse School of Economics (TSE).

ATKESON, A. (2020): "What will be the economic impact of COVID-19 in the US? Rough estimates of disease scenarios," Tech. rep., National Bureau of Economic Research.

BERGER, D. W., K. F. HERKENHOFF, AND S. MONGEY (2020): "An seir infectious disease model with testing and conditional quarantine," Tech. rep., National Bureau of Economic Research.

BETHUNE, Z. A. AND A. KORINEK (2020): "Covid-19 infection externalities: Trading off lives vs. livelihoods," *Covid Economics*, 11.

CHARI, V. V., R. KIRPALANI, AND C. PHELAN (2020): "The hammer and the scalpel: On the economics of indiscriminate versus targeted isolation policies during pandemics," Tech. rep., National Bureau of Economic Research.

CHETTY, R., J. N. FRIEDMAN, N. HENDREN, M. STEPNER, ET AL. (2020): "How did covid-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data," Tech. rep., National Bureau of Economic Research.

DORN, F., S. KHAILAIE, M. STOECKLI, S. C. BINDER, B. LANGE, S. LAUTENBACHER, A. PEICHL, P. VANELLA, T. WOLLMERSHAEUSER, C. FUEST, ET AL. (2020): "The Common Interests of Health Protection and the Economy: Evidence from Scenario Calculations of COVID-19 Containment Policies," *medRxiv*.

EICHENBAUM, M. S., S. REBELO, AND M. TRABANDT (2020): "The macroeconomics of epidemics," Tech. rep., National Bureau of Economic Research.

FARBOODI, M., G. JAROSCH, AND R. SHIMER (2020): "Internal and external effects of social distancing in a pandemic," *Covid Economics*, 9.

FAVERO, C. A., A. ICHINO, AND A. RUSTICHINI (2020): "Restarting the economy while saving lives under Covid-19," .

FERNÁNDEZ-VILLAVERDE, J. AND C. JONES (2020): "Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities," .

FERRETTI, L., C. WYMANT, M. KENDALL, L. ZHAO, A. NURTAY, L. ABELER-DÖRNER, M. PARKER, D. BONSALL, AND C. FRASER (2020): "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing," *Science*, 368.

GARIBALDI, P., E. R. MOEN, AND C. A. PISSARIDES (2020): "Modelling contacts and transitions in the SIR epidemics model," *Covid Economics*, 5.

GERLAGH, R. (2020): "Closed-form solutions for optimal social distancing in a SIR model of covid-19 suppression," .

GOLLIER, C. (2020): "Cost-benefit analysis of age-specific deconfinement strategiesChristian GollierToulouse School of Economics," Tech. rep., University of Toulouse-Capitole.

GONZALEZ-EIRAS, M. AND D. NIEPELT (2020): "On the Optimal Lockdown During an Epidemic," *Covid Economics*, 7.

HORNSTEIN, A. (2020): "Social Distancing, Quarantine, Contact Tracing, and Testing: Implications of an Augmented SEIR-Mode," Tech. rep., Technical Report, Federal Reserve Bank of Richmond.

JONES, C. J., T. PHILIPPON, AND V. VENKATESWARAN (2020): "Optimal mitigation policies in a pandemic: Social distancing and working from home," Tech. rep., National Bureau of Economic Research.

MICLO, L., D. SPIRO, AND J. WEIBULL (2020): "Optimal epidemic suppression under an ICU constraint," *arXiv preprint arXiv:2005.01327*.

MURRAY, C. J. ET AL. (2020): "Forecasting the impact of the first wave of the COVID-19 pandemic on hospital demand and deaths for the USA and European Economic Area countries," *medRxiv*.

OBIOLS-HOMS, F. (2020): “Precaution, social distancing and tests in a model of epidemic disease,” *Covid Economics*, 91.

PIGUILLEM, F. AND L. SHI (2020): “Optimal covid-19 quarantine and testing policies,” .

PUEYO, T. (2020): “Coronavirus: The Hammer and the Dance.” Tech. rep., blog post, Medium. URL https: //medium.com/@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56.

RACHEL, L. (2020): “An Analytical Model of Covid-19 Lockdowns: Equilibrium Mitigation Flattens the Epidemic Curve, Optimal Lockdown Does Not,” Tech. rep.

SCHERBINA, A. (2020): “Determining the optimal duration of the COVID-19 suppression policy: A cost-benefit analysis,” *Available at SSRN 3562053*.

UGAROV, A. (2020): “Inclusive Costs of NPI Measures for COVID-19 Pandemic: Three Approaches,” *medRxiv*.

WANG, Y. (2020): “An Analytical SIR model of Epidemics and A Sustainable Suppression Policy: Testing,” *Available at SSRN 3573979.*

A Appendix

A.1 The SIR Model

Assume the population has a mass of one and consists of three groups. I_t denotes the fraction of infectious individuals at time t. S_t denotes the fraction of the population susceptible to getting infected by the disease when meeting an infectious. R_t denotes the number of recovered individuals. Infectious and susceptible meet randomly, and the disease transmits at a certain rate. When there are no active control measures, the infection rate is β^0. The recovery rate, once infected, is θ. The initial fraction of infectious and susceptible is denoted by I_0 and S_0. Three differential equations describe the dynamics of the three groups:

$$\dot{I}_t = (\beta^0 S_t - \theta) I_t; \tag{33}$$

$$\dot{S}_t = -\beta^0 S_t I_t; \tag{34}$$

$$\dot{R}_t = \theta I_t, \tag{35}$$

where the dot superscript denotes the time derivative. When I_t is small compared to S_t, then $\dot{S}_t$ is small compared to S_t. As a consequence, S_t is approximately constant: $S_t \approx S$. The dynamic behavior of the infectious is approximately described by Equation (1):

$$\dot{I} = r^0 I_t,$$

where $r^0 = \beta^0 S - \theta$.

Note that I_t is small compared to S_t at the beginning of the pandemic, as well as after an extended period of effective control measures. Even if part of the population is immune, as long as I_t is small compared to the number of susceptible, the above approximation is valid. The lower fraction of susceptible is captured by a lower r^0.

A.1.1 The Detection Technology

$\tilde{X}(I, z)$ denotes the flow of detected cases. It is a production function with two inputs: the number of undetected cases I and the amount of resources allocated to case detection z. I assume $\tilde{X}(I, z)$ fulfills the following standard properties of a production function: $\tilde{X}(0, z) = 0$, $\tilde{X}(I, 0) = 0$, $\frac{\partial \tilde{X}(I,z)}{\partial I} > 0$, $\frac{\partial \tilde{X}(I,z)}{\partial z} > 0$, $\frac{\partial^2 \tilde{X}(I,z)}{\partial I^2} < 0$, $\frac{\partial^2 \tilde{X}(I,z)}{\partial z^2} < 0$, $\lim_{I \to \infty} \frac{\partial \tilde{X}(I,z)}{\partial I} = 0$. In particular, the concavity in I comes from the fact that, if resources are fixed, the resources per undetected case I decrease as I increases. An additional detection is carried out with fewer resources. Therefore, it is slower, and $\frac{\partial^2 \tilde{X}(I,z)}{\partial I^2} < 0$.

Assume the overall resources a country can allocate to case detection are fixed and constant: $z = \hat{z}$. Define $X(I) = \tilde{X}(I, \hat{z})$. It follows that $X(0) = 0, X'(I) > 0, X''(I) < 0$ and $\lim_{I\to\infty} X'(I) = 0$.

Lemma 2. *If $X''(I) < 0$ and $X'(I) > 0$ for all I, it follows that $\frac{X(I)}{I}$ is decreasing for all I.*

PROOF:
If $X''(.) < 0$ it follows that

$$\int_0^I (X'(\tilde{I}) - X'(I))d\tilde{I} > 0, \tag{36}$$

because $X'(.)$ is a decreasing and positive function. It follows that $X(I) - X'(I)I > 0$, which implies

$$\frac{d\,\frac{X(I)}{I}}{d\,I} = \frac{X'(I)I - X(I)}{I^2} < 0. \tag{37}$$

qed.

A.1.2 The Derivation of Equation (2) and (6) from the SIR Model

The mass of not quarantined infectious is I_t. Denote by J_t the overall mass of infected and by Q_t the mass of quarantined. It follows that $J_t = Q_t + I_t$ and $\dot{J}_t = \dot{Q}_t + \dot{I}_t$. The flow of new infections follows

$$\dot{J}_t = (\beta_0 S - \beta(p_t)S)I_t - \theta J_t \tag{38}$$

$$\dot{Q}_t = X(I_t) - \theta Q_t \tag{39}$$

$\beta(p_t)$ is the reduction in viral transmission due to application of policy p_t, where $p_t \in [0, 1]$. Denote $\beta_0 S - \theta = r^0$ and $\beta(p_t)S = r(p_t)$ to get

$$\dot{I}_t = (r^0 - r(p_t))I_t - X(I_t),$$

which gives Equation (2) when $p_t = 0$ and Equation (6) when detection is not used and $X(I) = 0$.

Note that, by Equation (38), the flow of new infections is $\beta_0 S - \beta(p_t)S$, which, by using the according definitions, is equal to $r^0 - r(p_t) + \theta$.

PROOF of Lemma 1:
$\frac{X(I)}{I}$ is strictly decreasing from ξ_0 to 0. It follows that there exits an I^* such that $\frac{X(I)}{I} = r^0$. For all $I < I^*$, $r^0 - \frac{X(I)}{I} < 0$. Therefore, $\dot{I} = \left(r^0 - \frac{X(I)}{I}\right) I < 0$.
qed.

A.2 Proofs Section 3

A.2.1 Proof Proposition 1

The minimum of the integral

$$\min_{p(.)} C(p(.)) = \int_{I_\epsilon}^{I_0} -\frac{c(p(I)) + vI(r^0 - r(p(I)) + \theta)}{(r^0 - r(p(I)))I - X(I)} dI \tag{40}$$

is at the point-wise minimum of each integrand. Note that I swapped the bounds. Change policy variable from p to r. For each I, the integrand is equal to

$$\frac{c(r) + vI(r^0 - r + \theta)}{\left(r - r^0 + \frac{X(I)}{I}\right) I}. \tag{41}$$

Note that $\dot{I} < 0$ by assumption. Therefore, the denominator has to be positive, which is the case when $r > r^0 - \frac{X(I)}{I}$. Also, $r \geq 0$ by definition. Note that $r \leq r^0 + \theta$. At this bound social contacts are zero. Assume that $c(r^0 + \theta) = \infty$. There are two cases:
First, if $r^0 - \frac{X(I)}{I} > 0$ it holds that $r \in \left(r^0 - \frac{X(I)}{I}, r^0 + \theta\right]$.
Second, if if $r^0 - \frac{X(I)}{I} \leq 0$ it holds that $r \in [0, r^0 + \theta]$.

Note that the integrand is finite, positive, and continuous for any interior r.

Lemma 3. *There exists a minimum of the integrand, and it is interior.*

PROOF:

Case 1, $r^0 - \frac{X(I)}{I} > 0$:
It follows that $r \in \left(r^0 - \frac{X(I)}{I}, r^0 + \theta\right]$. If r goes to the left limit, the integrand goes to infinity. If r goes to the right limit, the integrand goes to infinity as well. The integrand is finite, positive, and continuous for any interior r. It follows that there exists an interior minimum.

Case 2, $r^0 - \frac{X(I)}{I} < 0$:
It follows that $r \in [0, r^0 + \theta]$. At the left boundary, the integrand is equal to $\frac{v(r^0+\theta)}{\frac{X(I)}{I} - r^0}$. The minimum cannot be at zero, because the integrand is strictly decreasing in zero:

$$\frac{(c'(0) - vI)\left(0 - r^0 + \frac{X(I)}{I}\right) I - (c(0) + vI(r^0 - 0 + \theta))\, I}{\left(0 - r^0 + \frac{X(I)}{I}\right)^2 I^2} = \frac{-v\left(\frac{X(I)}{I} + \theta\right)}{\left(-r^0 + \frac{X(I)}{I}\right)^2} < 0. \tag{42}$$

If r goes to the right limit, the integrand goes to infinity. The integrand is finite, positive, and continuous for all r. It follows that there exists an interior minimum.

Case 3, $r^0 - \frac{X(I)}{I} = 0$:
It follows that $r \in [0, r^0 + \theta]$. As case 1. The integrand goes to infinity at both boundaries.

qed.

Any interior extremum fulfills the first order condition:

$$\frac{c(r) + vI(r^0 - r + \theta)}{\left(r - r^0 + \frac{X(I)}{I}\right) I} \left(\frac{c'(r) - vI}{c(r) + vI(r^0 - r + \theta)} - \frac{1}{r + \frac{X(I)}{I} - r^0}\right) = 0 \tag{43}$$

Cancel the left factor and change the choice variable back from r to p to get the optimality condition in Proposition 1.

Each interior extremum is a strict minimum. To see that, rearrange the first derivative of the integrand to:

$$\frac{1}{\left(r - r^0 + \frac{X(I)}{I}\right)^2 I} \left(-c(r) - vI(r^0 - r + \theta) + (c'(r) - vI)\left(r + \frac{X(I)}{I} - r^0\right)\right) \quad (44)$$

Take the derivative to get the second order condition. Note that it is equal to $a'b + ab'$ where a is the first factor above and b is the second factor. If the FOC holds, b is zero. Also, a is always positive. The sign of the SOC only depends on the sign of b'. $b' = c''(r)\left(r + \frac{X(I)}{I} - r^0\right)$, which is strictly greater than zero.

As each minimum is a strict minimum, and the function is continuous, there can only be one minimum. In particular, it fulfills the first order condition. As guessed, $\dot{I}_t < 0$ for all t and $\lim_{t\to\infty} I_t = 0$.

There can be no minimum where I_t does not converge to zero. In that case, the integral does not exist because the integrated cost becomes infinitely large. $\dot{I}_t \geq$ cannot be optimal because growing infections increase the health cost $vI(r^0 - r + \theta)$ and only transfers the cost of reducing infections to a later point in time.

qed.

A.2.2 Proof Proposition 2

Lemma 4. .
Consider the case where $\lim_{I\to 0} \frac{X(I)}{I} \geq r^0$. *It follows that:*
1) The optimal policy $r(I)$ *converges to zero as* I *converges to zero:*

$$\lim_{I\to 0} r(I) = 0. \quad (45)$$

2) For small I *the optimal policy* $r(I)$ *is approximately equal to*

$$r(I) \approx -\left(\frac{X(I)}{I} - r^0\right) + \sqrt{\left(\frac{X(I)}{I} - r^0\right)^2 + 2\frac{v}{c''(0)} I \left(\frac{X(I)}{I} + \theta\right)}. \quad (46)$$

In particular, $r(I) > 0$ *for* $I > 0$.

3) For small I *the growth rate under the optimal policy* $g(I)$ *is approximately equal to*

$$g(I) \approx -\sqrt{\left(r^0 - \frac{X(I)}{I}\right)^2 + 2\frac{v}{c''(0)} I \left(\frac{X(I)}{I} + \theta\right)}. \tag{47}$$

4) Under the optimal policy $r(I)$ *the growth rate converges to*

$$\lim_{I \to 0} g(I) = -(x^0 - r^0). \tag{48}$$

In particular, for

$$\lim_{I \to 0} \frac{X(I)}{I} = \infty, \text{ it holds that } \lim_{I \to 0} g(I) = -\infty. \tag{49}$$

The decay of the virus is accelerating as I *approaches zero.*

PROOF:

The optimality condition with r as the policy variable writes

$$\frac{c'(r) - vI}{c(r) + vI(r^0 - r + \theta)} = \frac{1}{r + \frac{X(I)}{I} - r^0}. \tag{50}$$

Taylor approximate the function $c(r)$ in the origin:

$$c(r) \approx \frac{1}{2} c''(0) r^2. \tag{51}$$

Use the approximation in the optimality condition to solve for Equation (46), which proofs point 2). Note that $r(I)$ is the solution of a quadratic equation. The second solution can be discarded as it violates $\dot{I} < 0$. Point 1) follows from taking the limit in Equation (46). Point 3) follows from using the definition of the growth rate. Point 4) follows from taking the limit in Equation (47).

qed.

Lemma 5. .

Consider the case where $\lim_{I \to 0} \frac{X(I)}{I} = \xi_0 < r^0$. *Assume that the cost function is quadratic:* $c(r) = \frac{1}{2} c''(0) r^2$ *It follows that:*

1) As I converges to zero, the optimal policy $r(I)$ converges to:

$$\lim_{I\to 0} r(I) = 2(r^0 - \xi_0). \tag{52}$$

In particular, if there is no test and trace $\xi_0 = 0$, and

$$\lim_{I\to 0} r(I) = 2r^0. \tag{53}$$

3) The optimal policy $r(I)$ is equal to

$$r(I) = r^0 - \frac{X(I)}{I} + \sqrt{\left(r^0 - \frac{X(I)}{I}\right)^2 + 2\frac{v}{c''(0)} I \left(\frac{X(I)}{I} + \theta\right)}. \tag{54}$$

3) The implied optimal growth rate $g(I)$ is equal to

$$g(I) = -\sqrt{\left(r^0 - \frac{X(I)}{I}\right)^2 + 2\frac{v}{c''(0)} I \left(\frac{X(I)}{I} + \theta\right)}. \tag{55}$$

In particular $r(I) > 0$ for all I.
4) Under the optimal policy $r(I)$ the growth rate converges to

$$\lim_{I\to 0} g(I) = -(r^0 - x^0). \tag{56}$$

PROOF:
As above. However, the cost function is quadratic by assumption and not by approximation.
qed.

Lemma 6. .
The optimal policy $r(I)$ is strictly increasing in I:

$$r'(I) > 0. \tag{57}$$

PROOF:

The optimal policy solves

$$\frac{c'(r(I)) - vI}{c(r(I)) + vI(r^0 - r(I) + \theta)} = \frac{1}{r(I) + \frac{X(I)}{I} - r^0}. \tag{58}$$

Differentiate with respect to I to obtain

$$r'(I) = \frac{v\left(\theta + \frac{X(I)}{I}\right) - (c'(r(I)) - vI)\frac{d\frac{X(I)}{I}}{dI}}{-g(r(I), I)c''(r(I))}. \tag{59}$$

The expression is positive because $v > 0$, $\frac{d\frac{X(I)}{I}}{dI} < 0$, $g(r(I), I) < 0$, $c''(.) > 0$, and $c'(r(I)) - vI > 0$. The last inequality follows from the FOC.
qed.

Point 1 of Proposition 2 follows directly from Proposition 1 and Lemma 6. Point 2 of Proposition 2 follows from point 1 in Lemma 4 and 5. Point 3 of Proposition 2 follows directly from Proposition 1 and from point 4 in Lemma 4 and 5.
qed.

A.2.3 Proof Proposition 3

Lemma 7. .
If $\xi_0 = \infty$, the total unit cost of suppression goes to zero as the mass of infectious goes to zero.

PROOF:
$\frac{dC}{dI}$ is the unit cost in the optimum. It is smaller or equal to the unit cost under any other policy that satisfies $\dot{I}(I) < 0$. In particular, take the policy $\tilde{r}(I) = 0$ for all $I < I^*/2$. It follows that

$$0 \leq \frac{c(r(I))}{\left(r(I) + \frac{X(I)}{I} - r^0\right) I} + \frac{v(r^0 - r(I) + \theta)}{\left(r(I) + \frac{X(I)}{I} - r^0\right)} \leq \frac{v(r^0 + \theta)}{\frac{X(I)}{I} - r^0}. \tag{60}$$

Take the limit on both sides to obtain the result.
qed.

Lemma 8. .

If $\xi_0 > r^0$, the economic unit cost of suppression goes zero, and the social unit cost from the flow of death goes to a constant, as the mass of infectious goes to zero.

PROOF:

Use the same argument as above. In the limit

$$0 \leq \lim_{I \to 0} \frac{c(r(I))}{\left(r(I) + \frac{X(I)}{I} - r^0\right) I} + \frac{v(r^0 + \theta)}{\xi_0 - r^0} \leq \frac{v(r^0 + \theta)}{\xi_0 - r^0}, \tag{61}$$

which proves the result.

qed.

Lemma 9. .

If $\xi_0 > r^0$, the total cost of suppression is bounded in the optimum.

PROOF:

The total cost of suppression at the optimum is smaller or equal to the total cost of suppression under any other policy that satisfies $\dot{I}(I) < 0$. In particular, take the policy $\tilde{r}(I) = 0$ for $I \leq I^*/2$ and $\tilde{r}(I) = r^0$ for $I > I^*/2$. It follows that

$$\int_0^{I_0} \frac{c(r(I)) + vI}{\left(r(I) + \frac{X(I)}{I} - r^0\right) I} dI \leq \int_0^{I^*/2} \frac{v(r^0 + \theta)}{\frac{X(I)}{I} - r^0} dI + \int_{I^*/2}^{I_0} \frac{c(r^0) + vI\theta}{X(I)} dI \tag{62}$$

Both integrals exist, which gives the result.

qed.

Lemma 10. .

If $\xi_0 < r^0$, and the cost function is quadratic, the economic unit cost of suppression goes to infinity and the social unit cost goes to a constant as the mass of infectious goes to zero.

PROOF:

Take the definition of the total unit cost and take the limit. Use the results from Lemma 5. Assume that $2(r^0 - \zeta_0) << r^0 + \theta$. Intuitively, it means the system is far enough from the right limit $r^0 + \theta$. Note that r close to the limit are at odds with the assumption of a quadratic cost:

$$\lim_{I\to 0}\frac{c(r(I))}{\left(r(I)+\frac{X(I)}{I}-r^0\right)I}+\lim_{I\to 0}\frac{v(r^0-r(I)+\theta)}{\left(r(I)+\frac{X(I)}{I}-r^0\right)}=\frac{c(2(r^0-\xi_0))}{r^0-\xi_0}\lim_{I\to 0}\frac{1}{I}+\frac{v(r^0+\theta-2(r^0-\xi_0))}{r^0-\xi_0}. \tag{63}$$

qed.

Lemma 11. .
If $\xi_0 < r^0$, and the cost function is quadratic, the total cost of suppression goes to infinity if the extinction threshold I_ϵ goes to zero, even in the optimum.

PROOF:
Take the expression for the total cost and use the optimality condition to get

$$C=\int_{I_\epsilon}^{I_0}\frac{c(r(I))+vI(r^0-r(I)+\theta)}{\left(r(I)+\frac{X(I)}{I}-r^0\right)I}dI=\int_{I_\epsilon}^{I_0}\frac{c'(r(I))-vI}{I}dI \tag{64}$$

The optimal policy is increasing and larger than zero in zero; therefore

$$C\geq\int_{I_\epsilon}^{I_0}\frac{c'(r(0))-vI}{I}dI \tag{65}$$

Take the limit for I_ϵ going to zero to get the result.
qed.

Lemma 7 to 9 prove the statements on case 1 in Proposition 3. Lemma 10 and 11 prove the statements on case 2.
qed.

A.3 Extensions

A.3.1 Discounting, Vaccine, and Cure

Suppose the planner discounts the future at a positive time-discount rate. Additionally, a vaccine or cure that immediately end the pandemic arrive stochastically at a positive and constant Poisson-rate. Together, the two phenomena give rise to a total discount factor of i. The planner's problem is:

$$\min_{r_t} C(r_t) = \int_0^\infty e^{-it}\left(c(r_t) + vI_t(r^0 - r_t + \theta)\right) dt \tag{66}$$

such that

$$\dot{I}_t = (r^0 - r_t)I_t - X(I_t). \tag{67}$$

Only consider solutions that converge to I_ϵ and $\dot{I}_t < 0$ for all t. Change variable to I. Note that

$$t(I) = \int_{I_0}^{I} \frac{1}{\dot{I}(I)} dI \tag{68}$$

It follows that

$$\min_{r(I)} C(r(I)) = \int_{I_0}^{I_\epsilon} e^{-i\int_{I_0}^{I}\frac{1}{\dot{I}(\tilde{I})}d\tilde{I}} \left(\frac{c(r(I)) + vI(r^0 - r(I) + \theta)}{\dot{I}(I)}\right) dI. \tag{69}$$

The solution to this problem is a control function $r(I)$.

Proposition 4. .

1. *For every I, the solution $r(I)$ of Problem (69) fulfills*

$$\frac{c'(r(I)) - vI}{c(r(I)) + vI(r^0 - r(I) + \theta)} + \\ + \frac{i}{(c(r(I)) + vI(r^0 - r(I) + \theta)(-g(r(I), I))} C(I) = \frac{1}{-g(r(I), I)}, \tag{70}$$

   where $C(I)$ denotes the value function and $g(r(I), I)$ the growth rate.

2. *The optimal policy for discount rate i, denote it by $r_i(I)$, is smaller than the optimal policy for discount rate $i = 0$, denoted by $r_{i=0}(I)$ and characterized in Proposition 1.*

3. *For i close to zero or I_0 close to I_ϵ, the optimal policy for discount rate i, $r_i(I)$, is close to the optimal policy for discount rate $i = 0$, $r_{i=0}(I)$. In the limit when $i = 0$ or $I = I_\epsilon$ the policies are equal.*

4. *For i close to zero or I_0 close to I_ϵ the solution $r(I)$ of Problem (69) exists.*

5. *There exist i, $X(.)$, and I_0, such that Problem (66) has local minima where prevalence converges to a constant steady state level I_{ss}: $\lim_{t\to\infty} I_t = I_{ss} \neq I_\epsilon$.*

6. *For i close to zero or I_0 close to I_ϵ, the solution $r(I)$ of Problem (69), converging to I_ϵ, is a global minimum of Problem (66).*

7. *For i close to zero it holds that optimal social distancing is increasing in prevalence: $r'(I) > 0$ for all $I \in [I_\epsilon, I_0]$.*

The proposition shows that for small enough discount rates the qualitative results in Proposition 1 and 2 do not change. This is the relevant case. The pandemic moves fast such that the relevant time discount rate is the daily or weekly rate. This rate is very low. Similarly, the daily or weekly probability for an effective cure or mass vaccine is currently very low. Point 2 shows that quantitatively, social distancing is less intense under discounting. This result is intuitive. Under discounting, costs can be reduced to some extend by deferring them into the future. In order to do that, suppression needs to progress slower, which is why r is smaller. The effect is driven by the second summand on the left in Condition (70).

Interestingly, as Point 5 shows, the suppression solution may not be the only local minimum of the problem. For some I_0, $X(I)$, and i, it does not even exist. One or several other local minima exist where prevalence converges to some steady-state value I_{ss}. In this case, the question of which local minimum is the global minimum is a quantitative question. In some cases, the global minimum is a path that converges to a steady-state level of prevalence. This second solution is the mitigation solution discussed in the conclusion. However, a steady-state level of prevalence is at odds with the initial assumption that the number of susceptible is approximately constant. As a consequence, mitigation needs to be studied in the full SIR model, which has already been done in the literature. Therefore, it is beyond the scope of this paper.

PROOF of Proposition 4:

Point 1:

Take the derivative of (69) with respect to each $r(I)$:

$$\left((c'(r(I)) - vI)\frac{1}{\dot{I}(I)} + \left(c(r(I)) + vI(r^0 - r(I) + \theta)\right)\frac{1}{\dot{I}(I)^2}I\right) e^{-i\int_{I_0}^{I}\frac{1}{\dot{I}(\tilde{I})}d\tilde{I}} +$$
$$+\int_{I}^{I_\epsilon} e^{-i\int_{I_0}^{\tilde{I}}\frac{1}{\dot{I}(\tilde{\tilde{I}})}d\tilde{\tilde{I}}}\left(-\frac{1}{\dot{I}(I)^2}Ii\right)\left(\frac{c(r(\tilde{I})) + v\tilde{I}(r^0 - r(\tilde{I}) + \theta)}{\dot{I}(\tilde{I})}\right)d\tilde{I} = 0. \quad (71)$$

The FOC simplifies to

$$(c'(r(I)) - vI)\frac{1}{-\dot{I}(I)} - \left(c(r(I)) + vI(r^0 - r(I) + \theta)\right)\frac{I}{\dot{I}(I)^2} + \frac{iI}{\dot{I}(I)^2}C(I) = 0, \quad (72)$$

where $C(I)$ is the value function.

To show that it is a local minimum consider the SOC. Rearrange the FOC to

$$\frac{1}{\left(r - r^0 + \frac{X(I)}{I}\right)^2 I}\left(-c(r) - vI(r^0 - r + \theta) + iC(I) + (c'(r) - vI)\left(r + \frac{X(I)}{I} - r^0\right)\right) \quad (73)$$

Take the derivative to get the second order condition. Note that it is equal to $a'b + ab'$ where a is the first factor above and b is the second factor. If the FOC holds, b is zero. Also, a is always positive. The sign of the SOC only depends on the sign of b'. $b' = c''(r)\left(r + \frac{X(I)}{I} - r^0\right)$, because $\frac{dC(I)}{dr(I)} = 0$. $b' > 0$, therefore, the SOC>0 and the FOC characterizes a local minimum.

Rewrite the FOC to get

$$\frac{c'(r(I)) - vI}{c(r(I)) + vI(r^0 - r(I) + \theta)} + \frac{i}{(c(r(I)) + vI(r^0 - r(I) + \theta)(-g(r(I), I))}C(I) = \frac{1}{-g(r(I), I)},$$

which is the condition in the proposition.

Point 2:

Rewrite the FOC to

$$(c'(r) - vI)\left(r + \frac{X(I)}{I} - r^0\right) - (c(r(I)) + vI(r^0 - r + \theta)) = -iC(I). \quad (74)$$

The right hand side is negative. For $r_{i=0}(I)$ the left hand side is equal to zero. The left hand side is increasing in r. Therefore, $r_i(I) < r_{i=0}(I)$.

Point 3:

$C(I)$ is bounded. If i goes to zero, $iC(I)$ goes to zero and Condition (70) goes to Condition (13). Therefore, $r_i(I)$ goes to $r_{i=0}(I)$. Similarly, when I goes to I_ϵ, $C(I)$ goes to zero. Therefore, $iC(I)$ goes to zero and the same argument applies.

Point 4:

The solution exists if $\dot{I} < 0$ for all $I \in [I_\epsilon, I_0]$. When I_0 is close to I_ϵ or i is close to 0, then r_i is close to $r_{i=0}$. By Proposition 1, the zero discount solution exists. Therefore $\dot{I}(r_{i=0}(I)) < 0)$, which by continuity is also true for all r close to $r_{i=0}$.

Point 5:

Use the Hamiltonian to solve Problem (66):

$$c'(r_t) - vI_t = \lambda_t I_t \tag{75}$$

$$\dot{\lambda}_t = (i + X'(I_t) + r_t - r^0)\lambda_t - v(r^0 - r + \theta) \tag{76}$$

$$\dot{I}_t = (r^0 - r_t)I_t - X(I_t) \tag{77}$$

$$\lim_{t\to\infty} e^{-it}\lambda_t = 0 \tag{78}$$

It follows that

$$\dot{r}_t = \frac{(i + X'(I_t) - X(I_t)/I_t)\,(c'(r_t) - vI_t) - vI_t\left(\theta + \frac{X(I)}{I}\right)}{c''(r_t)} \tag{79}$$

$$\dot{I}_t = (r^0 - r_t)I_t - X(I_t) \tag{80}$$

The two equations give two loci. $\dot{I}$ is zero if

$$r = r^0 - X(I)/I. \tag{81}$$

$\dot{r}$ is zero if

$$c'(r) = \frac{vI(\theta + i + X'(I))}{i - (X(I_t)/I_t - X'(I_t))}. \tag{82}$$

One or several steady states with positive I may exist, dependent on if the above system has a solution, i.e., the two loci cross at least once. Choose i and $X(.)$ such that the two loci cross at least once and I_0 equal to the corresponding steady-state prevalence I_{ss}.

Point 6:

Choose i small enough such that for all $I \in [I_\epsilon, I_0]$ it holds that $i-(X(I_t)/I_t - X'(I_t)) < 0$. Note that such an $i > 0$ exists because $(X(I)/I - X'(I) > 0$ for all $I > 0$. It follows that $\dot{r}_t < 0$ for all t. The only path fulfilling the boundary condition of the Hamiltonian converges to I_ϵ.

If a saddle path from I_ϵ to the steady state I_{ss} exists, denote by $C_{ss}(I_\epsilon)$ the value function to reach it. It is larger than zero. Denote by $C_\epsilon(I_0)$ the value function of reaching I_ϵ from I_0. Point 4 shows that it exists. $C_\epsilon(I_0)$ goes to zero if I_0 goes to I_ϵ. Choose I_0 low enough such that $C_{ss}(I_\epsilon) > C_\epsilon(I_0)$.

Point 7:

The total cost under the optimal policy is

$$C(I) = \int_I^{I_\epsilon} e^{-i \int_I^{\tilde{I}} \frac{1}{\dot{I}(\tilde{\tilde{I}})} d\tilde{\tilde{I}}} \left(\frac{c(r(\tilde{I})) + v\tilde{I}(r^0 - r(\tilde{I}) + \theta)}{\dot{I}(\tilde{I})} \right) d\tilde{I}. \tag{83}$$

The derivative is

$$\begin{aligned} C'(I) = & \frac{c(r(I) + vI(r^0 - r(I) + \theta)}{-\dot{I}(I)} - \\ & - \frac{i}{-\dot{I}(I)} \int_I^{I_\epsilon} e^{-i \int_I^{\tilde{I}} \frac{1}{\dot{I}(\tilde{\tilde{I}})} d\tilde{\tilde{I}}} \left(\frac{c(r(\tilde{I})) + v\tilde{I}(r^0 - r(\tilde{I}) + \theta)}{\dot{I}(\tilde{I})} \right) d\tilde{I}, \end{aligned} \tag{84}$$

which can be written as

$$C'(I) = \frac{c(r(I)) + vI(r^0 - r(I) + \theta)}{-\dot{I}(I)} - \frac{i}{-\dot{I}(I)} C(I). \tag{85}$$

Together with Condition (70) it gives

$$C'(I) = \frac{c'(r(I)) - vI}{I}. \tag{86}$$

Differentiate the Condition (70) with respect to I and use the result for $C'(I)$ to get

$$r' = \frac{v\left(\theta + \frac{X(I)}{I}\right) + \frac{c'(r) - vI}{I}\left(\frac{X(I)}{I} - X'(I) - i\right)}{\left(r + \frac{X(I)}{I} - r^0\right) c''(r)}. \tag{87}$$

In general, it is possible that r' is negative. However, if i is small enough r' is always positive. Choose i small enough such that for all $I \in [I_\epsilon, I_0]$ it holds that $X(I)/I - X'(I) - i > 0$. Note that such an $i > 0$ exists, because $(X(I)/I - X'(I)) > 0$ for all $I > 0$. Also, note that, by Condition 75, $c'(r) - vI > 0$.
qed.

A.3.2 Endogenous Choice of Detection

The function $\tilde{X}(I, z)$ is the detection function discussed in Section A.1.1. I is the number of undetected cases and z is the amount of resources spend for detection. The planner's problem is

$$\min_{r(.), z(.)} C(r(.), z(.)) = \int_{I_0}^{0} (c(r(I)) + z(I) + vI(r^0 - r(I) + \theta)) \frac{1}{\dot{I}(I)} dI, \tag{88}$$

where

$$\dot{I}(I) = I(r^0 - r(I) - \tilde{X}(I, z(I)). \tag{89}$$

Proposition 5. .

For each level of prevalence I, the optimal policy $\{r(I), z(I)\}$ solves

$$\frac{c'(r) - vI}{c(r) + z + vI(r^0 - r + \theta)} = \frac{1}{r + \frac{\tilde{X}(I,z)}{I} - r^0}, \tag{90}$$

$$\frac{\frac{I}{\frac{\partial \tilde{X}(I,z)}{\partial z}}}{c(r) + z + vI(r^0 - r + \theta)} = \frac{1}{r + \frac{\tilde{X}(I,z)}{I} - r^0}. \tag{91}$$

PROOF:

As for Proposition 1 in Section A.2.1.