COVID ECONOMICS
VETTED AND REAL-TIME PAPERS

ISSUE 41
3 AUGUST 2020

Covid Economics
Vetted and Real-Time Papers

Covid Economics, Vetted and Real-Time Papers, from CEPR, brings together formal investigations on the economic issues emanating from the Covid outbreak, based on explicit theory and/or empirical evidence, to improve the knowledge base.

Founder: Beatrice Weder di Mauro, President of CEPR
Editor: Charles Wyplosz, Graduate Institute Geneva and CEPR

Contact: Submissions should be made at https://portal.cepr.org/call-papers-covid-economics. Other queries should be sent to covidecon@cepr.org.

The Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,500 research economists based mostly in European universities. The Centre's goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers. CEPR's guiding principle is 'Research excellence with policy relevance'. A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society. CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Editorial Board

Ethics

Covid Economics will feature high quality analyses of economic aspects of the health crisis. However, the pandemic also raises a number of complex ethical issues. Economists tend to think about trade-offs, in this case lives vs. costs, patient selection at a time of scarcity, and more. In the spirit of academic freedom, neither the Editors of *Covid Economics* nor CEPR take a stand on these issues and therefore do not bear any responsibility for views expressed in the articles.

Submission to professional journals

The following journals have indicated that they will accept submissions of papers featured in *Covid Economics* because they are working papers. Most expect revised versions. This list will be updated regularly.

American Economic Review

American Economic Review, Applied Economics

American Economic Review, Insights

American Economic Review, Economic Policy

American Economic Review, Macroeconomics

American Economic Review, Microeconomics

American Journal of Health Economics

Canadian Journal of Economics

Econometrica*

Economic Journal

Economics of Disasters and Climate Change

International Economic Review

Journal of Development Economics

Journal of Econometrics*

Journal of Economic Growth

Journal of Economic Theory

Journal of the European Economic Association*

Journal of Finance

Journal of Financial Economics

Journal of International Economics

Journal of Labor Economics*

Journal of Monetary Economics

Journal of Public Economics

Journal of Political Economy

Journal of Population Economics

Quarterly Journal of Economics*

Review of Economics and Statistics

Review of Economic Studies*

Review of Financial Studies

(*) Must be a significantly revised and extended version of the paper featured in *Covid Economics*.

Covid Economics
Vetted and Real-Time Papers

Issue 41, 3 August 2020

Contents

Optimal vaccination and herd immunity[1]

Stefano Bosi,[2] Carmen Camacho[3] and David Desmarchelier[4]

Herd immunity is a central concept in epidemiology. It refers to a situation where the amount of recovered and immune individuals is high enough to protect susceptibles from contracting the disease. Herd immunity can be obtained naturally when individuals recover from the disease, or artificially after the administration of an appropriate vaccine. The present paper addresses the question of the amount of public spending in medical research to obtain a vaccine which maximizes welfare. Public spending is assumed to reduce the waiting time to discover a vaccine against an infectious disease evolving according to a SIR model. Both linear and logarithmic preferences are considered with and without time discounting. Worth to note, we show that if an economy has a sufficiently performing technology, then the government should invest as much as the initial public budget allows.

2   Full Professor of Economics, EPEE, Paris-Saclay University.
3   Senior Researcher in Economics, Paris School of Economics, Paris Jourdan Sciences Economiques.
4   Associate Professor of Economics, University of Lorraine, University of Strasbourg, CNRS, BETA.

1   Introduction

Herd immunity is the indirect protection conferred to susceptible individuals when there is a sufficiently large proportion of immune individuals (Randolph and Barreiro, 2020). Herd immunity can be attained either naturally or through a vaccination program, when and if a vaccine is available. When the recent COVID-19 became a pandemic in March 2020, most governments imposed a temporary and severe lockdown, aiming at slowing down the virus propagation. Others, on the contrary, did not impose any type of confinement aiming at attaining natural herd immunity as soon as possible. Whatever the temporary strategy, nothing seems to prevent this virus from spreading and causing thousands of deaths. As a consequence, governments and pharmaceutical firms have entered a race to discover a vaccine to artificially attain herd immunity. As reported by Callaway (2020), by April 2020 more than 90 vaccines were under study.

However, as the contemporary reader may have learned, a vaccine takes months to be discovered, tested, produced and commercialized. Hence, given the merciless advance of the COVID-19, policy makers may wonder whether and under which circumstances it is worth investing in a vaccine. Kwok et al. (2020) provide some elements to address the first question, estimating the minimum proportion of COVID-19 recovered individuals to attain herd immunity. This proportion ranges from 11.5% for Singapore to more than 80% for Spain, Denmark, Slovenia or Bahrain. For the USA, Sweden, Germany, Switzerland, Brazil and the Netherlands the share is around 70%; 59% for Italy and 65% for the UK. Assuming a mortality rate varying from 0.25% to 3%, the death toll that would ensure natural herd immunity seems difficult to accept. To complete this moral judgment, the present paper aims to study the questions of herd immunity and investment in vaccine research from an economic perspective. More precisely, we consider a simple economic model where the amount of public spending in medical research is chosen in order to maximize an intertemporal measure of welfare taking into account the SIR dynamics of an infectious disease. It is assumed that only healthy individuals are able to work and that production serves to finance both consumption and public spending. Moreover, public spending is assumed to increase the probability to discover a vaccine and all susceptibles are vaccinated as soon as the vaccine becomes available at no extra cost.

The literature on optimal vaccination is obviously not new. Let us mention Hethcote and Waltman (1973), which studies the optimal vaccination strategy that minimizes a cost function when an infectious disease evolves according to the SIR dynamics. Barrett and Hoel (2007) consider a closely related question but under SIS dynamics, where the only way to become immune is through vaccination. Brito et al. (1991) study the optimality of compulsory vaccination. In particular, they show that the market allocation is always better than compulsory vaccination. Indeed, if the agent does not care about whether others are vaccinated or not, she does not internalize the possible herd immunity, avoiding the possible free-rider problem. Note that the possibility to free-ride is a strong argument to implement a public compulsory vaccination campaign. Geoffard

and Philipson (1997) point out that despite the significant number of vaccines on the market, only smallpox has been eradicated. The authors explain this surprising outcome by the fact that vaccines imply a drop of infected individuals, entailing a drop in the demand for vaccines, which allows for a return of the infectious disease.

The present paper is closely related to Hethcote and Waltman (1973) and Barrett and Hoel (2007), although there are three major differences: (1) we consider a general equilibrium framework, (2) we assume that the vaccine has to be discovered, and (3) because of the interplay between the epidemic dynamics and public spending viewed as a control variable, the problem is non-convex. For this reason we consider that public spending is constant over time until the vaccine is discovered. This framework allows to discuss something new in the literature on optimal vaccination: the optimal public spending in medical research taking into account that the date of natural herd immunity depends on research investment and that it is a known probabilistic function.

Many infectious diseases evolve according to SIR dynamics. Hethcote (1976) pointed out that this is particularly the case for viral agent diseases such as measles, chicken pox, mumps, influenza and smallpox. More recently, Roda et al. (2020) have shown that the SIR model is also well suited to represent the spread of COVID-19.

In the SIR model, population is divided in three classes: susceptibles, infectives and recovered individuals. Susceptibles can contract the disease after a contact with an infected individual. After a given lapse of time, the infected individual recovers from the disease and develops a permanent immunity. Interestingly, as the share of recovered individuals increases, the probability that a susceptible meets an infected diminishes. If the share of recovered individuals is sufficiently high, then the share of infectives decreases over time and eventually, herd immunity results.

The present unified framework lies at the crossroad between economics and epidemiology, and serves to study the problem of a policy maker under linear and logarithmic preferences, with and without time discounting. When preferences are linear and the policy maker cares equally about the future and the present, we prove that it is always optimal to invest in vaccine research. Moreover, optimal public spending is higher when natural herd immunity occurs from the beginning. This interesting result obtains because when natural herd immunity occurs, labour supply increases over time, augmenting in turn the public budget and the vaccine spending possibilities. Somehow in contrast, when preferences are linear and time is discounted, or when preferences are logarithmic with no time discount, then it is optimal to engage in medical research if and only if research activities are sufficiently productive. This is equivalent to saying that policy makers should invest in the search for the vaccine if and only if the waiting time for the vaccine discovery is short enough.

The rest of the paper is organized as follows. Section 2 presents the SIR model. The market economy is presented in section 3. Sections 4 and 5 study the Ramsey criterion and the discounting criterion. Section 6 concludes the paper. All proofs are gathered in the appendix.

2 The SIR epidemiological model

This section presents and characterizes the SIR epidemiological model, which is used to explain the dynamics of a non-lethal disease. Population is divided into susceptible individuals (S), who have not contracted the illness yet; infected individuals (I) who are ill, and recovered individuals (R) who caught the disease in the past and who are now cured. Recovered individuals gain permanent immunity. Let us assume that population is constant and equal to N so that

$$S(t) + I(t) + R(t) = N$$

and

$$S'(t) + I'(t) + R'(t) = 0 \tag{1}$$

The shares of susceptibles, infectives and recovered individuals can be written as

$$s(t) \equiv \frac{S(t)}{N(t)}, \quad i(t) \equiv \frac{I(t)}{N(t)}, \quad r(t) \equiv \frac{R(t)}{N(t)} \tag{2}$$

Obviously, the sum of the shares is always one: $s(t) + i(t) + r(t) = 1$. As usual, the initial number of infectives, susceptibles and recovered are known and we can compute

$$s(0) = s_0, \quad i(0) = i_0, \quad r(0) = r_0$$

Besides, since the shares add to 1, $s_0 + i_0 + r_0 = 1$.

The key assumption here is that at every time t, each susceptible meets an infective with probability $i(t)$. If β measures the transmissibility rate, then the number of new infectives at time t is given by $\beta i(t) S(t)$. Furthermore, letting ρ be the recovery rate, then the number of recovered individuals at time t is $\rho I(t)$. With these in hand, we can describe the SIR dynamics:

Proposition 1 (SIR dynamics) *Under condition (1), the SIR epidemiological model can be written as*

$$\frac{s'(t)}{s(t)} = -\beta i(t) \tag{3}$$

$$\frac{i'(t)}{i(t)} = \beta s(t) - \rho \tag{4}$$

with $r'(t) = \rho i(t)$. Therefore, $s'(t) < 0$, $r'(t) > 0$ and

$$i'(t) < 0 \Leftrightarrow s(t) < \frac{\rho}{\beta} \tag{5}$$

for any t.

Proof. See Appendix A. ∎

Notice that the number of susceptible individuals always decreases with time and that of recovered always increases. The number of infected increases while

the share of susceptibles is high enough. When $s(t)$ falls below the threshold ρ/β, then, the share of infected decreases. Note that when $s(t) = \rho/\beta$, then the share of recovered equals the share of infected. Hence, any non-lethal disease will always spread and contaminate a maximum of individuals. After the peak, the share of infected will decrease although the disease will still propagate.

The following proposition provides equivalent, but independent, dynamics for r and s.

Proposition 2 (independent dynamics) *The SIR model in (3)-(4) is equivalent to the following pair of independent equations for r and s:*

$$r'(t) = \rho\left[1 - r(t) - s_0 e^{\frac{\beta}{\rho}[r_0 - r(t)]}\right] \tag{6}$$

$$s'(t) = \beta s(t)\left[r_0 + s(t) - 1 - \frac{\rho}{\beta}\ln\frac{s(t)}{s_0}\right] \tag{7}$$

There exists a critical value for s, s_1, verifying that $0 < s_1 < s_0$, and $s_1 < \rho/\beta$. $s(t)$ decreases monotonically from its initial value s_0 to s_1. Noteworthy, the disease is eradicated in the long run.

Proof. See Appendix B. ∎

Proposition 2 can also be read in terms of herd immunity. Indeed, since the disease completely disappears in the long run, we can compute the date of natural herd immunity which corresponds to the moment in time when the share of infectives starts decreasing.

Proposition 3 (natural herd immunity date) *Consider the SIR model studied in Proposition 2. The critical date of natural herd immunity, t_1, depends on the initial amount of infected:*
(1) If $s_0 \leq \rho/\beta$, then the share of infectives decreases forever. That is, natural herd immunity is attained at $t_1 = 0$.
(2) If $s_0 > \rho/\beta$, then the share of infectives increases for $t < t_1 = s^{-1}(\rho/\beta)$ and decreases after. Furthermore the epidemics reaches a maximum at t_1 with

$$i(t_1) = 1 - r\left(s^{-1}(\rho/\beta)\right) - \rho/\beta \tag{8}$$

Proof. See Appendix C. ∎

Proposition 3 shows that the trajectory of the share of infectives is a single-peaked curve, increasing first from 0 to t_1 and decreasing thereafter. Any public policy designed to control and eradicate the disease tries to flatten the curve by lowering this peak.

The following corollary underlines the link between infectiveness and the immunity date:

Corollary 4 (comparative statics) *The more infective the disease, the later the date of natural herd immunity.*

Proof. See Appendix D. ∎

The more infective the disease, the higher the share of recovered we need to ensure herd immunity and the longer it takes.

2.1 An approximated solution

We are unable to provide with the analytical solution of the dynamic system (6)-(7). However, using the information we have already obtained, we shall opt for an approximation which is extremely close to the exact solution of the system under certain assumptions. We know that r increases from r_0 to r_1 and s decreases from s_0 to s_1. At the same time, i moves continuously from i_0 to 0. According to Proposition 3, if $s_0 < \rho/\beta$, then i decreases forever. If, on the contrary, $s_0 > \rho/\beta$, then i increases first and from a given date on, it decreases steadily until it reaches its asymptotic value 0. As a result and according to (5), $s_1 \equiv \lim_{t\to\infty} s(t) = s_1 < \rho/\beta$.

Take the logarithm of infectives and approximate it with a first order Taylor polynomial

$$\ln i(t) \quad \approx \quad l_0 + l_1 t$$

Note that l_0 and l_1 can be computed using the initial condition i_0 and $\lim_{t\to\infty} i(t) = 0$, so that the log-linear approximation of i obtains

$$i(t) \approx i_0 e^{-\delta t} \equiv i_\delta(t) \tag{9}$$

where $\delta \equiv -l_1 > 0$.

The log-linear approximation is particularly well fit when the share of infectives decreases forever, that is, when $s_0 \leq \rho/\beta$. In this case, the approximated rate of decline of infectives is a negative constant:

$$\frac{i'(t)}{i(t)} = -\delta < 0$$

Given the quality of the fit when $s_0 \leq \rho/\beta$, and whenever necessary, we will consider the log-linear approximation (9). In particular, our approximation is well founded when the recovery rate is larger than the transmissibility rate, i.e. when $\rho \geq \beta$.

Note that in order to complete the log-linear approximation of i, we still need to identify δ. Proposition 5 below solves this problem using the asymptotic result $r_1 + s_1 = 1$.

Proposition 5 (log-linear SIR dynamics) *Under the log-linear approximation for i given in (9), the SIR model consists in*

$$
\begin{aligned}
s(t) &\approx s_0 e^{\frac{\beta}{\delta} i_0 \left(e^{-\delta t} - 1\right)} & (10) \\
i(t) &\approx i_0 e^{-\delta t} & \\
r(t) &\approx r_0 - \frac{\rho}{\delta} i_0 \left(e^{-\delta t} - 1\right) & (11)
\end{aligned}
$$

where δ is the solution to

$$r_0 + \frac{\rho}{\delta}i_0 + s_0 e^{-\frac{\beta}{\delta}i_0} = 1 \qquad (12)$$

Proof. See Appendix E. ∎

To illustrate the pertinence of the approximation, we numerically compare the approximated and the actual trajectories for the share of infectives, $i_\delta(t)$ and $i(t)$, respectively. For this exercise, we calibrate the model's parameters using quarterly data (see Table 1 below).

A	Scale parameter	1
β	Transmissibility rate	17/2
ρ	Recovery rate	6
r_0	Initial share of recovered	1/6
s_0	Initial share of susceptibles	1/6

Table 1. Calibration.

Using (12), we compute the value of δ associated to the previous calibration: $\delta = 1.96335$. Figure 1 shows that indeed the log-linear approximation is close to i.
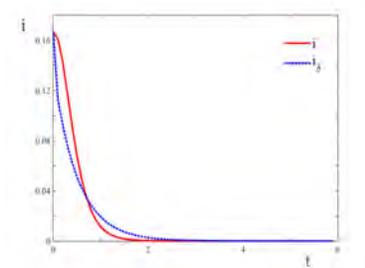
Fig. 1. Log-linear approximation.

2.2 Vaccine

Let us assume that the government considers to dedicate public funds to the discovery of a vaccine against the disease behind the epidemics. The vaccine will confer permanent immunity and all susceptibles will be vaccinated as soon as it becomes available at no additional cost. More precisely, we assume that the government spends a constant amount G from $t = 0$ until the moment T the vaccine is discovered. From date T on, the number of infectives decreases according to (30), where β will be 0 because the vaccine reduces the transmissibility to zero. Hence, substituting $\beta = 0$ in (4) we obtain $I'(t) = -\rho I(t)$ or, equivalently,

$$I(t) = I(T) e^{\rho(T-t)} \qquad (13)$$

The probability of finding the vaccine at time t follows a Poisson process and it depends on the amount invested. Intuitively, higher public spending increases the probability of discovering the vaccine. Accordingly, the probability of not finding the vaccine between time t and $t + dt$ will be $e^{-p(G)}$ for some function p. For simplicity reasons, we assume that public spending has a linear effect on p:

Assumption 1 $p\left(G\right) = \pi G$ *with* $\pi > 0$.

The probability that the vaccine is not discovered by time t is

$$e^{\int_0^t \ln e^{-\pi G} dt} = e^{-\pi G t}$$

Note that research investment G equals Ng, where g denotes research investment per capita. Then, with an abuse of notation, we can write p as a function of g, that is $p(g) = \pi Ng = \gamma g$ with $\gamma = \pi N \in (0, \infty)$. γ is the research efficiency of the economy, which depends in turn on the research technology itself, π, and the country's population size, N.

As advanced, observe that the average research time to discover a vaccine is a function of g:

$$T(g) = \int_0^\infty t e^{-\gamma g t} \left(1 - e^{-\gamma g}\right) dt = \frac{1 - e^{-\gamma g}}{(\gamma g)^2} \tag{14}$$

For later purposes, it is convenient to approximate this average research time. The simplest Taylor approximation for e^{-p} is the linear approximation:

$$e^{-p} = \sum_{n=0}^\infty \frac{(-p)^n}{n!} \approx 1 - p$$

Hence, $e^{-\gamma g} \approx 1 - \gamma g$. Using this approximation, we also obtain an approximation for the research time:

$$T(g) \approx \frac{1 - 1 + \gamma g}{(\gamma g)^2} = \frac{1}{\gamma g} \tag{15}$$

Figure 2 shows that this simple first-order Taylor approximation is an extraordinary good fit for two values of γ, $1/2$ and 12.
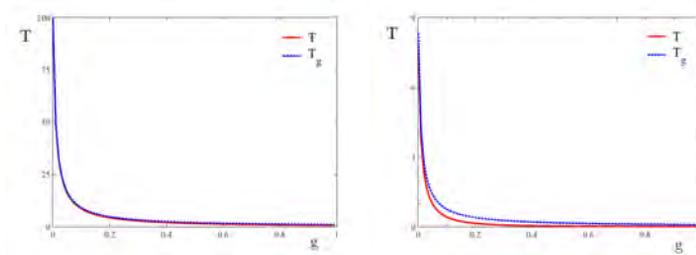
Fig. 2. T and T_g. Left panel: $\gamma = 1$. Right panel: $\gamma = 12$.

Derivating $T(g)$ using (14), it is easy to prove that $T'(g) < 0$. Unsurprisingly, the waiting time for a vaccine decreases with the public research effort.

The following lemma provides us with an upper bound for public investment when the disease is initially sufficiently spread.

Lemma 6 (research constraints) *Consider the critical date t_1 of natural herd immunity given by (6). Fix an arbitrary level of public spending per capita in research g. This given level of g also determines also the waiting time for a vaccine $T(g)$ and must satisfy the following constraints of non-negativity of consumption.*

(1) If $s_0 \le \rho/\beta$, the constraints of non-negativity of consumption $(c(t) \ge 0)$ on the interval $[0, T(g))$ are equivalent to $g \le A(1 - i_0)$.

(2) If $s_0 > \rho/\beta$, then $t_1 = s^{-1}(\rho/\beta)$. There are two cases.

(2.1) If $t_1 \le T(g)$, the constraints $c(t) \ge 0$ on the interval $[0, T(g))$ are equivalent to

$$g \le A\left[1 - r\left(s^{-1}(\rho/\beta)\right) - \rho/\beta\right]$$

(2.2) If $t_1 > T(g)$, the constraints $c(t) \ge 0$ on the interval $[0, T(g))$ are equivalent to

$$g \le A\left[1 - i(T(g))\right]$$

Proof. See Appendix F. ∎

3   Market economy

After describing firms and consumers, we describe the objective of the policy maker who will decide whether and how much to invest in the search of a vaccine taking into account overall welfare.

In this economy there are many identical firms indexed by j with no market power, that is they are price-takers. For mathematical tractability, we assume that all firms use the same linear production function.

Assumption 2 *The production function of firm j is a linear function of labor demand, for all j.*

Letting Y_j denote output of firm j, L_j its labor demand and A the common economy's total factor productivity, we can write by Assumption 1 that $Y_j = AL_j$. Each firm maximizes its profits $AL_j(t) - w(t) L_j(t)$, which implies that at equilibrium $w(t) = A$.

Let us assume that workers are fully insured against the potential illness, so that they receive the same wage, ω, whether they are healthy or ill. Hence, at equilibrium $w(t) L(t) = \omega(t) N$.

Their disposable income is given by

$$c(t) = y(t) = \begin{cases} \omega(t) - g = \frac{w(t)L(t)}{N} - g = \frac{AL(t)}{N} - g, & \text{if } 0 \leq t < T \\ \omega(t) = \frac{w(t)L(t)}{N} = \frac{AL(t)}{N}, & \text{otherwise} \end{cases} \tag{16}$$

where from the consumer perspective, $g \equiv G/N$ is the individual lump-sum tax used to finance public research to discover the vaccine.

The labor force is the number of healthy individuals, that is, $N - I(t)$. Thus, according to (13)

$$L(t) = \begin{cases} N - I(t), & \text{if } 0 \leq t < T \\ N - I(T) e^{\rho(T-t)}, & \text{otherwise} \end{cases}$$

Therefore, replacing $L(t)$ in (16), consumption can be written as

$$c(t) = \begin{cases} \frac{A[N-I(t)]}{N} - g = A[1 - i(t)] - g, & \text{if } 0 \leq t < T \\ \frac{A[N-I(T)e^{\rho(T-t)}]}{N} = A\left[1 - i(T) e^{\rho(T-t)}\right], & \text{otherwise} \end{cases} \tag{17}$$

Equation (17) demonstrates that there is a sudden rise of consumption at time T from $c(T^-) = A[1 - i(T)] - g$ to $c(T^+) = A[1 - i(T)]$. This stems from the fact that the research cost is not smoothed over time and that consumers must pay the cost of research while research is undertaken. In other words, the government cannot get indebted to enter the vaccine race.

In this setup, all households are identical, of constant size and they live forever. As a result, the welfare maximization problem of the government is equivalent to maximizing household's intertemporal utility. Accordingly the government maximizes the following utility functional with respect to the public spending in research

$$U(g) \equiv \int_0^\infty e^{-\eta\theta t} \left[u(c(t)) - (1-\eta) u(A)\right] dt \tag{18}$$

with $\eta \in \{0, 1\}$. θ is the subjective discount rate and $u = u(c)$ describes the household's instantaneous utility. According to (17), A is the asymptotic consumption per capita

$$A = \lim_{t \to \infty} \left(A\left[1 - i(T) e^{\rho(T-t)}\right]\right)$$

Thus, $u(A)$ is our reinterpretation of the bliss point considered in Ramsey (1928). Although the bliss point is a stationary state for Ramsey, it is here an asymptotic value.

Using (17) we can rewrite $U(g)$ as

$$
\begin{aligned}
U(g) \equiv\ & \int_0^T e^{-\eta\theta t} \left[u\left[A\left(1 - i\left(t\right)\right) - g \right] - \left(1 - \eta\right) u(A) \right] dt \\
& + \int_T^\infty e^{-\eta\theta t} \left[u\left[A\left(1 - i\left(T\right) e^{\rho(T-t)}\right) \right] - \left(1 - \eta\right) u(A) \right] dt
\end{aligned}
\tag{19}
$$

Let us make the following standard assumption regarding the utility function.

Assumption 3 *The utility function u is strictly increasing and concave.*

A standard functional form to measure utility is

$$
u(c) \equiv K \frac{c^{1 - \frac{1}{\varepsilon}}}{1 - \frac{1}{\varepsilon}}
\tag{20}
$$

defined for $c > 0$. This form has the appealing property of a constant elasticity of intertemporal substitution: $\varepsilon(c) = -u'(c) / [cu''(c)] = \varepsilon > 0$. We can normalize the constant K to one since $\arg\max_g U(g)$ does not depend on K.

In the remaining of the paper, the behavior of two policy makers is analyzed. The first policy maker values identically all moments in time, meaning that the future welfare of the population is to her as important as the present. We will characterize this policy maker as Ramsey in his 1928 paper, and compare the research effort of this policy maker with the choice of a second policy maker who discounts the future assigning higher value to present welfare than to the future's. In particular we consider the following two cases:

(1) The Ramsey criterion with $\eta = 0$. Here, agents do not discount future and the bliss point $u(A)$ is taken into account: $U(g) \equiv \int_0^\infty [u(c(t)) - u(A)] dt$.

(2) The discounting criterion with $\eta = 1$: $U(g) \equiv \int_0^\infty e^{-\theta t} u(c(t)) dt$. Contrary to case (1), agents discount the future without referring to a bliss point: the welfare functional boils down to the standard utility functional. Notice that in this simplified context, we do not consider the accumulation of capital as in the seminal models with discounting (Cass, 1965; Koopmans, 1965).

In each case, we solve the government's problem considering two different functional forms to describe household's preferences. Linear preferences will provide us with the explicit optimal research policy. On the downside, linear preferences are an extreme representation of the preferences of the individual over consumption. With these remarks in mind and for the two types of policy maker, we solve the linear problem obtaining the exact amount to invest in the search of a vaccine. Then, a more standard logarithmic case is analyzed, and optimal spending is characterized. We conclude that although linear preferences may distort to a certain extent the description of consumers' behavior, they provide a fairly accurate optimal policy regarding research investment.

4 Ramsey criterion

The government chooses the optimal research investment per capita, g, which maximizes the overall utility functional (19), that is

$$
\begin{aligned}
U(g) &= \int_0^{T(g)} \left[u\left(c\left(t\right)\right) - u\left(A\right)\right] dt + \int_{T(g)}^{\infty} \left[u\left(c\left(t\right)\right) - u\left(A\right)\right] dt \\
&= \int_0^{T(g)} \left[u\left(A\left[1 - i\left(t\right)\right] - g\right) - u\left(A\right)\right] dt \\
&\quad + \int_{T(g)}^{\infty} \left[u\left(A\left(1 - i\left(T\left(g\right)\right) e^{\rho\left[T(g)-t\right]}\right)\right) - u\left(A\right)\right] dt
\end{aligned}
\tag{21}
$$

subject to the epidemics dynamics in (6), (7) and the non-negativity of consumption, which according to (16) implies that for every t between 0 and $T(g)$

$$
c\left(t\right) = A\left[1 - i\left(t\right)\right] - g \geq 0, \text{ if } 0 \leq t < T\left(g\right)
$$

4.1 Linear preferences

In the following proposition, there are two important dates t_1 and t_2. First, t_1 is the natural herd immunity date, and it was defined in (8): $t_1 \equiv s^{-1}\left(\rho/\beta\right)$. Second, t_2 is the minimal solution to the following equation

$$
t = T\left(A\left[1 - i\left(t\right)\right]\right)
\tag{22}
$$

with $t \geq 0$. We can prove that a minimal positive solution to equation (22) exists. Indeed, t increases from 0 to ∞, and $T\left(A\left[1 - i\left(t\right)\right]\right)$ is a continuous and bounded function with $0 < T\left(A\right) < T\left(A\left[1 - i\left(t\right)\right]\right) \leq T\left(A\left[1 - i\left(t_1\right)\right]\right)$. Thus, the set of solutions to (22) is a compact set of positive real numbers with a minimum, which is positive.

Proposition 7 (linear utility) *The welfare functional (21) is given by*

$$
U\left(g\right) = \frac{A}{\rho}\left[r_0 + s\left(T\left(g\right)\right) - 1\right] - gT\left(g\right)
$$

(1) If $s_0 < \rho/\beta$, then the number of infectives decreases forever, optimal public spending is $g^* = A\left(1 - i_0\right)$ and the optimal research time to discover a vaccine

$$
T^* = \frac{1 - e^{-\gamma A\left(1 - i_0\right)}}{\left[\gamma A\left(1 - i_0\right)\right]^2}
\tag{23}
$$

(2) If $s_0 > \rho/\beta$ there is a peak of infectives associated to a natural herd immunity: $i\left(t_1\right) = 1 - r\left(t_1\right) - \rho/\beta$, with $t_1 \equiv s^{-1}\left(\rho/\beta\right)$.

(2.1) If $t_1 \leq t_2$, then optimal public spending is $g^* = A\left[1 - i\left(t_1\right)\right]$ and the optimal waiting time is

$$T^* = \frac{1 - e^{-\gamma A\left[1 - i\left(t_1\right)\right]}}{\left(\gamma A\left[1 - i\left(t_1\right)\right]\right)^2} \geq t_2 \geq t_1 \tag{24}$$

(2.2) If $t_1 > t_2$, optimal public spending is $g^* = A\left[1 - i\left(t_2\right)\right]$ and the optimal waiting time is

$$T^* = \frac{1 - e^{-\gamma A\left[1 - i\left(t_2\right)\right]}}{\left(\gamma A\left[1 - i\left(t_2\right)\right]\right)^2} = t_2 < t_1 \tag{25}$$

Proof. See Appendix G. ∎

In cases (1) and (2.1) of Proposition 7 the optimal time T^* to obtain the vaccine decreases in the efficacy of research γ since

$$\frac{\partial T^*}{\partial\left(\gamma A\right)} < 0$$

T^* also decreases in A because the upper bound for public spending increases in the labor productivity A. Hence a higher A allows the government for a larger investment, which improves research and speeds up the discovery of the vaccine.

Suppose that herd immunity is already attained at the beginning. Proposition 7 shows that even if herd immunity will naturally make decrease the number of infectives, it is optimal to invest to discover the vaccine. Indeed, a vaccine will speed up the recovery rate, increasing labor supply and as a result, consumption and production. However, note that g^* decreases with the initial number of infectives, i_0, because the larger i_0, the smaller labor, production and the public budget.

Proposition 7 approaches optimal investment according to the initial value of susceptibles. Alternatively, one can also analyze g^* as a function of the initial extent of the disease as measured by i_0. For a given i_0, and leaving constant all other parameters, optimal research investment is higher in case (1), where the number of susceptibles is low. Recall that in case (1) herd immunity is achieved from $t = 0$, the number of infectives decreases monotonically and labor supply increases. Then, why investing more than in the case when the disease is more severe? First, because the government has a larger budget to invest and second, because finding a vaccine will accelerate the cure and the recovery of the labor force. In contrast, in case (2.1), labor supply starts decreasing and it continues decreasing until t_1. Obviously, the spending possibilities are in this case lower than in case (1).

Comparing cases (2.1) and (2.2), optimal public spending is higher in case (2.2) when artificial herd immunity happens before natural herd immunity. Indeed, public investment is fostered by the possibility of increasing labor supply at the earliest.

4.2   Logarithmic preferences

Under logarithmic preferences the welfare functional (21) becomes

$$
\begin{aligned}
U\left(g\right) &= \int_0^\infty \left[\ln c\left(t\right) - \ln A\right] dt \\
&= \int_0^{T(g)} \ln\left[1 - i\left(t\right) - \frac{g}{A}\right] dt + \int_{T(g)}^\infty \ln\left(1 - i\left(T\left(g\right)\right) e^{\rho[T(g)-t]}\right) dt
\end{aligned}
$$

Proposition 8 (logarithmic utility) *Let us consider the approximations for $i(t)$ and $T(g)$ provided by (9) and (15). If the research productivity γ is sufficiently low, then the optimal investment in research is zero, i.e. $g^* = 0$. If on the contrary γ is sufficiently high, then g^* is positive but it remains bounded by initial output, that is $0 < g^* < y_0 = A\left(1 - i_0\right)$.*

Proof. See Appendix H. ∎

In our model, the government can never get indebted, even if it was expecting a great economic growth outburst in the future. The constant amount g^* is bounded by the initial output and it will be continuously invested until the vaccine is discovered. In the logarithmic case, if g tends to y_0, then initial consumption c_0 goes to zero and its marginal felicity, $u'\left(c_0\right)$, to $-\infty$.

In the logarithmic case, research is undertaken only beyond a productivity threshold (Proposition 8). Indeed, when the research productivity is low, the research effort has a significant opportunity cost in terms of consumption.

In order to verify that our approximations do not alter qualitatively the main results and that quantitatively speaking they are accurate enough, let us plot in Figure 3 actual welfare as a function of g for two distinct values of the productivity of medical research, namely $\gamma = 1/2$ and $\gamma = 12$. The remaining parameters of the model are calibrated as in Table 1.
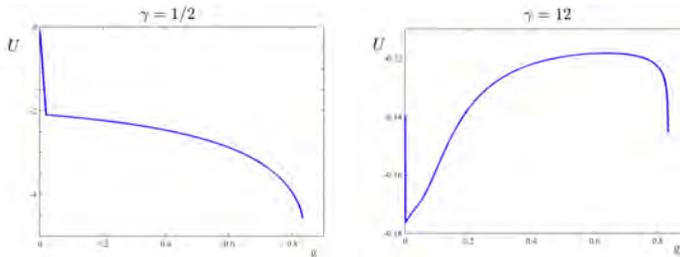
Fig. 3. $U\left(g\right)$ under the Ramsey criterion and logarithmic preferences.

Our simulations corroborate Proposition 8. $g^* = 0$ when γ is small, as in the left panel of Figure 3. On the contrary, when γ is large (right panel), then $g^* > 0$. For $\gamma = 12$ we obtain $g^* = 0.639$ which is indeed smaller than $y_0 = 0.833$.

5   The discounting criterion

In this section we study the second type of policy maker who discounts the future exponentially. The intertemporal utility function (19) becomes

$$
\begin{aligned}
U\left(g\right) &= \int_{0}^{T(g)} e^{-\theta t} u\left(c\left(t\right)\right) dt + \int_{T(g)}^{\infty} e^{-\theta t} u\left(c\left(t\right)\right) dt \\
&= \int_{0}^{T(g)} e^{-\theta t} u\left(A\left[1 - i\left(t\right)\right] - g\right) dt + \int_{T(g)}^{\infty} e^{-\theta t} u\left(A\left(1 - i\left(T\left(g\right)\right) e^{\rho(T(g)-t)}\right)\right) dt
\end{aligned}
$$
(26)

subject to the epidemics dynamics as described in (6), (7) and the positivity constraint for consumption

$$
c\left(t\right) = A\left[1 - i\left(t\right)\right] - g \geq 0, \text{ if } 0 \leq t < T\left(g\right)
$$

As in section 4, we shall study the policy maker's problem under two different descriptions of the household's preferences.

5.1   Linear preferences

Let $u\left(c\right) = c$ and $\rho > \delta$, then

Proposition 9 (linear utility) *Under the approximation for $i(t)$ given in (9), welfare is a function of g and it is given by*

$$
U\left(g\right) = \frac{A - g}{\theta} - \frac{Ai_0}{\delta + \theta} + \frac{g}{\theta} e^{-\theta T(g)} + \left(\frac{1}{\delta + \theta} - \frac{1}{\rho + \theta}\right) Ai_0 e^{-(\delta+\theta)T(g)}
$$

Under the approximation for $T(g)$ in (15), we find that

$$
\lim_{g \to 0^+} U'\left(g\right) \approx -\frac{1}{\theta} < 0
$$

Therefore, it is never optimal to invest in the vaccine research below a threshold. Finally, if the constant

$$
C \equiv \theta \frac{\rho - \delta}{\rho + \theta} \frac{\gamma Ai_0}{\left(\delta + \theta\right)^2} > 0
$$

is sufficiently large and

$$
\frac{\theta}{\delta + \theta} \frac{\rho - \delta}{\rho + \theta} \frac{i_0}{1 - i_0} e^{-\frac{\delta}{\gamma A(1-i_0)}} > e^{\frac{\theta}{\gamma A(1-i_0)}} - 1
$$
(27)

$$
\frac{\theta}{\gamma A\left(1 - i_0\right)} \left[1 + \frac{\rho - \delta}{\rho + \theta} \frac{i_0}{1 - i_0} e^{-\frac{\delta}{\gamma A(1-i_0)}}\right] > e^{\frac{\theta}{\gamma A(1-i_0)}} - 1
$$
(28)

then the optimal public spending is $g^* = A\left(1 - i_0\right)$.

Proof. See Appendix I. ∎

Recall that the approximation for $i(t)$ in (9) is only accurate when $s_0 < \rho/\beta$. Under this premise, we can compare the optimal policy of the time discounting policy maker in Proposition 9 with the altruistic policy maker in case (1) of Proposition 7, being both endowed with linear preferences. The most important and common point is that if engaging in medical research is optimal, then the amount invested g^* is identical in both cases independently of how the policy maker treats the future. However, while investing in medical research is always optimal under the Ramsey criterion, it is not the case under time preference. Indeed, according to Proposition 9, the efficient spending requires the constant C to be large enough. This happens when γ is large, that is, when the research sector is highly productive. According to approximation (15), the more productive is medical research, the shorter the waiting time for a vaccine. Thus, when agents are impatient and they discount the future, engaging in medical research is optimal only if the waiting time for a vaccine is sufficiently short. If the vaccine is expected to arrive in a relatively far future, governments will never engage in research independently of the effects and characteristics of the disease. A government who does not discount the future will always invest (when $s_0 < \rho/\beta$) even if the vaccine takes a long time to arrive.

As in section 4, we proceed next to illustrate Proposition 9 and to verify that our approximations do not alter qualitatively the main results. Let us plot in Figure 4 the actual welfare function using the original equations (6), (7) and (26) for the linear utility and two distinct values of the productivity of medical research, $\gamma = 1/2$ and $\gamma = 12$. Parameters are calibrated as shown in Table 1.
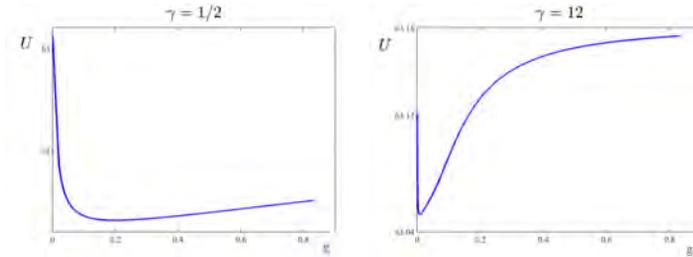
Fig. 4. $U(g)$ under the discounting criterion and linear preferences.

According to Proposition 9, if the policy maker invests in research, then it should invest beyond a certain minimum threshold. We find out that when the research sector is not very productive, then there is no investment in research and $g^* = 0$ (see the left panel of Figure 4). According to Proposition 9, when C is sufficiently large, then the policy maker will invest the maximum that is possible at $t = 0$. Since C increases with γ, we can identify the case $\gamma = 12$ with the case in which C is sufficiently high. In this regard, the right panel

of Figure 4 shows that when γ is high, then the optimal decision is to invest $g^* = y_0 = 0.833$. To complete the picture, let us mention that when $\gamma = 12$, parameter $\delta = 1.9633$ in i's approximation, and $C = 0.0034$.

5.2 Logarithmic preferences

Next, let us consider $u(c) = \ln c$. The welfare functional (26) is given by

$$
\begin{aligned}
U(g) &= \int_0^{T(g)} e^{-\theta t} \ln\left(A\left[1 - i(t)\right] - g\right) dt \\
&+ \int_{T(g)}^{\infty} e^{-\theta t} \ln\left(A\left(1 - i(T(g)) e^{\rho[T(g)-t]}\right)\right) dt
\end{aligned}
$$

This time, $U(g)$ cannot be computed explicitly, and we cannot obtain any insight using fair approximations. As a result, we shall proceed to compute numerically optimal research investment. Figure 5 shows our results for $\gamma = 1/2$ and $\gamma = 12$. The time discount parameter θ is set to $1/100$ and the remaining parameters take the same values as in section 4.

Figure 5 looks impressively close to Figure 3, when the policy maker did not discount the future. The behavior of a policy maker endowed with logarithmic preferences and who uses the true evolution of the epidemics is identical to the behavior of an altruistic policy maker who uses a good approximation. More precisely, it is optimal to engage in medical research only if productivity is sufficiently high. This is equivalent to saying that it is optimal to engage in medical research if and only if the time until the vaccine is discovered is short enough.
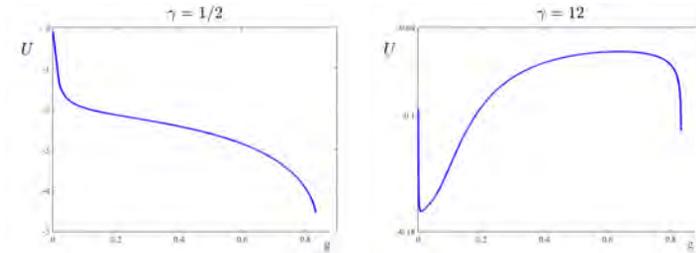
Fig. 5. $U(g)$ under the discounting criterion and logarithmic preferences.

6 Conclusion

This paper considers a unified framework at the crossroad between economics and epidemiology, whose purpose is to study optimal public spending in medical research to obtain a vaccine against an infectious disease evolving according

to a SIR dynamics. Exact optimal policies obtain when policy makers value equally all generations and they have linear preferences. In this case, it is always optimal to invest in medical research independently of the research productivity. Optimal public spending is higher when natural herd immunity has already been attained at the beginning of the planning period. This surprising result comes from the fact that natural herd immunity makes increase labor supply implying a larger public budget. In all other cases, analytical results are obtained using an approximation for the share of infected individuals.

Note that approximations are only legitimate when natural herd immunity is already attained at the beginning. This means that our policy maker starts considering the vaccine problem after the disease sprouts and it is sufficiently well spread. Both numerically and theoretically we find that even if herd immunity is attained, technologically advanced economies should optimally invest in medical research to accelerate the recovery of the population and the economy. Noteworthy, we show numerically that the optimal investment decision of a policy maker using an accurate approximation for the evolution of the epidemics, logarithmic preferences, and caring equally about the present and the future, is similar to that of a policy maker using the true evolution of the epidemics, and that of a policy maker who additionally discounts the future. That is, if an economy has a sufficiently performing technology, they should invest as much as their initial budget allows them to.

Appendix

A. Proof of Proposition 1
According to the assumptions of the SIR model, the three categories of population evolve in time according to:

$$S'(t) = -\beta \frac{I(t)}{N(t)} S(t) \tag{29}$$

$$I'(t) = \beta \frac{I(t)}{N(t)} S(t) - \rho I(t) \tag{30}$$

$$R'(t) = \rho I(t) \tag{31}$$

Note that (29)-(31) satisfy (1) since $S'(t) + I'(t) + R'(t) = 0$.

In order to reach equations (3) and (4) from the previous system, one needs to compute the time derivatives of the shares defined in (2), replace S', R' and I' using equations (29) to (31), and finally take into account that $s(t) + i(t) + r(t) = 1$. We obtain that $r'(t) = \rho i(t)$, $s'(t) = -\beta s(t) i(t)$ and

$$i'(t) = -r'(t) - s'(t) = [\beta s(t) - \rho] i(t) \tag{32}$$

∎

B. Proof of Proposition 2
Since s is monotonic, we can define $r(t) = \tilde{r}(s(t)) = r(s^{-1}(s))$, where $t = s^{-1}(s)$ is the inverse function of $s = s(t)$. Clearly, $r'(t) = \tilde{r}'(s)s'(t)$. Using $r'(t) = \rho i(t)$

and equation (3), we obtain

$$\tilde{r}'(s) = \frac{r'(t)}{s'(t)} = -\frac{\rho}{\beta}\frac{1}{s}$$

Integrating both sides of the expression above between 0 and t, that is between s_0 and $s(t)$, we obtain that

$$
\begin{aligned}
\int_{s_0}^{s(t)} \tilde{r}'(s)\,ds &= -\frac{\rho}{\beta}\int_{s_0}^{s(t)}\frac{1}{s}ds \\
\tilde{r}(s(t)) - \tilde{r}(s_0) &= -\frac{\rho}{\beta}\left[\ln s(t) - \ln s_0\right] \\
r(t) &= r_0 + \frac{\rho}{\beta}\ln s_0 - \frac{\rho}{\beta}\ln s(t) \quad (33)
\end{aligned}
$$

Therefore, replacing this expression for $r(t)$ in (3), we obtain the single differential equation (7). Moreover, (33) implies that

$$s(t) = s_0 e^{\frac{\beta}{\rho}[r_0 - r(t)]} \quad (34)$$

Finally, using that $r'(t) = \rho i(t)$, we obtain (6).

We prove next that $s(t)$ decreases from s_0 to a given critical value that we also compute. Let us define function ϕ as $\phi(s) \equiv q + \beta s - \rho \ln s$, where $q \equiv \rho \ln s_0 - \beta(1-r_0) < 0$. Then, the dynamic equation (7) can be written as

$$s'(t) = s(t)\phi(s(t)) \quad (35)$$

Equation $\phi(s) = 0$ has exactly two solutions s_1 and s_2 with $0 < s_1 < s_0 < s_2$ and $\phi(s) < 0$ if $s_1 < s \le s_0$. Observe that ϕ is a continuous and strictly convex function on $(0, +\infty)$. Indeed,

$$\phi''(s) = \frac{\rho}{s^2} > 0$$

Moreover, $\lim_{t \to 0^+}\phi(s) = +\infty$, $\lim_{t \to +\infty}\phi(s) = +\infty$ and

$$\phi(s_0) = -\beta i_0 < 0$$

Then, ϕ crosses the horizontal axis once on the left and once on the right of the point $s = s_0$. Note that $s_1 < 1$ because $s_0 < 1$ and $s_1 < s_0$.

According to (35) and the previous results, if $s_1 < s(t) \le s_0$, then $s'(t) < 0$. Thus, $s(t)$ decreases over time and converges to s_1. Since $r'(t) = \rho i(t)$, $r(t)$ increases over time and converges to $r_1 \le 1 - s_1$ with $\lim_{t \to \infty} r'(t) = 0$. If $r_1 < 1 - s_1$, we obtain $\lim_{t \to \infty} r'(t) = \rho(1 - r_1 - s_1) > 0$, which is a contradiction.

The share of infectives in total population is given by $i(t) \equiv I(t)/N(t) = 1 - r(t) - s(t)$. Since the population is constant over time, we obtain

$$\lim_{t \to \infty} I(t) = N\left[1 - \lim_{t \to \infty} r(t) - \lim_{t \to \infty} s(t)\right] = N(1 - r_1 - s_1) = 0$$

Finally, let us prove that the share of susceptibles falls below ρ/β. We know that the disease is eradicated in the long run, that is $\lim_{t\to\infty} i(t) = 0$. Hence, there exists a date τ such that $\dot{i}(\tau) < 0$. Therefore,

$$ i'(\tau) = [\beta s(\tau) - \rho] i(\tau) < 0 $$

that is $s(\tau) < \rho/\beta$. As shown above $s(t)$ is always decreasing, which implies that $s_1 < \rho/\beta$. ∎

C. Proof of Proposition 3.

According to Proposition 1, $i(t)$ increases if $s(t) > \rho/\beta$ and decreases if $s(t) < \rho/\beta$. Since $\dot{s}(t) < 0$ and $s_1 < \rho/\beta$, if $s_0 > \rho/\beta$, we have that $s(t)$ first decreases from s_0 to ρ/β and, then, from ρ/β to s_1. In this case, if t_1 denotes the critical date such that $s(t_1) = \rho/\beta$, and since $i'(t) = [\beta s(t) - \rho] i(t)$, we have that the share of infectives $i(t)$ always increases from 0 to t_1 and always decreases (to zero) from t_1 to infinite. We call t_1, the date of herd immunity and $s(t_1)$ the critical share.

If $s_0 > \rho/\beta$, $s^* \equiv s(\rho/\beta)$ is the critical share of susceptibles for a natural herd immunity and the peak of infectives is given by $i(t_1) = 1 - r(t_1) - s(t_1) = 1 - r\left(s^{-1}(\rho/\beta)\right) - \rho/\beta$.

Summarizing our results, herd immunity is attained at t_1, where

$$ t_1 = \begin{cases} 0, & \text{if} \quad s_0 \leq \rho/\beta \\ \\ s^{-1}\left(\frac{\rho}{\beta}\right) > 0, & \text{otherwise} \end{cases} $$

∎

D. Proof of Corollary 4.

A more infective disease means a lower ratio ρ/β. We observe that

$$ \frac{\partial t_1}{\partial(\rho/\beta)} = \frac{1}{s'(t_1)} < 0 $$

because $s(t)$ is always decreasing. ∎

E. Proof of Proposition 5.

According to (3) and (9), we have

$$ \frac{s'(t)}{s(t)} \approx -\beta i_0 e^{-\delta t} $$

Integrating on both sides,

$$ \int_0^t [\ln s(\tau)]' \, d\tau \quad \approx \quad -\beta i_0 \int_0^t e^{-\delta \tau} d\tau $$

$$ \ln \frac{s(t)}{s_0} \quad \approx \quad \frac{\beta}{\delta} i_0 \left(e^{-\delta t} - 1\right) $$

that is (10). Replacing (10) in (33), (11) obtains.

Finally, we notice that

$$r_1 = \lim_{t \to \infty} r(t) \approx r_0 + \frac{\rho}{\delta} i_0 \tag{36}$$

$$s_1 = \lim_{t \to \infty} s(t) \approx s_0 e^{-\frac{\beta}{\delta} i_0} \tag{37}$$

Since $\lim_{t \to \infty} [r(t) + s(t)] = 1$, then $r_1 + s_1 = 1$ and, replacing (36) and (37), we get equation (12). ■

F. Proof of Lemma 6.

According to (17), the non-negativity of consumption is equivalent to $c(t) = A[1 - i(t)] - g \geq 0$ for any $t \in [0, T(g))$. That is

$$g \leq A \min_{t \in [0, T(g))} [1 - i(t)] = A \left[1 - \max_{t \in [0, T(g))} i(t) \right]$$

(1) Consider the case $s_0 \leq \rho/\beta$. In this case, the date of natural herd immunity is $t_1 = 0$ and the number of infectives always decreases: $i'(t) < 0$ for any $t > 0$. Since $T(g) \geq t_1 = 0$, the non-negativity of consumption in the interval $[0, T(g))$ is equivalent to $g \leq A[1 - i(t_1)] = A(1 - i_0)$.

(2) Consider the case $s_0 > \rho/\beta$, that is $t_1 = s^{-1}(\rho/\beta)$.

(2.1) If $t_1 \leq T(g)$, the vaccine is discovered after t_1. Then, the non-negativity of consumption in the interval $[0, T(g))$ is equivalent to

$$g \leq A[1 - i(t_1)] = A[1 - r(t_1) - s(t_1)] = A\left[1 - r\left(s^{-1}(\rho/\beta)\right) - \rho/\beta\right]$$

(2.2) If $t_1 > T(g)$, the vaccine is discovered before the date of natural herd immunity t_1. Since $i'(t) > 0$ for any $t \leq T(g) (< t_1)$, then the non-negativity of consumption in the interval $[0, T(g))$ is equivalent to $g \leq A[1 - i(T(g))]$. ■

G. Proof of Proposition 7.

Using (17) we have

$$
\begin{aligned}
U(g) &= \int_0^{\infty} [c(t) - A] \, dt \\
&= -\int_{T(g)}^{\infty} Ai(T(g)) e^{\rho[T(g)-t]} dt - \int_0^{T(g)} [Ai(t) + g] \, dt \\
&= -\frac{A}{\rho} i(T(g)) - \int_0^{T(g)} [Ai(t) + g] \, dt
\end{aligned}
$$

Focus on the last integral. Since $r'(t) = \rho i(t)$, we obtain

$$
\begin{aligned}
-\int_0^{T(g)} [Ai(t) + g] \, dt &= -\frac{A}{\rho} \int_0^{T(g)} \dot{r}(t) \, dt - \int_0^{T(g)} g \, dt \\
&= \frac{A}{\rho} [r_0 - r(T(g))] - gT(g)
\end{aligned}
$$

Thus, we find

$$
\begin{aligned}
U(g) &= \frac{A}{\rho}\left[r_0 - r(T(g))\right] - gT(g) - \frac{A}{\rho} i(T(g)) \\
&= \frac{A}{\rho}\left[r_0 + s(T(g)) - 1\right] - gT(g)
\end{aligned}
$$

Note that

$$
U'(g) = \frac{A}{\rho} s'(T(g)) T'(g) - \left[T(g) + gT'(g)\right] > 0
$$

because $s'(T(g)) < 0$, $T'(g) < 0$ and

$$
T(g) + gT'(g) = \frac{(1+\gamma g)e^{-\gamma g} - 1}{(\gamma g)^2} < 0
$$

Notice that $T(g) + gT'(g) < 0$ implies $T'(g) < 0$.

We know from (17) that

$$
c(t) = \begin{cases} A[1 - i(t)] - g, & \text{if } 0 \le t < T \\ A\left[1 - i(T)e^{\rho(T-t)}\right], & \text{otherwise} \end{cases}
$$

Since $U'(g) > 0$, g should be such that $c(t) \ge 0$ for any t. Note that $c(t) > 0$ for any $t \ge T(g)$. Therefore, c should verify that $c(t) = A[1 - i(t)] - g \ge 0$ for any $t \in [0, T(g))$.

(1) According to (4), $i'(t)/i(t) = \beta s(t) - \rho$. Thus, if $s_0 < \rho/\beta$, the share of infectives decrease forever, optimal public spending is given by $g^* = A(1 - i_0)$ and according to (14), the optimal waiting time for a vaccine is given by (23).

(2) If $s_0 > \rho/\beta$, then there is a peak of infectives associated to a natural herd immunity: $i(t_1) = 1 - r(t_1) - \rho/\beta$ with $t_1 \equiv s^{-1}(\rho/\beta)$. We consider two sub-cases:

(2.1) If $t_1 \le t_2$, then $T^* \equiv T(g^*) \ge t_2 \ge t_1$. Indeed, $i(t_2) \le i(t_1)$, $g^* = A[1 - i(t_1)] \le A[1 - i(t_2)]$ and $T(g^*) = T(A[1 - i(t_1)]) \ge T(A[1 - i(t_2)]) = t_2 \ge t_1$. According to (14), the optimal waiting time is given by (24).

(2.2) If $t_1 > t_2$, then $T(g^*) = t_2 < t_1$. Indeed, we are looking for the maximal g^* compatible with $g \le A[1 - i(t)]$. The maximal g^* corresponds to the minimal $T(g^*)$ such that $T(g^*) \ge T(A[1 - i(t)])$ for any $t \le T(g^*)$ that is to the fixed point $T(g^*) = T(A[1 - i(T(g^*))]) = t_2$. Therefore, $g^* = T^{-1}(t_2) = T^{-1}T(A[1 - i(t_2)]) = A[1 - i(t_2)]$ is the optimal public spending, while the optimal waiting time for a vaccine is given by (25). ∎

H. Proof of Proposition 8.

$U'(g)$ is computed applying Leibniz' rule:

$$
\begin{aligned}
U'(g) &= T'(g)\ln\left[1 - \frac{g}{A[1 - i(T(g))]}\right] - \int_0^{T(g)} \frac{1}{A[1 - i(t)] - g} dt \\
&\quad - T'(g)V(g)i(T(g))
\end{aligned} \tag{38}
$$

where

$$V\left(g\right) \equiv \int_{T(g)}^{\infty} \frac{\rho + \frac{\dot{i}(T(g))}{i(T(g))}}{e^{\rho[t-T(g)]} - i\left(T\left(g\right)\right)} dt$$

Using (4), $V\left(g\right)$ writes as

$$
\begin{aligned}
V\left(g\right) &= \int_{T(g)}^{\infty} \frac{\beta s\left(T\left(g\right)\right)}{e^{\rho[t-T(g)]} - i\left(T\left(g\right)\right)} dt \\
&= \int_{T(g)}^{\infty} \frac{\beta s\left(T\left(g\right)\right) e^{\rho T(g)}}{e^{\rho t} - a} dt = \frac{\beta s\left(T\left(g\right)\right) e^{\rho T(g)}}{a\rho} \left[\ln\left(e^{\rho t} - a\right) - \rho t\right]_{T(g)}^{\infty}
\end{aligned}
$$

where $a \equiv e^{\rho T(g)} i\left(T\left(g\right)\right)$. Then

$$
\begin{aligned}
V\left(g\right) &= \frac{\beta s\left(T\left(g\right)\right) e^{\rho T(g)}}{a\rho} \left[\lim_{t\to\infty}\left[\ln\left(e^{\rho t} - a\right) - \rho t\right] - \ln\left(e^{\rho T(g)} - a\right) + \rho T\left(g\right)\right]_{T(g)}^{\infty} \\
&= \frac{\beta s\left(T\left(g\right)\right) e^{\rho T(g)}}{a\rho} \left[\lim_{t\to\infty} \ln e^{\ln\left(e^{\rho t} - a\right) - \rho t} - \ln\left[1 - i\left(T\left(g\right)\right)\right]\right]
\end{aligned}
$$

Noticing that

$$
\begin{aligned}
\lim_{t\to\infty} \ln e^{\ln\left(e^{\rho t} - a\right) - \rho t} &= \ln \lim_{t\to\infty} \left[e^{-\rho t} e^{\ln\left(e^{\rho t} - a\right)}\right] = \ln \lim_{t\to\infty} \left[e^{-\rho t}\left(e^{\rho t} - a\right)\right] \\
&= \ln \lim_{t\to\infty}\left(1 - a e^{-\rho t}\right) = \ln 1 = 0
\end{aligned}
$$

we obtain

$$V\left(g\right) = -s\left(T\left(g\right)\right) \frac{\beta}{\rho} \frac{\ln\left[1 - i\left(T\left(g\right)\right)\right]}{i\left(T\left(g\right)\right)}$$

Replacing $V\left(g\right)$ in (38), $U'\left(g\right)$ can be written as

$$
\begin{aligned}
U'\left(g\right) &= T'\left(g\right)\left(\ln\left[1 - \frac{g}{A\left[1 - i\left(T\left(g\right)\right)\right]}\right] + \frac{\beta}{\rho} s\left(T\left(g\right)\right) \ln\left[1 - i\left(T\left(g\right)\right)\right]\right) \\
&\quad - \int_0^{T(g)} \frac{1}{A\left[1 - i\left(t\right)\right] - g} dt
\end{aligned}
$$

Note that the first term is positive, while the second term with the integral is negative.

Under the approximation for $i(t)$ in (9)

$$
\begin{aligned}
\int_0^{T(g)} \frac{1}{A\left[1 - i\left(t\right)\right] - g} dt &\approx \int_0^{T(g)} \frac{1}{A - g - A i_0 e^{-\delta t}} dt \\
&= \left[\frac{1}{(A-g)\delta} \ln\left[1 - \frac{(A-g) e^{\delta t}}{A i_0}\right]\right]_0^{T(g)} \\
&= \frac{1}{(A-g)\delta} \ln \frac{(A-g) e^{\delta T(g)} - A i_0}{A\left(1 - i_0\right) - g}
\end{aligned}
$$

Finally, noticing that $i\left(T\left(g\right)\right) \approx i_0 e^{-\delta T(g)}$,

$$
\begin{aligned}
U'\left(g\right) &\approx T'\left(g\right)\left[\ln \frac{(A-g) e^{\delta T(g)} - A i_0}{A e^{\delta T(g)} - A i_0} + \frac{\beta}{\rho} s\left(T\left(g\right)\right) \ln\left[1 - i_0 e^{-\delta T(g)}\right]\right] \\
&\quad - \frac{1}{(A-g)\delta} \ln \frac{(A-g) e^{\delta T(g)} - A i_0}{A - g - A i_0}
\end{aligned}
$$

where $i_0 = i(0)$.

Replacing the approximation for $T(t)$ in (15):

$$U'(g) \approx \frac{1}{\gamma g^2}\left[\ln\frac{Ae^{\frac{\delta}{\gamma g}} - Ai_0}{(A-g)e^{\frac{\delta}{\gamma g}} - Ai_0} - \frac{\beta}{\rho}s\left(\frac{1}{\gamma g}\right)\ln\left(1 - i_0 e^{-\frac{\delta}{\gamma g}}\right)\right]$$
$$+ \frac{1}{(A-g)\delta}\ln\frac{A - g - Ai_0}{(A-g)e^{\frac{\delta}{\gamma g}} - Ai_0} \tag{39}$$

The term between the brackets is positive while the term on the second line is negative, and the sign of $U'(g)$ is *a priori* ambiguous.

To understand the role of research productivity γ, we compute the limits $\lim_{\gamma\to 0^+} U'(g)$ and $\lim_{\gamma\to\infty} U'(g)$ for all values of g. Consider first $\lim_{\gamma\to 0^+} U'(g)$. Since $s(T(g)) < s_0$,

$$U'(g) < \frac{1}{\gamma g^2}\left[\ln\frac{A - Ai_0 e^{-\frac{\delta}{\gamma g}}}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}} - \frac{\beta}{\rho}s_0\ln\left(1 - i_0 e^{-\frac{\delta}{\gamma g}}\right)\right] + \frac{1}{(A-g)\delta}\ln\frac{(A - g - Ai_0)e^{-\frac{\delta}{\gamma g}}}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}}$$

$$= \left[\frac{1}{(A-g)\delta} + \frac{\ln\frac{A - Ai_0 e^{-\frac{\delta}{\gamma g}}}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}} - \frac{\beta}{\rho}s_0\ln\left(1 - i_0 e^{-\frac{\delta}{\gamma g}}\right)}{\gamma g^2 \ln\frac{A - g - Ai_0}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}} - \delta g}\right]\ln\frac{(A - g - Ai_0)e^{-\frac{\delta}{\gamma g}}}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}}$$

Thus,

$$\lim_{\gamma\to 0^+} U'(g) \leq \frac{1}{A\delta}\left[\frac{1}{1 - \frac{g}{A}} + \frac{\ln\left(1 - \frac{g}{A}\right)}{\frac{g}{A}}\right]\left(\ln\frac{A - g - Ai_0}{A - g} - \lim_{\gamma\to 0^+}\frac{\delta}{\gamma g}\right)$$

It is easy to prove that

$$\frac{1}{1 - \frac{g}{A}} + \frac{\ln\left(1 - \frac{g}{A}\right)}{\frac{g}{A}} > 0 \text{ for } \frac{g}{A} \in (0,1)$$

$$\lim_{g\to 0^+}\left[\frac{1}{1 - \frac{g}{A}} + \frac{\ln\left(1 - \frac{g}{A}\right)}{\frac{g}{A}}\right] = 0$$

Therefore $\lim_{\gamma\to 0^+} U'(g) = -\infty$, for any $g > 0$.

Consider now $\lim_{\gamma\to\infty} U'(g)$. Since $s(T(G)) > s_1$

$$U'(g) > \frac{1}{\gamma g^2}\left[\ln\frac{A - Ai_0 e^{-\frac{\delta}{\gamma g}}}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}} - \frac{\beta}{\rho}s_1\ln\left(1 - i_0 e^{-\frac{\delta}{\gamma g}}\right)\right] - \frac{1}{(A-g)\delta}\ln\frac{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}}{(A - g - Ai_0)e^{-\frac{\delta}{\gamma g}}}$$

$$= \left[\frac{\ln\frac{A - Ai_0 e^{-\frac{\delta}{\gamma g}}}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}} - \frac{\beta}{\rho}s_1\ln\left(1 - i_0 e^{-\frac{\delta}{\gamma g}}\right)}{\frac{g}{A-g}\left(1 - \frac{\gamma g}{\delta}\ln\frac{A - g - Ai_0}{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}}\right)} - 1\right]\frac{1}{(A-g)\delta}\ln\frac{A - g - Ai_0 e^{-\frac{\delta}{\gamma g}}}{(A - g - Ai_0)e^{-\frac{\delta}{\gamma g}}} \tag{40}$$

Observe that

$$\lim_{\gamma \to \infty} \left[\frac{g}{A-g} \left(1 - \frac{\gamma g}{\delta} \ln \frac{A-g-Ai_0}{A-g-Ai_0 e^{-\frac{\delta}{\gamma g}}} \right) \right] = \frac{g}{A-g-Ai_0}$$

Therefore

$$\lim_{\gamma \to \infty} \left[\frac{\ln \frac{A-Ai_0 e^{-\frac{\delta}{\gamma g}}}{A-g-Ai_0 e^{-\frac{\delta}{\gamma g}}} - \frac{\beta}{\rho} s_1 \ln \left(1 - i_0 e^{-\frac{\delta}{\gamma g}} \right)}{\frac{g}{A-g} \left(1 - \frac{\gamma g}{\delta} \ln \frac{A-g-Ai_0}{A-g-Ai_0 e^{-\frac{\delta}{\gamma g}}} \right)} - 1 \right] = -\frac{A(1-i_0)-g}{g} \left[\psi(g) + \frac{\beta}{\rho} s_1 \ln(1-i_0) \right]$$

$$(41)$$

where

$$\psi(g) \equiv \frac{g}{A-Ai_0-g} + \ln \frac{A-Ai_0-g}{A-Ai_0}$$

Since $g < A(1-i_0)$, the limit (41) is positive if and only if

$$\psi(g) < -\frac{\beta}{\rho} s_1 \ln(1-i_0) \ (>0) \tag{42}$$

Moreover

$$\lim_{g \to 0+} \psi(g) = 0 \quad \text{and} \quad \psi'(g) = \frac{g}{(A-Ai_0-g)^2} > 0$$

Hence there exists a critical value of g, $\bar{g} \in (0, A(1-i_0))$, which does not depend on γ, and such that (42) is satisfied, that is, the limit (41) is positive.

Notice that

$$\frac{1}{(A-g)\delta} \ln \frac{A-g-Ai_0 e^{-\frac{\delta}{\gamma g}}}{(A-g-Ai_0) e^{-\frac{\delta}{\gamma g}}} > 0$$

Reconsidering (40), there exists a sufficiently high γ such that $U'(g) > 0$ for any $g \in (0, \bar{g})$. This implies that, for a sufficiently high γ, $g^* > 0$.

Finally, note that the optimal amount g^* satisfies that $g^* < A(1-i_0)$. Indeed, consumption non-negativity requires that $g \leq A(1-i_0)$ and, according to (39),

$$\lim_{g \to A(1-i_0)^-} U'(g) \approx Z + \frac{1}{A\delta i_0} \lim_{g \to A(1-i_0)^-} \ln[A(1-i_0)-g] = -\infty$$

where

$$Z \equiv \frac{1}{\gamma A^2 (1-i_0)^2} \left[\ln \frac{e^{\frac{\delta}{\gamma A(1-i_0)}} - i_0}{i_0 e^{\frac{\delta}{\gamma A(1-i_0)}} - i_0} - \frac{\beta}{\rho} s \left(\frac{1}{\gamma A(1-i_0)} \right) \ln \left[1 - i_0 e^{-\frac{\delta}{\gamma A(1-i_0)}} \right] \right]$$

$$- \frac{\ln(Ai_0) + \ln \left[e^{\frac{\delta}{\gamma A(1-i_0)}} - 1 \right]}{A\delta i_0}$$

is a finite real number. ∎

I. Proof of Proposition 9.

Using the approximation for $i(t)$ in (9)

$$
\begin{aligned}
U(g) &\approx \int_0^{T(g)} (A-g) e^{-\theta t} dt - \int_0^{T(g)} A i_0 e^{-(\delta+\theta)t} dt \\
&+ \int_{T(g)}^{\infty} A e^{-\theta t} dt - \int_{T(g)}^{\infty} A i\left(T(g)\right) e^{\rho T(g) - (\rho+\theta)t} dt \\
&= -\frac{A-g}{\theta}\left[e^{-\theta T(g)} - 1\right] + \frac{A i_0}{\delta+\theta}\left[e^{-(\delta+\theta)T(g)} - 1\right] + \frac{A}{\theta} e^{-\theta T(g)} - \frac{A i\left(T(g)\right)}{\rho+\theta} e^{-\theta T(g)} \\
&= \frac{A-g}{\theta} - \frac{A i_0}{\delta+\theta} + \frac{g}{\theta} e^{-\theta T(g)} + \left(\frac{1}{\delta+\theta} - \frac{1}{\rho+\theta}\right) A i_0 e^{-(\delta+\theta)T(g)}
\end{aligned}
$$

because $i\left(T(g)\right) \approx i_0 e^{-\delta T(g)}$.

If $T(g) \approx 1/(\gamma g)$, $U(g)$ becomes

$$
U(g) \approx \frac{1}{\theta}\left(A - g + g e^{-\frac{\theta}{\gamma g}}\right) + \frac{A i_0}{\delta+\theta}\left(\frac{\rho-\delta}{\rho+\theta} e^{-\frac{\delta+\theta}{\gamma g}} - 1\right) \tag{43}
$$

Its derivative is given by

$$
U'(g) \approx \frac{1}{\theta}\left[v_1(g) + C v_2(g) - 1\right] \tag{44}
$$

where

$$
C \equiv \theta \frac{\rho-\delta}{\rho+\theta} \frac{\gamma A i_0}{(\delta+\theta)^2} > 0
$$

and

$$
v_1(g) \equiv \left(1 + \frac{\theta}{\gamma g}\right) e^{-\frac{\theta}{\gamma g}} \text{ and } v_2(g) \equiv \left(\frac{\delta+\theta}{\gamma g}\right)^2 e^{-\frac{\delta+\theta}{\gamma g}}
$$

Note that

$$
U''(g) \approx \frac{1}{\theta}\left[v_1'(g) + C v_2'(g)\right]
$$

Moreover, $v_1(g) \in [0,1)$ and $\lim_{g\to 0+} v_1(g) = \lim_{g\to 0+} v_2(g) = 0$, which implies $\lim_{g\to 0+} U'(g) = -1/\theta < 0$.

Thus, if C is sufficiently large,

$$
U''(g) \approx \frac{C}{\theta} v_2'(g)
$$

$v_2(g)$ is a single-peaked function, increasing from 0 to $\max_g v(g) > 0$, and then, decreasing asymptotically towards 0 with a maximum at

$$
\hat{g} \equiv \frac{\delta+\theta}{2\gamma}
$$

Hence, if C is sufficiently large, $U(g)$ is convex for $g < \hat{g}$ and concave thereafter, approximately.

Summing up, the set of sufficient conditions for $g^* = A(1 - i_0)$ to be optimal when C is sufficiently large, are

$$
\begin{aligned}
U\left(A(1-i_0)\right) &> U(0) \\
U'\left(A(1-i_0)\right) &> 0
\end{aligned}
$$

■

References

[1] Barrett S. and M. Hoel (2007), "Optimal disease eradication", *Environment and Development Economics*, 12, 627–652.

[2] Brito D., Sheshinski E. and M. Intriligator (1991), "Externalities and compulsory vaccinations", *Journal of Public Economics*, 45, 69-90.

[3] Callaway E. (2020), "The race for coronavirus vaccines", *Nature*, 580, 576-577.

[4] Cass D. (1965), "Optimum growth in an aggregative model of capital accumulation", *Review of Economic Studies*, 32, 233-240.

[5] Geoffard P.-Y. and T. Philipson (1997), "Disease eradication: Private versus public vaccination", *The American Economic Review*, 87, 222-230.

[6] Hethcote H.W. (1976), "Qualitative analyses of communicable disease models", *Mathematical Biosciences*, 28, 335-356.

[7] Hethcote H.W. and P. Waltman (1973), "Optimal vaccination schedules in a deterministic epidemic model", *Mathematical Biosciences*, 18, 365-381.

[8] Koopmans T. (1965), "On the concept of optimal economic growth", in: *The Econometric Approach to Development Planning*, Rand-McNally, Chicago, 225-287.

[9] Kwok K., Lai F., Wong W., and J. Tang (2020), "Herd immunity - estimating the level required to halt the COVID-19 epidemics in affected countries", *Journal of Infection* 80, e32-e33.

[10] Roda W.C., Varughese M.B., Han D. and Li M.Y. (2020), "Why is it difficult to accurately predict the COVID-19 epidemic?", forthcoming in *Infectious Disease Modelling*, 271-281.

[11] Ramsey F.P. (1928), "A mathematical theory of saving", *Economic Journal*, 38, 543-559.

[12] Randolph H. and L. Barreiro. (2020), "Herd immunity: Understanding COVID-19", *Immunity*, 52, 737-741.

[13] Rossman J. (2020), "Herd immunity in Europe - are we close?", *The Conversation*, May 28.

The economic consequences of $\hat{R}$ = 1: Towards a workable behavioural epidemiological model of pandemics

Joshua S. Gans[1]

This paper reviews the literature on incorporating behavioural elements into epidemiological models of pandemics. While modelling behaviour by forward-looking rational agents can provide some insight into the time paths of pandemics, the non-stationary nature of Susceptible-Infected-Removed (SIR) models of viral spread makes characterisation of resulting equilibria difficult. Here I posit a shortcut that can be deployed to allow for a tractable equilibrium model of pandemics with intuitive com- parative statics and also a clear prediction that effective reproduction numbers (that is, R) will tend towards 1 in equilibrium. This motivates taking $\hat{R}$ = 1 as an equilibrium starting point for analyses of pandemics with behavioural agents. The implications of this for the analysis of widespread testing, tracing, isolation and mask-use is discussed.

1   Professor of Strategic Management, Rotman School of Management, University of Toronto.

1 Introduction

The workhorse model for the modelling of epidemics is the SIR (Susceptible-Infected-Removed) model of (30). It has been adopted to inform policy-makers in the management of the COVID-19 pandemic. The model is mechanistic in that people in the model do not make decisions that are reactive to current and predicted prevalence of an infectious disease in the population. As a key parameter, the basic reproduction number, $\mathcal{R}_0$, (a measure of the expected number of infections generated by a single infected person) is driven by people's choices regarding physical interactions. For this reason, the lack of behavioural elements has been a persistent source of criticism of such models.

This paper argues that, while a full behavioural model of pandemics is difficult to analyse as there is an element of non-stationarity in dynamic outcomes, there is value to be gained by analysing models that generate predictions that, for considerable lengths of time, the equilibrium reproduction number, $\hat{\mathcal{R}}$ is equal to 1 implying that the prevalence of an infectious disease/virus is constant over time with the number of those newly infected approximately equaling the number of those newly recovered in a given time period. For COVID-19, such outcomes have been observed empirically beyond the initial stages of outbreaks across many regions (see Figure 1).[1]

Models that can generate an $\hat{\mathcal{R}} = 1$ equilibrium exist in the literature. For the SIR model whereby infectious individuals who recover are removed from the susceptible pool, I show that an $\hat{\mathcal{R}} = 1$ outcome requires a special set of assumptions that are unlikely to generally hold. This is because individuals may base their behaviour on prevalence (i.e., the number of infected people they are likely to encounter) rather than on the ever falling set of susceptibles. That set, however, does impact on the reproduction number. Nonetheless, for the SIS model, whereby infectious individuals who recover remain susceptible to future infections, the $\hat{\mathcal{R}} = 1$ outcome is a natural equilibrium. This suggests that, when prevalence is relatively low, even for the SIR model, the number of susceptibles will not change at a rapid pace and thus, an $\hat{\mathcal{R}} = 1$ outcome provides an approximate outcome that may explain observed behaviour.

In what follows, I first present the standard (non-behavioural) SIR model. I then review various behavioural models that have been utilised in the literature deriving. I provide a graphical approach to describe the resulting equilibrium outcomes. A final section offers some predictions from this approach.

---

[1]Moreover, there is plenty of evidence that people act to mitigate their own infection risk apart from those mandated by governments. See (Farboodi et al.) and (26)).
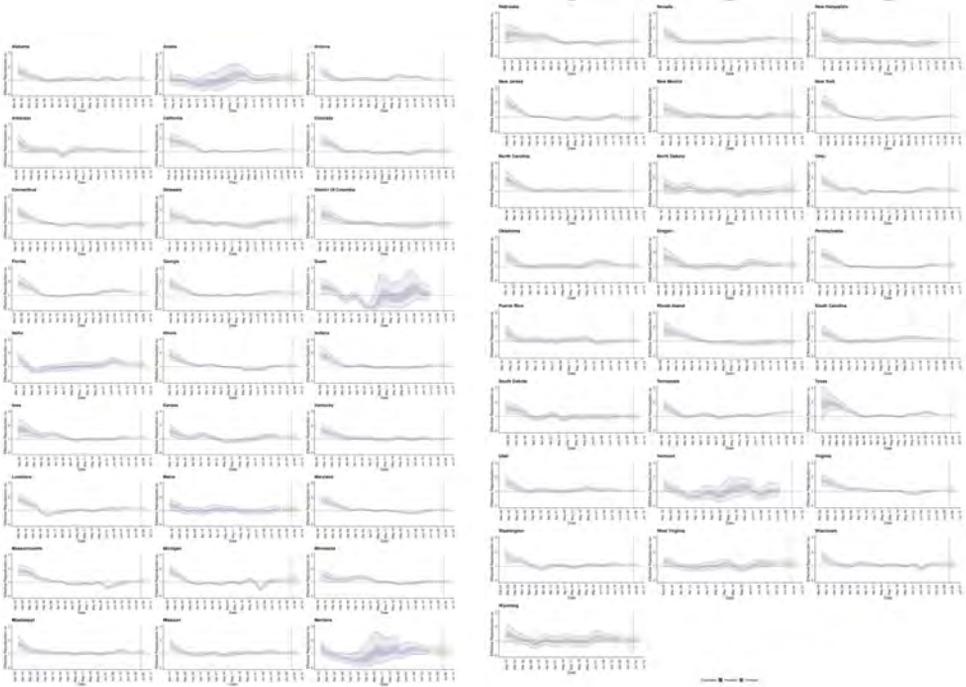
Figure 1: Estimated $\mathcal{R}_t$ for US States (COVID-19) from epiforecasts.io

2   The Standard SIR Model

Let $\{S(t), I(t), R(t)\}$ denote the shares (and levels) of the population (normalised to be of size 1 over a continuum of agents) who are either susceptible to the virus, infected with the virus or removed (i.e., recovered or dead) from the virus at time $t \geq 0$. It is assumed that time is discrete. In the SIR model, these variables are assumed to evolve according to the following dynamic equations:

$$S(t+1) - S(t) = -\beta S(t)I(t)$$

$$I(t+1) - I(t) = (\beta S(t) - \gamma)I(t)$$

$$R(t+1) - R(t) = \gamma I(t)$$

Here γ is the probability that an infected person will be removed in any given period while β is the probability that a susceptible person will become infected by an infected person in a given period. Observe that the number of infections in the population will be falling (i.e.,

$I(t+1) < I(t))$ if $\frac{\beta}{\gamma}S(t) < 1$ and will be rising (i.e., $I(t+1) > I(t))$ if $\frac{\beta}{\gamma}S(t) > 1$. The LHS of these inequalities is the effective reproduction number, $\mathcal{R}_t$. Since $S(0) \approx 1$, then $\mathcal{R}_0 = \frac{\beta}{\gamma}$. $\mathcal{R}_0$ is the basic reproduction number which has the interpretation as the total expected number of infections one infectious person will create over the life of their infection.

A few remarks about this model. First, there are two relevant state variables $\{I(t), S(t)\}$ and they co-evolve according to:

$$I(t) = 1 - S(t) + \frac{1}{\mathcal{R}_0}\log(S(t))$$

.

where it is assumed that $\{I(0), S(0)\} = \{0, 1\}$. Second, the share of the population that is eventually infected, $i \equiv 1 - S(\infty)$, is given by:

$$\mathcal{R}_0 = -\frac{\log(1-i)}{i}$$

.

Third, temporary changes to β can influence the eventual share of infected people, i, although regardless $i \geq \frac{1}{\mathcal{R}_0}$, the 'herd immunity' threshold. Fourth, the peak prevalence arises when $\bar{S} = \frac{1}{\mathcal{R}_0}$ and involves, at that point:

$$\bar{I} \equiv 1 - \frac{1 + \log(\mathcal{R}_0)}{\mathcal{R}_0}$$

This all implies that, for $\mathcal{R}_0 > 1$, (a) an equilibrium with $S = 1$ is locally unstable and (b) with temporary interventions that decrease β or increase γ, the absorbing states for i are characterised by $S(\infty) = [\underline{S}, \frac{1}{\mathcal{R}_0}]$ where $\underline{S}$ is defined by $\mathcal{R}_0 = -\frac{\log(\underline{S})}{1-\underline{S}}$; that is, either infections are kept at zero or they evolve to a point beyond the 'herd immunity' threshold.[2]

The standard SIR model is useful in that it relates the evolution of a pandemic according to $\mathcal{R}_0$ and how the underlying parameters associated with it may be impacted upon over the life of the pandemic. This can be useful for analysing the impact of non-pharmaceutical interventions that impact those underlying parameters. However, if, as is likely, those underlying parameters are not fixed but vary according in ways that relate to the underlying state variables, the standard SIR model will face challenges in being of predictive value.

---

[2]See (37) for details.

3 The Behavioural SIR Model

The fact that the standard SIR model lacked behavioural elements has not been lost on epidemiologists. In particular, it has been recognised that people might observe current prevalence (that is, $I(t)$) and modify their own behaviour so as to reduce infection risk. However, the mathematical epidemiologists have typically taken what economists would call a 'reduced-form' approach to this. For instance, they might posit a variable, $x \in [0, 1]$, that is a filter reducing the impact of β on new infections. That variable is then assumed to be a decreasing function of $I(t)$; e.g., $x(I(t))$.[3] A similar approach was used by (11).[4]

3.1 Literature Review

Work in economics to include behavioural elements in models of epidemics started in earnest with the study of the spread of AIDS. Following (36), (21) examined an SI model, whereby people can transition from susceptible to infected but cannot recover or become non-infectious, and examined the way in which increased prevalence would change the behaviour of a forward-looking rational agent. They showed that the incentives of infected agents – e.g., whether they altruistic or not – played an important role.[5] This line of research has continued with a mapping to empirical models by (27).

The pioneering treatment that first introduced forward-looking, rational economic agents into epidemiological models that could provide insights on COVID-19 was provided by (24). They examined SIR (in addition to SIS and SID models) to explore the different effects that prevention versus a treatment might have on the dynamics of epidemics. In doing this, they were able to clarify the externalities that may be present and the efficacy of various forms of interventions (including taxes and subsidies) to improve social welfare.[6] This approach inspired other analyses developing variants of their behavioural model including (38) who showed that agents will socially distance more when $\mathcal{R}_0$ is high (as they fear becoming infected) and (18) who showed that non-targeted lockdown policies may be worse than a

---

[3]See for example, (15) who also explore assumptions where $x(I(t), R(t))$ is decreasing in both variables, that they argue is a model of 'long-term awareness' in contrast to 'short-term awareness' where x is a function of $I(t)$ alone.

[4]This might be termed an 'old-timey' macro approach.

[5](31) also included behavioural elements in an SI model but his focus was on equilibrium outcomes in a broader matching game.

[6](9) uses an SIR model where agents can reduce their physical interactions and infected agents may be debilitated and so interact less. He focuses on myopic agents and analyses the impact of different matching functions on the resulting equilibrium outcomes.

decentralised behavioural outcome in terms of overall utility.[7][,][8,9]

A recent literature on COVID-19 has similarly built on these behavioural foundations with forward-looking rational agents including (14) (and by extension (32)) who provide a model of endogenous social distancing in a macroeconomic model; (Farboodi et al.) who examine how altruistic preferences (that capture the degree to which individuals choose to self-isolate if they know they are infected) impact on behaviour; (28) use a macro-model and highlight a 'fatalism' effect whereby, when prevalence is high, people do not socially distance as they are likely to become infected anyway; (4) who look at what optimal policies look like when the planner has a high degree of information regarding who is infected and who has recovered; (5) examine behavioural elements combined with frictions in spatial diffusion something also done by (3) using a structural model; (34) provides a finite time model but focuses on the case where agent value from economic activity depends on the activity of others introducing a complementarity and the possibility of multiple equilibria; (13) who look at the impact of limited information provided to agents and (6) do a variety of policy experiments.[10] The most careful analyses in this regard of the microeconomic foundations of the SIR model come from (42) and (37) who provide analyses that show the conditions under which endogenous social distancing will be too little, and potentially, too much compared with what might be socially optimal.

3.2 Model Setup

At the core of each of these models is a conception of a behavioural agent. An agent, n, chooses their level of activity, $x_n \in [0,1]$, which can be interpreted as their risk of interacting with another agent or preventative measures (such as wearing a mask). That activity gives them value in utility terms of $u_n(x_n)$ in each period where $u_n(.)$ is increasing, concave and independent across time periods. Agents have a common discount factor of $\delta < 1$. If an agent becomes infected, they incur an additional loss, L, in utility unless they die in which

---

[7]There is a literature that has examined behavioural SIS models with rational agents including (8) looks at how the provision of information impacts on agent's incentives to minimise risks of infection in an SIS model. When prevalence is low, agents may take more risks and make eradication impossible. (41) uses an SIS model where agents bear costs of reducing interactions. Agents are forward looking and understand the SIS dynamics. He examines the impact on a treatment that reduces transmission rates on social welfare and finds potential for a welfare-reducing rebound effect. (39) examines the appropriate mix of prevention and treatment while (25) looks at vaccine pricing where epidemiological effects are anticipated and influenced

[8]There is also a literature that focuses on incentives to be vaccinated using behavioural foundations. (19) uses an SIR model to consider an agent's choice of when to vaccine and finds that the market is efficient. (22) and (10) relaxes those conditions and finds inefficiency especially if individuals can independently acquire immunity

[9]See (35) and (23) for reviews.

[10]See (20) for further discussion of this literature.

case they can incur no utility thereafter. An infected agent has a probability, γ of becoming no longer infectious in each period they are infected. At that point, with probability ρ, they survive and become immune. Otherwise, they die. Either way they are part of R, the set of removed agents.

An agent's activity choices at t are determined by the condition, $\{S, I, R\}$, they are in at that time. If they are part of R and have not died, they are no longer infectious or at risk. Hence, they will set their activity, $x_{n,R} = 1$ and will earn an expected present discounted payoff of $\frac{u_n(1)}{1-\delta}$. In this, there is an implicit assumption that a recovery means a full recovery to the utility they would earn had the epidemic not emerged.

3.3 Infected Agent Activity

For an infected agent (a member of I), they are infectious and sick. Their instantaneous utility is $u_n(x_{n,I}) - L$ and their expected discounted payoff is:

$$V_{n,I}(t) = u_n(x_{n,I}(t)) - L + \delta(\gamma V_{n,R} + (1 - \gamma)V_{n,I}(t+1))$$

where here $V_{n,R} = \rho\frac{u_n(1)}{1-\delta}$. Note that, being self-interested, infected agents set $x_{n,I}(t) = 1$ in each period and, thus, their expected discounted payoff becomes:

$$V_{n,I} = \frac{u_n(1) - L + \delta(1 - \gamma)\rho\frac{u_n(1)}{1-\delta}}{1 - \delta\gamma}$$

This captures, in a stark way, a key externality that arises for infectious diseases when an infected person does not perceive a personal risk from social interactions. Of course, various factors could alter this stark result including that infected people may not be capable of or desire the same level of activity if they were healthy and that such activity may not be as valuable because others may avoid them if they knew they were infectious. For COVID-19, this was complicated by the fact that many of the infected were asymptomatic or pre-symptomatic and did not know they were infectious. In this situation, an agent may act as if they were still susceptible.

3.4 Susceptible Agent Activity

For both the infected and recovered, their choice of economic activity is not impacted upon by the state variables, $\{I(t), S(t)\}$. Thus, the key to the behavioural approach to epidemiology are the choices of the susceptible. Their instantaneous utility is $u_n(x_{(n,S)}(t))$ and their

expected discounted payoff is:

$$V_{n,S}(t) = u_n(x_{n,S}(t)) + \delta(p(x_{n,S}(t), I(t))V_{n,I}(t+1) + (1 - p(x_{n,S}(t), I(t)))V_{n,S}(t+1)$$

where $p(x_{n,S}(t), I(t))$ is probability that n becomes infected at time t (the consequences of which are felt at time $t + 1$). $p(.)$ is generally increasing in both of its arguments; i.e., a higher rate of infection in the population as well as a higher rate of activity by n raises the probability that n becomes infected. If $V_{n,I}(t+1) < V_{n,S}(t+1)$ this is not something that n wants and, thus, the increased risk of becoming infected will constrain the agent's choice of activity.

The structure of $p(x_{n,S}(t), I(t))$ depends upon how activity translates into an individual's risk of infection. The standard SIR model assumes that susceptible individuals face a probability, β, of becoming infected if they interact with an infected individual. What an 'interaction' precisely is, however, is potentially rich. For instance, if an agent visits a location where a number of other people are present, then β would be interpreted as the probability that at least one those people are infected. If a virus lingers or is spread on surfaces, then the probability that an agent becomes infected relates to the number of infected people who may be at a place in the past.[11]

Typically, the standard epidemiological models consider simpler environments. The simplest case assumes that an individual agent encounters one other member of the population at random in each period. In this situation, $x_{n,S}(t)$, is interpreted as the probability that n is matched with another person in period t who is infected with probability $I(t)$. Thus, the probability that n becomes infected is:

$$p(x_{n,S}(t), I(t)) = x_{n,S}(t)\beta I(t)$$

Of course, it is possible to imagine a slightly richer model whereby a susceptible understands that β might differ between alternative activities or that they can choose different populations with different $I(t)$ probabilities to interact with.[12] This structure presumes that $x_{m,I}(t) = 1$ for infected agents, $m \in I(t)$. If, for reasons of altruism or regulation, $x_{m,I}(t) < 1$, then the probability that n encounters an infected agent is $\frac{1}{I(t)} \int_0^{I(t)} x_{m,I}(t)dm$ so that $p(x_{n,S}(t), I(t)) = x_{n,S}(t)\beta \int_0^{I(t)} x_{m,I}(t)dm$.

---

[11](1) explore these issues by considering a variety of matching functions between susceptible agents and infecteds in an SIR model.

[12]See (16) for a review of these richer environments.

3.5 First-Order Effects

A susceptible individual, n, will choose $x_{n,S}(t)$ to maximise $V_{n,S}(t)$ holding the state variables and their future path as given. This gives rise to the marginal condition for the optimal choice $\hat{x}_{n,S}(t)$:

$$u'_n(\hat{x}_{n,S}(t)) = \beta I(t)\delta(V_{n,S}(t+1) - V_{n,I}) \tag{OPT}$$

This leads to a myriad of insights.

- (**Greater prevalence reduces susceptible activity**) Holding $V_{n,S}(t+1)$ fixed, as $I(t)$ increases $\hat{x}_{n,S}(t)$ falls. That is, the first-order effect of greater prevalence reduces an agent's activity as they forgo utility to reduce the risk of becoming infected.

- (**A more infectious virus reduces susceptible activity**) Holding $V_{n,S}(t+1)$ fixed, if the infectiousness of the virus (β) rises then $\hat{x}_{n,S}(t)$ falls. As will be noted below, this can reduce the rate of growth of the epidemic which stands in contrast to the clear prediction of the standard SIR model that a higher β will lead to faster epidemic spread and higher long term infections ((42)).

- (**Greater activity from infecteds reduces susceptible activity**) If, for some infecteds, $x_{m,I}(t) < 1$, it can be seen that $\hat{x}_{n,S}(t)$ may be higher. Thus, there is a strategic substitute between the activity choices of infected agents and susceptible agents (as noted by (29)).

- (**Activity is slower to return to normal as pandemic eases**) The future path of the epidemic is captured in the term, $V_{n,S}(t+1) - V_{n,I}$. Note, in particular, that if $I(t+1) > I(t)$, then $\hat{x}_{n,S}(t) \leq \hat{x}_{n,S}(t-1)$ while the opposite is true if $I(t+1) < I(t)$. This, as (37) shows, implies that a susceptible agent is going to engage in a smaller reduction in activity at the beginning of an epidemic than at the end for the same level of prevalence.[13] That is, for, $\underline{T} < \bar{T}$ where $I(\underline{T}) > I(\underline{T}-1)$, $I(\bar{T}) < I(\bar{T}-1)$ and $I(\underline{T}) = I(\bar{T})$, $\hat{x}_{n,S}(\underline{T}) > \hat{x}_{n,S}(\bar{T})$. Individuals will be more cautious at the end of a pandemic as the relative on-going value of being susceptible is higher.

- (**Complementarity between activity of susceptibles**) The interaction between a susceptible agent's decision on their own activity and the activity of other susceptible agents is potentially subtle. As will be described below, if susceptible agents reduce their activity at t, then this will reduce the share of the population infected at $t+1$. For

---

[13]The notion that at the onset of a pandemic, agents who expect a higher growth in infections tend to increase their activity and risk of infection is called the *fatalism* effect by (28).

an individual agent, therefore, a reduction in expected activity by other susceptibles increases $V_{n,S}(t+1)$ and hence, decreases their own choice of activity at time t as there is a greater value to not being infected. Thus, for susceptibles, their activity are strategic complements while at the same time constituting a negative externality on one another.

- (**Prospects for a vaccine or treatment have opposite effects on susceptible activity**) If a vaccine is expected at a future time, this increases $V_{n,S}(t+1)$ and hence, causes susceptible agents to reduce their activity; becoming more cautious so as to obtain the vaccine and not become infected. By contrast, if a treatment is expected at a future time, this, by either increasing ρ or decreasing L, causes $V_{n,I}$ to be higher and, thus, susceptibles to be less cautious of becoming infected and so increase their activity.

These insights are all implications of the first-order effects of changes in the environment on the behaviour of susceptible individuals. However, the full equilibrium effects can be harder to derive.

3.6 Equilibrium Analysis

To see this, we need to explore the evolution of the state variables under the behavioural assumptions that individual agents can influence their individual infection risk. Fortunately, the simple specification for $p(.)$ used above provides a natural way of aggregating into the expected path for the state variables, $\{I(t), S(t)\}$.

Let $X_S(t) \equiv \int_0^{S(t)} x_{n,S}(t) dn$. The expected number of new infecteds is equal to $\beta X_S(t) I(t)$ while each period $\gamma I(t)$ infecteds are removed. Thus,

$$I(t+1) - I(t) = (\beta X_S(t) - \gamma) I(t)$$

By construction, this also means the total number of susceptibles declines by:

$$S(t+1) - S(t) = -\beta X_S(t) I(t)$$

Note that if $x_{n,S}(t) = 1$ for all $n \in S(t)$, then $X_S(t) = S(t)$ and the above two equations become the same as the standard SIR model.

It can be seen here that the time path of $\{X_S(t), ...\}$ determines the net presented expected value of continuing to be susceptible and, thus, the incentives to undertake activity at time t. Thus, the equilibrium outcome would require solving for a multi-dimensional

fixed even with commonly used simplifying assumptions such as all agents being symmetric in preferences. Moreover, the set of susceptibles is being reduced in size over time at a rate that is endogenous to the activity choices of susceptibles themselves. This means that there is unlikely to be stationary equilibrium outcome that we usually look for in order to conduct comparative statics. For this reason, most studies of behavioural SIR models have used simulations to demonstrate potential outcomes rather than analytical solutions. For this reason, I propose here, instead, taking a shortcut that will permit an analytical solution albeit at the expense of not (usually) satisfying our usual equilibrium requirements.

4 An Analytical Shortcut

The analytical shortcut I propose here is to establish conditions under which $I(t+1) = I(t)$ for an interval of time. The condition is a simple one: $S(t+1) = S(t) = S$ for all t. It is immediately is apparent that this condition violates the laws of motion of the SIR model whenever $\gamma > 0$. As an accounting measure, it simply cannot be the case that some infected individuals are recovered (or strictly speaking) removed and $S(t)$ is not falling over time. Of course, this state of affairs is possible for the SIS model which is perhaps why much of the initial work integrating behavioural assumptions into epidemiology examined that environment. However, because we want the incentives of agents to reflect the possibility that they can be removed following an infection, I cannot simply follow the SIS model here. Instead, I have just been inspired by it.

4.1 Equilibrium Solution

The focus is on the equations governing the relationship between $X_S(t)$ and $I(t)$. The first equation is behavioural.

$$\hat{X}_{n,S}(I(t)) = \int_0^S \hat{x}_{n,S}(I(t))dn \tag{BEH}$$

This equation is how the aggregate activity of susceptible agents (now fixed at size S) is a function of $I(t)$ when individual agents are optimising. Note that $\hat{X}_{n,S}(I(t))$ is a non-decreasing function of $I(t)$ as discussed earlier.

The second equation comes from the SIR laws of motion.

$$I(t+1) = I(t) + (\beta X_S(t) - \gamma)I(t)$$

The number of infected agents is an increasing function of the aggregate activity, $X_S(t)$, of

those agents.

Essentially, these two equations describe a dynamic aggregate game involving choices of susceptible agents but under an assumption that the set of those agents is now fixed. The goal will be to characterise stationary Markov perfect equilibria of this game using a dynamic programming approach.

From the law of motion, we have:

$$X_S(t) = \frac{\frac{I(t+1)-I(t)}{I(t)} + \gamma}{\beta} \tag{EPI}$$

Setting this equal to $\hat{X}_S(I(t))$ equilibria in which $I(t+1) = I(t)$ for all t can be explored. When this condition is satisfied then $\hat{x}_{n,S}(t+1) = \hat{x}_{n,S}(t)$ for all t which carries over to, $\hat{X}_{n,S}(I(t))$. Importantly, this means that:

$$I(t+1) - I(t) = 0 = (\beta\hat{X}_S(I(t)) - \gamma)I(t) \implies \hat{X}_S(I^*) = \frac{1}{\mathcal{R}_0} \tag{EQM}$$

Importantly, this implies that the equilibrium effective reproduction number,

$$\hat{\mathcal{R}} = \hat{X}_S(I^*)\mathcal{R}_0 = 1$$

Thus, prevalence will neither rise nor decline in equilibrium and this pins down that equilibrium steady state of infected agents.[14]

4.2 Graphical Analysis

The analytical shortcut has the advantage that it permits a (familiar to economists) graphical analysis. Figure 2 shows the EPI and BEH lines in (X_S, I) space. BEH shows how the aggregate choice of activity level is determined by the prevailing share of infected agents and, as shown, earlier is typically downward sloping as agent's reduce activity more when there is a greater chance of encountering an infected agent. EPI shows how the number of infected agents relates to the aggregate choice of activity level by susceptibles. It is upward sloping as a higher X_S directly increases $I(t+1)$ in a linear fashion in the SIR model. Where the two curves intersect is the equilibrium outcome under the assumption that S is held fixed.

---

[14]It can readily be seen that this equilibrium exists if $\mathcal{R}_0 > 1$. When $I(t) = 0$, all agents set $\hat{x}_{n,S} = 1$ so that $\hat{X}_S(0) = 1$. At this point $X_S(0) = \frac{1}{\mathcal{R}_0}$ which is less than 1. On the other hand, if $I(t) = 1$, $X_S(1) = \frac{\frac{1-I(t-1)}{I(t-1)}+\gamma}{\beta} > 0$ while $\hat{X}_S(I(t)) \to 0$. As all of the relevant functions are continuous, there is a fixed point where $I(t) = I^*$.
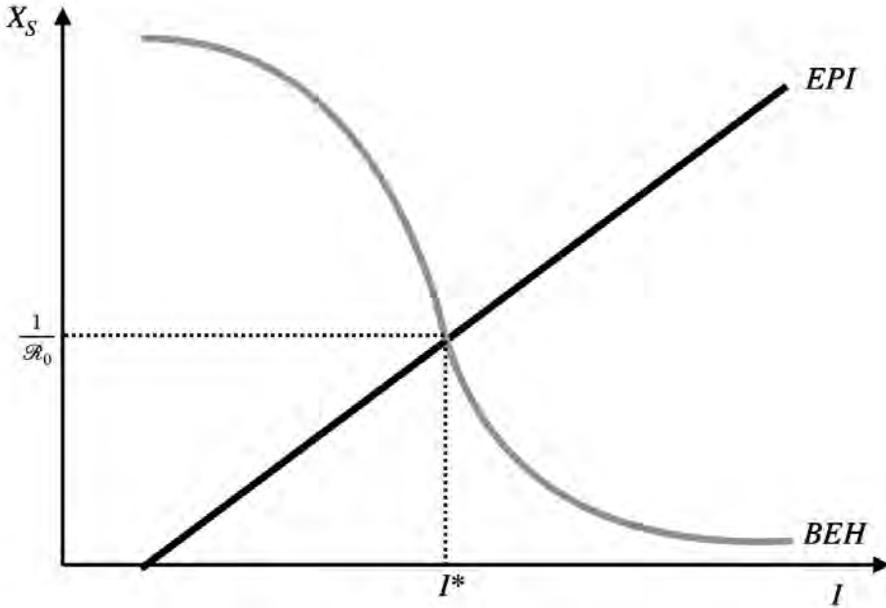
Figure 2: Equilibrium

This graphical approach also shows why the equilibrium is stable. Suppose that $I(t) < I^*$. Then $\hat{X}_S(t) > \frac{1}{S\mathcal{R}_0}$ and $I(t+1) > I(t)$. By contrast, if $I(t) > I^*$, $\hat{X}_S(t) < \frac{1}{S\mathcal{R}_0}$ and $I(t+1) < I(t)$. These processes only stop as $I(t) = I^*$.

This approach allows for intuitive comparative static analysis. Figure 3 shows what happens if there is an increase in baseline infectiousness, β. Firstly, as is well known in epidemiology, an increase in β means that more activity translates into higher infections at a faster rate; shifting the EPI line to the right. Second, from the analysis of behavioural responses, an increase in β causes susceptible agents to choose to be more careful and reduce their activity. Thus, the BEH curve shifts to the left. An increase in β has a negative equilibrium impact on aggregate activity from susceptibles but an ambiguous impact on the equilibrium number of infected agents. Figure 3 is drawn to show the case where an increase in β leads to a lower rate of infection in contrast to the standard epidemiological prediction. However, if the behavioural response is weaker, then the opposite comparative static is possible.

Interestingly, there are some unambiguous monotone comparative static results that can be derived. For instance, a change that impacts on $V_{n,I}$ only without changing anything
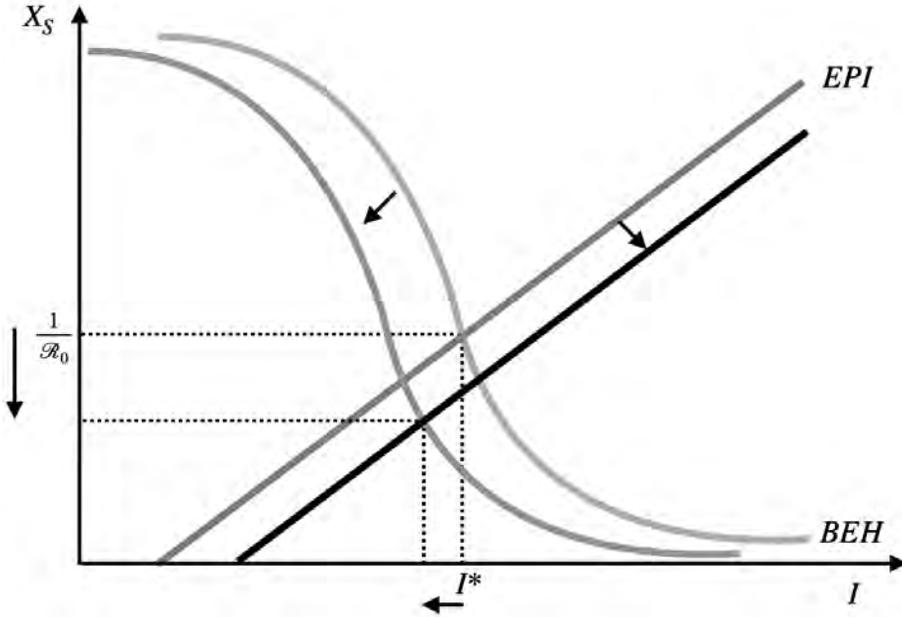
Figure 3: Increase in Infectiousness

else – e.g., a treatment that increases ρ or a measure that makes being infected less costly (i.e., reduces L) only impacts on BEH; shifting it up and to the right. Thus, the availability of a treatment causes both a higher equilibrium activity and a higher equilibrium level of infection.

4.3   Impact of Testing/Isolation

One policy that has received attention in COVID-19 is the increased use of testing (and contact tracing) to identify infected individuals earlier and isolate them to prevent them spreading the virus. This approach was adopted as a standard practice by many countries and appeared to be successful in reducing the scale of the COVID-19 pandemic. (See (20), Chapter 7 for more details). However, some recent work in economics has raised the possibility of unintended behavioural consequences from increased testing including testing giving infected people confidence to engage in activity because they can't get more infected ((40) and (12)), a reluctance to be tested for fear of being quarantined ((14) to the potential for a rebound effect that increases activity choices ((2)). The model presented here permits the

examination of these consequences.

The focus here is on the situation where tests immediately trigger isolation (say, because they are done by a public authority with enforcement power or have subsidies that induce isolation).[15] The first impact of testing (along with isolation) is one that is intended: it reduces the probability that a susceptible encounters an infected agent. This impacts on both the EPI and EPI equations. The epidemiological response is to shift the EPI curve to left as it directly reduces the probability that a susceptible agent will encounter and infected agent. However, this also leads to a shift outwards of the BEH curve. The reduction in the probability of encountering an infected agent, increases the incentives of susceptible agents to engage in activity for given level of prevalence. This is the effect identified by (2).[16] Thus, examining this impact alone, we would find a similar ambiguous comparative static as that for infectious but in the opposite direction to the movements depicted in Figure 3.

There is, however, a second impact of testing – and specifically, isolation – that has not been examined in the literature. Testing followed by isolation reduces the utility from becoming infected as an agent would not expect to be able to freely choose their activity level in that event. Formally, their utility becomes $u_n(0)$ rather than $u_n(1)$ in that case. While (14) focused on how agents may avoid tests altogether, if agents are tested, those tests themselves will cause the impact identified here. Specifically, with a reduction in the utility of becoming infected, agents will become more cautious. This will shift the BEH curve to the left countering the impact of increased testing on the likelihood of encountering an infected person.

Putting the two impact mechanisms together, we can explore further whether the ambiguity may be removed if BEH, on net, shifted to the left. To explore this, let α be the probability that an infected agent is isolated as a result of testing regime. Given this, we have:

$$p(x_{n,S}(t), I(t)) = x_{n,S}(t)\beta(1-\alpha)I(t)$$

$$V_{n,I} = \frac{(1-\alpha)u_n(1) + \alpha u_n(0) - L + \delta(1-\gamma)\rho\frac{u_n(1)}{1-\delta}}{1-\delta\gamma}$$

Note that:

---

[15]The situation where people may keep test outcomes private or not obtain tests is captured by the α below but the analysis does not inform on the issue created by that possibility as to whether it is desirable to have a testing regime relative to leaving individuals uninformed as to their infectiousness. The approach here could be used to analyse such cases but that is left to future work.

[16]They use a network rather than SIR model and so implicitly adopt the analytical shortcut proposed here.

$$V_{n,S} - V_{n,I} = \frac{(1-\delta\gamma)u(x_{n,S}) - (1-\delta)((1-\alpha)u_n(1) + \alpha u_n(0) - L) - \delta(1-\gamma)\rho u_n(1)}{(1-\delta\gamma)(1-\delta(1-x_{n,S}\beta(1-\alpha)I^*))}$$

(OPT) becomes:

$$u_n'(\hat{x}_{n,S}) - \beta(1-\alpha)I^*\delta\left(V_{n,S} - V_{n,I}\right) = 0$$

Taking the derivative of the LHS of (OPT) with respect to α we have:

$$\beta I^*\delta\left(V_{n,S} - V_{n,I}\right) - \beta(1-\alpha)I^*\delta\frac{\partial(V_{n,S} - V_{n,I})}{\partial\alpha}$$

The first term is the marginal benefit to more *risk* as a result of testing while the second term is the marginal benefit to more *caution*. Note that:

$$\frac{\partial(V_{n,S} - V_{n,I})}{\partial\alpha} = \frac{\delta x_{n,s}\beta I^*(1-\delta\gamma)(V_{n,S} - V_{n,I}) + (1-\delta)(u_n(1) - u_n(0))}{(1-\delta\gamma)(1-\delta(1-x_{n,S}\beta(1-\alpha)I^*))}$$

which is positive for $V_{n,S} \geq V_{n,I}$. Putting the two effects together, the impact of α on the marginal return to activity is positive if:

$$V_{n,S} - V_{n,I} \geq (1-\alpha)\frac{\delta x_{n,s}\beta I^*(1-\delta\gamma)(V_{n,S} - V_{n,I}) + (1-\delta)(u_n(1) - u_n(0))}{(1-\delta\gamma)(1-\delta(1-x_{n,S}\beta(1-\alpha)I^*))}$$

Notice that as $\alpha \to 1$, this always holds. By contrast for $\alpha \to 0$, this becomes:

$$(1-\delta\gamma)(V_{n,S} - V_{n,I})_{\alpha=0} \geq u_n(1) - u_n(0)$$

which may not hold. Thus, while it is possible, for low α, that there may be an unambiguous comparative static that testing will reduce equilibrium infections, for high α, ambiguity remains. In this case, an increase in α (i.e., the effectiveness of testing and isolating) leads to a shift upwards in the BEH curve. In this model, therefore, as testing increases the relative safety of interactions this causes activity to rise by more than the effect driven by the decrease in the utility of the infected. Hence, the ambiguity remains for this comparative static.

4.4   Mandated masks

Encouraging the use of masks has been a strategy increasingly deployed and even mandated for dealing with COVID-19. In some medical circles there is debate regarding whether mandated masks would encourage less social distancing and potentially have a immiseration

effect on infection rates ((33)). As was the case with testing, the analytical approach here can be used to provide insight on that potential.

Suppose that, if all but recovered agents wear masks, the probability that the virus infects a susceptible person in an interaction with an infected one is $1 - \alpha$; that is, a higher α means that a susceptible has more protection. Mask wearing is costly to individuals and, thus, it is assumed that all susceptible and infected agents, n, bear a cost, c_n, for each period they wear a mask. In this situation, the only difference between the impact of masks is this cost as well as the fact that infected people are not restricted in their activity and thus earn $u_n(1)$ while infected.

Thus, as was the case with testing, more effective masks (i.e., a higher α) leads to an increase in the returns to risky activity as well as a cautionary effect. The overall effect of masks is the impact of both. Note that, the impact of more effective masks on activity is positive if:

$$V_{n,S} - V_{n,I} \geq (1 - \alpha) \frac{\delta x_{n,s} \beta I^* (1 - \delta\gamma)(V_{n,S} - V_{n,I})}{(1 - \delta\gamma)(1 - \delta(1 - x_{n,S}\beta(1-\alpha)I^*))} \implies 1 \geq \delta$$

where the last implication assumes that $V_{n,S} \geq V_{n,I}$. Thus, masks will always move the BEH curve to the right meaning that, given that they move the EPI curve to the left, there is no unambiguous comparative static result with respect to masks. Compared with testing and isolation, the returns to being infected are higher with mask wearing and so this reduces one driver of caution.

5 The $\hat{\mathcal{R}} = 1$ Prediction

The analytical shortcut, whereby an equilibrium is analysed based on an assumption that S is fixed, gives rise to a prediction that $\hat{\mathcal{R}} = 1$. As noted earlier, this comports with the trends associated with the first few months of COVID-19 in a variety of countries that failed to suppress the pandemic. The question is: given that it is obtained using an analytical short-cut, how seriously should we take this prediction?

The potential error that arises from the short-cut can be seen by examining Figure 2. Note that rather than being constant, the share of susceptibles, S, will fall overtime. Indeed, if the level of infected persisted at I^*, S would fall by γI^* in each period. This means that the realised aggregate level of activity by susceptibles, X_S, would be expected to fall. This would not change the EPI curve as this change would be a movement along that curve. However, it would have an impact on the BEH curve. This is because the maximum value of X_S that can be generated by that relationship is S. Thus, a reduction in S may cause the

feasibility constraint to bind. Without modelling how agents take into account the change in S in their own decisions – through expectations of a lower I in the immediate future – this curve, as derived, is only an approximation of what might occur.

That said, there is one special case for which a full equilibrium of the environment (sans the analytical shortcut) coincides with the limited equilibrium outcome examined thusfar. (42) assumes that all agents are identical and that their activity choice $x \in 0, 1$. This gives rise to BEH as depicted in Figure 4. When $I(t) < I^*$, all agents choose $\hat{x} = 1$ and when $I(t) > I^*$, they choose $\hat{x} = 0$. He shows that I^* is independent of the share of susceptibles. The equilibrium arises when $I(t) = I^*$ and agents pursue a mixed strategy between $\{0, 1\}$. The total choosing $\hat{x} = 0$ averages $\frac{1}{S\mathcal{R}_0}$. Thus, $\hat{\mathcal{R}}_t = S\mathcal{R}_0 = 1$.

In Figure 4, it can be seen that as S falls, this reduces the maximum of BEH but otherwise leaves the line, and hence, equilibrium outcome in terms of infections and reproduction rate unchanged. This, of course, does not continue indefinitely. As (42) shows eventually $S\mathcal{R}_0 < 1$ in which case, the equilibrium moves down the EPI line until the pandemic eventually ends. Thus, compared with the standard SIR model, in this model, the pandemic emerges and hits a ceiling of infecteds at $I(t) = I^*$ and stays that way until $S < \frac{1}{\mathcal{R}_0}$. Thus, the curve is not so much flattened as 'pancaked' at I^*.

This at least provides comfort that the $\hat{\mathcal{R}} = 1$ prediction is the outcome of a possible full equilibrium model. It also gives insight as to why the simulations of (11) and (29) were able to generate outcomes whereby a $\mathcal{R}_t = 1$ outcome was observed for considerable periods of time when calibrated with parameters based on COVID-19. Put simply, during the early months of the pandemic S was so large compared to I that it would not be expected to change very much meaning that $\mathcal{R}_t$ appeared to be relatively constant over time and close to 1 especially relative to SIR simulations that did not include a behavioural element.

6   Conclusions

Standard epidemiological models of pandemics often do not consider how susceptible, infected and recovered people will change their behaviour over the life cycle of the pandemic. Economists have made progress in building behavioural elements into these models but the non-stationarity that is a key part of viral epidemics such as COVID-19 has prevented an easy characterisation of equilibrium paths of pandemics and the potential impact of interventions.

This paper argues that some analytical progress can be made on behavioural SIR models by taking inspiration for epidemiological models that do have stationary characteristics. In so doing, an equilibrium outcome is derived that allows intuitive comparative static outcomes on key variables such as infection rates and aggregate activity choices while at the same time
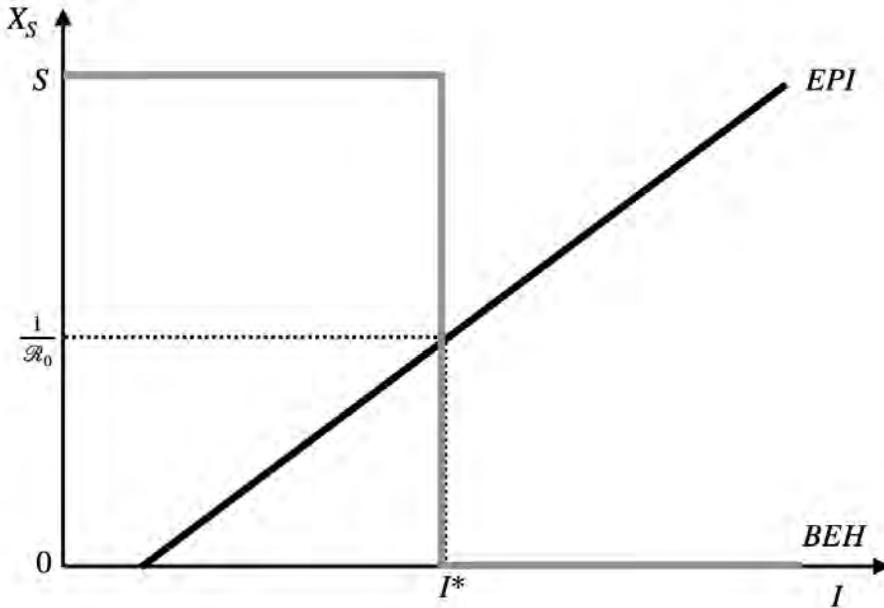
Figure 4: Binary Choice and Symmetric Agent Equilibrium

generating a prediction that during much of a pandemic, without intervention, the effective reproduction number, $\mathcal{R}_t$ will tend towards 1. At this point, the infection rate is neither rising nor falling. This is consistent with the outcomes in many regions with respect to COVID-19. Nonetheless, the model here falls short of the usual requirements for a full equilibrium outcome. It does, however, have the benefit of being upfront about this limitation and what precisely we are getting in return in terms of tractibility and potential insight.

In doing this, this paper makes the case for treating $\hat{\mathcal{R}}_t = 1$ to be an expected outcome that can be used to evaluate, for both policy analysis and empirical predictions regarding pandemics. While being upfront regarding its 'cargo cultish' logic (i.e., based on observations of effective reproduction numbers hovering around 1 rather than fully from primitive assumptions), as a shortcut it can provide some insight that might inform debates. For instance, (7) has argued that $\mathcal{R}$ being just below 1 should be a constraint that is met by policies that impose lockdowns and other behaviour during pandemics. However, if the expectation is that, absent interventions, that goal would be mostly achieved anyway, that target is arguably of limited use compared, say, to a target of achieving $I = 0$ prior to what would otherwise be the natural course of pandemics. Moreover, with regard to lockdowns,

the expectation that non-targeted lockdown activities may be adjusted to generate $\hat{\mathcal{R}}_t = 1$ requires us to not simply look at the epidemiological consequences of interventions (as (1) do) but also to whether the non-targeted activities are such that they would be unable to adjust so that $\hat{\mathcal{R}}_t = 1$ was feasible. In other words, the criteria for lockdowns is not simply about spread but about the scope for behavioural adjustment.

Nonetheless, this analysis here remains purely normative. While it is tempting to conclude that if testing or mask use led to more infections this would be welfare-reducing, we must also remember the purpose of those interventions is precisely to allow activity to be safer and hence, allow for more of these at the margin. Thus, even though it is possible to draw some possible welfare conclusions from the fact that the BEH curve does not take into account external effects and so likely lies above a suitably derived social curve, the reality is more nuanced and requires an embrace of dynamic impacts. In particular, as (37) has shown the cumulative nature of pandemic impacts suggest that a focus on instantaneous external effects is unlikely to provide the correct insight into optimal policy-making.

References

[1] Acemoglu, D., Chernozhukov, V., Werning, I., and Whinston, M. D. (2020a). A multi-risk sir model with optimally targeted lockdown. Technical report, National Bureau of Economic Research.

[2] Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. (2020b). Testing, voluntary social distancing and the spread of an infection. Technical report, National Bureau of Economic Research.

[3] Aguirregabiria, V., Gu, J., Luo, Y., and Mira, P. (2020). A dynamic structural model of virus diffusion and network production: A first report.

[4] Bethune, Z. and Korinek, A. (2020). Covid-19 infection externalities: Pursuing herd immunity or containment?

[5] Bisin, A. and Moro, A. (2020). Learning epidemiology by doing: The empirical implications of a spatial *sir* model with behavioral responses. *Available at SSRN 3625361*.

[6] Brotherhood, L., Kircher, P., Santos, C., and Tertilt, M. (2020). An economic model of the covid-19 epidemic: The importance of testing and age-specific policies.

[7] Budish, E. B. (2020). R < 1 as an economic constraint: Can we 'expand the frontier' in the fight against covid-19? *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2020-31).

[8] Chen, F. H. (2009). Modeling the effect of information quality on risk behavior change and the transmission of infectious diseases. *Mathematical biosciences*, 217(2):125–133.

[9] Chen, F. H. (2012). A mathematical analysis of public avoidance behavior during epidemics using game theory. *Journal of theoretical biology*, 302:18–28.

[10] Chen, F. H. and Toxvaerd, F. (2014). The economics of vaccination. *Journal of theoretical biology*, 363:105–117.

[11] Cochrane, J. C. (2020). An sir model with behavior. `https://johnhcochrane.blogspot.com/2020/05/an-sir-model-with-behavior.html`.

[12] Deb, R., Pai, M., Vohra, A., and Vohra, R. (2020). Testing alone is insufficient. *Available at SSRN 3593974*.

[13] Di Guilmi, C., Galanis, G., and Baskozos, G. (2020). A behavioural sir model and its implications for physical distancing.

[14] Eichenbaum, M. S., Rebelo, S., and Trabandt, M. (2020). The macroeconomics of epidemics. Technical report, National Bureau of Economic Research.

[15] Eksin, C., Paarporn, K., and Weitz, J. S. (2019). Systematic biases in disease forecasting–the role of behavior change. *Epidemics*, 27:96–105.

[16] Ellison, G. (2020). Implications of heterogeneous sir models for analyses of covid-19. Technical report, National Bureau of Economic Research.

[Farboodi et al.] Farboodi, M., Jarosch, G., and Shimer, R. Internal and external effects of social distancing in a pandemic. *Covid Economics 9*.

[18] Fenichel, E. P. (2013). Economic considerations for social distancing and behavioral based policies during an epidemic. *Journal of health economics*, 32(2):440–451.

[19] Francis, P. J. (1997). Dynamic epidemiology and the market for vaccinations. *Journal of Public Economics*, 63(3):383–406.

[20] Gans, J. S. (2020). *Economics in the Age of COVID-19*. MIT Press.

[21] Geoffard, P.-Y. and Philipson, T. (1996). Rational epidemics and their public control. *International economic review*, pages 603–624.

[22] Gersovitz, M. (2003). 25 births, recoveries, vaccinations, and externalities. *Economics for an imperfect world: Essays in honor of Joseph E. Stiglitz*, page 469.

[23] Gersovitz, M. (2011). The economics of infection control. *Annu. Rev. Resour. Econ.*, 3(1):277–296.

[24] Gersovitz, M. and Hammer, J. S. (2004). The economical control of infectious diseases. *The Economic Journal*, 114(492):1–27.

[25] Goodkin-Gold, M., Kremer, M., Snyder, C., and Williams, H. (2020). Vaccination policy in the short and long run. mimeo., Harvard.

[26] Goolsbee, A. and Syverson, C. (2020). Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. Technical report, National Bureau of Economic Research.

[27] Greenwood, J., Kircher, P., Santos, C., and Tertilt, M. (2019). An equilibrium model of the african hiv/aids epidemic. *Econometrica*, 87(4):1081–1113.

[28] Jones, C. J., Philippon, T., and Venkateswaran, V. (2020). Optimal mitigation policies in a pandemic: Social distancing and working from home. Technical report, National Bureau of Economic Research.

[29] Keppo, J., Quercioli, E., Kudlyak, M., Wilson, A., and Smith, L. (2020). The behavioral sir model, with applications to the swine flu and covid-19 pandemics. mimeo., Wisconsin.

[30] Kermack, W. O. and McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721.

[31] Kremer, M. (1996). Integrating behavioral choice into epidemiological models of aids. *The Quarterly Journal of Economics*, 111(2):549–573.

[32] Krueger, D., Uhlig, H., and Xie, T. (2020). Macroeconomic dynamics and reallocation in an epidemic. *Covid Economics 5*.

[33] Mantzari, E., Rubin, G. J., and Marteau, T. M. (2020). Is risk compensation threatening public health in the covid-19 pandemic? *BMJ*, 370.

[34] McAdams, D. (2020). Nash sir: An economic-epidemiological model of strategic behavior during a viral epidemic. *arXiv preprint arXiv:2006.10109*.

[35] Philipson, T. J. (2000). Economic epidemiology and infectious diseases. *Handbook of Health Economics*, 1:1761–1799.

[36] Philipson, T. J. and Posner, R. A. (1993). *Private Choices and Public Health*. Cambridge, Mass.: Harvard University Press.

[37] Rachel, L. (2020). An analytical model of covid-19 lockdowns. mimeo, London School of Economics.

[38] Reluga, T. C. (2010). Game theory of social distancing in response to an epidemic. *PLoS Comput Biol*, 6(5):e1000793.

[39] Rowthorn, B. R. and Toxvaerd, F. (2020). The optimal control of infectious diseases via prevention and treatment. Technical report.

[40] Taylor, C. (2020). Information and risky behavior: Model and policy implications for covid-19. mimeo, Stanford.

[41] Toxvaerd, F. (2019). Rational disinhibition and externalities in prevention. *International Economic Review*, 60(4):1737–1755.

[42] Toxvaerd, F. (2020). Equilibrium social distancing. *Covid Economics 15*.

Health vs economy: Politically optimal pandemic policy

Desiree A. Desierto[1] and Mark Koyama[2]

Pandemics have heterogeneous effects on the health and economic outcomes of members of the population. To stay in power, politician-policymakers have to consider the health vulnerability–economic vulnerability (HV–EV) profiles of their coalition. We show that the politically optimal pandemic policy (POPP) reveals the HV–EV profile of the smallest, rather than the largest, group in the coalition. The logic of political survival dictates that the preferences of the least loyal, most pivotal, members of the coalition determine policy.

1   Introduction

Research on Covid-19 has focused on the macroeconomic impacts of the pandemic and on deriving optimal policies. There has been comparatively little research from a political economy perspective.[1] We highlight the simple trade-off at the heart of policymaking during a pandemic: the trade-off between health and the economy. We propose a novel political economy explanation for the different policies taken by different governments in response to the pandemic.

Policies such as early, prolonged, or severe lockdowns may limit disease contagion but impose economic costs. To the extent that politician-policymakers are accountable to citizens, equilibrium pandemic policy depends on citizens' preferences, which are likely to be heterogeneous. We demonstrate that the politically optimal pandemic policy will be most responsive to the preferences of the least loyal, or most pivotal, members of the ruling coalition.

There were striking differences across countries in their initial policy responses to Covid-19. Some countries imposed severe lockdowns relatively early on. Others waited until confirmed deaths from Covid-19 were in the hundreds or higher. To illustrate, Figure 1 plots the stringency index calculated by OxCGRT against total Covid-19 deaths for Austria compared to Sweden and Italy compared to UK. Both Austria and Italy were more aggressive in implementing lockdowns, closing schools and non-essential businesses. Similar patterns are evident for other countries.

Figure 2 highlights this heterogeneity in lockdown policy still more starkly. It plots the stringency index for nine countries at the point at which they reached 100 total Covid-19 deaths.[2] There is considerable variation between the aggressive pandemic response of Denmark and Israel compared to the United States and Japan. What explains these sharp differences in policy responses?

We apply the selectorate theory of Bueno de Mesquita, Morrow, Siverson, and Smith (2003) to provide a rigorous yet simple explanation of this variation. In this framework, the political survival of the incumbent politician-policymaker depends on the loyalty of a coalition of supporters drawn from the 'selectorate'. We then specify how selectorate members generate income and, in particular, posit that pandemic policy affects income in two ways – by suppressing potential productivity, but also increasing the probability that the productivity is realized by mitigating the severity of the pandemic. The net effect of the policy on selectorate members' income differs, as some are more sensitive to economic conditions while others are more susceptible to the disease. This heterogeneity then affects the profile of coalition members, which ultimately determines the incumbent's policy response.

To motivate our analysis we highlight the following stylized facts. First, the heterogeneity in risk

---

[1]One exception is Pulejo and Querubın (2020) who find that reelection concerns are associated with a less stringent response to Covid-19.

[2]In the online appendix we report the stringency index for other countries and for different death thresholds.

CEPR PRESS

Figure 1: Contrasting Pandemic Policies
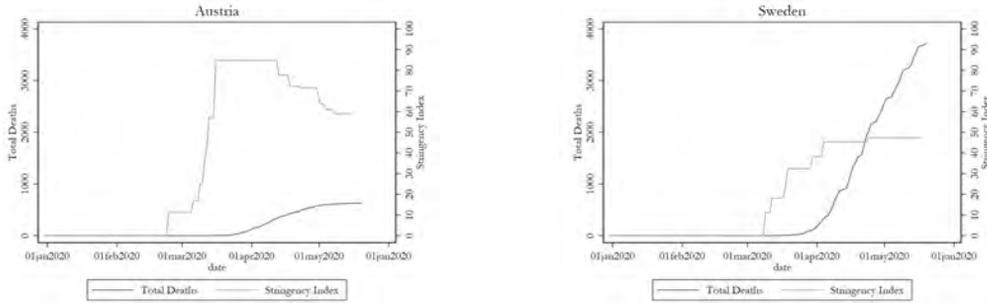
Contrasting pandemic policies in Austria and Sweden. Both scaled to 4000 total Covid-19 deaths. Data: (Max Roser and Hasell 2020).
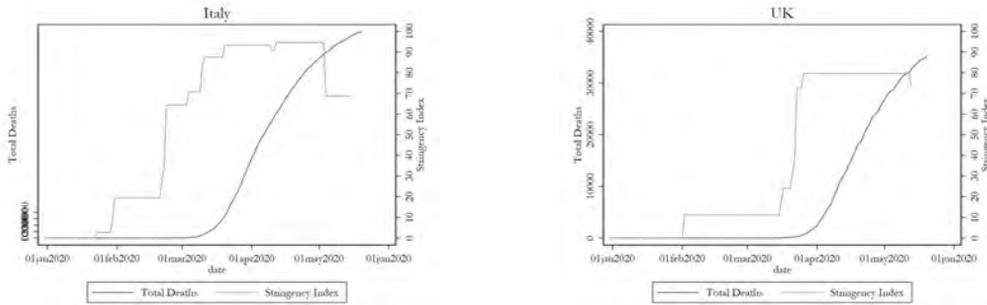
Contrasting pandemic policies in Italy and UK Both scaled to 40000 total Covid-19 deaths. Data: (Max Roser and Hasell 2020).

exposure is stark. Individuals under the age of 30 without preexisting conditions, are extremely unlikely to develop a severe infection. However, the case fatality rate for individuals aged over 70 or with comorbidities is high. Thus we can distinguish between those with a high health vulnerability and those with a low health vulnerability to the pandemic.

Second, the economic risk engendered by Covid-19 is also highly dichotomous. Many small businesses have been forced to close and face the threat of collapse whereas many low-skilled workers, particularly in the service sector, have been made unemployed.[3] Dingel and Neiman (2020) find that approximately 37% of jobs can be done from home. However, these tend to be high skilled jobs in corporate management, journalist, academia, law, and the technology sector. In contrast, workers in the service sector or in construction or manufacturing require the economy to remain open in order to keep their livelihoods. This

---

[3]Similiarly the Daily Telegraph notes "The great class divide is now between those who can work remotely (Heath 2020).

Figure 2: A snapshot of the stringency index at the point when 100 deaths were reached. Data: (Max Roser and Hasell 2020).
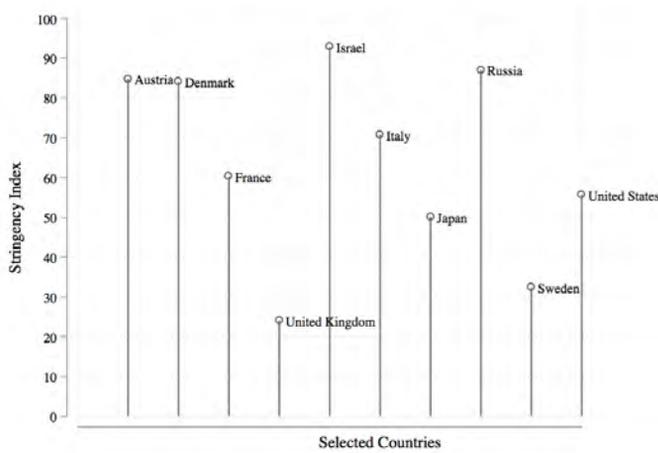
provides a sharp distinction between individuals who have high economic exposure to the pandemic and those who have a comparatively low economic exposure to the pandemic.

For tractability, the framework we propose is static. Thus it abstracts away from modeling the dynamics of the disease—which is the focus of models that combine a macroeconomic model of the economy with a Susceptible Infectious Recovered (SIR) model of disease spread. Nevertheless, our model provides a framework for conducting simple comparative statics with respect to new information about disease severity or the cost of pandemic policies. Future research can also incorporate other variables such as cultural attitudes (individualism versus collectivism) that might affect compliance to pandemic policies.

2   The Model

Consider a population, the 'selectorate', of size $S = \sum_i S_i$, $i \in \{K\}$, that is composed of K groups, each of type i. The government needs to maintain the support of a coalition of selectorate members of size $W \subseteq S$. Let w_i denote the number of i-type members in the coalition, such that $\sum_i w_i = W$, and let i also index the size of these groups such that $w_1 > w_2 > ...w_K$.

To keep the loyalty of the coalition, the government provides a combination of transfers and public policy – in this context, pandemic policy. The policy affects all of the selectorate, while the transfers are given only to coalition members. The disposable income of a selectorate member of type i during a

pandemic is

$$y_i = (1 - \tau)[\rho_i Y_{\rho i} + (1 - \rho_i)\bar{Y}] + \frac{t_i}{w_i}\mathbb{1}_W, \tag{1}$$

where τ is the tax rate, $Y_{\rho i}$ is type i's potential income, ρ_i the probability that she realizes the potential income, $\bar{Y}$ the minimum level of income received if she is unable to realize potential income, $Y_{\rho i} > \bar{Y}$, t_i is total transfers to coalition members of type i, which is distributed equally among w_i, and $\mathbb{1}_W$ an indicator variable that takes on one if the selectorate member is in the coalition.

Denote pandemic policy as $p > 0$, with larger values associated with more aggressive policy, e.g. earlier lockdowns. Pandemic policy has heterogeneous effects across selectorate types. In particular, ρ_i is a function $\rho_i \equiv \rho(p, \eta_i \frac{\sum_i (S_i \rho_i)}{S})$, $\eta_i \geq 0$, while $Y_{\rho i}$ is a function $Y_{\rho i} \equiv Y_\rho(p, \epsilon_i \sum_i (S_i y_i))$, $\epsilon_i \geq 0$. Letting partial derivatives with respect to p be $\frac{\partial \rho_i}{\partial p} > 0$ and $\frac{\partial Y_{\rho i}}{\partial p} < 0$, then ρ_i can capture the health of i (enabling her to realize potential output $Y_{\rho i}$), which is directly (positively) influenced by a more aggressive pandemic policy, and indirectly through some contagion effect, which we simply capture by average health $\frac{\sum_i (S_i \rho_i)}{S}$ of the selectorate, scaled by a 'health vulnerability' parameter η_i. Meanwhile, potential output $Y_{\rho i}$ is negatively impacted by more aggressive pandemic policy, directly and indirectly through total income $\sum_i (S_i y_i)$ of the selectorate, scaled by an 'economic vulnerability' parameter ϵ_i.

Optimal pandemic policy and transfer amounts are thus determined from the following game that is played at each time $t = 1, 2, ... \infty$.

1. The incumbent government I forms a coalition of size W composed of K groups, each of size w_i. The size of each type in the coalition is fixed at $w_1 > w_2 > ... w_K$. A political challenger C nominates her own coalition, also of size W, with $w_i \; \forall \; i \in \{K\}$ the same size as in I's coalition, and at least one member belonging to I's coalition. I and C propose pandemic policy p and transfers t_i to each type in the coalition.

2. Each selectorate member chooses between I and C. I is deposed if at least one member of her coalition chooses C.

3. The pandemic policy of the chosen leader is implemented, her transfers allocated, and incomes taxed at rate τ.

We characterize a stationary equilibrium in which the incumbent stays in power.

First, note that each selectorate member would prefer a pandemic policy and level of transfer that would maximize her disposable income y_i. Thus, the best that any challenger C can offer to her nominated coalition is to choose p and t_i as though she herself were simultaneously maximizing each member's disposable income, subject to the government budget constraint, which we specify as $\tau\left[\sum_i (S_i[(\rho_i Y_{\rho i}) + (1 - \rho_i)\bar{Y}]\right] + R = \kappa p + \sum_i (w_i t_i)$, where $\tau\left[\sum_i (S_i[(\rho_i Y_{\rho i}) + (1 - \rho_i)\bar{Y}]\right]$ is total tax

revenues, R is other revenues, κ is the cost of implementing pandemic policy p, and $\sum_i (w_i t_i)$ is total transfers.

C offers to each coalition member of type i, policy p and transfer $\frac{t_i}{w_i}$ that solve

$$Max_{p,\frac{t_i}{w_i}}(1-\tau)[\rho_i Y_{\rho i} + (1-\rho_i)\bar{Y}] + \frac{t_i}{w_i} \; ;$$

$$s.t. \; \tau \Big[\sum_i (S_i[(\rho_i Y_{\rho i}) + (1-\rho_i)\bar{Y}]\Big] + R = \kappa p + \sum_i (w_i t_i) \; .$$

Re-writing the budget constraint as $t_i = \frac{\tau}{w_i}\Big[\sum_i (S_i[(\rho_i Y_{\rho i}) + (1-\rho_i)\bar{Y}]\Big] + \frac{1}{w_i}\Big[R - \kappa p - \sum_{-i}(w_{-i}t_{-i})\Big]$ and plugging into the maximand turns the problem into an unconstrained one:

$$Max_p (1-\tau)[\rho_i Y_{\rho i} + (1-\rho_i)\bar{Y}] + \frac{\tau}{w_i^2}\Big[\sum_i (S_i[(\rho_i Y_{\rho i}) + (1-\rho_i)\bar{Y}]\Big] + \frac{1}{w_i^2}\Big[R - \kappa p - \sum_{-i}(w_{-i}t_{-i})\Big] \quad (2)$$

Thus, focusing on interior solutions, in equilibrium, C proposes pandemic policy p_C and transfers t_{iC} to $w_i \; \forall i \in \{K\}$ that simultaneously satisfy K FOCs, each of the form:

$$
\begin{aligned}
F_i \equiv \quad & \Big[\frac{d\rho_i}{dp}(Y_{\rho i} - \bar{Y}) + \frac{dY_{\rho i}}{dp}\rho_i\Big]\Big[(1-\tau) + \frac{\tau S_i}{w_i^2}\Big] \\
& + \frac{\tau}{w_i^2}\Big[\sum_{-i}(S_{-i}[\frac{d\rho_{-i}}{dp}(Y_{\rho-i} - \bar{Y}) + \frac{dY_{\rho-i}}{dp}\rho_{-i}])\Big] \\
& - \frac{\kappa}{w_i^2} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad = 0.
\end{aligned}
\quad (3)
$$

These give p_C which, when plugged into the budget constraint, solves for $t_C^i \; \forall i \in \{K\}$. The value of C's offer for a member of type i is therefore:

$$V_C^i = y_i(p_C, \frac{t_{iC}}{w_i}) + \delta\Big[\frac{w_i}{S}V_I^i + (1 - \frac{w_i}{S})V_o^i\Big], \quad (4)$$

where δ is the discount rate, $\frac{w_i}{S}$ is the probability that a selectorate member of type i is included in the coalition of whoever is the incumbent, in which case she obtains the value $V_I^i = \frac{y_i(p_I,\frac{t_{iI}}{w_i})}{1-\delta}$, with p_I the pandemic policy of the incumbent, and $V_o^i = \frac{y_i(p_I,0)}{1-\delta}$ is the value of being outside the incumbent's coalition, in which case i is subject to the incumbent's pandemic policy, but no transfers are received.[4] Since the same pandemic policy is implemented for all selectorate members, whether or not they are in the coalition,

---

[4]In the canonical selectorate framework in which members are homogeneous, the probability that any one member is included in the coalition is $\frac{W}{S}$. Here, with W composed of heterogeneous groups, $\frac{\sum_i w_i}{S} = \frac{W}{S}$, with $\frac{W}{S}$ still capturing the type of regime, with values close to one-half approximating more democratic regimes.

then in equilibrium, $p_C = p_I = p$.

For the incumbent to stay in power, she has to match the value V_C^i from the challenger, i.e. $V_I^i = V_C^i$, which means $V_I^i = y_i(p_C, \frac{t_{iC}}{w_i}) + \delta \left[\frac{w_i}{S} V_I^i + (1 - \frac{w_i}{S}) V_o^i \right]$. Imposing $p_C = p_I = p$, substituting the above expressions for V_I^i and V_o^i, and re-arranging obtain:

$$\frac{y_i(p, \frac{t_I^i}{w_i})}{1 - \delta} - \left(\frac{1}{1 - \frac{\delta w_i}{S}} \right) \left[y_i\left(p, \frac{t_C^i}{w_i}\right) + \left(1 - \frac{w_i}{S}\right) \left(\frac{y_i(p, \frac{t_I^i}{w_i})}{1 - \delta} \right) \right] = 0. \tag{5}$$

It is straightforward to show that the incumbent gives transfers that are less than what the challenger would give. That is, re-writing the above as $\frac{y_i(p, \frac{t_I^i}{w_i})}{1-\delta} \left[1 - \left(\frac{1}{1 - \frac{\delta w_i}{S}} \right) \left(1 - \frac{w_i}{S}\right) \right] = \left(\frac{1}{1 - \frac{\delta w_i}{S}} \right) y_i(p, \frac{t_C^i}{w_i})$, it is obvious that $y_i(p, \frac{t_I^i}{w_i}) \leq y_i(p, \frac{t_C^i}{w_i})$ since:

$$\left(\frac{1}{1 - \frac{\delta w_i}{S}} \right) \left[\frac{1 - \delta}{\left[1 - \left(\frac{1}{1 - \frac{\delta w_i}{S}} \right) \left(1 - \frac{w_i}{S}\right) \right]} \right] > 1$$

or, simplifying, $S > w_i$.

More relevant for our analysis is the level of pandemic policy that the incumbent implements. Since the incumbent, to stay in power, would have to provide pandemic policy that is the same as what a challenger would provide, the optimal policy is given by the conditions $F_i = 0 \; \forall i \in \{K\}$. The following results can then be obtained.

Theorem 1 *The **politically optimal pandemic policy** (POPP) chosen by the incumbent government considers the effect of the policy on each group in its coalition, weighing them according to group size.*

Specifically, in equilibrium, for any pair (i, j) of groups in the coalition, where $P_i \equiv \frac{d\rho_i}{dp}(Y_{\rho i} - \bar{Y}) + \frac{dY_{\rho i}}{dp}\rho_i$ and $P_j \equiv \frac{d\rho_j}{dp}(Y_{\rho j} - \bar{Y}) + \frac{dY_{\rho j}}{dp}\rho_j$ denote the total marginal effect of the policy on i and j, respectively, $w_i^2 P_i = w_j^2 P_j$.

All proofs are in the appendix.

Theorem 2 *The effect of the POPP on each group in the incumbent's coalition is decreasing in group size – w_K is most, while w_1 is least, affected.*

Let the pair of parameters (η_i, ϵ_i) denote the Health Vulnerability-Economic Vulnerability (HV–EV) profile of a selectorate member of type i.

Theorem 3 *The POPP most closely reveals the HV–EV profile of the least loyal group in the incumbent's coalition.*

		Economic Vulnerability	
		Low	High
Health Vulnerability	Low	Moderate pandemic policy	Least stringent pandemic policy
	High	Most stringent pandemic policy	Moderate pandemic policy

Table 1

The model predicts that the least loyal members of the ruling coalition – those with the smallest probability of remaining in the coalition, determine pandemic policy. For this group, the total marginal effect P of the policy is largest. From the proofs of Theorems 1 and 3, one can write $P \equiv \left(\frac{\partial \rho}{\partial p} + \eta \frac{\partial(\frac{\sum_i S_i \rho_i}{S})}{\partial p} \right)(Y_\rho - \bar{Y}) + \left(\frac{\partial Y\rho}{\partial p} + \epsilon \frac{\partial(\sum_i S_i y_i)}{\partial p} \right)\rho$. Thus, when the health vulnerability of the least loyal group is high, while its economic vulnerability is low, then pandemic policy would be aggressive, as this would induce larger total health effect $\frac{\partial(\frac{\sum_i S_i \rho_i}{S})}{\partial p}$ which, when weighted by high η, would generate large P. Conversely, if the group's health vulnerability is low, while its economic vulnerability is high, a less aggressive pandemic policy would induce larger aggregate economic effect $\frac{\partial(\sum_i S_i y_i)}{\partial p}$ which, when weighted by high ϵ, would induce large P. Table 1 summarizes the predicted relationship between the health-economic vulnerability (HV-EV) profile of the least loyal members of the coalition and the politically optimal pandemic policy (POPP).

3   Applications

We now employ our model to explore cross-country variation in pandemic policy. The model predicts that pandemic policy is responsive to the preferences of the least loyal members of the coalition. As the taxonomy in Table 1 illustrates, if such members are have high health vulnerability but low economic vulnerability, the incumbent will choose more stringent policies. If such members are relatively more concerned about economic outcomes, that is they have low health vulnerability and high economic vulnerability, the incumbent will choose less stringent policies.

As there is no unified database that would permit us to conduct an econometric investigation, we rely on detailed case studies to operationalize the concept of least loyal or pivotal coalition members. Specifically, we first compare otherwise similar countries such as Brazil and Argentina, and Denmark and Sweden, who choose very different pandemic policies, before applying our model to the UK and the United States which

have comparable policies.[5] Several caveats are in order before proceeding. First, in no cases do we attempt to explain the entirety of the variation in pandemic policies; other factors both cultural and idiosyncratic are also clearly relevant. Second, identifying the most pivotal members of the incumbent coalition is not always obvious and relies on a degree of country-specific knowledge. Thus in all cases, we regard our analysis here as a starting point for further work.

Argentina and Brazil Argentina under Alberto Fernández adopted an aggressive pandemic policy. A full lockdown was imposed on March 19, at which point Argentina had only 128 official Covid-19 cases. This was kept in place until July 17th. The result has been that (thus far) the outbreak has been contained—-87,030 cases and 1,694 confirmed deaths as of July 9. In contrast, Brazil was slow to impose major restrictives and the federal government sought to limit local lockdown measures. President Bolsanaro fired two health ministers and opposed local attempts to impose lockdowns, for instance, by the mayor of San Paulo (see Appendix Figure 4). .

Existing explanations of this policy divergence point to the personality of Javier Bolsanaro and also to his lack of a party base. For instance, the Guardian notes that "Bolsonaro split with the party that brought him to power, whereas Fernández is a product of one of Latin America's most enduring and powerful national movements" (Goñi 2020).Another relevant factor was that Bolsanaro's policy response was shaped by the desire of opposing the lockdown policies of two prominent governors: of the State of Sao Paulo (Joao Doria), and of the State of Rio de Janeiro (Wilson José Witzel). But to better understand this divergence it is more informative to look at the respective coalitions that elected Bolsanaro and Fernández.

Bolsanaro was elected following a sharp recession. He was able to win power from the Workers' Party by bidding away the support of working class voters. The voters who swung the electors to Bolsanaro were non-ideological working class voters who voted on grounds of the economy and law and order (see Hunter and Power 2019).[6] These voters would be highly vulnerable to the economic consequences of a stringent pandemic policy (i.e. high EV). Our model suggests that Brazil's sluggish response to Covid-19 is a response to the concerns of the group that were pivotal in his election.

In contrast, Fernández was elected by a leftwing coalition. The working class poor—those with the highest EV—were already likely to vote for Everyone's Front–which comprised various social democratic, Peronist, and Communist parities. Our model predicts that the Fernández government simply does not have to respond to their demands. Rather policies will more responsive to more marginal members of the coalition who were more vulnerable to the pandemic itself. In the case of Everyone's Front, these may

---

[5]Note we focus on the *initial* response to Covid-19. We do not consider the speed at which countries reopened as these policies are likely endogenous to the number of cases.

[6]Specifically, while affluent voters in Brazil have long been concerned with crime in the 2018 election "Poorer segments, who not only lack access to such options (e,g. gated communities) but also typically reside in areas of greater crime, sought credible promises of protection as well" (Hunter and Power 2019, 74).

have been voters with Peronist learnings, who tend to be older than the other members of the Everyone's Front coalition.

Denmark and Sweden   Both Denmark and Sweden are governed by left-wing coalitions that won power in 2019. Both are Scandinavian social democracies with high trust in government and high levels of social capital. While there are social and cultural differences between these two countries, these do not seem strong enough to explain the very different responses to Covid-19.

Denmark shutdown on March 11 2020—the second country in Europe to do so. This was only twelve days after the first confirmed case was announced on February 27. On March 13 Denmark closed its borders. In contrast, Sweden refrained from imposing an official lockdown, instead issuing guidance on social distances and risk avoidance. Schools, cafes, and restaurants remained open (see Appendix Figure 5).

Our model provides an explanation of this divergence. In Sweden, the smallest group in the governing coalition is the Green Party. The voters of the latter skew young. Hence they have a low HV but they are likely to have a relatively high EV. Other explanations of Swedish policy rest on the discretion given to the Public Health Agency of Sweden (Folkhälsomyndigheten) and Anders Tegnell.[7] Nonetheless, this explanation does not explain why there was little pressure on the Public Health Agency from the government and why criticism for not adopting a more aggressive pandemic policy has come from parties on political right.[8] Future developments will allow us to further falsify or validate our model.

The UK and USA   The UK stands out from Figure 2 as having been slow to implement lockdown policies. The per capita fatality rates from Covid in the UK are also among the highest in the world. In the press, the reluctance to impose a strict lockdown early on has been attributed to the liberal political preferences of Prime Minister Boris Johnson. While these factors may have been relevant, our model points to the importance of the new electoral coalition that was responsible for providing the Johnson government with its large majority in December 2019.

Johnson's victory rested on voters in the north of England who traditionally voted Labour. In the words of the Economist: "The party of the rich buried Labour under the votes of working-class northerners and Midlanders." ("Victory for Boris Johnson's all-new Tories" 2019). From the perspective of our model, northern working class voters are the newest and most pivotal part of the current governing coalition. Indeed in his election results speech, Johnson acknowledged: "You may only have lent us your vote and you may not think of yourself as a natural Tory . . . And if that is the case, I am humbled that you have put

---

[7]Importantly, this discretionary authority is protected in the constitution and as such may seem to be outside the considerations of our model (Jonung 2020).

[8]See Lindeberg (2020).

your trust in me, and that you have put your trust in us. And I, and we, will never take your support for granted. And I will make it my mission to work night and day, flat out, to prove you right in voting for me this time, and to earn your support in the future."[9] These voters are particularly sensitive to downturns in the economy. Any lockdown policy imposed by Johnson had to carry the support of this key constituency. As a result, even had he been more alert to the threat posed by the pandemic, it is unlikely he would have been able to impose an earlier, more severe, lockdown policy.

The case of the United States is similar (see Appendix Figure 6). President Donald Trump was relatively slow to acknowledge the case of the threat posed by COVID-19. While this has been attributed to his personality, it also reflects his focus on the economy for the 2020 election. Indeed our model suggests that this focus on the economy is driven by concern over the least loyal members of the Trump electoral coalition.

Our framework can account for otherwise puzzling observations. For instance, a much remarked upon finding in the United States is that Covid-19 is more likely to kill Republicans than Democrats, as the former are older on average and more likely to have significant comorbidities. The Washington Post carried the headline "Republicans are endangering their own supporters and destroying Trump's electoral map" [10] Indeed Johnson, Pollock, and Rauhaus (2020) hypothesized that predicted fatalities from Covid-19 could swing elections in favor of the Democrats.

However, our model explains why Trump's policies have tended to favor reopening the economy despite fears of a second-wave. The incumbent should be most responsive, not to the concerns of his or her base— elderly, conservative, Republican voters in Southern states are unlikely to defect to the Democrats—but to the concerns of the most pivotal members of the coalition. Viewed through this lens, it is entirely rational for Trump's policies to mirror the concerns of blue-collar workers in states such as Wisconsin, Ohio, and Pennsylvania who were critical to his 2016 election win.

Numerous other arguments have been proposed to explain variation in the pandemic response. For example, it has been claimed that populist leaders have been especially slow in their response. But while this claim fits the examples of Bolsanaro and Trump, it does not fit other populist leaders like Rodrigo Duterte in the Philippines who reacted aggressively (Billing 2020). Moreover, this explanation does not explain why populist leaders would be slow to react. Our model suggests that part of the answer may lay in their electoral base, particularly, in the relative HV-EV vulnerability of the most pivotal supporters.

The electoral concerns we have discussed here can also apply to the decision to open up following a lockdown, and the nature of the economic stimulus passed in response to the pandemic. We leave an analysis of these policies to subsequent research.

---

[9]https://www.news24.com/news24/columnists/guestcolumn/transcript-boris-johnsons-election-victory-speech-in-full-20191213.

[10]see Rubin (2020) and Cadelago (2020).

CEPR PRESS

4 Concluding Comments

We have developed a model of the tradeoffs facing an incumbent leader facing a pandemic. Responding to the pandemic imposes heterogeneous costs on members of the incumbent's ruling coalition. Specifically, policies like lockdowns impose economic costs that are disproportionately bourn by some groups. The benefits of these policies disproportionately benefit those individuals most at risk to disease. We show that the incumbent's chosen is most responsible to the most pivotal members of the incumbent's ruling coalition. We discuss how this model sheds lights on variation in pandemic policies across otherwise similar countries.

References

Billing, Lynzy (Apr. 2020). "Duterte's Response to the Coronavirus: 'Shoot Them Dead'". *Foreign Policy*. `https://foreignpolicy.com/2020/04/16/duterte-philippines-coronavirus-response-shoot-them-dead/` visited on.

Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson, and Alastair Smith (2003). *The Logic of Political Survival*. Cambridge, MA: MIT Press, 2003.

Cadelago, Christopher (May 2020). "Study: Elderly Trump voters dying of coronavirus could cost him in November". *Politico*. `https://www.politico.com/news/2020/04/23/how-coronavirus-could-upend-2020-battlegrounds-204708` visited on.

Dingel, Jonathan I and Brent Neiman (Apr. 2020). *How Many Jobs Can be Done at Home?* Working Paper 26948. National Bureau of Economic Research, Apr. 2020. doi:`10.3386/w26948`. `http://www.nber.org/papers/w26948`.

Goñi, Uki (May 2020). "'You can't recover from death': Argentina's Covid-19 response the opposite of Brazil's". *Guardian*. `https://www.theguardian.com/world/2020/may/10/argentina-covid19-brazil-response-bolsonaro-fernandez` visited on.

Heath, Allister (July 2020). "The home-working revolution will derail the middle-class gravy train". *The Daily Telegraph*. `https://digitaleditions.telegraph.co.uk/data/298/reader/reader.html?#!preferred/0/package/298/pub/298/page/68/article/64478` visited on.

Hunter, Wendy and Timothy J. Power (2019). "Bolsonaro and Brazil's Illiberal Backlash." *Journal of Democracy* 30.1, 68–82.

Johnson, Andrew F., Wendi Pollock, and Beth Rauhaus (2020). "Mass casualty event scenarios and political shifts: 2020 election outcomes and the U.S. COVID-19 pandemic". *Administrative Theory & Praxis* 42.2, 249–264.

Johnson, Boris (2019). "Boris Johnson's Victory Speech". 2019. `https : / / www . news24 . com / news24 / columnists / guestcolumn / transcript - boris - johnsons - election - victory - speech - in - full - 20191213`.

Jonung, Lars (June 2020). *Sweden's Constitution Decides Its Covid-19 Exceptionalism*. Working Paper 11. Lund University, Department of Economics School of Economics and Management, June 2020.

Lindeberg, Rafaela (June 2020). "Sweden's PM Hits Back at Criticism Over Covid Strategy". *Bloomberg*. `https : / / www . bloomberg . com / news / articles / 2020 - 06 - 14 / sweden - s - pm - hits - back - at - criticism-over-outlier-covid-strategy` visited on.

Max Roser Hannah Ritchie, Esteban Ortiz-Ospina and Joe Hasell (2020). "Coronavirus Pandemic (COVID-19)". *Our World in Data*. Https://ourworldindata.org/coronavirus.

Pulejo, Massimo and Pablo Querubın (2020). "Electoral Concerns Reduce Restrictive Measures During the COVID-19 Pandemic". Working Paper. 2020.

Rubin, Jennifer (July 2020). "Republicans are endangering their own supporters and destroying Trump's electoral map". *Washington Post*. 07/13/20.

"Victory for Boris Johnson's all-new Tories" (Dec. 2019). *The Economist*. `https://www.economist.com/ leaders/2019/12/13/victory-for-boris-johnsons-all-new-tories` visited on.

A Appendix (Not For Publication)

Proofs

Proof of Theorem 1

Proof Condition (3) can be expressed as $F_i \equiv P_i\left[(1-\tau) + \frac{\tau S_i}{w_i^2}\right] + \frac{\tau}{w_i^2}[\sum_{-i}(S_{-i}P_{-1})] - \frac{\kappa}{w_i^2} = 0$, where $P_i \equiv \frac{d\rho_i}{dp}(Y_{\rho i} - \bar{Y}) + \frac{dY_{\rho i}}{dp}\rho_i$ is the total marginal effect of the pandemic policy p on coalition member of type i. This implies that, for, say a pair of members with types 1 and 2, $P_1[w_1^2(1-\tau) + \tau S_1] + \tau[S_2 P_2 + S_3 P_3 + ...S_K P_K] = P_2[w_2^2(1-\tau) + \tau S_2] + \tau[S_1 P_1 + S_3 P_3 + ...S_K P_K]$. This reduces to $P_1[w_1^2(1-\tau)] = P_2[w_2^2(1-\tau)]$, or $w_1^2 P_1 = w_2^2 P_2$. More generally, for any pair (i,j) of coalition members of type i and j, $w_i^2 P_i = w_j^2 P_j$.

Proof of Theorem 2

Proof The proof of Theorem 1 also implies that $w_1^2 P_1 = w_2^2 P_2 = ... = w_K^2 P_K$. Since $w_1 > w_2 > ...w_K$, it must be that $P_1 < P_2 < ...P_K$, with P_i denoting the total marginal effect of the pandemic policy on a member of group w_i.

Proof of Theorem 3

Proof By Theorem 2, $P_K - P_1$ is largest, and P_1/P_K closest to zero, than other pairs of total marginal effects. Set $P_1/P_K \approx 0$ and $\frac{1}{P_K} \approx 0$. Then $\frac{1}{\frac{d\rho_K}{dp}(Y_{\rho K}-\bar{Y})+\frac{dY_{\rho K}}{dp}\rho_K} \approx 0$, which is implicit in (η_K, ϵ_K), since $\frac{d\rho_K}{dp} = \frac{\partial\rho_K}{\partial p} + \eta_K \frac{\partial(\frac{\sum_i S_i \rho_i}{S})}{\partial p}$ and $\frac{dY_{\rho K}}{dp} = \frac{\partial Y_{\rho K}}{\partial p} + \epsilon_K \frac{\partial(\sum_i S_i y_i)}{\partial p}$ are functions of η_K and ϵ_K.

B   Additional Figures

Figure 3: The stringency index for selected countries at 10, 100, 200, and 500 deaths respectively.

CEPR PRESS
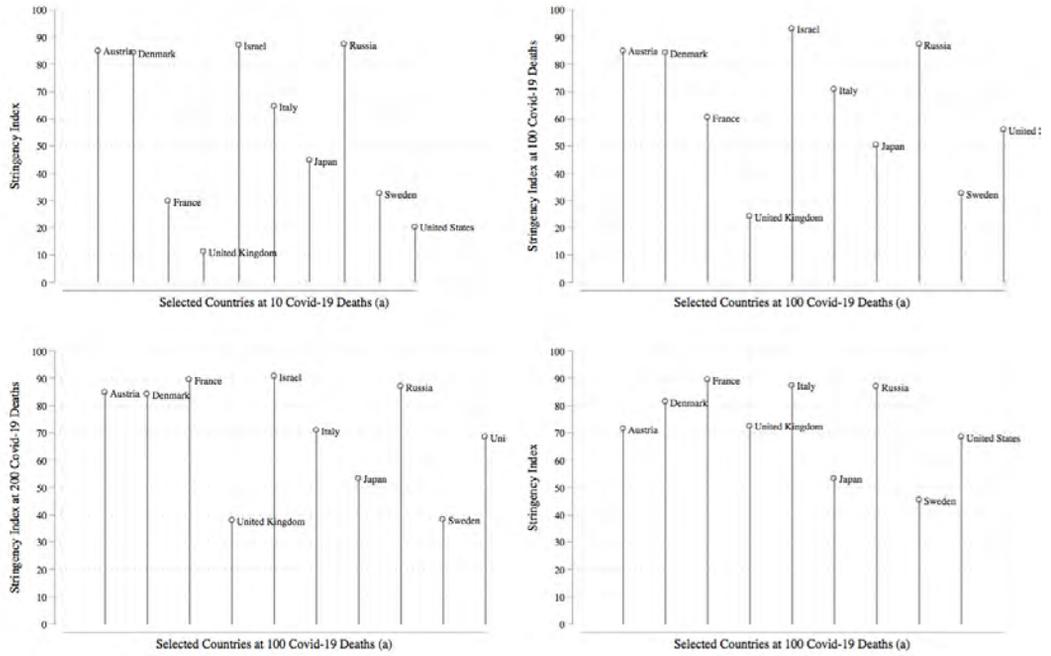
Figure 4: The contrasting pandemic policy of Argentina and Brazil.
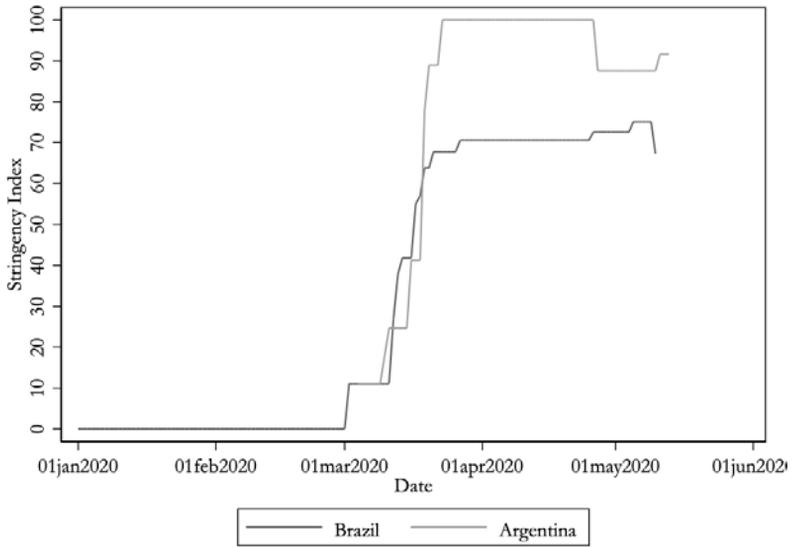
Figure 5: The contrasting pandemic policy of Denmark and Sweden.
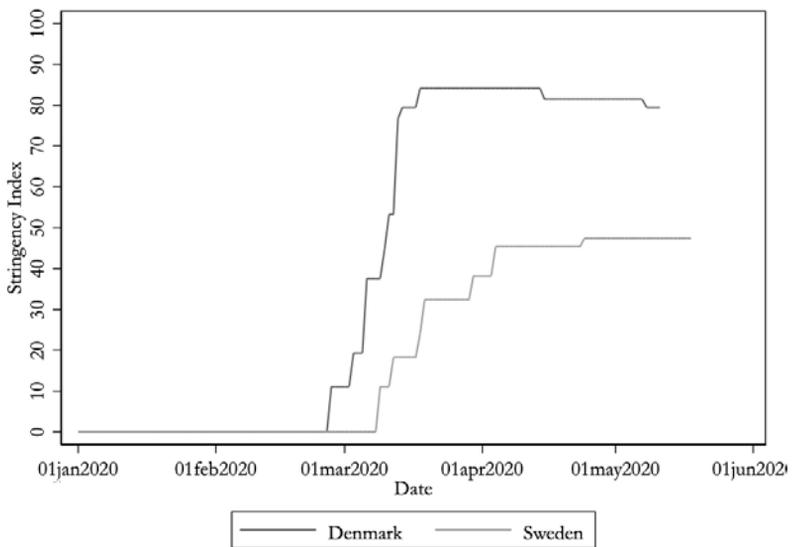
Figure 6: The similar pandemic policy of the UK and USA.
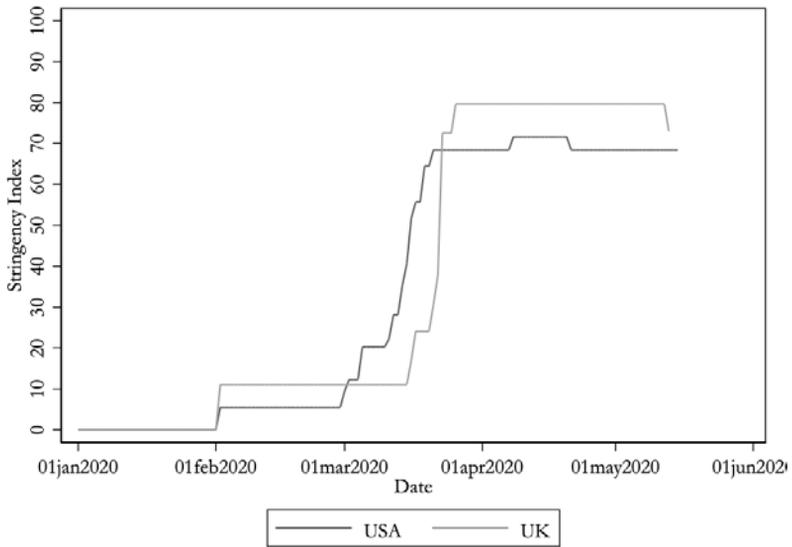
Lockdown and voting behaviour: A natural experiment on postponed elections during the COVID-19 pandemic

Tommaso Giommoni[1] and Gabriel Loumeau[2]

The goal of this paper is to study the impact of the lockdown policy on voting behaviour, during the COVID-19 pandemic. We focus on France, where a differential lockdown was implemented across departments, based on the local diffusion of the disease. In particular, the country has been divided in two areas, red and green, subject to a "hard" and a "soft" lockdown, respectively. To measure voting behaviour, before and after the policy, we rely on 2020 French municipal elections: the first round took place before the introduction of the restrictions, while the second round was delayed after the end of the lockdown. We estimate a Spatial Regression-Discontinuity-Design model comparing the difference in outcomes between the two electoral rounds, at the border of red and green areas. The main results suggest that lockdown regulations significantly affected electoral outcomes. First, in localities under a harder lockdown, the incumbent's vote share is higher as well as the consensus for Green parties. Second, voter turnout is larger where more stringent restrictions are adopted. These results suggest that lockdown measures strongly lead citizens to rally around the local incumbent politicians.

1   Postdoctoral researcher, ETH Zurich.
2   Postdoctoral researcher, ETH Zurich.

1 Introduction

The COVID-19 pandemic represents an unprecedented global crisis and poses massive challenges to contemporary democracies. Governments are faced with a global health problem that threatens the life of millions of people in the world and seriously undermines economic development. A common reaction to the insurgence of the pandemic has been the implementation of a set of social containment measures, known as the "lockdown", with the goal of limiting the spread of the disease.[1] On the one hand, these measures represent an effective tool to contain the diffusion of the virus, on the other hand, they limit individual liberty and constrain economic activity (Fetzer et al., 2020).

There is already evidence that these unparalleled restrictions affect individuals political preferences, as they raise the support for the incumbent (Bol et al., 2020; De Vries et al., 2020), as well as social trust and confidence towards the institutions (Amat et al., 2020; Esaiasson et al., 2020). Nevertheless, due to the impossibility of hosting electoral events during the pandemic, very little is still known on how lockdown measures impact voting behaviour. Studying the relation between lockdown restrictions and electoral outcomes is of vital importance in order to understand who will be in charge of managing the post-pandemic period and, consequently, which policies will be implemented.

To fill this gap in the literature, this paper studies the short-term effects of lockdown measures on voting behaviour and electoral outcomes. We rely on a natural experiment taking place in France during the COVID-19 outbreak. Similarly to the other European countries, France was severely affected by the pandemic with 173,304 confirmed COVID-19 cases and 30,120 total casualties, on July 15th 2020. The French government introduced the lockdown before the surge of the pandemic, on March 17th. This consisted of a series of restrictions which include the closure of most public establishments, such as schools and universities, and the ban of non necessary movements. After the pandemic reached the contagion peak in April, the government decided to start loosening the restrictions but *the relaxation was not uniform across French departments*. In particular, the French territory was divided into two areas, a green one and a red one, depending on the local severity of the pandemic. Assignment to either areas was based on an arbitrary threshold defined by four local health criteria.[2] Departments in the green zone experienced a faster relaxation – "soft" lockdown – than those in the red

---

[1]Lockdown measures have been implemented by many European countries: Italy (March 9), Spain (March 15), Austria (March 17), France (March 17), Denmark (March 18), Germany (March 22), Greece (March 23), the Netherlands (March 24), and the United Kingdom (March 24). Moreover, the recent resurgence of the disease in specific areas, e.g., Atlanta, Barcelona, two German Kreis, has forced governments to reintroduce lockdown measures.

[2]The four criteria are: (i) number of cases per 100,000 inhabitants, (ii) rate of positive virologic tests, (iii) reproduction number, R0, (iv) hospital bed occupancy rate.

one – "hard" lockdown. In the green zone, movements within 100 Km were allowed again and the economic activity restarted, while in the red zone, the standard restrictions remained in force. This division remained effective from May 11th to June 2nd. Figure 1 shows the map of France with the border between the red area, North-Eastern France, and the green area. By focusing on municipalities around the lockdown border, we are able to analyze the impact of restriction measures on electoral outcomes in areas with a similar local health situation. We do not observe, indeed, significant differences in the number of cases or deaths by COVID-19 in those departments –in the two zones– that are closely located to the border.[3]

Figure 1: Differential lockdown from May 11$^{\text{th}}$ to June 2$^{\text{nd}}$, 2020
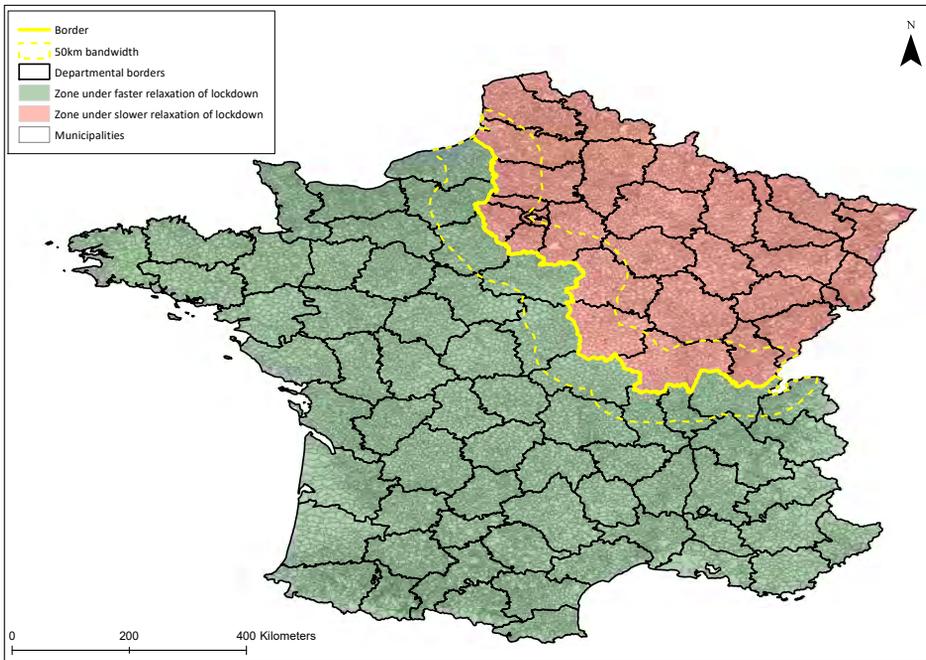
Notes: The yellow bold line represents the border between the red and the green area, subject to different lockdown policies. Black borders represent departments and grey ones indicate municipalities. Dotted yellow lines represent the 50 Km bandwidths from the border.

---

[3]More precisely, we formally test for this assumption by looking at the data on the number of cases and deaths by COVID-19 during the periods preceding the first and second round of elections at the departmental level. We do not observe large differences in the severity of the health crisis between the two zones for the departments close to the border. Given that the disease spreads locally in a smooth manner, the absence of observed differences at the departmental level likely indicates that municipalities at the border indeed experience similar health situations. Therefore, in the sample of municipalities under consideration, the treatment mostly captures the difference in lockdown and we can interpret the main results as, primarily, due to the different containment measures. For more details, see Section 2.

To measure voting behaviour of French citizens before and after the lockdown, we rely on the 2020 French municipal elections, whose implementation was interrupted by the pandemic.[4] The first round of the polls took place on March 15th, while the introduction of the lockdown two days after the elections, on March 17th, impeded the conduct of the second round, that was delayed until the conclusion of the restrictions, on June 28th. This unique setting allows us to evaluate the voting behaviour of French voters just before –first round– and just after –second round– the introduction of the lockdown. Furthermore, we naturally control for fixed local characteristics and, given the short time span between the two rounds and the limit on mobility, we deem implausible the presence of endogenous spatial sorting as a response to the lockdown.

Overall, this institutional setting gives rise to two sources of variation we can exploit in order to estimate the causal effect of lockdown restrictions on electoral outcomes. On the one hand, the spatial variation between the red and the green zone allows us to study the effect of the severity of lockdown measures in areas with a similar local health situation; on the other hand, the time variation between the first and the second round can be used to assess the impact of the introduction of these restrictions, accounting for fixed local characteristics. Our identification strategy is structured as a spatial Regression-Discontinuity-Design (RDD), where the distance to the border represents our forcing variable and the dependent variable is the difference in outcomes between the two electoral rounds. Therefore, our strategy is based on the comparison between cities located across the border, and then subject to different rules, before and after the implementation of the lockdown.

The main results suggest that the severity of the lockdown significantly affects electoral outcomes and voting behaviour. First, we find that localities subject to "hard" restrictions have significantly different voting patterns compared to those under "soft" measures. Cities in the red zone, indeed, support more the local incumbent, whose vote share is between 2 and 4.5 percentage points higher. The higher visibility of the mayors in the red zone and the multiple contacts with the citizens may have contributed to increase their local popularity and electoral success. This result is also in line with the "rally around the flag hypothesis" (Baum, 2002; Mueller, 1970), claiming that the trust in leaders tends to raise in the wake of disasters: voters in the areas with longer restrictions may either feel safer or be more aware of the health crisis, and therefore grateful for the policy.[5] Moreover, we find that "hard" lockdown measures lead

---

[4]The first round of the municipal elections took place after a vivid debate. On the days before the first round, several parliament members publicly asked to postpone the first round due to contagion risks. However, on March 12th (i.e., three days before the elections), President Macron consulted the presidents of the main French political parties who all agreed to maintain the first round on March 15th as initially planned. The *Union of French Mayors* also publicly advocated in favor of maintaining the initial calendar.

[5]Chanley (2002) shows an important example of the "rally around the flag hypothesis" documenting that in the days following the 11 September 2001 attacks, trust in the U.S. government rose to levels not seen since the

to stronger support to Green parties, suggesting that a higher awareness of the health crisis persuade voters to be more concerned about environmental issues, potentially connected with the pandemic (Grant and Tilley, 2019). Second, we find that the enforcement of the "hard" lockdown raises political participation, measured with voter turnout. Being subject to a longer lockdown seems to mobilize voters, motivated to choose the right leader, rather than discourage them to participate, in fear of the disease.[6]

This paper relates to several strands of literature. First, we contribute to the recent literature on the political effect of lockdown measures during the COVID-19 pandemic. Bol et al. (2020) rely on an online survey covering 15 Western European countries and show that the lockdown led to higher consensus -vote intention- for the national incumbent, trust in the government and satisfaction with democracy. De Vries et al. (2020) find similar conclusions looking at the introduction of restrictions abroad. Moreover, Sibley et al. (2020) use experimental data to show that restrictive measures increased trust in science, politicians and police as well as patriotism in New Zeland.

More generally, our paper is related to the fast growing literature on the effect of the COVID-19 pandemic on political outcomes. Merkley et al. (2020) show that the pandemic led to an increase in consensus for the government in Canada (see also Harell (2020)). Leininger and Schaub (2020) study the local elections of the German state of Bavaria at the beginning of the pandemic –before the introduction of the lockdown– and find that the incidence of the disease advantaged the incumbent party. Moreover, Adam-Troian et al. (2020) study the first round of 2020 French municipal elections and find that areas more afflicted by the pandemic showed higher support for conservative parties. At last, there is evidence that the health crisis positively affected social trust and confidence toward institutions (Esaiasson et al., 2020) as well as the support for authoritarian and technocratic government (Amat et al., 2020).

Finally, we relate to the contributions studying the impact of natural disasters and catastrophic events on political outcomes. A set of studies analyzes the effect of recent epidemics on political preferences and electoral results: there is evidence on the Ebola outbreak (Beall et al., 2016; Campante et al., 2020; Maffioli, 2018) and on the HIV/AIDS epidemic (Mansour et al., 2020). Furthermore, many papers focus on the impact of natural disasters such as earthquakes and hurricanes. This literature suggests that the policy response is vital in determining the impact on incumbent's support: if the reaction is considered to be inadequate, voters tend to punish the leaders (Akarca and Tansel, 2016; Eriksson, 2016), while a positive response leads to an increase in their consensus (Bechtel and Hainmueller, 2011; Healy and Malhotra, 2009).

1960s.

[6] In these regards, Zeitoun et al. (2020) show that there is no association between the level of participation to the (first round of) 2020 municipal elections in France and the subsequent spread of the disease at the local level.

To the best of our knowledge, we are the first to study the causal impact of the COVID-19 lockdown on electoral outcomes, making use of electoral –post lockdown– data rather than survey or experimental evidence. Moreover, we advance the literature as, thanks to this unique setting, we study within-country differences in lockdown measures and we are able to assess the impact of the severity of these restrictions on electoral outcomes. Finally, we are among the first to focus on local incumbents, rather than on national politicians.

The remainder of the paper is structured as follows. Section 2 describes the institutional setting and the data. Section 3 presents the empirical strategy. Section 4 discusses the main results and the robustness checks and Section 5 concludes.

2 Institutional background and data

2.1 COVID-19 pandemic in France and lock-down measures

In France, the daily number of confirmed COVID-19 cases started to increase significantly in the first days of March. Figure 2 display the evolution of the COVID-19 in France from January 1st to July 15th, 2020. Following this increase, the French government announced, on March 16th (i.e., one day after the first round of the municipal elections), a strict lock-down starting on March 17th at 12am. Individuals were only allowed outside of their homes for specific reasons, (e.g., grocery shopping), most of public establishments have been closed as well as schools and institutes of higher education and most of religious gatherings have been forbidden. Police forces controlled the respect of the lock-down, and non-compliance was fined.[7]

Thanks to the lockdown measures, the pandemic peak was reach in early April. However, whereas the situation was mostly under control in the South-West half of France, the number of cases was still high in the North-Eastern half of the country. Figure 1 displays the two zones. Consequently, the government decided to relax the lockdown deferentially as of May 11th. The zones were defined using four indicators at the department level (NUTS3 region): (i) number of cases per 100,000 inhabitants, (ii) rate of positive virologic tests, (iii) reproduction number, R0, (iv) hospital bed occupancy rate. The blue vertical lines, in Figure 2, show the evolution of the lockdown policy over time.

In the green zone, movements were allowed within 100km of one's residential location and the economic activity started again. In the red zone, on the contrary, all the restrictions remained in force and, *except for Paris*, the lock-down was relaxed only on June 2nd. In Paris' region, the lockdown was lifted on June 22nd: between June 2nd and June 22nd, the lock-down in this region only slightly loosened relative to before, and the region was attributed a yellow

---

[7]In the standard case, amount of the fine was €135.

Figure 2: Evolution of Covid-19 confirmed cases in France
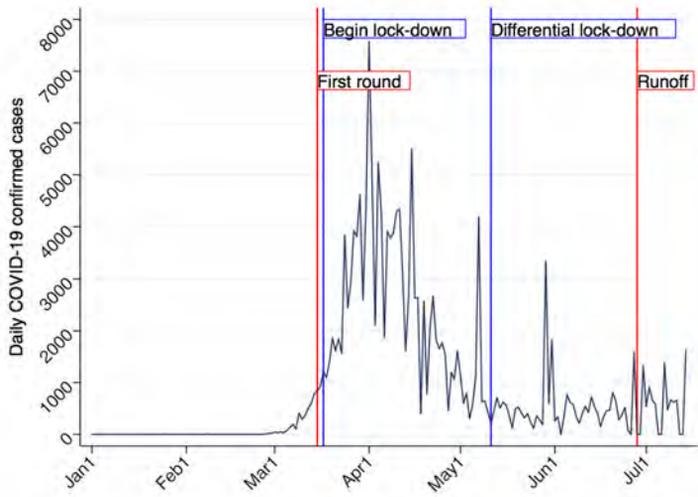
Notes: The plot shows the total number of confirmed COVID-19 cases in France starting from January 1st 2020. The red lines indicate the dates of 2020 local elections (first round -March 15th-, runoff -June 28th-), the blue lines indicate the dates of the modification in the lockdown policy (introduction of the lockdown -March 17th-, first relaxation of the lockdown -May 11th-). The source is the French Government data portal (`https://www.data.gouv.fr/fr/`).

color by the Health Ministry (Figure 5, Appendix A).

2.2 Municipal election

French municipal elections allow local constituents to elect the city council, with the mayor chairing the council. Elections take place every six years. The voting procedure differs by municipality size. On the one hand, in municipalities with less than 1,000 inhabitants, a two rounds block vote with panachage (or majority-at-large voting) takes place. All candidates that received more than 50% of the votes are elected to the council. Those that did not reach this threshold compete in a second round for the remaining seats. On the other hand, in municipalities with more than 1,000 inhabitants, a proportional representation with a premium for the majority takes place. Voters choose between different lists and can neither add nor remove candidates from the lists. If a list obtains the absolute majority from the first round, no second round takes place. Otherwise, lists that have obtained at least 10% of the vote remain for a second round. The first half of the seats is attributed to the lists with the most

vote, whereas the remaining one is attributed proportionally to all lists that received at least 5% of the votes. Due to the highly personalized nature of the elections in municipalities with less than 1,000 inhabitants (which leads to many candidates being elected in the first round) and to the substantial differences in the two electoral systems, we focus in this paper only on municipalities with 1,000 inhabitants or more.

Based on the decision by the Ministry of the Interieur on July 16th, 2019, the 2020 French municipal elections had been scheduled on March 15th for the first round, and on March 22nd for the runoff. Regardless of the development of the pandemic in France, the first round was maintained and took place as planned on March 15th, one day before the lockdown started. However, given the COVID-19 outbreak, the second round was postponed after the lockdown on June 28th. All candidates elected during the first round took office immediately, whereas incumbents remained in power when runoffs were required. Overall, a runoff took place in 14% of all polling stations, and in 20% of all polling stations located in municipalities with 1,000 or more inhabitants. The red vertical lines, in Figure 2, show the timeline of the 2020 municipal elections.

At the time of the first round, the COVID-19 situation was very similar across the future differential lockdown border, as revealed in Panel a of Figure 3. Except clusters in the north of Paris and in Alsace, the number of deaths was relatively small (i.e., between 1 and 9 deaths per 100,000 inhabitants).[8] Hence, even if voters could possibly expect the pandemic to develop across the country, they could not reasonably anticipate that France would be divided in differential lockdown zones; and even less so, forecast the location of the boundary.

Panel b of Figure 3 displays, then, the number of deaths per 100,000 inhabitants by COVID-19 during the full period between mid-March and mid-July. As one would expect, given that lockdown zones were defined based on the local evolution of the pandemic, the North-Eastern half of France is significantly more impacted by the virus. However, although the departments on the north side of the lockdown border are indeed more impacted, the difference in the number of deaths per 100,000 inhabitants remains relatively small if we consider departments located around the border. This observation is worth noting to interpret our results, as it implies that the health situation did not differ largely around the border, at the time of the second round. Hence, pure COVID-19 rational, aside from the lockdown, could not reasonably explain large differences in voting outcomes around the border and we can interpret the main results of this analysis as, mostly, due to the different lockdown policies.

---

[8]Looking at the departmental hospitalization rate per 100,000 inhabitants, instead of the death rate, we obtain a qualitatively similar picture.

Figure 3: Number of deaths by COVID-19 (by department)

(a) End of March   (b) March to July
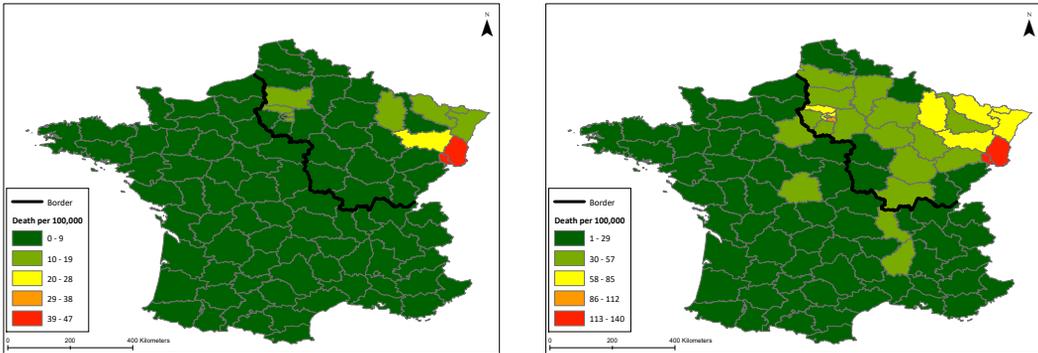
Notes: Death count is reported per 100,000 inhabitants. "End of March" corresponds to the period from March 19th until March 31st, whereas "March to July"jee refers to the period from March 19th until July 14th, 2020. *Source:* French Ministry of Health.

2.3 Data

The sample in analysis consists in all municipalities that host a runoff in 2020 French municipal elections, with 1,000 or more inhabitants. The data used in this paper comes from mostly three sources. All elections data is provided by the French centralized public data platform (`https://www.data.gouv.fr/fr/`). Election data is recorded at the polling station level. Information about the differential lockdown and health outcomes is obtained from the open data platform of the Health Ministry (`https://www.santepubliquefrance.fr/`). Finally, all economic and demographic municipal information is provided by the National Institute of Statistics and Economic Studies (`https://www.insee.fr/en/accueil`). Based on this data, we use the geographic information software ArcGIS to derive all spatial information.

Table 6, in the Appendix A, shows the descriptive statistics. Panel A shows the electoral outcomes expressed in difference between the two electoral rounds: the average difference in incumbent's vote share is 6.7%, while the one in turnout is -2.8%. Moreover, the average difference in margin of victory and number of white votes is 6.4% and 2.4%, respectively.

3 Empirical approach

In this section, we detail the identification strategy used to identify the causal effect of the lockdown measures on voting outcomes.

3.1 Estimation strategy

Consider $p \in P$ polling stations within $i \in I$ municipalities. All polling stations within a municipality are assigned the same location relative to the border.[9] Location of polling stations within a municipality are defined by the infinite set of municipal border points, $\overline{\boldsymbol{L}}$. Further, consider $\overline{\boldsymbol{B}}$ as the infinite set of border points constituting the border between the differential lockdown zones. Let us then define the subsets $\boldsymbol{L} \in \overline{\boldsymbol{L}}$ and $\boldsymbol{B} \in \overline{\boldsymbol{B}}$ of border points $\boldsymbol{l}_{pi} = (l^x_{pi}, l^y_{pi})$ and $\boldsymbol{b}_i = (b^x, b^y)$, such that the euclidean distance to the lockdown border $d_{pi} = ||\boldsymbol{l}_{pi} - \boldsymbol{b}_i||$ is minimized. Hence, this distance will be equal to 0 for polling station in a municipality at the border. Finally, define two zones $\mathcal{A}^+$ and $\mathcal{A}^-$ as the treatment and the control areas.

Location relative to the border acts as the forcing variable. Assignment into treatment is then a function of a municipality's location relative to the border. Formally, treatment status T_{pi} of polling station p in municipality i is defined as $T_{pi} = \mathbb{1}\left[\boldsymbol{L}_{pi} \in \mathcal{A}^+\right]$. To account for all fixed characteristics of polling station, and by extension of municipalities, we study the difference in outcomes between the two electoral rounds, which we denote with $\Delta \boldsymbol{Y}$. We then focus on the discontinuity of the expected outcomes at the geographical border:

$$\tau(\boldsymbol{d}) \equiv \mathbb{E}[\Delta \boldsymbol{Y}_1 - \Delta \boldsymbol{Y}_0 | d_p = 0] = \lim_{\boldsymbol{d} \to 0}[\Delta \boldsymbol{Y} | l_p \in \mathcal{A}^+] - \lim_{\boldsymbol{d} \to 0}[\Delta \boldsymbol{Y} | l_p \in \mathcal{A}^-]. \tag{1}$$

Parametric specification — We approach the spatial RD design in both a parametric and non-parametric way. In the parametric approach, we define the conditional expectations in (1) as $\mathbb{E}[\Delta \boldsymbol{Y}_0 | \boldsymbol{d}] = \alpha + f(\boldsymbol{d})$ and $\mathbb{E}[\Delta \boldsymbol{Y}_1 | \boldsymbol{d}] = \alpha + \tau + f(\boldsymbol{d})$, where $f(\boldsymbol{d})$ refers to flexible polynomials of shortest distance to the border. Allowing for asymmetric control distance functions insures that a kink is not misinterpreted as a discontinuity Lee and Lemieux (2010). The regression model is then:

$$\Delta \boldsymbol{Y} = \alpha + \boldsymbol{T}\tau + f(\boldsymbol{d}) + \boldsymbol{\epsilon} \tag{2}$$

The credibility of the model rests on the valid specification of the control functions. Therefore, we run various specifications with different order of polynomials and with different bandwidth sizes.

Non-parametric specification — For the non-parametric specification, we estimate local-polynomial regression-discontinuities with robust confidence intervals and optimal bandwidth selection following Calonico et al. (2014). The performance of standard local polynomial estimators may be seriously limited by their sensitivity to the specific bandwidth employed. Hence,

---

[9] Given the average size of French municipalities, i.e., approx. 5km^2, this assumption is expected to have a little impact, if at all, on the spatial distribution of polling stations relative to the border.

we employ mean squared error optimal bandwidths, which are valid given the robust approach in Calonico et al. (2014).

3.2 Validity of the approach

Causal identification in Regression Discontinuity designs rest on two key assumptions: (i) no selective sorting takes place, (ii) all underlying variables – besides treatment – vary smoothly at the border. The spatial RD design is a special case as assignment to treatment is very difficult to manipulate for municipalities as regional borders are very stable over time. Plus, the suddenty and unpredictableness of the spread of the COVID19 made it very unlikely for individuals and municipalities to anticipate and strategically react to the pandemic.

We approach the second assumption in two ways. First, in this paper, our main outcomes of interest are variable differences between the first and the second round of 2020 French municipal elections across the lockdown border. Hence, all fixed municipal characteristics are differenced out by design. Over extensive exploratory balance tests, this approach offers the key advantage that all fixed characteristics are accounted for as opposed to all *observed* fixed characteristics.

Second, smoothness tests are nonetheless required to insure the smooth variation of the underlying voting variables. Two key concerns arise. First, are runoff more likely to occur in either treatment or control zones? Second, are incumbent more likely to be present in the first and second rounds on either side of the border? A positive answer to any of these questions would endanger our estimation strategy. To answer these questions, we run smoothness tests around the border using 1km and 2km bins symmetrically on both side of the border. Results are displayed in Table 1. Overall, we observe no significant difference around the border in the density of runoffs, in the density of incumbent running for reelection, and in the density of incumbent present in the runoff. As always absence of evidence is not evidence of absence; however, these results are in line with the assumption of smooth underlying variables.

Table 1: Smoothness tests

	Runoff		Incumbent runs		Incumbent in runoff	
Bin size	1km	2km	1km	2km	1km	2km
RD Estimate	6.598	12.313	12.157	13.983	12.462*	15.740
	(11.452)	(24.156)	(7.904)	(14.461)	(7.407)	(14.269)
Obs.	88	49	70	47	71	47

Notes: The unit of analysis are bands of 1km or 2km width within 50km of the border on either side.

4 Results

4.1 Spatial regression discontinuity design analysis

We start by presenting the results of the spatial RD analysis of incumbent's vote share and turnout.[10] Following the identification strategy presented in Section 3.1, we compare the difference in outcomes between the two rounds of the election on both sides of the lockdown border.

Table 2, Panel a, presents the results for incumbent's vote share. The sample focuses on those municipalities where the incumbent runs in the first and in the second round. However, as tested in Section 3.2, no differences in the likelihood of the incumbent being present in either the first or second rounds of the election are observed around the lockdown border. Three main specifications are displayed. Columns (1-3) focus on the parametric estimation of the lockdown effect *without controlling for covariates* and using different polynomial orders. Columns (4-6) adopt a similar strategy but include economic, demographic and geographic municipal covariates.[11] It also includes a dummy equal to unity for Paris region, i.e., the "yellow zone" (Figure 5, in Appendix A). The AIC is used to compared the different parametric specifications. Finally, columns (7-8) present the results of a non-parametric approach à la Calonico et al. (2014) without and with municipal covariates, respectively. Overall, incumbent's vote share appears between 2 and 4.5 percentage points higher in the zone with the longer lockdown. Controlling for municipal covariates appears to increase the magnitude of the effect. To assess the robustness of the parametric results to the choice of the bandwidth, Table 7 (Appendix A), Panel a, reports results under the parametric specification using a wider range of bandwidth sizes. The bandwidth choice appears to have minor impacts on both the estimated coefficients and the significance levels derived. These results may derive from the higher visibility that the mayors in the red zone received as well as the multiple contacts with the citizens during the health crisis. Moreover, these findings argue in favor of the "rally around the flag hypothesis" (Baum, 2002; Mueller, 1970), as trusts in the local incumbent appears to increase in the wake of the pandemic. In this setting voters subject to longer restrictions could either feel safer or be more aware of the health crisis, and therefore grateful to the administration for the policy.

Table 2, Panel b, focuses on the voter turnout. Overall, both the parametric and non-parametric approaches point at a positive effect of lockdown on voter turnout. An estimated causal effect between 1.5 and 3 percentage points is observed. Controlling for covariates does

---

[10]We also studied the local margin of victory and the share of white votes. The lockdown appears to have had no significant effects on these outcomes. The analysis of these outcomes is available upon request.

[11]In particular, the following covariates are included: number of old people (those in the age groups 60-74, 75-89, more than 90), number of female individuals, population growth rate between 2011 and 2016 (in percentage) and median income in 2017. We also include department fixed effects.

Table 2: Spatial RD design analysis of turnout

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Approach	Parametric			Parametric			Non-parametric	
Bandwidth	50	50	50	50	50	50	Opt.	Opt.
Polynomial order	2	3	4	2	3	4	Opt.	Opt.
Covariates	No	No	No	Yes	Yes	Yes	No	Yes
	(a) Incumbent's vote share							
Treatment	1.585**	5.904***	1.543	3.303***	3.534**	4.836**	4.592***	2.947*
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)	(1.753)	(1.704)
Adj. R^2	0.04	0.10	0.12	0.22	0.29	0.29	-	-
AIC	11,084	10,969	10,940	10,742	10,603	10,588	-	-
Obs.	1,645	1,645	1,645	1,645	1,645	1,645	2,586	2,586
	(b) Voter turnout							
Treatment	1.301***	1.438***	0.393	1.990***	2.919***	2.343***	1.985**	3.841***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)	(0.827)	(0.756)
Adj. R^2	0.09	0.09	0.09	0.26	0.26	0.26	-	-
AIC	22,504	22,505	22,482	21,664	21,658	21,657	-	-
Obs.	4,009	4,009	4,009	4,009	4,009	4,009	5,998	5,998

Notes: Unit of analysis is the polling station. Table 7 (Appendix) reports results under the parametric specification using a wider range of bandwidth sizes.

not appear to have any systematic impact on the estimated coefficients. As for incumbent's vote share, we assess the robustness of the parametric results to the choice of the bandwidth in Table 7 (Appendix A), Panel b, by using a wider range of bandwidth sizes. The bandwidth choice appears to have minor impacts on both the estimated coefficients and the significance levels derived. This results further confirm the "rally around the flag hypothesis" in this setting, suggesting that the longer lockdown seems to mobilize voters, motivated to choose the right leader, rather than discourage them to participate, in fear of the disease.

Figure 4 present non-parametric discontinuity plots of incumbent's vote share (Panel a) and voter turnout (Panel b). A third order polynomial specification using a symmetric 50km bandwidth is used. Bin size are selected using the IMSE-optimal evenly-spaced method with spacing estimators. Confirming the results in Table 2, a positive jump at the lockdown border is observed in both outcomes.

CEPR PRESS

Figure 4: Non-parametric discontinuity plots

(a) Incumbent's vote share

(b) Voter turnout
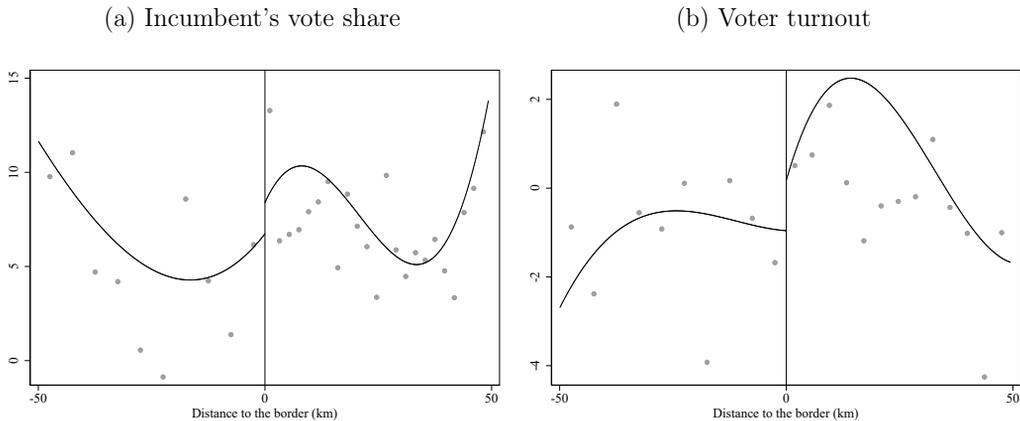
Notes: Third order polynomial with 50km bandwidth used. Bin size selected using the IMSE-optimal evenly-spaced method with spacing estimators.

4.2 Evolution in vote share by political groups

In our analysis, we have set party considerations aside so far. We now investigate whether the lockdown has affected parties' performance heterogeneously. As our estimation strategy relies on polling stations close to the lockdown border, we can only rigorously analyze parties with a good spatial coverage in both the first and second round. Extreme left and extreme right parties in France are highly concentrated in specific locations. The communist party is, for example, present in former blue collar municipalities at the periphery of the Paris metropolitan areas. Similarly, the "Rassemblement National", i.e., the main extreme right nationalist party, is mostly present in the North and South East of the country. President Macron's party "La République en Marche" (center right) is also characterized by a weak spatial coverage in the 2020 municipal elections for mostly two reasons. First, the party is very new as it was created for the 2017 presidential elections. Second, it performed badly at the 2020 municipal elections and in many cases was not present in the second round.

For these reasons, we focus on three political groups: (i) left-wing parties; (ii) Greens and affiliated parties, (iii) right-wing parties.[12] All three groups have proposed candidates in most municipalities studied and have often been able to qualify for the second round. We test for

---

[12]"Left-wing" and "right-wing" parties refer to all parties categorized as "diverse left" and "diverse right", respectively, by the French Ministry of the Interieur. Green and affiliated refers to the Green party and the lists labeled "Liste Union de la Gauche" (for which a direct translation would be: List of the United Left) which regroups left-wing parties alongside the Green party.

the smoothness of the presence of each party in the second round in Table 8 (Appendix A). No significant difference is observed around the lockdown border. We also look at different bandwidth sizes to test the robustness of the results to bandwidth selection. In Table 3, we display results using 60 km and 50 km bandwidths.

The Green and affiliated parties appear to have gained the most from the lockdown. A treatment effect between 11 and 22 percentage points is estimated.[13] We interpret this very substantial gains as the result of the perception of the pandemic as a natural disaster. The possible link between the increased frequency of epidemics in the last 20 years and human activity would substantiate this perception. Right-wing parties also appear to have gained vote share between the two rounds. Discontentment against the health crisis may have induced voters of the presidential party to move to the ideologically nearest party, i.e., right-wing parties.

Table 3: Evolution of the vote share by political groups

Bandwidth	60	60	60	50	50	50
Polynomial order	3	4	5	3	4	5
	(a) Left-wing					
Treatment	5.522***	5.380**	3.828	2.782	6.667**	5.993
	(1.858)	(2.399)	(2.804)	(2.601)	(2.965)	(3.707)
Adj. R^2	0.062	0.061	0.062	0.151	0.169	0.187
AIC	15523.89	15527.00	15527.88	6194.26	6178.74	6161.65
Obs.	2,152	2,152	2,152	853	853	853
	(b) Green and affiliated					
Treatment	6.348	11.190***	22.667***	51.350***	55.147***	68.686***
	(4.149)	(4.272)	(5.203)	(6.172)	(5.245)	(8.184)
Adj. R^2	0.164	0.192	0.193	0.373	0.580	0.591
AIC	14764.93	14702.01	14701.43	1759.72	1653.38	1647.73
Obs.	1,916	1,916	1,916	271	271	271
	(c) Right-wing					
Treatment	3.894***	3.723***	3.070**	5.924***	6.514***	7.839***
	(1.180)	(1.283)	(1.367)	(1.432)	(1.402)	(1.822)
Adj. R^2	0.080	0.081	0.086	0.136	0.147	0.172
AIC	15625.44	15625.59	15614.78	6303.34	6294.70	6269.39
Obs.	2,188	2,188	2,188	895	895	895

Notes: Unit of analysis is the polling station. Results of smoothness test around the border *by party* are displayed in Table 8.

---

[13]Preferred specification selected based on the AIC value. Overall, the large coefficient estimated using a bandwidth of 50km appears to be mostly driven by outliers given the relatively small number of observations analyzed. Hence, in this particular case, the use of a larger bandwidth appears reasonable.

CEPR PRESS

4.3 Additional analysis and robustness checks

Local Difference-in-Difference (DiD) analysis — As a robustness check, we run in this section the main analysis with a local difference-in-differences strategy. More precisely, we estimate two different models with the general goal of confirming the results obtained with the spatial RDD analysis. Importantly, in both exercises we only consider municipalities close to the border that marks the two policy areas.

As a first analysis, we estimate a standard difference-in-difference model relying on the variation induced by the differential relaxation of the lockdown policy. Therefore, we compare administrative units located in the red and the green area before -in the first round- and after -in the runoff- the modification of the measures.[14] This analysis allows us to include municipal fixed effects in order to capture for municipal-specific and time invariant characteristics. Columns 1-3 of Table 4 show the corresponding results using three different bandwidths -50, 40 and 30 kilometers-: Panel a confirms the positive effect on the vote share to the incumbent, although the effect is not always statistically significant. Further, Panel b reports the same results for the variable voter turnout and we are not able to confirm the positive effect.[15] As a second analysis, we estimate an OLS model using as dependent variable the difference between the value in the runoff and the value in the first round.[16] The corresponding results are shown in columns 4-6 of Table 4, Panel a and b: the positive effect of lockdown measures on incumbent's vote share is confirmed, while a positive, but weak, effect emerges when we consider voter turnout as dependent variable.

Naive spatial regression discontinuity design — Instead of using the difference in outcomes between the first and second round of the municipal elections, we now study the

---

[14]The estimated model is as follows:

$$y_{pir} = \beta Post_r + \gamma Red_{pir} * Post_r + \delta X_{pir} + \epsilon_{pir} \quad (3)$$

where y_{pir} represents the dependent variable, for instance, incumbent's vote share, of polling station p, in municipality i in round r, $Post_r$ indicates the runoff and Red_{pir} is the treatment dummy indicating the municipalities in the red zone. Finally X_{pir} includes municipal fixed effects and a set of controls at the polling station level (*i.e.* number of registered voters and turnout during the first round).

[15]It is important to mention that we obtain similar results if we include polling station fixed effects instead of municipality fixed effects.

[16]The estimated model is as follows:

$$\Delta y_{pi} = \beta Red_{pi} + \gamma X_{pi} + \eta_{pi} \quad (4)$$

where Δy_{pi} is the difference between the value in the runoff and the value in first round ($y_{pi,2} - y_{pi,1}$), Red_{pi} is the treatment dummy indicating the municipalities in the red zone and X_{pir} includes a set of municipal-specific or polling station-specific controls (municipal controls include: distance to the border, latitude/longitude, population, age structure, median income in 2017, characteristics of the mayor -gender, year of birth, job category-; polling station controls include: number of registered voters and turnout during the first round).

Table 4: Local Difference-in-difference analysis

Bandwidth	50	40	30	50	40	30
	DiD			OLS		
	(a) Incumbent's vote share					
Treatment	0.015	2.120*	2.306	1.473	2.943*	3.692***
	(1.967)	(1.059)	(1.794)	(1.245)	(1.506)	(1.137)
N	3290	1602	1078	1645	801	539
R^2	0.599	0.632	0.617	0.322	0.181	0.242
	(b) Turnout					
Treatment	-0.781	-0.351	0.877	2.115***	0.521	0.095
	(1.312)	(0.759)	(0.887)	(0.672)	(0.647)	(0.563)
N	8018	2892	1750	3988	1443	875
R^2	0.920	0.942	0.941	0.331	0.186	0.220
Municipality FE	Yes	Yes	Yes	-	-	-
Municipal controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Unit of analysis is the polling station. The sample includes all French municipalities that experienced a runoff in 2020 municipal election with population higher than 1,000 inhabitants. In Panel A the sample is limited to cities where the incumbent runs in the first round and in the runoff. Standard errors are clustered at the department level.

raw results of the second round around the border. To distinguish it from the spatial RDD specification used above, we label this approach "naive" spatial RDD. As the raw second round results are used, this approach is more prone to violate the smoothness identifying assumption due to differences in fixed municipal characteristics. Nonetheless, it is informative to investigate whether the results in the Table 2 hold under this specification, even if only to know whether fixed municipal characteristics are indeed a concern in the present setting. Table 5 presents the results. Two key results appear. First, the magnitude of the estimated coefficients increase for both incumbent's vote share and voter turnout relative to the results using the difference between the election rounds. Second, the parametric approach of the effect of the lockdown on incumbent's vote share appears much less precisely estimated. This, however, is not the case using the non-parametric approach.

Rolling regressions — To assess the robustness of our results to the exclusion of regional units around the differential lockdown border, we run parametric rolling regressions. We use a 50km bandwidth and various polynomial orders while alternatively excluding all departments around the border. Results are displayed in Appendices C and B. Overall, results appear robust to the exclusion of any department.

Table 5: Naive RD design analysis

Approach	Parametric			Parametric			Non-parametric	
Bandwidth	50	50	50	50	50	50	Opt.	Opt.
Polynomial order	2	3	4	2	3	4	Opt.	Opt.
Covariates	No	No	No	Yes	Yes	Yes	No	Yes
	(a) Incumbent's vote share							
Treatment	-2.794**	0.125	0.678	1.893	-1.424	2.734	17.632***	6.877**
	(1.215)	(1.877)	(2.796)	(2.009)	(2.619)	(3.328)	(4.096)	(3.201)
Adj. R^2	0.02	0.05	0.05	0.09	0.12	0.12	-	-
AIC	13,078	13,034	13,037	12,964	12,913	12,909	-	-
Obs.	1,675	1,675	1,675	1,675	1,675	1,675	2,616	2,616
	(b) Voter turnout							
Treatment	1.157	3.621***	6.046***	5.523***	7.419***	7.138***	6.241***	5.095***
	(0.858)	(1.213)	(1.528)	(1.195)	(1.538)	(1.811)	(1.391)	(1.243)
Adj. R^2	0.08	0.09	0.10	0.23	0.24	0.24	-	-
AIC	28,509	28,497	28,420	27,819	27,743	27,743	-	-
Obs.	4,009	4,009	4,009	4,009	4,009	4,009	5,998	5,998

Notes: Unit of analysis is the polling station.

Additional analysis — To complete our analysis of the lockdown effects on voting outcomes, we present two complementary analysis in Table 9 (Appendix A). We first look at incumbent's vote share and turnout in all municipalities jointly (i.e., above and below 1,000 inhabitants). Estimated coefficient point in the same direction with similar magnitude as in Table 2, however, results appear much less significant. Given the electoral rules in small municipalities (i.e., two rounds block vote with panachage), this result is not surprising as many candidates were already elected to the council in the first round. Second, we focus on electoral outcomes in the yellow zone relative to the red zone. To do so, we consider as treated all municipalities in the yellow zone, and as control all municipalities in the green zone within 50km of the yellow zone. Overall, the lockdown effect in the yellow zone on incumbent's vote share and voter turnout is similar in sign and magnitude to the effect on the red zone. However, the coefficient are less precisely estimated.

5 Concluding remarks

The COVID-19 outbreak represents an unprecedented global crisis which is likely to deeply affect economic, social and political outcomes. Most governments faced the worst phases of the pandemic imposing a set of social containment measures, the lockdown. In many countries different forms of lockdown measures remained in force also after the peak, as the contagion curve showed weak signs of abating. There are also cases in which these restrictions have been reintroduced after an initial relaxation, as was recently the case in Atlanta, Barcelona

or in specific areas of Germany, and we may expect these precautionary measures to be used extensively in the near future in many areas.[17] Nonetheless, despite the relevance of these policies, very little is still known on their impact on voting behaviour.

This paper is among the first to provide a causal evidence of the impact of the introduction of lockdown measures on voting and electoral outcomes. We focus on France which introduced a differential lockdown, dividing the country between a red area, subject to a "hard" lockdown, and a green area, subject to a "soft" one. Moreover, the occurrence of 2020 French municipal elections allows us to measure voting behaviour before and after the implementation of the restriction measures. Therefore, we can exploit this unique setting to estimate a spatial RDD model using as dependent variable the difference in outcomes between the two electoral rounds: what we compare are cities located across the border, subject to different rules, before and after the implementation of the lockdown.

Our findings suggest that the severity of the lockdown significantly affects electoral outcomes and voting behaviour. First, we find that areas subject to the longer lockdown display higher support for the local incumbent, as well as for the Green party. Second, it emerges that municipalities in the red area have a higher level of voter turnout. These results seem to support the "rally around the flag" hypothesis: voters subject to the "hard" lockdown, indeed, either feel safer or are more aware of the health crisis, and therefore grateful for the policy. This reinforces local incumbents and mobilize voters at the polls.

The results of this paper show that voters experiencing containment measures increase their loyalties for political institutions. This legitimates the imposition of such limitations and may have important consequences in terms of political stability in the aftermath of the crisis. In terms of external validity, lockdown policies have been imposed to millions of individuals in many countries but the impact of these restrictions on the approval for governments clearly depends on many local factors such as the management of the health crisis as well as the policies implemented.

---

[17]The city of Atlanta was reopened at the end of May 2020, but a new surge in Coronavirus cases led the administration to reintroduce the lockdown in mid July (https://www.ft.com/content/255b99c1-c08c-3df9-874b-400a8cf27e04). Similar episodes happened in Barcelona (https://www.theguardian.com/world/2020/jul/15/global-report-barcelona-facing-new-lockdown-as-tokyo-raises-alert-level) and in the district of Guetersloh, in the state of North Rhine-Westphalia (https://www.bbc.com/news/world-europe-53217861).

References

Jais Adam-Troian, Eric Bonetto, Florent Varet, Thomas Arciszewski, and Théo Guiller. Pathogen threat increases electoral success for conservative parties: Results from a natural experiment with covid-19 in france. 2020.

Ali T Akarca and Aysit Tansel. Voter reaction to government incompetence and corruption related to the 1999 earthquakes in turkey. *Journal of Economic Studies*, 2016.

Francesc Amat, Andreu Arenas, Albert Falcó-Gimeno, and Jordi Muñoz. Pandemics meet democracy. experimental evidence from the covid-19 crisis in spain. 2020.

Matthew A Baum. The constituent foundations of the rally-round-the-flag phenomenon. *International Studies Quarterly*, 46(2):263–298, 2002.

Alec T Beall, Marlise K Hofer, and Mark Schaller. Infections and elections: Did an ebola outbreak influence the 2014 us federal elections (and if so, how)? *Psychological Science*, 27 (5):595–605, 2016.

Michael M Bechtel and Jens Hainmueller. How lasting is voter gratitude? an analysis of the short-and long-term electoral returns to beneficial policy. *American Journal of Political Science*, 55(4):852–868, 2011.

Damien Bol, Marco Giani, André Blais, and Peter John Loewen. The effect of covid-19 lockdowns on political support: Some good news for democracy? *European Journal of Political Research*, 2020.

Sebastian Calonico, Matias D. Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

Filipe R Campante, Emilio Depetris-Chauvin, and Ruben Durante. The virus of fear: The political impact of ebola in the us. Technical report, National Bureau of Economic Research, 2020.

Virginia A Chanley. Trust in government in the aftermath of 9/11: Determinants and consequences. *Political psychology*, 23(3):469–483, 2002.

Catherine E De Vries, Bert N Bakker, Sara Hobolt, and Kevin Arceneaux. Crisis signaling: How italy's coronavirus lockdown affected incumbent support in other european countries. *Available at SSRN 3606149*, 2020.

CEPR PRESS

Lina M Eriksson. Winds of change: Voter blame and storm gudrun in the 2006 swedish parliamentary election. *Electoral Studies*, 41:129–142, 2016.

Peter Esaiasson, Jacob Sohlberg, Marina Ghersetti, and Bengt Johansson. How the coronavirus crisis affects citizen trust in government institutions and in unknown others–evidence from "the swedish experiment". 2020.

Thiemo Fetzer, Lukas Hensel, Johannes Hermle, and Christopher Roth. Coronavirus perceptions and economic anxiety. *Review of Economics and Statistics*, pages 1–36, 2020.

Zack P Grant and James Tilley. Fertile soil: explaining variation in the success of green parties. *West European Politics*, 42(3):495–516, 2019.

Allison Harell. How canada's pandemic is shifting political views. report for the institute for research on public policy, 2020.

Andrew Healy and Neil Malhotra. Myopic voters and natural disaster policy. *American Political Science Review*, pages 387–406, 2009.

David S. Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, 2010.

Arndt Leininger and Max Schaub. Voting at the dawn of a global pandemic. 2020.

Elisa M Maffioli. The political economy of health epidemics: evidence from the ebola outbreak. *Available at SSRN 3383187*, 2018.

Hani Mansour, Daniel I Rees, and James M Reeves. Voting and political participation in the aftermath of the hiv/aids epidemic. Technical report, National Bureau of Economic Research, 2020.

Eric Merkley, Aengus Bridgman, Peter John Loewen, Taylor Owen, Derek Ruths, and Oleg Zhilin. A rare moment of cross-partisan consensus: Elite and public response to the covid-19 pandemic in canada. *Canadian Journal of Political Science/Revue canadienne de science politique*, pages 1–8, 2020.

John E Mueller. Presidential popularity from truman to johnson. *The American Political Science Review*, 64(1):18–34, 1970.

Chris G Sibley, Lara M Greaves, Nicole Satherley, Marc S Wilson, Nickola C Overall, Carol HJ Lee, Petar Milojev, Joseph Bulbulia, Danny Osborne, Taciano L Milfont, et al. Effects of

the covid-19 pandemic and nationwide lockdown on trust, attitudes toward government, and well-being. *American Psychologist*, 2020.

Jean-David Zeitoun, Matthieu Faron, Jerome Fourquet, Marc Lavielle, et al. Reciprocal association between participation to a national election and the epidemic spread of covid-19 in france: nationwide observational and dynamic modeling study. *medRxiv*, 2020.

A Appendix: Supporting material

Figure 5: Differential lockdown from June 2nd to June 22nd, 2020
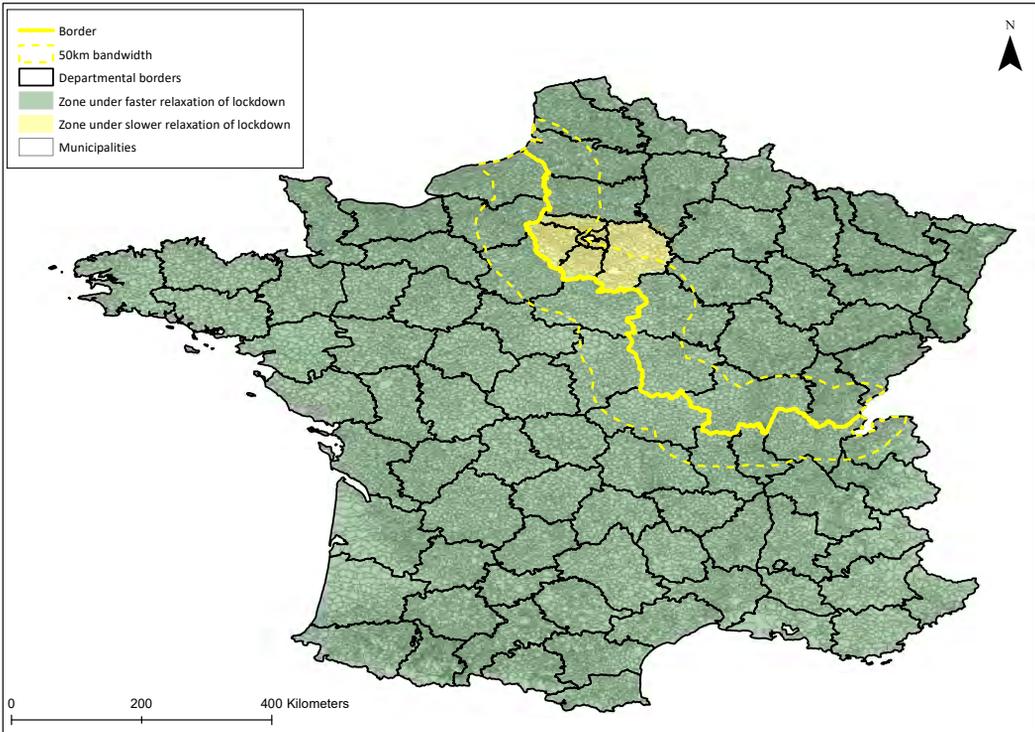
Table 6: Descriptive statistics

	Average value	Standard deviation	N
Panel A: Electoral outcomes (polling station)			
Incumbent's vote share (diff)	6.762	9.581	1,737
Turnout (diff)	-2.892	6.334	4,462
Victory margin (diff)	6.436	12.062	4,376
White votes (diff)	2.446	6.287	4,462
Number candidates (diff)	-4.417	3.309	4,462
Panel B: Controls variables (municipality)			
Population (0-14)	69,713.4	117,554.6	4,525
Population (15-29)	114,806.3	197,925.7	4,525
Population (30-44)	107,374.4	188,941.2	4,525
Population (45-59)	87,667.5	154,789.4	4,525
Population (60-74)	66,597.3	118,610.3	4,525
Population (75-89)	31,981.1	55,902.7	4,525
Population (90+)	5,382.5	9,437.7	4,525
Female population	255,901.8	446,613.9	4,525
Pop. growth rate (2011-2016)	1.640	5.989	4,525
Median income (2017)	23,378.4	4,452.2	4,480

Notes: The sample under consideration includes municipalities with population higher or equal than 1,000 inhabitants that have a runoff in 2020 French municipal elections. The variables in Panel A are expressed in difference between the two electoral round. The variable *Incumbent's vote share (diff)* is only defined for municipalities in which the incumbent runs in both rounds. Population variables are defined according to 2016 population census.

Table 7: Spatial RD design analysis of turnout

Approach	Parametric								
Bandwidth	60	60	60	50	50	50	40	40	40
Poly. order	2	3	4	2	3	4	2	3	4
	(a) Incumbent's vote share								
Treatment	3.057***	1.541	3.658**	3.303***	3.534**	4.836**	2.436	4.597**	1.414
	(0.987)	(1.161)	(1.574)	(1.209)	(1.439)	(1.888)	(1.590)	(1.877)	(2.215)
Adj. R^2	0.21	0.23	0.24	0.22	0.29	0.29	0.11	0.12	0.14
AIC	13,156	13,099	13,081	10,742	10,603	10,588	5,195	5,187	5,176
Obs.	2,019	2,019	2,019	1,645	1,645	1,645	801	801	801
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	(b) Voter turnout								
Treatment	1.315***	1.999***	2.505***	1.990***	2.919***	2.343***	1.500**	-0.078	0.033
	(0.453)	(0.636)	(0.763)	(0.543)	(0.715)	(0.829)	(0.682)	(0.860)	(0.954)
Adj. R^2	0.22	0.22	0.22	0.26	0.26	0.26	0.09	0.10	0.10
AIC	27,389	27,375	27,375	21,664	21,658	21,657	7,900	7,880	7,884
Obs.	4,927	4,927	4,927	4,009	4,009	4,009	1,446	1,446	1,446
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Unit of analysis is the polling station.

Table 8: Smoothness test for presence in second round by party

Bin size	1km			2km		
Political party	Left-wing	Greens	Right-wing	Left-wing	Greens	Right-wing
RD Estimate	14.541	-0.788	5.057	20.049*	-3.728	4.425
	(8.906)	(39.632)	(7.924)	(10.715)	(54.511)	(12.023)
Obs.	53	28	61	36	22	42

Notes: The unit of analysis are bands of 1km or 2km width within 50km of the border on either side.

Table 9: Additional anylsis

Approach	Parametric			Parametric			Non-parametric	
Bandwidth	50	50	50	50	50	50	Opt.	Opt.
Polynomial order	2	3	4	2	3	4	Opt.	Opt.
Covariates	No	No	No	Yes	Yes	Yes	No	Yes
Spatial RD design analysis in all municipalities								
(a) Incumbent's vote share								
Treatment	2.563**	5.061***	-1.411	0.349	-0.232	0.621	2.936	0.793
	(1.231)	(1.790)	(2.492)	(2.005)	(2.313)	(2.734)	(2.923)	(2.826)
Adj. R^2	0.05	0.07	0.09	0.19	0.22	0.23	-	-
AIC	12,697	12,665	12,622	12,353	12,288	12,268	-	-
Obs.	1,737	1,737	1,737	1,731	1,731	1,731	2,794	2,783
(b) Voter turnout								
Treatment	1.884***	1.187	-0.468	-1.527*	-1.563	-1.782	1.980	1.137
	(0.649)	(0.949)	(1.240)	(0.890)	(1.138)	(1.329)	(1.516)	(1.427)
Adj. R^2	0.01	0.02	0.02	0.13	0.13	0.13	-	-
AIC	29,105	29,062	29,044	28,204	28,197	28,198	-	-
Obs.	4,462	4,462	4,462	4,432	4,432	4,432	7,004	6,918
Spatial RD design analysis in the yellow zone								
(a) Incumbent's vote share								
Treatment	1.075	6.308***	-0.176	0.863	1.679	2.800*	7.614***	27.001***
	(0.703)	(0.941)	(1.185)	(1.088)	(1.220)	(1.501)	(2.539)	(5.410)
Adj. R^2	0.06	0.11	0.14	0.27	0.33	0.34	-	-
AIC	9,639	9,553	9,516	9,272	9,159	9,138	-	-
Obs.	1,444	1,444	1,444	1,444	1,444	1,444	2,268	2,268
(b) Voter turnout								
Treatment	0.924**	0.884	-1.594**	1.989***	2.716***	1.263	0.370	0.486
	(0.379)	(0.549)	(0.653)	(0.617)	(0.771)	(0.880)	(1.615)	(1.564)
Adj. R^2	0.09	0.09	0.10	0.27	0.27	0.28	-	-
AIC	20,405	20,408	20,363	19,583	19,583	19,575	-	-
Obs.	3,657	3,657	3,657	3,657	3,657	3,657	5,153	5,153

Notes: Unit of analysis is the polling station.

B Appendix: Rolling regressions on incumbent's vote share

Parametric regression on incumbent's vote share (1)

	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth	50	50	50	50	50	50
Polynomial order	2	3	4	2	3	4
Covariates	No	No	No	Yes	Yes	Yes
Excluded department number						
Number: 1	1.653**	5.903***	1.149	3.568***	3.311**	4.291**
	(0.739)	(1.120)	(1.584)	(1.248)	(1.464)	(1.882)
Number: 3	1.096	5.057***	0.625	3.059**	2.259	3.532*
	(0.754)	(1.135)	(1.620)	(1.206)	(1.434)	(1.848)
Number: 14	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 18	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 23	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 26	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 27	3.306***	7.695***	0.970	5.072***	5.142***	5.892***
	(0.789)	(1.183)	(1.723)	(1.323)	(1.518)	(2.020)
Number: 28	0.667	5.287***	1.949	2.599**	3.061**	5.413***
	(0.708)	(1.115)	(1.607)	(1.187)	(1.439)	(1.924)
Number: 36	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 38	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 39	1.456**	6.098***	1.502	3.045**	3.834**	4.875**
	(0.729)	(1.171)	(1.742)	(1.349)	(1.642)	(2.162)

Parametric regression on incumbent's vote share (2)

	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth	50	50	50	50	50	50
Polynomial order	2	3	4	2	3	4
Covariates	No	No	No	Yes	Yes	Yes
Excluded department number						
Number: 41	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 42	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 43	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 45	1.058	6.817***	1.927	4.666***	4.358***	5.028***
	(0.745)	(1.152)	(1.617)	(1.330)	(1.501)	(1.948)
Number: 58	0.967	5.290***	-0.505	1.918**	2.727***	3.126***
	(0.644)	(0.838)	(1.032)	(0.892)	(0.968)	(1.202)
Number: 60	1.740**	6.023***	1.706	3.518**	3.786**	5.026**
	(0.721)	(1.119)	(1.628)	(1.407)	(1.659)	(2.066)
Number: 61	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 62	1.606**	5.894***	1.546	3.050**	3.269**	4.722**
	(0.717)	(1.107)	(1.597)	(1.220)	(1.465)	(1.894)
Number: 63	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 69	1.660**	4.548***	2.076	1.866	4.330***	4.400**
	(0.737)	(1.147)	(1.624)	(1.177)	(1.447)	(1.905)
Number: 71	1.605**	6.049***	1.734	3.501***	3.580**	4.980***
	(0.725)	(1.119)	(1.613)	(1.287)	(1.475)	(1.917)
Number: 72	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 73	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)

Parametric regression on incumbent's vote share (3)

	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth	50	50	50	50	50	50
Polynomial order	2	3	4	2	3	4
Covariates	No	No	No	Yes	Yes	Yes
Excluded department number						
Number: 74	1.217*	4.919***	3.140*	3.820***	3.583**	5.182***
	(0.723)	(1.109)	(1.625)	(1.271)	(1.430)	(1.884)
Number: 75	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 76	2.213***	6.646***	1.297	4.022***	3.580**	5.670***
	(0.744)	(1.147)	(1.678)	(1.338)	(1.601)	(2.125)
Number: 77	1.345*	5.566***	1.538	2.900**	3.127**	4.988***
	(0.759)	(1.161)	(1.658)	(1.250)	(1.482)	(1.932)
Number: 78	1.891**	6.422***	1.294	3.496***	3.718**	4.475**
	(0.820)	(1.354)	(1.862)	(1.333)	(1.629)	(2.094)
Number: 80	1.630**	6.047***	1.647	3.374***	3.655**	4.950**
	(0.722)	(1.123)	(1.628)	(1.225)	(1.465)	(1.924)
Number: 89	1.397*	5.844***	1.661	3.642***	3.464**	5.123***
	(0.745)	(1.165)	(1.642)	(1.260)	(1.510)	(1.956)
Number: 91	1.763**	6.994***	2.647	3.320***	3.755**	5.339**
	(0.779)	(1.260)	(1.946)	(1.199)	(1.514)	(2.141)
Number: 92	1.968***	5.247***	1.791	3.167***	2.793*	4.701**
	(0.724)	(1.126)	(1.608)	(1.219)	(1.479)	(1.893)
Number: 93	1.585**	5.904***	1.543	3.303***	3.534**	4.836**
	(0.716)	(1.106)	(1.597)	(1.209)	(1.439)	(1.888)
Number: 94	2.616***	4.878***	2.395	3.527***	3.676**	5.032***
	(0.718)	(1.127)	(1.620)	(1.243)	(1.447)	(1.925)
Number: 95	0.903	6.234***	1.373	2.590**	3.895***	4.747**
	(0.737)	(1.116)	(1.622)	(1.238)	(1.500)	(1.922)

C   Appendix: Parametric rolling regressions on voter turnout

Parametric regression on voter turnout (1)

	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth	50	50	50	50	50	50
Polynomial order	2	3	4	2	3	4
Covariates	No	No	No	Yes	Yes	Yes
Excluded department number						
Number: 1	1.300***	1.471**	0.492	2.018***	2.970***	2.393***
	(0.390)	(0.576)	(0.746)	(0.553)	(0.731)	(0.850)
Number: 3	1.299***	1.435**	0.347	1.880***	2.713***	2.097**
	(0.394)	(0.584)	(0.744)	(0.548)	(0.731)	(0.842)
Number: 14	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 18	1.413***	0.987*	-0.045	1.344**	1.771**	1.373*
	(0.361)	(0.523)	(0.648)	(0.544)	(0.709)	(0.783)
Number: 23	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 26	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 27	0.476	0.645	0.728	1.274**	2.231***	2.688***
	(0.401)	(0.590)	(0.787)	(0.560)	(0.749)	(0.895)
Number: 28	1.399***	1.619***	0.517	2.090***	3.085***	2.449***
	(0.404)	(0.597)	(0.795)	(0.571)	(0.764)	(0.904)
Number: 36	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 38	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 39	1.189***	1.187**	0.001	1.719***	2.539***	1.869**
	(0.373)	(0.552)	(0.720)	(0.558)	(0.735)	(0.845)

Parametric regression on voter turnout (2)

	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth	50	50	50	50	50	50
Polynomial order	2	3	4	2	3	4
Covariates	No	No	No	Yes	Yes	Yes
Excluded department number						
Number: 41	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 42	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 43	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 45	1.591***	1.789***	1.177	2.308***	3.447***	3.297***
	(0.407)	(0.611)	(0.800)	(0.598)	(0.821)	(0.940)
Number: 58	1.450***	1.778***	0.730	2.343***	3.596***	3.125***
	(0.370)	(0.540)	(0.704)	(0.534)	(0.699)	(0.809)
Number: 60	1.297***	1.419**	0.440	2.443***	3.435***	2.854***
	(0.373)	(0.553)	(0.721)	(0.575)	(0.762)	(0.865)
Number: 61	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 62	1.336***	1.387**	0.440	1.714***	2.560***	2.135***
	(0.372)	(0.550)	(0.716)	(0.535)	(0.706)	(0.827)
Number: 63	1.291***	1.437***	0.410	1.984***	2.917***	2.359***
	(0.373)	(0.551)	(0.716)	(0.543)	(0.715)	(0.830)
Number: 69	1.338***	1.806***	-0.367	1.858***	2.603***	1.453*
	(0.393)	(0.586)	(0.755)	(0.561)	(0.717)	(0.869)
Number: 71	1.258***	1.483***	0.345	1.968***	2.979***	2.390***
	(0.377)	(0.564)	(0.742)	(0.564)	(0.750)	(0.872)
Number: 72	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)
Number: 73	1.301***	1.438***	0.393	1.990***	2.919***	2.343***
	(0.373)	(0.550)	(0.716)	(0.543)	(0.715)	(0.829)

Parametric regression on voter turnout (3)

	(1)	(2)	(3)	(4)	(5)	(6)
Bandwidth	50	50	50	50	50	50
Polynomial order	2	3	4	2	3	4
Covariates	No	No	No	Yes	Yes	Yes
Excluded department number						
Number: 74	1.427***	1.499***	0.280	2.369***	3.032***	2.044**
	(0.387)	(0.554)	(0.727)	(0.626)	(0.741)	(0.846)
Number: 75	0.518	1.212**	1.231*	1.966***	2.985***	2.302***
	(0.375)	(0.548)	(0.716)	(0.544)	(0.718)	(0.830)
Number: 76	1.414***	1.925***	0.126	1.343**	2.640***	1.095
	(0.370)	(0.556)	(0.744)	(0.534)	(0.709)	(0.851)
Number: 77	1.265***	1.278**	0.263	2.132***	2.819***	2.309***
	(0.378)	(0.557)	(0.724)	(0.551)	(0.721)	(0.838)
Number: 78	1.825***	2.546***	1.084	2.181***	3.572***	2.991***
	(0.405)	(0.620)	(0.797)	(0.580)	(0.773)	(0.903)
Number: 80	1.260***	1.371**	0.272	2.144***	2.860***	2.285***
	(0.378)	(0.557)	(0.720)	(0.579)	(0.731)	(0.843)
Number: 89	1.198***	1.340**	-0.197	1.828***	2.697***	1.729**
	(0.374)	(0.553)	(0.709)	(0.548)	(0.725)	(0.830)
Number: 91	0.711*	1.441**	0.290	1.969***	3.243***	2.631***
	(0.389)	(0.589)	(0.766)	(0.546)	(0.740)	(0.865)
Number: 92	1.392***	1.315**	0.427	1.659***	2.305***	2.262***
	(0.374)	(0.554)	(0.718)	(0.540)	(0.716)	(0.829)
Number: 93	1.332***	1.385**	0.444	2.037***	2.897***	2.376***
	(0.372)	(0.550)	(0.716)	(0.542)	(0.715)	(0.829)
Number: 94	1.451***	1.356**	0.277	1.975***	3.452***	2.134**
	(0.375)	(0.557)	(0.724)	(0.556)	(0.722)	(0.830)
Number: 95	1.771***	1.200**	0.528	2.582***	2.993***	2.465***
	(0.376)	(0.550)	(0.717)	(0.549)	(0.731)	(0.843)

The stochastic reproduction rate of a virus[1]

Richard Holden[2] and D.J. Thornton[3]

We consider an SIR model where the probability of infections between infected and susceptible individuals are viewed as Poisson trials. The probabilities of infection between pairwise susceptible-infected matches are thus order statistics. This implies that the reproduction rate is a random variable. We derive the first two moments of the distribution of R_t conditional on the information available at time $t-1$ for Poisson trials drawn from an arbitrary parent distribution with finite mean. We show that the variance of R_t is increasing in the proportion of susceptible individuals in the population, and that ex ante identical populations can exhibit large differences in the path of the virus. This has a number of implications for policy during pandemics. We provide a rationale for why shelter-in-place orders may be a better containment measure than mandating the use of masks because of their impact on the variance of the reproduction rate.

1 Introduction

The COVID-19 pandemic has shown that the effective reproduction rate of the virus $\mathcal{R}_t$ is a crucial determinant not only if public health, but also of public policy. Social distancing, shelter-in-place and other containment measures aim to stop the spread of the virus, essentially by attempting to push $\mathcal{R}_t$ below 1.

It has been clear to epidemiologists, public health officials, and economists that with $\mathcal{R}_t > 1$ the virus spreads exponentially—overwhelming health systems and leading to substantial loss of life.

Despite the critical importance of the reproduction rate it is treated in epidemiological models as a parameter, rather than as a random variable. Yet the primitives of so-called SIR models are, from a behavioral perspective, the proportion of the population that is infected and susceptible, the rate at which those sub-populations interact, and the probability that a given interaction leads an infected individual to infect a susceptible individual.

Those primitives *determine* the reproduction rate, yet $\mathcal{R}_t$ is typically modeled as an exogenous variable. The point-of-view we take in this paper is to derive $\mathcal{R}_t$ from primitives and explicitly model the stochastic nature of those primitives. This leads one to the seemingly obvious conclusion that $\mathcal{R}_t$ is a random variable—and as such it has an associated probability distribution. An immediate consequence of this is that the rate of susceptibility, infection, and recovery in a population are also random variables with their own probability distributions.

The first main result we derive is that the variance of $\mathcal{R}_t$ at any point in time, conditional on the information available that time, is increasing in the proportion of the population that is susceptible at that time. This implies that early on in an epidemic–when the susceptible proportion of the population is low–the variance of infection rates is the highest. This, in turn, implies that two *ex ante* identical jurisdictions could have very different infection rates simply due to chance, and this can create persistent difference in the path of the virus in those jurisdictions.

An optimal policy for controlling $\mathcal{R}_t$ must account for the complete distribution of $\mathcal{R}_t$ and the amount of mass of its probability distribution that lies on the part of the support above $\mathcal{R}_t = 1$. All else equal the amount of mass in this "danger zone" is larger when the proportion of the population that is susceptible is larger—i.e. at the onset of the virus. This implies that the optimal containment policy is stricter earlier in the evolution of the virus.

The remainder of the paper proceeds as follows. In section 2 we outline the statement

of the problem. Section 3 characterizes the first two moments of the distribution of $\mathcal{R}_t$. Section 4 derives a number of implications for decision-makers and concludes. Proofs of results not contained in the text are relegated to an appendix.

2 The Model

2.1 Statement of the Problem

There are N agents that are infinitely lived. Time is discrete and indexed by $t = 0, 1, 2, \ldots$. At each time t an agent is in one of three states: *susceptible*, *infected*, or *recovered*. There are S_t susceptible agents in the population, there are I_t infected agents, and the number of recovered individuals is $R_t = N - S_t - I_t$, where $S_t, I_t, R_t \in \mathbb{N}$.

Following the standard convention, we shall refer to I_t as the *prevalence* of the disease at time t.

At time t each agent is pairwise random matched with $\phi = \phi(N)$ other agents *independently* with equal probability among agents irrespective of health status, where $\phi(N)$ captures the intensity of social interactions in the population. We assume that $\phi \colon \mathbb{N} \to \mathbb{N}$ is an integer valued function satisfying $0 \leq \phi(N) \leq N$ for all $N \in \mathbb{N}$. This allows us to consider both the simple case where $\phi(N) = d \in \mathbb{N}$ for all N, and more complex cases where for example $\phi(N) = \lfloor \log(N) \rfloor$ (here, $\lfloor \cdot \rfloor$ denotes the *floor* function, though we routinely omit floor and ceiling signs for readability). For notational convenience later on we also define here $m = \phi/N$ to be the proportion of the population that an individual matches up with in a given time period.

An interaction structure in the population can be formed by considering the *random regular graph* formed by the well known *configuration model* [6, 50] with N individuals or vertices and degree sequence $(d_1, \ldots d_N) = (\phi, \ldots, \phi)$, conditioning on getting a simple graph. In the limit of a large population size, and for small $\phi(N)$, matches formed by the configuration model are approximately independent (this technique is used in [19] for example). Some notes on the reproductive rate with the independence condition removed are provided in the appendix.

If an infected agent j is matched with a susceptible agent k then with probability p_j the susceptible agent becomes infected. We assume that p_j is drawn from a distribution $\mathcal{D}$ on support $[0, 1]$ with finite mean μ, and that the draw for each infected-susceptible match is stochastically independent.

Recovered agents are assumed to be immune to infection and are not contagious, and infected agents spontaneously recover with probability $\gamma \geq 0$.

2.2 Discussion

This paper contributes to the literature on heterogeneity (or 'heterogeneous infection rates' if you prefer) in epidemic models, much in the spirit of the now classic [22]. In the existing literature, studies have focused on heterogeneity at the group level [29, 9, 3, 14, 8, 45, 44] (subdividing a heterogeneous population into homogeneous groups) as well as individual level [28, 42, 33, 23], whilst other authors have taken the approach of incorporating exogenous shocks into their models. Such shock are usually modeled by introducing discontinuities into the deterministic SIR (as well as SEIR, SIS, and SI) systems of differential equations via stochastic jumps or 'noise' [46, 48, 53, 54, 35]. Several studies built on variants of the SIR model have been applied directly to analysing the Covid-19 epidemic [52, 44, 40, 17], with some recent models taking into account the economic costs of various policy interventions (e.g. [1, 5, 11, 17, 13]). A significant amount of work has been done to describe the behavior of epidemics on networks with heterogeneous individuals [34, 30, 4, 31, 20, 47, 51, 41, 7, 16, 18], often making use of results from bond percolation or applying mean field approximations (e.g. see [36]).

Models which account for individual heterogeneity have become especially important in light of the accumulating theoretical [29, 27, 25, 33, 7] and empirical [10, 49, 15, 21, 4, 32] evidence that heterogeneity in a population strongly influences (and in particular, lowers) the probability of epidemic invasion.

Most of the models in the literature–with the exception of a few [8, 41, 23, 7] involve continuous time. In the way of discrete SIR models some work has been done (see [2] for example), though the disease-spreading mechanism that we employ has, as far as we are aware, not been considered in the epidemiological literature. Moreover, whilst many papers calculate the basic reproductive number R_0 under their models, this parameter –and consequently, the effective reproductive number $\mathcal{R}_t$ –often fail to accurately capture the threshold properties of epidemics, and can vary dramatically according to the method by which they are calculated [8, 24, 38]. For example, the observations of of super-spreading events (SSE's) which is well documented for the SARS outbreak of 2003 [39, 26, 43, 28, 12] represent tail end draws from the (random) reproductive rate. Our main innovation is the explicit focus on the reproductive rate as a random variable, and our novel modeling approach of infected-susceptible matches as non-identical Bernoulli "trials"–or *Poisson trials* [37]. The paper perhaps closest to ours is [28].

Before we proceed with the analysis we pause briefly to remark that our main result about the variance of the reproduction rate at a particular point in time is stated as being conditional on the information available to an observer at that time. There are two reasons for this. The first is conceptual. Any policy maker should care about the variance at a

given time accounting for the information that she has. Calculating the unconditional variance is to discard useful information. The second reason we focus on the conditional variance is practical. One can analytically derive the unconditional variance, but it has a somewhat unwieldy recursive structure. In any case, we view it as being of limited interest for policy makers concerned with the evolution of a virus.

3   Analysis

3.1   The Distribution of the Reproduction Rate

We seek to characterize the distribution of the reproduction rate of the virus at time t, which we denote $\mathcal{R}_t$.

Notice that a given infected agent has ϕ pairwise random matches in the population and that infections can only occur when they are matched with a susceptible agent (which have proportion S_t/N in the population). Assuming there is independence between matches (more on this assumption later), the expected number of susceptible matches of a given infected individual is therefore $\phi S_t/N$. Recalling that we defined $m = \phi/N$ we can write this as mS_t

Denote the probability that given such a match infection occurs as p_j. We define $\mathcal{R}_t$ as the random variable counting the number of people that are infected by a given infected person at time t. That is, we have

$$\mathcal{R}_t = \sum_{j=1}^{mS_t} \mathbb{I}_j, \tag{1}$$

where $\mathbb{I}_j$ is an indicator variable taking the value 1 if match j leads to an infection and 0 otherwise. That is, we define $\mathbb{I}_j$ by the conditional distribution

$$\mathbb{P}(\mathbb{I}_j = x \mid p_j = p) = xp + (1 - x)(1 - p).$$

Recalling that the mean of p_j is finite and denoted by μ, we may calculate the expectation of the indicator variable $\mathbb{I}_j$ by the *law of iterated expectations* as

$$E(\mathbb{I}_j) = E\big(E(\mathbb{I}_j \mid p_j)\big) = E\big(\mathbb{P}(\mathbb{I}_j = 1 \mid p_j)\big) = \mathbb{E}(p_j) = \mu. \tag{2}$$

We also ignore potential integer problems in the upper limit of the summation in the definition of $\mathcal{R}_t$, one could avoid this by explicitly taking $\lfloor mS_t \rfloor$ as the upper limit.

CEPR PRESS

Defining $\mathcal{R}_t$ in this way gives us something analogous to a discrete-time SIR model which we will discuss how to specify shortly. By way of reference, we describe the classic discrete-time SIR model here. Let $\Delta S_{t+1} = S_{t+1} - S_t$ and define ΔI_{t+1} and ΔR_{t+1} similarly. The classic discrete-time SIR model can be described as

$$\Delta S_{t+1} = -\beta I_t S_t \tag{3}$$

$$\Delta I_{t+1} = \beta I_t S_t - \gamma I_t \tag{4}$$

$$\Delta R_{t+1} = I_t \gamma, \tag{5}$$

where β is the average rate of infection. Our model is substantially more complex than this since we account for the distribution of $\mathcal{R}_t$ and not just the point estimate β. As such, our rates of susceptibility, infection, and recover will also be random variables determined by some initial state vector (S_0, I_0, R_0).

Let $\mathcal{R}_t^{pop}$ denote the population-level reproductive rate, that is, a random variable whose distribution is equal to the number of new infections in the population from time t to time $t+1$. We saw that in the calculation of $\mathcal{R}_t$, the expected number of susceptible matches of a given infected individual is mS_t. Since there are I_t infected individuals in the population at time t, and each infected-susceptible match is assumed to be independent, we have

$$\mathcal{R}_t^{pop} = \sum_{j=1}^{I_t m S_t} \mathbb{I}_j. \tag{6}$$

Taking $\phi(N) = \phi \in \mathbb{N}$ is sufficient for the independence properties we require, henceforth we assume $\phi \in \mathbb{N}$ is a constant. As a result, our SIR equations become

$$\Delta S_{t+1} = -\sum_{j=1}^{m I_t S_t} \mathbb{I}_j. \tag{7}$$

$$\Delta I_{t+1} = \sum_{j=1}^{m I_t S_t} \mathbb{I}_j - \gamma I_t \tag{8}$$

$$\Delta R_{t+1} = \gamma I_t. \tag{9}$$

We suppress the dependence of the indicator variables on the time period t for notational convenience, in reality each new time period yields a new set of indicator variables $\mathbb{I}_j$ which are independent of those in any other time period. Notice by Wald's identity we have

$$\mathbb{E}\left(\mathcal{R}_t^{pop}\right) = \mathbb{E}\left(\sum_{j=1}^{m I_t S_t} \mathbb{I}_j\right) = \mathbb{E}(mS_t I_t)\mathbb{E}(\mathbb{I}_j) = m\mu\mathbb{E}(S_t I_t),$$

and hence

$$\mathbb{E}(\Delta S_{t+1}) = -m\mu\mathbb{E}(S_t I_t).$$
$$\mathbb{E}(\Delta I_{t+1}) = m\mu\mathbb{E}(S_t I_t) + \gamma\mathbb{E}(I_t)$$
$$\mathbb{E}(\Delta R_{t+1}) = \gamma\mathbb{E}(I_t).$$

That is, our model coincides with the standard model in expectation, with average infection rate $\beta = m\mu$.

Consider a policy maker who, at time t, knows the current state $\theta_{t-1} = \{S_{t-1}, I_{t-1}\}$ of susceptibility and infection in the population at all time periods before t. They will use this information to inform their policy. The key parameters of interest therefore are the mean and variance of $\mathcal{R}_t$, conditioned on knowing the history θ_{t-1}. That is, we want to find expressions for $\mathbb{E}(\mathcal{R}_t \mid \theta_{t-1})$ and $\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1})$.

We begin by calculating $\mathbb{E}(\mathcal{R}_t \mid \theta_{t-1})$. By noting that S_t is independent of the infection event captured by $\mathbb{I}_j$ for all $t, j \in \mathbb{N}$, we can calculate the conditional expectation of $\mathcal{R}_t$ by *Wald's identity* and (2) as

$$\begin{aligned} \mathbb{E}(\mathcal{R}_t \mid \theta_{t-1}) &= \mathbb{E}(mS_t \mid \theta_{t-1})E(\mathbb{I}_j \mid \theta_{t-1}) \\ &= m\mathbb{E}(S_t \mid \theta_{t-1})E(\mathbb{I}_j) \\ &= m\mu\mathbb{E}(S_t \mid \theta_{t-1}). \end{aligned} \tag{10}$$

Therefore, the mean of $\mathcal{R}_t$ depends crucially on the mean of S_t, the number of susceptible individuals in the population at time t. We now derive the mean and variance of S_t conditional on knowing the history of susceptibility and infection. This will allow us to derive the conditional mean and variance of $\mathcal{R}_t$. We work towards the theorem below in what follows.

Theorem 1 *For all $t \in \mathbb{N}$*

    1.

$$\mathbb{E}(\mathcal{R}_t \mid \theta_{t-1}) = m\mu S_{t-1} + m^2\mu^2 I_{t-1}S_{t-1}.$$

    2.

$$\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1}) = m\mu(1-\mu)S_{t-1} + m^2\mu^2(1-\mu)(1+m\mu)I_{t-1}S_{t-1}.$$

We use the next result several times in the proofs that follow, so we state it here as a lemma.

Lemma 1 *For all $t \in \mathbb{N}$, we have*

$$\mathbb{E}(\Delta S_t \mid \theta_{t-1}) = -m\mu I_{t-1}S_{t-1}.$$

Proof. The result can be obtained by a single application of Wald's identity. We thus have

$$\begin{aligned}
\mathbb{E}(\Delta S_t \mid \theta_{t-1}) &= \mathbb{E}\left(-\sum_{j=1}^{mI_{t-1}S_{t-1}} \mathbb{I}_j \;\middle|\; \theta_{t-1} \right) \\
&= -\mathbb{E}(mI_{t-1}S_{t-1} \mid \theta_{t-1})\mathbb{E}(\mathbb{I}_j \mid \theta_{t-1}) \\
&= -m\mu I_{t-1}S_{t-1},
\end{aligned}$$

as desired. ∎

We now prove Part 1. of Theorem 1.

Proof of Part 1. We can compute $\mathbb{E}(S_t \mid \theta_{t-1})$ directly by noting that $S_t = S_{t-1} + \Delta S_t$. This gives us that

$$\begin{aligned}
\mathbb{E}(S_t \mid \theta_{t-1}) &= \mathbb{E}\left(S_{t-1} + \Delta S_t \mid \theta_{t-1} \right) \\
&= S_{t-1} + \mathbb{E}\left(\Delta S_t \mid \theta_{t-1} \right).
\end{aligned}$$

Now applying Lemma 1, we have

$$\mathbb{E}(S_t \mid \theta_{t-1}) = S_{t-1} - m\mu I_{t-1}S_{t-1}. \tag{11}$$

Applying (10) yields the desired result. ∎

We will need to establish a few more lemmas in order to prove Part 2. of Theorem 1. We being by proving an equation for the conditional variance of $\mathcal{R}_t$ in terms of the conditional expectation and variance of S_t.

Lemma 2 *For all $t \in \mathbb{N}$, we have*

$$\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1}) = m\mu(1 - \mu)\mathbb{E}(S_t \mid \theta_{t-1}) + m^2\mu^2 \, \mathrm{Var}(S_t \mid \theta_{t-1}). \tag{12}$$

Proof. First, we have

$$
\begin{aligned}
\mathcal{R}_t^2 &= \left(\sum_{j=1}^{mS_t} \mathbb{I}_j \right) \left(\sum_{l=1}^{mS_t} \mathbb{I}_l \right) \\
&= \sum_{j=l} \mathbb{I}_j \mathbb{I}_l + \sum_{j \neq l} \mathbb{I}_j \mathbb{I}_l. \\
&= \sum_{j=l} \mathbb{I}_j + \sum_{j \neq l} \mathbb{I}_j \mathbb{I}_l.
\end{aligned}
$$

There are $(mS_t)^2$ terms in the expansion of $\mathcal{R}_t^2$. Exactly mS_t of these fall under the sum where $j = l$, and the remaining $(mS_t)^2 - mS_t$ of them fall under the sum where $j \neq l$. Hence by Wald's identity and (2), we have

$$
\begin{aligned}
\mathbb{E}\left(\mathcal{R}_t^2 \mid \theta_{t-1}\right) &= \mathbb{E}\left(\sum_{j=l} \mathbb{I}_j \mid \theta_{t-1} \right) + \mathbb{E}\left(\sum_{j \neq l} \mathbb{I}_j \mathbb{I}_l \mid \theta_{t-1} \right) \\
&= \mathbb{E}(mS_t \mid \theta_{t-1})\mathbb{E}(\mathbb{I}_j \mid \theta_{t-1}) + \mathbb{E}\left((mS_t)^2 - mS_t \mid \theta_{t-1}\right)\mathbb{E}(\mathbb{I}_j \mathbb{I}_l \mid \theta_{t-1}) \\
&= \mu \mathbb{E}(mS_t \mid \theta_{t-1}) + \mu^2 \mathbb{E}\left((mS_t)^2 - mS_t \mid \theta_{t-1}\right) \\
&= m\mu(1-\mu)\mathbb{E}(S_t \mid \theta_{t-1}) + m^2\mu^2\mathbb{E}(S_t^2 \mid \theta_{t-1}). \tag{13}
\end{aligned}
$$

Where we have used independence of indicator variables on the second sum. Hence utilizing (10) and (13), the variance of $\mathcal{R}_t$ is given by

$$
\begin{aligned}
\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1}) &= \mathbb{E}(\mathcal{R}_t^2 \mid \theta_{t-1}) - (\mathbb{E}(\mathcal{R}_t \mid \theta_{t-1}))^2 \\
&= m\mu(1-\mu)\mathbb{E}(S_t \mid \theta_{t-1}) + m^2\mu^2\mathbb{E}(S_t^2 \mid \theta_{t-1}) - (m\mu\mathbb{E}(S_t \mid \theta_{t-1}))^2 \\
&= m\mu(1-\mu)\mathbb{E}(S_t \mid \theta_{t-1}) + m^2\mu^2\,\mathrm{Var}(S_t \mid \theta_{t-1}),
\end{aligned}
$$

as required. ∎

We require one more lemma before we complete the proof of Theorem 1.

Lemma 3 *For all $t \in \mathbb{N}$ we have*

$$
\mathbb{E}\left((\Delta S_t)^2 \mid \theta_{t-1}\right) = m\mu(1-\mu)I_{t-1}S_{t-1} + (m\mu I_{t-1}S_{t-1})^2.
$$

Proof. The proof is similar to the proof of Lemma 2. We first note that

$$(\Delta S_t)^2 = \left(-\sum_{j=1}^{mI_{t-1}S_{t-1}} \mathbb{I}_j\right)\left(-\sum_{l=1}^{mI_{t-1}S_{t-1}} \mathbb{I}_l\right)$$
$$= \sum_{j=l} \mathbb{I}_j\mathbb{I}_l + \sum_{j\neq l} \mathbb{I}_j\mathbb{I}_l.$$
$$= \sum_{j=l} \mathbb{I}_j + \sum_{j\neq l} \mathbb{I}_j\mathbb{I}_l.$$

There are $(mI_{t-1}S_{t-1})^2$ terms in the expansion of $(\Delta S_t)^2$. Exactly $mI_{t-1}S_{t-1}$ of these fall under the sum where $j = l$, and the remaining $(mI_{t-1}S_{t-1})^2 - mI_{t-1}S_{t-1}$ of them fall under the sum where $j \neq l$. Hence by (2), we have

$$\mathbb{E}\left((\Delta S_k)^2 \mid \theta_{t-1}\right) = \mathbb{E}\left(\sum_{j=l} \mathbb{I}_j \mid \theta_{t-1}\right) + \mathbb{E}\left(\sum_{j\neq l} \mathbb{I}_j\mathbb{I}_l \mid \theta_{t-1}\right).$$
$$= \mu\mathbb{E}(mI_{t-1}S_{t-1} \mid \theta_{t-1}) + \mu^2\mathbb{E}\left((mI_{t-1}S_{t-1})^2 - mI_{t-1}S_{t-1} \mid \theta_{t-1}\right)$$
$$= m\mu(1-\mu)I_{t-1}S_{t-1} + (m\mu I_{t-1}S_{t-1})^2, \tag{14}$$

as required. ∎

We now prove Part 2. of Theorem 1.

Proof of Part 2. We want to find a closed-form expression for $\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1})$. We will require (11) which we proved in Part 1. of Theorem 1. We recall this equation here as

$$\mathbb{E}\left(S_t \mid \theta_{t-1}\right) = S_{t-1} - m\mu I_{t-1}S_{t-1}.$$

Hence by Lemma 2 it only remains to calculate $\mathbb{E}\left(S_t^2 \mid \theta_{t-1}\right)$. Noting that

$$S_t^2 = (S_{t-1} + \Delta S_t)^2 = S_{t-1}^2 + 2S_{t-1}\Delta S_t + (\Delta S_t)^2$$

we can use Lemma 1 and Lemma 3 to compute

$$\mathbb{E}(S_t^2 \mid \theta_{t-1}) = S_{t-1}^2 + 2S_{t-1}\mathbb{E}(\Delta S_t \mid \theta_{t-1}) + \mathbb{E}\left((\Delta S_t)^2 \mid \theta_{t-1}\right)$$
$$= S_{t-1}^2 + 2S_{t-1}\left(-m\mu I_{t-1}S_{t-1}\right) + \left(m\mu(1-\mu)I_{t-1}S_{t-1} + (m\mu I_{t-1}S_{t-1})^2\right)$$
$$= m\mu(1-\mu)I_{t-1}S_{t-1} + S_{t-1}^2 - 2m\mu I_{t-1}S_{t-1}^2 + (m\mu I_{t-1}S_{t-1})^2. \tag{15}$$

Further, we have by (11) that

$$
\begin{aligned}
(\mathbb{E}(S_t \mid \theta_{t-1}))^2 &= (S_{t-1} - m\mu I_{t-1} S_{t-1})^2 \\
&= S_{t-1}^2 - 2m\mu I_{t-1} S_{t-1}^2 + (m\mu I_{t-1} S_{t-1})^2
\end{aligned}
\tag{16}
$$

Now using (15) and (16) we can calculate

$$
\begin{aligned}
\operatorname{Var}(S_t \mid \theta_{t-1}) &= \mathbb{E}(S_t^2 \mid \theta_{t-1}) - (\mathbb{E}(S_t \mid \theta_{t-1}))^2 \\
&= \left(m\mu(1-\mu) I_{t-1} S_{t-1} + S_{t-1}^2 - 2m\mu I_{t-1} S_{t-1}^2 + (m\mu I_{t-1} S_{t-1})^2 \right) \\
&\quad - \left(S_{t-1}^2 - 2m\mu I_{t-1} S_{t-1}^2 + (m\mu I_{t-1} S_{t-1})^2 \right) \\
&= m\mu(1-\mu) I_{t-1} S_{t-1}.
\end{aligned}
\tag{17}
$$

Finally, combining (17) and (11) together with Lemma 2, we have

$$
\begin{aligned}
\operatorname{Var}(\mathcal{R}_t \mid \theta_{t-1}) &= m\mu(1-\mu)\mathbb{E}(S_t \mid \theta_{t-1}) + m^2\mu^2 \operatorname{Var}(S_t \mid \theta_{t-1}) \\
&= m\mu(1-\mu)\left(S_{t-1} - m\mu I_{t-1} S_{t-1}\right) + m^2\mu^2\left(m\mu(1-\mu) I_{t-1} S_{t-1}\right) \\
&= m\mu(1-\mu)S_{t-1} - m^2\mu^2(1-\mu)(1-m\mu) I_{t-1} S_{t-1},
\end{aligned}
$$

completing the proof. ∎

The first consequence of Theorem 1 is the following important Corollary.

Corollary 1 *For sufficiently large N, the variance of $\mathcal{R}_t$ given θ_{t-1} is increasing in S_{t-1}.*

In particular the proof of Corollary 1 which we present below shows that the conditional variance of $\mathcal{R}_t$ is increasing in S_{t-1} when t is small, since at the beginning of the SIR process, infection rates are low and susceptibility is high. This implies that if variance is a concern to policy makers, then containment policies should be stricter earlier on in the spread of the virus.

Proof. Note firstly that an immediate decrease in S_{t-1} translates into a commensurate increase in I_{t-1}, that is, an individual who is no longer susceptible must have gotten infected. As such, it makes sense to think about an *increase* in S_{t-1} as resulting from a commensurate *decrease* in I_{t-1}. One can think of this as having an additional susceptible individual at time $t-2$ who did not get infected by any of the I_{t-2} infected individuals at that time period.

Now, suppose that we fix $S_{t-1} \in \mathbb{N}$ and $I_{t-1} \in \mathbb{N}$ with $S_{t-1} + I_{t-1} \leq N$. Consider what happens when the number of susceptible individuals at time $t-1$ increases by 1.

Formally, let $S'_{t-1} = S_{t-1} + 1$ and let $I'_{t-1} = I_{t-1} - 1$, now consider what happens when $S_{t-1} \to S'_{t-1}$ and $I_{t-1} \to I'_{t-1}$. Letting $\theta'_{t-1} = \{S'_{t-1}, I'_{t-1}\}$, we have

$$
\begin{aligned}
\mathrm{Var}(\mathcal{R}_t \mid \theta'_{t-1}) &= m\mu(1-\mu)S'_{t-1} - m^2\mu^2(1-\mu)(1-m\mu)I'_{t-1}S'_{t-1} \\
&= m\mu(1-\mu)\left(S_{t-1}+1\right) - m^2\mu^2(1-\mu)(1-m\mu)\left(I_{t-1}-1\right)\left(S_{t-1}+1\right) \\
&= \mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1}) + m\mu(1-\mu) - m^2\mu^2(1-\mu)(1-m\mu)\left(I_{t-1}-S_{t-1}-1\right).
\end{aligned}
\tag{18}
$$

Now (18) tells us that $\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1})$ is increasing in S_{t-1} if and only if

$$
m\mu(1-\mu) - m^2\mu^2(1-\mu)(1-m\mu)\left(I_{t-1}-S_{t-1}-1\right) > 0.
\tag{19}
$$

Supposing $m \neq 0$ such that meetings occur, and that $\mu \neq 0, 1$ such that the Poisson trials are not degenerate, (19) holds if and only if

$$
1 - m\mu(1-m\mu)(I_{t-1}-S_{t-1}-1) > 0.
\tag{20}
$$

If $I_{t-1} = S_{t-1} + 1$ then this equation holds trivially. Else writing $m = \frac{\phi}{N}$ and multiplying throughout by N we can write this as

$$
N > \phi\mu\left(1 - \frac{\phi\mu}{N}\right)(I_{t-1}-S_{t-1}-1).
\tag{21}
$$

Noting that the right hand side of (21) is $O(\phi(N))$ this inequality will hold for N sufficiently large whenever $\phi(N) = o(N)$ (and in fact we have assumed that $\phi(N)$ is a constant), completing the proof. ∎

We also present a second corollary regarding the asymptotic conditional variance of $\mathcal{R}_t$.

Corollary 2 *As $N \to \infty$, we have* $\mathrm{Var}(\mathcal{R}_t \mid \theta_{t-1}) \to m\mu(1-\mu)S_{t-1} - m^2\mu^2(1-\mu)I_{t-1}S_{t-1}.$

Proof. The proof follows immediately from noting that

$$
1 - m\mu = 1 - \frac{\phi\mu}{N}
$$

which converges to 1 as $N \to \infty$. ∎

3.2 Poisson trials and Superspreaders

The Poisson trials framework we have used allows us to observe that as the population size N grows, the realized distribution of infection probabilities converges to the parent distribution.

Theorem 2 *Let $X_1, \ldots, X_N$ be random variables drawn from an absolutely continuous distribution F, and let $\xi_k = \inf\{x \colon F(x) \geq k\}$ denote the k-th quantile of F, for $0 < k < 1$. Similarly, let $\hat{\xi}_{Nk} = \inf\{x \colon F_N(x) \geq k\}$ denote the k-th quantile of the sample distribution $X_1, \ldots, X_N$, where $F_N(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(X_i \leq x)$, and $\mathbb{I}(A)$ is an indicator variable for the event A. Then*

$$\mathbb{P}(|\hat{\xi}_{Nk} - \xi_k| > \epsilon) \leq 2 \exp(-2N\delta_\epsilon^2),$$

where $\delta_\epsilon = \min\{F(\xi_k + \epsilon) - k, k - F(\xi_k - \epsilon)\}$. That is, $\hat{\xi}_{Nk} \xrightarrow{P} \xi_k$ exponentially fast.

Proof. This is a well known fact but we provide a proof of one of the bounds for completeness. Let $\epsilon > 0$, and note that

$$\mathbb{P}(|\hat{\xi}_{Nk} - \xi_k| > \epsilon) = \mathbb{P}(\hat{\xi}_{Nk} > \xi_k + \epsilon) + \mathbb{P}(\hat{\xi}_{Nk} < \xi_k + \epsilon).$$

By definition of the sample quantile $\hat{\xi}_{Nk}$, we have

$$
\begin{aligned}
\mathbb{P}(\hat{\xi}_{Nk} > \xi_k + \epsilon) &= \mathbb{P}(k > F_N(\xi_p + \epsilon)) \\
&= \mathbb{P}\left(Nk > \sum_{i=1}^{N} \mathbb{I}(X_i \leq \xi_k + \epsilon) \right) \\
&= \mathbb{P}\left(\sum_{i=1}^{N} \mathbb{I}(X_i > \xi_k + \epsilon) > N(1 - k) \right).
\end{aligned}
$$

Now let $Y_i = \mathbb{I}(X_i > \xi_k + \epsilon)$. Then

$$\mathbb{E}(Y_i) = \mathbb{P}(X_i > \xi_p + \epsilon) = 1 - F(\xi_p + \epsilon).$$

Hence

$$
\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{N} \mathbb{I}(X_i > \xi_k + \epsilon) > N(1-k) \right) &= \mathbb{P}\left(\sum_{i=1}^{N} Y_i - N(1 - F(\xi_k + \epsilon)) > N(F(\xi_p + \epsilon) - k) \right) \\
&= \mathbb{P}\left(\sum_{i=1}^{N} Y_i - \sum_{i=1}^{N} \mathbb{E}(Y_i) > N\delta_1 \right),
\end{aligned}
$$

where $\delta_1 = F(\xi_k + \epsilon) - k$. Hence by Hoeffding's inequality, we have

$$\mathbb{P}(\hat{\xi}_{Nk} > \xi_k + \epsilon) = \mathbb{P}\left(\sum_{i=1}^{N} Y_i - \sum_{i=1}^{N} \mathbb{E}(Y_i) > N\delta_1\right) \leq \exp(-2N\delta_1^2).$$

A similar method shows that

$$\mathbb{P}(\hat{\xi}_{Nk} < \xi_k + \epsilon) \leq \exp(-2N\delta_1^2),$$

and putting the two together yields the desired result. ∎

Suppose for a moment that infection probabilities are individual-specific rather than match-specific. That is, each infected individual has the same probability of infecting any susceptible individual that they meet. Corollary 1 implies that, if we denote a person who has a very high probability of infecting other as being a *superspreader*, then however one defines that in terms of the threshold probability of infecting a susceptible person, the chance of a superspreader being present in a population grows exponentially in the size of the population. More generally when probabilities are match-specific as we have assumed, if we denote an event which causes a large number of new cases as being a *superspreading event*, then the probability of a superspreading event is increasing exponentially fast in the population size.

4 Implications and Conclusion

4.1 Path Dependence

An immediate implication that flows from our model is the possibility of *path dependence*, where two *ex ante* identical populations have persistent differences in observed infection rates over time. To see this, note the the infection rate evolves as follows.

Lemma 4 *For all $t \in \mathbb{N}$, we have*

$$I_t = I_0(1-\gamma)^t - \sum_{k=1}^{t}(1-\gamma)^{t-k}\Delta S_k.$$

Proof. The proof is by induction. For $t = 0$, the result holds trivially. Now suppose the equation holds for some $t \geq 0$, and consider the $t + 1$-th case. We have from Equations (7)

and (8) that

$$I_{t+1} = I_t + \sum_{j=1}^{mI_tS_t} \mathbb{I}_j - \gamma I_t$$

$$= (1-\gamma)I_t + \sum_{j=1}^{mI_tS_t} \mathbb{I}_j$$

$$= (1-\gamma)I_t - \Delta S_{t+1}.$$

Now applying the induction hypothesis yields

$$I_{t+1} = (1-\gamma)\left(I_0(1-\gamma)^t - \sum_{k=1}^{t}(1-\gamma)^{t-k}\Delta S_k \right) - \Delta S_{t+1}$$

$$= I_0(1-\gamma)^{t+1} - \sum_{k=1}^{t}(1-\gamma)^{t+1-k}\Delta S_k - \Delta S_{t+1}$$

$$= I_0(1-\gamma)^{t+1} - \sum_{k=1}^{t+1}(1-\gamma)^{t+1-k}\Delta S_k.$$

Hence the lemma is true for $t+1$ and is therefore true for all $t \in \mathbb{N}$ by induction. ∎

We can see from Lemma 4 that the realized number of infections (the number of "successful" Poisson trials) at $t = 1$ has a persistent effect on the infection rate for all $t > 1$. That is, a bad round of early draws has persistent effect on the path of infection. Indeed, ΔS_t depends crucially on ΔS_k for every $k < t$, and so by Lemma 4 I_t depends crucially on I_k for $k < t$. This implies that a bad early realisation of trials from Δs_1 increases the number of infected people in the population, making transmission more likely since the probability of an infected-susceptible match in the population increases.

If the population is large enough, then the law of large numbers implies that the proportion of successful draws from the Poisson trials will converge to the mean of the trials. In particular, the number of new infected individuals in a given time period will be roughly equal to the mean, which we gave in Lemma 1 as $m\mu I_{t-1}S_{t-1}$.

If $I_0 = \frac{1}{N}$, then standard branching process methods confirm that there are essentially 2 equilibria–either the infection persists forever in the population or $i_t \to 0$ as $N, t \to \infty$.

4.2 Containment Policy

An overarching message from our analysis is that it is incomplete to focus only on the mean of the reproduction rate. Because a virus grows exponentially when $\mathcal{R} > 1$ a policy

maker should consider the mass of the distribution of $\mathcal{R}$ above 1. It is this region that we have referred to as the *danger zone*, which is depicted in the following figure.
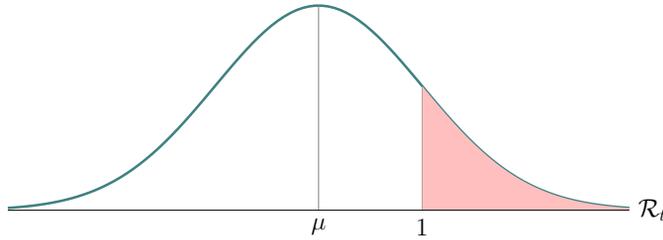
Figure 1: The danger zone.

We end this section with a suggestion for how one might measure the sensitivity the SIR system to perturbations. Fix the time at t and the let $t_d(\phi) > t$ denote the first time period at which the expected number of total infections is greater than $2I_t$ when the number of matchings per individual in each time period is $\phi > 0$. That is, $t_d(\phi)$ denotes the number of time periods it takes for the expected number of total infections to double. We set $t_d(\phi) = \infty$ if infections are never expected to double.

We define the *stability index* of the virus at time t as

$$\mathcal{S}_t = \frac{1}{t_d(\phi + 1)} - \frac{1}{t_d(\phi)}. \tag{22}$$

The stability index measures how responsive the system is with respect to the number of matches per time period. It is one way to account for the mass of $\mathcal{R}_t$ which lies above 1.

A stability index of 0 indicates that the expected increase in the number of infections does not change substantially when the ϕ is increased by 1–in other words, the system is very stable to small perturbations of the number of matchings per time period. On the other hand, an index of 1 indicates that by increasing ϕ by 1, the number of infections is expected to double in the very next period whereas as ϕ currently stands, the number of infections is expected never to double. In other words the system is highly unstable to small perturbations of the number of matchings per time period.

We compute the stability index numerically when the underlying distribution $\mathcal{D}$ for the Poisson trials is $U[0, 1]$ (so $\mu = \frac{1}{2}$). In particular we look at the cases when $N = 100,000$, $\gamma = 0.2$, $t = 1$ and $\phi = 1, 2, 3$. We find

$$(t_d(1), t_d(2), t_d(3) \, t_d(4)) = (\infty, \infty, 1, 1),$$

and hence for $\phi = 1$, $\mathcal{S}_t = 0$, for $\phi = 2$, $\mathcal{S}_t = 1$, and for $\phi = 3$, $\mathcal{S}_t = 0$. The reason the process takes off only when $\phi > 2$ is because this implies $\phi\mu > 1$ and so the infection grows exponentially.

We can also compute the conditional mean and variance (as in Theorem 1) of the reproductive rate numerically. Consider again the case when $\mathcal{D} = U[0,1]$, $\phi = 2$, and $\gamma = 0.2$. We find

$$\mathbb{E}(\mathcal{R}_t \mid \theta_{t-1}) = \frac{1}{N}S_{t-1} + \frac{1}{N^2}I_{t-1}S_{t-1}$$
$$\text{Var}(\mathcal{R}_t \mid \theta_{t-1}) = \frac{1}{4N}S_{t-1} + \frac{N+1}{8N^3}I_{t-1}S_{t-1},$$

Taking $t = 20$ and $N = 100000$ (for the purpose of this example), we have $(S_{t-1}, I_{t-1}) = (44860.3, 40686.1)$, therefore

$$\mathbb{E}(\mathcal{R}_{20} \mid \theta_{19}) = 0.631$$
$$\text{Var}(\mathcal{R}_{20} \mid \theta_{19}) = 0.135,$$

Finally, in general, given a critical value c, numerical mean and variance calculations such as the one above allow us to construct a "confidence interval" for $\mathcal{R}_t$ as

$$CI = \left[\mathbb{E}(\mathcal{R}_t \mid \theta_{t-1}) - c\sqrt{\text{Var}(\mathcal{R}_t \mid \theta_{t-1})}, \mathbb{E}(\mathcal{R}_t \mid \theta_{t-1}) + c\sqrt{\text{Var}(\mathcal{R}_t \mid \theta_{t-1})} \right],$$

giving us another measure for how much of the mass of $\mathcal{R}_t$ lies above 1. Taking $c = 1$ and using the example above we find $CI = [0.264, 0.998]$ and so in this case the system appears to be fairly stable.

4.3   Shelter-in-Place Orders versus Masks

Our model also speaks to the differential effectiveness of alternative containment policies. Given that the variance of the reproductive rate is increasing in S_{t-1}, consider a policy maker who seeks to implement a containment policy early in the spread of the virus. Such a planner has two lines of attack. The first is to try and lower the number of meetings an individual has per time period (ϕ), for example, by a shelter-in-place order. The second is to try and lower the infectiousness of the virus (μ), perhaps by mandating the widespread use of masks.

When t is small, I_{t-1} is approximately 1, and S_{t-1} is approximately N. Hence

$$(m\mu)^2 I_{t-1} S_{t-1} = \frac{\phi\mu}{N^2} I_{t-1} S_{t-1} \approx \frac{\phi^2 \mu^2}{N^2} N,$$

which is close to 0 for large enough N. It follows that early on in the spread of the virus,

$$E(\mathcal{R}_t \mid \theta_{t-1}) \approx m\mu S_{t-1} = \frac{\phi\mu}{N} S_{t-1},$$

and

$$Var(\mathcal{R}_t \mid \theta_{t-1}) \approx m\mu(1-\mu)S_{t-1} = \frac{\phi\mu}{N}(1-\mu)S_{t-1}.$$

While lowering ϕ or μ both result in a lower expected reproductive rate, the effect of such measures on the variance are less straightforward. The conditional variance of the reproductive rate is increasing in ϕ and thus a reduction in ϕ will indeed lower the variance of $\mathcal{R}_t$.

However, the variance is increasing in μ up until $\mu = 1/2$, at which point it is decreasing in μ, leading to the perhaps counterintuitive result that the more contagious a virus is, the less effective it is to lower its infectiousness. The reason for this is that efforts to decrease μ may have the unintended result of also increasing the variance in the spread of the virus. By contrast, ϕ does not suffer from this problem, so a reduction in ϕ (e.g. shelter-in-place) will unambiguously result in a reduction in both the mean and variance of the reproductive rate of the virus.

4.4   Conclusion

By modeling the reproduction rate as something that emerges from primitives of how infections occur we have highlighted the importance of considering the whole distribution of the reproduction rate for understanding the spread of a virus, and the optimal policy response to it.

References

[1] Daron Acemoglu, Victor Chernozhukov, Iván Werning, and Michael D Whinston. A multi-risk sir model with optimally targeted lockdown. Technical report, National Bureau of Economic Research, 2020.

[2] Linda J.S. Allen. Some discrete-time SI, SIR, and SIS epidemic models. *Mathematical Biosciences*, 124(1):83–105, nov 1994.

[3] Frank Ball and Owen D. Lyne. Stochastic multitype sir epidemics among a population partitioned into households. *Advances in Applied Probability*, 33(1):99–123, 2001.

[4] Shweta Bansal, Bryan T Grenfell, and Lauren Ancel Meyers. When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of The Royal Society Interface*, 4(16):879–891, jul 2007.

[5] David W Berger, Kyle F Herkenhoff, and Simon Mongey. An seir infectious disease model with testing and conditional quarantine. *COVID Economics*, 13:1–30.

[6] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[7] Karol Capała and Bartłomiej Dybiec. Epidemics spread in heterogeneous populations. *The European Physical Journal B*, 90(5), may 2017.

[8] Paul C Cross, Philip L.F Johnson, James O Lloyd-Smith, and Wayne M Getz. Utility of r0 as a predictor of disease invasion in structured populations. *Journal of The Royal Society Interface*, 4(13):315–324, nov 2006.

[9] O. Diekmann, J.A.P. Heesterbeek, and J.A.J. Metz. On the definition and the computation of the basic reproduction ratio r 0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28(4), jun 1990.

[10] Greg Dwyer, Joseph S. Elkinton, and John P. Buonaccorsi. Host heterogeneity in susceptibility and disease dynamics: Tests of a mathematical model. *The American Naturalist*, 150(6):685–707, dec 1997.

[11] Martin S Eichenbaum, Sergio Rebelo, and Mathias Trabandt. The macroeconomics of epidemics. Working Paper 26882, National Bureau of Economic Research, March 2020.

[12] Masao Fukui and Chishio Furukawa. Power laws in superspreading events: Evidence from coronavirus outbreaks and implications for sir models. *COVID Economics*, 30:1–43, 2020.

[13] Pietro Garibaldi, Espen R. Moen, and Christopher A Pissarides. Modelling contacts and transitions in the sir epidemics model. Technical report, Centre For Economic Policy Research, April 2020.

[14] Wayne Getz, James Lloyd-Smith, Paul Cross, Shirli Bar-David, Philip Johnson, Travis Porco, and María Sánchez. Modeling the invasion and spread of contagious diseases in heterogeneous populations. In *Disease Evolution*, pages 113–144. American Mathematical Society, jul 2006.

[15] K. Glass, J. Kappey, and B. T. Grenfell. The effect of heterogeneity in measles vaccination on population immunity. *Epidemiology and Infection*, 132(4):675–683, jul 2004.

[16] Wei Gou and Zhen Jin. How heterogeneous susceptibility and recovery rates affect the spread of epidemics on networks. *Infectious Disease Modelling*, 2(3):353 – 367, 2017.

[17] Harrison Hong, Neng Wang, and Jinqiang Yang. Implications of stochastic transmission rates for managing pandemic risks. Working Paper 27218, National Bureau of Economic Research, May 2020.

[18] Babak Jamshidi, Sayed Mohammad Reza Alavi, and Gholam Ali Parham. The distribution of the number of the infected individuals in a stochastic SIR model on regular rooted trees. *Communications in Statistics - Simulation and Computation*, pages 1–18, apr 2019.

[19] Svante Janson, Malwina Luczak, Peter Windridge, and Thomas House. Near-critical SIR epidemic on a random graph with given degrees. *Journal of Mathematical Biology*, 74(4):843–886, jul 2016.

[20] Brian Karrer and M. E. J. Newman. A message passing approach for general epidemic models. *Phys. Rev. E 82, 016101*, 2010.

[21] Matthew J. Kauffman and Erik S. Jules. Heterogeneity shapes invasion: Host size and environment influence susceptibility to a nonnative pathogen. *Ecological Applications*, 16(1):166–175, feb 2006.

[22] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 115(772):700–721, 1927.

[23] Lingcai Kong, Jinfeng Wang, Weiguo Han, and Zhidong Cao. Modeling heterogeneity in direct infectious disease transmission in a compartmental model. *International Journal of Environmental Research and Public Health*, 13(3):253, feb 2016.

[24] Jing Li, Daniel Blakeley, and Robert J. Smith. The failure of r_0. *Computational and Mathematical Methods in Medicine*, 2011:1–17, 2011.

[25] X. Li, L. Cao, and G. F. Cao. Epidemic prevalence on random mobile dynamical networks: individual heterogeneity and correlation. *The European Physical Journal B*, 75(3):319–326, mar 2010.

[26] M. Lipsitch. Transmission dynamics and control of severe acute respiratory syndrome. *Science*, 300(5627):1966–1970, jun 2003.

[27] Alun L Lloyd, Ji Zhang, and A.Morgan Root. Stochasticity and heterogeneity in host–vector models. *Journal of The Royal Society Interface*, 4(16):851–863, jun 2007.

[28] James O Lloyd-Smith, Sebastian J Schreiber, P Ekkehard Kopp, and Wayne M Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066):355–359, 2005.

[29] Robert M. May and Roy M. Anderson. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences*, 72(1):83–111, nov 1984.

[30] Joel C. Miller. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. *Phys. Rev. E*, 76:010101, Jul 2007.

[31] Joel C. Miller. Spread of infectious disease through clustered populations. *Journal of The Royal Society Interface*, 6(41):1121–1134, mar 2009.

[32] Franco M. Neri, Anne Bates, Winnie S. Fchtbauer, Francisco J. Pérez-Reche, Sergei N. Taraskin, Wilfred Otten, Douglas J. Bailey, and Christopher A. Gilligan. The effect of heterogeneity on invasion in spatial epidemics: From theory to experimental evidence in a model system. *PLoS Computational Biology*, 7(9):e1002174, sep 2011.

[33] Franco M. Neri, Francisco J. Prez-Reche, Sergei N. Taraskin, and Christopher A. Gilligan. Heterogeneity in susceptible–infected–removed (sir) epidemics on lattices. *Journal of The Royal Society Interface*, 8(55):201–209, 2011.

[34] Mark E. J. Newman. The spread of epidemic disease on networks. *Phys. Rev. E 66, 016128*, 2002.

[35] Olusegun M. Otunuga. Closed-form probability distribution of number of infections at a given time in a stochastic sis epidemic model. *Heliyon*, 5(9):e02499, 2019.

[36] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, apr 2001.

[37] Siméon-Denis Poisson. *Recherches sur la probabilité des jugements en matiére criminelle et en matiére civile ; précédées des Régles générales du calcul des probabilités.* Bachelier, Paris, 1837.

[38] Benjamin Ridenhour, Jessica M. Kowalik, and David K. Shay. Unraveling r0: Considerations for public health applications. *American Journal of Public Health*, 104(2):e32–e41, feb 2014.

[39] Steven Riley, Christophe Fraser, Christl A. Donnelly, Azra C. Ghani, Laith J. Abu-Raddad, Anthony J. Hedley, Gabriel M. Leung, Lai-Ming Ho, Tai-Hing Lam, Thuan Q. Thach, Patsy Chau, King-Pan Chan, Su-Vui Lo, Pak-Yin Leung, Thomas Tsang, William Ho, Koon-Hung Lee, Edith M. C. Lau, Neil M. Ferguson, and Roy M. Anderson. Transmission dynamics of the etiological agent of sars in hong kong: Impact of public health interventions. *Science*, 300(5627):1961–1966, 2003.

[40] Arkaprabha Sau, Santanu Phadikar, and Ishita Bhakta. Estimation of time dependent reproduction number for the ongoing COVID-2019 pandemic. *SSRN Electronic Journal*, 2020.

[41] Marialisa Scatà, Alessandro Di Stefano, Pietro Liò, and Aurelio La Corte. The impact of heterogeneity and awareness in modeling epidemic spreading on multiplex networks. *Scientific Reports*, 6(1), nov 2016.

[42] Kieran J. Sharkey. Deterministic epidemiological models at the individual level. *Journal of Mathematical Biology*, 57(3):311–331, feb 2008.

[43] Zhuang Shen, Fang Ning, Weigong Zhou, Xiong He, Changying Lin, Daniel P. Chin, Zonghan Zhu, and Anne Schuchat. Superspreading SARS events, beijing, 2003. *Emerging Infectious Diseases*, 10(2):256–260, feb 2004.

[44] Herman M. Singer. Short-term predictions of country-specific covid-19 infection rates based on power law scaling exponents. *arXiv preprint arXiv:2003.11997*, 2020.

[45] Alexis A. Toda. Susceptible-infected-recovered (sir) dynamics of covid-19 and economic impact. *COVID Economics*, 1:43–63, 2020.

[46] Elisabetta Tornatore, Stefania [Maria Buccellato], and Pasquale Vetro. Stability of a stochastic sir system. *Physica A: Statistical Mechanics and its Applications*, 354:111 – 126, 2005.

[47] Erik M. Volz, Joel C. Miller, Alison Galvani, and Lauren Ancel Meyers. Effects of heterogeneous and clustered contact patterns on infectious disease dynamics. *PLoS Computational Biology*, 7(6):e1002042, jun 2011.

[48] Peter J. Witbooi. Stability of an seir epidemic model with independent stochastic perturbations. *Physica A: Statistical Mechanics and its Applications*, 392(20):4928 – 4936, 2013.

[49] M. E. J. Woolhouse, C. Dye, J.-F. Etard, T. Smith, J. D. Charlwood, G. P. Garnett, P. Hagan, J. L. K. Hii, P. D. Ndhlovu, R. J. Quinnell, C. H. Watts, S. K. Chandiwana, and R. M. Anderson. Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proceedings of the National Academy of Sciences*, 94(1):338–342, 1997.

[50] N. Wormald. *Some problems in the enumeration of labelled graphs*. Ph.d. thesis, Newcastle University, 1978.

[51] Hui Yang, Ming Tang, and Thilo Gross. Large epidemic thresholds emerge in heterogeneous networks of heterogeneous nodes. *Scientific Reports*, 5(1), aug 2015.

[52] Chong You, Yuhao Deng, Wenjie Hu, Jiarui Sun, Qiushi Lin, Feng Zhou, Cheng Heng Pang, Yuan Zhang, Zhengchao Chen, and Xiao-Hua Zhou. Estimation of the time-varying reproduction number of covid-19 outbreak in china. *International Journal of Hygiene and Environmental Health*, 228:113555, 2020.

[53] Xianghua Zhang and Ke Wang. Stochastic seir model with jumps. *Applied Mathematics and Computation*, 239:133 – 143, 2014.

[54] Xiaojing Zhong and Feiqi Deng. Dynamics of a stochastic multigroup SEIR epidemic model. *Journal of Applied Mathematics*, 2014:1–9, 2014.

5   Appendix A

We make some remarks here about the matching technology when the independence assumption of matches is violated. The main effect of losing the independence of matches is that it invalidates $\mathcal{R}_t^{pop}$. Recall that $\mathcal{R}_t^{pop}$ is defined as the population-level reproductive rate, that is, a random variable whose distribution is equal to the number of new infections in the population from time t to time $t + 1$. In general, we can define the changes in

S_t, I_t, and r_t by the discrete Markov process:

$$\Delta S_{t+1} = -\mathcal{R}_t^{pop} \tag{23}$$

$$\Delta I_{t+1} = \mathcal{R}_t^{pop} - I_t\gamma \tag{24}$$

$$\Delta R_{t+1} = I_t\gamma, \tag{25}$$

We want to work out a general form for $\mathcal{R}_t^{pop}$ without independence of matchings. Suppose we fix an interaction structure in the population at time t by the configuration model or some other method. Let m_k $(k = 1, \ldots, n_t)$ denote the number of susceptible individuals who interact with exactly k *different* infected individuals. Then denoting by $[l]$ the set $\{1, 2, \ldots, l\}$, we can write the population reproduction rate as

$$\mathcal{R}_t^{pop} = S_t \left(\sum_{j_1=0}^{m_1} \mathbb{I}_{j_1} + \sum_{\substack{j_i=0 \\ i\in[2]}}^{m_2} \mathbb{I}_{j_1\cup j_2} + \sum_{\substack{j_i=0 \\ i\in[3]}}^{m_3} \mathbb{I}_{j_1\cup j_2\cup} + \cdots + \sum_{\substack{j_i=0 \\ i\in[n_t]}}^{m_{n_t}} \mathbb{I}_{\bigcup_{k\le n_t} j_k} \right) \tag{26}$$

$$= S_t \sum_{l=1}^{n_t} \sum_{\substack{j_i=0 \\ i\in[l]}}^{m_l} I_{\bigcup_{k\le l} j_k}. \tag{27}$$

Note that for an arbitrary collection of events $\bigcup_k A_k$, the indicator variable of their union satisfies

$$\mathbb{I}_{\bigcup_k A_k} = 1 - \prod_k (1 - \mathbb{I}_{A_k}). \tag{28}$$

Which allows us to rewrite the above as

$$\mathcal{R}_t^{pop} = \sum_{l=1}^{n_t} \sum_{j_i=0, i\in[l]}^{m_l} \left(1 - \prod_{k\le l} (1 - \mathbb{I}_{j_k}) \right) S_t. \tag{29}$$

It is clear that this makes the system of equations in Equations (23) to (25) particularly difficult to work with let alone to solve in some kind of closed form. At this point it is worth looking into the matching technology in the hope that $\mathcal{R}_t^{pop}$ can be simplified. Indeed, under the assumption of a large N and a small $\phi(N)$ as outlined in Section 2.1, matches are approximately independent. This means that $m_k \to 0$ for all $k \ge 2$ It follows that the number of infected-susceptible matches in the population is $m_1 = I_t m S_t$, which gives us

$$\mathcal{R}_t^{pop} = \sum_{j=1}^{m_1 I_t S_t} \mathbb{I}_j,$$

as used in the paper.

6 Appendix B

A natural question is whether or not there exists a closed form expression for the unconditional mean and variance $\mathbb{E}(\mathcal{R}_t)$ and $\mathrm{Var}(\mathcal{R}_t)$. We present some results here which indicate that whilst in theory such an expression exists, it is rather difficult to write down. First, recall that the mean of $\mathcal{R}_t$ depends entirely on the mean of S_t. Our first lemma establishes an equation for the mean of S_t in terms of lower order terms $I_k S_k$ for $k < t$.

Lemma 5 *For all $t \in \mathbb{N}$,*

$$\mathbb{E}(S_t) = S_0 + \frac{m}{2} \sum_{k=0}^{t-1} \mathbb{E}(I_k S_k).$$

Proof. First, note that for all $t \in \mathbb{N}$, we have

$$S_t = \sum_{k=0}^{t} \Delta S_k, \tag{30}$$

where we let $\Delta S_0 = S_0$. Hence it follows that

$$\mathbb{E}(S_t) = S_0 + \sum_{k=1}^{t} \mathbb{E}(\Delta S_k). \tag{31}$$

Note then that since $\Delta S_{k+1} = \sum_{k=1}^{m I_k S_k} \mathbb{I}_j$, we have by *Wald's identity* that

$$\mathbb{E}(\Delta S_{k+1}) = \mathbb{E}(m I_k S_k)\mathbb{E}(\mathbb{I}_j) = \frac{m}{2}\mathbb{E}(I_k S_k). \tag{32}$$

Finally then substitution (32) into (31) and relabelling the index, we have

$$\mathbb{E}(S_t) = S_0 + \frac{m}{2} \sum_{k=0}^{t-1} \mathbb{E}(I_k S_k), \tag{33}$$

proving the lemma. ∎

We can see from Lemma 5 that the expectation of S_t depends entirely upon the expectation of $I_t S_t$.

We have from Lemma 4 an expression for I_t in terms of ΔS_k for $k < t$. This allows us to find a closed form expression for $I_t S_t$ from which we can calculate its mean. However,

before we are ready to compute the expectation of $I_t S_t$, we need to know how to compute $\mathbb{E}(\Delta S_k \Delta S_l)$ for both $l = k$ and $l \neq k$. We establish this with the following lemma.

Lemma 6 *Let $k, l \in \mathbb{N}$. Then*

$$\mathbb{E}(\Delta S_k \Delta S_l) = \frac{(Nd)^2}{4} \mathbb{E}\left(I_{k-1} S_{k-1} I_{l-1} S_{l-1} \right).$$

Proof. Notice firstly that ΔS_k and ΔS_l are not independent. Indeed, suppose $l < k$, then S_{k-1} appears in the upper limit of the sum for ΔS_k, and one has that $S_{k-1} = \sum_{j=0}^{k-1} \Delta s_j$. We have

$$\Delta S_k \Delta S_l = \left(\sum_{j=1}^{mI_{k-1}S_{k-1}} \mathbb{I}_j \right) \left(\sum_{q=1}^{mI_{l-1}S_{l-1}} \mathbb{I}_j \right).$$

Noting that this is a sum over $(mI_{k-1}S_{k-1})(mI_{l-1}S_{l-1})$ products of independent indicator variables, we apply Wald's identity to get

$$\mathbb{E}(\Delta S_k \Delta S_l) = \mathbb{E}\left((mI_{k-1}S_{k-1})(mI_{l-1}S_{l-1}) \right) \mathbb{E}(\mathbb{I}_j \mathbb{I}_l) \tag{34}$$

$$= (Nd)^2 \mathbb{E}\left(I_{k-1} S_{k-1} I_{l-1} S_{l-1} \right) \mathbb{E}(\mathbb{I}_j) \mathbb{E}(\mathbb{I}_l) \tag{35}$$

$$= \frac{(Nd)^2}{4} \mathbb{E}\left(I_{k-1} S_{k-1} I_{l-1} S_{l-1} \right), \tag{36}$$

completing the proof. ∎

We are now ready to provide a recursive formula for $I_t S_t$, which will allow us to compute its expectation.

Lemma 7 *The expectation $\mathbb{E}(I_t S_t)$ satisfies the recursive equation*

$$\mathbb{E}(I_t S_t) = \left(1 - \gamma + \frac{m}{4} \right) \mathbb{E}(I_{t-1} S_{t-1}) - \frac{(Nd)^2}{4} \mathbb{E}\left((I_{t-1} S_{t-1})^2 \right)$$

$$- \frac{(Nd)^2}{4} \sum_{k=1}^{t-1} \left(1 + (1-\gamma)^{t-k} \right) \mathbb{E}(S_{t-1} I_{t-1} S_{k-1} I_{k-1}).$$

Proof. Recall that in the proof of Lemma 4 we used the fact that $I_t = (1 - \gamma)I_{t-1} + \Delta S_t$.

Hence we may write

$$I_t S_t = ((1-\gamma)I_{t-1} + \Delta S_t)\left(S_0 + \sum_{k=1}^{t-1}\Delta S_k + \Delta S_t\right)$$

$$= (1-\gamma)I_{t-1}S_{t-1} - S_0\Delta S_t - \Delta S_t\sum_{k=1}^{t-1}\Delta S_k - (\Delta S_t)^2 + (1-\gamma)I_{t-1}\Delta S_t.$$

Now using Lemma 4, we have

$$I_{t-1} = I_0(1-\gamma)^t - \sum_{k=1}^{t-1}(1-\gamma)^{t-k}\Delta S_k,$$

and hence that

$$I_t S_t = (1-\gamma)I_{t-1}S_{t-1} - S_0\Delta S_t - \Delta S_t\sum_{k=1}^{t-1}\Delta S_k - (\Delta S_t)^2$$

$$+ I_0(1-\gamma)^t\Delta S_t - \Delta S_t\sum_{k=1}^{t-1}(1-\gamma)^{t-k}\Delta S_k.$$

Which we can rewrite by collecting ΔS_t terms as

$$I_t S_t = (1-\gamma)I_{t-1}S_{t-1} - \left(S_0 + I_0(1-\gamma)^t\right)\Delta S_t - \mathbb{E}\left((\Delta S_t)^2\right) - \sum_{k=1}^{t-1}(1+(1-\gamma)^{t-k})\mathbb{E}(\Delta S_t\Delta S_k).$$

Finally, applying Lemma 6 and (32), we have the desired result. ∎

An alternative but less useful expression for $\mathbb{E}(I_t S_t)$ in terms of lower order terms can be derived by using the following lemma. This lemma is a little more involved but establishes $I_t S_t$ in terms of I_k and S_k for $k < t$.

Lemma 8 *For all $t \in \mathbb{N}$ we have*

$$I_t S_t = S_0 I_0(1-\gamma)^t + \sum_{k=1}^{t}\left(I_0(1-\gamma)^t - S_0(1-\gamma)^{t-k}\right)\Delta S_k$$

$$- \sum_{k=2}^{t}\sum_{l=1}^{k-1}\left[(1-\gamma)^{t-k} + (1-\gamma)^{t-l}\right]\Delta S_k\Delta S_l - \sum_{k=1}^{t}(1-\gamma)^{t-k}(\Delta S_k)^2.$$

Proof. Recall from (30) that $S_t = S_0 + \sum_{k=1}^{t}\Delta S_k$. Substituting this, and the equation we

established in Lemma 4 into the expression $I_t S_t$, we have for any $t \in \mathbb{N}$,

$$
I_t S_t = \left(I_0 (1-\gamma)^t - \sum_{k=1}^{t} (1-\gamma)^{t-k} \Delta S_k \right) \left(S_0 + \sum_{l=1}^{t} \Delta S_l \right)
$$

$$
= S_0 I_0 (1-\gamma)^t - S_0 \sum_{k=1}^{t} (1-\gamma)^{t-k} \Delta S_k + I_0 (1-\gamma)^t \sum_{l=1}^{t} \Delta S_l - \sum_{k=1}^{t} \sum_{l=1}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l
$$

$$
= S_0 I_0 (1-\gamma)^t + \sum_{k=1}^{t} \left(I_0 (1-\gamma)^t - S_0 (1-\gamma)^{t-k} \right) \Delta S_k - \sum_{k=1}^{t} \sum_{l=1}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l.
$$

Now we work with the double sum to put it into the form given in the lemma. we have

$$
\sum_{k=1}^{t} \sum_{l=1}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l = \sum_{\substack{k,l \\ k=l}}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l + \sum_{\substack{k,l \\ k\neq l}}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l
$$

$$
= \sum_{k=1}^{t} (1-\gamma)^{t-k} (\Delta S_k)^2 + \sum_{\substack{k,l \\ k\neq l}}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l.
$$

Focusing on the second summation, we note that by summing over the upper and lower triangles of an $l \times k$ matrix, we can write

$$
\sum_{\substack{k,l \\ k\neq l}}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l = \sum_{k=2}^{t} \sum_{l=1}^{k-1} (1-\gamma)^{t-k} \Delta S_k \Delta S_l + \sum_{l=2}^{t} \sum_{k=1}^{l-1} (1-\gamma)^{t-k} \Delta S_k \Delta S_l.
$$

Then by swapping the labels l and k on the second sum, this gives

$$
\sum_{\substack{k,l \\ k\neq l}}^{t} (1-\gamma)^{t-k} \Delta S_k \Delta S_l = \sum_{k=2}^{t} \sum_{l=1}^{k-1} (1-\gamma)^{t-k} \Delta S_k \Delta S_l + \sum_{k=2}^{t} \sum_{l=1}^{k-1} (1-\gamma)^{t-l} \Delta S_k \Delta S_l
$$

$$
= \sum_{k=2}^{t} \sum_{l=1}^{k-1} \left[(1-\gamma)^{t-k} + (1-\gamma)^{t-l} \right] \Delta S_k \Delta S_l,
$$

completing the proof. ∎

Putting $I_t S_t$ in the above form allows us to apply the expectation operator in a clean way. In particular, applying Lemma 6 and (32), we can evaluate the expectation of $I_t S_t$

using Lemma 8 as

$$\mathbb{E}(I_t S_t) = S_0 I_0 (1-\gamma)^t + \sum_{k=1}^{t} \left(I_0(1-\gamma)^t - S_0(1-\gamma)^{t-k} \right) \frac{m}{2} \mathbb{E}(I_{k-1} S_{k-1})$$

$$- \sum_{k=2}^{t} \sum_{l=1}^{k-1} \left[(1-\gamma)^{t-k} + (1-\gamma)^{t-l} \right] \mathbb{E}(\Delta S_k \Delta S_l)$$

$$, - \sum_{k=1}^{t} (1-\gamma)^{t-k} \frac{1}{4} \left((m I_{k-1} S_{k-1} + (m I_{k-1} S_{k-1})^2 \right).$$

The CoRisk-Index: A data-mining approach to identify industry-specific risk assessments related to COVID-19 in real time[1]

Fabian Stephany,[2] Niklas Stoehr,[3] Philipp Darius,[4] Leonie Neuhäuser,[5] Ole Teutloff[6] and Fabian Braesemann[7]

While the coronavirus spreads, governments are attempting to reduce contagion rates at the expense of negative economic effects. Market expectations plummeted, foreshadowing the risk of a global economic crisis and mass unemployment. Governments provide huge financial aid programmes to mitigate the economic shocks. To achieve higher effectiveness with such policy measures, it is key to identify the industries that are most in need of support. In this study, we introduce a data-mining approach to measure industry-specific risks related to COVID-19. We examine company risk reports filed to the U.S. Securities and Exchange Commission (SEC). This alternative data set can complement more traditional economic indicators in times of the fast-evolving crisis as it allows for a real-time analysis of risk assessments. Preliminary findings suggest that the companies' awareness towards corona-related business risks is ahead of the overall stock market developments. Our approach allows to distinguish the industries by their risk awareness towards COVID-19. Based on natural language processing, we identify corona-related risk topics and their perceived relevance for different industries.

The preliminary findings are summarised as an up-to-date online index. The CoRisk-Index tracks the industry-specific risk assessments related to the crisis, as it spreads through the economy. The tracking tool is updated

weekly. It could provide relevant empirical data to inform models on the economic effects of the crisis. Such complementary empirical information could ultimately help policymakers to effectively target financial support in order to mitigate the economic shocks of the crisis.

1 | INTRODUCTION

With COVID-19 ("coronavirus") reaching the level of a pandemic, governments and companies around the world are also exposed to the resulting risks for the highly interconnected global economy. To slow down the spread of COVID-19, governments in China, Europe, the US, and beyond are taking drastic measures, such as travel warnings, border and store closures, regional lockdowns, and curfews.

These measures are having drastic consequences for personal freedom and the economy. Some sectors, such as airlines or hotels, are facing a nearly complete breakdown of demand. As a first reaction to the expected economic shocks, the global stock markets have collapsed (see Fig. 1B). In an attempt to mitigate the general economic downturn, governments all over the world are providing considerable financial support. The US government, for example, is preparing an aid package of 2.2$ trillion in response to the virus [1]. The German government plans to take up to 156 billion Euros of additional debt (equivalent to half of the federal budget for 2020) to support the economy [2].

Many of these immediate aid packages are not targeted to specific industries, but are meant to support businesses in all parts of the economy. While such general programmes can help to stabilise financial markets in the short term, it is paramount for their long-term effectiveness to concentrate the support to those areas of the economy that are most in need. For this reason, governments have to identify the industries that are most severely affected by the coronavirus pandemic. However, in the current situation, policymakers lack reliable and up-to-date empirical data, which would allow to assess industry-specific economic risks in real-time. Such information would be crucial to effectively target financial support as the crisis hits the economy and to mitigate the economic shocks.

The study presented here investigates a potential data source that could provide an empirical basis to identify industry-specific economic risks related to COVID-19 and to inform models on the economic effects of the current crisi. We examine company risk reports (10-K reports) filed to the U.S. Securities and Exchange Commission (SEC) and introduce a data-mining approach to measure firms' risk assessments.[1] In collecting all reports published from 30th January 2020 — the day the term *coronavirus* first appeared in a 10-K report — we can assess and track the reported risk perceptions related to COVID-19 for different industries.

Preliminary findings suggest that the company risk reports' show a forward-looking awareness of potential economic risks associated with the corona-crisis, which is leading stock market developments. Moreover, the awareness towards COVID-19 differs substantially between industries. For example, while 78 % of the firms in retail have mentioned the coronavirus as a potential economic risk, only 23 % of the businesses in financial services have done so. Thirdly, based on natural language processing, we can identify specific corona-related risk topics and their perceived relevance for the different industries. Lastly, our approach allows us distinguish the industries by their reported risk awareness towards COVID-19. The empirical information provided could help to inform macro-economic models on the effects of the corona-crisis [3, 4, 5] and, thus, help to inform policymakers to better target current economic support programmes to industries that report most severe risks at the current phase of the crisis.The preliminary findings presented here are summarised in one compound index. The *CoRisk*-Index[2] tracks industry-specific risk assessments related to COVID-19 in real-time. It is available on an interactive online dashboard.

---

[1]SEC filings represent financial statements of publicly listed companies including a risk assessment. SEC filings are imperative to comply with legal and insurance requirements and therefore should contain the most relevant risks. As a result, companies have a strong incentive to neither under- nor overestimate risks. Moreover, analysing the most recent two-month period of SEC filings represents a random sample of all companies since companies are obliged to report to the SEC at a fixed, but randomly assigned date, independent of their industry.

[2]`http://oxford.berlin/CoRisk`

A S&P 1200 Global Index (index points)

Feb, 20:
Start of stock
market crash

Mar, 09:
'Black Monday'

B Corona-sentence Text Negativity in 10-K reports
(avg. daily share of negative words per 'corona' sentence)

Jan, 30: WHO
declares public
health emergency

Apr, 26 – May, 04:
15 States abandon
stay–at–home order
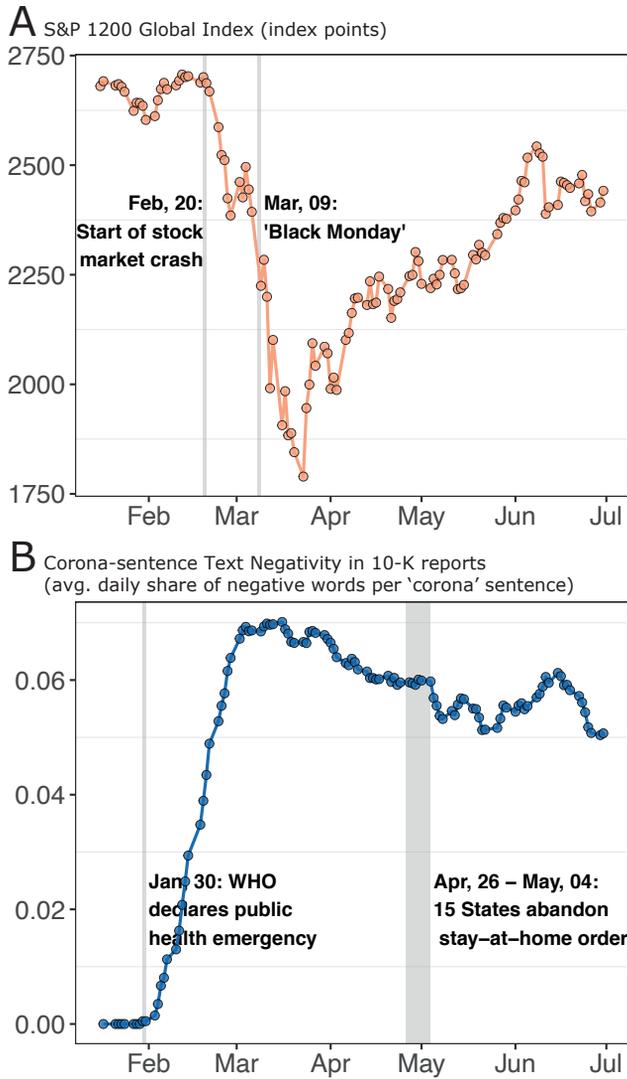
FIGURE 1   (A) S&P Global 1200 Stock Index from 30/01 – 30/06/2020.   (B) Text 'Negativity' (sentiment) in sentences mentioning coronavirus of 10-k reports. The negative sentiment surges before the global stock market crash and peaks before "Black Monday's" most severe stock market losses.

This tool is constantly updated. It allows researchers, policymakers, and the public to estimate potential risk factors of COVID-19 in individual sectors ofthe economy. As the risk filings database is updated on a daily basis, the online tool will allow tracking the potential impact ofthe crisis as it unfolds and spreads through the economy. With more firms providing risk reports that describe the potential or real impacts ofthe crisis, the tool will be refined to allow for a more granular perspective on individual industries and sub-sectors.
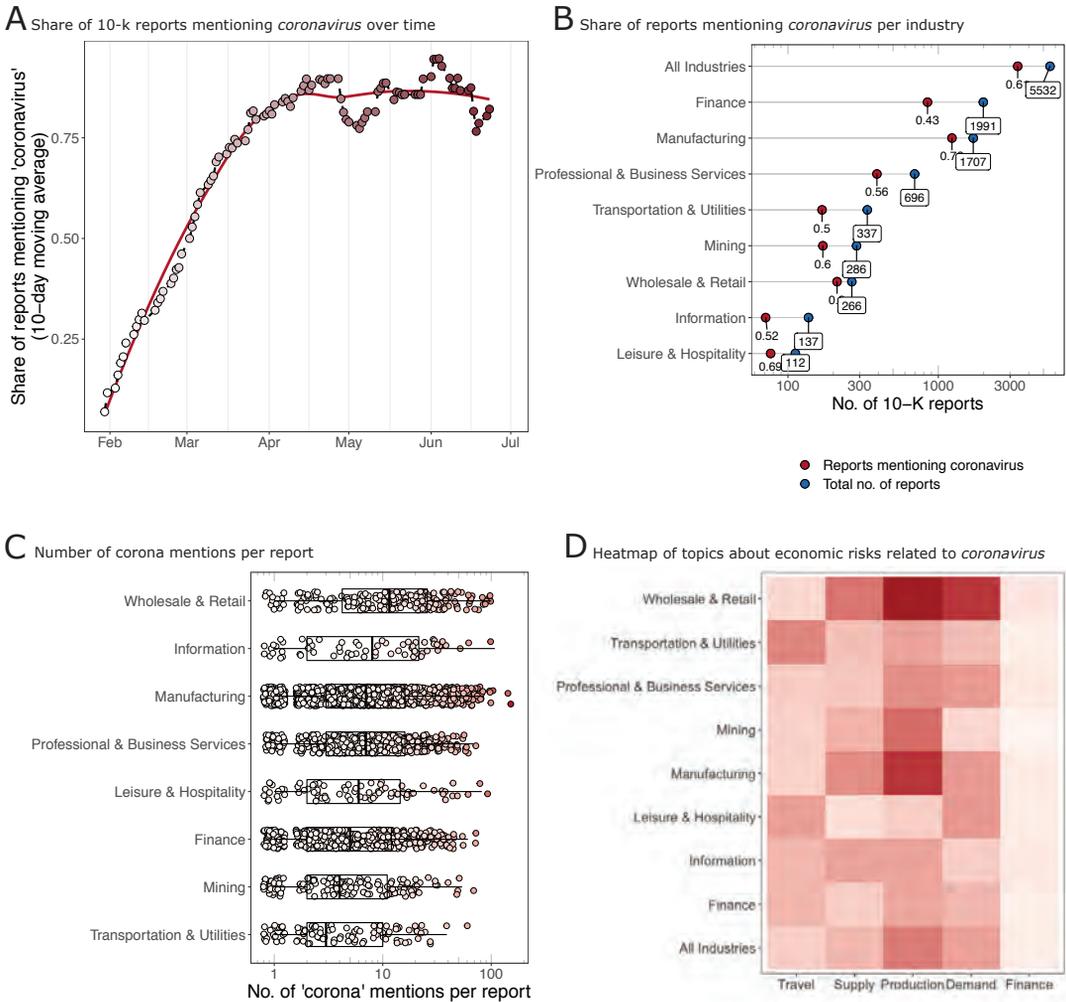
FIGURE 2 (**A**) Share of 10-K reports mentioning *coronavirus* over time: with the global spread of the crisis, the share of firms that mention the pandemic as a risk factor increases sharply.(**B**) Number of 10-K reports filed since 30/01/2020 and share of reports mentioning *coronavirus*. (**C**) Number of *coronavirus* mentions per report. (**D**) Heatmap of relevant economic topics mentioned in 10-K reports related to *coronavirus*.

The study is structured as follows: The next section introduces relevant related research focusing on the economic consequences of infectious diseases and risk assessment in highly interconnected economies. In section 3 we discuss our data and methods before presenting preliminary results in section 4. Lastly, we discuss policy implications, methodological limitations, and planned extensions of our approach.
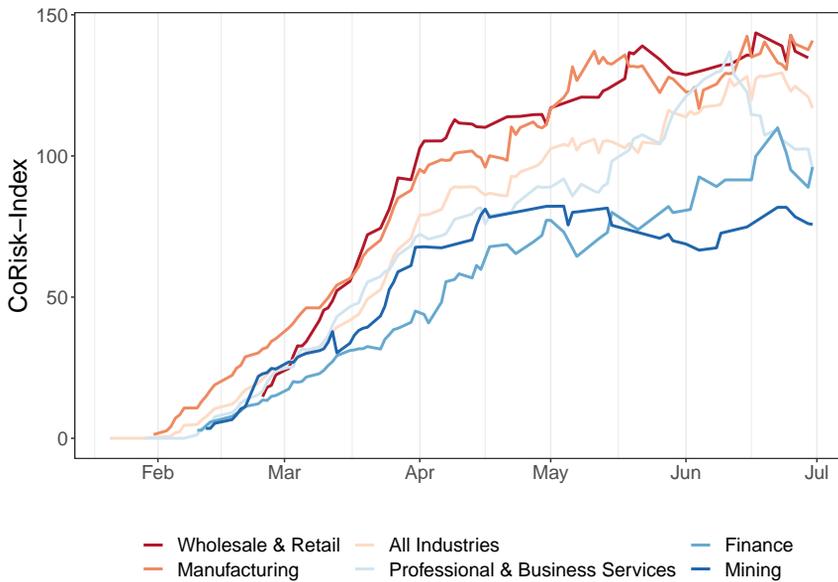
FIGURE 3   The *CoRisk*-Index: the index is updated weekly. It is a compound measure of the share of corona-mentioning firms per industry,the average number of corona-mentions per report and the industry-specific corona-sentence text negativity.

2 | RELATED WORK

In this section, we review related work on the economic consequences of COVID-19 and studies on previous epidemics. Moreover, we discuss the assessment of economic risks via reports such as the 10-K reports required by the U. S. Securities and Exchange Commission (SEC).

2.1 | Studies on the economic effects of COVID-19

Global pandemics of infectious diseases are not a new phenomenon. Throughout the last century, the world experienced several global and regional outbreaks: Most severely, millions died during the spread of the Spanish flu in 1918-1920. More recently, smaller epidemics spread around many countries, such as SARS in 2002, the swine flu in 2009 and Ebola in 2014. Nonetheless, the current coronavirus pandemic is, in recent history, unprecedented in its global social and economic consequences. Governments are taking drastic measures while researchers attempt to provide urgently needed policy advice. Much of the coronavirus-related social science and public health research focuses on disease transmission, global spread, and different interventions (see for example [6, 7, 8, 9, 10]).

A rapidly growing body of literature investigates the (potential) economic consequences induced by the COVID-19 pandemic [3, 4, 5, 11, 12, 13].[3] Many of these studies aim to assess the potential economic consequences in presenting

---

[3]At this point we want to highlight that it is not the aim of this study to provide a model of risk forecasting that is competing with established macro-economic approaches as described in this section. In contrast, the data-driven methodology presented here aims to explore an alternative data source that could help to inform such economic models. The advantage of the data we provide here is the high time resolution. More traditional sources of empirical information used to calibrate macro-economic models usually include, for example, unemployment rates. While the value of such statistics is undisputed, they are reported with a time-lag. We perceive the purpose of the index provided here

simulation-based macroeconomic models. For example, Dorn et al. [13] use scenario calculations to estimate the economic costs of the pandemic for the case of Germany. The authors estimate the financial consequences for the state budget and the employment effects depending on the length of the economic shutdown. Moreover, they include differential adverse effects by sectors, based on press releases and the provisional Ifo business climate index for March 2020. They conclude that the travel and restaurant industry is likely to face a complete shutdown, whereas the pharmaceutical, logistics and health sectors are likely to continue to operate at full capacity. Ludvigson et al. [5] aim to estimate the macroeconomic consequences of the pandemic in investigating the impact of disasters in the recent U. S. history. Baldwin and Mauro [14, 12] collect a great variety of perspectives on the potential economic implications of COVID-19. Topics range from impacts on trade, economic policy measures, monetary policy, and finance to labour market effects. These contributions rely on simulations, scenarios, descriptive statistics and qualitative arguments.

Regarding the economic impact by sector, Gopinath [15] identifies manufacturing and the services sector as disproportionately affected in China, based on the Purchasing Managers' Index. Ramelli and Wagner [16] uses Google search intensity to measure attention paid to COVID-19 and stock market data to reveal the economic impact by sector. Their analysis shows that the energy, retail, and transportation sector experienced the largest losses in the United States and China, whereas health care gained considerably in both countries. The analysis by Huang et al. [17] (also based on stock market data) confirms that the services sectors seem to be the most severely affected in China, and del Rio-Chanona et al. [3] provide quantitative predictions of first-order supply and demand shocks to the U. S. economy on the level of individual industries.

In sum, the contributions presented in this section provide valuable insights, quantitative scenario calculations, and timely policy recommendations. Most of the macroeconomic analyses are, however, based on assumptions-driven simulations and models, as up-to-date empirical data to assess the immediate economic consequences or industry-specific risks are lacking. These studies could benefit from the alternative data set we are exploring in this study.

2.2 | Historical pandemics

While research on recent epidemics are limited to simulations [18, 19], or specific sectors [20], the study of historical pandemics might provide informative empirical assessments of pandemic-related economic effects. Studies on the 1918 Spanish Flu confirm the primordial effectiveness of non-pharmaceutical (such as, for example, social distancing) interventions even if these come at the cost of economic slowdowns [21, 22]. Moreover, based on the historical data, researchers find a correlation between mortality rates and declines in GDP, consumption and returns on stocks [23] and an increase in poverty [24]. However due to the global scope of the crisis for the highly inter-connected world economy of 2020, insights derived from past epidemics are of limited use in identifying the various industry-related risks during the COVID-19 pandemic.

Both, the research on COVID-19 and the historical pandemics rely on stock market information to quantify the economic effects of infectious diseases. However, stock market information comes with several drawbacks. Most importantly, stock markets are prone to irrational herd behaviour and prices capture a variety of information signals into one aggregated index. Examining current stock market dynamics reveals a general economic downturn, but it does not allow to isolate the sector-specific COVID-19 risks. Therefore, we propose to use SEC reports which include risk statements. We argue that these reports represent a promising real-time measure of industry-specific business risks. Furthermore, the analysis of report statements discussing 'coronavirus' allows to isolate the business risks exclusively associated with the COVID-19 outbreak.

---

to be a complementary data source that could be compared with official statistics on the economic effects of the crisis over time.

2.3 | Assessing economic risks via business reports

Since the great recession hit the world economy in 2008, risk has been a crucial topic in governance and finance. While risk assessments of the financial system led to diverse measures to make the world economy less vulnerable to economic shocks originating in the financial sector, a health crisis, such as the current pandemic, poses different risks to the economy. While government measures against the spread of the disease hinder the population from working and consuming, which results in businesses interrupting production, many economies face demand and supply shocks at the same time. In particular, as different industries rely on distinct input factor compositions and supply chains, the sectors of the economy react differently to shocks [25]. Regarding the COVID-19 crisis, we expect sectors whose operations are more connected to supply from manufacturers in China to publicly report corona-related risk earlier than others. These sectors are also highly connected and interdependent within the national economy and risk might spread between sectors [26, 27].

Most risk assessment approaches focus on quantitative probability-based methods and financial data [28, 29]. The data published in such quantified risk assessments is often made available retrospectively, which makes a real-time evaluation of risks difficult. In contrast to such assessments, we investigate the annual 10-K reports filed to the U.S. Securities and Exchange Commission (SEC), which provide verbal corporate risk disclosure and financial statements. Besides, SEC filings are imperative for legal and insurance requirements and they need to contain the most relevant risks to protect the company from legal liabilities. Depending on the volume of publicly traded stocks, companies with a public float of over 700 million USD are obliged to report within 60, while smaller companies have 75 or 90 days after the end of the fiscal year. Furthermore, the reports inform investment decisions and risk governance at the same time and, thus, companies, as rational agents, are likely to communicate moderate risk assessments [30]. In fact, prior work has underlined the forward-looking nature of the reports, since they allowed a more effective prediction of volatility on stock market returns than the compared approaches [31]. Correspondingly, we expect the 10-K reports to also provide forward-looking information on risk assessments during the observed time period and in particular on rising business risk in relation to the spread of COVID-19.

Moreover, the filing companies are categorised into sectors using the "Standard Industry Classifier", which allows a real-time analysis of industry-specific risk developments. In this study we apply natural language processing to extract corona-related risk information from the reports and analyse sector-specific differences in risk awareness and disclosure (details in section 3 and in the appendix A).

2.4 | Hypotheses

Based on the literature on the potential economic effects of pandemics and on the information available in business risks reports, we derive the working hypothesis[4] of this study:

> *Working hypothesis: SEC 10-K reports contain corona-related information, which allow to track the industry-specific economic risk assessments in near real-time as the economic crisis unfolds.*

This working hypothesis is split into two operationalised hypotheses. A core assumption of the investigation of the risk reports is that they contain economic meaningful information. Accordingly, the sentiment of the reports should reflect overall market conditions:

---

[4]While the results provided in this working paper are largely descriptive, we present the research hypotheses, which have guided our data exploration.

*H*1: *10-K risk reports contain economic relevant information on short-term market expectations, i. e. they are related to overall stock market trends in the current crisis.*

Industries are affected differently during the crisis, depending on their business model. As introduced in [32], the economic crisis will unfold in several phases with different characteristics. Those sectors of the economy that are more vulnerable to supply chain interruptions and short-term collapses of consumer demand should be more affected in the early phase of the crisis than other sectors:

*H*2: *Corona-related risk factors in 10-K reports differ between industries.*

To test the hypotheses, we use web-mining techniques to collect data from SEC 10-K reports and conduct text mining to extract information relevant for the individual hypotheses, as outlined in the next section. Section 4 presents the results of the analysis.

3 | METHODS

The 10-K SEC filings, as legal reports to publicly communicate corporate risks and financial statements, provide a valuable and innovative text data source for risk assessment.[5]

We use a web scraper to collect all 10-K filings published since 30/01/2020 from the SEC's "Electronic Data Gathering, Analysis and Retrieval System" (EDGAR) database. The filing documents contain the company names and central index keys (CIK) as unique firm identifiers and the Standard Industry Classifier (SIC) that allow linking and comparing the filings data with individual business and industry data. Each 10-K filing contains a Risk Factor section (Item 1A) under which the reporting company is obliged to disclose all types of risks their business might be facing to adequately warn investors. Companies are required to use "plain English" in describing these risk factors, avoiding overly technical jargon that would be difficult for a layperson to follow. The text of each 10-k report builds the document by which a company is represented in our analysis. The entire text is set to lower case for further analysis, before the occurrence of the main two keyword tokens, *'corona'* and *'covid'*, is counted. Any word containing one of the two main tokens is likewise counted. Hereby the measure 'corona-keyword count per report' is created. Similarly, each company that reports one of the keywords at least once is included in the 'share of corona-mentioning firms per industry'.

After text pre-processing, we apply different Natural Language Processing tools to analyse the reports. Different sectors are facing different challenges, therefore companies are reporting about different corona-related risks. We aim to capture these risk topics via a keyword search on predefined topics. In order to explore possible topics, we used Latent Dirichlet Allocation (LDA) for unsupervised topic modelling, similar to [33]. We only apply the topic model to corona-related paragraphs in the risk sections. We additionally examine the most frequent words and bi-grams in the documents. Using this exploratory analysis, we define a set of topics, which are specified by keywords. We then conduct a keyword search to count how much these terms are mentioned in the different industries in order to estimate the topic prevalence. The resulting topic heatmap (see section 4) reports the share of sentences per topic per 1,000 corona-containing sentences for the different industries. Moreover, we measure the sentiment of corona-related sentences via the share of negative words [34]. A more detailed description of the different methods can be found in the appendix A.

---

[5]The data source explored here is not meant to replace any established macro-economic statistics, but rather to provide a complementary alternative source of data to identify immediate industry-specific risk perceptions.

4 | PRELIMINARY RESULTS

The main results of our analysis are displayed in Figure 1 as well as in Figure 2.[6] The overall sentiment of the 10-K reports becomes more negatively sharply (Fig. 1B), closely related to the collapsed stock markets after 15th February (Fig. 1A), implying a forward-looking nature of the 10-K sentiment measure. The stock markets show a V-shaped behaviour with a rapid decrease in February in March and fast increase since then, while the negative text sentiment is characterised by an inverse V-shaped dynamic. Since its' peak in March, it slowly went down towards levels of 5 %, which correspond to the overall share of negative words in 10-K reports.

Differences in corona-related risk assessment by industries are displayed in Figure 2. Overall, From the end of January 2020, an increasing number of 10-K reports refers to the coronavirus (Fig. 2A), indicating an increasing awareness of COVID-19 as a potential economic risk. Not all sectors of the economy show a similar awareness of the potential business impacts of the pandemic (Fig. 2B), which provides evidence in favour of hypothesis $H2$. For example, while 70 % of the firms in manufacturing mention corona-related risks, only 43 % of the firms in the financial sector consider such risk factors. Other sectors that show a high awareness of corona-related risks are retail (0.8 %), or hospitality (69 %).[7].

The firms, moreover, differ substantially with regards to the intensity with which they discuss the potential impacts of the coronavirus to their businesses in the 10-K reports. As Fig. 2C shows, some firms, in particular in retail and manufacturing, mention terms related to the pandemic (*'corona'*, *'coronavirus'*, *'COVID'*) substantially more often than other firms.

The 10-K reports moreover provide information about specific types of risk perceptions. Fig. 2D provides a heatmap with important topics per sector. The rows represent five relevant topics (demand, finance, production, supply, and travel), which have been derived from unsupervised topic modelling and subsequent uni- and bi-gram search.[8] The cells of the heatmap are coloured according to the topic relevance, i. e. the number of topic-specific keywords per industry. Additionally, we have applied a hierarchical clustering algorithm on the data to identify related topics and sectors, which report about similar topics (not displayed in the figure).

The figure reveals that most firms either report demand and financial risk factors or production and supply chain risks.For example, supply chain problems represent the biggest reported risk component for firms in the retail industry, while manufacturing firms report both supply and production risks, as well as finance and potential demand risks. While reports of the mining industry consider demand-related issues as the biggest risk factor, other sectorsdo not report extensively about any of the specific risk types.

These findings support hypothesis $H2$ — corona-related risk factors reported in 10-K reports differ between industries with regard to occurrence and topical context. The reports are, thus, a valid data source to identify sectors that face particular risk factors in the current early phase of the crisis.

While a one-dimensional categorisation of risk assessments tends to over-simplify the crisis firms are facing, it allows to compare the different industries and to identify those parts of the economy, which currently report more or less severe effects due to the immediate economic consequences of the pandemic. The data-driven assessment

---

[6]At this point we want to highlight, again, that the results reported here are preliminary. With more data becoming available over the next weeks and, hence, with a refined methodology, we expect to be able to report more in-depth analyses and conclusive findings. These will be constantly updated on the online dashboard.

Moreover, we want to emphasise that all quantitative findings are build on data of U. S. firms. It remains to be seen in how far the findings could be extrapolated to other economies.

[7]In some of the industries, only very few firms have provided risk reports in the current observation period. In particular in these groups, the results are likely to change as more data become available.

[8]Details of the text mining approach and the final keywords that define the topics are provided in the appendix.

reveals, in particular, that manufacturing and retail are among the industries that report to be most vulnerable to the changing economic environment. Not only decreasing consumer demand but particularly problems along the supply chain mark substantial risk factors for those industries. On the other hand, it is mostly firms from the information and communication services sector, which are less dependent on the physical transport of goods, that are currently reporting fewer risks.

The extent to which the industries are affected by corona-related business risks is likely to change over time. Accordingly, the risk categorisation presented here is a static measure that needs constant updating and refinement as the crisis unfolds. To do this, the study is supplemented by an online dashboard, which tracks the main findings with respect to the risk categorisation over time. A static visualisation of the tracking tool is displayed in Fig. 3. Due to limited data availability, the time-tracking could, so far, only be conducted for the largest industries; all other industries are aggregated in 'other industries'. The *CoRisk*-Index is a compound measure (i. e. geometric mean) of the share of firms that have reported corona-related risks (see Fig. 2B) the average number of corona-keywords per report (see Fig. 2C) and the industry-specific text negativity (see Fig. 1A), aggregated weekly. The different industries can be dinstiguished by their perceived risks they are reporting.With more reports being released over time, we will update and refine the *CoRisk*-Index to allow for more fine-grained industry-specific analyses. The index starts with zero, indicating no corona-related business risks in the reports. We will track the *CoRisk*-Index over the course of the corona crisis until the index reaches zero again.

5 | CONCLUSION

5.1 | Summary

As the COVID-19 pandemic unfolds, governments are attempting to reduce contagion rates at the expense of personal freedom and negative economic effects. Sizeable cyclical and fiscal policy packages are prepared in order to counterbalance the dooming global economic downturn. In order to ensure an effective usage of public crisis spending, it is paramount to understand in detail which industries are most affected by the pandemic and currently most in need of support. This study introduces a data-mining approach to measure the reported business risks induced by the current COVID-19 pandemic. We examine company risk reports filed to the U.S. Securities and Exchange Commission (SEC). Harnessing this data set enables a real-time analysis of potential risk factors. Preliminary findings show that the companies' risk awareness is preempting stock market developments. While stock prices typically condense the market's multiple signals, the 10-K reports allow to isolates the company risk perceptions associated to the COVID-19 outbreak. Moreover, this risk awareness differs substantially between industries, both in magnitude as well as in nature. Based on natural language processing techniques, we can identify specific corona-related risk-topics and their relevance for the different industries: supply chain- and production related issues seem to be mostly relevant for retail and manufacturing, while several industries have reported demand- or finance related risk factors. We summarise the corona-related business risk perceptions per industry in a compound index published online.[9] The *CoRisk*-Index, which will be updated constantly over the course of the crisis, provides an up-to-date database to identify the industries that report most substantial risk factors in the different phases of the unfolding economic crisis. The online dashboard can provide a data source to inform economic models and to provide empirical information that could help to assess potential policies, which aim to effectively target financial support and to mitigate the economic shocks of the current crisis.

Governments are eager to counterbalance the dooming global economic crisis induced by the COVID-19 pandemic

---

[9] http://oxford.berlin/CoRisk

with cyclical and fiscal policy packages of enormous volumes. Democratic accountability demands this public crisis stimulus to be spend as effectively as possible. Our data-driven analysis of the 10-K SEC filings provides an alternative data source, which could be used to calibrate macroeconomic models and thus help identifying industries that are likely to be most severely affected by COVID-19. In particular the manufacturing and retail industries are currently reporting most severe effects from the sharply decreasing consumer demand and interruptions of the supply chain, and might face substantial problems through the unfolding of the crisis.[10]

Other parts of the economy, in particular information-processing service industries, are currently reporting less severe corona-related risks. However, as the shock will transmit throughout the tightly inter-connected economy over time, these industries are likely to face larger challenges at a later stage in time. Nonetheless, at the current stage, their core businesses seem to be less directly affected by supply chain interruptions and collapsing consumer demand than other industries.

5.2 | Methodological Limitations

Our approach is based on the risk assessments in the U. S. Securities and Exchange Commission (SEC) filing reports. Thus, the value of the approach relies crucially on company self-reporting. While the firms are unlikely to provide a risk prediction with a high forecasting accuracy, it might still be worth exploring the reports as alternative data source to measure risk perceptions. As the reports serve as legal and insurance requirements against financial risks, but also as a basis for investment decisions of investors, companies are implicitly encouraged to neither over- nor understate the risks they are facing. Nevertheless, our results are limited to this self-assessment. As many of the implications of the corona crisis are still uncertain, our approach thus reflects a way to approximate potential implications on current estimations of experts in the different sectors, represented by the companies, and does not include risks that are unforeseeable for themselves at a given point in time. Moreover, the data to assess the precision of these estimations does not exist yet, as there has not been a pandemic with comparable global economic consequences in recent time. In the future, we will evaluate our results by looking at employment data in different sectors. Moreover, the pandemic continues to spread and will soon affect all industry sectors and all countries of the world. As more and more companies will report on related risks, the count of corona mentions alone will lose its information-value as a measure to differentiate between endangered industry sectors. Then, more granular measures, such as the identified topic categories, will become more important to distinguish between different natures of risks.

The exploration of alternative data sources that are meant to complement or now-cast established economic statistics always comes with uncertainty. For example, it is not yet clear how the short-term risks described by the different industries will translate into long-term economic outcomes, such as bankruptcies. Nonetheless, we believe that the reports could be a reliable source of empirical information about the issues faced by different industries in the current situation, and they might be used to inform forecasting models on industry-specific economic effects of the crisis, as they help fill a data gap. Models that incorporate alternative data sources such as the one presented here could then be beneficial for developing economic support packages that are currently provided by governments.

All technical methods serve the higher purpose of providing timely and comprehensible insights into the industry-specific effects caused by the global outbreak of the coronavirus. To mitigate susceptibility to errors and increase reproducibility, we mostly draw from more basic technical methods. This can be seen in the discovery of Corona-

---

[10]At this point, we want to emphasise again that the findings are preliminary and might change over time, as more data becomes available and as the research methodology is refined. Moreover, the risk categorisation presented in this study is not meant as a direct policy recommendation. Instead, the tool is meant to provide an empirical source of information for real-time tracking of industry-specific risk perceptions, which can help to inform policy makers in times of a fast-evolving economic crisis.

relevant keywords which is based on the matching of regular expressions to avoid error-prone text pre-processing. Reduction of technical complexity, however, comes at the cost of diminished modelling fidelity and potential accuracy of results. For instance, the sentiment classifier has not been fine-tuned on text snippets discussing financial risks in particular. Further, it may be questioned whether a generalist sentiment score yields a reliable measure for the assessment of risks. Additionally, the LDA-based topic modelling approach lacks in interpretability and robustness, and we therefore only use it for exploration until now. The topics and keywords we use for estimating topic prevalence are therefore hand-coded which limits the detection of topics to predefined terms.

In general, the findings presented here should be considered as preliminary. Sensitivity checks are needed to validate their robustness, and numerous extensions will help to better assess the potential value of the exploration of 10-K reports as a complementary data source to measure industry-specific risk factors.

We expect some of these limitations to be mitigated with more reports to be analysed in the near future. Adjusted results will be published on the online dashboard and in a refined version of the working paper.

5.3 | Future Work

The following impetus for future work builds directly on the identified methodological limitations and caveats mentioned throughout the paper. Rather than harnessing generalist sentiment analysis, future efforts should explore risk classifiers trained on labelled financial risk assessments and entity sentiment analysis [34, 35, 36, 37].

The robustness of the results needs to be checked in more detail. In particular, we will compare historical and unemployment data with risk measures (text negativity) extracted from 10-K reports to investigate the correlation between risk reports and overall economic trends. Moreover, we will weight the index by firm size and compare the composition of the index with the overall U.S. economy in order to investigate the representativeness of the sample.

Likewise, the content of other online platforms, such as Wikipedia, as a source for risk assessment could be explored [38, 39]. Disentangling international cooperation patterns [40, 41] and mining industry-specific key technologies [36] will be highly beneficial. This could allow an estimation of operational risk channels and risk spillovers propagating between industries and countries due to global supply chains and peer-tier co-operations. Considering more financial data could include credit risks (financial exposures) of companies. Besides the comparison with additional financial signals, we plan to include unemployment rates in our analysis as an exogenous variable to the SEC risk assessment.

Acknowledgements

CEPR PRESS

References

[1] Werner E, DeBonis M, Kane P, Wagner J. House leaders look to expedite $2.2 trillion relief package but face possibility that one GOP lawmaker may delay passage - The Washington Post. Washington Post 2020 Mar;`https://www.washingtonpost.com/business/2020/03/26/senate-trump-coronavirus-economic-stimulus-2-trillion/`.

[2] Zacharakis Z. Wirtschaftsmaßnahmen: Was das alles kostet. Die Zeit 2020 Mar;`https://www.zeit.de/wirtschaft/2020-03/wirtschaftsmassnahmen-corona-krise-hilfspakete-kosten`.

[3] del Rio-Chanona RM, Mealy P, Pichler A, Lafond F, Farmer D. Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective. arXiv preprint arXiv:200406759 2020;.

[4] Beland LP, Brodeur A, Wright T. The Short-Term Economic Consequences of COVID-19: Exposure to Disease, Remote Work and Government Response. IZA Institute of Labour Economics 2020;p. 92.

[5] Ludvigson SC, Ma S, Ng S. Covid19 and the Macroeconomic Effects of Costly Disasters. National Bureau of Economic Research; 2020.

[6] Wu JT, Leung K, Bushman M, Kishore N, Niehus R, Salazar PMd, et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nature Medicine 2020 Mar;p. 1–5. `https://www.nature.com/articles/s41591-020-0822-7`, publisher: Nature Publishing Group.

[7] Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Potential for global spread of a novel coronavirus from China. Journal of Travel Medicine 2020 Mar;27(2). `https://academic.oup.com/jtm/article/27/2/taaa011/5716260`, publisher: Oxford Academic.

[8] Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. London: Imperial College COVID-19 Response Team, March 2020;16.

[9] Fong MW, Gao H, Wong JY, Xiao J, Shiu EY, Ryu S, et al. Nonpharmaceutical Measures for Pandemic Influenza in Nonhealthcare Settings-Social Distancing Measures. Emerging infectious diseases 2020;26(5).

[10] Nicolaides C, Avraam D, Cueto-Felgueroso L, González MC, Juanes R. Hand-Hygiene Mitigation Strategies Against Global Disease Spreading through the Air Transportation Network. Risk Analysis 2019;n/a(n/a). `https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13438`, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/risa.13438.

[11] Lewis D, Mertens K, Stock JH. US Economic Activity During the Early Weeks of the SARS-Cov-2 Outbreak. National Bureau of Economic Research; 2020.

[12] Baldwin R, Mauro BWd. Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes. CEPR Press; 2020.

[13] Dorn F, Fuest C, Göttert M, Krolage C, Lautenbacher S, Link S, et al. Die volkswirtschaftlichen Kosten des Corona-Shutdown für Deutschland: Eine Szenarienrechnung. ifo Institut 2020;ifo Schnelldienst Vorabdruck. `http://www.ifo.de/publikationen/2020/aufsatz-zeitschrift/die-volkswirtschaftlichen-kosten-des-corona-shutdown`, library Catalog: www.ifo.de.

[14] Baldwin R, Mauro BWd. Economics in the Time of COVID-19. CEPR Press; 2020.

[15] Gopinath G. Limiting the economic fallout of the coronavirus with large targeted policies. In: Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes CEPR Press; 2020.

[16] Ramelli S, Wagner A. What the stock market tells us about the consequences of COVID-19. In: Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes CEPR Press; 2020.

[17] Huang Y, Lin C, Wang P, Xu Z. Saving China from the coronavirus and economic meltdown: Experiences and lessons. In: Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes CEPR Press; 2020.

[18] Keogh-Brown MR, Smith RD, Edmunds JW, Beutels P. The macroeconomic impact of pandemic influenza: estimates from models of the United Kingdom, France, Belgium and The Netherlands. The European Journal of Health Economics 2010 Dec;11(6):543–554. `https://doi.org/10.1007/s10198-009-0210-1`.

[19] Buetre B, Kim Y, Tran QT, Gunasekera D. Avian Influenza: Potential Economic Impact of a Pandemic on Australia. Australian Commodities: Forecasts and Issues 2006;13(2):351. `https://search.informit.com.au/documentSummary;dn=079724205036869;res=IELAPA`, publisher: Australian Bureau of Agricultural and Resource Economics.

[20] Rassy D, Smith RD. The economic impact of H1N1 on Mexico's tourist and pork sectors. Health Economics 2013;22(7):824–834. `https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.2862`, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.2862.

[21] Hatchett RJ, Mecher CE, Lipsitch M. Public health interventions and epidemic intensity during the 1918 influenza pandemic. Proceedings of the National Academy of Sciences 2007;104(18):7582–7587. Publisher: National Acad Sciences.

[22] Morse SS. Pandemic influenza: studying the lessons of history. Proceedings of the National Academy of Sciences 2007;104(18):7313–7314. Publisher: National Acad Sciences.

[23] Barro RJ, Ursua JF, Weng J. The Coronavirus and the Great Influenza Epidemic - Lessons from the 'Spanish Flu' for the Coronavirus's Potential Effects on Mortality and Economic Activity. Rochester, NY: Social Science Research Network; 2020.

[24] Karlsson M, Nilsson T, Pichler S. The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: An investigation into the consequences of an extraordinary mortality shock. Journal of Health Economics 2014 Jul;36:1–19. `http://www.sciencedirect.com/science/article/pii/S0167629614000344`.

[25] Bogataj D, Bogataj M. Measuring the supply chain risk and vulnerability in frequency space. International Journal of Production Economics 2007 Jul;108(1):291–301. `http://www.sciencedirect.com/science/article/pii/S0925527306003264`.

[26] Atalay E, Hortacsu A, Roberts J, Syverson C. Network structure of production. Proceedings of the National Academy of Sciences 2011;108(13):5199–5202. Publisher: National Acad Sciences.

[27] Acemoglu D, Carvalho VM, Ozdaglar A, Tahbaz-Salehi A. The network origins of aggregate fluctuations. Econometrica 2012;80(5):1977–2016. Publisher: Wiley Online Library.

[28] Aven T. The risk concept—historical and recent development trends. Reliability Engineering & System Safety 2012 Mar;99:33–44. `http://www.sciencedirect.com/science/article/pii/S0951832011002584`.

[29] Aven T, Renn O. Some foundational issues related to risk governance and different types of risks. Journal of Risk Research 2019 Mar;0(0):1–14. `https://doi.org/10.1080/13669877.2019.1569099`, publisher: Routledge _eprint: https://doi.org/10.1080/13669877.2019.1569099.

[30] Richman LD, Bakst DS, Gray RF, Hermsen ML, Pinedo AT, Schuette DA. SEC adopts rules to modernize and simplify disclosure. Journal of Investment Compliance 2019 Jan;20(2):1–8. `https://doi.org/10.1108/JOIC-04-2019-0022`, publisher: Emerald Publishing Limited.

[31] Kogan S, Levin D, Routledge BR, Sagi JS, Smith NA. Predicting risk from financial reports with regression. In: Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics on - NAACL '09 Boulder, Colorado: Association for Computational Linguistics; 2009. p. 272. `http://portal.acm.org/citation.cfm?doid=1620754.1620794`.

[32] Bénassy-Quéré A, Marimon R, Pisani-Ferry J, Reichlin L, Schoenmaker D, Mauro BWd. COVID-19: Europe needs a catastrophe relief plan. In: Mitigating the COVID Economic Crisis: Act Fast and Do Whatever It Takes CEPR Press; 2020.

[33] Dyer T, Lang M, Stice-Lawrence L. The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. Journal of Accounting and Economics 2017 Nov;64(2):221–245. `http://www.sciencedirect.com/science/article/pii/S0165410117300484`.

[34] Loughran T, Mcdonald B. When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. The Journal of Finance 2011;p. 46.

[35] Falck F, Marstaller J, Stoehr N, Maucher S, Ren J, Thalhammer A, et al. Measuring Proximity Between Newspapers and Political Parties: The Sentiment Political Compass. Policy & Internet 2019;`https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.222`.

[36] Stoehr N, Braesemann F, Frommelt M, Zhou S. Mining the Automotive Industry: A Network Analysis of Corporate Positioning and Technological Trends. In: Complex Networks XI Springer Proceedings in Complexity, Cham: Springer International Publishing; 2020. p. 297–308. `https://doi.org/10.1007/978-3-030-40943-2_25`.

[37] Darius P, Stephany F. "Hashjacking" the Debate: Polarisation Strategies of Germany's Political Far-Right on Twitter. In: Weber I, Darwish KM, Wagner C, Zagheni E, Nelson L, Aref S, et al., editors. Social Informatics Lecture Notes in Computer Science, Cham: Springer International Publishing; 2019. p. 298–308.

[38] Stephany F, Braesemann F. An exploration of wikipedia data as a measure of regional knowledge distribution. In: International Conference on Social Informatics Springer; 2017. p. 31–40.

[39] Stephany F, Braesemann F, Graham M. Coding together-coding alone: The role of trust in collaborative programming. Information, Communication and Society 2019;Publisher: SocArXiv.

[40] Braesemann F, Stoehr N, Graham M. Global networks in collaborative programming. Regional Studies, Regional Science 2019;6(1):371–373. `https://doi.org/10.1080/21681376.2019.1588155`.

[41] Stoehr N, Yilmaz E, Brockschmidt M, Stuehmer J. Disentangling Interpretable Generative Parameters of Random and Real-World Graphs. arXiv:191005639 2019;NeurIPS, Workshop on Graph Representation Learning. `http://arxiv.org/abs/1910.05639`.

[42] Rank and Filed, SEC filings for humans;. `http://rankandfiled.com/#/`.

[43] Aroussi R, yfinance: Yahoo! Finance market data downloader;. `https://github.com/ranaroussi/yfinance`.

[44] TextBlob: Simplified Text Processing — TextBlob 0.15.2 documentation;. `https://textblob.readthedocs.io/en/dev/index.html`.

[45] Blei DM. Probabilistic topic models. Communications of the ACM 2012;55(4):77–84. Publisher: ACM New York, NY, USA.

A | METHODOLOGICAL AND TECHNICAL APPENDIX

A.1 | Scraping of 10-K SEC Filings

We use web crawlers and scrapers written in the general programming language Python for finding and extracting the relevant risk section in the 10k reports. Via the SEC online interface[11] past SEC filings are queried since the first emergence of 'coronavirus' in a report on January 30th 2020. A web crawler allows to find the relevant risk factor section in the 10-K reports and stores each paragraph for later text processing and analysis.

---

[11] `https://searchwww.sec.gov/EDGARFSClient/jsp/EDGAR_MainAccess.jsp`

A.2 | Collection of Stock Market Data

The collection of stock market data follows the incentive to provide reference data for the company SEC filings. In both, SEC filings and financial markets, each company is assigned a unique Central Index Key (CIK) and a ticker number. Based on the CIK, we identified a list of 13,737 companies using the data registry provided by Rank and Filed [42]. This list was then filtered by the list of companies of which we obtained SEC filings. Using the Yahoo Finance API, we finally retrieved stock prices for all remaining companies in the list [43].

A.3 | Discovery of Relevant Keywords

Since this study examines the attention attributed to COVID-19 in the SEC filings, the discovery mechanism of relevant COVID-19 mentions is of central importance. To mitigate susceptibility to errors due to word splitting, stemming and other text preprocessing, we decided for the most simple approach based on the matching of regular expressions. We scanned the reports for the two relatively unambiguous terms "corona" and "covid", also accounting for "coronavirus" and "covid-19" without duplication.

A.4 | Sentiment Analysis on 10-K SEC Filings

Sentiment analysis on the SEC filings serves the purpose of a more contextualised semantic understanding of the risk assessment concerning COVID-19. We selected the paragraph of the risk report, with the highest number of coronavirus mentions and calculated the sentiment based on the TextBlob API [44] using the code developed in [35],

A.5 | Topic Modelling on 10-K SEC Filings

We use unsupervised learning techniques to identify the main topics that companies mention when describing coronavirus-related risks. Latent Dirichlet Allocation (LDA) is a Bayesian computational linguistic technique that identifies the latent topics in a corpus of documents [45]. This statistical model falls into the category of generative probabilistic modelling: a generative process which defines a joint probability distribution over the observed random variable, i.e. the words of the documents, and the hidden random variables, i.e. the topic structure. In other words, LDA uses the probability of words that co-occur within documents to identify sets of topics and their associated words [33]. The number of topics has to be defined in advance. LDA is a frequently used technique to identify main topics in a corpus. Nevertheless, the interpretation of these topics can sometimes be difficult. We thus perform LDA for explorative purposes in our research only. By additionally exploring the most common words and bi-grams we then define topics and the defining keywords by hand. We then detect how often companies mention these keywords in the corona related risk sections to estimate how important the topics are in different sectors. In particular, we perform the following steps:

Sample restriction

Referring to section A.3, we filter all sentences from the risk sections that mention either "corona" and "covid", thereby also accounting for "coronavirus" and "covid-19".

Text preparation

Before we train the LDA model we prepare the documents to achieve better performance of the method. We remove all common English stopwords, which are frequent words such as "is," "the," and "and" as well as those words which

appear in at least 80% of the documents. These words are not useful in classifying topics as they are too frequent and therefore decrease performance. Moreover, we delete all words that do not occur in at least 2 documents.

Topic modelling with LDA

We turn the documents into numerical "Bag of words" feature vectors, disregarding word order. We then use LDA to extract the topic structure. Like any unsupervised topic model, this requires setting the number of topics a priori. We selected this key parameter based on semantic coherence, evaluating a range of 2 to 8 topics leading to a final model of 4 topics. The top 10 terms of each topic are displayed in Table 2.

Topic keyword frequencies

The derived topics give a good insight into the general narratives of risks used in the documents. Nevertheless, they are hard to interpret, as early corona-related risk reports are still generic in that various risk factors are covered. We make use of our insights from the topic modelling to define 5 main topics defined by keywords, displayed in Table 1. The choice of keywords is additionally informed by clustering the most frequently used words and bi-grams in the documents. We then measure the frequency of these keywords per topic per industry in all of the documents to get more insights into the nature of risks different sectors are facing due to the pandemic.

TABLE 1 Topics and keywords

Topic	Keywords
Production	business operation, business disruption, product, work stoppage, labor disruption, labor, work, manufacturing operation, labor shortage, employee productivity, product development, business activity
Supply	manufacturing facility, manufacture facility, contract manufacturer, service provider, logistic provider, supply disruption, party manufacturer, supply disruption, facility, supply, transportation delay, delivery delay, supplier, business partner, supply chain, material shortage
Travel	air travel, travel, travel restriction, airline industry, travel disruption
Demand	store closure, distribution channel, market condition, consumer spend, market acceptance, consumer confidence, consumer demand, consume, store, customer, store traffic
Finance	operating result, cash flow, stock price, estate value, credit availability, performance problem

TABLE 2 LDA with 4 topics.

Topic number	Top 10 words
Topic 0:	impact extent including outbreak uncertain future highly results developments depend
Topic 1:	operations including health outbreak business supply products economic public result
Topic 2:	outbreak spread countries impact including china business potential economic government
Topic 3:	china outbreak novel covid adversely wuhan strain business december recent