Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig

Mathematical modelling and empirical data analysis of the Covid-19 pandemic

(revised version: July 2020)

by

Hoang Duc Luu and Jürgen Jost

Preprint no.: 80 2020

Mathematical modelling and empirical data analysis of the Covid-19 pandemic

Luu Hoang Duc [*], Jürgen Jost [†]

Abstract

The SIR model is the basic mathematical model for epidemics, but it needs some modification to capture the dynamics of the current Covid-19 pandemic. Here, we consider contact rates that depend negatively on the total number Γ of infections, as a result of the social distancing measures. Under general assumptions, the recovery and death rates are proved to be increasing functions of Γ, as witnessed in empirical data. Population structure is another issue, for instance concerning the contact number distribution and age structure of the population.
We develop and describe such models and show how the coefficients can be estimated and what the effects of delays are. We find that a simple linear regression is adequate for modelling the decay of the epidemic. We also discuss on the possibility of extending the model taking into account the time delays corresponding to the incubation period and the duration of the disease, which should be considered as random variables.

1 Introduction

In the current Covid-19 pandemic, it is important to understand the available models for the spreading of epidemics, to identify their conceptual and/or empirical shortcomings, and to develop new ones that overcome those problems. These should enable a more accurate modelling of the pandemic and help to assess the effects of various policies aimed at constraining it.
The mathematical basis of epidemic models is the classical SIR model [16] that models the dynamics between four subgroups, the susceptible, the infected, the recovered and the deceased individuals. Susceptible individuals can get infected, depending on the contact rate with infected ones, and infected individuals in turn will recover or die. While this model provides the main dynamical insights, it needs of course be enriched to capture real epidemics. Important effects that have to be taken into account are time delays caused by incubation periods, variations of the model coefficients, population structures and network effects. The coefficients may in particular vary as the result of policies designed and implemented to constrain the spreading of the epidemic. The most important one here is the contact rate between infected and susceptible individuals. Thus, one may isolate infected individuals, but since such individuals may be asymptomatic or for some other reasons not diagnosed, one has to take the general contact rate in the population as a proxy or move to a more finegrained analysis that takes the population structure or the network structure of social contacts into account. In this contribution, we develop and describe such models and show how the coefficients can be estimated and what the effects of delays are. A simple linear regression is displayed that models the decay of the epidemic.

[*]Luu H. Duc is with the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany, & Institute of Mathematics, Viet Nam Academy of Science and Technology, Viet Nam, duc.luu@mis.mpg.de, lhduc@math.ac.vn

[†]Jürgen Jost is with the Max-Planck-Institute for Mathematics in the Sciences, Leipzig, Germany & the Santa Fe Institute, Santa Fe, USA, jjost@mis.mpg.de.

2 Variational SIR model with containment efficiency and control strategies

We start by describing the classical SIR model for epidemic spreading [16], although, as we shall see, the model cannot be taken as such to model the current pandemic, but will need some modifications. We look at the following quantities that depend on time t.

- S_t: the number of susceptibles,
- I_t: the number of active infected cases,
- D_t: the number of death cases,
- R_t: the number of recovered cases,
- Γ_t: the total infections, as reported daily. Then $\Gamma_t = I_t + D_t + R_t$.
- N_t: the total population. Then $N_t = S_t + \Gamma_t$ (neglecting normal births and deaths).

We assume that we are at a stage of the epidemic where $D_t \ll N_t$ and without vital dynamics, so that we may replace N_t by a constant population size N. We may therefore normalize the other quantitities and simply write S, I, R, D, Γ for the normalized data $\frac{S}{N}, \frac{I}{N}, \frac{R}{N}, \frac{D}{N}, \frac{\Gamma}{N}$. The classical SIR model (here augmented by the death rate) for normalized data then is

$$\frac{dS_t}{dt} = -\beta S_t I_t \tag{2.1}$$

$$\frac{dI_t}{dt} = \beta S_t I_t - \mu I_t - \gamma I_t \tag{2.2}$$

$$\frac{dR_t}{dt} = \gamma I_t, \quad \frac{dD_t}{dt} = \mu I_t. \tag{2.3}$$

where linear relations are assumed between the numbers of new deaths, recovered and active infected cases. It is assumed that recovered patients are immune and therefore will not become susceptible again. We call parameters β, μ, γ respectively the contact rate, the death rate and the recovery rate.

Several variational models have been proposed from previous studies, for e.g. by considering $\beta \frac{S_t I_t}{N_t}$ in the right hand side of (2.1), which is βSI, as a functional response $F(S, I)$ where

- for the Holling type II functional response (see [13]) $F(S, I) = \frac{\beta SI}{m_1 + S}$;
- for the nonlinear functional response (see [7], [27], [33]): $F(S, I) = \frac{\beta SI^l}{1 + mI^h}$;
- for the Beddington-DeAngelis functional response (see [2]): $F(S, I) = \frac{\beta SI}{1 + m_1 S + m_2 I}$.

To keep the analysis simple, we keep equations (2.1)-(2.3) in the normalized data unchanged, but accept the possibility that the parameters β, μ, γ can vary in time, and in fact, can be modified by appropriate policies. For the Covid-19 disease, empirical studies [4], [17], [26], [22] and also the analysis in Section 3 suggest that for simplicity, we may keep the recovery rate γ constant at γ_0 (neglecting the time lag and the random fluctuation effects). The contact rate β and the death rate μ, however, should not be assumed constant since β is the average contact rate of the whole population which might change during the course of time, while μ depends on the public health care system, on the age structure of the population and also on how the disease spreads to the old age group. In fact, it is precisely the purpose of the social distancing policies and the improvement of the public health care system to decrease β and μ. On the one hand, we would like to work with

a time varying β_t which, among possible other factors, depends on the social distancing policies. On the other hand, we need to keep the model intrinsically closed.

Motivated by [11], we consider the contact rate β and the death rate μ as functions of the total infections, i.e. $\beta = \beta(\Gamma), \mu = \mu(\Gamma)$. Specifically, consider the general model for the normalized data

$$\begin{aligned} \frac{dS_t}{dt} &= -\beta_0 G(\Gamma_t) S_t I_t && (2.4)\\ \frac{dI_t}{dt} &= \beta_0 G(\Gamma_t) S_t I_t - \mu_0 P(\Gamma_t) I_t - \gamma_0 I_t && (2.5)\\ \frac{dR_t}{dt} &= \gamma_0 I_t, \quad \frac{dD_t}{dt} = \mu_0 P(\Gamma_t) I_t && (2.6) \end{aligned}$$

where $S + I + R + D = S + \Gamma = 1$ and $\beta_0, \mu_0, \gamma_0 > 0$ are constant. The contact function $G(\Gamma)$ is assumed to be differentiable and satisfy

$$G(\Gamma) > 0, \quad G(0) = 1, \quad G'(\Gamma) \leq 0. \tag{2.7}$$

Meanwhile the death function $P(\Gamma)$ is assumed to be differentiable such that

$$P(\Gamma) > 0, \quad P(0) = 1, \quad P'(\Gamma) \leq 0. \tag{2.8}$$

Denote by $\kappa_0 := \frac{\beta_0}{\mu_0 + \gamma_0}$ the basic reproduction number, it is well-known that for covid-19, $\kappa_0 \in (2, 3)$. By similar arguments as in [11], one derives from (2.5) and (2.6) with $S = 1 - \Gamma$ that

$$\frac{d\Gamma_t}{dt} = \beta_0 G(\Gamma_t)(1 - \Gamma_t) I_t \Rightarrow I_t = \frac{1}{\beta_0 G(\Gamma_t)(1 - \Gamma_t)} \frac{d\Gamma_t}{dt}. \tag{2.9}$$

Inserting (2.9) into (2.5) yields

$$\frac{dI_t}{dt} = \frac{d\Gamma_t}{dt} - \frac{\Big(\mu_0 P(\Gamma_t) + \gamma_0\Big)}{\beta_0 (1 - \Gamma_t) G(\Gamma_t)} \frac{d\Gamma_t}{dt},$$

which implies, after neglecting the time dependence and integrating both sides, that I and Γ satisfy the relation

$$I = I(\Gamma) = \Gamma - \frac{1}{\beta_0} \int \frac{\Big(\mu_0 P(\Gamma) + \gamma_0\Big) d\Gamma}{(1 - \Gamma) G(\Gamma)} + C \tag{2.10}$$

where the constant C is determined from the initial relation that $I(0) = 0$.

An advantage of the model (2.4)-(2.6) is that it possesses an asymptotic state of the total infections, as shown below.

Theorem 2.1 *Assume that $I(1) < 0$ and $\left(\frac{P}{G}\right)'(\Gamma) \geq 0$. Then there exists a final stationary state Γ^∞ of the total infections which is asymptotically stable.*

Proof: A direct computation using (2.7) shows that

$$\begin{aligned} I'(\Gamma) &= 1 - \frac{1}{\beta_0} \frac{\Big(\mu_0 P(\Gamma) + \gamma_0\Big)}{(1 - \Gamma) G(\Gamma)}; && (2.11)\\ I'(0) &= 1 - \frac{\mu_0 + \gamma_0}{\beta_0 G(0)} = 1 - \frac{1}{\kappa_0} > 0, \qquad I'(1) = -\infty; \\ I''(\Gamma) &= -\frac{1}{\beta_0} \frac{\mu_0 (1 - \Gamma)(P'(\Gamma) G(\Gamma) - P(\Gamma) G'(\Gamma)) - \gamma_0 (1 - \Gamma) G'(\Gamma) + G(\Gamma)(\mu_0 P(\Gamma) + \gamma_0)}{(1 - \Gamma)^2 G(\Gamma)^2} < 0, \end{aligned}$$

where the last inequality is due to (2.7), (2.8) and the assumption $\left(\frac{P}{G}\right)'(\Gamma) \geq 0$. Hence $I(\Gamma)$ is a concave function on $\Gamma \in [0,1]$ which attains the maximum at some $\Gamma^{\max} \in (0,1)$ such that $I'(\Gamma^{\max}) = 0$. If we assume further that $I(1) < 0$, then there exists a unique $\Gamma^{\infty} \in (\Gamma^{\max}, 1)$ such that $I(\Gamma^{\infty}) = 0$. As a consequence,

$$I(\Gamma) > 0, \quad \forall \Gamma \in (0, \Gamma^{\infty}) \qquad \text{and} \qquad I(\Gamma) < 0, \quad \forall \Gamma \in (\Gamma^{\infty}, 1). \tag{2.12}$$

Next, inserting (2.10) into (2.9), one derives an ODE for Γ in time

$$\frac{d\Gamma_t}{dt} = \beta_0 G(\Gamma_t)(1-\Gamma_t) I(\Gamma_t) \tag{2.13}$$

From the above arguments, system (2.13) has three equilibria $\{0, \Gamma^{\infty}, 1\}$. Using (2.12), one concludes that 0 and 1 are unstable fixed points while Γ^{∞} is the unique fixed point of system (2.13) that is asymptotically stable, and it represents the expected total infections (normalized) (w.r.t. the whole population) at the end of the pandemic.

□

The asymptotic state of the pandemic can be observed in Figure 1 for the case progression in selected countries worldwide.

Example 2.2 [11] chooses $G(\Gamma) = \frac{1}{1+\alpha\Gamma}$ and $P(\Gamma) \equiv 1$ on $[0,1]$ for $\alpha > 0$, which satisfies (2.7) and (2.8). Then the function $I(\Gamma)$ in (2.10) can be written in the explicit form

$$I(\Gamma) = \Gamma - \frac{1}{\kappa_0} \int \Big(\frac{1+\alpha}{1-\Gamma} - \alpha \Big) d\Gamma = (1 + \frac{\alpha}{\kappa_0})\Gamma + \frac{1+\alpha}{\kappa_0} \log(1-\Gamma). \tag{2.14}$$

In this case $I(0) = 0$ and $I(1) = -\infty$. A direct computation shows that

$$\Gamma^{\max} = \frac{\kappa_0 - 1}{\kappa_0 + \alpha}, \qquad h(\Gamma^{\infty}) := -\frac{\kappa_0 \Gamma^{\infty} + \log(1-\Gamma^{\infty})}{\Gamma^{\infty} + \log(1-\Gamma^{\infty})} = \alpha. \tag{2.15}$$

It follows from (2.15) that $\Gamma^{\max}$ and Γ^{∞} are decreasing functions of α. Hence to reduce Γ^{∞} and $\Gamma^{\max}$, it is important to increase the intensity α of the control function $\alpha\Gamma$.

Example 2.3 Empirical data from the Covid19 pandemic suggest to use the control function $G(\Gamma) = e^{-\alpha\Gamma}$ and $P(\Gamma) = e^{-\delta\Gamma}$ for $\Gamma \in [0,1]$ with $\alpha > \delta > 0$, which satisfies (2.7) and (2.8). Then function $I(\Gamma)$ in (2.10) has the form

$$I(\Gamma) = \Gamma - \frac{\mu_0}{\beta_0} \int \frac{e^{(\alpha-\delta)\Gamma} d\Gamma}{1-\Gamma} - \frac{\gamma_0}{\beta_0} \int \frac{e^{\alpha\Gamma} d\Gamma}{1-\Gamma} + C \tag{2.16}$$

where the constant C is determined by the initial condition $I(0) = 0$. A direct computation shows that

$$I'(\Gamma^{\max}) = 0 \Leftrightarrow h(\alpha, \Gamma^{\max}) := \frac{\mu_0 e^{(\alpha-\delta)\Gamma^{\max}} + \gamma_0 e^{\alpha\Gamma^{\max}}}{(1-\Gamma^{\max})} = \beta_0, \tag{2.17}$$

where $h(\alpha, \Gamma^{\max})$ is an increasing function of α and $\Gamma^{\max}$. As a result, the solution $\Gamma^{\max}$ of (2.17) is a decreasing function of α, and one can decrease the theoretical value $\Gamma^{\max}$ by increasing the intensity α of the control function $e^{\alpha\Gamma}$.

Example 2.4 Another possible candidate are the control and death functions of the form $G(\Gamma) = (1-\Gamma)^{\alpha}$, $P(\Gamma) = (1-\Gamma)^{\delta}$ for $\Gamma \in [0,1]$ and some $\alpha > \delta > 0$, which also satisfies (2.7) and (2.8). Like Example 2.2, the function $I(\Gamma)$ can also be solved explicitly as

$$\begin{aligned} I(\Gamma) &= \Gamma - \frac{1}{\beta_0} \int \Big[\mu_0 (1-\Gamma)^{\delta-\alpha-1} + \gamma_0 (1-\Gamma)^{-\alpha-1} \Big] d\Gamma + C \\ &= \Gamma + \frac{\mu_0}{(\alpha-\delta)\beta_0} \Big[1 - (1-\Gamma)^{\delta-\alpha} \Big] + \frac{\gamma_0}{\alpha\beta_0} \Big[1 - (1-\Gamma)^{-\alpha} \Big]. \end{aligned} \tag{2.18}$$

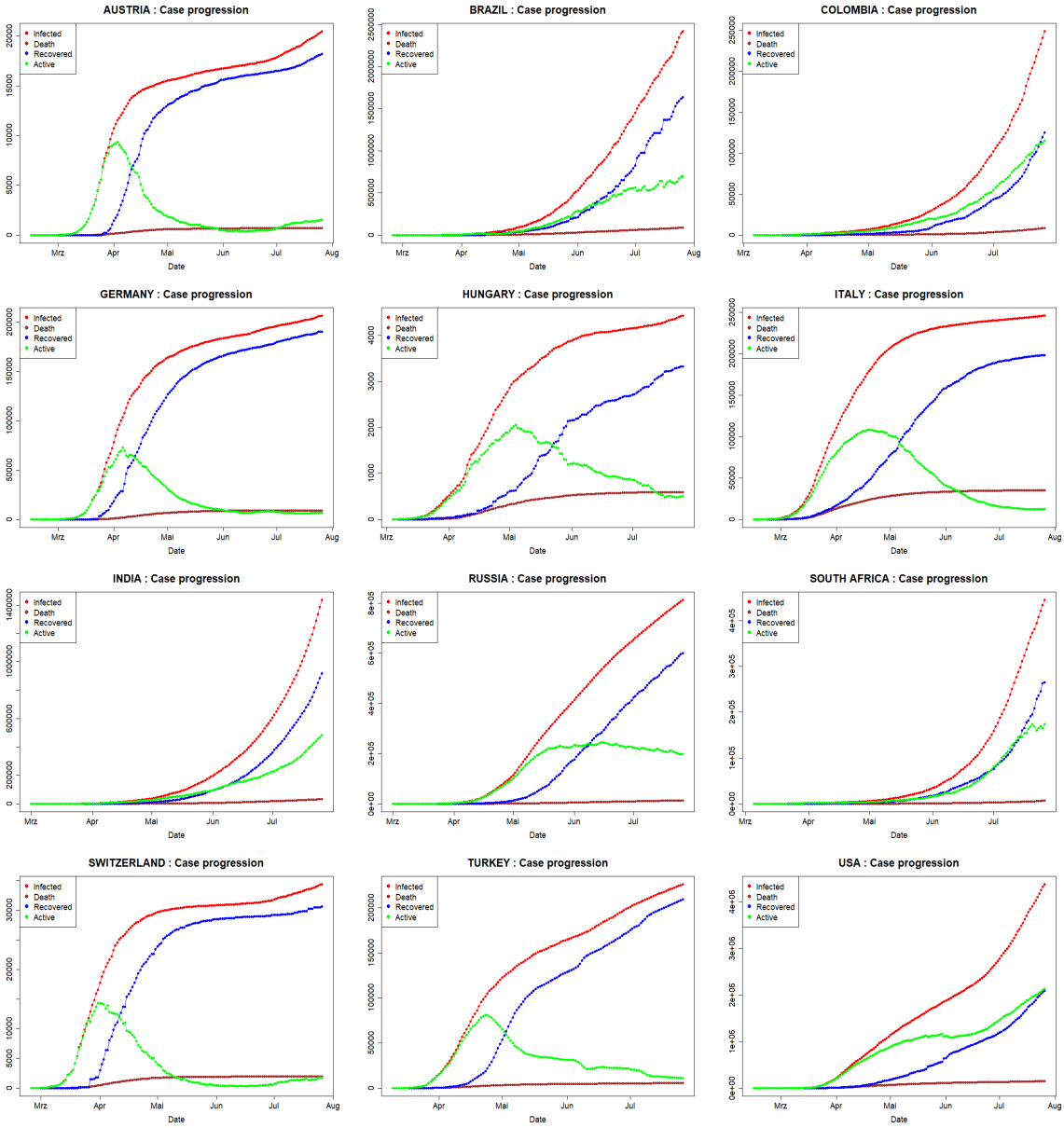

Figure 1: Case progression in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometers.

A direct computation shows that $I(1) = -\infty$ and

$$I'(\Gamma^{\max}) = 0 \Leftrightarrow h(\alpha, \Gamma^{\max}) := \mu_0(1-\Gamma^{\max})^{\delta-\alpha-1} + \gamma_0(1-\Gamma^{\max})^{\delta-\alpha-1} \tag{2.19}$$

which is an increasing function of α and Γmax. Hence $\Gamma^{\max}$ is a decreasing function of α, and an increase in the exponent α of the control function $(1-\Gamma)^{\alpha}$ leads to a decrease of $\Gamma^{\max}$.

2.1 Death and recovery ratio

In practice, there often is a lack of data reported from the beginning phase of the pandemic for the recovered cases, and hence the estimate for the Γ versus I relation cannot be directly checked. In

this scenarios, one can apply (2.6) and (2.9) to obtain

$$\frac{dD_t}{d\Gamma_t} = \frac{\mu_0}{\beta_0}\frac{P(\Gamma)}{(1-\Gamma_t)G(\Gamma_t)}, \qquad \frac{dR_t}{d\Gamma_t} = \frac{\gamma_0}{\beta_0}\frac{1}{(1-\Gamma_t)G(\Gamma_t)}. \tag{2.20}$$

Hence by a similar method as above, one can derive the equations for the $\Gamma - D$ and $\Gamma - R$ relations

$$D = D(\Gamma) = \frac{\mu_0}{\beta_0}\int \frac{P(\Gamma)d\Gamma}{(1-\Gamma)G(\Gamma)} + C, \qquad R = R(\Gamma) = \frac{\gamma_0}{\beta_0}\int \frac{d\Gamma}{(1-\Gamma)G(\Gamma)} + C \tag{2.21}$$

where C is determined by the initial conditions $D(0) = 0$ and $R(0) = 0$. Equation (2.20) can also be used to estimate the shape of the function $G(\Gamma)$ as

$$\log\Big\{(1-\Gamma_t)\Big(\frac{D_{t+1}-D_t}{\Gamma_{t+1}-\Gamma_t}\Big)\Big\} = \log\frac{\mu_0}{\beta_0} + \log\Big(\frac{P(\Gamma_t)}{G(\Gamma_t)}\Big). \tag{2.22}$$

Theorem 2.5 *Assume the equations* (2.20) *and* (2.21) *for the deceased and recovered groups. Then the recovery ratio* $\frac{R(\Gamma)}{\Gamma}$ *is an increasing function of* Γ. *If in addition* $\big(\frac{P}{G}\big)'(\Gamma) \geq 0$ *then the death ratio* $\frac{D(\Gamma)}{\Gamma}$ *is an increasing function of* Γ.

Proof: Since the two rations behave in the same way, it is enough to prove that $\Big[\frac{D(\Gamma)}{\Gamma}\Big]' > 0$ for $\Gamma \in (0,1)$. Observe that

$$\Big[\frac{D(\Gamma)}{\Gamma}\Big]' = \frac{D'(\Gamma)\Gamma - D(\Gamma)}{\Gamma^2} \quad \text{where} \tag{2.23}$$

$$\Big[D'(\Gamma)\Gamma - D(\Gamma)\Big]' = D''(\Gamma)\Gamma = \frac{\mu_0}{\beta_0}\Gamma\frac{\Big[(1-\Gamma)(P'(\Gamma)G(\Gamma) - P(\Gamma)G'(\Gamma)) + G(\Gamma)\Big]}{(1-\Gamma)^2G(\Gamma)^2} > 0 \tag{2.24}$$

for all $\Gamma \in [0,1)$, where the last equality is due to (2.20), the assumption $\big(\frac{P}{G}\big)'(\Gamma) \geq 0$ and the conditions (2.7), (2.8). From (2.24), the numerator in the right hand side of (2.23) is increasing in Γ, thus $D'(\Gamma)\Gamma - D(\Gamma) \geq D'(0)0 - D(0) = 0$. This shows that the death rate $\frac{D(\Gamma)}{\Gamma}$ is increasing in Γ. As a consequence, due to the fact that Γ_t is increasing in time, the time varying current death rate $\frac{D(\Gamma_t)}{\Gamma_t}$ is also increasing in time and will approach the limit $\frac{D(\Gamma^\infty)}{\Gamma^\infty}$ as Γ_t tends to Γ^∞.

□

Figure 2 and figure 3 show the time development of the current death and recovery ratios for the selected countries. While the recovery rate seems to increase over time in every country, the death rate shows an increasing trend most of the time until it reaches a top level, then stabilizes and gradually decreases, which indicates that there might be a change in the control functions P, G and leads to a violation of the assumption $\big(\frac{P}{G}\big)'(\Gamma) \geq 0$.

3 Estimating the contact rate

Since model (2.4)-(2.6) is for normalized data, to test the model, we use equation (2.9) in the discrete version for the original data to compute the contact rate

$$\beta(\Gamma_t) := \frac{\Gamma_{t+1} - \Gamma_t}{(1 - \frac{\Gamma_t}{N})I_t} = \beta_0 G\Big(\frac{\Gamma_t}{N}\Big) \tag{3.1}$$

As such, a constant value in the right hand side of (3.1) would show that there is no control effect and the contact rate would be constant as presented in the classical SIR model. With Covid-19, we observe, however, three phases in the development: first the contact rate looks quite volatile and

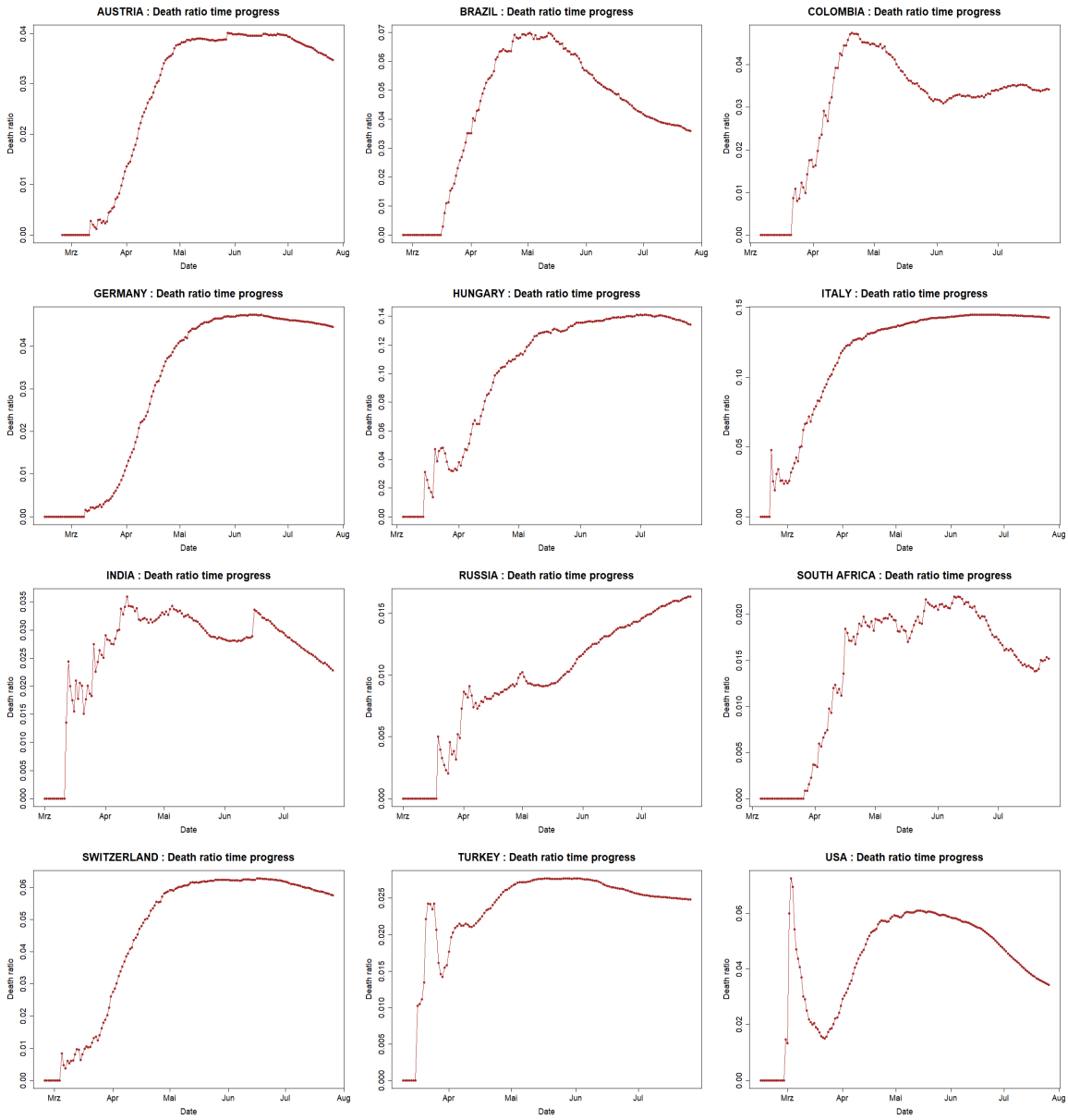

Figure 2: Current death ratio in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

stays in the high value area. In the second phase, it starts to decrease considerably. And finally, in the stable phase, the contact rate stays in a low value area and only fluctuates a little.

The way $\beta(\Gamma_t)$ decreases also varies from country to country and needs to be quantified. [11] suggests a model as in Example 2.2, in which the contact rate has the form $\beta(\Gamma_t) = \frac{\beta_0}{1+\alpha\Gamma_t}$. For checking this model, it is convenient to take the inverse of the contact rate and study its relation to the total infections Γ_t

$$\frac{1}{\beta(\Gamma_t)} = \frac{1}{\beta_0} + \frac{\alpha}{\beta_0}\Gamma_t. \tag{3.2}$$

As shown in Figure 4, the relation between Γ_t and the inverse contact rate $\frac{1}{\beta(\Gamma_t)}$ in (3.2) in the selected countries does not show exactly this linear dependence. In fact, the inverse contact rate

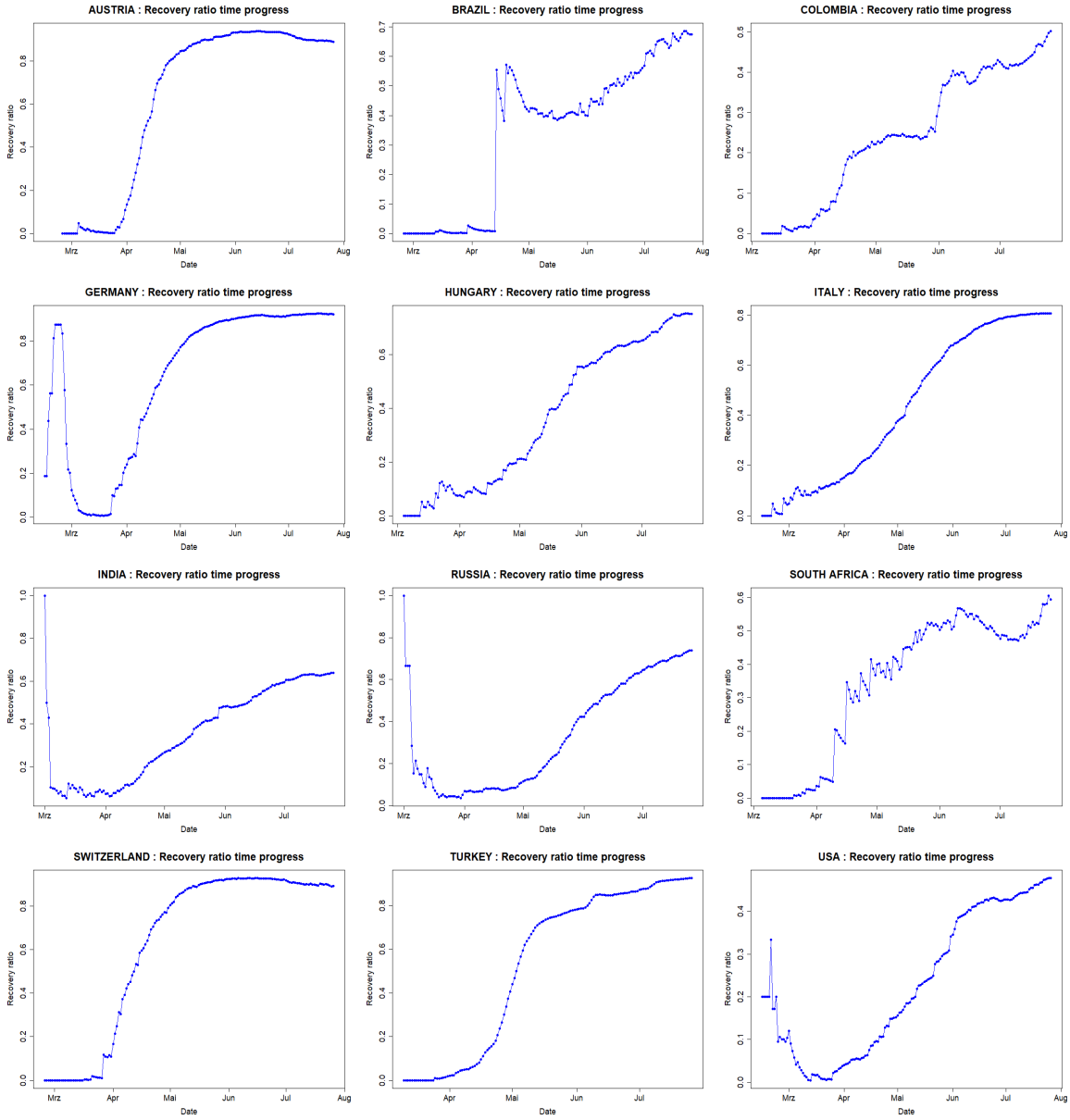

Figure 3: Current recovery ratio in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

seems to depend linearly on the total infections only in the first half period, while in the second half it rather grows exponentially.

On the other hand, Example 2.3 suggests a model where the contact rate has the form $\beta(\Gamma_t) = \beta_0 e^{-\alpha\Gamma_t}$ with $\alpha > 0$, which can be checked by considering the logarithmic scaled contact rate

$$\log \beta(\Gamma_t) = \log \beta_0 - \alpha\Gamma_t. \tag{3.3}$$

As shown in Figure 5, the linear relation in (3.3) is confirmed in the selected countries in the first and second phases. The volatility in Phase 3 can be explained as the technical issue that arises from taking the logarithm of small and rapidly changing new infections $\Gamma_{t+1} - \Gamma_t$, and from the fact that in reality, the social distancing policies were gradually relaxed in Phase 3 once the active cases decreased to a level that was considered safe for the public health system. In this scenario,

the control parameter α can be switched off to zero (see also [11]). In countries like Russia, we see $\beta(\Gamma_t)$ dropping like a piecewise constant function, while in others like Turkey or the US, the picture is more complex. There, $\beta(\Gamma_t)$ first drops exponentially, then stays with little change at the bottom before bouncing back to a higher level as a result of relaxing social distancing measures. In other countries like Brazil, Colombia, India and South Africa, the contact rate currently looks unchanged, indicating that there is no effective policy to stop the disease from spreading.

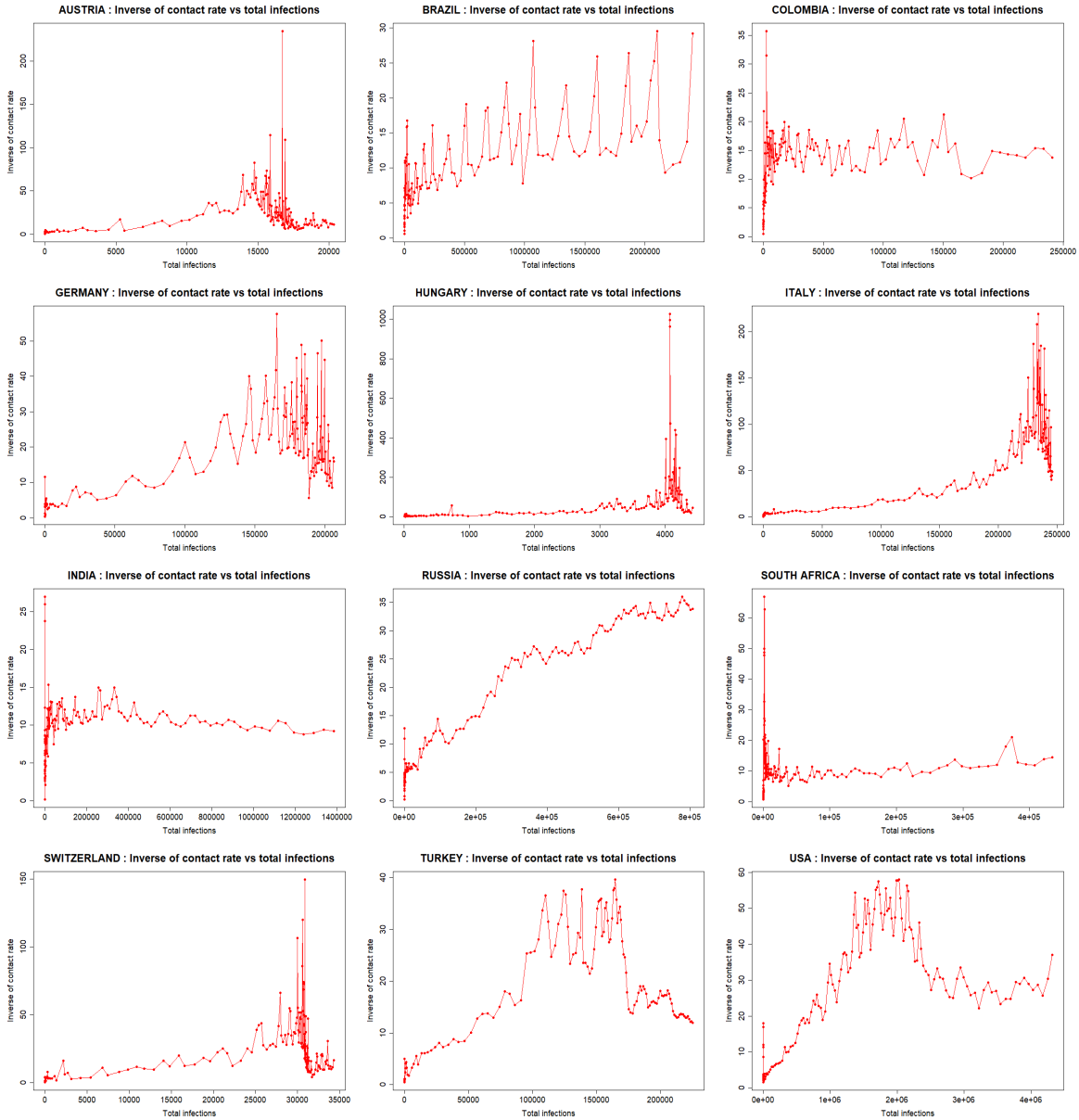

Figure 4: Inverse contact rate versus total infections in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

Remark 3.1 From the computations in (2.11) and (2.24), the assumption (2.7) that $G(\Gamma)$ is a decreasing function can be relaxed to the assumption that $G(\Gamma)(1-\Gamma)$ is a decreasing function of Γ. Since this quantity also contributes to the susceptible group S_t in the contact process with

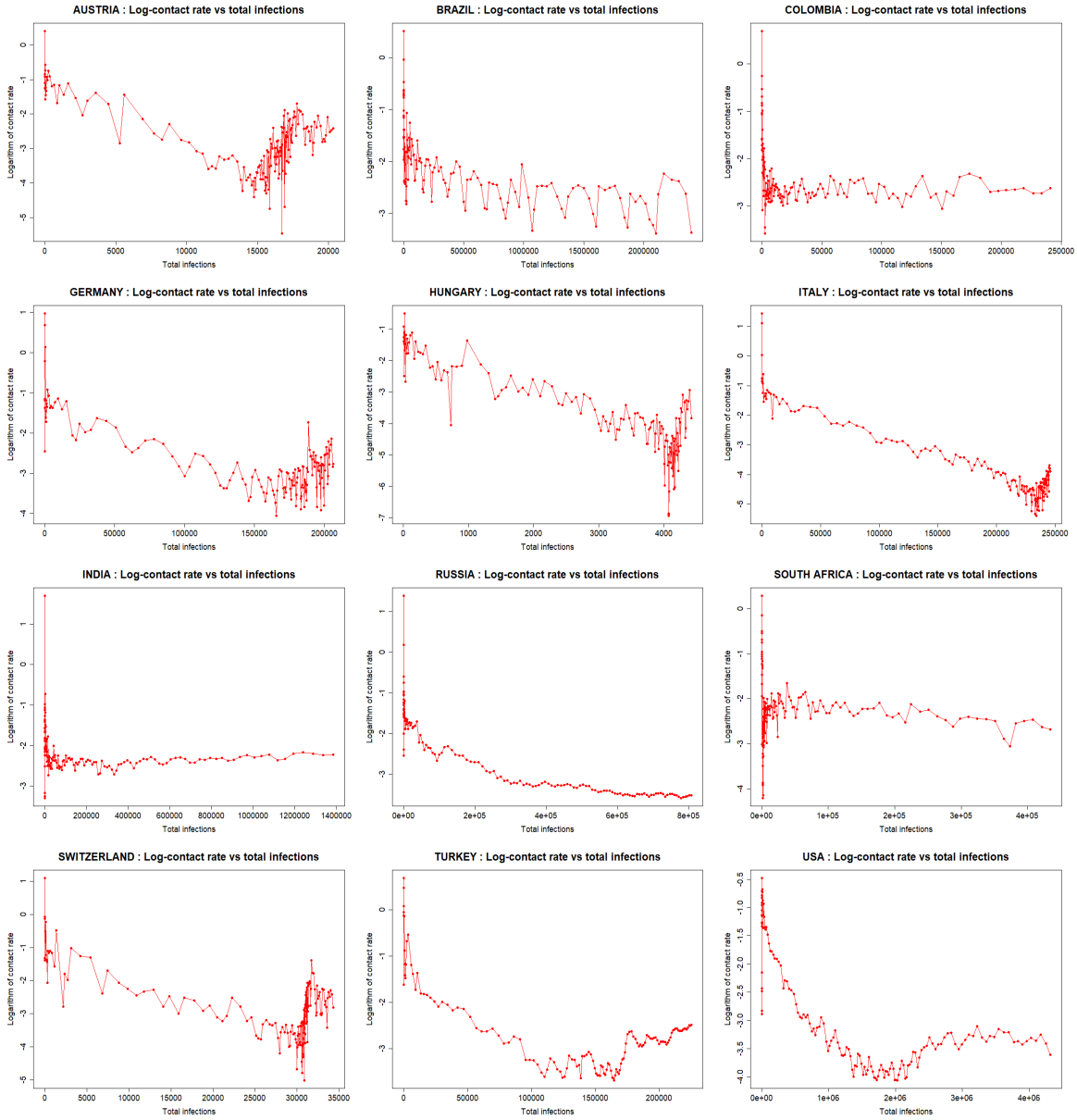

Figure 5: Contact rate (logarithmic) versus total infections in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

the active group I_t as shown in (2.4), in particular $1 - \Gamma_t = S_t$, one can in general assume that $G(\Gamma)(1-\Gamma) = f(S)$ is an increasing function of the susceptible group S. The generalized equation (2.9) then has the form

$$\frac{d\Gamma_t}{dt} = \beta_0 f(S_t) I_t. \tag{3.4}$$

We may also look at its discrete version

$$\log\left(\frac{\Gamma_{t+1} - \Gamma_t}{I_t}\right) = \log(\beta_0) + \log(f(S_t)). \tag{3.5}$$

Figure 6 presents the relation between the left hand side of (3.5) and the normalized susceptible numbers in a logarithmic scale. It is not surprising that the relation looks almost the same as that

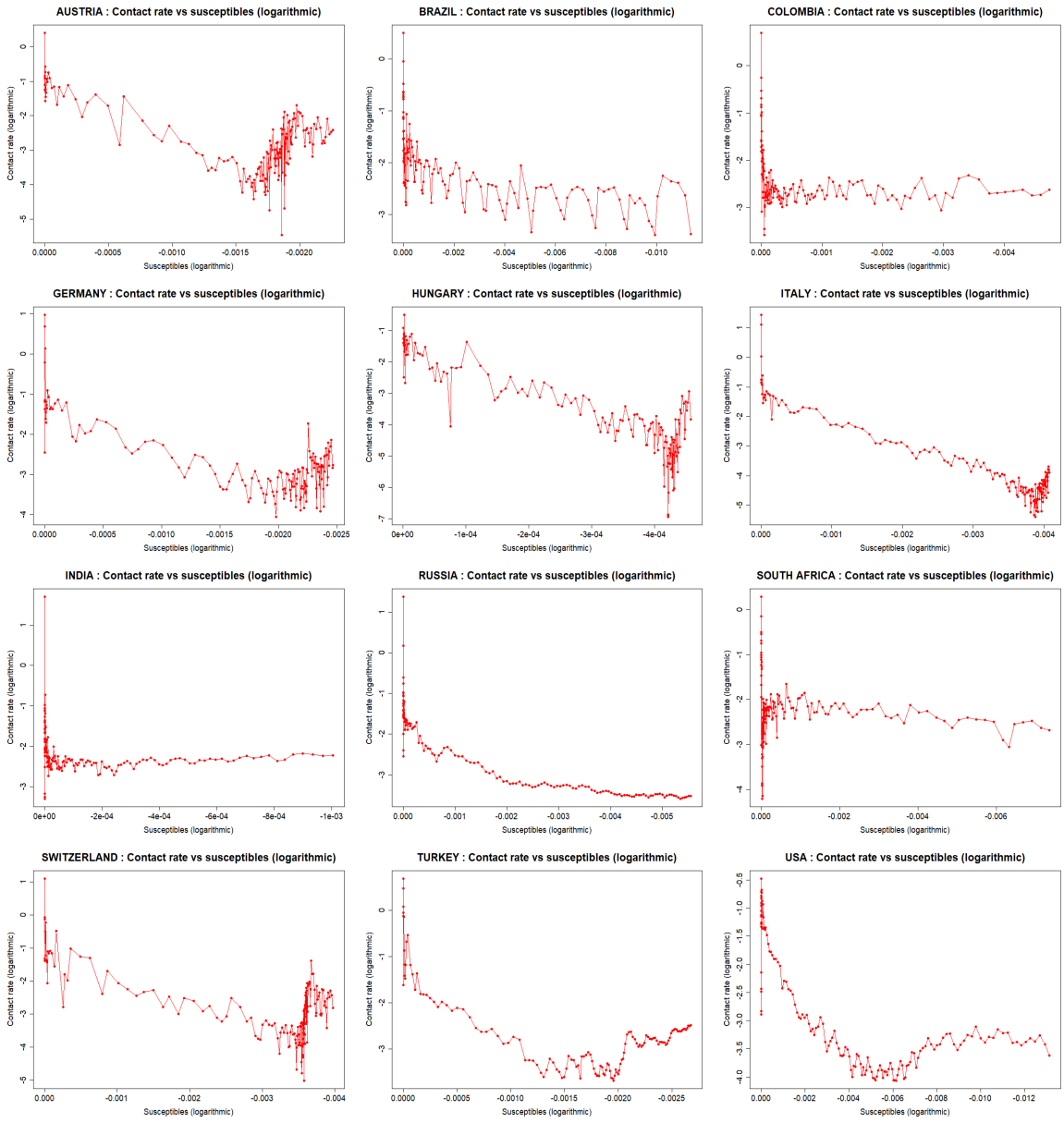

Figure 6: Contact rate versus susceptibles (logarithmic scale) in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

in Figure 5. This is because one obtains from (3.1) and (3.3) that

$$\begin{aligned}\log\beta(\Gamma_t) = \log\left(\frac{\Gamma_{t+1}-\Gamma_t}{I_t}\right) - \log(S_t) &= \log\beta_0 - \alpha\Gamma_t \\ &\approx \log\beta_0 + \alpha\log(1-\Gamma_t) \approx \log\beta_0 + \alpha\log(S_t)\end{aligned}$$

or

$$\log\left(\frac{\Gamma_{t+1}-\Gamma_t}{I_t}\right) \approx \log\beta_0 + \alpha\log(1-\Gamma_t) \approx \log\beta_0 + (1+\alpha)\log(S_t), \tag{3.6}$$

and hence we can approximate $f(S) = S^{1+\alpha}$, which matches with the control function in Example 2.4. Despite this similarity, the choice of model (3.1) addresses the negative dependence of the

contact rate on the increasing progression of the total infections Γ_t, which might be used to evaluate the containment efficiency of the control strategies.

3.1 Relation between total numbers and their log-growth rates

Another important issue is that, due to the lack of information (sometimes unreliable) in recovery and active case reports in several countries (e.g. Netherlands, Sweden, UK), it is impossible to estimate the contact rate using the previous analysis. Thus one needs to draw conclusions based on the growth rate of the total infections and the total deaths, on the possible final state of the total infections and total deaths. As shown below, our model implies a relation between $\log\left(\frac{d\log\Gamma_t}{dt}\right)$ versus Γ_t and also $\log\left(\frac{d\log D_t}{dt}\right)$ versus D_t.

First, observe from Theorem 2.5 that $\frac{I(\Gamma)}{\Gamma}$ is decreasing in Γ, hence it follows from (2.13) that $\frac{d\log\Gamma_t}{dt} = \frac{d\Gamma_t}{\Gamma_t dt}$ is a decreasing function of Γ_t, and so is $\log\left(\frac{d\Gamma_t}{\Gamma_t dt}\right)$. In addition, for $\Gamma_t \ll 1$ one can approximate

$$\begin{aligned}\log\left(\frac{d\Gamma_t}{\Gamma_t dt}\right) &= \log\beta_0 + \log(1-\Gamma_t) + \log G(\Gamma_t) + \log\frac{I(\Gamma_t)}{\Gamma_t} \\ &\approx \log\beta_0 - \Gamma_t + \log G(\Gamma_t) + \log\frac{I(\Gamma_t)}{\Gamma_t}. \end{aligned} \tag{3.7}$$

Using (2.10) and the fact that $I(0) = 0$, one can use Taylor's expansion to approximate, provided that $\Gamma_t \ll 1$,

$$\frac{I(\Gamma_t)}{\Gamma_t} \approx I'(0) + \frac{1}{2}I''(0)\Gamma_t = (1 - \frac{1}{\kappa_0}) + \frac{1}{2\kappa_0}(G'(0)-1)\Gamma_t.$$

As a result, it follows from (3.7) that for $\Gamma_t \ll 1$

$$\begin{aligned}\log\left(\frac{d\log\Gamma_t}{dt}\right) &\approx \log\Big[\beta_0(1-\frac{1}{\kappa_0})\Big] - \Gamma_t + \log G(\Gamma_t) + \log\Big[1 + \frac{G'(0)-1}{2(\kappa_0-1)}\Gamma_t\Big] \\ &\approx \log\Big[\beta_0(1-\frac{1}{\kappa_0})\Big] + \Big[\frac{G'(0)-1}{2(\kappa_0-1)} - 1\Big]\Gamma_t + \log G(\Gamma_t). \end{aligned} \tag{3.8}$$

Hence, under the assumption $G(\Gamma_t) = e^{-\alpha\Gamma_t}$, (3.7) yields

$$\log\left(\frac{d\log\Gamma_t}{dt}\right) \approx \log\Big[\beta_0(1-\frac{1}{\kappa_0})\Big] - (\alpha+1)\Big[\frac{1}{2(\kappa_0-1)} + 1\Big]\Gamma_t. \tag{3.9}$$

Figure 7 shows relation (3.9) in the selected countries, the steeper the curve, the faster it would reach the saturation phase. It is important to note that the linear regression is only broken at the saturation phase, when Γ_t approaches the theoretical limit Γ^∞ with $I(\Gamma^\infty) = 0$ thus making $\log\frac{I(\Gamma_t)}{\Gamma_t}$ in (3.7) to deteriorate very fast to negative values and the linear regression (3.9) is no longer true.

Figure 8 describes also a similar relation between the logarithm of the log-growth rate of the total deaths $\log\left(\frac{d\log D_t}{dt}\right)$ and the total deaths D_t, although it would be complicated to derive a quantitatve analysis for it as above using (2.20) and (2.21). The patterns in Figure 7 and 8 show that in general

$$\log\left(\frac{d\log\Gamma_t}{dt}\right) = H^{-1}(\Gamma_t), \qquad \log\left(\frac{d\log D_t}{dt}\right) = K^{-1}(D_t), \tag{3.10}$$

where H, K are decreasing sigmoid-type functions.

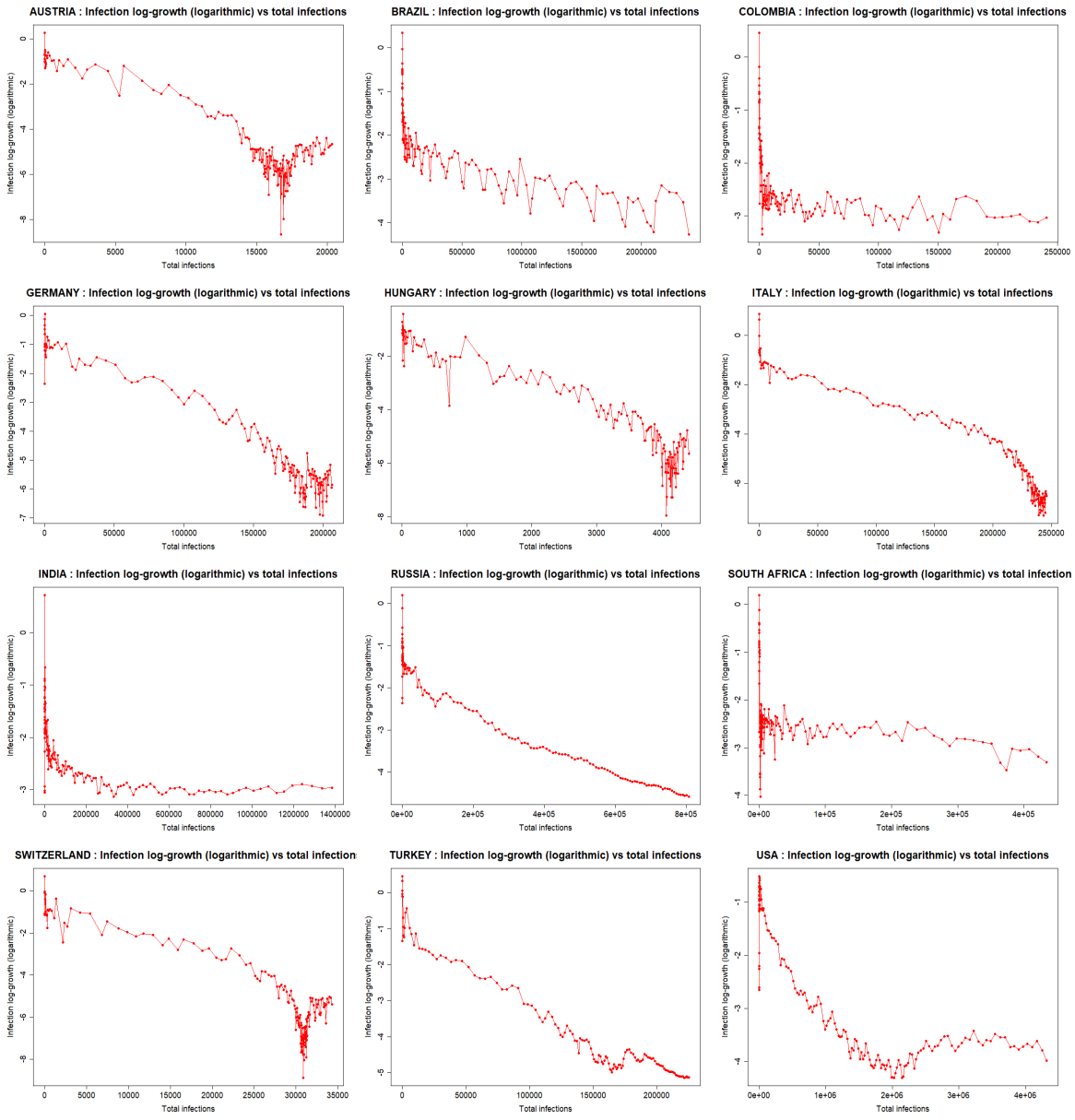

Figure 7: Total infection log-growth (logarithmic) versus total infections in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

3.2 An epidemic model with a scale-free contact number distribution

Since contacts take place locally, it might seem natural to include some spatial structure to obtain a more realistic model. This has been abundantly discussed in the literature, and in particular, there are many models for epidemic spread on networks. In this section, we provide a possible explanation of the observed behavior by considering a structured population. The population is structured by the contact numbers of individuals. Thus, the local structure of contacts is only implicitly contained in the model. This makes the model analytically tractable, while allowing for capturing some of the observed phenomena. – In a later section, we shall also consider an age-structured population.

A possible explanation of the decrease of the contact rate comes from an epidemic model with a

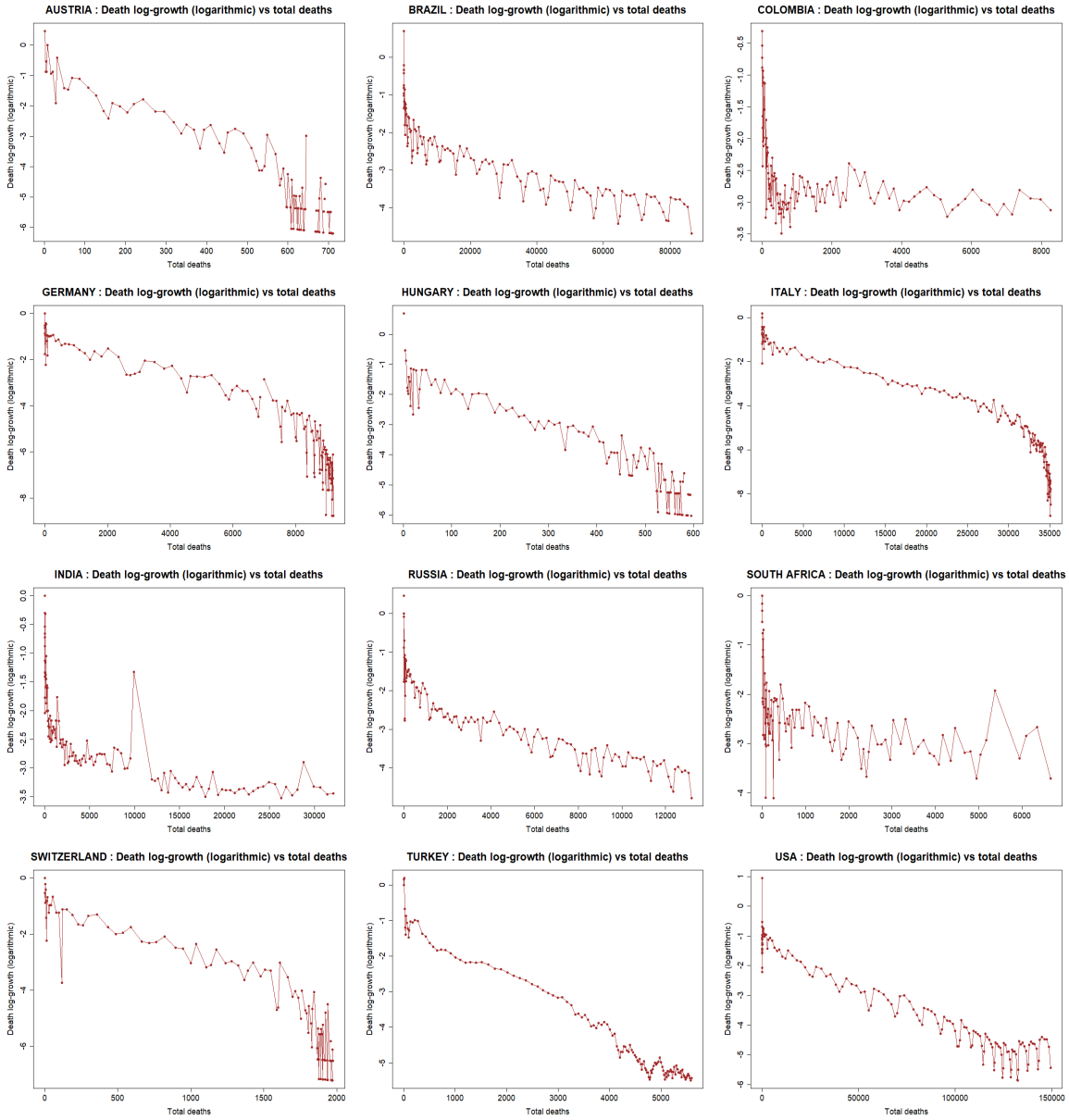

Figure 8: Death log-growth (logarithmic) versus total deaths in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

scale-free contact number distribution. In fact, Zipf's law, many social distributions are scale-free. Thus, we develop here a model where the number of people with k contacts follows such a scale-free distribution. This can be seen as a special case of a structured population model (see e.g. [14]). Also, there are network models available which attempt to explain the transmission mechanism [5], [18].

Thus, let $N_k = N_k(t)$ be the number of persons in the population with k contacts at time t. Then $N = \sum_{k=1}^{M} N_k$ is the whole population with $P(k) = \frac{N_k}{N}$ the probability that a person is connected to k other persons, where M is the maximal contact number in the population. The average contact is $\langle k \rangle := \sum_{i=1}^{M} iP(i)$, which is also called the average connectivity of the individuals in the contact network [23].

It is often assumed that this connectivity distribution follows a power law [25], [23], i.e. $P(k) \approx k^{-\lambda}$ with some exponent $\lambda \in (2,3]$. Denote by β_0, μ_0, γ_0 respectively the transmission, death and recovery coefficients, which characterize the disease and are usually assumed as independent of the contact networks in regions or countries. Following [12] and [23], we assume that the mixing pattern of the population is random, and the probability that a contact of a person is a person with i contacts depends only on the connectivity distribution of the population, i.e. this probability is given by $\frac{iP(i)}{\sum_{j=1}^{M} jP(j)}$. In particular, there are no assortativity or other effects that depend on a specific contact network topology. For simplicity of the model, assume that N_k is slowly changing over time (but might be abruptly set to zero if the network connectivity topology is changed due to the social distancing policies). An SIR model (which can be generalized to an SEIR model [28]) for such a population then has different equations for each k for the part of the population with k contacts, because persons with more contacts have a higher chance of getting infected and in turn can also spread the disease to more others. As we shall explore below, this suggests the strategy of reducing the average contact rate. These equations are of the form

$$\frac{dS_k}{dt} = -\beta_0 k S_k \sum_{j=1}^{M} \frac{jP(j)}{\sum_{j=1}^{M} jP(j)} \frac{I_j}{N_j} = -\frac{\beta_0}{N\langle k\rangle} k S_k \sum_{j=1}^{M} jI_j \tag{3.11}$$

$$\frac{d\Gamma_k}{dt} = \beta_0 k S_k \sum_{j=1}^{M} \frac{jP(j)}{\sum_{j=1}^{M} jP(j)} \frac{I_j}{N_j} = \frac{\beta_0}{N\langle k\rangle} k S_k \sum_{j=1}^{M} jI_j, \tag{3.12}$$

where S_k, I_k, Γ_k are the numbers of susceptible, infected, and total infections with degree k for $k = 1, \ldots, M$. The total numbers of susceptible, infected, recovered and death individuals are denoted by S, I, Γ. $\Gamma_k = I_k + D_k + R_k$ counts the total infections with k contacts, and $\Gamma = I + D + R$ the total number of infections, where $D_k(D), R_k(R)$ are the number of deaths (respectively recovered) with k contacts. Then $N_k = S_k + \Gamma_k$ and $N = S + \Gamma$.

By summing all equations (3.12) over k, one obtains

$$\frac{dS}{dt} = -\frac{\beta_0}{N\langle k\rangle} \sum_{k=1}^{M} kS_k \sum_{k=1}^{M} kI_k \tag{3.13}$$

$$\frac{d\Gamma}{dt} = \frac{\beta_0}{N\langle k\rangle} \sum_{k=1}^{M} kS_k \sum_{j=1}^{M} jI_j. \tag{3.14}$$

According to the data, $\Gamma \le \rho N$ where $\rho < 0.03$, hence it makes sense to assume that Γ_k almost contributes only a small part in N_k, so that in distribution $\frac{S_k}{S} \approx \frac{N_k}{N}$. As a result

$$\frac{1}{N\langle k\rangle} \sum_{k=1}^{M} kS_k = \frac{S}{N} \frac{\sum_{k=1}^{M} k\frac{S_k}{S}}{\sum_{k=1}^{M} k\frac{N_k}{N}} \approx \frac{S}{N}.$$

Therefore, equation (3.14) yields an approximation

$$\frac{d\Gamma}{dt} = \beta_0 \frac{S}{N} \sum_{i=1}^{M} kI_k = \beta_0 \frac{S}{N} I\Big(\sum_{i=1}^{M} k\frac{I_k}{I}\Big) = \beta_0 \frac{S}{N} I\langle k, I\rangle, \tag{3.15}$$

where $\langle k, I\rangle := \sum_{i=1}^{M} k\frac{I_k}{I}$ is the average contact of the infected group. The contact rate $\beta(\Gamma)$ in (3.1) can then be estimated as

$$\beta(\Gamma) = \frac{d\Gamma}{\frac{S}{N} I dt} \approx \beta_0 \langle k, I\rangle. \tag{3.16}$$

Since β_0, μ_0, γ_0 are epidemic coefficients independent of the network, the aim is then to minimize $\langle k, I\rangle$. Observing that

$$\langle k, I\rangle \leq M \sum_{k=1}^{M} \frac{I_k}{I} = M,$$

one way to do that is by social distancing policies, which aim to disable any type of meeting with more than m individuals, so that the maximum contact would be a time dependent number M_t which gradually decreases from M_0 to only $m \ll M_0$, hence finally $\langle k, I\rangle \leq m$. Therefore, using (3.16) one obtains

$$\beta(\Gamma) \leq \beta_0 m. \tag{3.17}$$

In particular, if the distribution of the currently infected group follows a power law with the time varying maximum contact M_t and the exponent $\lambda \in (2, 3)$, it follows from (3.16) that

$$\log \beta(\Gamma) \approx \log\Big(\beta_0 \sum_{k=1}^{M_t} k k^{-\lambda}\Big) \approx \log\Big(\beta_0 C_\lambda M_t^{2-\lambda}\Big) = \log(\beta_0 C_\lambda) - (\lambda - 2)\log M_t. \tag{3.18}$$

It is not clear how to relate $\log M_t$ with Γ_t, hence it remains to show quantitatively how the mechanism that works on the contact network affects the contact rate as in the form of (3.1)-(3.5), which is confirmed by empirical data. The intensity of the social distancing policies varies from country to country, and so we witness a sharp decrease of $\langle k, I\rangle = \langle k, I\rangle_t$ from close to M_0 to a much smaller m over time in Asian countries like China, Taiwan, South Korea, Thailand, while there is a more gradual decrease of $\langle k, I\rangle_t$ in European countries. In countries like Brazil, Colombia, India or South Africa, with either no or very little effective social distancing policies, M_t can remain little changed over time, resulting in a constant contact rate.

4 Estimating the death rate and the recovery rate

We test the relation (2.6) in the discrete form

$$\mu(\Gamma_t) := \frac{D_{t+1} - D_t}{I_t} = \mu_0 P\Big(\frac{\Gamma_t}{N_t}\Big) \Leftrightarrow \log \mu(\Gamma_t) = \log \mu_0 + \log P\Big(\frac{\Gamma_t}{N_t}\Big). \tag{4.1}$$

Figure 9 shows the relation of the logarithmic death rate $\log \mu(\Gamma_t)$ and the total infection Γ_t. While in some countries like Austria, Germany and Switzerland the death rate has some time lag before fluctuating around a constant rate (thus δ can be switched off to zero), other countries in Figure 9 show a remarkable downtrend regression (4.1), which suggests a model for the death rate of the form $\mu(\Gamma_t) = \mu_0 e^{-\delta\Gamma_t}$ as in Example 2.3. A possible explanation might come from differences in the age structure between the entire population and the infected group. In principle, old people are more vulnerable and would take social distancing measures more seriously than the young group. As a result, while the number of novel deaths is proportional to the currently infected old aged group, it is possible that the currently infected old aged group is decreasing exponentially relatively to the currently infected group. In other words

$$\mu(\Gamma_t) := \frac{D_{t+1} - D_t}{I_t} = \frac{D_{t+1} - D_t}{I_t^{\text{old}}} \times \frac{I_t^{\text{old}}}{I_t} = \mu_0 \times e^{-\delta\Gamma_t}. \tag{4.2}$$

On the other hand, figure 10 shows the relation of the logarithmic recovery rate in the discrete form

$$\gamma_0(\Gamma_t) := \frac{R_{t+1} - R_t}{I_t} = \gamma_0 \Leftrightarrow \log \gamma_0(\Gamma_t) = \log \gamma_0, \tag{4.3}$$

which implies that the recovery rate only fluctuates around a constant value and thus can be considered as constant. This is not surprising given the fact that almost 80% of the infected group of Covid1-19 will recover after a time lag of 2 to 3 weeks regardless of the social distancing policies or the public health care system.

In countries like Brazil, Colombia, India and South Africa, the recent situation is that the contact rate stays little changed (so that α is set to zero) while the death rate decreases exponentially ($\delta > 0$). As a result, $\frac{P(\Gamma_t)}{G(\Gamma_t)} = e^{-\delta\Gamma_t}$ is no longer an increasing function of Γ_t and we could observe the death ratio $\frac{D_t}{\Gamma_t}$ to top at some point and to decline slowly after that.

Both figure 9 and figure 10 show some noisy fluctuations in the first and in the saturation phase, which comes from the technical issue of taking the logarithm of the small number of novel deaths and novel recoveries.

4.1 A model with an age-structured population

Above, we have structured the population by contact numbers. Here, instead, we structure it according to age cohorts. Empirical data show that the contact, mortality and recovery rates vary considerably between age groups. Younger people may have more contacts, but older ones tend to be more severely affected by the disease. Therefore, it makes a difference whether contacts preferably take place between people in the same age group or between different age groups. Thus, we need to develop a model with an age-structured population. This may then suggest more targeted containment measures that vary between the different age groups.

An age-structured population can be modelled as follows (see [14, p. 288]). For simplicity, we only consider two groups, young people (below 50) and old people (over 50), but it is not difficult to extend the model to finer age divisions. Denote by $(S^o, \Gamma^o, I^o, D^o, R^o)$ and $(S^y, \Gamma^y, I^y, D^y, R^y)$ the corresponding variables in the two groups, where we assume for simplicity that the distribution of the two group stays static so that $S^o + \Gamma^o = N^o = \lambda \leq \frac{1}{2}, S^y + \Gamma^y = N^y = 1 - \lambda \geq \frac{1}{2}$. The mechanism of contacts between the susceptible $S^{y/o}$ and the infected $I^{y/o}$ in the two groups is then characterized by parameters $(\beta^{y,y}, \beta^{y,o}, \beta^{o,y}, \beta^{o,o})$. Likewise, we have parameters (μ^o, γ^o) and (μ^y, γ^y) for the death and recovery rates of the two groups. System (2.1)-(2.3) can be rewritten as

$$\begin{aligned}
\frac{dS_t^y}{dt} &= -S_t^y(\beta^{y,y}I_t^y + \beta^{y,o}I_t^o)e^{-\alpha\Gamma_t} && (4.4)\\
\frac{dI_t^y}{dt} &= S_t^y(\beta^{y,y}I_t^y + \beta^{y,o}I_t^o)e^{-\alpha\Gamma_t} - \mu^y I_t^y - \gamma^y I_t^y && (4.5)\\
\frac{dR_t^y}{dt} &= \gamma^y I_t^y, \quad \frac{dD_t^y}{dt} = \mu^y I_t^y; && (4.6)\\
\frac{dS_t^o}{dt} &= -S_t^o(\beta^{o,y}I_t^y + \beta^{o,o}I_t^o)e^{-\alpha\Gamma_t} && (4.7)\\
\frac{dI_t^o}{dt} &= S_t^o(\beta^{o,y}I_t^y + \beta^{o,o}I_t^o)e^{-\alpha\Gamma_t} - \mu^o I_t^o - \gamma^o I_t^o && (4.8)\\
\frac{dR_t^o}{dt} &= \gamma^o I_t^o, \quad \frac{dD_t^o}{dt} = \mu^o I_t^o; && (4.9)
\end{aligned}$$

where $\Gamma_t = \Gamma_t^y + \Gamma_t^o, S_t = S_t^y + S_t^o, I_t = I_t^y + I_t^o, D_t = D_t^y + D_t^o, R_t = R_t^y + R_t^o$ are respectively the total infections, the susceptible, the currently infected, the total deaths and total recoveries. α is a control varible which can be switched off to zero. Moreover, it is reasonable to assume

$$\beta^{y,y} \geq \beta^{y,o}, \beta^{o,y}, \beta^{o,o} \geq 0; \quad \mu^o \gg \mu^y; \quad \gamma^y \simeq \gamma^o. \tag{4.10}$$

When social distancing control policies are in force, the total infections $\Gamma^o, \Gamma^y, \Gamma$ take a negligible amount of the population and we may approximate $S_t^y \approx 1 - \lambda, S_t^o \approx \lambda$. Hence the equations (4.5),

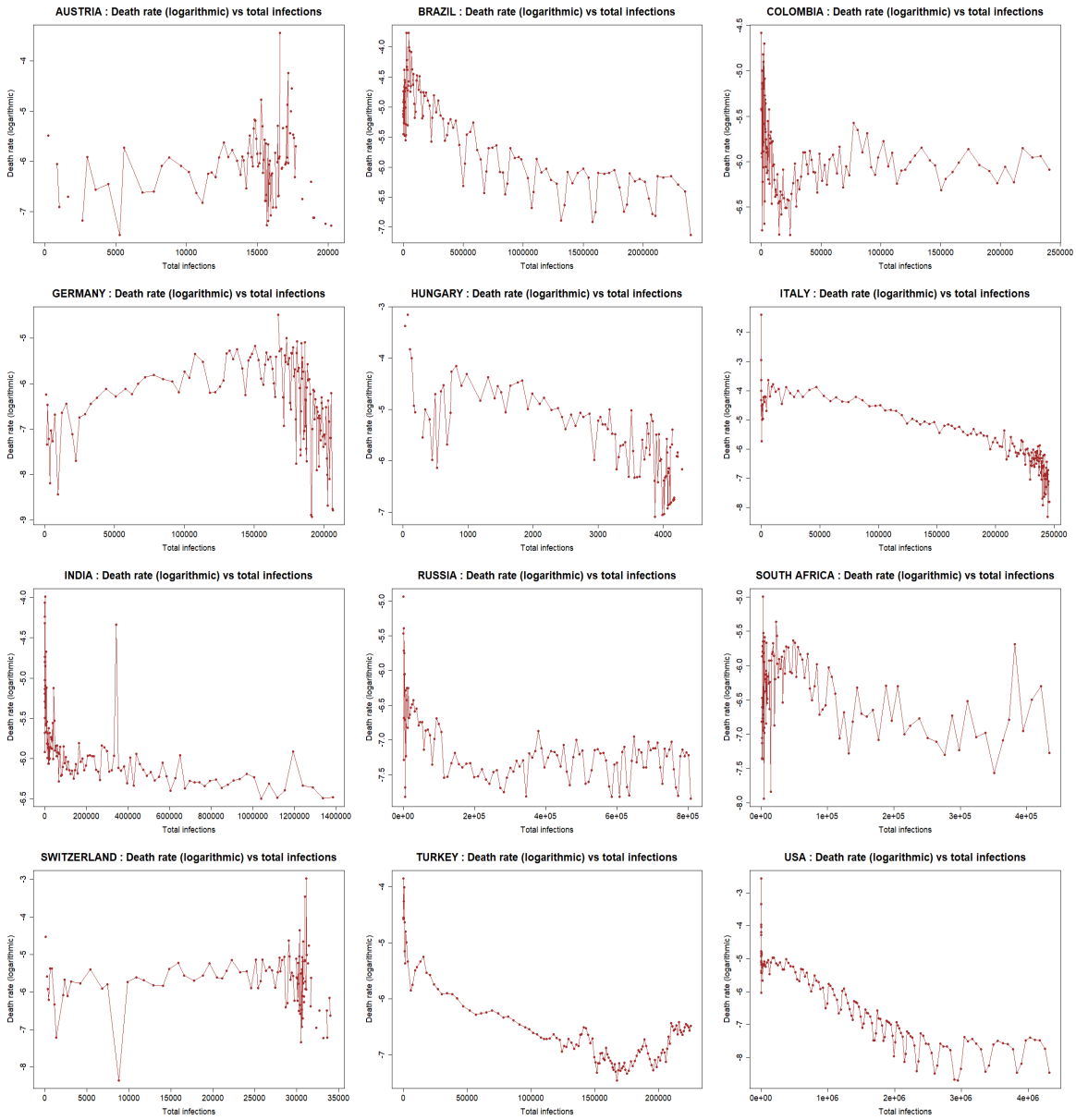

Figure 9: Death rate (logarithmic) versus total infections in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

(4.8) for currently infected groups can be approximated by

$$\frac{dI_t^y}{dt} = \Big((1-\lambda)\beta^{y,y}e^{-\alpha\Gamma_t} - \mu^y - \gamma^y\Big)I_t^y + (1-\lambda)\beta^{y,o}e^{-\alpha\Gamma_t}I_t^o \tag{4.11}$$

$$\frac{dI_t^o}{dt} = \lambda\beta^{o,y}e^{-\alpha\Gamma_t}I_t^y + \Big(\lambda\beta^{o,o}e^{-\alpha\Gamma_t} - \mu^o - \gamma^o\Big)I_t^o. \tag{4.12}$$

Wehn in addition to the social distancing measures, contact between the two age groups are strictly prevented, we can assume that

$$\beta^{0,y} \ll 1, \tag{4.13}$$

so that the term $\lambda\beta^{o,y}e^{-\alpha\Gamma_t}I_t^y$ in (4.12) can be neglected. In that case, system (4.11)-(4.12) can be

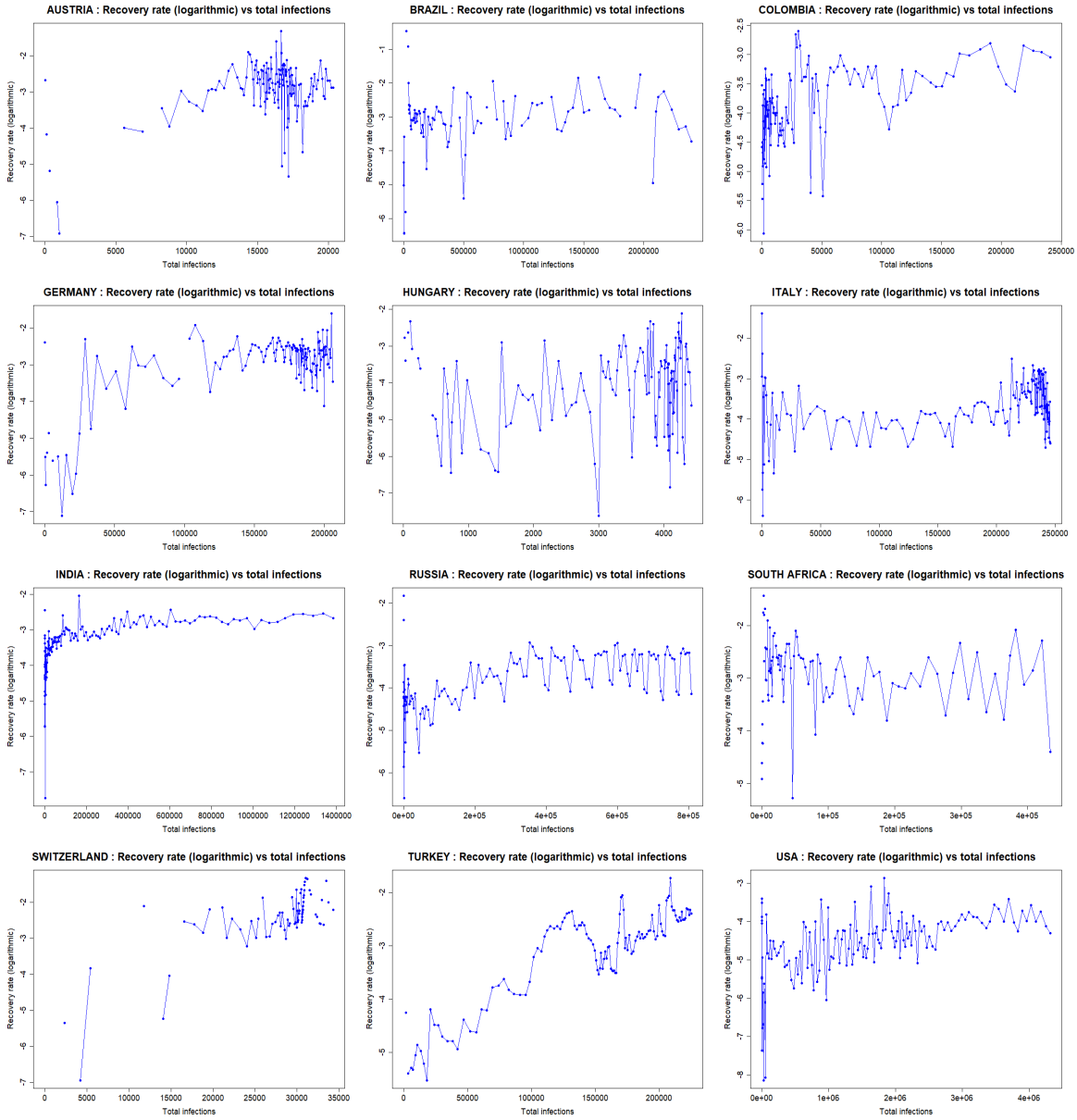

Figure 10: Recovery rate (logarithmic) versus total infections in countries: Austria, Brazil, Colombia, Germany, Hungary, Italy, India, Russia, South Africa, Switzerland, Turkey, USA. Data source: Worldometer.

approximated by a triangular time varying system

$$\frac{dI_t^y}{dt} = \Big((1-\lambda)\beta^{y,y}e^{-\alpha\Gamma_t} - \mu^y - \gamma^y\Big)I_t^y + (1-\lambda)\beta^{y,o}e^{-\alpha\Gamma_t}I_t^o \tag{4.14}$$

$$\frac{dI_t^o}{dt} = \Big(\lambda\beta^{o,o}e^{-\alpha\Gamma_t} - \mu^o - \gamma^o\Big)I_t^o. \tag{4.15}$$

As a result, if we assume further that

$$(1-\lambda)\beta^{y,y}e^{-\alpha\Gamma_t} - \mu^y - \gamma^y \geq \lambda\beta^{o,o}e^{-\alpha\Gamma_t} - \mu^o - \gamma^o + \delta, \tag{4.16}$$

then it is easy to solve the triangular system (4.14)-(4.15) explicitly and prove that there exists a

generic constant C such that

$$\frac{I_t^o}{I_t^y} \leq Ce^{-\delta t}. \tag{4.17}$$

Hence (4.17) and the assumption $\mu^o \gg \mu^y$ in (4.10) yield

$$\frac{dD_t}{I_t} = \frac{dD_t^y + dD_t^o}{I_t^y + I_t^o} = \frac{\mu^y I_t^y + \mu^o I_t^o}{I_t^y + I_t^o} \approx \frac{\mu^y + \mu^o \frac{I_t^o}{I_t^y}}{1 + \frac{I_t^o}{I_t^y}} \approx \frac{\mu^y + C\mu^o e^{-\delta t}}{1 + Ce^{-\delta t}} \approx \mu^y + C\mu^o e^{-\delta t}, \tag{4.18}$$

which explains the empirical evidence that the death rate decreases exponentially for some time and then stays constant after that. Similarly

$$\frac{dR_t}{I_t} = \frac{dR_t^y + dR_t^o}{I_t^y + I_t^o} = \frac{\gamma^y I_t^y + \gamma^o I_t^o}{I_t^y + I_t^o} \approx \gamma^y, \tag{4.19}$$

which explains the empirical evidence that the recovery rate only fluctuates around a constant value.

Finally, since (4.17) implies that $\frac{I_t^o}{I_t} \approx Ce^{-\delta t}$ and $\frac{I_t^y}{I_t} \approx 1$, it follows from (4.5) and (4.8) that

$$\frac{d\Gamma_t}{S_t I_t dt} \approx \Big((1-\lambda)\beta^{y,y} + \lambda\beta^{o,y}\Big)e^{-\alpha\Gamma_t},$$

which matches (3.3).

One could also combine an age-structured model with another one that is structured by some other criterion, like the contact number above, but we do not pursue that here.

5 Discussion: the effect of the time lag

To make the model more realistic, we also need to include factors like the incubation time that induce delays into the epidemic. In particular, because of the varying incubation time, the number $I_{t+1} - I_t$ of new infected cases depends on the frequency of social contacts between the susceptible and infected groups at a time $t - \tau$ in the past. Likewise, the death rate and the recovery rate should be considered as a function of the infections at a time in the past. We therefore propose a variational SIR discrete model which takes into account three time delay factors. In empirical models, the distribution of the random delay time of course should represent the available data as closely as possible, in order to make proper quantitative predictions.

Infection mechanism with time delay

We now develop some more details. Denoting by τ the incubation time, then τ is a random variable with a distribution $\mathbb{P}_\tau$ and is bounded from above by a non-random constant r (for other studies, τ is often assume to be non-random for simplicity, see e.g. [1], [8], [3], [9], [21],[32], [34]). It is clear that all the social contacts in the past between the susceptible group and the infected group need some incubation time for the susceptible group to show symptoms of infection. Therefore, from time t to time $t+1$, the new infected cases $\Gamma_{t+1} - \Gamma_t$ depend on the expected number of all social contacts between the susceptible group and the infected group at a random time in the past $t - \tau$. The explicit formula is given by

$$\Gamma_{t+1} - \Gamma_t \;=\; \beta_0 \sum_{i=1}^{r} G(\Gamma_{t-i}) S_{t-i} I_{t-i} \mathbb{P}_\tau(\tau = i) = \beta_0 \mathbb{E}_{\mathbb{P}_\tau} G(\Gamma_{t-\tau}) S_{t-\tau} I_{t-\tau}. \tag{5.1}$$

Death and recovery mechanism with time delay

Also, assume that the change of deaths (recoveries) over time also depends on some death random interval (recovery random interval) τ_D with distribution $\mathbb{P}_{\tau_D}$ and bounded by r (τ_R, with distribution $\mathbb{P}_{\tau_R}$ and bounded by r). In addition, assume that the number of deaths (recoveries) with death interval τ_D is also proportional to the infected number $I_{t-\tau_D}$ (respectively $I_{t-\tau_R}$). Then we derive the following equations which help to determine the death and recovery rates

$$D_{t+1} - D_t = \mu_0 \mathbb{E}_{\mathbb{P}_{\tau_D}} P(\Gamma_{t-\tau_D}) I_{t-\tau_D} = \mu \sum_{i=1}^{r} P(\Gamma_{t-i}) I_{t-i} \mathbb{P}_{\tau_D}(\tau_D = i) \tag{5.2}$$

and

$$R_{t+1} - R_t = \gamma_0 \mathbb{E}_{\mathbb{P}_{\tau_R}} I_{t-\tau_R} = \gamma_0 \sum_{i=1}^{r} I_{t-i} \mathbb{P}_{\tau_R}(\tau_R = i). \tag{5.3}$$

While the delay effect would result in a more difficult problem on the asymptotic stability of the epidemic model, determining the time lags quantitatively is also another challenge. Practically, τ, τ_D and τ_R are often assumed to be lognormal, Weibull or Gamma distributed [19], [20], [29]. However, their mean and standard deviation values are either unknown or very uncertain. Recent studies [4], [15], [17], [19], [20], [29], [31] with different data sets show that the incubation time τ ranges from 2 to 14 days, with mean value varying from 4.6 to 6.5 days. The mean time from the onset of the symptom to death also fluctuates from 7 to 21 days [15], [20], [29], [30], while the mean time from being infected to recovered ranges from 2 to 3 weeks [4], [20], [29]. This lack of detailed and accurate information really poses a data issue before reaching any reliable model or prediction, as discussed in [6].

Disclaimer

We evaluate the data provided by WORLDOMETER (https://www.worldometers.info/coronavirus/) on a daily basis, concerning the numbers of persons infected, deceased or recovered from the COVID-19 virus in different countries. We exclude any liability with regard to the quality and accuracy of the data in use. The data are usually based on information provided by national governments or authorities. The number of cases reported may be significantly lower than the number of people actually infected, and these discrepancies may be different in different countries.

References

[1] Akimenko, V. (2017). An age-structured SIR epidemic model with fixed incubation period of infection. Computers and Mathematics with Applications, 73: 1485–1504.

[2] Beddington, J. Mutual interference between parasites or predators and its effect on searching efficiency. J. Animal Ecology, Vol. 44, No. 1 (1975), 331–340.

[3] Beretta, E.; Hara, T.; Ma, W.; Takeuchi, Y. (2001). Global asymptotic stability of an SIR epidemic model with distributed time delay. Nonlinear Analysis, 47: 4107–4115.

[4] Bi, Q. et al. (2020). Epidemiology and transmission of Covid-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. The Lancet. DOI: 10.1016/S1473-3099(20)30287-5.

[5] Brockmann, D.; Helbing D. (2013). The hidden geometry of complex network-driven contagion phenomena. Science, 342: 1337–1342.

[6] Callaghan, S. (2020). COVID-19 is a data science issue. (Editorial) Patterns, 1: 1–3.

[7] Capasso, V.; Serio, G. (1978). A generalization of the Kermack - Mckendrick deterministic epidemic model. Math. Biosci., 42 (1978), 43–61.

[8] Cook, K.; van den Driessche, P. (1996). Analysis of an SEIRS epidemic model with two delays. J. Math. Biol., 35: 240–260.

[9] Enatsu, Y.; Nakata, Y.; Muroya, Y. (2011). Global stability of SIR epidemic models with a wide class of nonlinear incidence rates and distributed delays. Discrete and Continuous Dynamical Systems B, 15: 61–74.

[10] Gray, A.; Greenhalgh, D.; Hu, L.; Mao, X.; Pan, J. (2011). A stochastic differential equation SIS epidemic model. SIAM J. Appl. Math. 71(3): 876–902.

[11] Gros, C.; Valenti, R.; Schneider, L.; Valenti, K.; Gros, D. (2020). Containment efficiency and control strategies for the Corona pandemic costs. Preprint: arXiv: 2004.00493v2.

[12] Huang, C.; Cao J.; Wen. F.; Yang. X. (2016). Stability analysis of SIR model with distributed delay on complex networks. Plos One, 11(8): e0158813.

[13] Huo, L.; Jiang, J.; Gong S.; He, B. Dynamical behavior of a rumor transmission model with Holling-type II functional response in emergency event. Phys. A, 450 (2016), 228–240.

[14] H.Inaba, Age-structured population dynamics in demography and epidemiology, Springer, 2017

[15] Jung, S. et al. (2020). Real-time estimation of the risk of death from novel coronavirus (Covid-19) infection: inference using exported cases. Journal of Medicine, 9(253). DOI: 10.3390/jcm9020523.

[16] Kermack, W. O.; McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society A. 115 (772): 700–721. Bibcode:1927RSPSA.115..700K. doi:10.1098/rspa.1927.0118.

[17] Kuchrski, A. J. et al. (2020). Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The Lancet, 20: 553–558.

[18] Lang, J.; Miller, J. (2018). Analytic models for SIR disease spread on random spatial networks. Journal of Complex Networks, 6: 948–970.

[19] Lauer, S. et al. (2020). The incubation period of coronavirus disease 2019 (Covid-19) from publicly reported confirmed cases: Estimation and application. Annals of the Internal Medicine, 172(9), 577–583.

[20] Linton, N. et al. (2020). Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data. Journal of Clinical Medicine, 9(538). DOI: 10.3390/jcm9020538.

[21] Ma, W.; Takeuchi, Y.; Hara, T.; Beretta, E. (2002). Permanence of an SIR epidemic model with distributed time delays. Tohoku Math. J., 54: 581–591.

[22] Maier, B.; Brockmann, D. (2020). Effective containment explains subexponential growth in recent confirmed Covid-19 cases in China. Science, 10.1126/science.abb4557

[23] May, R. M.; Lloyd, A. L. (2001). Infection dynamics on scale-free networks. Physical Review E, 64: 066112.

[24] Miller, J.C. (2012). A note on the derivation of epidemic final sizes. Bulletin of Mathematical Biology. 74 (9). section 4.1. doi:10.1007/s11538-012-9749-6. PMC 3506030. PMID 22829179.

[25] Moreno, Y.; Pastor-Satorras, R.; Vespignani, A. (2002). Epidemic outbreaks in complex heterogenous networks. Eur. Phys. J. B., 26: 521–529.

[26] Prem, K. et al. (2020). The effect of control strategies to reduce social mixing on outcomes of the Covid-19 epidemic in Wuhan, China: a modelling study. The Lancet. DOI: 10.1016/S2468-2667(20)30073-6.

[27] Ruan, S. ; Wang, W. Dynamical behavior of an epidemic model with a nonlinear incidence rate. J. Differential Equations., 188 (2003), 135–163.

[28] Small, M.; Tse, Chi K. (2005). Small world and scale free model of tranmission of SARS. International Journal of Bifurcation and Chaos, 5(5): 1745–1755.

[29] Verity, R. et al. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet. DOI: 10.1016/S1473-3099(20)30243-7.

[30] Wilson, N. et al. Case-fatality risk estimates for Covid-19 calculated by using a lag time for fatality. Emerging Infectious Diseases., 26 (6), (2020), 1339–1341.

[31] Wu, J.; Leung K.; Leung, G. (2020). Nowcasting and forecasting the potential demestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. The Lancet. DOI: 10.1016/S0140-6736(20)30260-9

[32] Yan, P.; Liu, Sh. (2006). SEIR epidemic model with delay. ANZIAM J. 48: 119–134.

[33] Yang, Q.; Jiang, D.; Shi, N.; Ji, C. The ergodicity and extinction of stochastically perturbed SIR and SEIR epidemic models with saturated incidence. J. Math. Anal. Appl., 388 (2012), 248–271.

[34] Yin, Zh.; Yu, Y.; Lu, Zh. (2020). Stability analysis of an age-structured SEIRS model with time delay. Mathematics, 8(455). DOI: 10.3390/math8030455.