

# Fleeing lockdown and its impact on the size of epidemic outbreaks in the source and target regions - a COVID-19 lesson

Maria Vittoria Barbarossa (✉ [barbarossa@fias.uni-frankfurt.de](mailto:barbarossa@fias.uni-frankfurt.de))

Frankfurt Institute for Advanced Studies <https://orcid.org/0000-0001-8788-1709>

Norbert Bogya

Bolyai Institute, University of Szeged

Attila Dénes

Bolyai Institute, University of Szeged <https://orcid.org/0000-0003-1827-7932>

Gergely Röst

Bolyai Institute, University of Szeged <https://orcid.org/0000-0001-9476-3284>

Hridya Vinod Varma

Interdisciplinary Center for Scientific Computing, Heidelberg University

Zsolt Vizi

Bolyai Institute, University of Szeged

---

## Research Article

**Keywords:** disease dynamics, COVID-19, final size, lockdown, two patches, control measures

**Posted Date:** September 24th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-82993/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Fleeing lockdown and its impact on the size of epidemic outbreaks in the source and target regions – a COVID-19 lesson

Maria Vittoria Barbarossa<sup>1,3,\*</sup>, Norbert Bogya<sup>2</sup>, Attila Dénes<sup>2</sup>, Gergely Röst<sup>2</sup>, Hridya Vinod Varma<sup>3</sup>, and Zsolt Vizi<sup>2</sup>

<sup>1</sup>Frankfurt Institute for Advanced Studies, Frankfurt, D-60438, Germany

<sup>2</sup>Bolyai Institute, University of Szeged, Szeged, H-6720, Hungary

<sup>3</sup>Interdisciplinary Center for Scientific Computing, Heidelberg, D-69120, Germany

\*barbarossa@fias.uni-frankfurt.de

## ABSTRACT

The COVID-19 pandemic forced authorities worldwide to implement moderate to severe restrictions in order to slow down or suppress the spread of the disease. It has been observed in several countries that a significant number of people fled a city or a region just before strict lockdown measures were implemented. This behavior carries the risk of seeding a large number of infections all at once in regions with otherwise small number of cases. In this work, we investigate the effect of fleeing on the size of an epidemic outbreak in the region under lockdown, and also in the region of destination. We propose a mathematical model that is suitable to describe the spread of an infectious disease over multiple geographic regions. Our approach is flexible to characterize the transmission of different viruses. As an example, we consider the COVID-19 outbreak in Italy. Projection of different scenarios shows that (i) timely and stricter intervention could have significantly lowered the number of cumulative cases in Italy, and (ii) fleeing at the time of lockdown possibly played a minor role in the spread of the disease in the country.

## Introduction

A novel coronavirus (SARS-CoV-2) causing the severe acute respiratory illness COVID-19 appeared in China at the end of 2019. The first cases were identified in Wuhan, a city with over 11 million inhabitants, capital and largest city of Hubei Province. Wuhan being the political, economic, financial, commercial, cultural and educational centre of Central China and a major domestic and international transportation hub, it was expected that the disease could easily spread to other parts of China and to other countries. On January 23, 2020, in an attempt to control the spread of the disease, the Chinese government introduced a lockdown in Wuhan and other cities in Hubei. The residents of Wuhan were informed at 2 am that from 10 am of the same day, all public transport would be suspended. Wuhan residents would thereafter not be allowed to leave the city without permission. This notice was followed by an exodus from Wuhan. It was reported that around 300,000 people left Wuhan by train alone before the start of the lockdown<sup>1</sup>. It is important to note that prior to the lockdown of Hubei on January 23, almost all detected cases in other major cities were exported from Wuhan<sup>2</sup>.

Similar events happened in Italy, where the first lockdown measures were introduced on February 21, with ten municipalities of the province of Lodi (Lombardy) and one in the province of Padua (Veneto) being quarantined<sup>3</sup>. This measure affected around 50,000 people. Extension of the quarantine zone to a large area of Northern Italy, including the whole region of Lombardy and 14 major provinces in Emilia-Romagna, Veneto, Piedmont and Marche, was announced on March 8. Travel from and to the affected areas was restricted, though not completely banned. A leakage of a draft of the decree on the night of March 7, published by *Corriere della Sera*, resulted in a panic in the affected areas. Major newspapers reported of hundreds of people crowding onto trains and buses to leave the extended quarantine area in Milan<sup>4</sup>. Overall, thousands of people reportedly left the north for the southern regions, including Sicily<sup>5</sup> and Sardinia<sup>6</sup>.

A few mathematical studies on the Chinese data have investigated the effects of mobility from quarantined regions from different points of view. Chen et al.<sup>7</sup> analysed the correlation of case numbers and population migration data of Wuhan and Hubei to find that the number of infected cases was highly correlated with the emigrated populations from Wuhan. Li et al.<sup>8</sup> applied a nonautonomous mechanistic model to study the effects of the lockdown and predict the virus transmission in Wuhan and Beijing. Most studies (see, e.g. Zhao et al.<sup>2</sup>, Lau et al.<sup>9</sup>, Lai et

al.<sup>10</sup>, Boldog et al.<sup>11</sup> ) investigate the association between the domestic/international travel load and the number of cases exported from Wuhan to other regions in China. These studies, however, do not seem to consider the large exodus induced by the lockdown measures.

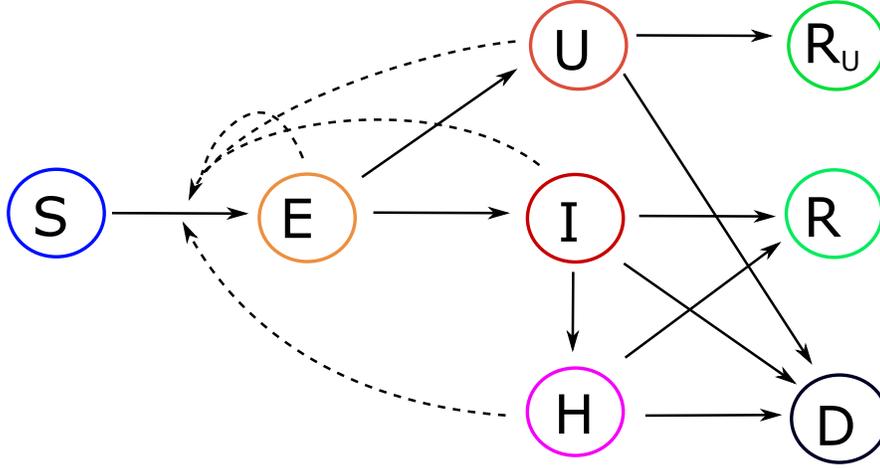
We present here a mathematical framework for studying the effects of a potential lockdown-induced exodus, together with the time and strength of intervention measures, on the outbreak size in different geographic regions. We assume that the outbreak starts in one specific area, simply denoted as *region A*, where lockdown is implemented at time  $T$ . This induces people from region A to flee the lockdown and move to another geographic *region B*. If region B was disease-free up to time  $T$ , the lockdown-induced migration from region A can lead to importation of the infection in a previously uninfected area. We formulate and study a compartmental epidemic model for the spread of an infectious disease in the two different geographic regions, establishing a new type of intervention-dependent final size relation. The latter can be used to estimate the final epidemic size even in the case of a change in the population size due to the migration just before the lockdown. This model is calibrated on the example of the Italian outbreak. By means of numerical simulations we project different scenarios showing that (i) timely and stricter intervention could have significantly lowered the number of cumulative cases in Italy, and (ii) migration of at the time of lockdown possibly played a minor role in the spread of the disease in the country. The approach presented in this study can be adapted to describe the spread of an infectious disease over multiple geographic regions or to describe the transmission of viruses others than SARS-CoV-2.

## Methods

### Modeling transmission dynamics

The mathematical model proposed in this study is based on a system of ordinary differential equations (ODEs) that describes interactions between different groups of individuals in the population. The proposed approach extends the known *S-E-I-R* (susceptibles–exposed–infected–recovered) model for disease dynamics<sup>12</sup>, and is suitable to describe the spread of several infectious diseases, though we shall specifically talk about COVID-19 later. Before describing the lockdown scenario and other intervention measures aimed at mitigating the outbreak, we introduce the core model for the (uncontrolled) spread of the virus in a population. The ODE approach that we use assumes that the population is homogeneous and well-mixed within a region. Individuals are classified according to their status with respect to the virus spread in the community.

At the beginning of the epidemic, that is, when first cases are reported, the majority of the population is assumed to be susceptible ( $S$ ), and hence can be infected. The time between exposure to the virus (becoming infected) and symptom onset, on average  $1/\alpha$  days, is known as "exposed" or "presymptomatic" period. To better understand the role of transmission from infected people with mild or no symptoms, we distinguish between transmission from people who are infected, might have very mild symptoms, but remain undetected ( $U$ ) and transmission from people who are infected but still in the presymptomatic period ( $E$ ). Detected infectives might require hospitalization ( $H$ ) or not ( $I$ ). Detection of cases is assumed to occur either after the latent phase (with probability  $\rho$ ) or post mortem (with probability  $\sigma$ ). Duration of infection can be different for the three different groups. Undetected infections lead to undetected recoveries, which cannot be reported unless testing for ongoing (virus detection) or previous (antibody detection) infections is performed. Individuals who recovered from a detected ( $R$ ) or an undetected ( $R_U$ ) infection, as well as patients who died from the infection ( $D$ ), are removed from the chain of transmission. Susceptible individuals can be infected via contacts with presymptomatic (transmission rate  $\beta_E$ ), undetected cases (transmission rate  $\beta_U$ ), detected but not hospitalized ( $\beta_I$ ) and hospitalized cases ( $\beta_H$ ). We assume that undetected infectives, due to possibly absent or unspecific symptoms, do not restrict their contacts to others, and therefore have higher transmission rates than detected infected individuals. Hospitalized cases are properly isolated, and hence their transmission rate is assumed to be the lowest. The interpretation of model parameters, all assumed to be nonnegative, is summarized in Table 1. The dynamics of the core model described above and shown in Figure 1 is



**Figure 1. Model structure for the transmission dynamics of an infectious disease, on the example of COVID-19.** Solid arrows indicate transition from one compartment to another, dashed arrows indicate virus transmission due to contact with infectives. Upon infection, susceptible ( $S$ ) individuals enter a latent/presymptomatic phase ( $E$ ). After symptoms onset, infections may be detected ( $I$ ) or remain undetected ( $U$ ). Detected cases might become severe and require hospitalization ( $H$ ). Infected individuals who recovered from a detected ( $R$ ) or an undetected ( $R_U$ ) infection, as well as patients who died ( $D$ ) upon infections, are removed from the chain of transmission.

given by the following system of differential equations:

$$\begin{aligned}
 \dot{S}(t) &= -\lambda(t)S(t) && \text{susceptibles} \\
 \dot{E}(t) &= \lambda(t)S(t) - \alpha E(t) && \text{exposed/presymptomatic} \\
 \dot{U}(t) &= (1 - \rho)\alpha E(t) - \gamma_U U(t) && \text{undetected infectives} \\
 \dot{I}(t) &= \rho\alpha E(t) - (\gamma_I + \delta_I + \eta)I(t) && \text{detected, non-hospitalized infectives} \\
 \dot{H}(t) &= \eta I(t) - (\gamma_H + \delta_H)H(t) && \text{hospitalized infectives} \\
 \dot{R}(t) &= \gamma_I I(t) + \gamma_H H(t) && \text{recovered from detected infection} \\
 \dot{R}_U(t) &= (1 - \sigma)\gamma_U U(t) && \text{recovered from undetected infection} \\
 \dot{D}(t) &= \delta_I I(t) + \delta_H H(t) + \sigma\gamma_U U(t) && \text{deceased,}
 \end{aligned} \tag{1}$$

where

$$\lambda(t) = \frac{\beta_E E(t) + \beta_I I(t) + \beta_U U(t) + \beta_H H(t)}{N - D(t)}$$

and  $N$  is the total population and is assumed to be constant for the duration of the observations, that is demographic variations other than disease-induced deaths are not considered.

### Numerical integration, model calibration and sensitivity analysis

The functionalities for solving the epidemic model (1) and fitting the parameters to the data were implemented in Python language: we use `odeint` function from the `scipy.integrate` package<sup>13</sup> for integrating the model equations and the non-linear least-squares minimization from the `lmfit` module<sup>14</sup> for estimating the parameters. We selected the L-BFGS-B method<sup>15</sup> for the optimization procedure, which is a very efficient algorithm for solving large scale problems. In case the time series were incomplete, the days corresponding to the missing values were omitted in the minimization. For all methods, we used the default values for options such as tolerances and step sizes. The confidence intervals of the estimated model parameters were computed based on the standard error of the estimated covariance matrix. For the plots in Figures 2–5, integration in MATLAB<sup>®</sup> by means of available ODE routines (`ode45`, `ode23s`) was used. Numerical approximation of the final size formula was verified by comparing the numerical solution of the implicit equation (4) (computed using `fsolve` from the `scipy.optimize` package)

**Table 1.** Interpretation of parameters in model (1).

Parameters	Description (unit)
$\beta_E$	Transmission rate from presymptomatics to susceptibles (1/(days $\times$ contact))
$\beta_U$	Transmission rate from undetected infectives to susceptibles (1/(days $\times$ contact))
$\beta_I$	Transmission rate from non-hospitalized cases to susceptibles (1/(days $\times$ contact))
$\beta_H$	Transmission rate from hospitalized cases to susceptibles (1/(days $\times$ contact))
$1/\alpha$	Duration of latency period (days)
$\gamma_U$	Recovery rate for undetected infectives (1/days)
$\gamma_I$	Recovery rate for non-hospitalized cases (1/days)
$\gamma_H$	Recovery rate for hospitalized cases (1/days)
$\rho$	Probability of detection on symptoms onset
$\sigma$	Probability of detection post mortem
$\eta$	Hospitalization rate (1/days)
$\delta_I$	Disease-induced death rate for non-hospitalized cases (1/days)
$\delta_H$	Disease-induced death rate for hospitalized cases (1/days)

with the numerical integration of the ODE model (1) over a sufficiently long time interval (cf. **Supplementary Material**). Parameter sensitivity analysis of model quantities of interest, such as the basic reproduction number, the number of hospitalized cases at the outbreak peak and the final size, were investigated by means of Partial Rank Correlation Coefficients (PRCC) analysis<sup>16</sup>. Latin Hypercube Sampling<sup>17</sup> was used to generate a representative sample set of tuples of parameters from the parameter ranges indicated in the **Supplementary Material**. Using PRCC, we can rank the effect that each parameter has on the outcome, when other parameters are simultaneously varying in the given ranges. Calculation of PRCC allows us to determine the strength of statistical relationships that exist between each input parameter and the output variable. Parameters with PRCC larger (respectively, smaller) than zero are positively (respectively, negatively) correlated with the quantity of interest.

## Results

### Reproduction number

Possibly among the most interesting quantities to identify in the early phase of an outbreak, the *basic reproduction number* (denoted by  $\mathcal{R}_0$ ) is used in mathematical epidemiology as an indicator for the severity of the disease spread. This  $\mathcal{R}_0$  is a metric which indicates the average number of secondary infections generated in a fully susceptible population by one infected host over the course of the infection. In most epidemiological models,  $\mathcal{R}_0 > 1$  leads to a disease outbreak, whereas for  $\mathcal{R}_0 < 1$  the disease will not spread in the population.

The basic reproduction number of system (1) can be calculated analytically (cf. **Supplementary Material**), e.g., by means of the next-generation matrix approach<sup>18</sup>, and is given by

$$\mathcal{R}_0 = \underbrace{\frac{\beta_E}{\alpha}}_{=: \mathcal{R}_0^E} + \underbrace{\frac{\rho\beta_I}{\gamma_I + \delta_I + \eta}}_{=: \mathcal{R}_0^I} + \underbrace{\frac{\rho\beta_H\eta}{(\gamma_I + \delta_I + \eta)(\gamma_H + \delta_H)}}_{=: \mathcal{R}_0^H} + \underbrace{\frac{\beta_U(1-\rho)}{\gamma_U}}_{=: \mathcal{R}_0^U}. \quad (2)$$

The summands of  $\mathcal{R}_0$  account for contacts with presymptomatic cases ( $\mathcal{R}_0^E$ ), undetected infectives ( $\mathcal{R}_0^U$ ), detected non-hospitalized ( $\mathcal{R}_0^I$ ) and hospitalized ( $\mathcal{R}_0^H$ ) cases. The reproduction number evolves in time with time-dependent parameters, which might vary e.g. because of intervention measures, and with variations in the susceptible portion of the population. As the outbreak evolves, people who go through infection are either immunized or die, and the susceptible population decreases (vaccination-induced immunity is not considered here but would also contribute to lowering the susceptible population). The *effective reproduction number*  $\mathcal{R}_t$  at time  $t$  can be obtained by the same formula, substituting the values of time-dependent parameters in (2) and multiplying this value by the susceptible fraction of the population,  $S(t)/(N - D(t))$ .

### Modeling the lockdown at time $T$ and migration from region A

Inspired by the history of the COVID-19 outbreak in China and Italy, we use the previously presented transmission network (1) to describe an outbreak spreading from one geographic region A, to the rest of the country (here considered as one single region, B). The country itself is assumed to be isolated, that is, migration inward/outward

is not considered here. The outbreak is assumed to have started in region A, that is, numerous cases were detected in a short time during the early phase of the epidemic. The prevalence (cases/population) of COVID-19 in region A being much higher than in the rest of the country (region B), the government decides that region A is going to be confined at some time  $T$ . This motivates many people to move from region A to region B, just before region A is effectively isolated from the rest of the country. We adapt the mathematical model (1) to reproduce this situation and understand how the lockdown time  $T$ , and the migration of a fraction  $\phi \in (0,1)$  of the *unconstrained population* (that is, everyone except detected, hospitalized and deceased persons) from region A to region B could affect the evolution of the epidemic in the two regions. Detected cases, hospitalized cases and deceased individuals, cannot leave region A at any time. Transitions between region A and region B, which are assumed to be in balance before the lockdown, are stopped after time  $T$ . Migration is modeled as an impulse at time  $T$ , simplifying a possible continuous transition from A to B in the days immediately ahead the lockdown. This means we impose the conditions

$$\begin{aligned} Z^B(T_+) &= \phi Z^A(T_-) + Z^B(T_-), \quad Z^A(T_+) = (1 - \phi)Z^A(T_-), \quad \text{for } Z = S, E, U, R, R_U, \quad \text{and} \\ X^j(T_+) &= X^j(T_-), \quad \text{where } X = I, H, D \quad \text{and } j = A, B, \end{aligned} \quad (3)$$

where  $T_-$  denotes the time just before migration occurred and  $T_+$  denotes the time of migration, respectively.

After the lockdown ( $t \geq T$ ) the outbreak evolves independently in the two regions. Further, concurrent with the lockdown, other non-pharmaceutical intervention measures, such as encouraged social distancing, school closure, shut down of many economic activities, increased testing activity, improved hospital capacity etc., might be introduced in both regions, possibly in a different manner or with different timing.

### Final size

The final size relation is an analytic formula to predict the total number of individuals in the population who are infected over the course of the epidemic<sup>12</sup>. By means of analytical methods (cf. **Supplementary Material**), we derived a final size relation taking into account the lockdown time ( $T$ ) and the induced exodus at the time of the lockdown. This formula provides a reliable approximation of the cumulative cases, when the total population and the parameters do not vary too much after time  $T$ . Assume the status of the system, including the numbers of cases, hospitalizations, deaths, recoveries, and active population ( $\mathcal{N}$ ), in one region right after the lockdown time (denoted below by the subscript  $T_+$ ) can be obtained from available data or approximated by means of the model (1). Then the final size formula

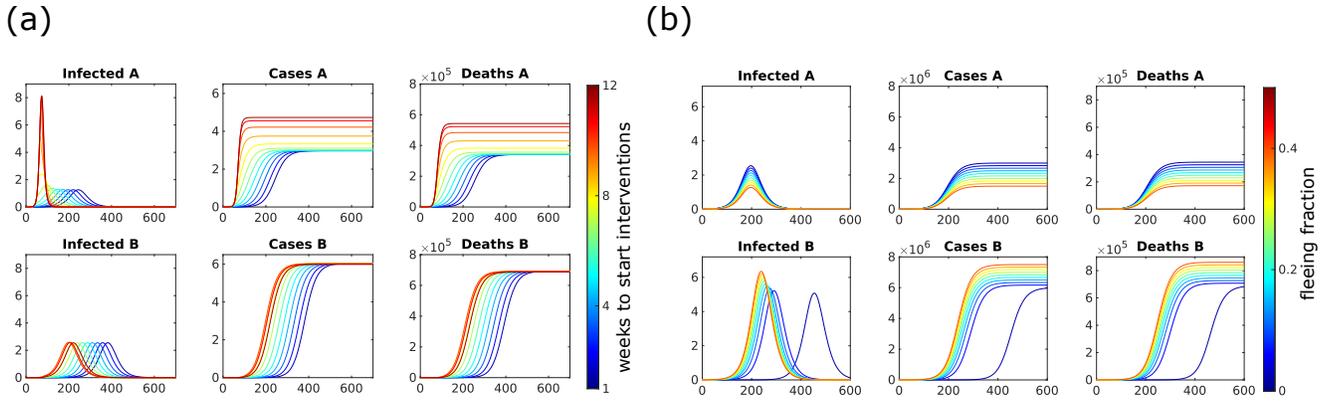
$$\begin{aligned} \ln S_{T_+} - \ln S_\infty &= \Phi(\mathcal{N} - S_\infty) + \tilde{b}_1(\chi - 1)I_{T_+} + (\chi - 1)\tilde{b}_2H_{T_+} + (\chi(\tilde{b}_3 - \tilde{a}_4) + \Phi - \tilde{b}_3)D_{T_+} \\ &\quad + (\chi(\tilde{b}_4 - \tilde{a}_2) - \tilde{b}_4)R_{T_+} - \chi\tilde{a}_1U_{T_+} - \chi\tilde{a}_3R_{T_+}^U, \end{aligned} \quad (4)$$

allows us to determine the number of cases occurring over the complete course of the outbreak, without requiring model integration after time  $T$ . Observe that the fleeing ( $\phi$ ) is included in the relation (4) into  $\mathcal{N}$  and the terms with  $T_+$ -subscript, following the conditions (3). Here  $S_\infty$  denotes the portion of population which remains susceptible at the end of the outbreak. If  $S_0 \approx N_0$  is the susceptible (total) population at the beginning of the outbreak, then  $S_0 - S_\infty$  corresponds to the cumulative number of cases at the end of the outbreak. The coefficients  $\chi, \Phi, \tilde{a}_j, \tilde{b}_j$ ,  $j = 1, \dots, 4$  in equation (4) are obtained from the parameters of model (1) as detailed in the **Supplementary Material**. Formula (4) is used below to analyze the sensitivity of the total number of cases with respect to model parameters.

### How lockdown and intervention measures affect the severity of the outbreak

The proposed model (1) with the lockdown scenario (cf. **Methods**) is used to investigate the sensitivity of case and death counts with respect to the intervention parameters, such as the time of lockdown of the source region A, the fraction of population fleeing region A due to the lockdown announcement, and the strength and timing of intervention measures in both regions. We consider a hypothetical scenario which is suitable to describe various infectious diseases and populations, though inspired by COVID-19 spreading in a fully susceptible population. Region A, where the outbreak starts, is assumed to have a population of 25 million. Though the population size in region B can vary, this region is always assumed to be disease-free until lockdown of region A occurs. Region B is parametrized as region A, unless explicitly mentioned. We set the parameter values such that the initial reproduction number in region A is  $\mathcal{R}_0 \approx 3.8$ , while the *beginning of the outbreak* is defined to occur with 20 cumulative detected cases and two deaths. For further details on the parametrization used for this section, we refer to the **Supplementary Material**.

Figure 2(a) shows how the cumulative number of detected cases and deaths depend on the lockdown time  $T$ . At the time of lockdown, which varies here between one and twelve weeks after the beginning of the outbreak in region A,

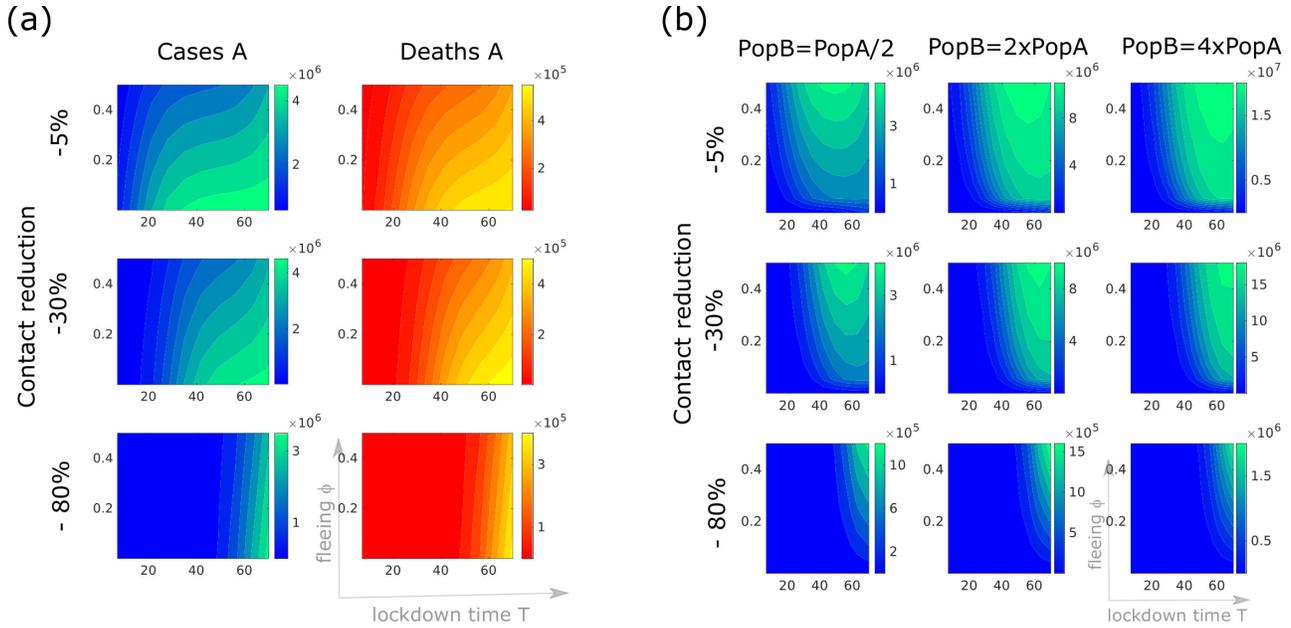


**Figure 2. Detected infected cases and deaths depending on (a) the lockdown time  $T$  and (b) the fleeing fraction of exposed/undetected individuals at the time of the lockdown.** An initial outbreak starts off with 20 cases and two deaths in region A (population 25 million), where lockdown is established at time  $T$ , immediately followed by lockdown-induced migration of a fraction of the unconstrained population from region A to the disease-free region B (population 50 million). (a) The lockdown time  $T$  varies from one to twelve weeks after the initial reporting, while the fleeing fraction is fixed to 1% of the unconstrained population in region A; (b) 0.01% to 50% of the unconstrained population moves from region A to region B at the fixed time  $T = 21$  days after the initial reporting. At the time of lockdown, control measures are applied in both regions in order to restrict contacts by 60% and reduce transmission.

1% of the unconstrained population flees from region A to region B. Lockdown is coupled with stronger control measures, both in region A and region B, which assume reduction in contact rates by 60% to be maintained until the end of the observations (here about two years after beginning of the outbreak). This reduces the reproduction number in both regions, but not enough to prevent an outbreak also in region B ( $\mathcal{R}_T^B \approx 1.5$ ). The panels in Figure 2(a) show that in both regions, the time of the lockdown strongly affects both the number of cases and the peak of the epidemic wave. Figure 2(b) shows the cumulative number of detected cases and deaths, depending on the fraction of exposed and undetected infected individuals leaving region A for region B. Here the lockdown time  $T$  is fixed at three weeks after the beginning of the outbreak, whereas the migrating population  $\phi$  varies (between 0.01% and 50% of the unconstrained population in region A at the lockdown time). Increasing the fleeing fraction  $\phi$ , the peak in cases decreases in region A, whereas in region B, it increases and advances in time.

Figure 3(a) shows the variation in cumulative detected cases and deaths in region A depending on the lockdown time  $T$ , the fraction  $\phi$  of population leaving region A and the effectiveness of intervention measures in reducing contacts. We project here the number of cases and deaths predicted by the model two months after lockdown, while control measures are maintained for the whole period. While migration does not particularly affect the evolution of the outbreak in region A in case of early intervention, the later the lockdown and the weaker the control measures, the higher will be the number of cases and deaths. In case of late lockdown time (about two months after the beginning of the outbreak), a significant fleeing population reduces the number of cases/deaths in region A, in agreement with what observed in Figure 2. Opposite is the effect of migration from region A on cumulative cases (and deaths, not shown here) in region B. Figure 3(b) shows the importance of applying stronger control measures in region B when region A is locked. When contact restrictions are more severe, the size of the population of region B does not importantly affect the severity of the outbreak. If contact restriction is minimal in region B, a higher populated target region allows a longer and more severe outbreak.

One of the most difficult questions to answer during an ongoing epidemic concerns the duration of restrictive control measures. How long should these restrictions be in place? Several criteria could be applied to determine the end of the restrictive measures<sup>19</sup>. Here we assume that a region is released one week after reaching the peak in the daily incidence of reported cases. As in the above simulations, we assume that the outbreak spreads from region A to region B when the former is isolated. The result is shown in Figure 4. In both regions, at the time of lockdown, control measures are applied to restrict contacts by 5% to 60%, and maintained for 7 days after the peak in the daily incidence in the respective region is reached. More restrictive measures postpone (and lower, cf. Figure 3(b)) the peak of the outbreak, however this prolongs the time that the control measures should be in place. On the other hand, migration and late intervention accelerate the speed of the epidemic in region B, and hence the peak comes



**Figure 3. Sensitivity of cases in region B depending on the lockdown time  $T$ , the fraction  $\phi$  of population leaving region A, the strength of intervention measures in reducing contacts, and the population size of region B.** An initial outbreak starts off with 20 detected cases and two deaths in region A (population 25 million) which is locked at time  $T$  (horizontal axes indicate days after the initial reporting) and lockdown-induced migration (vertical axes) occurs. At the time of lockdown, control measures are applied to restrict contacts by 5% (first row), 30% (second row) or 80% (third row), and maintained for two months. Color codes represent cases/deaths projected at the end of the two months following model (1): (a) the number of detected cases (left column) and deaths (right column) in region A, (b) detected case numbers in region B depending on the size of the population in region B with respect to that of region A (columns).

earlier (is much higher, brings more cases and more deaths, cf. Figure 2) and makes the lockdown time shorter.

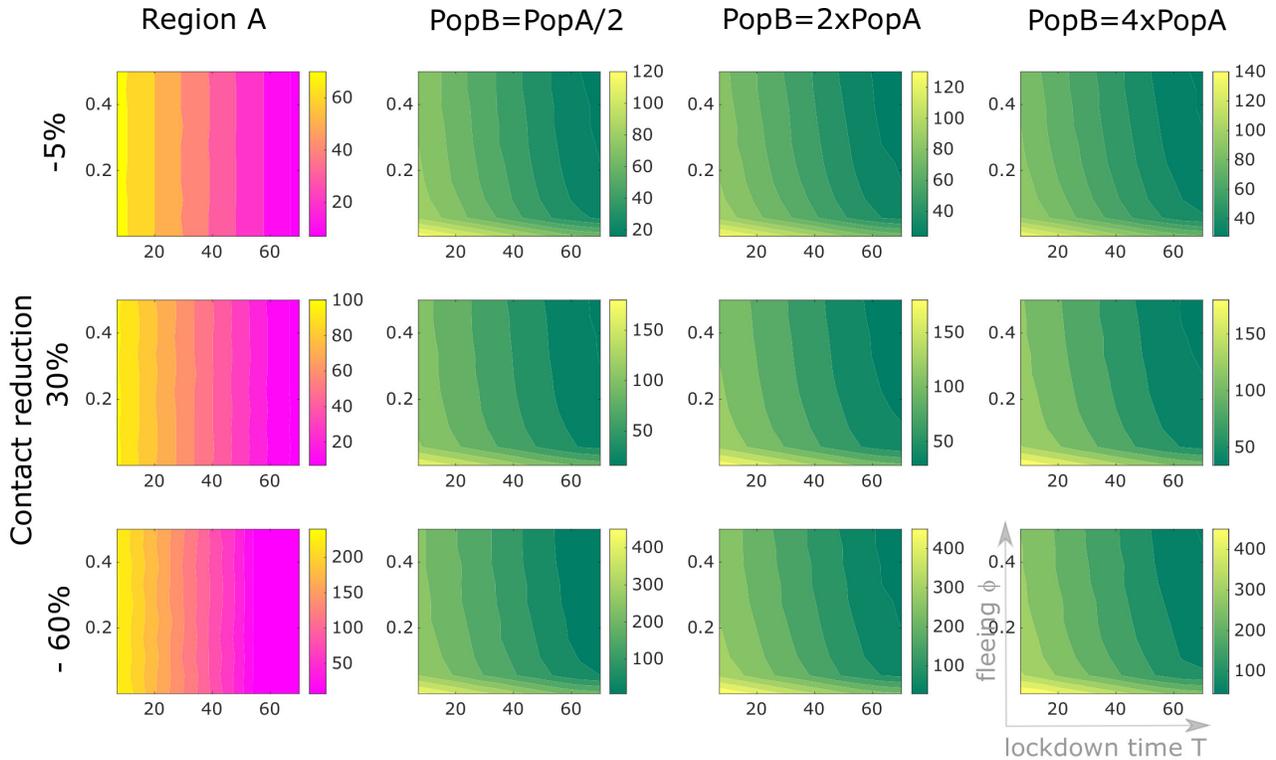
So far we have assumed that control measures in region B are applied when region A is isolated. Figure 5 illustrates the damage that could be induced by a delayed reaction in region B combined with the application of not sufficiently strong control measures. The panels show the number of cases in region B three months after lockdown of region A occurred (here  $T$  is fixed on day 21 since the beginning of the outbreak), assuming that the interventions in region B is delayed by  $T_B$  days with respect to region A.

### COVID-19 in Italy as an example

In the previous section we used the model (1) to study the spread of a potential outbreak on two hypothetical geographic regions, without specific application to real data. Here we consider the recent COVID-19 outbreak in Italy and apply model (1) to study the lockdown and the early outbreak dynamics. The COVID-19 outbreak hit Italy in early 2020 as the first country in Europe. While preliminary control measures were introduced locally at the end of February as first cases were detected, at the beginning of March a large area including Lombardy and several provinces in Emilia-Romagna, Veneto, Piedmont and Marche were declared as “red zone” and isolated<sup>20</sup>. Shortly after, even more restrictive control measures were applied on the whole country<sup>3</sup>.

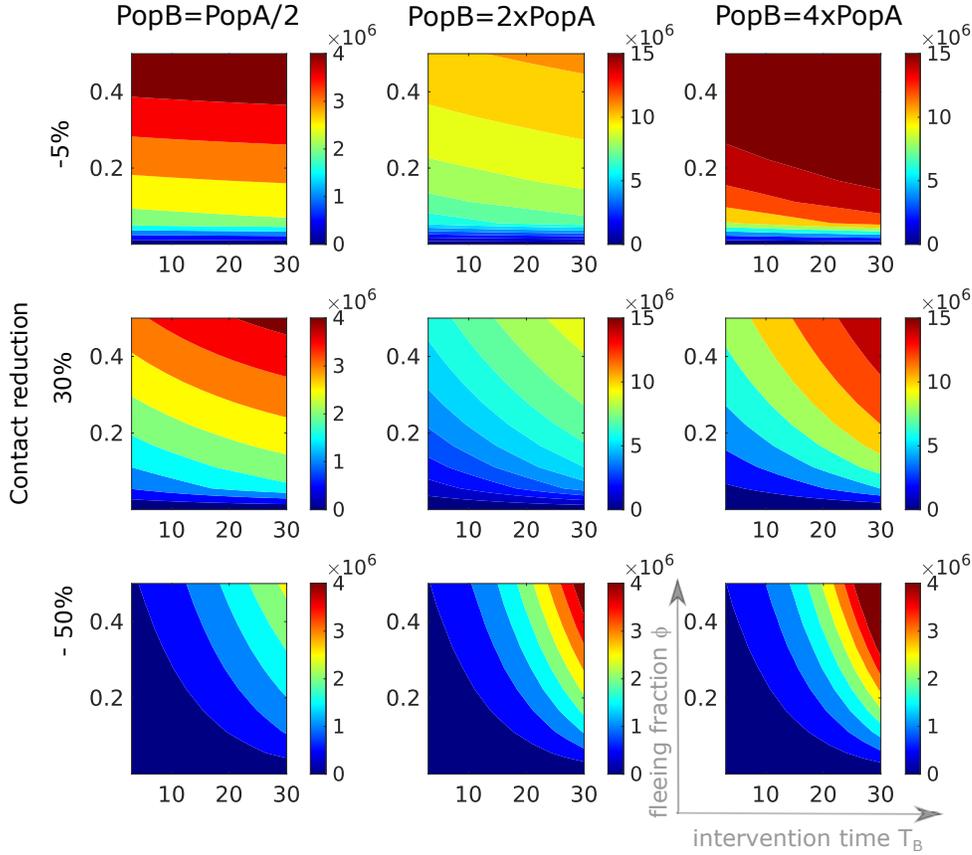
#### Data and setting for parametrization

The publicly available data-set provided by the Italian Ministry of Health (Ministero della Salute) and the national Civil Protection Department (Dipartimento della Protezione Civile)<sup>21</sup> was used for this study. Time series for hospitalized (totale\_ospedalizzati), daily case incidence (nuovi\_positivi) and cumulative cases (totale\_casi), as well as deaths (deceduti) were obtained from the *dati-regioni* repository<sup>21</sup>. Data for region A was prepared accounting for the isolated provinces<sup>3,22</sup>. As hospitalization and death counts are available at regional level, but not at the level of provinces, we define region A as the area composed of the whole regions of Lombardy, Emilia-Romagna,



**Figure 4. How long should control measures be in place?** An initial outbreak starts off with 20 cases and two deaths in region A (population 25 million) which is then locked at time  $T$  (horizontal axes indicate days after the initial reporting) and lockdown-induced migration (vertical axes) occurs. In both regions, at the time of lockdown, control measures are applied to restrict contacts by 5% (first row), 30% (second row) or 60% (third row), and maintained for 7 days after the peak in the daily incidence in the respective region is reached. Color codes represent the duration in days of the control period in region A (first column) and region B (second to fourth columns), also depending on the population size in region B (being half, twice or four times the population in region A).

Veneto, Piedmont and Marche (about 25 million people). In reality, only fourteen largely-neighboring provinces in Emilia-Romagna, Veneto, Piedmont and Marche were isolated, together with Lombardy which is the most populous region in Italy and accounts for the majority of COVID-19 cases. Region B is defined as the rest of the country (about 35 million people). The model is calibrated on daily incidence, deaths and hospitalized time series for the two regions. Based on estimates from previous studies, we fix the average duration of the latent phase ( $1/\alpha \approx 5.5$  days<sup>23,24</sup>), and the duration of the infectious periods in undetected (ca. 7 days<sup>24,25</sup>) and in detected cases (ca. 10 days<sup>26,27</sup>). For model parameters that cannot be obtained from other references nor estimated reliably with the available data, we fix values based on plausible assumptions. For example, we fix the ratio between  $\beta_E$  (respectively,  $\beta_I$  and  $\beta_H$ ) and  $\beta_U$  and estimate the latter. It is unclear which amount of secondary cases resulted from presymptomatic transmission, with estimates ranging from 6.4%<sup>28</sup> to 44%<sup>29</sup>. Here we assume that  $\beta_E = 0.2\beta_U$ , accounting for a two days infectious period in the presymptomatic phase. As undetected infectives have possibly mild or no symptoms, we assume that they do not restrict their contacts to others, and therefore have higher transmission rates than detected cases ( $\beta_I = 0.25\beta_U$ ). Hospitalized patients are supposed to have very limited contacts compared to undetected cases ( $\beta_H = 0.1\beta_U$ ). The detection ratio  $\rho$  is estimated based on Table 1 of the very recent report by the Italian National Institute of Statistics and the Ministry of Health on the seroprevalence in Italy<sup>30</sup>. From the data shown in this report, we could compute a seroprevalence 4.53% (weighted average) for region A and 0.96% for region B, as of July 20, 2020. As of the same date, the cumulative cases in region A and region B report about 182,000 cases detected in region A and about 62,000 cases detected in region B<sup>21</sup>. Rounding the seroprevalence to 5% for region A, this leads to 1.25 million estimated COVID-19 cases for region A, out of which we know only 182,000, henceforth a detection ratio  $\rho_A \approx 14.5\%$ . Analogously, rounding to 1% seroprevalence for region B, we obtain  $\rho_B \approx 17.5\%$ . We assume that the detected cases are hospitalized with a rate  $\eta = p_H/2$ , where the probability  $p_H \in (0, 1)$  of being hospitalized two days after symptoms onset is estimated from the data. Once the detected cases are hospitalized, they are assumed

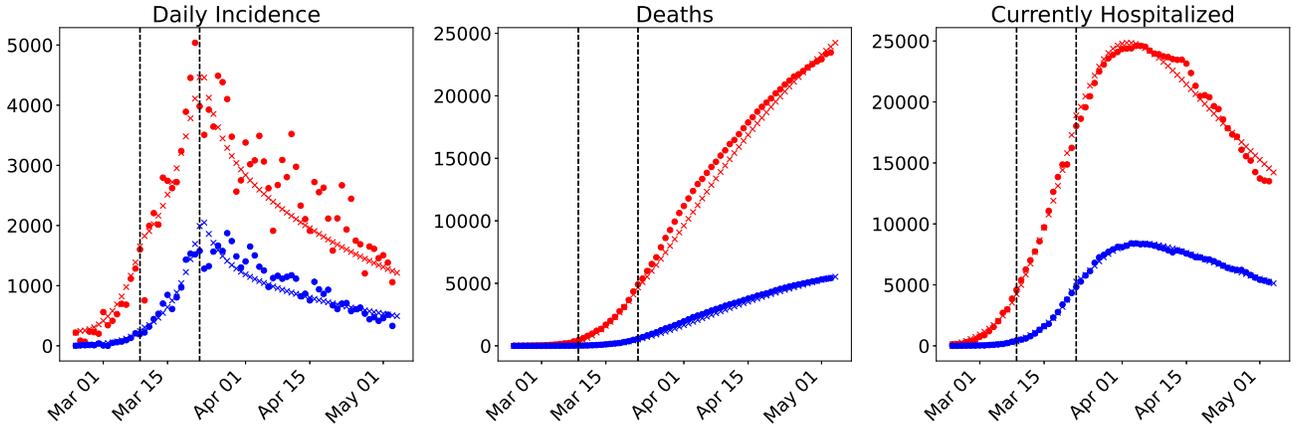


**Figure 5.** Cumulative detected cases in region B three months after lockdown of region A depending on the fraction  $\phi$  of population leaving region A, the intervention time  $T_B$  in region B, the effectiveness of intervention measures in reducing contacts, and the population size of region B. Region A is isolated 21 days after the beginning of the outbreak and migration towards region B occurs. For each panel, the vertical axis denotes the variation in the fleeing fraction from A ( $\phi$ , between 0.1% and 0.5). On the horizontal axis, the reaction time  $T_B$  indicates how many days passed for control measures in region B to be applied since isolation of region A. Control measures are applied in region B to restrict contacts by 5% (first row), 30% (second row) or 50% (third row), and maintained until the end of the simulations. Cumulative detected cases are projected also depending on the size of population in region B with respect to that in region A (columns). Note the different scaling of the color legend in the panels.

to either recover with a probability  $p_{HR} \in (0, 1)$  after 10 days, or die with a probability  $(1 - p_{HR})$ . This leads to a recovery rate of  $\gamma_H = p_H p_{HR} / 10$  in the hospitals and a death rate of  $\delta_H = p_H (1 - p_{HR}) / 10$ . The probability  $p_{HR}$  is estimated from the data. Among those detected cases that do not get hospitalized, we assume that  $p_{IR} = 99\%$  of the cases recover after on average 10 days from symptoms onset, while the remaining 1% die. This means that we can interpret the recovery rate of the detected cases that are not hospitalized as  $\gamma_I = (1 - p_H) p_{IR} / 10$  and the death rate as  $\delta_I = (1 - p_H) (1 - p_{IR}) / 10$ . We assume that  $\sigma = 0.1\%$  cases are detected post mortem.

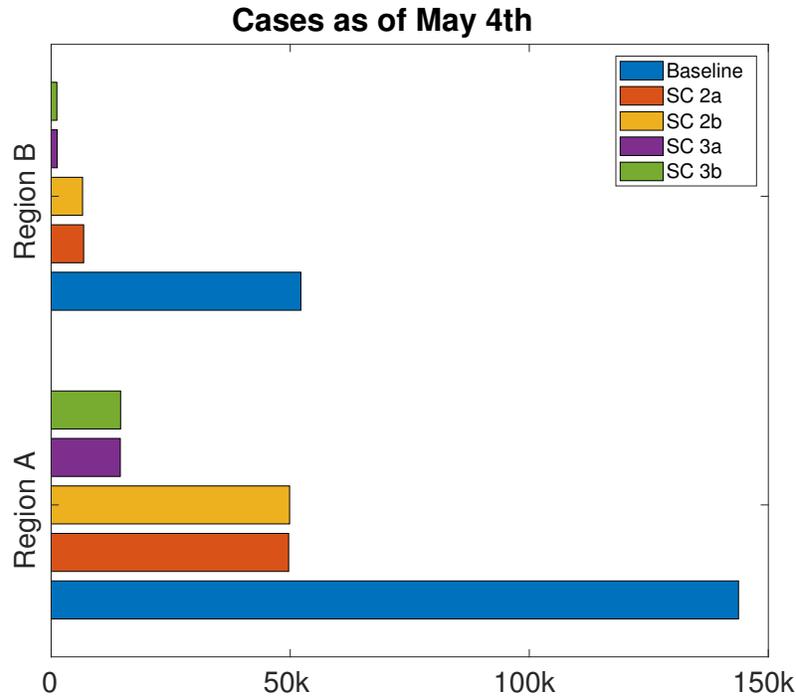
The initial time point for simulations is set to February 24, 2020. Most of the initial conditions for solving the ODE system, such as the number of deaths or hospitalized cases could be obtained directly from the data. However, there is no data available for the exposed and undetected infections. These values are approximated making use of the detection ratio  $\rho$  and the average latency period  $1/\alpha$ , with  $\rho E(0)$  roughly corresponding to the case incidence six days later, and  $U(0)$  corresponding to the undetected fraction  $(1/\rho - 1)I(0)$  of cases at the beginning of the observations. The initial conditions used for simulations are specified in the **Supplementary Material**.

To parametrize model (1) taking into account the numerous intervention measures adopted in the early phase of the outbreak we define three time periods: before lockdown (February 24 – March 8), under initial lockdown measures



**Figure 6. Model fit for the early dynamics of the COVID-19 outbreak in Italy.** Reported data (dots) and model results (crosses) for daily incidence (left panel), deaths (middle panel) and hospitalized cases (right panel) in region A (Lombardy, Emilia-Romagna, Marche, Piedmont, Veneto; red) and region B (rest of the country; blue). Parameter values are estimated as indicated in three different time intervals (separated by vertical lines in the figure): pre-lockdown (February 24 – March 8), first lockdown measures (March 9 – 21), and extended lockdown measures (March 22 – May 4), cf. **Supplementary Material**.

(March 9 – 21) and under stricter lockdown measures introduced later (March 22 – May 4). At the lockdown time  $T$  (March 9), we assume that 0.5% ( $\phi = 0.005$ ) of the unconstrained population in region A (roughly 130,000 people) migrate to region B. Given the available data, the parameter  $\phi$  cannot otherwise be reliably estimated. Nevertheless, the estimated number is plausible, given the reported large migration from the north of the country<sup>5,6</sup>. The model parameters for the two regions were estimated for each time period independently. When available, the estimated parameters from the previous time period were given as initial guesses for the optimizer in the next time period. The estimated and fixed model parameters for the three time periods are summarized in the **Supplementary Material**. The result of the model fit is given in Figure 6. The parameter estimation results highlight a clear reduction in the contact rates ( $\beta_U$ ), lowering the basic reproduction number in region A from  $\mathcal{R}_0 = 3.7$  at the beginning of the outbreak to  $\mathcal{R}_t = 0.76$  at the end of March, in agreement with estimates obtained by other groups<sup>31</sup>. The decrease in the hospitalization rate  $\eta$  and increase in the mortality rates  $\delta_I, \delta_H$  can be explained with the quickly rising number of hospitalizations and deaths in the second half of March and beginning of April. Fitting data for both region A and region B, we clearly see that control measures were effectively in place also in region B, in which the reproduction number was also lowered to  $\mathcal{R}_t = 0.7$  at the end of March. Compared to the abstract scenarios considered in the previous section, we are here in a setting where the population in region B is about 1.5 times larger than in region A, however not fully disease-free at the time of the outbreak. First cases outside of region A were reported at the end of February, quickly isolated, and could be traced back to clusters in Lombardy or Veneto<sup>3</sup>. While control measures were introduced more severely in region A, awareness and other major interventions such as school closure were applied on the whole Italian territory even before the lockdown on March 9. After May 4, the lockdown was released and most severe control measures were relaxed in more steps. As the focus of this work was concerned with the isolation of a specific region and not with the full parametrization of the ongoing outbreak, we do not consider here the relaxations of the control measures, nor data after May 4. Starting from the obtained data-fit, we simulate different possible scenarios for a hypothetical intervention with stricter and/or earlier measures. We compare (Figure 7) the number of cumulative cases in region A and region B as of May 4 for different scenarios, with *baseline* representing the parameter setting as of the fit in Figure 6. For all other scenarios we assume that most severe restrictions (parameter set as of March 22, cf. **Supplementary Material**) are in place from the lockdown time. In scenario 2 (SC2) we keep the lockdown time as for the data fit and assume that strict intervention measures are in place from March 9, whereas in scenario 3 (SC3) we advance the lockdown by one week to March 2. In both scenarios we consider a fraction (0.5%) of unconstrained individuals (a) to not flee (b) to flee region A immediately before the lockdown. The projected scenarios show that stricter interventions from the beginning of the lockdown could have reduced the number of cumulative cases as of May 4 by at least one third, and that importation of cases has far less impact than timely intervention.



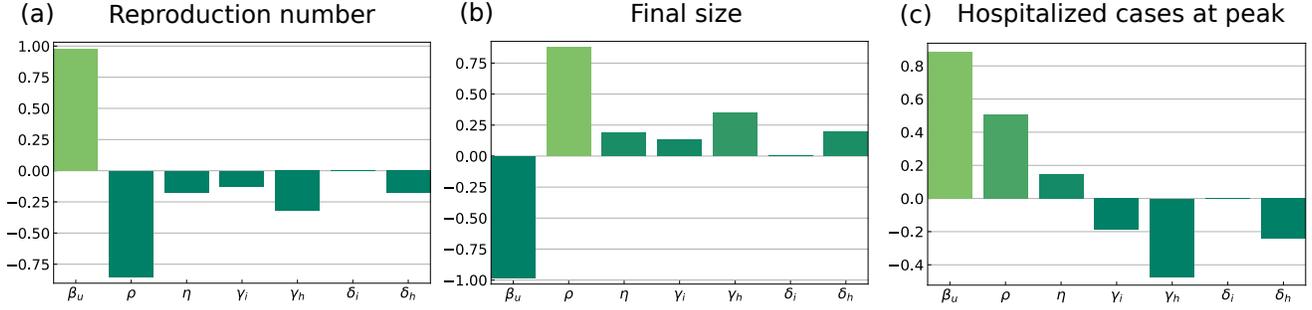
**Figure 7.** Scenario comparison for the cumulative cases of COVID-19 in Italy as of May 4, 2020. Cumulative detected cases in region A and region B are simulated for different scenarios: (baseline) the setting as of fit in Figure 6; For all other scenarios severe restrictions (parameter set as of March 22) are in place from the lockdown time: (SC2) lockdown on March 9 (a) with and (b) without fleeing population at lockdown; (SC3) lockdown anticipated to March 2 (a) with and (b) without fleeing population at lockdown.

### Sensitivity analysis

To quantify the effect that model parameters, in particular those which might be affected by control measures, have on critical quantities of the epidemic, we performed a global sensitivity analysis using PRCC on seven model parameters (cf. **Methods** and **Supplementary Material**). Figure 8 shows the result of the sensitivity analysis of the reproduction number  $\mathcal{R}_0$ , the final size ( $S_\infty$ ) as of formula (4), and the number of hospitalized cases at the outbreak peak in the target region A (analogous results for region B, not shown here). Panels in Figure 8 show that  $\beta_U$  and  $\rho$  have the strongest correlation with all the three model outputs. Therefore, a small increase in the transmission rate of undetected cases or a small decrease in the detection ratio can significantly increase the spread of the disease ( $\mathcal{R}_0$ ) and the number of hospitalized cases at the outbreak peak eventually leaving much fewer susceptibles in the population at the end of the epidemic ( $S_\infty$ ). Recovery, hospitalization and death rates are negatively correlated with the reproduction number (positively with  $S_\infty$ ). The number of hospitalized cases at the outbreak peak is also strongly affected by the recovery rate for the  $H$  compartment ( $\gamma_H$ ).

### Discussion

A major epidemic outbreak might require authorities to introduce various control measures, from encouraging social distancing to closing educational and economical activities. The goal is to limit the number of victims and to prevent the health care system from collapsing. In the present study, we have established a model for the spread of an epidemic outbreak considering the effects of an exodus of large masses of people from infected areas before restrictive control measures are introduced. Such a phenomenon was reported in the occasion of the COVID-19 outbreak, first in Wuhan, China and then in Northern Italy. The compartmental model introduced in our work differentiates infected people based on the severity of their infection, accounting also for presymptomatic transmission, undetected infections and hospitalized cases. We consider two geographic regions, with the disease being initially confined in region A, where also severe control measures are applied and from where people flee to a previously (ideally) uninfected region B. In contrast to previous studies for the spread of COVID-19, we consider here the geographic



**Figure 8.** Partial rank correlation coefficients (PRCCs) of seven main model parameters for (a) reproduction number, (b) final size ( $S_\infty$ ) and (c) the number of hospitalized cases at peak in region A. Parameters with PRCC larger than zero are positively correlated with the quantity of interest, whereas parameters with negative PRCC are negatively correlated. Parameters were varied within the ranges given in the **Supplementary Material**.

distribution of cases, but relax the classification of the individuals into age groups<sup>24,32,33</sup>. Calculating the basic reproduction number (2), we identified its four components that account for transmissions from presymptomatic, undetected infectives, detected but not hospitalized, and hospitalized cases. Further we have derived a new type of final size relation (4), a formula serving for the estimation of the total number of infectives during the whole duration of the epidemic. A final size formula in non-autonomous epidemic models was previously obtained, for example, in a model for Ebola Virus Disease<sup>34</sup>. The novelty in this work is that the final size relation obtained here covers models in which not only the parameters, but also the population size varies at some specific time point  $T$  (in the considered example due to a fleeing fraction of individuals from one region to another). The result presented here is novel and flexible, so that it is suitable to study various outbreaks of infectious diseases. Further, the same approach could be used for describing migration out of region A and distribution of a fraction  $\phi_i$  of individuals moving from region A to a region  $B_i$ ,  $i = 1, \dots, n$  for an arbitrary number  $n \in \mathbb{N}$  of destination regions.

By means of numerical simulations, we have assessed the short-term effect of various parameters on the number of cases and the number of deaths. Figures 2–5 show that the earlier the lockdown happens, the slower and less dramatic the epidemic evolves, in agreement with previous studies<sup>7–9</sup>. A late lockdown importantly advances the peak in the number of currently infected people and leads to significantly more deaths and total cases. The fraction of people moving from region A to region B, however, rather affects the latter region. The more individuals moving from region A to region B, the earlier the peak in the infective curve of the target region. The strength and timing of intervention measures in region B is crucial to contain the number of cases and deaths due to importation of the disease from region A (Figure 3(b)). We have first assumed that control measures are applied at the same time of the lockdown both in region A and in region B, simulating the behavior of people fleeing isolation and joining their relatives, rather than avoiding restrictive measures. Delays of up to a month in intervention in region B, compared to the lockdown time of region A, can be balanced out with the maintenance of stricter control measures also in region B. More restrictive measures postpone (and lower, cf. Figure 3(b)) the peak of the outbreak, however this prolongs the time intervention measure should be in place, if reaching the peak of the infections curve is the criterion for controls to be relaxed.

As an example, we have fitted our model to the Italian COVID-19 data (Figure 6), the most affected areas in the country defining region A, and the rest of the country corresponding to region B. Projection of different scenarios (Figure 7) shows that timely and strict intervention could have importantly lowered the number of cumulative cases (in region A by at least 60%) and that migration at the time of lockdown possibly played a minor role for the spread of the disease in Italy. A limitation of our work is that in our setting, migration from region A to region B is assumed to occur as an impulse at time  $T$ . Given the available data, the fleeing fraction ( $\phi$ ) could not be reliably estimated and it was fixed to obtain a plausible value for the reported large migration from the north of the country. While it was initially observed that a large number of people fled the quarantined regions immediately before these were confined, it is yet debated if the exodus was not rather diluted over the time starting on February 23<sup>35</sup>. A repeated or continuous migration from region A to region B could be included in the model simulations, though this would increase uncertainty in the parametrization. Similarly, the assumption that all unconstrained compartments move in the same proportion from region A to region B could be relaxed, and different fleeing scenarios could be assumed. In all cases, the final size formula (4) would still hold true, as long as the parameters before/after intervention are

known and the values for detected cases, hospitalized, recovered and deceased people at time  $T$  can be estimated from the early outbreak data. In this study, we focused on the timing and strength of intervention, but did not take into account relaxation of control measures. Our setting is realistic for short-term intervention scenarios, such as two to three months, as was the case in many countries during the COVID-19 pandemic. For long-term predictions, assuming maintenance of strict control measures over several months leads to an unrealistically low estimate for the number of total cases and deaths.

## Acknowledgements

The work of MVB was supported by the LOEWE focus CMMS. AD was supported by the grant NKFIH PD 128363 and the János Bolyai Research Scholarship of HAS. NB was supported by the grant EFOP-3.6.2-16-2017-00015 and ZsV by the grant NKFIH KKP 129877. The work of GR was supported by NKFIH FK 124016 and TUDFO/47138-1/2019-ITM.

## Contributions

Conceptualization: MVB, GR; Modeling: MVB, GR; Mathematical analysis: MVB, AD; Data curation and analysis: NB, HVV; Literature research: AD, MVB, HVV; Software development: ZV, NB, HVV; Calibration: HVV. Numerical simulations: HVV, ZV, MVB. Writing (original draft): MVB, AD.

## Additional information

**Accession codes:** not applicable; **Competing interests:** no competing interests.

## References

1. BBC News. *China coronavirus: Lockdown measures rise across Hubei province* Jan 23 2020 <https://www.bbc.com/news/world-asia-china-51217455>(last accessed Aug 7 2020)
2. Zhao S. *et al.* Quantifying the association between domestic travel and the exportation of novel coronavirus (2019-nCoV) cases from Wuhan, China in 2020: a correlational analysis. *J. Travel Med.* **27(2)**, March 2020, taaa022.
3. Wikipedia. *COVID-19 pandemic lockdown in Italy* [https://en.wikipedia.org/wiki/COVID-19\\_pandemic\\_lockdown\\_in\\_Italy](https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdown_in_Italy) (last accessed July 29, 2020)
4. Corriere della Sera. *Coronavirus, la grande fuga da Milano prima del decreto che isola la Lombardia* March 8, 2020 [https://www.corriere.it/cronache/20\\_marzo\\_08/coronavirus-grande-fuga-milano-prima-decreto-che-isola-lombardia-9b87a15a-60df-11ea-8d61\438e0a276fc4.shtml](https://www.corriere.it/cronache/20_marzo_08/coronavirus-grande-fuga-milano-prima-decreto-che-isola-lombardia-9b87a15a-60df-11ea-8d61\438e0a276fc4.shtml) (last accessed Aug 11, 2020)
5. La Stampa. *Fuga dal Nord: ventimila persone rientrate in Sicilia nel weekend* March 11, 2020 <https://www.lastampa.it/topnews/primo-piano/2020/03/11/news/fuga-dal-nord-ventimila-persone-rientrate-in-sicilia-nel-weekend-1.38580593> (last accessed Aug 11, 2020)
6. La Stampa. *In fuga dal Nord Italia, 11 mila individuati in Sardegna: per tutti scattata la quarantena obbligatoria* March 11, 2020 <https://www.lastampa.it/cronaca/2020/03/11/news/in-fuga-dal-nord-italia-11-mila-individuati-in-sardegna-per-tutti-scattata-la-quarantena-obbligatoria-1.38580219> (last accessed Aug 11, 2020)
7. Chen, Z.-L. *et al.* Distribution of the COVID-19 epidemic and correlation with population emigration from Wuhan, China. *Chin. Med. J. (Engl.)* **133(9)**, 1044–1050 (2020).
8. Li, X., Zhao, X. & Sun, Y. The lockdown of Hubei Province causing different transmission dynamics of the novel coronavirus (2019-nCoV) in Wuhan and Beijing, *medRxiv* 2020.02.09.20021477v2 (2020)
9. Lau, H. *et al.* The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China, *J. Travel Med.* **27(3)**, taaa037 (2020).
10. Lai, S. *et al.* Assessing spread risk of Wuhan novel coronavirus within and beyond China, January–April 2020: a travel network-based modelling study, *medRxiv* 2020.02.04.20020479 (2020).
11. Boldog, P. *et al.* Risk Assessment of Novel Coronavirus COVID-19 Outbreaks Outside China *J. Clin. Med.* **9(2)**, 571 (2020)

12. Brauer, F., Castillo-Chavez, C. & Feng, Z. *Mathematical models in epidemiology*, Springer, 2019.
13. Virtanen, P. *et al.* *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*. *Nature Methods* 17.3 (2020): 261–272.
14. Newville, M. *et al.* *lmfit/lmfit-py 1.0.1*. *Zenodo*, <https://zenodo.org/record/3814709> (May 2020)
15. Zhu, C., Byrd, R. H., Lu, P. & Nocedal, J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM T. Math. Software* 23(4), 550–560 (1997).
16. Blower, S. M. & Dowlatabadi, H. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *International Statistical Review/Revue Internationale de Statistique*, 229–243 (1994).
17. McKay, M. D., Beckman, R. J. & Conover, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 42(1), 55–61 (2000)
18. Diekmann, O., Heesterbeek, J. A. P. & Roberts, M. G. The construction of next-generation matrices for compartmental epidemic models. *J. R. Soc. Interface* 7, 873–885 (2010).
19. Rawson, T., Brewer, T., Veltcheva, D., Huntingford, C., & Bonsall, M. B. How and when to end the COVID-19 lockdown: an optimization approach. *Frontiers in Public Health* 8, 262 (2020).
20. Davidson, H., Tondo, L. & Yu, V. Coronavirus: quarter of Italy’s population put in quarantine as virus reaches Washington DC. *The Guardian*, 8 March 2020. <https://www.theguardian.com/world/2020/mar/08/coronavirus-italy-quarantine-virus-reaches-washington-dc> (last accessed Aug 13, 2020)
21. Dipartimento della Protezione Civile. *Dati COVID-19 Italia* (last accessed Aug 4, 2020). <https://github.com/pcm-dpc/COVID-19>
22. La Repubblica. *Stretta del governo, Lombardia e 14 province in isolamento. Conte: “Cambiare lo stile di vita”. Speranza: “Pugno duro contro gli irresponsabili”* March 8 2020 (accessed July 29, 2020) [https://www.repubblica.it/cronaca/2020/03/07/news/coronavirus\\_chiusa\\_la\\_lombardia\\_e\\_11\\_province\\_-250570150/?ref=RHPPTP-BH-I250571072-C12-P2-S1.12-T1](https://www.repubblica.it/cronaca/2020/03/07/news/coronavirus_chiusa_la_lombardia_e_11_province_-250570150/?ref=RHPPTP-BH-I250571072-C12-P2-S1.12-T1)
23. an der Heiden, M. & Hamouda, O. *Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting*. *Epi. Bull.* 17 (2020).
24. Barbarossa, M. V. *et al.* Modeling the spread of COVID-19 in Germany: Early assessment and possible scenarios. *PLoS ONE* 15(9): e0238559 (2020).
25. Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M. & Davie, N. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* 5(5), e261–e270 (2020).
26. Wölfel R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* 581(7809), 465–469 (2020).
27. RKI *SARS-CoV-2 Steckbrief zur Coronavirus-Krankheit-2019 (COVID-19)* Robert Koch-Institute, Berlin, Germany 2020. Available online: [https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Steckbrief.html#doc13776792bodyText11](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html#doc13776792bodyText11) (last accessed Aug 11, 2020).
28. Wei, W. E. *et al.* Presymptomatic Transmission of SARS-CoV-2—Singapore, January 23–March 16, 2020. *MMWR*, 69(14), 411 (2020)
29. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med.* 26(5):672–675 (2020)
30. Istat – Ministero della Salute, *Indagine di sieroprevalenza sul SARS-CoV-2, Anno 2020 (dati provvisori)* available online, Aug 3, 2020. <https://www.istat.it/it/files//2020/08/ReportPrimiRisultatiIndagineSiero.pdf>
31. Guzzetta, G. *et al.* The impact of a nation-wide lockdown on COVID-19 transmissibility in Italy. *arXiv* 2004.12338 (2020).
32. Weitz, J. S. *et al.* Modeling shield immunity to reduce COVID-19 epidemic spread. *Nat. Med.* 26, 849–854 (2020).
33. Röst, G. *et al.* Early phase of the COVID-19 outbreak in Hungary and post-lockdown scenarios, *Viruses* 12(7), 708 (2020). doi: 10.3390/v12070708

34. Barbarossa, M. V., Dénes, A., Kiss, G., Nakata, Y., Röst, G. & Vizi, Z. Transmission dynamics and final epidemic size of Ebola virus disease outbreaks with varying interventions. *PLOS ONE* **10(7)**, e0131398 (2015).
35. D'Alessandro J., Coronavirus, l'illusione della grande fuga da Milano. Ecco i veri numeri degli spostamenti verso sud. *La Repubblica*, Apr 23, 2020 (last accessed July 8, 2020).

## Supplementary material

### Fleeing lockdown and its impact on the size of epidemic outbreaks in the source and target regions – a COVID-19 lesson

*MV Barbarossa\**, *N Bogya*, *A Dénes*, *G Röst*, *H Vinod Varma*, *Zs Vizi*;  
 (\*) *barbarossa@fias.uni-frankfurt.de*

#### Computation of the basic reproduction number $\mathcal{R}_0$

For the analytical computation of the basic reproduction number of the system

$$\begin{aligned}
 \dot{S}(t) &= -\lambda(t)S(t) && \text{susceptibles} \\
 \dot{E}(t) &= \lambda(t)S(t) - \alpha E(t) && \text{exposed/presymptomatic} \\
 \dot{U}(t) &= (1-\rho)\alpha E(t) - \gamma_U U(t) && \text{undetected infectives} \\
 \dot{I}(t) &= \rho\alpha E(t) - (\gamma_I + \delta_I + \eta)I(t) && \text{detected, non-hospitalized infectives} \\
 \dot{H}(t) &= \eta I(t) - (\gamma_H + \delta_H)H(t) && \text{hospitalized infectives} \\
 \dot{R}(t) &= \gamma_I I(t) + \gamma_H H(t) && \text{recovered from detected infection} \\
 \dot{R}_U(t) &= (1-\sigma)\gamma_U U(t) && \text{recovered from undetected infection} \\
 \dot{D}(t) &= \delta_I I(t) + \delta_H H(t) + \sigma\gamma_U U(t) && \text{deceased,}
 \end{aligned} \tag{1}$$

where

$$\lambda(t) = \frac{\beta_E E(t) + \beta_I I(t) + \beta_U U(t) + \beta_H H(t)}{N - D(t)},$$

we use the next-generation matrix approach and split the system into a vector of compartments which are going through the infection  $x = (E, I, H, U)$  and a vector of non-infectious compartments  $y = (S, R, R_U, D)$ . Then the system can be rewritten as

$$\begin{aligned}
 x'_i &= \mathcal{F}_i(x, y) - \mathcal{V}_i(x, y), & i = 1, \dots, 4, \\
 y'_j &= g_j(x, y), & j = 1, \dots, 4.
 \end{aligned} \tag{2}$$

In the  $(x, y)$ -notation the initial disease-free equilibrium (DFE) of the system is  $P_0 = (0, y^*)$ , where  $y^* = (S_0, 0, 0, 0)$ . The linearized system at the DFE can be then written by means of the linearization of  $\mathcal{F}$  and  $\mathcal{V}$ , that is the  $4 \times 4$ -matrices

$$F_{ij} = \frac{\partial \mathcal{F}_i}{\partial x_j}(0, y^*), \quad V_{ij} = \frac{\partial \mathcal{V}_i}{\partial x_j}(0, y^*),$$

obtaining  $x' = (F - V)x$ . From the theory of positive matrices, one obtains that the basic reproduction number  $\mathcal{R}_0$  is the spectral radius of the next-generation matrix,  $K := FV^{-1}$ .

For our model we have:

$$F = \begin{pmatrix} \beta_E & \beta_I & \beta_H & \beta_U \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \alpha & 0 & 0 & 0 \\ -\rho\alpha & \gamma_I + \delta_I + \eta & 0 & 0 \\ 0 & -\eta & \gamma_H + \delta_H & 0 \\ -(1-\rho)\alpha & 0 & 0 & \gamma_U \end{pmatrix} \tag{3}$$

Hence

$$FV^{-1} = \begin{pmatrix} \mathcal{R}_0 & * & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where  $\mathcal{R}_0$  is

$$\mathcal{R}_0 = \frac{\beta_E}{\alpha} + \frac{\rho\beta_I}{\gamma_I + \delta_I + \eta} + \frac{\rho\beta_H\eta}{(\gamma_I + \delta_I + \eta)(\gamma_H + \delta_H)} + \frac{\beta_U(1-\rho)}{\gamma_U}.$$

## Computation of the final size formula

We present in this section all steps necessary to obtain the final size formula for the system (1). Let us assume that for  $t$  sufficiently large  $N - D(t) \equiv \mathcal{N}$  is constant over time. We define the functions  $V(t)$  and  $W(t)$  as

$$V(t) = \ln S(t) + a_1 U(t) + a_2 R(t) + a_3 R^U(t) + a_4 D(t)$$

and

$$W(t) = \ln S(t) + b_1 I(t) + b_2 H(t) + b_3 D(t) + b_4 R(t),$$

where

$$a_1 = \frac{\beta_E}{\mathcal{N}(1-\rho)\alpha}, \quad a_2 = \frac{\beta_I - a_4 \delta_I \mathcal{N}}{\gamma_I \mathcal{N}}, \quad a_3 = \frac{\beta_U + \mathcal{N}(+a_1 \gamma_U - a_4 \sigma \gamma_U)}{(1-\sigma)\gamma_U \mathcal{N}}, \quad a_4 = \frac{\beta_I \gamma_H - \beta_H \gamma_I}{\mathcal{N}(\gamma_H \delta_I - \gamma_I \delta_H)} \quad (4)$$

and

$$b_1 = \frac{\beta_E}{\mathcal{N}\rho\alpha}, \quad b_2 = \frac{\beta_U \delta_H + \sigma \gamma_U (-\beta_H + b_4 \gamma_H \mathcal{N})}{(\gamma_H + \delta_H) \mathcal{N} \sigma \gamma_U}, \quad (5)$$

$$b_3 = \frac{\beta_U}{\mathcal{N} \sigma \gamma_U}, \quad b_4 = \frac{\left(\frac{\beta_I}{\mathcal{N}} + b_1(\gamma_I + \delta_I + \eta) - b_3 \delta_I\right)(\gamma_H + \delta_H) - \eta \left(-\frac{\beta_H}{\mathcal{N}} + b_3 \delta_H\right)}{\eta \gamma_H + \gamma_I(\gamma_H + \delta_H)}.$$

The functions  $V$  and  $W$  are invariants for the system (1). Indeed, differentiation along the solutions yields

$$\begin{aligned} \dot{V}(t) &= -\frac{\beta_E E(t)}{\mathcal{N}} - \frac{\beta_I I(t)}{\mathcal{N}} - \frac{\beta_U U(t)}{\mathcal{N}} - \frac{\beta_H H(t)}{\mathcal{N}} + a_1(1-\rho)\alpha E(t) - a_1 \gamma_U U(t) \\ &\quad + a_2 \gamma_I I(t) + a_2 \gamma_H H(t) \\ &\quad + a_3(1-\sigma)\gamma_U U(t) + a_4 \delta_I I(t) + a_4 \delta_H H(t) + a_4 \sigma \gamma_U U(t) \end{aligned} \quad (6)$$

$$\begin{aligned} &= E \left( -\frac{\beta_E}{\mathcal{N}} + a_1(1-\rho)\alpha \right) + I \left( -\frac{\beta_I}{\mathcal{N}} + a_2 \gamma_I + a_4 \delta_I \right) \\ &\quad + U \left( -\frac{\beta_U}{\mathcal{N}} - a_1 \gamma_U + a_3(1-\sigma)\gamma_U + a_4 \sigma \gamma_U \right) + H \left( -\frac{\beta_H}{\mathcal{N}} + a_2 \gamma_H + a_4 \delta_H \right) = 0, \end{aligned}$$

and

$$\begin{aligned} \dot{W}(t) &= -\frac{\beta_E E(t)}{\mathcal{N}} - \frac{\beta_I I(t)}{\mathcal{N}} - \frac{\beta_U U(t)}{\mathcal{N}} - \frac{\beta_H H(t)}{\mathcal{N}} + b_1 \rho \alpha E(t) - b_1(\gamma_I + \delta_I + \eta) I(t) \\ &\quad + b_2 \eta I(t) - b_2(\gamma_H + \delta_H) H(t) + b_3 \delta_I I(t) + b_3 \delta_H H(t) + b_3 \sigma \gamma_U U(t) + b_4 \gamma_I I(t) + b_4 \gamma_H H(t), \end{aligned} \quad (7)$$

$$\begin{aligned} &= E \left( -\frac{\beta_E}{\mathcal{N}} + b_1 \rho \alpha \right) + I \left( -\frac{\beta_I}{\mathcal{N}} - b_1(\gamma_I + \delta_I + \eta) + b_2 \eta + b_4 \gamma_I + b_3 \delta_I \right) \\ &\quad + U \left( -\frac{\beta_U}{\mathcal{N}} + b_3 \sigma \gamma_U \right) + H \left( -\frac{\beta_H}{\mathcal{N}} - b_2(\gamma_H + \delta_H) + b_3 \delta_H + b_4 \gamma_H \right) = 0. \end{aligned}$$

In particular, we have the relations  $V(T+) = V(\infty)$  and  $W(T+) = W(\infty)$ , that is

$$\ln S_{T+} + \tilde{a}_1 U_{T+} + \tilde{a}_2 R_{T+} + \tilde{a}_3 R_{T+}^U + \tilde{a}_4 D_{T+} = \ln S_\infty + \tilde{a}_2 R_\infty + \tilde{a}_3 R_\infty^U + \tilde{a}_4 D_\infty \quad (8)$$

$$\ln S_{T+} + \tilde{b}_1 I_{T+} + \tilde{b}_2 H_{T+} + \tilde{b}_3 D_{T+} + \tilde{b}_4 R_{T+} = \ln S_\infty + \tilde{b}_3 D_\infty + \tilde{b}_4 R_\infty, \quad (9)$$

where  $\tilde{a}_j, \tilde{b}_j, j = 1, 2, 3$  are the coefficients in (5), with values of parameters after intervention at time  $T$ . Further we have the relation

$$R_\infty^U = \underbrace{\frac{(1-\rho)(1-\sigma)}{\rho + \sigma - \sigma\rho}}_{=: \mu} (R_\infty + D_\infty) = \mu(R_\infty + D_\infty). \quad (10)$$

Next we subtract (8) from (9)

$$\begin{aligned} &\tilde{b}_1 I_{T+} + \tilde{b}_2 H_{T+} + (\tilde{b}_3 - \tilde{a}_4) D_{T+} + (\tilde{b}_4 - \tilde{a}_2) R_{T+} - \tilde{a}_1 U_{T+} - \tilde{a}_3 R_{T+}^U \\ &= [(\tilde{b}_3 - \tilde{a}_4) D_\infty + (\tilde{b}_4 - \tilde{a}_2) R_\infty - \tilde{a}_3 R_\infty^U] \\ &\stackrel{(10)}{=} [(\tilde{b}_3 - \tilde{a}_4 - \tilde{a}_3 \mu) D_\infty + (\tilde{b}_4 - \tilde{a}_2 - \tilde{a}_3 \mu) R_\infty], \end{aligned} \quad (11)$$

and define

$$A_T := \frac{\tilde{b}_1 I_{T_+} + \tilde{b}_2 H_{T_+} + (\tilde{b}_3 - \tilde{a}_4) D_{T_+} + (\tilde{b}_4 - \tilde{a}_2) R_{T_+} - \tilde{a}_1 U_{T_+} - \tilde{a}_3 R_{T_+}^U}{\tilde{b}_3 - \tilde{a}_4 - \tilde{a}_3 \mu}, \quad b := \frac{\tilde{b}_4 - \tilde{a}_2 - \tilde{a}_3 \mu}{\tilde{b}_3 - \tilde{a}_4 - \tilde{a}_3 \mu}. \quad (12)$$

We observe that

$$P_T^+ := P(T_+) := \mathcal{N} + D(T_+) \underbrace{=}_{(10)} S_\infty + (1 + \mu) R_\infty + (1 + \mu) D_\infty, \quad (13)$$

hence

$$D_\infty = \frac{P_T^+ - S_\infty}{1 + \mu} - R_\infty. \quad (14)$$

Solving (11) for  $D_\infty$  and using (13) we obtain

$$D_\infty = A_T - b R_\infty. \quad (15)$$

From the relations (14) and (15), and solving for  $R_\infty$  we get

$$R_\infty = \frac{P_T^+ - S_\infty}{(1 + \mu)(1 - b)} - \frac{A_T}{1 - b}. \quad (16)$$

We substitute this expression in the relation (15) and obtain

$$D_\infty = -\frac{b(P_T^+ - S_\infty)}{(1 + \mu)(1 - b)} + \frac{A_T}{1 - b}. \quad (17)$$

Now use the relations (17) and (16) in (9):

$$\begin{aligned} \ln S_{T_+} + \tilde{b}_1 I_{T_+} + \tilde{b}_2 H_{T_+} + \tilde{b}_3 D_{T_+} + \tilde{b}_4 R_{T_+} &= \ln S_\infty + \tilde{b}_3 \left( -\frac{b(P_T^+ - S_\infty)}{(1 + \mu)(1 - b)} + \frac{A_T}{1 - b} \right) + \tilde{b}_4 \left( \frac{P_T^+ - S_\infty}{(1 + \mu)(1 - b)} - \frac{A_T}{1 - b} \right) \\ &= \ln S_\infty + \frac{\tilde{b}_4 - b\tilde{b}_3}{(1 + \mu)(1 - b)} (P_T^+ - S_\infty) + \frac{(\tilde{b}_3 - \tilde{b}_4) A_T}{1 - b}. \end{aligned}$$

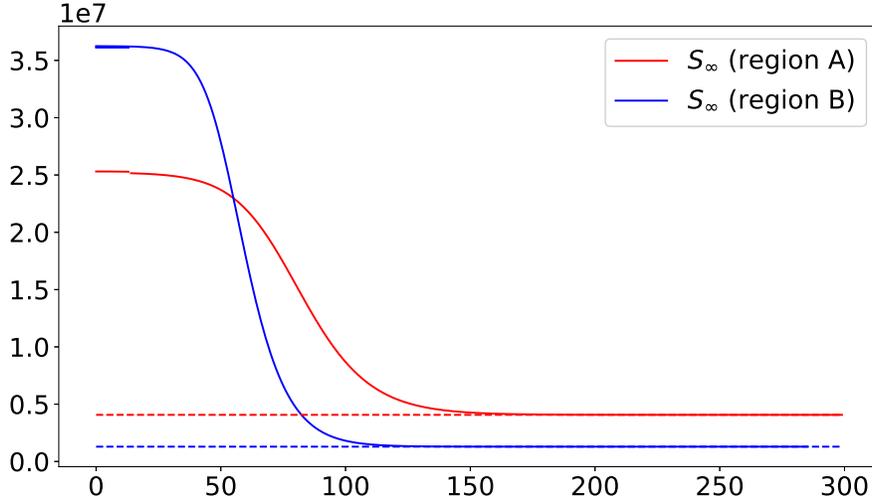
Define  $\chi = \frac{\tilde{b}_3 - \tilde{b}_4}{(1 - b)(\tilde{b}_3 - \tilde{a}_4 - \tilde{a}_3 \mu)}$  and  $\Phi = \frac{\tilde{b}_4 - b\tilde{b}_3}{(1 + \mu)(1 - b)}$ . Finally we substitute back  $A_T$  from (12), use the relation  $P_{T_+} = \mathcal{N} + D_{T_+}$ . We obtain

$$\begin{aligned} \ln S_{T_+} + \tilde{b}_1 I_{T_+} + \tilde{b}_2 H_{T_+} + \tilde{b}_3 D_{T_+} + \tilde{b}_4 R_{T_+} &= \ln S_\infty + \Phi(\mathcal{N} + D_{T_+} - S_\infty) + \chi \left( \tilde{b}_1 I_{T_+} + \tilde{b}_2 H_{T_+} + (\tilde{b}_3 - \tilde{a}_4) D_{T_+} + (\tilde{b}_4 - \tilde{a}_2) R_{T_+} - \tilde{a}_1 U_{T_+} - \tilde{a}_3 R_{T_+}^U \right) \\ &= \ln S_\infty + \Phi(\mathcal{N} - S_\infty) + \tilde{b}_1 \chi I_{T_+} + \chi \tilde{b}_2 H_{T_+} \\ &\quad + (\chi(\tilde{b}_3 - \tilde{a}_4) + \Phi) D_{T_+} + \chi(\tilde{b}_4 - \tilde{a}_2) R_{T_+} - \chi \tilde{a}_1 U_{T_+} - \chi \tilde{a}_3 R_{T_+}^U. \end{aligned} \quad (18)$$

Rearranging the terms we obtain the final size relation as in the main manuscript,

$$\begin{aligned} \ln S_{T_+} - \ln S_\infty &= \Phi(\mathcal{N} - S_\infty) + \tilde{b}_1(\chi - 1) I_{T_+} + (\chi - 1) \tilde{b}_2 H_{T_+} + (\chi(\tilde{b}_3 - \tilde{a}_4) + \Phi - \tilde{b}_3) D_{T_+} \\ &\quad + (\chi(\tilde{b}_4 - \tilde{a}_2) - \tilde{b}_4) R_{T_+} - \chi \tilde{a}_1 U_{T_+} - \chi \tilde{a}_3 R_{T_+}^U. \end{aligned}$$

On the example of the model parametrized on the Italian data, and assuming that parameters are maintained in both region A and region B as of the first intervention on March 9, 2020, we verify and visualize in Figure S1 the convergence of the analytical final size formula with the numerical approximation.



**Figure S1.** Final size approximation (straight lines) and numerical solution of the susceptible curve of the model (1). Parametrization as for the Italian data (Table 2) assuming parameters remain constant after first intervention measures (March 8) both in region A and region B.

### Parameter settings in simulations

For numerical investigation of sensitivity of cases and deaths counts with respect to lockdown-time, fleeing fraction and intervention measures (Figures 2–5) the setting in Table 1 was used. Unless otherwise mentioned, initial values for region A are:  $E(0) = 100$ ,  $I(0) = 8$ ,  $H(0) = 10$ ,  $U(0) = 100$ ,  $R(0) = R_U(0) = 0$ ,  $D(0) = 2$ , and  $S(0) = Pop_A - I(0) - D(0) - U(0) - H(0) - E(0) - R(0) - R_U(0)$ . In region B the initial population is assumed to be fully susceptible. Further

- In Figure 2(a)  $Pop_A = 25e6$ ,  $Pop_B = 50e6$ ,  $\phi = 0.01$  and the lockdown time  $T$  varies between one and twelve weeks after the beginning of the outbreak.
- In Figure 2(b) the same setting as for Figure 2(a) is used, but the lockdown time  $T$  is fixed at 3 weeks after the beginning of the observations, whereas the migrating fraction  $\phi$  varies between 0.01% and 50%.
- In Figure 3(a) and Figure 3(b),  $Pop_A = 25e6$  and  $Pop_B = [0.5, 2, 4] \times Pop_A$ . The fleeing fraction of exposed/undetected cases  $\phi$  varies between 0.01% and 50%. Numbers are projected at  $t = T + 60$  days and the control measures (reducing the transmission rates by respectively 5%, 30% and 80%) and maintained for two months. Figure 4 is produced with the same setting but running simulations for up to 500 days.
- In Figure 5  $Pop_A = 25e6$  and  $Pop_B = [0.5, 2, 4] \times Pop_A$ . Region A is isolated on day  $T = 21$  since beginning of the outbreak. Intervention in region B occurs  $T_B$  days (varying from 3 to 30 days) after  $T$ . Cumulative detected cases are projected at  $t = T + 90$  days and the control measures in B (reducing the transmission rates by respectively 5%, 30% and 50%) are maintained until the end of the simulations.

For the numerical simulations of the Italian dataset parameter values as in Table 2 were used. Initial values for integration of the ODE system are: for region A,  $E^A(0) = 9,517$ ,  $I^A(0) = 94$ ,  $H^A(0) = 125$ ,  $U^A(0) = 554$ ,  $R^A(0) = 0$ ,  $R_U^A(0) = 0$  and  $D^A(0) = 7$ ; for region B,  $E^B(0) = 486$ ,  $I^B(0) = 0$ ,  $H^B(0) = 2$ ,  $U^B(0) = 5$ ,  $R^B(0) = 1$ ,  $R_U^B(0) = 0$  and  $D^B(0) = 0$ . Initial susceptible population for both regions were then computed as the difference between the population size and the sum of all other compartments in the region, with  $Pop_A(0) = 25,307,582$  and  $Pop_B(0) = 36,124,240$ . For Partial Rank Correlation Coefficients (PRCC) analysis (Figure 8) parameters were varied in the ranges reported in Table 3, and otherwise fixed as in Table 2.

**Table 1.** Parameter description and values used for numerical simulations of model (1)

Parameters	Value (unit)	Description
$\beta_U$	0.5 (1/(days× contact))	Transmission rate from undetected infectives to susceptibles
$\beta_I$	0.05(1/(days× contact))	Transmission rate from non-hospitalized cases to susceptibles
$\beta_E$	0.01 (1/(days× contact))	Transmission rate from presymptomatics to susceptibles
$\beta_H$	0.015 (1/(days× contact))	Transmission rate from hospitalized cases to susceptibles
$1/\alpha$	5.5 (days) <sup>1</sup>	Duration of latency period
$\gamma_U$	0.125 (1/days) <sup>3</sup>	Recovery rate for undetected infectives
$\gamma_I$	0.1 (1/days) <sup>4</sup>	Recovery rate for non-hospitalized cases
$\gamma_H$	0.1 (1/days) <sup>4</sup>	Recovery rate for hospitalized cases
$\rho$	0.2	Fraction of detected cases (after latent period)
$\sigma$	0.001	Probability of detection post mortem
$\eta$	0.333 (1/days)	Hospitalization rate
$\delta_I$	0.0004 (1/days)	Disease-induced death rate for non-hospitalized cases
$\delta_H$	0.0167 (1/days)	Disease-induced death rate for hospitalized cases

## References

1. an der Heiden, M. & Hamouda, O. Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting. *Epi. Bull.* **17** (2020).
2. Barbarossa, M. V. *et al.* Modeling the spread of COVID-19 in Germany: Early assessment and possible scenarios. *PLoS ONE* 15(9): e0238559 (2020).
3. Prem, K., Liu, Y., Russell, T. W., Kucharski, A. J., Eggo, R. M. & Davie, N. *The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study.* *Lancet Public Health* **5(5)**, e261–e270 (2020).
4. RKI *SARS-CoV-2 Steckbrief zur Coronavirus-Krankheit-2019 (COVID-19)* Robert Koch-Institute, Berlin, Germany 2020. Available online: [https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Steckbrief.html#doc13776792bodyText11](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html#doc13776792bodyText11) (last accessed Aug 11, 2020).
5. Istat- Ministero della Salute, *Indagine di sieroprevalenza sul SARS-CoV-2, Anno 2020 (dati provvisori)* available online, Aug 3, 2020. <https://www.istat.it/it/files//2020/08/ReportPrimiRisultatiIndagineSiero.pdf>

**Table 2.** Parameter values used for fitting model (1) on Italian data. When multiple rows are given, the first indicates values for region A and the second for region B.

	<b>24.02 – 08.03</b>	<b>09.03 – 21.03</b>	<b>22.03 – 04.05</b>	<b>Comments</b>
$\beta_U$	$0.5091 \pm 0.0048$	$0.2969 \pm 0.0056$	$0.1037 \pm 0.0006$	estimated
	$0.7016 \pm 0.0564$	$0.4834 \pm 0.0208$	$0.1021 \pm 0.0018$	
$\beta_I$	0.1273	0.0742	0.0259	$= 0.25\beta_U$
	0.1754	0.1209	0.0255	
$\beta_E$	0.1018	0.0594	0.0207	$= 0.20\beta_U$
	0.1403	0.0967	0.0204	
$\beta_H$	0.0509	0.0297	0.0104	$= 0.10\beta_U$
	0.0702	0.0483	0.0102	
$\alpha$	1 / 5.5	1 / 5.5	1 / 5.5	assumed <sup>1,2</sup>
$\gamma_U$	1 / 7	1/7	1/7	assumed <sup>2,3</sup>
$p_H$	75%	$72.00 \pm 2.57\%$	$49.42 \pm 0.88\%$	estimated
	50%	$69.93 \pm 8.81\%$	$33.74 \pm 1.42\%$	
$p_{HR}$	85%	$68.74 \pm 1.46\%$	$78.27 \pm 0.17\%$	estimated
	90%	$81.77 \pm 8.60\%$	$83.89 \pm 0.48\%$	
$p_{IR}$	99%	99%	99%	assumed
$\gamma_I$	0.0248	0.0277	0.0501	$= (1 - p_H)p_{IR}/10$
	0.0495	0.0298	0.0656	
$\gamma_H$	0.0850	0.0687	0.0783	$= p_H p_{HR}/10$
	0.0900	0.0818	0.0839	
$\rho$	14.5%	14.5%	14.5%	estimated <sup>5</sup>
	17.5%	17.5%	17.5%	
$\sigma$	0.001	0.001	0.001	assumed
$\eta$	0.3750	0.3600	0.2471	$= p_H/2$
	0.2500	0.3496	0.1687	
$\delta_I$	0.0003	0.0003	0.0005	$= (1 - p_H)(1 - p_{IR})/10$
	0.0005	0.0003	0.0007	
$\delta_H$	0.0150	0.0313	0.0217	$= p_H(1 - p_{HR})/10$
	0.0100	0.0182	0.0161	

**Table 3.** Parameter ranges used for sensitivity analysis (PRCC)

<b>Parameter</b>	<b>Range</b>	<b>Parameter</b>	<b>Range</b>
$\beta_U$	[0.01, 1]	$\rho$	[0.1, 0.9]
$\eta$	[0.1, 0.5]	$\gamma_I$	[0.009, 0.09]
$\gamma_H$	[0.01, 0.1]	$\delta_I$	[0.0001, 0.001]
$\delta_H$	[0.005, 0.05]		

# Figures

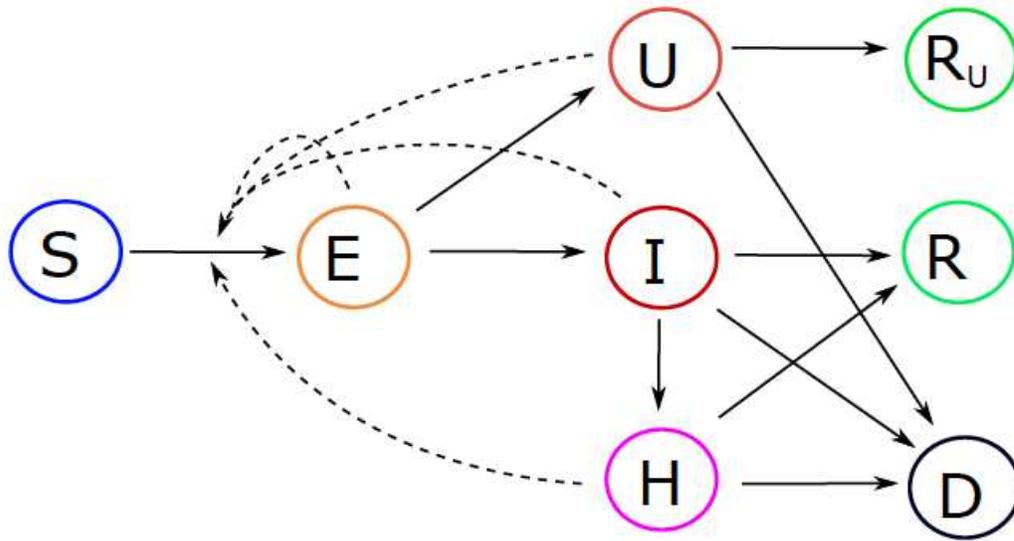


Figure 1

Model structure for the transmission dynamics of an infectious disease, on the example of COVID-19. Solid arrows indicate transition from one compartment to another, dashed arrows indicate virus transmission due to contact with infectives. Upon infection, susceptible (S) individuals enter a latent/presymptomatic phase (E). After symptoms onset, infections may be detected (I) or remain undetected (U). Detected cases might become severe and require hospitalization (H). Infected individuals who recovered from a detected (R) or an undetected (RU) infection, as well as patients who died (D) upon infections, are removed from the chain of transmission.

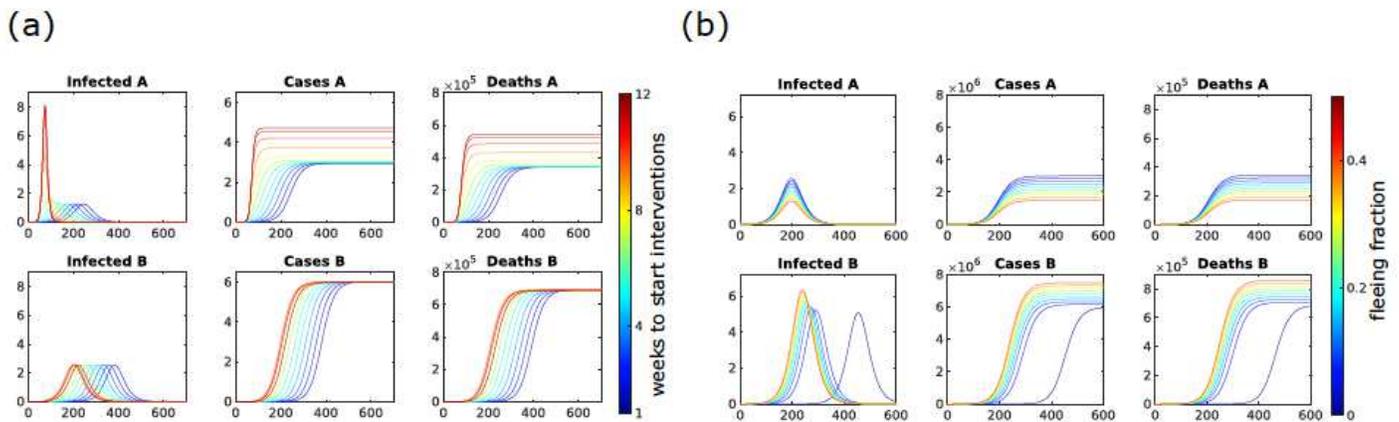
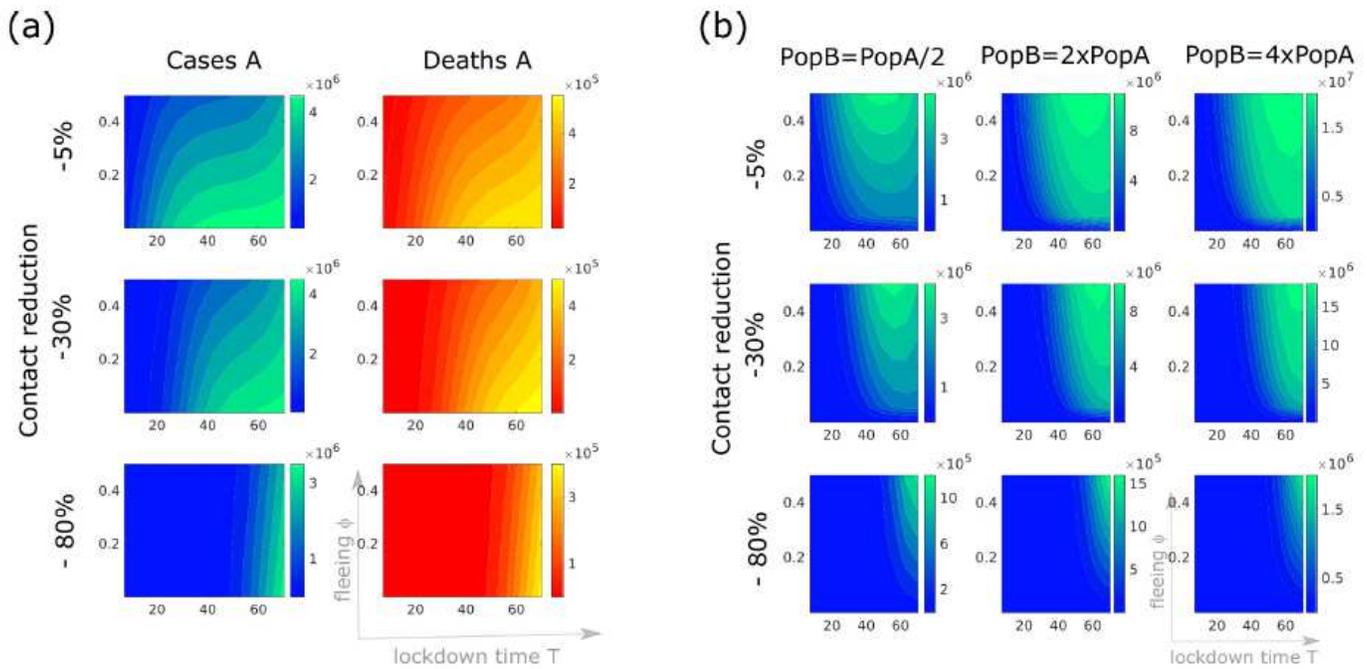


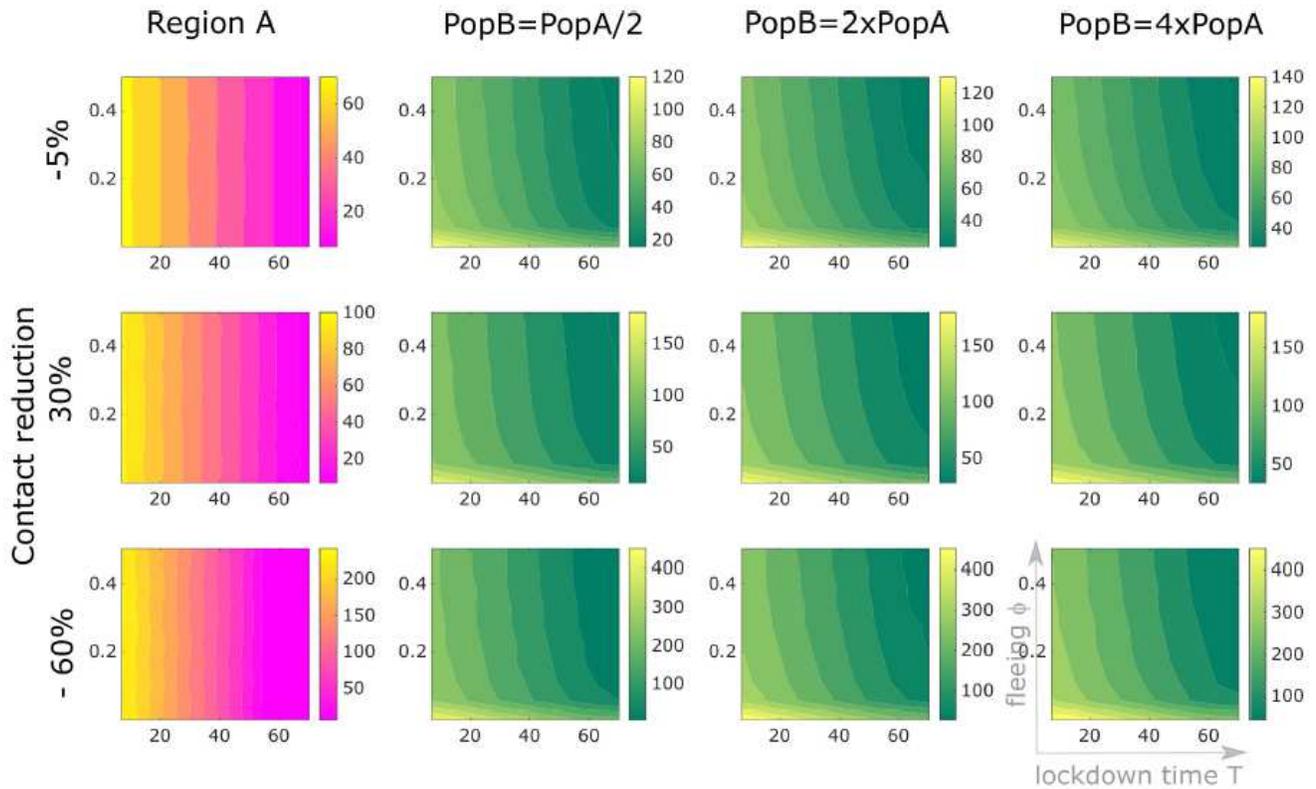
Figure 2

Detected infected cases and deaths depending on (a) the lockdown time  $T$  and (b) the fleeing fraction of exposed/undetected individuals at the time of the lockdown. An initial outbreak starts off with 20 cases and two deaths in region A (population 25 million), where lockdown is established at time  $T$ , immediately followed by lockdown-induced migration of a fraction of the unconstrained population from region A to the disease-free region B (population 50 million). (a) The lockdown time  $T$  varies from one to twelve weeks after the initial reporting, while the fleeing fraction is fixed to 1% of the unconstrained population in region A; (b) 0.01% to 50% of the unconstrained population moves from region A to region B at the fixed time  $T = 21$  days after the initial reporting. At the time of lockdown, control measures are applied in both regions in order to restrict contacts by 60% and reduce transmission.



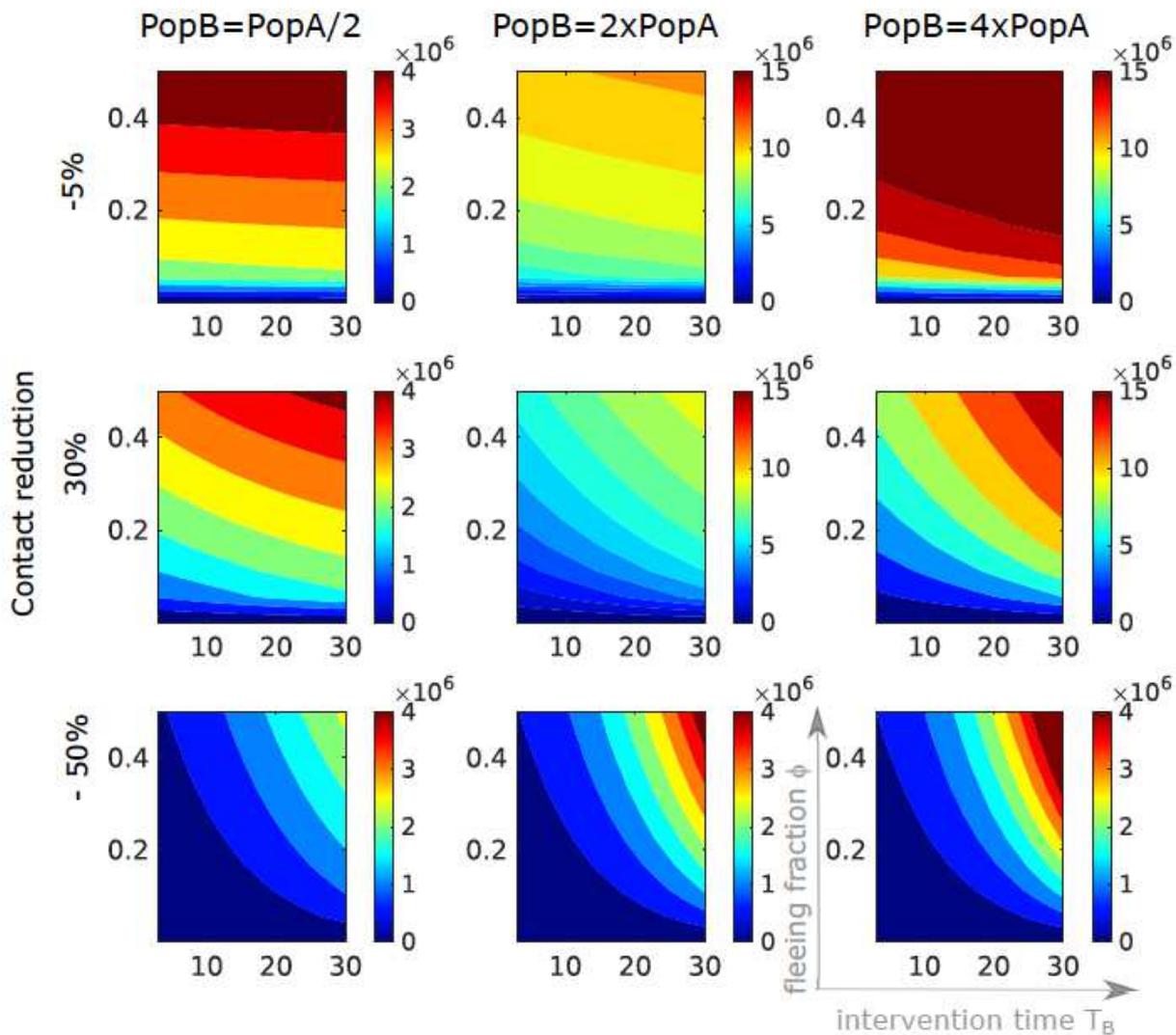
**Figure 3**

Sensitivity of cases in region B depending on the lockdown time  $T$ , the fraction of population leaving region A, the strength of intervention measures in reducing contacts, and the population size of region B. An initial outbreak starts off with 20 detected cases and two deaths in region A (population 25 million) which is locked at time  $T$  (horizontal axes indicate days after the initial reporting) and lockdown-induced migration (vertical axes) occurs. At the time of lockdown, control measures are applied to restrict contacts by 5% (first row), 30% (second row) or 80% (third row), and maintained for two months. Color codes represent cases/deaths projected at the end of the two months following model (1): (a) the number of detected cases (left column) and deaths (right column) in region A, (b) detected case numbers in region B depending on the size of the population in region B with respect to that of region A (columns).



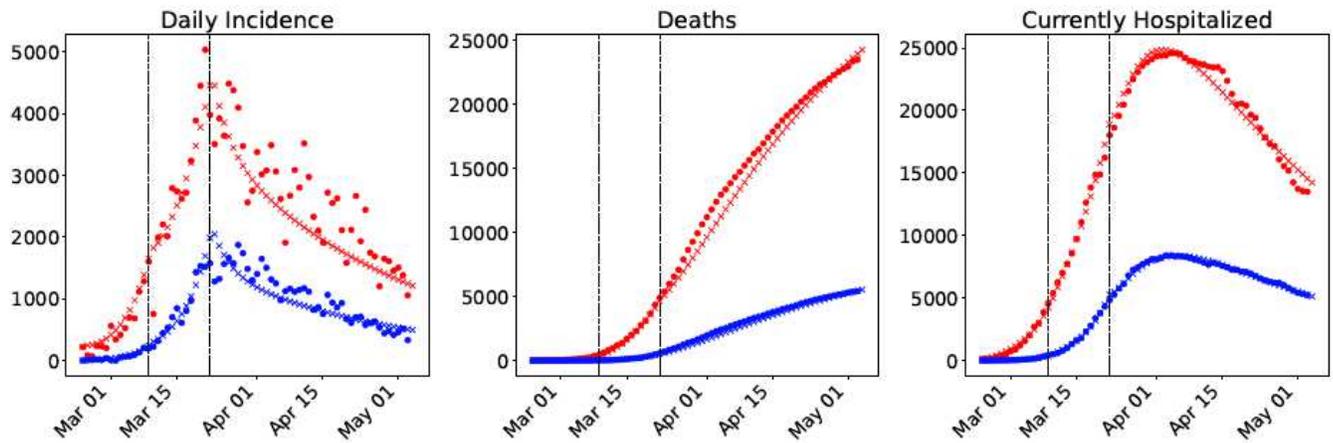
**Figure 4**

How long should control measures be in place? An initial outbreak starts off with 20 cases and two deaths in region A (population 25 million) which is then locked at time  $T$  (horizontal axes indicate days after the initial reporting) and lockdown-induced migration (vertical axes) occurs. In both regions, at the time of lockdown, control measures are applied to restrict contacts by 5% (first row), 30% (second row) or 60% (third row), and maintained for 7 days after the peak in the daily incidence in the respective region is reached. Color codes represent the duration in days of the control period in region A (first column) and region B (second to fourth columns), also depending on the population size in region B (being half, twice or four times the population in region A).



**Figure 5**

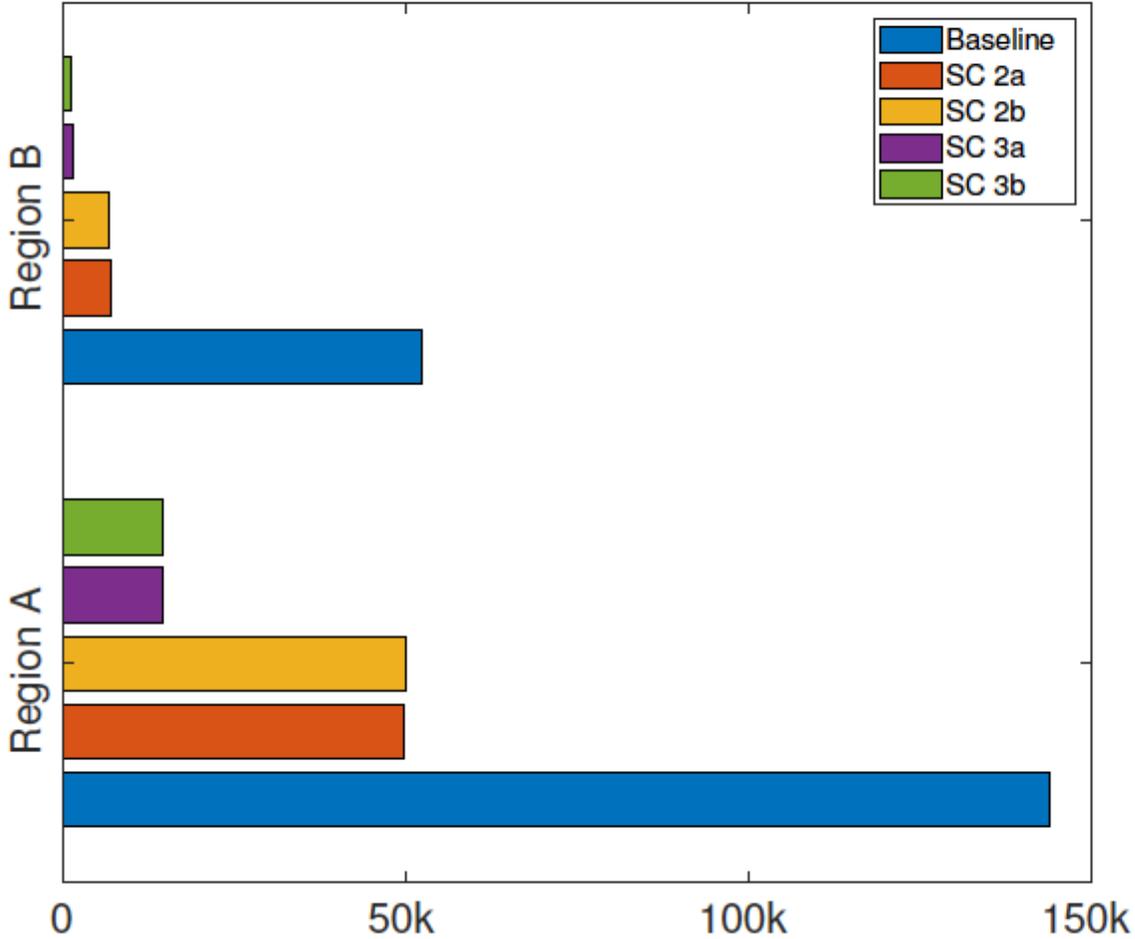
Cumulative detected cases in region B three months after lockdown of region A depending on the fraction  $\phi$  of population leaving region A, the intervention time  $T_B$  in region B, the effectiveness of intervention measures in reducing contacts, and the population size of region B. Region A is isolated 21 days after the beginning of the outbreak and migration towards region B occurs. For each panel, the vertical axis denotes the variation in the fleeing fraction from A ( $\phi$ , between 0.1% and 0.5). On the horizontal axis, the reaction time  $T_B$  indicates how many days passed for control measures in region B to be applied since isolation of region A. Control measures are applied in region B to restrict contacts by 5% (first row), 30% (second row) or 50% (third row), and maintained until the end of the simulations. Cumulative detected cases are projected also depending on the size of population in region B with respect to that in region A (columns). Note the different scaling of the color legend in the panels.



**Figure 6**

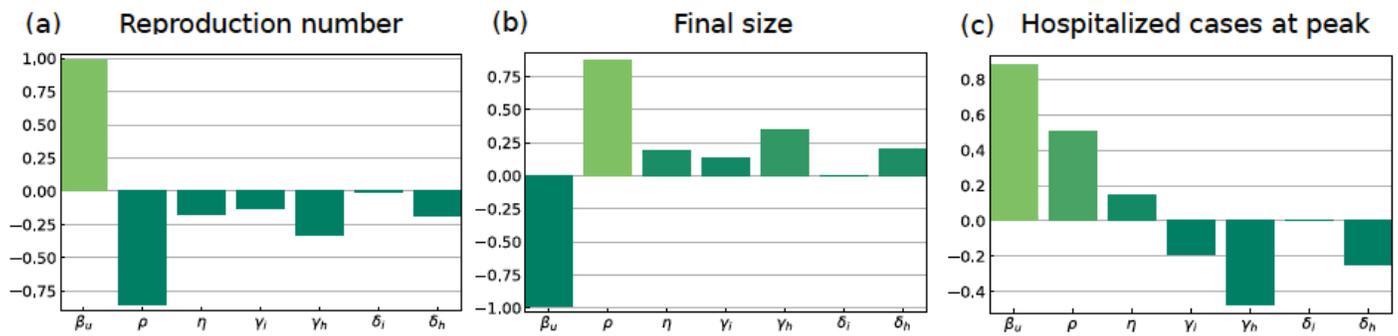
Model fit for the early dynamics of the COVID-19 outbreak in Italy. Reported data (dots) and model results (crosses) for daily incidence (left panel), deaths (middle panel) and hospitalized cases (right panel) in region A (Lombardy, Emilia-Romagna, Marche, Piedmont, Veneto; red) and region B (rest of the country; blue). Parameter values are estimated as indicated in three different time intervals (separated by vertical lines in the figure): pre-lockdown (February 24 – March 8), first lockdown measures (March 9 – 21), and extended lockdown measures (March 22 – May 4), cf. Supplementary Material.

### Cases as of May 4th



**Figure 7**

Scenario comparison for the cumulative cases of COVID-19 in Italy as of May 4, 2020. Cumulative detected cases in region A and region B are simulated for different scenarios: (baseline) the setting as of fit in Figure 6; For all other scenarios severe restrictions (parameter set as of March 22) are in place from the lockdown time: (SC2) lockdown on March 9 (a) with and (b) without fleeing population at lockdown; (SC3) lockdown anticipated to March 2 (a) with and (b) without fleeing population at lockdown.



**Figure 8**

Partial rank correlation coefficients (PRCCs) of seven main model parameters for (a) reproduction number, (b) final size ( $S_1$ ) and (c) the number of hospitalized cases at peak in region A. Parameters with PRCC larger than zero are positively correlated with the quantity of interest, whereas parameters with negative PRCC are negatively correlated. Parameters were varied within the ranges given in the Supplementary Material.