Research article

What can we learn from COVID-19 data by using epidemic models with unidentified infectious cases?

Quentin Griette[1], Jacques Demongeot[2] and Pierre Magal[1,*]

[1] Université de Bordeaux, IMB, UMR 5251, Talence F-33400, France
   CNRS, IMB, UMR 5251, Talence F-33400, France

[2] Université Grenoble Alpes, AGEIS EA7407, La Tronche F-38700, France

* **Correspondence:** Email: pierre.magal@u-bordeaux.fr.

Abstract: The COVID-19 outbreak, which started in late December 2019 and rapidly spread around the world, has been accompanied by an unprecedented release of data on reported cases. Our objective is to offer a fresh look at these data by coupling a phenomenological description to the epidemiological dynamics. We use a phenomenological model to describe and regularize the reported cases data. This phenomenological model is combined with an epidemic model having a time-dependent transmission rate. The time-dependent rate of transmission involves changes in social interactions between people as well as changes in host-pathogen interactions. Our method is applied to cumulative data of reported cases for eight different geographic areas. In the eight geographic areas considered, successive epidemic waves are matched with a phenomenological model and are connected to each other. We find a single epidemic model that coincides with the best fit to the data of the phenomenological model. By reconstructing the transmission rate from the data, we can understand the contributions of the changes in social interactions (contacts between individuals) on the one hand and the contributions of the epidemiological dynamics on the other hand. Our study provides a new method to compute the instantaneous reproduction number that turns out to stay below 3.5 from the early beginning of the epidemic. We deduce from the comparison of several instantaneous reproduction numbers that the social effects are the most important factor in understanding the epidemic wave dynamics for COVID-19. The instantaneous reproduction number stays below 3.5, which implies that it is sufficient to vaccinate 71% of the population in each state or country considered in our study. Therefore, assuming the vaccines will remain efficient against the new variants and adjusting for higher confidence, it is sufficient to vaccinate 75 − 80% to eliminate COVID-19 in each state or country.

Keywords: SARS-CoV-2; social changes; transmission rate; identification of parameters

1. Introduction

Since the first cases occurred in early December 2019, the COVID-19 crisis has been accompanied by unprecedented data release. The first cluster of cases was reported on December 31, 2019, by WHO (World Health Organization) [1]. Chinese authorities confirmed on January 7 that this cluster was caused by a novel coronavirus [2]. The disease then rapidly spread throughout the world; a case was identified in the U.S.A. as early as January 19, 2020, for instance [3]. According to the WHO database [4], the first cases in Japan date back to January 14, in Italy to January 29 (even though the cluster of cases was announced on January 21, 2020 [5]), in France to January 24, 2020, etc. The spread of the epidemic across countries was monitored, and the data was made publicly available at the international level by recognized scientific institutions such as WHO [4] and the Johns Hopkins University [6], who collected data provided by national health agencies. To the best of our knowledge, this is the first time in history that such detailed epidemiological data have been made publicly available on a global scale; this opens up new questions and new challenges for the scientific community.

Modeling efforts in order to analyze and predict the dynamics of the epidemics were initiated from the start [7–9]. Forecasting the propagation of the epidemic is, in particular, a key challenge in infectious disease epidemiology. It has quickly become clear to medical doctors and epidemiologists that covert cases (asymptomatic or unreported infectious cases) play an important role in the spread of the COVID-19. An early description of an asymptomatic transmission in Germany was reported by Rothe et al. [10]. It was also observed on the Diamond Princess cruise ship in Yokohama in Japan [11] that many of the passengers were tested positive for the virus but never presented any symptoms. On the French aircraft carrier Charles de Gaulle, clinical and biological data for all 1739 crew members were collected on arrival at the Toulon harbor and during quarantine: 1121 crew members (64%) were tested positive for COVID-19 using RT-PCR, and among these, 24% were asymptomatic [12]. The importance of covert cases in the silent propagation of the epidemic was highlighted by Qiu [13]. Models accounting for asymptomatic transmission, which agree with reported cases data, have been used from the start of the epidemic [7, 14–16]. The implementation of such models depends, however, on the *a priori* knowledge of some characteristic parameters of the host-pathogen interaction, among which is the ascertainment rate. Nishiura and collaborators [17] estimated this ascertainment rate as 9.2% on a 7.5-days detection window, based on testing data of repatriated Japanese nationals from Wuhan. This was corrected later to 44% for non-severe cases [18]. An early review of SARS-CoV-2 facts can be found in the work of Bar-On et al. [19].

To describe the spread of COVID-19 mathematically, Liu et al. [7] first took into account the infection of susceptible individuals by contacts with unreported infectious individuals. A new method using the number of reported cases in SIR models was also proposed in the same work. This method and the model were extended in several directions by the same group [14–16, 20] to include non-constant transmission rates and a period of exposure. More recently, the method was extended and successfully applied to a Japanese age-structured dataset by Griette, Magal, and Seydi [21]. The method was also extended to investigate the predictability of the outbreak in several countries, including China, South Korea, Italy, France, Germany, and the United Kingdom by Liu, Magal, and Webb [20].

Phenomenological models were extensively used in the literature even before the SARS-CoV-2 pandemic to describe reported cases data, see e.g., [22, 23] for the 2003 SARS outbreak, and also

[24, 25], to cite a few. In the case of the SARS-CoV-2 epidemic, articles related to phenomenological models are particularly numerous, see e.g., [26–35]. More precisely, Castro et al. [27] investigate the possibility of predicting the turning point of an epidemic wave. Many studies use phenomenological models to issue short-term predictions on the epidemic [30, 32–34]. But these models can also be used to reconstruct the evolution of the epidemic a posteriori [26, 28, 35].

In previous works [7, 14–16], we have replaced the data with the phenomenological model, and we use this continuous description as the output of the epidemic model. This allows us to understand how to express part of the initial distribution and some parameters (e.g., the transmission rate) from the data and the given parameters of the model. By using this approach, we obtain an explicit formula for the time-dependent transmission rate expressed by using some given parameters of the model and some parameters of the phenomenological model. In [36], we used a Bernoulli-Verhulst phenomenological model to describe a single epidemic wave and compute a time-dependent transmission rate.

There are many potential phenomenological models to represent a single epidemic wave [24, 26, 28]. However, except in the case of the logistic equation, there is usually no explicit formula for the solution. The explicit formula in the case considered here permits to develop a comprehensive statistical analysis. Phenomenological models also serve to regularize the data, which is a complex question. Indeed the idea is to get rid of the stochastic oscillations (due for example to the way the data are collected, or the stochastic nature of the contact between individuals). Some phenomenological models also redistribute the reported cases to dampen the fluctuations in the data. Let us stress here the fact that some oscillations in cases data may not be random and might correspond to complex transmission dynamics (delayed infection, peak in contact numbers during the day, etc.) This highlights one of the drawbacks of phenomenological models: while they allow a precise description of epidemic waves, they might also hide some valuable information on how the disease is transmitted in the population.

A key parameter in understanding the dynamics of the COVID-19 epidemic is the transmission rate, defined as the fraction of all possible contacts between susceptible and infected individuals that effectively result in a new infection per unit of time. Estimating the average transmission rate is one of the most crucial challenges in the epidemiology of communicable diseases. In practice, many factors can influence the actual transmission rate, (i) the coefficient of susceptibility; (ii) the coefficient of virulence; (iii) the number of contacts per unit of time [37]; (iv) the environmental conductivity [38]. Let us remark that the rate of contacts per units of time can also be investigated by agent models [39].

Epidemic models with time-dependent transmission rates have been considered in several articles in the literature. The classical approach is to fix a function of time that depends on some parameters and to fix these parameters by using the best fit to the data. In Chowell et al. [40] a specific form was chosen for the rate of transmission and applied to the Ebola outbreak in Congo. Huo et al [41] used a predefined transmission rate which is a Legendre polynomial depending on a tunable number of parameters. Let us also mention that kinetic model idea has been used to understand this problem in the paper by Dimarco et al. [42]. Here we are going the other way around. We reconstruct the transmission rate from the data by using the model without choosing a predefined function for the transmission rate. Such an approach was used in the early 70s by London and Yorke [43, 44] who used a discrete-time model and discussed the time-dependent rate of transmission in the context of measles, chickenpox, and mumps. More recently, several authors [36, 45, 46] used both an explicit formula and algorithms to reconstruct the transmission rate. These studies allow us to understand that the regularization of the data is a complex problem and is crucial in order to rebuild a meaningful time-dependent transmission

rate.

In the present paper, we apply a new method to compute the transmission rate from cumulative reported cases data. While the use of a predefined transmission rate $\tau(t)$ as a function of time can lead to very good fits of the data, here we are looking for a more intrinsic relationship between the data and the transmission rate. Therefore we propose a different approach and use a two-step procedure. Firstly, we use a phenomenological model to describe the data and extract the general trend of the epidemiological dynamics while removing the insignificant noise. Secondly, we derive an explicit relationship between the phenomenological model and the transmission rate. In other words, we compute the transmission rate directly from the data. As a result, we can reconstruct an estimation of the state of the population at each time, including covert cases. Our method also provides new indicators for the epidemiological dynamics that are related to the reproductive number.

2. Methods

2.1. COVID-19 data and phenomenological model

We regularized the time series of cumulative reported cases by fitting standard curves to the data to reconstruct the time-dependent transmission rate. We first identified the epidemic waves for each of the eight geographic areas. A Bernoulli–Verhulst curve was then fitted to each epidemic wave using the Levenberg–Marquardt algorithm [47]. We reported the detailed output of the algorithm in the supplementary material, including confidence bounds on the parameters. The model was completed by filling the time windows between two waves with straight lines. Finally, we applied a Gaussian filter with a standard deviation of 7 days to the curve to obtain a smooth model.

2.1.1. Data sources

We used reported cases data for 8 different geographic areas, namely California, France, India, Israel, Japan, Peru, Spain, and the UK. Apart from California State, for which we used data from the COVID tracking project [48], the reported cases data were taken from the WHO database [4].

2.1.2. Phenomenological model used for multiple epidemic waves

To represent the data, we used a phenomenological model to fit the curve of cumulative rate cases. Such an idea is not new since it was already proposed by Bernoulli [49] in 1760 in the context of the smallpox epidemic. Here we used the so-called Bernoulli–Verhulst [50] model to describe the epidemic phase. Bernoulli [49] investigated an epidemic phase followed by an endemic phase. This appears clearly in Figures 9 and 10 of the paper by Dietz, and Heesterbeek [51] who revisited the original article of Bernoulli. We also refer to Blower [52] for another article revisiting the original work of Bernoulli. Several works comparing cumulative reported cases data and the Bernoulli–Verhulst model appear in the literature (see [22, 24, 25]). The Bernoulli–Verhulst model is sometimes called Richard's model, although Richard's work came much later in 1959.

The phenomenological model deals with data series of new infectious cases decomposed into two successive phases: 1) endemic phases followed by 2) epidemic phases.

Endemic phase: During the endemic phase, the dynamics of new cases appears to fluctuate around an average value independently of the number of cases. Therefore the average cumulative number of

cases is given by

$$\mathrm{CR}(t) = N_0 + (t - t_0) \times a, \text{ for } t \in [t_0, t_1],$$ (2.1)

where t_0 denotes the beginning of the endemic phase, and a is the average value of the daily number of new cases.

We assume that the average daily number of new cases is constant. Therefore the daily number of new cases is given by

$$\mathrm{CR}'(t) = a.$$ (2.2)

Epidemic phase: In the epidemic phase, the new cases are contributing to produce secondary cases. Therefore the daily number of new cases is no longer constant, but varies with time as follows

$$\mathrm{CR}(t) = N_{\mathrm{base}} + \frac{e^{\chi(t-t_0)} N_0}{\left[1 + \frac{N_0^\theta}{N_\infty^\theta} \left(e^{\chi\theta(t-t_0)} - 1 \right) \right]^{1/\theta}}, \text{ for } t \in [t_0, t_1].$$ (2.3)

In other words, the daily number of new cases follows the Bernoulli–Verhulst [49, 50] equation. Namely, by setting

$$N(t) = \mathrm{CR}(t) - N_{\mathrm{base}},$$ (2.4)

we obtain

$$N'(t) = \chi N(t) \left[1 - \left(\frac{N(t)}{N_\infty} \right)^\theta \right],$$ (2.5)

completed with the initial value

$$N(t_0) = N_0.$$

In the model, $N_{\mathrm{base}} + N_0$ corresponds to the value $\mathrm{CR}(t_0)$ of the cumulative number of cases at time $t = t_0$. The parameter $N_\infty + N_{\mathrm{base}}$ is the maximal value of the cumulative reported cases after the time $t = t_0$. $\chi > 0$ is a Malthusian growth parameter, and θ regulates the speed at which $\mathrm{CR}(t)$ increases to $N_\infty + N_{\mathrm{base}}$.

Regularize the junction between the epidemic phases: Because the formula for $\tau(t)$ involves derivatives of the phenomenological model regularizing $\mathrm{CR}(t)$ (see Eqs (2.12)–(2.15)), we need to connect the phenomenological models of the different phases as smoothly as possible. Let $t_0, \ldots, t_n$ denote the $n + 1$ breaking points of the model, that is, the times at which there is a transition between one phase and the next one. We let $\widetilde{\mathrm{CR}}(t)$ be the global model obtained by placing the phenomenological models of the different phases side by side.

More precisely, $\widetilde{\mathrm{CR}}(t)$ is defined by Eq (2.3) during an epidemic phase $[t_i, t_{i+1}]$, or during the initial phase $(-\infty, t_0]$ or the last phase $[t_n, +\infty)$. During an endemic phase, $\widetilde{\mathrm{CR}}(t)$ is defined by Eq (2.1). The parameters are chosen so that the resulting global model $\widetilde{\mathrm{CR}}$ is continuous. We define the regularized model by using the convolution formula:

$$\mathrm{CR}(t) = \int_{-\infty}^{+\infty} \mathcal{G}(t - s) \times \widetilde{\mathrm{CR}}(s) \mathrm{d}s = (\mathcal{G} * \widetilde{\mathrm{CR}})(t),$$ (2.6)

where

$$\mathcal{G}(t) := \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}$$

is the Gaussian function with mean 0 and variance σ^2. The parameter σ controls the trade-off between smoothness and precision: increasing σ reduces the variations in CR(t) and reducing σ reduces the distance between CR(t) and $\widetilde{\text{CR}}(t)$. In any case the resulting function CR(t) is very smooth (as well as its derivatives) and close to the original model $\widetilde{\text{CR}}(t)$ when σ is not too large. In the results section (Section 3), we fix $\sigma = 7$ days.

Numerically, we will need to compute some $t \to \text{CR}(t)$ derivatives. Therefore it is convenient to take advantage of the convolution Eq (2.6) and deduce that

$$\frac{\mathrm{d}^n \text{CR}(t)}{\mathrm{d}t^n} = \int_{-\infty}^{+\infty} \frac{\mathrm{d}^n \mathcal{G}(t-s)}{\mathrm{d}t^n} \times \widetilde{\text{CR}}(s)\mathrm{d}s, \tag{2.7}$$

for $n = 1, 2, 3$.

2.2. Epidemic model

To reconstruct the transmission rate, we used the underlying mathematical model described by the flowchart presented in Figure 1.
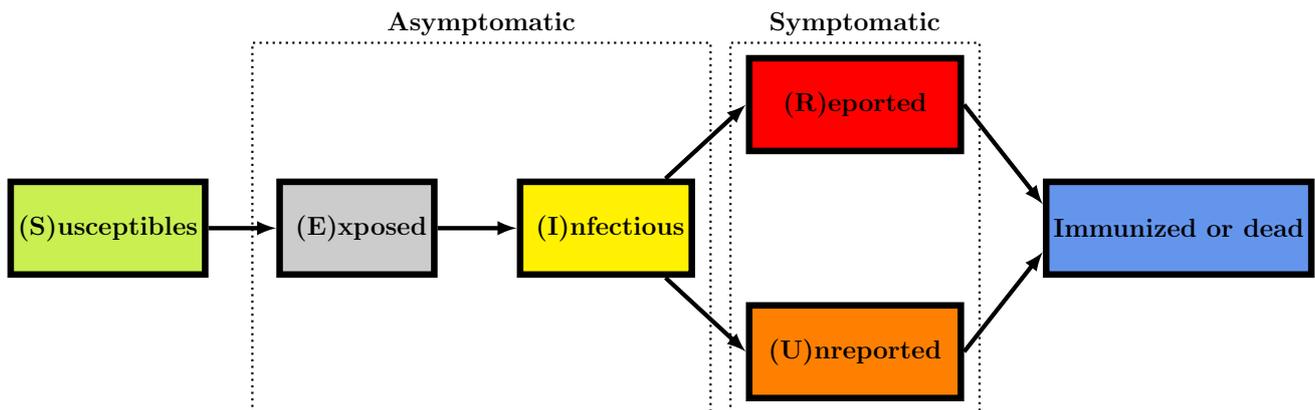
Figure 1. Flowchart for the model.

The model itself includes five parameters whose values were taken from the literature: the average length of the noninfectious incubation period (1 day, (E)xposed); the average length of the infectious incubation period (3 days, (I)nfectious); the average length of the symptomatic period (7 days, (R)eported or (U)nreported); the ascertainment rate (0.8). Additional parameters appear in the initial condition and could not be computed from the initial number of unreported individuals. The transmission rate was computed from the regularized data and the assumed parameters according to a methodology adapted from Demongeot et al. [36].

Many epidemiological models are based on the SIR or SEIR model, which is classical in epidemic modelling. We refer to [53, 54] for the earliest articles devoted to such a question and to [55–65] for more models. In this chapter, we will compare the following SEIUR model to the cumulative reported

cases data

$$
\begin{cases}
S'(t) = -\tau(t)\left[I(t) + \kappa\,U(t)\right]S(t), \\
E'(t) = \tau(t)\left[I(t) + \kappa\,U(t)\right]S(t) - \alpha\,E(t), \\
I'(t) = \alpha\,E(t) - \nu\,I(t), \\
U'(t) = \nu\,(1 - f)\,I(t) - \eta\,U(t), \\
R'(t) = \nu\,f\,I(t) - \eta\,R(t),
\end{cases}
\tag{2.8}
$$

where at time t, $S(t)$ is the number of susceptible, $E(t)$ the number of exposed (not yet capable of transmitting the pathogen), $I(t)$ the number of asymptomatic infectious, $R(t)$ the number of reported symptomatic infectious and $U(t)$ the number of unreported symptomatic infectious. This system is supplemented by initial data

$$
S(t_0) = S_0,\ E(t_0) = E_0,\ I(t_0) = I_0,\ U(t_0) = U_0,\ \text{and } R(t_0) = R_0.
\tag{2.9}
$$

In this model, $\tau(t)$ is the rate of transmission, $1/\alpha$ is the average duration of the exposure period, $1/\nu$ is the average duration of the asymptomatic infectious period, and for simplicity, we subdivide the class of symptomatic patients into the fraction $0 \le f \le 1$ of patients showing some severe symptoms, and the fraction $1 - f$ of patients showing some mild symptoms assumed to be not detected. The quantity $1/\eta$ is the average duration of the symptomatic infectious period. In the model, we assume that the average time of infection is the same for Reported and Unreported infectious individuals. We refer to [66, 67] for more information about this topic. Finally, we assume that reported symptomatic individuals do not contribute significantly to the transmission of the virus.

The cumulative number of reported cases $CR(t)$ is connected to the epidemic model by the following relationship

$$
CR(t) = CR_0 + \nu\,f\,CI(t),\ \text{for } t \ge t_0,
\tag{2.10}
$$

where

$$
CI(t) = \int_{t_0}^{t} I(\sigma)\,d\sigma.
\tag{2.11}
$$

Given and estimated parameters

We assume that the following parameters of the model are known

$$
S_0, U_0, R_0, f, \kappa, \alpha, \nu, \eta.
$$

The goal of our method is to focus on the estimation of the three remaining parameters. Namely, knowing the parameters mentioned above, we plan to identify

$$
E_0, I_0, \tau(t).
$$

Computation of the rate of transmission

The transmission rate is fully determined by the parameters $\kappa, \alpha, \nu, \eta, f, S_0, E_0, I_0, U_0$, and the data that are represented by the function $t \to CR(t)$, by using the three following equations

$$\tau(t) = \frac{1}{I(t) + \kappa U(t)} \times \frac{CE''(t) + \alpha CE'(t)}{E_0 + S_0 - CE'(t) - \alpha CE(t)}, \tag{2.12}$$

where

$$I(t) = \frac{CR'(t)}{\nu f}, \tag{2.13}$$

$$CE(t) = \frac{1}{\alpha \nu f} \left[CR'(t) - \nu f I_0 + \nu (CR(t) - CR_0) \right], \tag{2.14}$$

$$U(t) = e^{-\eta(t-t_0)} U_0 + \int_{t_0}^{t} e^{-\eta(t-s)} \frac{(1-f)}{f} CR'(s) ds. \tag{2.15}$$

2.2.1. Instantaneous reproduction number computed for COVID-19 data

We have only a single epidemic phase in the standard SI epidemic model because the epidemic exhausts the susceptible population. Here, the changes of regimes (epidemic phase versus endemic phase) are partly due to the decay in the number of susceptible. But these changes are also influenced by the changes in the transmission rate. These changes in the transmission rate are due to the limitation of contacts between individuals or to changes in climate (in summer) or other factors influencing transmissions.

In this section, we will observe that the main factors for the changes in the epidemic regimes are the changes in the transmission rate. To investigate this for the COVID-19 data, we use our method to compute the transmission rate, and we consider the *instantaneous reproduction number*

$$R_e(t) = \frac{\tau(t)S(t)}{\eta\nu}(\eta + \nu(1 - f)), \tag{2.16}$$

and the *quasi-instantaneous reproduction number*

$$R_e^0(t) = \frac{\tau(t)S_0}{\eta\nu}(\eta + \nu(1 - f)), \tag{2.17}$$

in which the transmission varies, but the size of the susceptible population remains constant equal to S_0. We refer to subsection S8 for detailed computations to obtain the Eq (2.16).

The comparison between $R_e(t)$ and $R_e^0(t)$ permits us to understand the contribution of the decay of the susceptible population in the variations of $R_e(t)$. Another interesting aspect is that $R_e^0(t)$ is proportional to the transmission rate $\tau(t)$. Therefore plotting $R_e^0(t)$ permits us to visualize the variation of $t \to \tau(t)$ only.

2.2.2. Computation of the initial value of the epidemic model

Based on Eq (2.4), we can recover the initial number of asymptomatic infectious $I_0 = I(t_0)$ and the initial number of exposed $E_0 = E(t_0)$ for an epidemic phase starting at time t_0. Indeed by definition,

we have $CR'(t) = \nu f I(t)$ and therefore

$$I_0 = \frac{CR'(t_0)}{\nu f} = \frac{\chi N_0 \left(1 - \left(\frac{N_0}{N_\infty}\right)^\theta\right)}{\nu f}.$$

Estimated initial number of infected

The initial number of asymptomatic infectious is given by

$$I_0 = \frac{CR'(t_0)}{\nu f}. \tag{2.18}$$

In the special case of the Bernoulli–Verhulst model we obtain

$$I_0 = \frac{\chi}{\nu f} N_0 \left(1 - \left(\frac{N_0}{N_\infty}\right)^\theta\right). \tag{2.19}$$

By differentiating Eq (2.5) we deduce that

$$
\begin{aligned}
N''(t) &= \chi N'(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right) - \frac{\chi\theta}{N_\infty^\theta} N(t) (N(t))^{\theta-1} N'(t) \\
&= \chi N'(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right) - \frac{\chi\theta}{N_\infty^\theta} (N(t))^\theta N'(t),
\end{aligned}
$$

therefore

$$CR''(t) = N''(t) = \chi^2 N(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right)\left(1 - (1+\theta)\left(\frac{N(t)}{N_\infty}\right)^\theta\right).$$

By using the third equation in Eq (2.8) we obtain

$$E_0 = \frac{I'(t_0) + \nu I(t_0)}{\alpha} = \frac{CR''(t_0) + \nu CR'(t_0)}{\alpha} = \frac{N''(t_0) + \nu N'(t_0)}{\alpha}.$$

Estimated initial number of exposed

The initial number of exposed is given by

$$E_0 = \frac{CR''(t_0) + \nu CR'(t_0)}{\alpha}. \tag{2.20}$$

In the special case of the Bernoulli–Verhulst model, we obtain

$$E_0 = \frac{\chi}{\alpha \nu f} N_0 \left(1 - \left(\frac{N_0}{N_\infty}\right)^\theta\right)\left(\chi + \nu - \chi(1+\theta)\left(\frac{N_0}{N_\infty}\right)^\theta\right). \tag{2.21}$$

2.2.3. Theoretical formula for $\tau(t)$

We first remark that the S-equation of model (2.8) can be written as

$$\frac{d}{dt} \ln(S(t)) = \frac{S'(t)}{S(t)} = -\tau(t)[I(t) + \kappa U(t)],$$

therefore by integrating between t_0 and t we get

$$S(t) = S_0 \exp\left(-\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma\right).$$

Next we plug the above formula for $S(t)$ into the E-equation of model (2.8) and obtain

$$E'(t) = S_0 \exp\left(-\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma\right)\tau(t)\ [I(t) + \kappa U(t)] - \alpha\, E(t)$$

$$= -S_0 \frac{d}{dt}\left(-\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma\right)\exp\left(-\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma\right)$$

$$- \alpha\, E(t),$$

and by integrating this equation between t_0 and t we obtain

$$E(t) = E_0 + S_0\left[1 - \exp\left(-\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma\right)\right] - \alpha \int_{t_0}^{t} E(\sigma)d\sigma. \tag{2.22}$$

Define the cumulative numbers of exposed, infectious and unreported individuals by

$$\mathrm{CE}(t) := \int_{t_0}^{t} E(\sigma)d\sigma,\ \mathrm{CI}(t) := \int_{t_0}^{t} I(\sigma)d\sigma,\ \text{and}\ \mathrm{CU}(t) := \int_{t_0}^{t} U(\sigma)d\sigma,$$

and note that $\mathrm{CE}'(t) = E(t)$. We can rewrite the Eq (2.22) as

$$S_0 \exp\left(-\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma\right) = E_0 + S_0 - \mathrm{CE}'(t) - \alpha\, \mathrm{CE}(t).$$

By taking the logarithm of both sides we obtain

$$\int_{t_0}^{t} \tau(\sigma)\ [I(\sigma) + \kappa U(\sigma)]\, d\sigma = \ln(S_0) - \ln\left(E_0 + S_0 - \mathrm{CE}'(t) - \alpha\, \mathrm{CE}(t)\right),$$

and by differentiating with respect to t:

$$\tau(t) = \frac{1}{I(t) + \kappa U(t)} \times \frac{\mathrm{CE}''(t) + \alpha \mathrm{CE}'(t)}{E_0 + S_0 - \mathrm{CE}'(t) - \alpha\, \mathrm{CE}(t)}. \tag{2.23}$$

Therefore we have an explicit formula giving $\tau(t)$ as a function of $I(t)$, $U(t)$ and $\mathrm{CE}(t)$ and its derivatives. Next we explain how to identify those three remaining unknowns as a function of $\mathrm{CR}(t)$ and its derivatives. We first recall that, from Eq (2.10), we have

$$\mathrm{CR}(t) = \mathrm{CR}(t_0) + \nu f\, \mathrm{CI}(t).$$

The I-equation of model (2.8) can be rewritten as

$$\alpha E(t) = I'(t) + \nu I(t),$$

and by integrating this equation between t_0 and t we obtain

$$\alpha\,\mathrm{CE}(t) = \mathrm{CI}'(t) - I_0 + \nu\,\mathrm{CI}(t) = \frac{1}{\nu f}(\mathrm{CR}'(t) + \nu\mathrm{CR}(t) - \nu\mathrm{CR}(t_0)). \tag{2.24}$$

Finally, by applying the variation of constants formula to the U-equation of system (2.8) we obtain

$$U(t) = \mathrm{e}^{-\eta(t-t_0)}U_0 + \int_{t_0}^t \mathrm{e}^{-\eta(t-s)}\nu\,(1-f)\,I(s)\mathrm{d}s$$

$$= \mathrm{e}^{-\eta(t-t_0)}U_0 + \int_{t_0}^t \mathrm{e}^{-\eta(t-s)}\frac{1-f}{f}\mathrm{CR}'(s)\mathrm{d}s. \tag{2.25}$$

From these computations we deduce that $\tau(t)$ can be computed thanks to Eq (2.23) from $\mathrm{CR}(t)$, α, ν, η, κ, f and U_0. The following theorem is a precise statement of this result.

Theorem 2.1. *Let $S_0 > 0$, $E_0 > 0$, $I_0 > 0$, $U_0 > 0$, $\mathrm{CR}_0 \geq 0$, α, ν, η and $f > 0$ be given. Let $t \mapsto \tau(t) \geq 0$ be a given continuous function and $t \to I(t)$ be the second component of system (2.8). Let $\widehat{\mathrm{CR}} : [t_0, \infty) \to \mathbb{R}$ be a twice continuously differentiable function. Then*

$$\widehat{\mathrm{CR}}(t) = \mathrm{CR}_0 + \nu f \int_{t_0}^t I(s)\,ds, \forall t \geq t_0, \tag{2.26}$$

if and only if $\widehat{\mathrm{CR}}$ satisfies

$$\widehat{\mathrm{CR}}(t_0) = \mathrm{CR}_0, \tag{2.27}$$

$$\widehat{\mathrm{CR}}'(t_0) = \nu f I_0, \tag{2.28}$$

$$\widehat{\mathrm{CR}}''(t_0) + \nu\widehat{\mathrm{CR}}'(t_0) = \alpha \nu f E_0, \tag{2.29}$$

$$\widehat{\mathrm{CR}}'(t) > 0, \forall t \geq t_0, \tag{2.30}$$

$$\nu f (E_0 + S_0) - \left[\widehat{\mathrm{CR}}''(t) + \nu\widehat{\mathrm{CR}}'(t)\right] - \alpha\left[\widehat{\mathrm{CR}}'(t) - \nu f I_0 + \nu\widehat{\mathrm{CR}}(t)\right] > 0, \forall t \geq t_0, \tag{2.31}$$

and $\tau(t)$ is given by

$$\tau(t) = \frac{1}{\widehat{I}(t) + \kappa\widehat{U}(t)} \times \frac{\widehat{\mathrm{CE}}''(t) + \alpha\widehat{\mathrm{CE}}'(t)}{E_0 + S_0 - \widehat{\mathrm{CE}}'(t) - \alpha\widehat{\mathrm{CE}}(t)}, \tag{2.32}$$

where

$$\widehat{I}(t) := \frac{\widehat{\mathrm{CR}}'(t)}{\nu f}, \tag{2.33}$$

$$\widehat{\mathrm{CI}}(t) := \frac{1}{\nu f}\left[\widehat{\mathrm{CR}}(t) - \widehat{\mathrm{CR}}(t_0)\right], \tag{2.34}$$

$$\widehat{\mathrm{CE}}(t) := \frac{1}{\alpha}\left[\widehat{\mathrm{CI}}'(t) - I_0 + \nu\widehat{\mathrm{CI}}(t)\right] = \frac{1}{\alpha \nu f}\left[\widehat{\mathrm{CR}}'(t) - \nu f I_0 + \nu\left(\widehat{\mathrm{CR}}(t) - \mathrm{CR}_0\right)\right], \tag{2.35}$$

$$\widehat{U}(t) := \mathrm{e}^{-\eta(t-t_0)}U_0 + \int_{t_0}^t \mathrm{e}^{-\eta(t-s)}\frac{(1-f)}{f}\widehat{\mathrm{CR}}'(s)\mathrm{d}s. \tag{2.36}$$

Proof. Assume first that $\widehat{\mathrm{CR}}(t)$ satisfies Eq (2.26). Then by using the first equation of system (2.8) we deduce that

$$S_0 \exp\left(-\int_{t_0}^t \tau(\sigma)\,[I(\sigma) + \kappa U(\sigma)]\,d\sigma\right) = E_0 + S_0 - E(t) - \alpha\mathrm{CE}(t). \tag{2.37}$$

Therefore

$$\int_{t_0}^t \tau(\sigma)\,[I(\sigma) + \kappa U(\sigma)]\,d\sigma = \ln\left[\frac{S_0}{E_0 + S_0 - E(t) - \alpha\mathrm{CE}(t)}\right]$$
$$= \ln(S_0) - \ln[E_0 + S_0 - E(t) - \alpha\mathrm{CE}(t)],$$

and by taking the derivative of both sides we obtain

$$\tau(t)\,[I(t) + \kappa U(t)] = \frac{E'(t) + \alpha E(t)}{E_0 + S_0 - E(t) - \alpha\mathrm{CE}(t)},$$

which is equivalent to

$$\tau(t) = \frac{E(t)}{I(t) + \kappa U(t)} \times \frac{\dfrac{E'(t)}{E(t)} + \alpha}{E_0 + S_0 - E(t) - \alpha\mathrm{CE}(t)}.$$

By using the fact that $E(t) = \mathrm{CE}'(t)$ and $I = \mathrm{CR}'(t)/(\nu f)$, we deduce Eq (2.32). By differentiating Eq (2.26), we get Eqs (2.28) and (2.30). Equation (2.29) is a consequence of the E-component of Eq (2.8). We get Eq (2.31) by combining Eqs (2.37) and (2.35) (since $\widehat{\mathrm{CE}}(t) = \mathrm{CE}(t)$).

Conversely, assume that $\tau(t)$ is given by Eq (2.31) and all the Eqs (2.27)–(2.36) hold. We define $\widehat{I}(t) = \widehat{\mathrm{CR}}'(t)/\nu f$ and $\widehat{\mathrm{CI}}(t) = \left(\widehat{\mathrm{CR}}(t) - \mathrm{CR}_0\right)/\nu f$. Then, by using Eq (2.27), we deduce that

$$\widehat{\mathrm{CI}}(t) = \int_{t_0}^t \widehat{I}(\sigma)\,d\sigma, \tag{2.38}$$

and by using Eq (2.28), we deduce

$$\widehat{I}(t_0) = I_0. \tag{2.39}$$

Moreover, from Eq (2.31) and $\widehat{I}(t) = \widehat{\mathrm{CR}}'(t)/\nu f$ we deduce that

$$\tau(t) = \frac{1}{\widehat{I}(t) + \kappa\widehat{U}(t)} \times \frac{\widehat{\mathrm{CE}}''(t) + \alpha\widehat{\mathrm{CE}}'(t)}{E_0 + S_0 - \widehat{\mathrm{CE}}'(t) - \alpha\widehat{\mathrm{CE}}(t)}. \tag{2.40}$$

Multiplying Eq (2.40) by $\widehat{I}(t) + \kappa\widehat{U}(t)$ and integrating, we obtain

$$\int_{t_0}^t \tau(\sigma)\left[\widehat{I}(\sigma) + \kappa\widehat{U}(\sigma)\right]d\sigma = \ln\left(E_0 + S_0 - \widehat{\mathrm{CE}}'(t_0) - \alpha\widehat{\mathrm{CE}}(t_0)\right) \\ - \ln\left(E_0 + S_0 - \widehat{\mathrm{CE}}'(t) - \alpha\widehat{\mathrm{CE}}(t)\right), \tag{2.41}$$

where the right-hand side is well defined thanks to Eq (2.31). By combining Eqs (2.27), (2.28) and (2.35) we obtain

$$\widehat{\mathrm{CE}}(t_0) = 0, \tag{2.42}$$

and by taking the derivative in Eq (2.35) we obtain

$$\widehat{CE}'(t_0) = \frac{1}{\alpha \nu f} \left[\widehat{CR}''(t) + \nu \widehat{CR}'(t) \right]$$

therefore by using Eq (2.29) we deduce that

$$\widehat{CE}'(t_0) = E_0. \tag{2.43}$$

In particular, $E_0 + S_0 - \widehat{CE}'(t_0) - \alpha\widehat{CE}(t_0) = S_0$ and, by taking the exponential of Eq (2.41), we obtain

$$S_0 e^{-\int_{t_0}^t \tau(\sigma)\left[\widehat{I}(\sigma)+\kappa\widehat{U}(\sigma)\right]d\sigma} = E_0 + S_0 - \widehat{CE}'(t) - \alpha\widehat{CE}(t),$$

which, differentiating both sides, yields

$$-S_0 e^{-\int_{t_0}^t \tau(\sigma)\left[\widehat{I}(\sigma)+\kappa\widehat{U}(\sigma)\right]d\sigma} \tau(t)\left[\widehat{I}(t) + \kappa\widehat{U}(t)\right] = -\widehat{CE}''(t) - \alpha\widehat{CE}'(t)$$
$$= -\widehat{E}'(t) - \alpha\widehat{E}(t),$$

and therefore

$$\widehat{E}'(t) = \tau(t)\widehat{S}(t)\left[\widehat{I}(t) + \kappa\widehat{U}(t)\right] - \alpha\widehat{E}(t), \tag{2.44}$$

where $\widehat{E}(t) := \widehat{CE}'(t)$ and $\widehat{S}(t) := S_0 e^{-\int_{t_0}^t \tau(\sigma)\left[\widehat{I}(\sigma)+\kappa\widehat{U}(\sigma)\right]d\sigma}$. Differentiating the definition of $\widehat{S}(t)$, we get

$$\widehat{S}'(t) = -\left[\widehat{I}(t) + \kappa\widehat{U}(t)\right]\widehat{S}(t). \tag{2.45}$$

Next the derivative of Eq (2.35) can be rewritten as

$$\widehat{I}'(t) = \frac{1}{\nu f}\widehat{CR}''(t) = \alpha\widehat{CE}'(t) - \nu\frac{1}{\nu f}\widehat{CR}'(t) = \alpha\widehat{E}(t) - \nu\widehat{I}(t). \tag{2.46}$$

Finally, differentiating Eq (2.36) yields

$$\widehat{U}'(t) = \nu(1 - f)\widehat{I}(t) - \eta\widehat{U}(t). \tag{2.47}$$

By combining Eqs (2.44)–(2.47) we see that $(\widehat{S}(t), \widehat{E}(t), \widehat{I}(t), \widehat{U}(t))$ satisfies Eq (2.8) with the initial condition $(\widehat{S}(t_0), \widehat{E}(t_0), \widehat{I}(t_0), \widehat{U}(t_0)) = (S_0, E_0, I_0, U_0)$. By the uniqueness of the solutions of Eq (2.8) for a given initial condition, we conclude that $(\widehat{S}(t), \widehat{E}(t), \widehat{I}(t), \widehat{U}(t)) = (S(t), E(t), I(t), U(t))$. In particular, $CR(t)$ satisfies Eq (2.26). The proof is completed. □

Remark 2.2. *The condition Eq (2.31) is equivalent to*

$$E_0 + S_0 - \widehat{CE}'(t) - \alpha\widehat{CE}(t) > 0, \ \forall t \geq t_0.$$

Remark 2.3. *The present computations have been previously done, in a different context, by Hadeler [68].*

2.3. Computing the explicit formula for $\tau(t)$ during an epidemic phase

In this section, we assume that the curve of cumulative reported cases is given by the Bernoulli–Verhulst formula

$$N(t) := \mathrm{CR}(t) - N_{\mathrm{base}} = \frac{e^{\chi(t-t_0)} N_0}{\left[1 + \frac{N_0^\theta}{N_\infty^\theta} \left(e^{\chi\theta(t-t_0)} - 1\right)\right]^{1/\theta}}, \text{ for } t \in [t_0, t_1],$$

and we recall that

$$N'(t) = \chi N(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right).$$

Then we can compute an explicit formula for the components of the system (2.8). By definition we have

$$I(t) = \frac{\mathrm{CR}'(t)}{\nu f} = \frac{\chi}{\nu f} N(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right), \tag{2.48}$$

which gives

$$I'(t) = \frac{\mathrm{CR}''(t)}{\nu f} = \frac{\chi^2}{\nu f} N(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right) \left(1 - (1+\theta)\left(\frac{N(t)}{N_\infty}\right)^\theta\right),$$

so that by using the I-component in the system (2.8) we get

$$E(t) = \frac{1}{\alpha}(I'(t) + \nu I(t)) = \frac{1}{\alpha\nu f}\left(\mathrm{CR}''(t) + \nu\mathrm{CR}'(t)\right).$$

By integration, we get

$$\begin{aligned}
\mathrm{CE}(t) &= \frac{1}{\alpha \nu f}\left[(\mathrm{CR}'(t) - \mathrm{CR}_0') + \nu\left[\mathrm{CR}(t) - \mathrm{CR}(t_0)\right]\right], \\
&= \frac{1}{\alpha \nu f}\left[\chi N(t)\left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right) - \nu f I_0 + \nu\left[N(t) - N_0\right]\right], \\
&= \frac{1}{\alpha \nu f}\left[N(t)\left(\chi + \nu - \chi\left(\frac{N(t)}{N_\infty}\right)^\theta\right) - \nu f I_0 - \nu N_0\right],
\end{aligned}$$

and since

$$\nu f I_0 = \mathrm{CR}'(t_0) = N'(t_0) = \chi N_0\left(1 - \left(\frac{N_0}{N_\infty}\right)^\theta\right),$$

we obtain

$$\mathrm{CE}(t) = \frac{1}{\alpha \nu f}\left[N(t)\left(\chi + \nu - \chi\left(\frac{N(t)}{N_\infty}\right)^\theta\right) - N_0\left(\chi + \nu - \chi\left(\frac{N_0}{N_\infty}\right)^\theta\right)\right].$$

Note also that we have explicit formulas for $E(t) = \mathrm{CE}'(t)$ and $E'(t) = \mathrm{CE}''(t)$,

$$E(t) = \mathrm{CE}'(t) = \frac{\chi}{\alpha\nu f}\left[N(t)\left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right)\left(\chi + \nu - \chi(1+\theta)\left(\frac{N(t)}{N_\infty}\right)^\theta\right)\right] \tag{2.49}$$

and

$$E'(t) = CE''(t) = \frac{\chi^2}{\alpha v f} N(t) \left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right)$$

$$\times \left[\chi + v - (\chi(2+\theta)+v)(1+\theta)\left(\frac{N(t)}{N_\infty}\right)^\theta + \chi(1+\theta)(1+2\theta)\left(\frac{N(t)}{N_\infty}\right)^{2\theta}\right].$$

Next, recall the U-equation of Eq (2.8), that is,

$$U'(t) = v(1-f)I(t) - \eta U(t),$$

therefore by the variation of constant formula we have

$$U(t) = e^{-\eta(t-t_0)}U(t_0) + \int_{t_0}^t e^{-\eta(t-s)}(1-f)vI(s)\mathrm{d}s$$

$$= e^{-\eta(t-t_0)}U_0 + \int_{t_0}^t e^{-\eta(t-s)}\frac{1-f}{f}CR'(s)\mathrm{d}s. \tag{2.50}$$

Explicit formula for the transmission rate during an epidemic phase

The transmission rate $\tau(t)$ can be computed as

$$\tau(t) = \frac{\chi N(t)\left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right)}{I(t) + \kappa U(t)} \times \frac{\left[A\left(\frac{N(t)}{N_\infty}\right)^{2\theta} - B\left(\frac{N(t)}{N_\infty}\right)^\theta + C\right]}{E_0 + S_0 - E(t) - \alpha CE(t)}, \tag{2.51}$$

where

$$N(t) = \frac{e^{\chi(t-t_0)}N_0}{\left[1 + \frac{N_0^\theta}{N_\infty^\theta}\left(e^{\chi\theta(t-t_0)} - 1\right)\right]^{1/\theta}}, \quad \text{for } t \ge t_0, \tag{2.52}$$

---

and

$$A := \chi^2(1+\theta)(1+2\theta), \tag{2.53}$$
$$B := \chi(1+\theta)[\chi(2+\theta)+v+\alpha], \tag{2.54}$$
$$C := (\alpha+\chi)(\chi+v), \tag{2.55}$$

and $I(t)$ is given by Eq (2.48), $E(t)$ by Eq (2.49) and $U(t)$ by Eq (2.50).

2.3.1. Compatibility conditions for the positivity of the transmission rate

Recall from Eq (2.51):

$$\tau(t) = \frac{\chi N(t)\left(1 - \left(\frac{N(t)}{N_\infty}\right)^\theta\right)}{I(t) + \kappa U(t)} \times \frac{\left[A\left(\frac{N(t)}{N_\infty}\right)^{2\theta} - B\left(\frac{N(t)}{N_\infty}\right)^\theta + C\right]}{E_0 + S_0 - E(t) - \alpha CE(t)}.$$

Here we require that the numerator and the denominator of the last fraction stay positive for all times.

Positivity of the numerator: The model is compatible with the data if the transmission rate $\tau(t)$ stays positive for all times $t \in \mathbb{R}$. The numerator

$$p(N) := AN^2 - BN + C$$

is a second-order polynomial with $N \in (0, 1)$. Let $\Delta := B^2 - 4AC$ be the discriminant of $p(N)$. Since $p'(0) = -B < 0$ and

$$p'(N) = 0 \Leftrightarrow N = \frac{B}{2A}$$

we have two cases: 1) $\frac{B}{2A} \geq 1$; or 2) $0 < \frac{B}{2A} < 1$.

Case 1: If $\frac{B}{2A} \geq 1$, $p(N)$ is non-negative for all $N \in [0, 1]$ if and only if

$$p(1) > 0 \Leftrightarrow A + C - B > 0. \tag{2.56}$$

Substituting A, B, C by their expression, we get

$$\begin{aligned}
A + C - B &= \chi^2(1 + \theta)(1 + 2\theta) + (\alpha + \chi)(\chi + \nu) - \chi(1 + \theta)(\chi(2 + \theta) + \alpha + \nu) \\
&= \chi^2 + 2\chi^2\theta + \chi^2\theta + 2\chi^2\theta^2 + \alpha\chi + \alpha\nu + \chi^2 + \chi\nu \\
&\quad - 2\chi^2 - \chi\theta - 2\chi^2\theta - \chi^2\theta^2 - \alpha\chi - \nu\chi - \alpha\chi\theta - \nu\chi\theta \\
&= \chi^2\theta^2 + \alpha\nu - \alpha\chi\theta - \nu\chi\theta \\
&= (\alpha - \chi\theta)(\nu - \chi\theta).
\end{aligned}$$

Case 2: If $\frac{B}{2A} < 1$, $p(N)$ is non-negative for all $N \in [0, 1]$ if and only if

$$p\left(\frac{B}{2A}\right)(1) > 0 \Leftrightarrow \Delta < 0 \Leftrightarrow B^2 - 4AC < 0. \tag{2.57}$$

Lemma 2.4. $\Delta < 0 \Rightarrow A + C - B > 0$.

Proof. We have

$$\Delta < 0 \Rightarrow B^2 - 4AC \leq (B - 2A)^2 \Leftrightarrow B^2 - 4AC \leq B^2 - 4AB + 4A^2$$

and after simplifying the result follows. $\qquad \square$

Positivity of the denominator: Next we turn to the denominator in the expression of τ, i.e., we want to ensure

$$E_0 + S_0 - E(t) - \alpha CE(t) > 0 \text{ for all } t \in \mathbb{R}. \tag{2.58}$$

We let $Y := \frac{N(t)}{N_\infty}$ and observe that $E(t) + \alpha CE(t)$ can be written as

$$\begin{aligned}
E(t) + \alpha CE(t) = \frac{1}{\alpha\nu f}[\chi N_\infty Y(1 - Y^\theta)(\chi + \nu - \chi(1 + \theta)Y^\theta) \\
+ \alpha N_\infty Y(\chi + \nu - \chi Y^\theta) - \alpha N_\infty Y_0(\chi + \nu - Y_0^\theta)]
\end{aligned}$$

$$= \frac{N_\infty}{\alpha \nu f} Y[(\chi + \alpha)(\chi + \nu) - \chi(\alpha + \nu + \chi(2 + \theta))Y^\theta + \chi^2(1 + \theta)Y^{2\theta}]$$
$$- \frac{N_0}{\nu f}\left(\chi + \nu - Y_0^\theta\right),$$

since we know that $A > 0$. Therefore Eq (2.58) becomes

$$Y[(\chi + \alpha)(\chi + \nu) - \chi(\chi + \nu + \chi(1 + \theta) + \alpha)Y^\theta + \chi^2(1 + \theta)Y^{2\theta}]$$
$$\leq \frac{\alpha \nu f}{N_\infty}\left[E_0 + S_0 + \frac{N_0}{\nu f}\left(\chi + \nu - \left(\frac{N_0}{N_\infty}\right)^\theta\right)\right].$$

We let

$$g(Y) := Y[(\chi + \alpha)(\chi + \nu) - \chi(\alpha + \nu + \chi(2 + \theta))Y^\theta + \chi^2(1 + \theta)Y^{2\theta}]$$

and notice that

$$g'(Y) = (\chi + \alpha)(\chi + \nu) - \chi(1 + \theta)(\alpha + \nu + \chi(2 + \theta))Y^\theta + \chi^2(1 + 2\theta)(1 + \theta)Y^{2\theta},$$

is exactly $p(N) := AN^2 - BN + C$.

Therefore, assuming that $A + C - B > 0$, the derivative $g'(Y)$ is positive and g is strictly increasing. So we only have to check the final value $g(1)$. We get

$$\frac{\alpha \nu f}{N_\infty}\left(S_0 + E_0 + \frac{N_0}{\nu f}\left(\chi + \nu - \left(\frac{N_0}{N_\infty}\right)^\theta\right)\right)$$
$$\geq (\chi + \alpha)(\chi + \nu) - \chi(\alpha + \nu - \chi(2 + \theta)) + \chi^2(1 + \theta)$$
$$= \chi^2 + \alpha\nu + \alpha\chi + \nu\chi + \chi^2 + \chi^2\theta - \alpha\chi - \nu\chi - 2\chi^2 - \chi^2\theta$$
$$= \alpha\nu.$$

Compatibility for the positivity

The SEIUR model is compatible with the data only when $\tau(t)$ stays positive for all $t \geq t_0$. Therefore the following two conditions should be met:

$$(\nu - \chi\theta)(\alpha - \chi\theta) \geq 0 \tag{2.59}$$

and

$$f + \frac{1}{\nu}\frac{N_0}{S_0 + E_0}\left(\chi + \nu - \left(\frac{N_0}{N_\infty}\right)^\theta\right) \geq \frac{N_\infty}{S_0 + E_0}. \tag{2.60}$$

2.3.2. Computing the explicit formula for $\tau(t)$ during an endemic phase

Recall that during an endemic phase, the cumulative number of cases is assumed to be a line. Therefore,

$$CR(t) = A(t - t_0) + B$$

and

$$CR'(t) = A \text{ and } CR''(t) = 0.$$

Therefore

$$I(t) = \frac{\text{CR}'(t)}{\nu f} = \frac{A}{\nu f} \qquad (2.61)$$

and

$$E(t) = \frac{I'(t) + \nu I(t)}{\alpha} = \frac{A}{\alpha f}. \qquad (2.62)$$

Hence

$$\text{CE}(t) = \frac{A}{\alpha f}(t - t_0). \qquad (2.63)$$

Moreover

$$U(t) = e^{-\eta(t-t_0)} U_0 + \int_{t_0}^{t} e^{-\eta(t-s)} \nu(1-f) I(s) ds,$$

and we obtain

$$U(t) = e^{-\eta(t-t_0)} U_0 + \frac{(1-f)A}{\eta f}\left(1 - e^{-\eta(t-t_0)}\right). \qquad (2.64)$$

By combining Eqs (2.12) and (2.61)–(2.64) we obtain the following explicit formula.

Explicit formula for the transmission rate during an endemic phase

The transmission rate $\tau(t)$ can be computed as

$$\tau(t) = \frac{1}{\frac{A}{\nu f} + \kappa\left(e^{-\eta(t-t_0)} U_0 + \frac{1-f}{\eta f} A\left(1 - e^{-\eta(t-t_0)}\right)\right)} \times \frac{A}{f S_0 - A(t-t_0)}, \qquad (2.65)$$

with the compatibility condition

$$t_0 \le t < \frac{f S_0}{A} + t_0.$$

Remark 2.5. *The above transmission rate corresponds to a constant number of daily infected A. Therefore it is impossible to maintain such a constant flux of new infected whenever the number of susceptible individuals is finite. The time $t = \dfrac{f S_0}{A} + t_0$ corresponds to the maximal time starting from t_0 during which we can maintain such a regime.*

3. Results

3.1. Phenomenological model applied to COVID-19 data

Our method to regularize the data was applied to the eight geographic areas. The resulting curves are presented in Figure 2. The blue background color regions correspond to epidemic phases, and the yellow background color regions to endemic phases. We added a plot of the daily number of cases (black dots) and the derivative of the regularized model for comparison, even though the daily number of cases is not used in the fitting procedure. The figures show in general, an extremely good agreement between the time series of reported cases (top row, black dots) and the regularized model (top row, blue curve). The match between the daily number of cases (bottom row, black dots) and

the derivative of the regularized model (bottom row, blue curve) is also excellent, even though it is not a part of the optimization process. Of course, we lose some of the information like the extremal values ("peaks") of the daily number of cases. This is because we focus on an averaged value of the number of cases. More information could be retrieved by studying statistically the variation around the phenomenological model. However, we leave such a study for future work. The relative error between the regularized curve and the data may be relatively high at the beginning of the epidemic because of the stochastic nature of the infection process and the small number of infected individuals but quickly drops below 1% (see the supplementary material for more details).
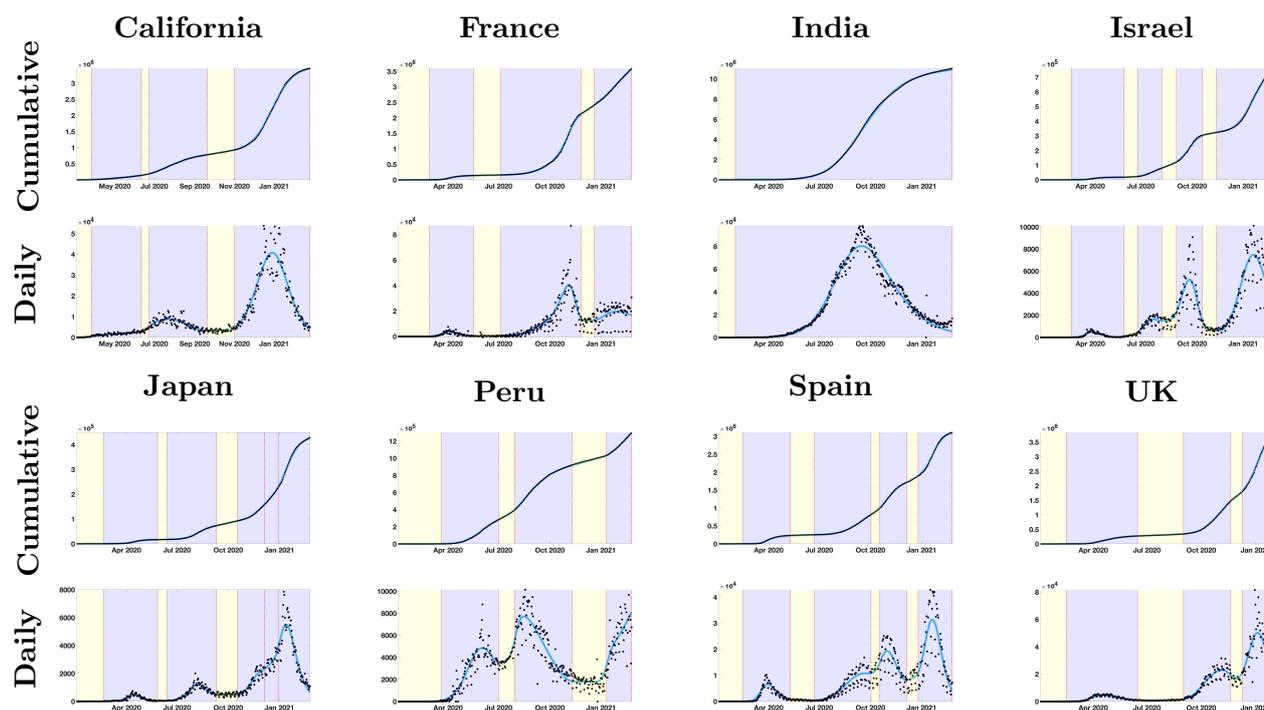
Figure 2. In the top rows, we plot the cumulative number of reported cases (black dots) and the best fit of the phenomenological model (blue curve). In the bottom rows, we plot the daily number of reported cases (black dots) and the first derivative of the phenomenological model (blue curve).

3.2. *Bounds for the value of non-identifiable parameters*

Even if some parameters of the mathematical model are not identifiable, we were able to gain some information on possible values for those parameters. Indeed, a mathematical model with a negative transmission rate $\tau(t)$ cannot be consistent with the real phenomenon. Therefore, parameter values which produce such negative transmission rates cannot be compatible with the data. Using this argument, we found that the average incubation period cannot exceed eight days. The actual value of the upper bound is highly variable across countries and epidemic waves. We report the values of the upper bound in Section 11 of the supplementary material.

3.3. Instantaneous reproduction number computed for COVID-19 data

Our analysis allows us to compute the instantaneous transmission rate $\tau(t)$. We use this transmission rate to compute two different indicators of the epidemiological dynamics for each geographic area, the instantaneous reproduction number and the quasi-instantaneous reproduction number. Both coincide with the basic reproduction number R_0 on the first day of the epidemic. The instantaneous reproduction number at time t, $R_e(t)$, is the basic reproduction number corresponding to an epidemic starting at time t with a constant transmission rate equal to $\tau(t)$ and with an initial population of susceptibles composed of $S(t)$ individuals (the number of susceptible individuals remaining in the population). The quasi-instantaneous reproduction number at time t, $R_e^0(t)$, is the basic reproduction number corresponding to an epidemic starting at time t with a constant transmission rate equal to $\tau(t)$ and with an initial population of susceptibles composed of S_0 individuals (the number of susceptible individuals at the start of the epidemic). The two indicators are represented for each geographic area in the top row of Figure 3 (black curve: instantaneous reproduction number; green curve: quasi-instantaneous reproduction number).

There is one interpretation for $R_e(t)$ and another for $R_e^0(t)$. The instantaneous reproduction number indicates if, given the current state of the population, the epidemic tends to persist or die out in the long term (note that our model assumes that recovered individuals are perfectly immunized). The quasi-instantaneous reproduction number indicates if the epidemic tends to persist or die out in the long term, provided the number of susceptible is the total population. In other words, we forget about the immunity already obtained by recovered individuals. Also, it is directly proportional to the transmission rate and therefore allows monitoring of its changes. Note that the value of $R_e^0(t)$ changed drastically between epidemic phases, revealing that $\tau(t)$ is far from constant. In any case, the difference between the two values starts to be visible in the figures one year after the start of the epidemic.

We also computed the reproduction number by using the method described in Cori et al. [69], which we denote $R_e^c(t)$. The precise implementation is described in the supplementary material. It is plotted in the bottom row of Figure 3 (green curve), along with the instantaneous reproduction number $R_e(t)$ (green curve).

Remark 3.1. *In the bottom of Figure 3, we compare the instantaneous reproduction numbers obtained by our method in black and the classical method of Cori et al. [69] in green. We observe that the two approaches are not the same at the beginning. This is because the method of Cori et al. [69] does not take into account the initial values I_0 and E_0 while we do. Indeed the method of Cori et al. [69] assumes that I_0 and E_0 are close to 0 at the beginning when it is viewed as a Volterra equation reformulation of the Bernoulli–Kermack–McKendrick model with the age of infection. Our method, on the other hand, does not require such an assumption since it provides a way to compute the initial states I_0 and E_0.*

Remark 3.2. *It is essential to "regularize" the data to obtain a comprehensive outcome from SIR epidemic models. In general, the rate of transmission in the SIR model (applying identification methods) is not very noisy and meaningless. For example, at the beginning of the first epidemic wave, the transmission rate should be decreasing since peoples tend to have less and less contact while to epidemic growth. The standard regularization methods (like, for example, the rolling weekly average method) have been tested for COVID-19 data in Demongeot, Griette, and Magal [36]. The outcome in terms of transmission rate is very noisy and even negative transmission (which is impossible).*

Regularizing the data is not an easy task, and the method used is very important in order to obtain a meaningful outcome for the models. Here, we tried several approaches to link an epidemic phase to the next endemic phase. So far, this regularization procedure is the best one.
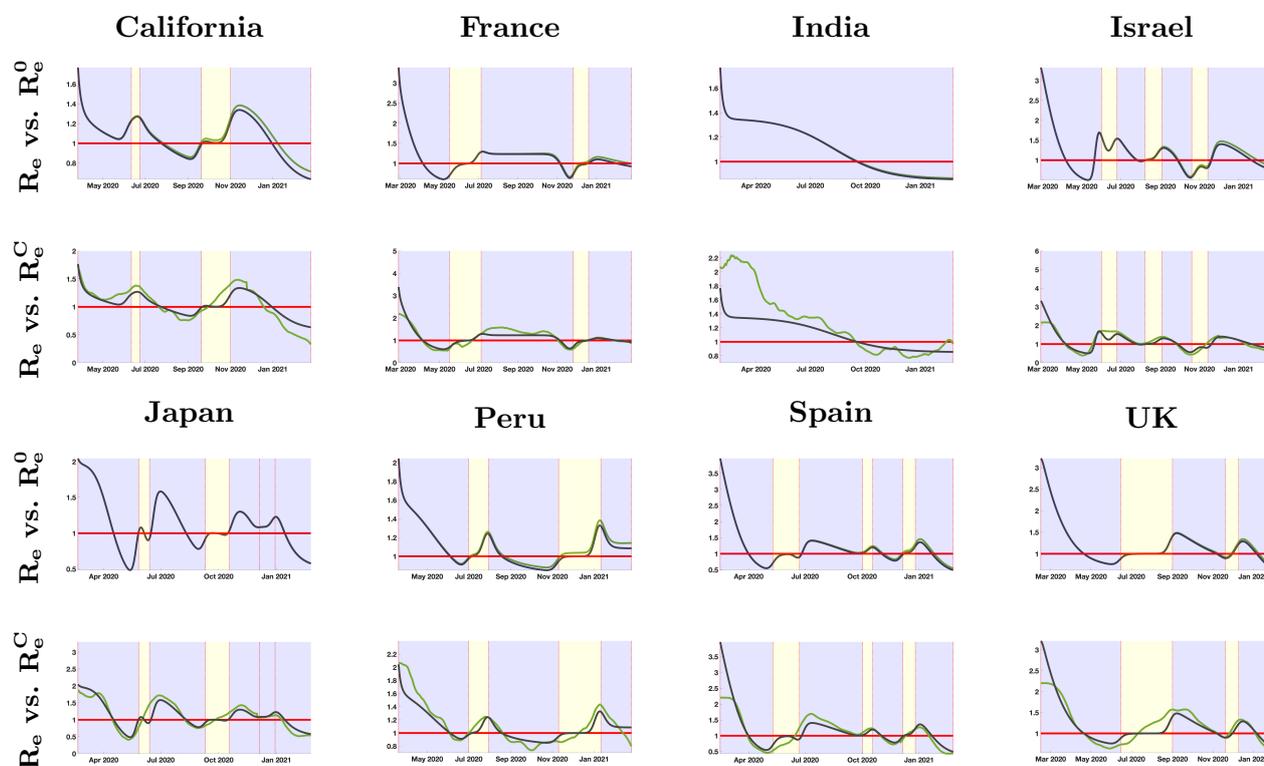
Figure 3. In the top rows, we plot the instantaneous reproduction number $R_e(t)$ (in black) and the quasi instantaneous reproduction number $R_e^0(t)$ (in green). In the bottom rows, we plot the instantaneous reproduction number $R_e(t)$ (in black) and the one obtained by the standard method [69, 70] $R_e^c(t)$ (in green).

3.4. Consequences for vaccination

It is essential to "regularize" the data to obtain a comprehensive outcome from SIR epidemic models. In general, the rate of transmission in the SIR model (applying identification methods) is not very noisy and meaningless. For example, at the beginning of the first epidemic wave, the transmission rate should be decreasing since peoples tend to have less and less contact while to epidemic growth. The standard regularization methods (like, for example, the rolling weekly average method) have been tested for COVID-19 data in Demongeot, Griette, and Magal [36]. The outcome in terms of transmission rate is very noisy and even negative transmission (which is impossible). Regularizing the data is not an easy task, and the method used is critical in order to obtain a meaningful outcome for the models. Here, we tried several approaches to link an epidemic phase to the next endemic phase. So far, this regularization procedure is the best one we tested.

4. Discussion

In this article, we presented a new phenomenological model to describe cumulative reported cases data. This model allows us to handle multiple epidemic waves and fits the data for the eight geographic areas considered very well. The use of Bernoulli-Verhulst curves to fit an epidemic wave is not necessary. We expect that a number of different phenomenological models could be employed for the same purpose; however, our method has the advantage of involving a limited number of parameters. Moreover, the Bernoulli-Verhulst model leads to an explicit algebraic formula for the compatibility conditions of non-identifiable parameters. It is far from obvious that the same computations can be carried out with other models. Our method also provides a very smooth curve with controlled upper bound for the first (four) derivatives, and we use the regularity obtained to compute the transmission rate. We refer to Demongeot, Griette, and Magal [36] for several examples of problems that may occur when using other methods to regularize the data (rolling weekly average, etc.).

The first goal of the article was to understand how to connect successive epidemic waves. As far as we know, this is new compared to the existing literature. A succession of epidemic waves separated by a short period of time with random transmissions is regularly observed in the COVID-19 epidemic data. But several consecutive epidemic phases may happen without endemic transition. An illustration of this situation is provided by the case of Japan, where the parameters of the Bernoulli-Verhulst model changed three times during the last epidemic phases (without endemic interruption). Therefore we subdivide this last epidemic wave into three epidemic phases.

Another advantage of our method is the connection with an epidemiological model. Our study provides a way to explain the data by using a single epidemic model with a time-dependent transmission rate. More precisely, we find that there exists precisely one model that matches the best fit to the data. The fact that the transmission rate corresponding to the data is not constant is, therefore, meaningful. This means that the depletion of susceptible hosts due to natural epidemiological dynamics is not sufficient to explain the reduction in the epidemic spread. Indeed, due to the social changes involving the distancing between individuals, the transmission rate should vary to take into account the changes in the number of contacts per unit of time. The variations in the observed dynamics of the number of cases mainly result from the modification of people's behavior. In other words, the social changes in the population have a stronger impact on the propagation of the disease than the pure epidemiological dynamics. By computing the transmission rate and the associated reproduction numbers, we propose a new method to quantify those social changes. Other factors may also influence the dynamics of the COVID-19 outbreak (temperature, humidity, etc.) and should be taken into account. However, the correlation between the dates of the waves and the mitigation measures imposed by local governments suggests that the former phenomenon takes a more significant role in the epidemiological dynamics.

Precisely because it involves an epidemiological model, our method provides an alternative, robust way to compute indicators for the future behavior of the epidemic: the instantaneous and quasi-instantaneous reproductive numbers $R_e(t)$ and $R_e^0(t)$. It is natural to compare them to an alternative in the literature, sometimes called "effective reproductive number". The method of Cori et al. [69] is a popular framework to estimate its value. Compared with this standard method, our indicators perform better near the beginning of the epidemic and close to the last data point and are

less variable in time. That we require an *a priori* definition of epidemic waves can be considered as an advantage and a drawback. It is a drawback because the computed value of the indicator may slightly depend on the choice of the dates of the epidemic waves. On the other hand, this flexibility also allows testing different scenarios for the future evolution of the epidemic. Thanks to the explicit formula for $R_e(t)$ expressed in function of the parameters, we can also explore the dependency to the parameters (see supplementary material Section S6).

It appears from our results that the instantaneous reproduction number in almost every geographic area considered is less than 3.5. Therefore, an efficient policy to eliminate the COVID-19 would be to vaccinate a fraction of $75 - 80\%$ of the population. Once this threshold is reached, the situation should go back to normal in all the geographic areas considered in this study. This proportion can even be reduced at the expense of partially maintaining the social distancing and the other anti-COVID measures for a sufficiently long period of time.

With a few modifications, our method could also include several other features. It is likely, for instance, that the vaccination of a large part of the population has an impact on the epidemiological dynamics, and this impact is not taken into account for the time being. Different distributions of serial intervals could be taken into account by replacing the mathematical model of ordinary differential equations with integral equations. What we have shown is that the coupling of a phenomenological model to describe the data, with an epidemiological model to take into account the nature of the underlying phenomenon, should provide us with a new, untapped source of information on the epidemic.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. WHO, Disease Outbreak News. Pneumonia of unknown cause–China, January 5, 2020. Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON229 (Accessed on 30 Jun 2021).

2. WHO, Disease Outbreak News. Novel coronavirus–China, January 12, 2020. Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON233 (Accessed on 30 Jun 2021).

3. M. L. Holshue, C. DeBolt, S. Lindquist, K. H. Lofy, J. Wiesman, H. Bruce, et al, First case of 2019 novel Coronavirus in the United States, *New Engl. J. Med.*, **382** (2020), 929–936. doi: 10.1056/NEJMoa2001191.

4. Data from WHO. Available from: https://covid19.who.int/WHO-COVID-19-global-data.csv (Accessed on 30 Jun 2021).

5. E. Anzolin, A. Amante, First Italian dies of coronavirus as outbreak flares in north, *Reuters*, 2020.

6. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, *Lancet Infect. Dis.*, **20** (2020), 533–534. doi: 10.1007/s00285-021-01630-1.

7. Z. Liu, P. Magal, O. Seydi, G. F. Webb, Understanding unreported cases in the COVID-19 epidemic outbreak in Wuhan, China, and the importance of major public health interventions, *Biology*, **9** (2020), 50. doi: 10.3390/biology9030050.

8. A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, et al., Early dynamics of transmission and control of COVID-19: a mathematical modelling study, *Lancet Infect. Dis.*, **20** (2020), 553–558. doi: 10.1016/S1473-3099(20)30144-4.

9. G. T. Patrick, C. W. Walker, W. Oliver, The global impact of COVID-19 and strategies for mitigation and suppression, *Imperial Rep.*, 2020.

10. C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, et al., Transmission of 2019-nCoV infection from an asymptomatic contact in Germany, *New Engl. J. Med.*, **382** (2020), 970–971. doi : 10.1056/NEJMc2001468.

11. K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020, *Eurosurveillance*, **25** (2020), 2000180. doi: 10.2807/1560-7917.ES.2020.25.10.2000180.

12. O. Bylicki, N. Paleiron, F. Janvier, An Outbreak of Covid-19 on an Aircraft Carrier, *New Engl. J. Med.*, **384** (2021), 976–977. doi: 10.1056/NEJMoa2019375.

13. J. Qiu, Covert coronavirus infections could be seeding new outbreaks, *Nature (Lond.)*, 2020. doi: 10.1038/d41586-020-00822-x.

14. Z. Liu, P. Magal, O. Seydi, G. F. Webb, Predicting the cumulative number of cases for the COVID-19 epidemic in China from early data, *Math. Biosci. Eng.*, **17** (2020), 3040–3051. doi: 10.3934/mbe.2020172.

15. Z. Liu, P. Magal, O. Seydi, G. F. Webb, A COVID-19 epidemic model with latency period, *Infect. Dis. Model.*, **5** (2020), 323–337. doi: 10.1016/j.idm.2020.03.003.

16. Z. Liu, P. Magal, O. Seydi, G. F. Webb, A model to predict COVID-19 epidemics with applications to South Korea, Italy, and Spain, *SIAM News*, **53** (2020).

17. H. Nishiura, T. Kobayashi, Y. Yang, K. Hayashi, T. Miyama, R. Kinoshita, et al., The rate of underascertainment of novel coronavirus (2019-nCoV) infection: estimation using Japanese passengers data on evacuation flights, *J. Clin. Med.*, 2020. doi: 10.3390/jcm9020419.

18. R. Omori, K. Mizumoto, H. Nishiura, Ascertainment rate of novel coronavirus disease (COVID-19) in Japan, *Int. J. Infect. Dis.*, **96** (2020), 673–675. doi: 10.1016/j.ijid.2020.04.080.

19. Y. M. Bar-On, A. Flamholz, R. Phillips, R. Milo, Science Forum: SARS-CoV-2 (COVID-19) by the numbers, *eLife*, **9** (2020), e57309. doi: 10.7554/eLife.57309.

20. Z. Liu, P. Magal, G. F. Webb, Predicting the number of reported and unreported cases for the COVID-19 epidemics in China, South Korea, Italy, France, Germany and United Kingdom, *J. Theor. Biol.*, **509** (2021), 110501. doi: 10.1016/j.jtbi.2020.110501.

21. Q. Griette, P. Magal, O. Seydi, Unreported cases for age dependent COVID-19 outbreak in Japan, *Biology*, **9** (2020), 132. doi: 10.3390/biology9060132.

22. G. Zhou, G. Yan, Severe acute respiratory syndrome epidemic in Asia, *Emerg. Infect. Dis.*, **9** (2003), 1608–1610. doi: 10.3201/eid0912.030382.

23. Y. Hsieh, G. Yan, H. Chang, J. Lee, SARS epidemiology modeling, *Emerg. Infect. Dis.*, **10** (2004), 1165–1167. doi: 10.3201/eid1006.031023.

24. X. S. Wang, J. Wu, Y. Yang, Richards model revisited: validation by and application to infection dynamics, *J. Theoret. Biol.*, **313** (20212), 12–19. doi: 10.1016/j.jtbi.2012.07.024.

25. Y. H. Hsieh, Richards model: a simple procedure for real-time prediction of outbreak severity, in *Modeling and dynamics of infectious diseases*, (2009), 216–236.

26. A. Attanayake, S. Perera, S. Jayasinghe, Phenomenological Modelling of COVID-19 Epidemics in Sri Lanka, Italy, the United States, and Hebei Province of China, *Comp. Math. Meth. Med.*, (2020), 6397063. doi:10.1155/2020/6397063.

27. M. Castro, S. Ares, J. Cuesta, S. Manrubia, The turning point and end of an expanding epidemic cannot be precisely forecast, *Proc. Natl. Acad. Sci.*, **117** (2020), 26190–26196. doi: 10.1073/pnas.2007868117.

28. A. M. Ćmiel, B. Ćmiel, A simple method to describe the COVID-19 trajectory and dynamics in any country based on Johnson cumulative density function fitting, *Sci. Rep.*, **11** (2021), 1–10. doi: 10.1038/s41598-021-97285-5.

29. D. Faranda, I. Castillo, O. Hulme, A. Jezequel, J. S. W. Lamb, Y. Sato, et al, Asymptotic estimates of SARS-CoV-2 infection counts and their sensitivity to stochastic perturbation *Chaos*, **30** (2020), 051107. doi: 10.1063/5.0008834.

30. J. Mazurek, The evaluation of COVID-19 prediction precision with a Lyapunov-like exponent, *Plos One*, **16** (2021), e0252394. doi: 10.1371/journal.pone.0252394.

31. L. Palatella, F. Vanni, D. Lambert, A phenomenological estimate of the true scale of CoViD-19 from primary data, *Chaos Solitons Fractals*, **146** (2021), 110854. doi: 10.1016/j.chaos.2021.110854.

32. K. Roosa, Y. Lee, R. Luo, A. Kirpich, R. Rothenberg, J. M. Hyman, et al., Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020, *Infect. Dis. Model.*, **5** (2020), 256–263. doi: 10.1016/j.idm.2020.02.002.

33. R. Singh, M. Rani, A. Bhagavathula, R. Sah, A. Rodriguez-Morales, H. Kalita, et al., Prediction of the COVID-19 pandemic for the top 15 affected countries: advanced autoregressive integrated moving average (ARIMA) model, *JMIR Publ. Health Surv.*, **6** (2020), e19115. doi: 10.2196/19115.

34. P. Wang, X. Zheng, J. Li, B. Zhu, Prediction of epidemic trends in COVID-19 with logistic model and machine learning technics, *Chaos Solitons Fractals*, **139** (2020), 110058. doi: 10.1016/j.chaos.2020.110058.

35. Y. Zou, S. Pan, P. Zhao, L. Han, X. Wang, L. Hemerik ,et al., Outbreak analysis with a logistic growth model shows COVID-19 suppression dynamics in China, *PLoS One*, **15** (2020), e0235247. doi: 10.1371/journal.pone.0235247.

36. J. Demongeot, Q. Griette, P. Magal, SI epidemic model applied to COVID-19 data in mainland China, *R. Soc. Open Sci.*, **7** (2020), 201878. doi: 10.1098/rsos.201878.

37. P. Magal, S. Ruan, Susceptible-infectious-recovered models revisited: from the individual level to the population level, *Math. Biosci.*, **250** (2014), 26–40. doi: 10.1016/j.mbs.2014.02.001.

38. J. Demongeot, Y. Flet-Berliac, H. Seligmann, Temperature decreases spread parameters of the new Covid-19 case dynamics, *Biology*, **9** (2020), 94. doi: 10.3390/biology9050094.

39. Y. Wang, B. Li, R. Gouripeddi, J. C. Facelli, Human activity pattern implications for modeling SARS-CoV-2 transmission, *Comp. Meth. Prog. Biomed.*, **199** (2021), 105896. doi: 10.1016/j.cmpb.2020.105896.

40. G. Chowell, P. Diaz-Dueñas, J. C. Miller, A. Alcazar-Velazco, J. M. Hyman, P. W. Fenimore, et al., Estimation of the reproduction number of dengue fever from spatial epidemic data, *Math. Biosci.*, **208** (2007), 571–589. doi: 10.1016/j.mbs.2006.11.011.

41. X. Huo, J. Chen, S. Ruan, Estimating asymptomatic, undetected and total cases for the COVID-19 outbreak in Wuhan: a mathematical modeling study, *BMC Infect. Dis.*, **21** (2021), 1–18. doi: 10.1186/s12879-021-06078-8.

42. G. Dimarco, B. Perthame, G. Toscani, M. Zanella, Kinetic models for epidemic dynamics with social heterogeneity, *J. Math. Biol.*, **83** (2021), 1–32. doi: 10.1007/s00285-021-01630-1.

43. W. P. London, J. A. Yorke, Recurrent outbreaks of measles, chickenpox and mumps: I. Seasonal variation in contact rates, *Am. J. Epidemiol.*, **98** (1973), 453–468. doi: 10.1093/oxfordjournals.aje.a121575.

44. J. A. Yorke, W. P. London, Recurrent outbreaks of measles, chickenpox and mumps: II. Systematic differences in contact rates and stochastic effects, *Am. J. Epidemiol.*, **98** (1973), 469–482. doi: 10.1093/oxfordjournals.aje.a121576.

45. A. Bakhta, T. Boiveau, Y. Maday, O. Mula, Epidemiological forecasting with model reduction of compartmental models. Application to the COVID-19 pandemic, *Biology*, **10**, 2021. doi: 10.3390/biology10010022.

46. Q. Griette, J. Demongeot, P. Magal, A robust phenomenological approach to investigate COVID-19 data for France, *Math. Appl. Sci. Eng.*, 2021. doi: 10.5206/mase/14031.

47. H. P. Gavin, The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems, *Dep. Civ. Environ. Eng.*, Duke University, (2019), 1-19.

48. The COVID Tracking Project at *The Atlantic*, June 30, 2021. Available from: https://covidtracking.com/.

49. D. Bernoulli, Essai d'une nouvelle analyse de la petite Vérole, & des avantages de l'Inoculation pour la prévenir, *Mémoire Académie Royale des Sciences, Paris*, 1760.

50. P. F. Verhulst, Notice sur la loi que la population poursuit dans son accroissement, *Corresp. Mathématique Phys.*, (1938),113–121.

51. K. Dietz, J. A. P. Heesterbeek, Daniel Bernoulli's epidemiological model revisited, *Math. Biosci.*, **180** (2002), 1–21. doi: 10.1016/S0025-5564(02)00122-0.

52. D. Bernoulli, S. Blower, An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it, *Rev. Med. Virol.*, **14** (2004), 275. doi: 10.1002/rmv.443.

53. B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, et al., Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, *J. Clin. Med.*, **9** (2020), 462. doi: 10.3390/jcm9020462.

54. J. T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P. M. de Salazar, et al., Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China, *Nat. Med.*, **26** (2020), 506–510. doi: 10.1038/s41591-020-0822-7.

55. R. M. Anderson, R. M. May. *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1992.

56. N. T. J. Bailey, *The Mathematical Theory of Epidemics*, Hafner Publishing Co., New York, 1957.

57. F. Brauer, C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, Springer, New York, $2^{n}d$ edition, 2012.

58. F. Brauer, C. Castillo-Chavez, Z. Feng, *Mathematical Models in Epidemiology*, Springer, New York, 2019.

59. F. Brauer, P. van den Driessche, J. Wu, *Mathematical Epidemiology*, Springer, Berlin, Germany, 2008.

60. S. Busenberg, K. Cooke, *Vertically Transmitted Diseases*, Springer-Verlag, Berlin, 1993.

61. O. Diekmann, H. Heesterbeek, T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*, Princeton University Press, Princeton, NJ, 2013.

62. H. W. Hethcote, The mathematics of infectious diseases, *SIAM Rev.*, **42** (2000), 599–653. doi: 10.1137/S0036144500371907.

63. M. J. Keeling, P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, Princeton, NJ, 2008.

64. J. D. Murray, Mathematical Biology, *Biomathematics*, Springer-Verlag, Berlin, 1989.

65. H. R. Thieme, *Mathematics in Population Biology*, Princeton University Press, Princeton, NJ, 2003.

66. H. Kawasuji, Y. Takegoshi, M. Kaneda, A. Ueno, Y. Miyajima, K. Kawago, et al., Transmissibility of COVID-19 depends on the viral load around onset in adult and symptomatic patients, *PLoS ONE*, (2020), e0243597. doi: 10.1371/journal.pone.0243597.

67. S. E. Kim, H. S. Jeong, Y. Yu, S. U. Shin, S. Kim, T. H. Oh, et al., Viral kinetics of SARS-CoV-2 in asymptomatic carriers and presymptomatic patients, *Int. J. Infect. Dis,*, **95** (2020), 441–443. doi: 10.1016/j.ijid.2020.04.083.

68. K. P. Hadeler, Parameter identification in epidemic models, *Math. Biosci.*, **229** (2011), 185–189. doi: 10.1016/j.mbs.2010.12.004.

69. A. Cori, N. M. Ferguson, C. Fraser, S. Cauchemez, A new framework and software to estimate time-varying reproduction numbers during epidemics, *Am. J. Epidemiol.*, **178** (2013), 1505–1512. doi: 10.1093/aje/kwt133.

70. J. Scire, S. A. Nadeau, T. Vaughan, B. Gavin, S. Fuchs, J. Sommer, et al., Reproductive number of the COVID-19 epidemic in Switzerland with a focus on the Cantons of Basel-Stadt and Basel-Landschaft, *Swiss Med. Wkly.*, **150** (2020), w20271. doi: 10.4414/smw.2020.20271.

71. O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28** (1990), 365–382. doi: 10.1007/BF00178324.

72. P. van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48. doi: 10.1016/S0025-5564(02)00108-6.

Supplementary Material

S1. Table of estimated parameters for the phenomenological model

S1.1. California

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Mar 26, 2020 - Jun 11, 2020	$N_0 = 7.34 \times 10^3$	fitted	$N_0 \in [4.16 \times 10^3, 1.05 \times 10^4]$
	$N_{base} = 1.14 \times 10^{-5}$	fitted	$N_{base} \in [-4.33 \times 10^3, 4.33 \times 10^3]$
	$N_\infty = 3.24 \times 10^5$	fitted	$N_\infty \in [2.52 \times 10^5, 3.96 \times 10^5]$
	$\chi = 4.14 \times 10^4$	fitted	$\chi \in [7.74 \times 10^2, 8.20 \times 10^4]$
	$\theta = 4.62 \times 10^{-7}$	fitted	$\theta \in [2.39 \times 10^{-8}, 9.00 \times 10^{-7}]$
Period 2: Endemic phase Jun 11, 2020 - Jun 23, 2020	$a = 3.81 \times 10^3$	computed	
	$N_0 = 1.36 \times 10^5$	computed	
Period 3: Epidemic phase Jun 23, 2020 - Sep 20, 2020	$N_0 = 1.57 \times 10^5$	fitted	$N_0 \in [5.96 \times 10^4, 2.55 \times 10^5]$
	$N_{base} = 2.45 \times 10^4$	fitted	$N_{base} \in [-7.58 \times 10^4, 1.25 \times 10^5]$
	$N_\infty = 8.22 \times 10^5$	fitted	$N_\infty \in [7.36 \times 10^5, 9.08 \times 10^5]$
	$\chi = 5.54 \times 10^{-2}$	fitted	$\chi \in [5.32 \times 10^{-3}, 1.05 \times 10^{-1}]$
	$\theta = 7.18 \times 10^{-1}$	fitted	$\theta \in [-3.59 \times 10^{-2}, 1.47]$
Period 4: Endemic phase Sep 20, 2020 - Nov 01, 2020	$a = 3.66 \times 10^3$	computed	
	$N_0 = 7.76 \times 10^5$	computed	
Period 5: Epidemic phase Nov 01, 2020 - Feb 25, 2021	$N_0 = 6.27 \times 10^4$	fitted	$N_0 \in [4.95 \times 10^4, 7.59 \times 10^4]$
	$N_{base} = 8.67 \times 10^5$	fitted	$N_{base} \in [8.45 \times 10^5, 8.88 \times 10^5]$
	$N_\infty = 2.66 \times 10^6$	fitted	$N_\infty \in [2.64 \times 10^6, 2.67 \times 10^6]$
	$\chi = 6.36 \times 10^{-2}$	fitted	$\chi \in [5.73 \times 10^{-2}, 6.98 \times 10^{-2}]$
	$\theta = 1.02$	fitted	$\theta \in [8.79 \times 10^{-1}, 1.16]$

Table S1. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in California from January 03 2020 to February 25 2021.

S1.2. France

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 27, 2020 - May 17, 2020	$N_0 = 3.61 \times 10^{-4}$	fitted	$N_0 \in [-3.77, 3.77]$
	$N_{base} = 0.00$	fixed	
	$N_\infty = 1.43 \times 10^5$	fitted	$N_\infty \in [-1.58 \times 10^4, 3.01 \times 10^5]$
	$\chi = 1.17 \times 10^2$	fitted	$\chi \in [-1.09 \times 10^7, 1.09 \times 10^7]$
	$\theta = 7.29 \times 10^{-4}$	fitted	$\theta \in [-6.84 \times 10^1, 6.84 \times 10^1]$
Period 2: Endemic phase May 17, 2020 - Jul 05, 2020	$a = 3.14 \times 10^2$	computed	
	$N_0 = 1.39 \times 10^5$	computed	
Period 3: Epidemic phase Jul 05, 2020 - Nov 26, 2020	$N_0 = 1.50 \times 10^4$	fitted	$N_0 \in [1.36 \times 10^4, 1.65 \times 10^4]$
	$N_{base} = 1.40 \times 10^5$	fitted	$N_{base} \in [1.33 \times 10^5, 1.46 \times 10^5]$
	$N_\infty = 1.99 \times 10^6$	fitted	$N_\infty \in [1.97 \times 10^6, 2.01 \times 10^6]$
	$\chi = 3.68 \times 10^{-2}$	fitted	$\chi \in [3.60 \times 10^{-2}, 3.76 \times 10^{-2}]$
	$\theta = 6.55$	fitted	$\theta \in [5.52, 7.58]$
Period 4: Endemic phase Nov 26, 2020 - Dec 20, 2020	$a = 1.28 \times 10^4$	computed	
	$N_0 = 2.11 \times 10^6$	computed	
Period 5: Epidemic phase Dec 20, 2020 - Feb 25, 2021	$N_0 = 2.73 \times 10^5$	fitted	$N_0 \in [-2.43 \times 10^3, 5.48 \times 10^5]$
	$N_{base} = 2.15 \times 10^6$	fitted	$N_{base} \in [1.86 \times 10^6, 2.43 \times 10^6]$
	$N_\infty = 2.13 \times 10^6$	fitted	$N_\infty \in [1.88 \times 10^6, 2.39 \times 10^6]$
	$\chi = 5.88 \times 10^{-2}$	fitted	$\chi \in [-6.11 \times 10^{-2}, 1.79 \times 10^{-1}]$
	$\theta = 5.47 \times 10^{-1}$	fitted	$\theta \in [-9.19 \times 10^{-1}, 2.01]$

Table S2. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in France from January 03 2020 to February 25 2021.

S1.3. ***India***

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 01, 2020 - Feb 25, 2021	$N_0 = 5.83 \times 10^2$	fitted	$N_0 \in [3.45 \times 10^2, 8.20 \times 10^2]$
	$N_{base} = 1.97 \times 10^4$	fitted	$N_{base} \in [5.36 \times 10^3, 3.39 \times 10^4]$
	$N_\infty = 1.10 \times 10^7$	fitted	$N_\infty \in [1.10 \times 10^7, 1.11 \times 10^7]$
	$\chi = 4.89 \times 10^{-2}$	fitted	$\chi \in [4.59 \times 10^{-2}, 5.20 \times 10^{-2}]$
	$\theta = 5.12 \times 10^{-1}$	fitted	$\theta \in [4.71 \times 10^{-1}, 5.54 \times 10^{-1}]$

Table S3. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in India from January 03 2020 to February 25 2021.

S1.4. Israel

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 27, 2020 - Jun 01, 2020	$N_0 = 1.08 \times 10^{-2}$	fitted	$N_0 \in [-3.85 \times 10^{-2}, 6.02 \times 10^{-2}]$
	$N_{base} = 4.27 \times 10^1$	fitted	$N_{base} \in [-3.36 \times 10^1, 1.19 \times 10^2]$
	$N_\infty = 1.71 \times 10^4$	fitted	$N_\infty \in [1.70 \times 10^4, 1.72 \times 10^4]$
	$\chi = 9.18 \times 10^{-1}$	fitted	$\chi \in [1.71 \times 10^{-1}, 1.67]$
	$\theta = 1.05 \times 10^{-1}$	fitted	$\theta \in [1.55 \times 10^{-2}, 1.94 \times 10^{-1}]$
Period 2: Endemic phase Jun 01, 2020 - Jun 25, 2020	$a = 2.04 \times 10^2$	computed	
	$N_0 = 1.70 \times 10^4$	computed	
Period 3: Epidemic phase Jun 25, 2020 - Aug 08, 2020	$N_0 = 2.48 \times 10^3$	fitted	$N_0 \in [3.43 \times 10^2, 4.61 \times 10^3]$
	$N_{base} = 1.95 \times 10^4$	fitted	$N_{base} \in [1.70 \times 10^4, 2.20 \times 10^4]$
	$N_\infty = 8.66 \times 10^4$	fitted	$N_\infty \in [7.78 \times 10^4, 9.55 \times 10^4]$
	$\chi = 2.93 \times 10^{-1}$	fitted	$\chi \in [-2.61 \times 10^{-1}, 8.48 \times 10^{-1}]$
	$\theta = 2.04 \times 10^{-1}$	fitted	$\theta \in [-2.43 \times 10^{-1}, 6.50 \times 10^{-1}]$
Period 4: Endemic phase Aug 08, 2020 - Sep 03, 2020	$a = 1.54 \times 10^3$	computed	
	$N_0 = 7.97 \times 10^4$	computed	
Period 5: Epidemic phase Sep 03, 2020 - Oct 20, 2020	$N_0 = 4.59 \times 10^4$	fitted	$N_0 \in [2.88 \times 10^4, 6.31 \times 10^4]$
	$N_{base} = 7.38 \times 10^4$	fitted	$N_{base} \in [5.53 \times 10^4, 9.23 \times 10^4]$
	$N_\infty = 2.35 \times 10^5$	fitted	$N_\infty \in [2.19 \times 10^5, 2.52 \times 10^5]$
	$\chi = 5.05 \times 10^{-2}$	fitted	$\chi \in [3.77 \times 10^{-2}, 6.34 \times 10^{-2}]$
	$\theta = 3.45$	fitted	$\theta \in [1.96, 4.93]$
Period 6: Endemic phase Oct 20, 2020 - Nov 14, 2020	$a = 8.90 \times 10^2$	computed	
	$N_0 = 3.04 \times 10^5$	computed	
Period 7: Epidemic phase Nov 14, 2020 - Feb 25, 2021	$N_0 = 3.16 \times 10^3$	fitted	$N_0 \in [2.16 \times 10^3, 4.17 \times 10^3]$
	$N_{base} = 3.23 \times 10^5$	fitted	$N_{base} \in [3.21 \times 10^5, 3.25 \times 10^5]$
	$N_\infty = 4.87 \times 10^5$	fitted	$N_\infty \in [4.79 \times 10^5, 4.95 \times 10^5]$
	$\chi = 8.28 \times 10^{-2}$	fitted	$\chi \in [7.22 \times 10^{-2}, 9.34 \times 10^{-2}]$
	$\theta = 7.06 \times 10^{-1}$	fitted	$\theta \in [5.69 \times 10^{-1}, 8.43 \times 10^{-1}]$

Table S4. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in Israel from January 03 2020 to February 25 2021.

S1.5. Japan

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 20, 2020 - May 27, 2020	$N_0 = 5.83$	fitted	$N_0 \in [1.91, 9.74]$
	$N_{base} = 3.25 \times 10^2$	fitted	$N_{base} \in [2.55 \times 10^2, 3.95 \times 10^2]$
	$N_\infty = 1.63 \times 10^4$	fitted	$N_\infty \in [1.62 \times 10^4, 1.64 \times 10^4]$
	$\chi = 1.48 \times 10^{-1}$	fitted	$\chi \in [1.30 \times 10^{-1}, 1.65 \times 10^{-1}]$
	$\theta = 8.29 \times 10^{-1}$	fitted	$\theta \in [6.88 \times 10^{-1}, 9.70 \times 10^{-1}]$
Period 2: Endemic phase May 27, 2020 - Jun 13, 2020	$a = 7.07 \times 10^1$	computed	
	$N_0 = 1.65 \times 10^4$	computed	
Period 3: Epidemic phase Jun 13, 2020 - Sep 10, 2020	$N_0 = 1.49 \times 10^2$	fitted	$N_0 \in [8.52 \times 10^1, 2.13 \times 10^2]$
	$N_{base} = 1.75 \times 10^4$	fitted	$N_{base} \in [1.73 \times 10^4, 1.78 \times 10^4]$
	$N_\infty = 6.02 \times 10^4$	fitted	$N_\infty \in [5.93 \times 10^4, 6.10 \times 10^4]$
	$\chi = 1.19 \times 10^{-1}$	fitted	$\chi \in [1.03 \times 10^{-1}, 1.35 \times 10^{-1}]$
	$\theta = 6.28 \times 10^{-1}$	fitted	$\theta \in [5.04 \times 10^{-1}, 7.52 \times 10^{-1}]$
Period 4: Endemic phase Sep 10, 2020 - Oct 18, 2020	$a = 5.36 \times 10^2$	computed	
	$N_0 = 7.27 \times 10^4$	computed	
Period 5: Epidemic phase Oct 18, 2020 - Dec 05, 2020	$N_0 = 6.33 \times 10^3$	fitted	$N_0 \in [4.64 \times 10^3, 8.01 \times 10^3]$
	$N_{base} = 8.68 \times 10^4$	fitted	$N_{base} \in [8.48 \times 10^4, 8.88 \times 10^4]$
	$N_\infty = 9.10 \times 10^4$	fitted	$N_\infty \in [7.75 \times 10^4, 1.05 \times 10^5]$
	$\chi = 5.60 \times 10^{-2}$	fitted	$\chi \in [4.74 \times 10^{-2}, 6.46 \times 10^{-2}]$
	$\theta = 2.58$	fitted	$\theta \in [1.00, 4.16]$
Period 6: Epidemic phase Dec 05, 2020 - Dec 30, 2020	$N_0 = 1.23 \times 10^5$	fitted	$N_0 \in [-2.43 \times 10^5, 4.90 \times 10^5]$
	$N_{base} = 3.43 \times 10^4$	fitted	$N_{base} \in [-3.33 \times 10^5, 4.01 \times 10^5]$
	$N_\infty = 3.49 \times 10^5$	fitted	$N_\infty \in [-2.92 \times 10^7, 2.99 \times 10^7]$
	$\chi = 1.78 \times 10^{-2}$	fitted	$\chi \in [-3.59 \times 10^{-2}, 7.15 \times 10^{-2}]$
	$\theta = 7.84$	fitted	$\theta \in [-1.28 \times 10^3, 1.30 \times 10^3]$
Period 7: Epidemic phase Dec 30, 2020 - Feb 25, 2021	$N_0 = 2.00 \times 10^4$	fitted	$N_0 \in [1.59 \times 10^3, 3.84 \times 10^4]$
	$N_{base} = 2.05 \times 10^5$	fitted	$N_{base} \in [1.85 \times 10^5, 2.25 \times 10^5]$
	$N_\infty = 2.29 \times 10^5$	fitted	$N_\infty \in [2.11 \times 10^5, 2.47 \times 10^5]$
	$\chi = 7.98 \times 10^{-1}$	fitted	$\chi \in [-2.54, 4.13]$
	$\theta = 9.61 \times 10^{-2}$	fitted	$\theta \in [-3.15 \times 10^{-1}, 5.07 \times 10^{-1}]$

Table S5. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in Japan from January 03 2020 to February 25 2021.

S1.6. Peru

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Mar 20, 2020 - Jul 01, 2020	$N_0 = 8.36 \times 10^2$	fitted	$N_0 \in [2.63 \times 10^2, 1.41 \times 10^3]$
	$N_{base} = 3.00 \times 10^{-5}$	fitted	$N_{base} \in [-1.74 \times 10^3, 1.74 \times 10^3]$
	$N_\infty = 3.61 \times 10^5$	fitted	$N_\infty \in [3.44 \times 10^5, 3.79 \times 10^5]$
	$\chi = 1.08 \times 10^{-1}$	fitted	$\chi \in [7.59 \times 10^{-2}, 1.41 \times 10^{-1}]$
	$\theta = 4.20 \times 10^{-1}$	fitted	$\theta \in [2.41 \times 10^{-1}, 5.98 \times 10^{-1}]$
Period 2: Endemic phase Jul 01, 2020 - Jul 30, 2020	$a = 3.67 \times 10^3$	computed	
	$N_0 = 2.83 \times 10^5$	computed	
Period 3: Epidemic phase Jul 30, 2020 - Nov 10, 2020	$N_0 = 1.86 \times 10^5$	fitted	$N_0 \in [-2.61 \times 10^4, 3.98 \times 10^5]$
	$N_{base} = 2.03 \times 10^5$	fitted	$N_{base} \in [-1.11 \times 10^4, 4.18 \times 10^5]$
	$N_\infty = 7.69 \times 10^5$	fitted	$N_\infty \in [5.65 \times 10^5, 9.72 \times 10^5]$
	$\chi = 4.84 \times 10^{-1}$	fitted	$\chi \in [-6.23, 7.20]$
	$\theta = 5.95 \times 10^{-2}$	fitted	$\theta \in [-7.74 \times 10^{-1}, 8.93 \times 10^{-1}]$
Period 4: Endemic phase Nov 10, 2020 - Jan 11, 2021	$a = 1.80 \times 10^3$	computed	
	$N_0 = 9.16 \times 10^5$	computed	
Period 5: Epidemic phase Jan 11, 2021 - Feb 25, 2021	$N_0 = 3.23 \times 10^5$	fitted	
	$N_{base} = 7.04 \times 10^5$	fitted	
	$N_\infty = 7.00 \times 10^6$	fitted	
	$\chi = 1.36 \times 10^{-2}$	fitted	
	$\theta = 3.67 \times 10^1$	fitted	

Table S6. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in Peru from January 03 2020 to February 25 2021.

S1.7. Spain

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 15, 2020 - May 10, 2020	$N_0 = 5.19 \times 10^{-4}$	fitted	$N_0 \in [-5.00 \times 10^{-3}, 6.04 \times 10^{-3}]$
	$N_{base} = 5.77 \times 10^2$	fitted	$N_{base} \in [-4.50 \times 10^2, 1.60 \times 10^3]$
	$N_\infty = 2.32 \times 10^5$	fitted	$N_\infty \in [2.30 \times 10^5, 2.34 \times 10^5]$
	$\chi = 9.80 \times 10^{-1}$	fitted	$\chi \in [-1.26 \times 10^{-1}, 2.09]$
	$\theta = 9.75 \times 10^{-2}$	fitted	$\theta \in [-1.83 \times 10^{-2}, 2.13 \times 10^{-1}]$
Period 2: Endemic phase May 10, 2020 - Jun 22, 2020	$a = 5.67 \times 10^2$	computed	
	$N_0 = 2.28 \times 10^5$	computed	
Period 3: Epidemic phase Jun 22, 2020 - Oct 02, 2020	$N_0 = 2.38 \times 10^3$	fitted	$N_0 \in [1.39 \times 10^3, 3.36 \times 10^3]$
	$N_{base} = 2.50 \times 10^5$	fitted	$N_{base} \in [2.48 \times 10^5, 2.53 \times 10^5]$
	$N_\infty = 9.89 \times 10^5$	fitted	$N_\infty \in [9.02 \times 10^5, 1.08 \times 10^6]$
	$\chi = 9.29 \times 10^{-2}$	fitted	$\chi \in [7.07 \times 10^{-2}, 1.15 \times 10^{-1}]$
	$\theta = 3.84 \times 10^{-1}$	fitted	$\theta \in [2.38 \times 10^{-1}, 5.29 \times 10^{-1}]$
Period 4: Endemic phase Oct 02, 2020 - Oct 18, 2020	$a = 1.09 \times 10^4$	computed	
	$N_0 = 8.14 \times 10^5$	computed	
Period 5: Epidemic phase Oct 18, 2020 - Dec 06, 2020	$N_0 = 1.68 \times 10^5$	fitted	$N_0 \in [-3.50 \times 10^4, 3.72 \times 10^5]$
	$N_{base} = 8.20 \times 10^5$	fitted	$N_{base} \in [6.12 \times 10^5, 1.03 \times 10^6]$
	$N_\infty = 9.85 \times 10^5$	fitted	$N_\infty \in [8.01 \times 10^5, 1.17 \times 10^6]$
	$\chi = 3.15 \times 10^{-1}$	fitted	$\chi \in [-1.05, 1.68]$
	$\theta = 2.02 \times 10^{-1}$	fitted	$\theta \in [-7.15 \times 10^{-1}, 1.12]$
Period 6: Endemic phase Dec 06, 2020 - Dec 26, 2020	$a = 9.15 \times 10^3$	computed	
	$N_0 = 1.72 \times 10^6$	computed	
Period 7: Epidemic phase Dec 26, 2020 - Feb 25, 2021	$N_0 = 5.94 \times 10^4$	fitted	$N_0 \in [3.86 \times 10^4, 8.02 \times 10^4]$
	$N_{base} = 1.84 \times 10^6$	fitted	$N_{base} \in [1.81 \times 10^6, 1.87 \times 10^6]$
	$N_\infty = 1.30 \times 10^6$	fitted	$N_\infty \in [1.28 \times 10^6, 1.32 \times 10^6]$
	$\chi = 1.30 \times 10^{-1}$	fitted	$\chi \in [9.90 \times 10^{-2}, 1.60 \times 10^{-1}]$
	$\theta = 7.84 \times 10^{-1}$	fitted	$\theta \in [5.50 \times 10^{-1}, 1.02]$

Table S7. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in Spain from January 03 2020 to February 01 2021.

S1.8. United Kingdom

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 15, 2020 - Jun 15, 2020	$N_0 = 2.65 \times 10^{-2}$	fitted	$N_0 \in [-8.82 \times 10^{-2}, 1.41 \times 10^{-1}]$
	$N_{base} = 1.12 \times 10^2$	fitted	$N_{base} \in [-4.82 \times 10^2, 7.06 \times 10^2]$
	$N_\infty = 2.86 \times 10^5$	fitted	$N_\infty \in [2.84 \times 10^5, 2.88 \times 10^5]$
	$\chi = 1.76$	fitted	$\chi \in [-1.46, 4.98]$
	$\theta = 2.76 \times 10^{-2}$	fitted	$\theta \in [-2.38 \times 10^{-2}, 7.90 \times 10^{-2}]$
Period 2: Endemic phase Jun 15, 2020 - Sep 01, 2020	$a = 9.43 \times 10^2$	computed	
	$N_0 = 2.70 \times 10^5$	computed	
Period 3: Epidemic phase Sep 01, 2020 - Nov 20, 2020	$N_0 = 7.85 \times 10^3$	fitted	$N_0 \in [3.63 \times 10^3, 1.21 \times 10^4]$
	$N_{base} = 3.36 \times 10^5$	fitted	$N_{base} \in [3.28 \times 10^5, 3.43 \times 10^5]$
	$N_\infty = 2.14 \times 10^6$	fitted	$N_\infty \in [1.93 \times 10^6, 2.36 \times 10^6]$
	$\chi = 2.41 \times 10^{-1}$	fitted	$\chi \in [2.16 \times 10^{-2}, 4.60 \times 10^{-1}]$
	$\theta = 1.32 \times 10^{-1}$	fitted	$\theta \in [-9.25 \times 10^{-3}, 2.74 \times 10^{-1}]$
Period 4: Endemic phase Nov 20, 2020 - Dec 10, 2020	$a = 1.61 \times 10^4$	computed	
	$N_0 = 1.48 \times 10^6$	computed	
Period 5: Epidemic phase Dec 10, 2020 - Feb 01, 2021	$N_0 = 2.26 \times 10^5$	fitted	$N_0 \in [1.16 \times 10^5, 3.35 \times 10^5]$
	$N_{base} = 1.58 \times 10^6$	fitted	$N_{base} \in [1.46 \times 10^6, 1.70 \times 10^6]$
	$N_\infty = 2.42 \times 10^6$	fitted	$N_\infty \in [2.34 \times 10^6, 2.51 \times 10^6]$
	$\chi = 8.57 \times 10^{-2}$	fitted	$\chi \in [5.14 \times 10^{-2}, 1.20 \times 10^{-1}]$
	$\theta = 1.08$	fitted	$\theta \in [4.85 \times 10^{-1}, 1.68]$

Table S8. In this table we list the parameters of the phenomenological model which gives the best fit to the cumulative number of cases data in United Kingdom from January 03 2020 to February 01 2021.

S2. Plot of the multiple Bernoulli–Verhulst models fitted to each epidemic phase

In Figure S1, we present the details of the fit of the Bernoulli–Verhulst models to the successive epidemic waves in the 8 geographic areas considered. Each epidemic wave is associated with a different color.
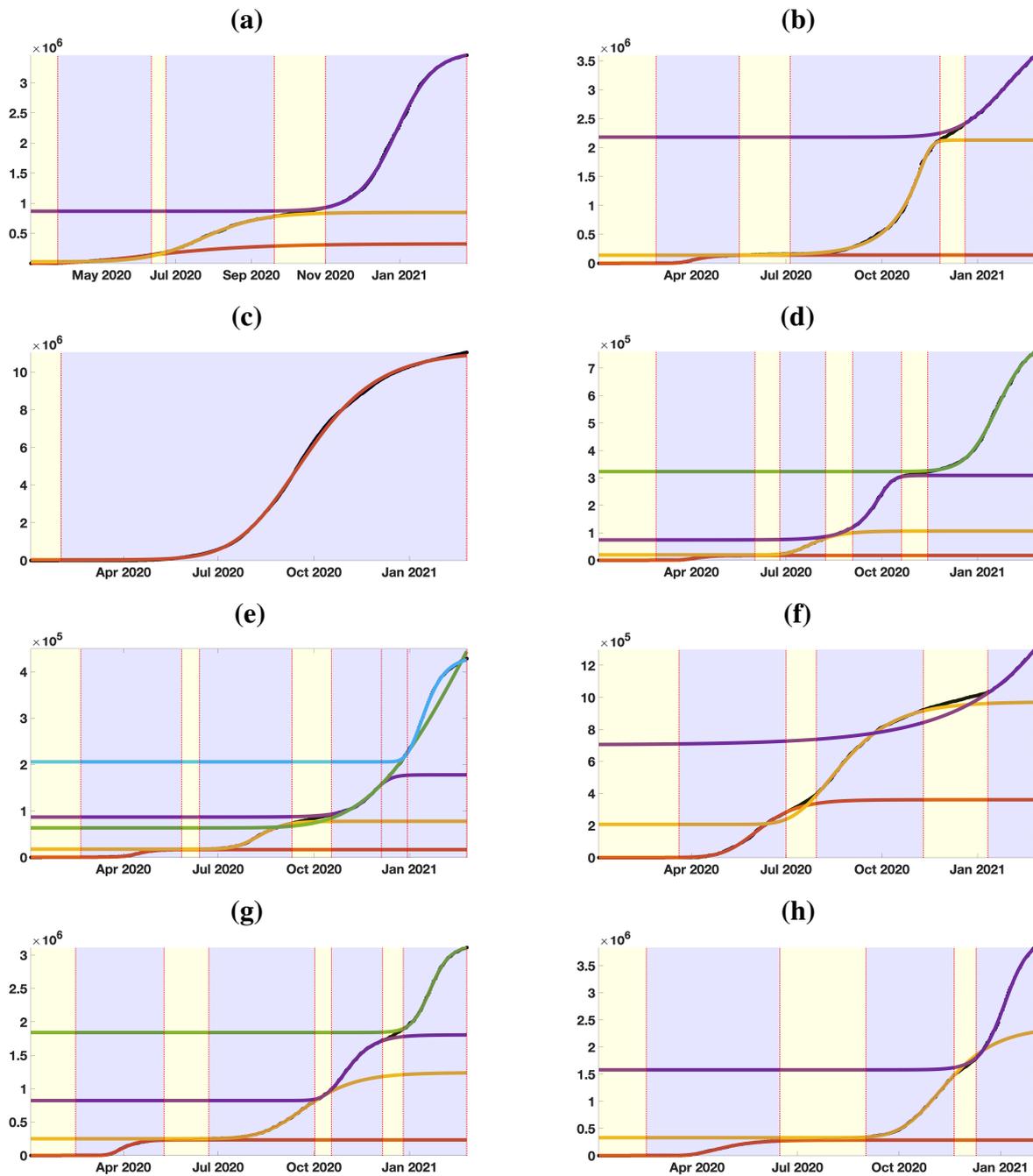
Figure S1. In this figure, we plot the cumulative number of cases (black dots) and the best fit of Bernoulli–Verhulst for each epidemic wave for (a) California; (b) France; (c) India; (d) Israel; (e) Japan; (f) Peru; (g) Spain; (h) United Kingdom.

S3. Relative error of the fitted curve compared to the data in each geographic area

S3.1. California
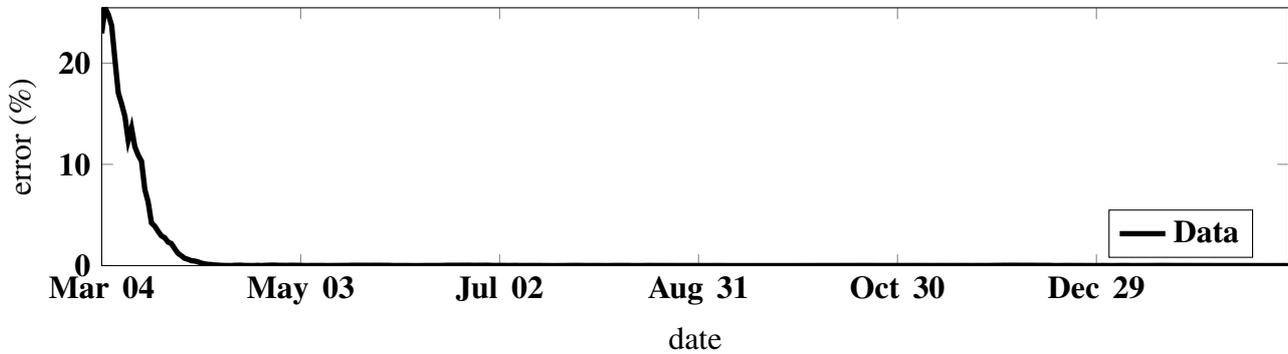
Figure S2. Relative error between the data and the model for California State, expressed in percent.

S3.2. France
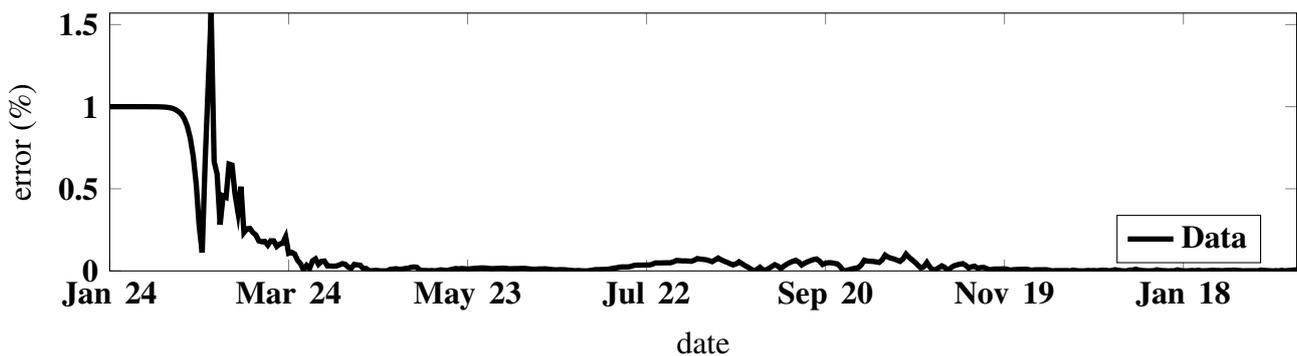
Figure S3. Relative error between the data and the model for France, expressed in percent.

S3.3. India
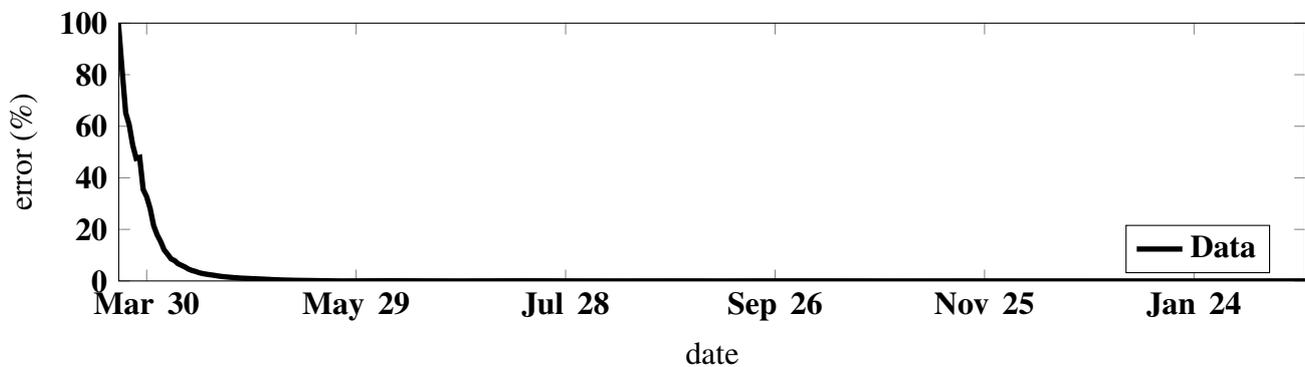
Figure S4. Relative error between the data and the model for India, expressed in percent.

S3.4. *Israel*
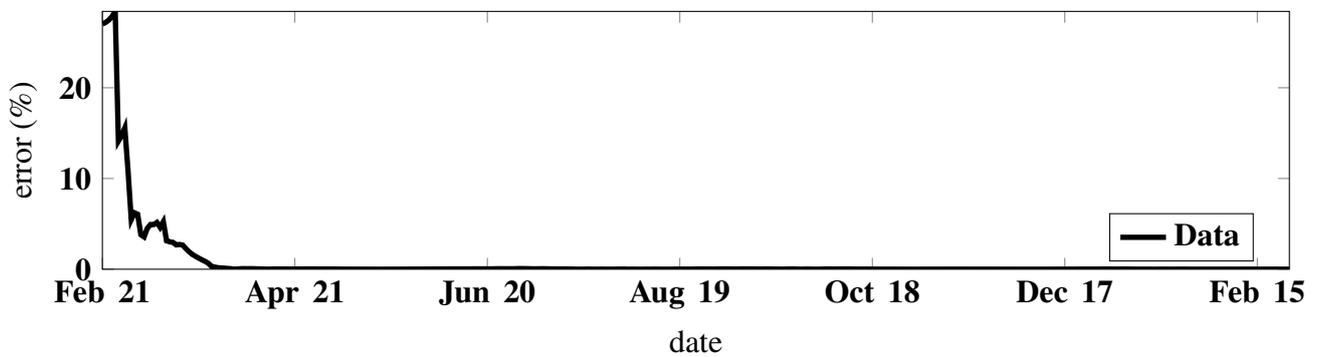
Figure S5. Relative error between the data and the model for Israel, expressed in percent.

S3.5. *Japan*
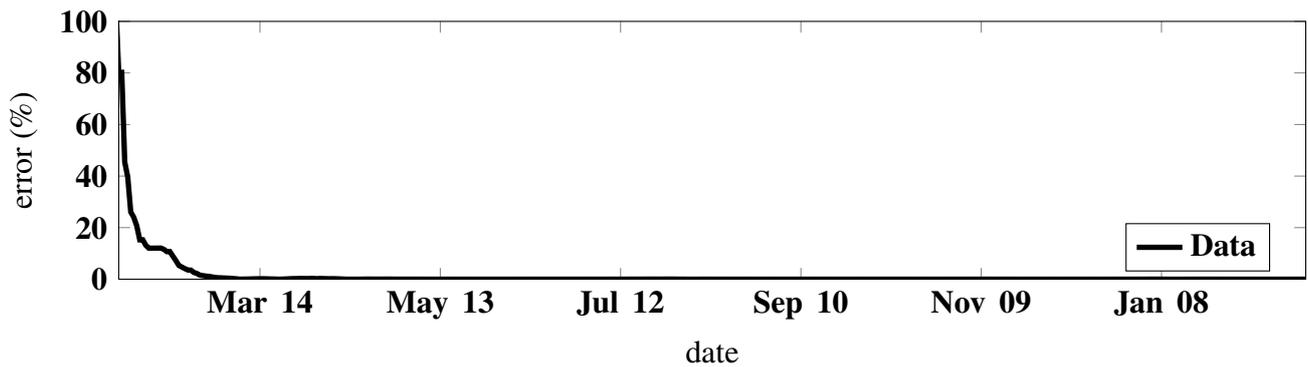
Figure S6. Relative error between the data and the model for Japan, expressed in percent.

S3.6. *Peru*
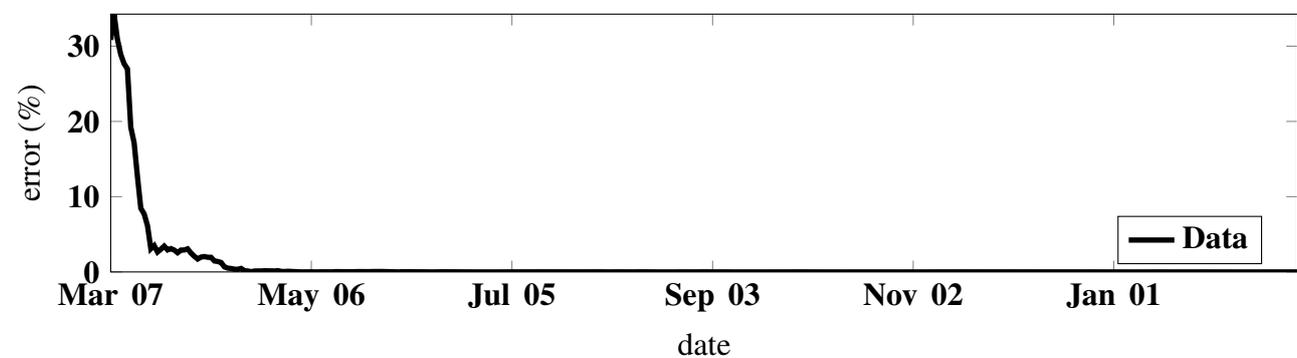
Figure S7. Relative error between the data and the model for Peru, expressed in percent.

S3.7. Spain
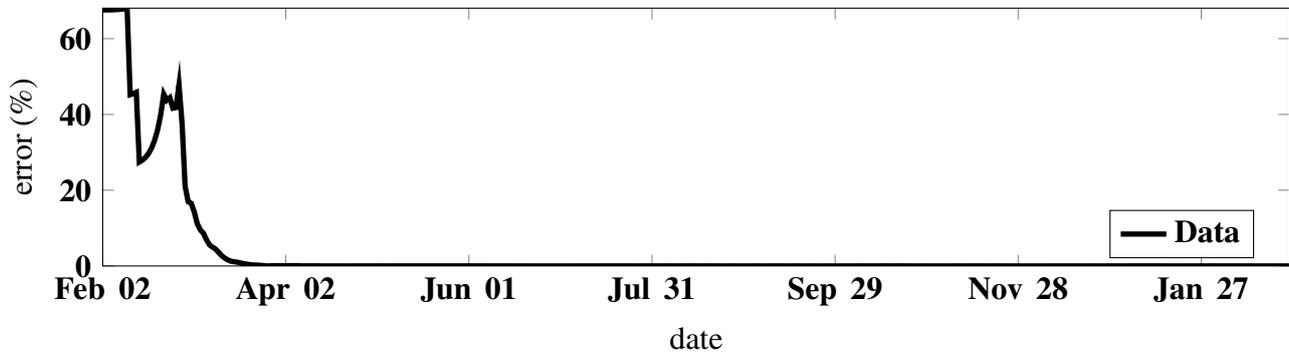
Figure S8. Relative error between the data and the model for Spain, expressed in percent.

S3.8. United Kingdom
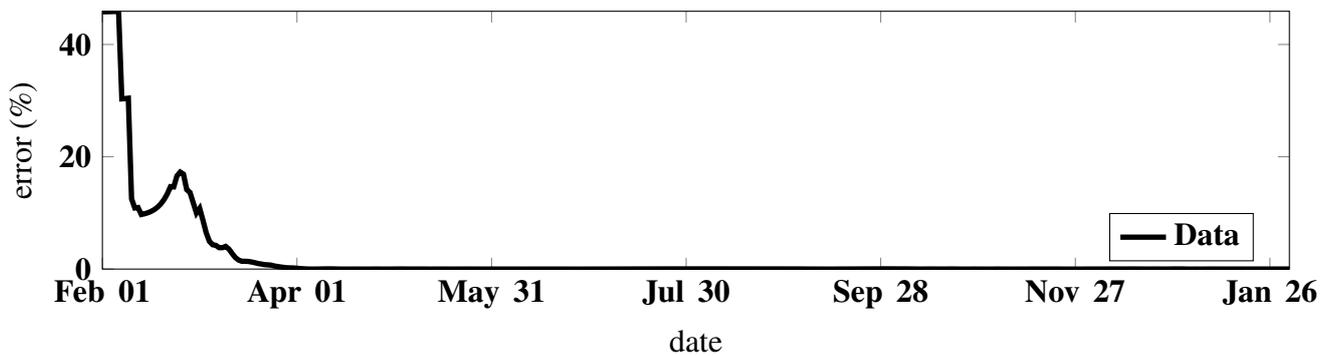
Figure S9. Relative error between the data and the model for UK, expressed in percent.

S4. Table of estimated parameters for the phenomenological model

S4.1. California

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Mar 26, 2020 - Jun 11, 2020	$N_0 = 7.34 \times 10^3$	fitted	$N_0 \in [4.16 \times 10^3, 1.05 \times 10^4]$
	$N_{base} = 1.14 \times 10^{-5}$	fitted	$N_{base} \in [-4.33 \times 10^3, 4.33 \times 10^3]$
	$N_\infty = 3.24 \times 10^5$	fitted	$N_\infty \in [2.52 \times 10^5, 3.96 \times 10^5]$
	$\chi = 4.14 \times 10^4$	fitted	$\chi \in [7.74 \times 10^2, 8.20 \times 10^4]$
	$\theta = 4.62 \times 10^{-7}$	fitted	$\theta \in [2.39 \times 10^{-8}, 9.00 \times 10^{-7}]$
Period 2: Endemic phase Jun 11, 2020 - Jun 23, 2020	$a = 3.81 \times 10^3$	computed	
	$N_0 = 1.36 \times 10^5$	computed	
Period 3: Epidemic phase Jun 23, 2020 - Sep 20, 2020	$N_0 = 1.57 \times 10^5$	fitted	$N_0 \in [5.96 \times 10^4, 2.55 \times 10^5]$
	$N_{base} = 2.45 \times 10^4$	fitted	$N_{base} \in [-7.58 \times 10^4, 1.25 \times 10^5]$
	$N_\infty = 8.22 \times 10^5$	fitted	$N_\infty \in [7.36 \times 10^5, 9.08 \times 10^5]$
	$\chi = 5.54 \times 10^{-2}$	fitted	$\chi \in [5.32 \times 10^{-3}, 1.05 \times 10^{-1}]$
	$\theta = 7.18 \times 10^{-1}$	fitted	$\theta \in [-3.59 \times 10^{-2}, 1.47]$
Period 4: Endemic phase Sep 20, 2020 - Nov 01, 2020	$a = 3.66 \times 10^3$	computed	
	$N_0 = 7.76 \times 10^5$	computed	
Period 5: Epidemic phase Nov 01, 2020 - Feb 25, 2021	$N_0 = 6.27 \times 10^4$	fitted	$N_0 \in [4.95 \times 10^4, 7.59 \times 10^4]$
	$N_{base} = 8.67 \times 10^5$	fitted	$N_{base} \in [8.45 \times 10^5, 8.88 \times 10^5]$
	$N_\infty = 2.66 \times 10^6$	fitted	$N_\infty \in [2.64 \times 10^6, 2.67 \times 10^6]$
	$\chi = 6.36 \times 10^{-2}$	fitted	$\chi \in [5.73 \times 10^{-2}, 6.98 \times 10^{-2}]$
	$\theta = 1.02$	fitted	$\theta \in [8.79 \times 10^{-1}, 1.16]$

Table S9. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in California from January 03 2020 to February 25 2021.

S4.2. France

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 27, 2020 - May 17, 2020	$N_0 = 3.61 \times 10^{-4}$	fitted	$N_0 \in [-3.77, 3.77]$
	$N_{base} = 0.00$	fixed	
	$N_\infty = 1.43 \times 10^5$	fitted	$N_\infty \in [-1.58 \times 10^4, 3.01 \times 10^5]$
	$\chi = 1.17 \times 10^2$	fitted	$\chi \in [-1.09 \times 10^7, 1.09 \times 10^7]$
	$\theta = 7.29 \times 10^{-4}$	fitted	$\theta \in [-6.84 \times 10^1, 6.84 \times 10^1]$
Period 2: Endemic phase May 17, 2020 - Jul 05, 2020	$a = 3.14 \times 10^2$	computed	
	$N_0 = 1.39 \times 10^5$	computed	
Period 3: Epidemic phase Jul 05, 2020 - Nov 26, 2020	$N_0 = 1.50 \times 10^4$	fitted	$N_0 \in [1.36 \times 10^4, 1.65 \times 10^4]$
	$N_{base} = 1.40 \times 10^5$	fitted	$N_{base} \in [1.33 \times 10^5, 1.46 \times 10^5]$
	$N_\infty = 1.99 \times 10^6$	fitted	$N_\infty \in [1.97 \times 10^6, 2.01 \times 10^6]$
	$\chi = 3.68 \times 10^{-2}$	fitted	$\chi \in [3.60 \times 10^{-2}, 3.76 \times 10^{-2}]$
	$\theta = 6.55$	fitted	$\theta \in [5.52, 7.58]$
Period 4: Endemic phase Nov 26, 2020 - Dec 20, 2020	$a = 1.28 \times 10^4$	computed	
	$N_0 = 2.11 \times 10^6$	computed	
Period 5: Epidemic phase Dec 20, 2020 - Feb 25, 2021	$N_0 = 2.73 \times 10^5$	fitted	$N_0 \in [-2.43 \times 10^3, 5.48 \times 10^5]$
	$N_{base} = 2.15 \times 10^6$	fitted	$N_{base} \in [1.86 \times 10^6, 2.43 \times 10^6]$
	$N_\infty = 2.13 \times 10^6$	fitted	$N_\infty \in [1.88 \times 10^6, 2.39 \times 10^6]$
	$\chi = 5.88 \times 10^{-2}$	fitted	$\chi \in [-6.11 \times 10^{-2}, 1.79 \times 10^{-1}]$
	$\theta = 5.47 \times 10^{-1}$	fitted	$\theta \in [-9.19 \times 10^{-1}, 2.01]$

Table S10. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in France from January 03 2020 to February 25 2021.

S4.3. India

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 01, 2020 - Feb 25, 2021	$N_0 = 5.83 \times 10^2$	fitted	$N_0 \in [3.45 \times 10^2, 8.20 \times 10^2]$
	$N_{base} = 1.97 \times 10^4$	fitted	$N_{base} \in [5.36 \times 10^3, 3.39 \times 10^4]$
	$N_\infty = 1.10 \times 10^7$	fitted	$N_\infty \in [1.10 \times 10^7, 1.11 \times 10^7]$
	$\chi = 4.89 \times 10^{-2}$	fitted	$\chi \in [4.59 \times 10^{-2}, 5.20 \times 10^{-2}]$
	$\theta = 5.12 \times 10^{-1}$	fitted	$\theta \in [4.71 \times 10^{-1}, 5.54 \times 10^{-1}]$

Table S11. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in India from January 03 2020 to February 25 2021.

S4.4. Israel

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 27, 2020 - Jun 01, 2020	$N_0 = 1.08 \times 10^{-2}$	fitted	$N_0 \in [-3.85 \times 10^{-2}, 6.02 \times 10^{-2}]$
	$N_{base} = 4.27 \times 10^1$	fitted	$N_{base} \in [-3.36 \times 10^1, 1.19 \times 10^2]$
	$N_\infty = 1.71 \times 10^4$	fitted	$N_\infty \in [1.70 \times 10^4, 1.72 \times 10^4]$
	$\chi = 9.18 \times 10^{-1}$	fitted	$\chi \in [1.71 \times 10^{-1}, 1.67]$
	$\theta = 1.05 \times 10^{-1}$	fitted	$\theta \in [1.55 \times 10^{-2}, 1.94 \times 10^{-1}]$
Period 2: Endemic phase Jun 01, 2020 - Jun 25, 2020	$a = 2.04 \times 10^2$	computed	
	$N_0 = 1.70 \times 10^4$	computed	
Period 3: Epidemic phase Jun 25, 2020 - Aug 08, 2020	$N_0 = 2.48 \times 10^3$	fitted	$N_0 \in [3.43 \times 10^2, 4.61 \times 10^3]$
	$N_{base} = 1.95 \times 10^4$	fitted	$N_{base} \in [1.70 \times 10^4, 2.20 \times 10^4]$
	$N_\infty = 8.66 \times 10^4$	fitted	$N_\infty \in [7.78 \times 10^4, 9.55 \times 10^4]$
	$\chi = 2.93 \times 10^{-1}$	fitted	$\chi \in [-2.61 \times 10^{-1}, 8.48 \times 10^{-1}]$
	$\theta = 2.04 \times 10^{-1}$	fitted	$\theta \in [-2.43 \times 10^{-1}, 6.50 \times 10^{-1}]$
Period 4: Endemic phase Aug 08, 2020 - Sep 03, 2020	$a = 1.54 \times 10^3$	computed	
	$N_0 = 7.97 \times 10^4$	computed	
Period 5: Epidemic phase Sep 03, 2020 - Oct 20, 2020	$N_0 = 4.59 \times 10^4$	fitted	$N_0 \in [2.88 \times 10^4, 6.31 \times 10^4]$
	$N_{base} = 7.38 \times 10^4$	fitted	$N_{base} \in [5.53 \times 10^4, 9.23 \times 10^4]$
	$N_\infty = 2.35 \times 10^5$	fitted	$N_\infty \in [2.19 \times 10^5, 2.52 \times 10^5]$
	$\chi = 5.05 \times 10^{-2}$	fitted	$\chi \in [3.77 \times 10^{-2}, 6.34 \times 10^{-2}]$
	$\theta = 3.45$	fitted	$\theta \in [1.96, 4.93]$
Period 6: Endemic phase Oct 20, 2020 - Nov 14, 2020	$a = 8.90 \times 10^2$	computed	
	$N_0 = 3.04 \times 10^5$	computed	
Period 7: Epidemic phase Nov 14, 2020 - Feb 25, 2021	$N_0 = 3.16 \times 10^3$	fitted	$N_0 \in [2.16 \times 10^3, 4.17 \times 10^3]$
	$N_{base} = 3.23 \times 10^5$	fitted	$N_{base} \in [3.21 \times 10^5, 3.25 \times 10^5]$
	$N_\infty = 4.87 \times 10^5$	fitted	$N_\infty \in [4.79 \times 10^5, 4.95 \times 10^5]$
	$\chi = 8.28 \times 10^{-2}$	fitted	$\chi \in [7.22 \times 10^{-2}, 9.34 \times 10^{-2}]$
	$\theta = 7.06 \times 10^{-1}$	fitted	$\theta \in [5.69 \times 10^{-1}, 8.43 \times 10^{-1}]$

Table S12. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in Israel from January 03 2020 to February 25 2021.

S4.5. Japan

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 20, 2020 - May 27, 2020	$N_0 = 5.83$	fitted	$N_0 \in [1.91, 9.74]$
	$N_{base} = 3.25 \times 10^2$	fitted	$N_{base} \in [2.55 \times 10^2, 3.95 \times 10^2]$
	$N_\infty = 1.63 \times 10^4$	fitted	$N_\infty \in [1.62 \times 10^4, 1.64 \times 10^4]$
	$\chi = 1.48 \times 10^{-1}$	fitted	$\chi \in [1.30 \times 10^{-1}, 1.65 \times 10^{-1}]$
	$\theta = 8.29 \times 10^{-1}$	fitted	$\theta \in [6.88 \times 10^{-1}, 9.70 \times 10^{-1}]$
Period 2: Endemic phase May 27, 2020 - Jun 13, 2020	$a = 7.07 \times 10^1$	computed	
	$N_0 = 1.65 \times 10^4$	computed	
Period 3: Epidemic phase Jun 13, 2020 - Sep 10, 2020	$N_0 = 1.49 \times 10^2$	fitted	$N_0 \in [8.52 \times 10^1, 2.13 \times 10^2]$
	$N_{base} = 1.75 \times 10^4$	fitted	$N_{base} \in [1.73 \times 10^4, 1.78 \times 10^4]$
	$N_\infty = 6.02 \times 10^4$	fitted	$N_\infty \in [5.93 \times 10^4, 6.10 \times 10^4]$
	$\chi = 1.19 \times 10^{-1}$	fitted	$\chi \in [1.03 \times 10^{-1}, 1.35 \times 10^{-1}]$
	$\theta = 6.28 \times 10^{-1}$	fitted	$\theta \in [5.04 \times 10^{-1}, 7.52 \times 10^{-1}]$
Period 4: Endemic phase Sep 10, 2020 - Oct 18, 2020	$a = 5.36 \times 10^2$	computed	
	$N_0 = 7.27 \times 10^4$	computed	
Period 5: Epidemic phase Oct 18, 2020 - Dec 05, 2020	$N_0 = 6.33 \times 10^3$	fitted	$N_0 \in [4.64 \times 10^3, 8.01 \times 10^3]$
	$N_{base} = 8.68 \times 10^4$	fitted	$N_{base} \in [8.48 \times 10^4, 8.88 \times 10^4]$
	$N_\infty = 9.10 \times 10^4$	fitted	$N_\infty \in [7.75 \times 10^4, 1.05 \times 10^5]$
	$\chi = 5.60 \times 10^{-2}$	fitted	$\chi \in [4.74 \times 10^{-2}, 6.46 \times 10^{-2}]$
	$\theta = 2.58$	fitted	$\theta \in [1.00, 4.16]$
Period 6: Epidemic phase Dec 05, 2020 - Dec 30, 2020	$N_0 = 1.23 \times 10^5$	fitted	$N_0 \in [-2.43 \times 10^5, 4.90 \times 10^5]$
	$N_{base} = 3.43 \times 10^4$	fitted	$N_{base} \in [-3.33 \times 10^5, 4.01 \times 10^5]$
	$N_\infty = 3.49 \times 10^5$	fitted	$N_\infty \in [-2.92 \times 10^7, 2.99 \times 10^7]$
	$\chi = 1.78 \times 10^{-2}$	fitted	$\chi \in [-3.59 \times 10^{-2}, 7.15 \times 10^{-2}]$
	$\theta = 7.84$	fitted	$\theta \in [-1.28 \times 10^3, 1.30 \times 10^3]$
Period 7: Epidemic phase Dec 30, 2020 - Feb 25, 2021	$N_0 = 2.00 \times 10^4$	fitted	$N_0 \in [1.59 \times 10^3, 3.84 \times 10^4]$
	$N_{base} = 2.05 \times 10^5$	fitted	$N_{base} \in [1.85 \times 10^5, 2.25 \times 10^5]$
	$N_\infty = 2.29 \times 10^5$	fitted	$N_\infty \in [2.11 \times 10^5, 2.47 \times 10^5]$
	$\chi = 7.98 \times 10^{-1}$	fitted	$\chi \in [-2.54, 4.13]$
	$\theta = 9.61 \times 10^{-2}$	fitted	$\theta \in [-3.15 \times 10^{-1}, 5.07 \times 10^{-1}]$

Table S13. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in Japan from January 03 2020 to February 25 2021.

S4.6. Peru

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Mar 20, 2020 - Jul 01, 2020	$N_0 = 8.36 \times 10^2$	fitted	$N_0 \in [2.63 \times 10^2, 1.41 \times 10^3]$
	$N_{base} = 3.00 \times 10^{-5}$	fitted	$N_{base} \in [-1.74 \times 10^3, 1.74 \times 10^3]$
	$N_\infty = 3.61 \times 10^5$	fitted	$N_\infty \in [3.44 \times 10^5, 3.79 \times 10^5]$
	$\chi = 1.08 \times 10^{-1}$	fitted	$\chi \in [7.59 \times 10^{-2}, 1.41 \times 10^{-1}]$
	$\theta = 4.20 \times 10^{-1}$	fitted	$\theta \in [2.41 \times 10^{-1}, 5.98 \times 10^{-1}]$
Period 2: Endemic phase Jul 01, 2020 - Jul 30, 2020	$a = 3.67 \times 10^3$	computed	
	$N_0 = 2.83 \times 10^5$	computed	
Period 3: Epidemic phase Jul 30, 2020 - Nov 10, 2020	$N_0 = 1.86 \times 10^5$	fitted	$N_0 \in [-2.61 \times 10^4, 3.98 \times 10^5]$
	$N_{base} = 2.03 \times 10^5$	fitted	$N_{base} \in [-1.11 \times 10^4, 4.18 \times 10^5]$
	$N_\infty = 7.69 \times 10^5$	fitted	$N_\infty \in [5.65 \times 10^5, 9.72 \times 10^5]$
	$\chi = 4.84 \times 10^{-1}$	fitted	$\chi \in [-6.23, 7.20]$
	$\theta = 5.95 \times 10^{-2}$	fitted	$\theta \in [-7.74 \times 10^{-1}, 8.93 \times 10^{-1}]$
Period 4: Endemic phase Nov 10, 2020 - Jan 11, 2021	$a = 1.80 \times 10^3$	computed	
	$N_0 = 9.16 \times 10^5$	computed	
Period 5: Epidemic phase Jan 11, 2021 - Feb 25, 2021	$N_0 = 3.23 \times 10^5$	fitted	
	$N_{base} = 7.04 \times 10^5$	fitted	
	$N_\infty = 7.00 \times 10^6$	fitted	
	$\chi = 1.36 \times 10^{-2}$	fitted	
	$\theta = 3.67 \times 10^1$	fitted	

Table S14. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in Peru from January 03 2020 to February 25 2021.

S4.7. Spain

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 15, 2020 - May 10, 2020	$N_0 = 5.19 \times 10^{-4}$	fitted	$N_0 \in [-5.00 \times 10^{-3}, 6.04 \times 10^{-3}]$
	$N_{base} = 5.77 \times 10^2$	fitted	$N_{base} \in [-4.50 \times 10^2, 1.60 \times 10^3]$
	$N_\infty = 2.32 \times 10^5$	fitted	$N_\infty \in [2.30 \times 10^5, 2.34 \times 10^5]$
	$\chi = 9.80 \times 10^{-1}$	fitted	$\chi \in [-1.26 \times 10^{-1}, 2.09]$
	$\theta = 9.75 \times 10^{-2}$	fitted	$\theta \in [-1.83 \times 10^{-2}, 2.13 \times 10^{-1}]$
Period 2: Endemic phase May 10, 2020 - Jun 22, 2020	$a = 5.67 \times 10^2$	computed	
	$N_0 = 2.28 \times 10^5$	computed	
Period 3: Epidemic phase Jun 22, 2020 - Oct 02, 2020	$N_0 = 2.38 \times 10^3$	fitted	$N_0 \in [1.39 \times 10^3, 3.36 \times 10^3]$
	$N_{base} = 2.50 \times 10^5$	fitted	$N_{base} \in [2.48 \times 10^5, 2.53 \times 10^5]$
	$N_\infty = 9.89 \times 10^5$	fitted	$N_\infty \in [9.02 \times 10^5, 1.08 \times 10^6]$
	$\chi = 9.29 \times 10^{-2}$	fitted	$\chi \in [7.07 \times 10^{-2}, 1.15 \times 10^{-1}]$
	$\theta = 3.84 \times 10^{-1}$	fitted	$\theta \in [2.38 \times 10^{-1}, 5.29 \times 10^{-1}]$
Period 4: Endemic phase Oct 02, 2020 - Oct 18, 2020	$a = 1.09 \times 10^4$	computed	
	$N_0 = 8.14 \times 10^5$	computed	
Period 5: Epidemic phase Oct 18, 2020 - Dec 06, 2020	$N_0 = 1.68 \times 10^5$	fitted	$N_0 \in [-3.50 \times 10^4, 3.72 \times 10^5]$
	$N_{base} = 8.20 \times 10^5$	fitted	$N_{base} \in [6.12 \times 10^5, 1.03 \times 10^6]$
	$N_\infty = 9.85 \times 10^5$	fitted	$N_\infty \in [8.01 \times 10^5, 1.17 \times 10^6]$
	$\chi = 3.15 \times 10^{-1}$	fitted	$\chi \in [-1.05, 1.68]$
	$\theta = 2.02 \times 10^{-1}$	fitted	$\theta \in [-7.15 \times 10^{-1}, 1.12]$
Period 6: Endemic phase Dec 06, 2020 - Dec 26, 2020	$a = 9.15 \times 10^3$	computed	
	$N_0 = 1.72 \times 10^6$	computed	
Period 7: Epidemic phase Dec 26, 2020 - Feb 25, 2021	$N_0 = 5.94 \times 10^4$	fitted	$N_0 \in [3.86 \times 10^4, 8.02 \times 10^4]$
	$N_{base} = 1.84 \times 10^6$	fitted	$N_{base} \in [1.81 \times 10^6, 1.87 \times 10^6]$
	$N_\infty = 1.30 \times 10^6$	fitted	$N_\infty \in [1.28 \times 10^6, 1.32 \times 10^6]$
	$\chi = 1.30 \times 10^{-1}$	fitted	$\chi \in [9.90 \times 10^{-2}, 1.60 \times 10^{-1}]$
	$\theta = 7.84 \times 10^{-1}$	fitted	$\theta \in [5.50 \times 10^{-1}, 1.02]$

Table S15. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in Spain from January 03 2020 to February 01 2021.

S4.8. United Kingdom

Period	Parameters value	Method	95% Confidence interval
Period 1: Epidemic phase Feb 15, 2020 - Jun 15, 2020	$N_0 = 2.65 \times 10^{-2}$	fitted	$N_0 \in [-8.82 \times 10^{-2}, 1.41 \times 10^{-1}]$
	$N_{base} = 1.12 \times 10^2$	fitted	$N_{base} \in [-4.82 \times 10^2, 7.06 \times 10^2]$
	$N_\infty = 2.86 \times 10^5$	fitted	$N_\infty \in [2.84 \times 10^5, 2.88 \times 10^5]$
	$\chi = 1.76$	fitted	$\chi \in [-1.46, 4.98]$
	$\theta = 2.76 \times 10^{-2}$	fitted	$\theta \in [-2.38 \times 10^{-2}, 7.90 \times 10^{-2}]$
Period 2: Endemic phase Jun 15, 2020 - Sep 01, 2020	$a = 9.43 \times 10^2$	computed	
	$N_0 = 2.70 \times 10^5$	computed	
Period 3: Epidemic phase Sep 01, 2020 - Nov 20, 2020	$N_0 = 7.85 \times 10^3$	fitted	$N_0 \in [3.63 \times 10^3, 1.21 \times 10^4]$
	$N_{base} = 3.36 \times 10^5$	fitted	$N_{base} \in [3.28 \times 10^5, 3.43 \times 10^5]$
	$N_\infty = 2.14 \times 10^6$	fitted	$N_\infty \in [1.93 \times 10^6, 2.36 \times 10^6]$
	$\chi = 2.41 \times 10^{-1}$	fitted	$\chi \in [2.16 \times 10^{-2}, 4.60 \times 10^{-1}]$
	$\theta = 1.32 \times 10^{-1}$	fitted	$\theta \in [-9.25 \times 10^{-3}, 2.74 \times 10^{-1}]$
Period 4: Endemic phase Nov 20, 2020 - Dec 10, 2020	$a = 1.61 \times 10^4$	computed	
	$N_0 = 1.48 \times 10^6$	computed	
Period 5: Epidemic phase Dec 10, 2020 - Feb 01, 2021	$N_0 = 2.26 \times 10^5$	fitted	$N_0 \in [1.16 \times 10^5, 3.35 \times 10^5]$
	$N_{base} = 1.58 \times 10^6$	fitted	$N_{base} \in [1.46 \times 10^6, 1.70 \times 10^6]$
	$N_\infty = 2.42 \times 10^6$	fitted	$N_\infty \in [2.34 \times 10^6, 2.51 \times 10^6]$
	$\chi = 8.57 \times 10^{-2}$	fitted	$\chi \in [5.14 \times 10^{-2}, 1.20 \times 10^{-1}]$
	$\theta = 1.08$	fitted	$\theta \in [4.85 \times 10^{-1}, 1.68]$

Table S16. In this table we list the values of the parameters of the phenomenological model which give the best fit to the cumulative number of cases data in United Kingdom from January 03 2020 to February 01 2021.

S5. Additional information for the results section

Period	Interpretation	Parameters value	Method
U_0	Number of unreported symptomatic infectious at time t_0	1	Fixed
R_0	Number of reported symptomatic infectious at time t_0	0	Fixed
$\tau(t)$	Transmission rate	Eqs (2.27)–(2.32)	Computed
f	Fraction of reported symptomatic infectious	0.8	Fixed
κ	Fraction of unreported symptomatic infectious capable to transmit the pathogen	1	Fixed
$1/\alpha$	Average duration of the exposed period	1 days	Fixed
$1/\nu$	Average duration of the asymptomatic infectious period	3 days	Fixed
$1/\eta$	Average duration of the symptomatic infectious period	7 days	Fixed

Table S17. In this table we list the values of the parameters of the epidemic model used for the simulations.

S5.1. California

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Mar 26, 2020	Fixed
S_0	Number of susceptibles at time t_0	3.95×10^7	Fixed
E_0	Number of exposed at time t_0	7.91×10^2	Computed
I_0	Number of asymptotic infectious at time t_0	2.06×10^3	Computed

Table S18. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the Californian data for the first, the second and the third epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 5.23 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 8.21 \times 10^{-3}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 2.52 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 2.08 \times 10^{-2}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.54 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 6.72 \times 10^{-2}$

S5.2. France

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Feb 27, 2020	Fixed
S_0	Number of susceptibles at time t_0	6.50×10^7	Fixed
E_0	Number of exposed at time t_0	4.27×10^1	Computed
I_0	Number of asymptotic infectious at time t_0	6.30×10^1	Computed

Table S19. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the French data for the first, the second and the third epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.17 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 2.19 \times 10^{-3}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 4.15$ days	$f \geq \frac{N_\infty}{S_0} = 3.06 \times 10^{-2}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 3.11 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 3.28 \times 10^{-2}$

S5.3. India

Figure 4 of the main text is devoted to the reproduction number of the model. The instantaneous reproduction number $t \to R_e(t)$ is decreasing from February 01, 2020 until February 25, 2021.

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Feb 01, 2020	Fixed
S_0	Number of susceptibles at time t_0	1.39×10^9	Fixed
E_0	Number of exposed at time t_0	4.29×10^1	Computed
I_0	Number of asymptomatic infectious at time t_0	1.12×10^2	Computed

Table S20. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the Indian data for the first single wave, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 3.99 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 7.93 \times 10^{-3}$

S5.4. Israel

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Feb 27, 2020	Fixed
S_0	Number of susceptibles at time t_0	8.74×10^6	Fixed
E_0	Number of exposed at time t_0	4.16	Computed
I_0	Number of asymptomatic infectious at time t_0	6.25	Computed

Table S21. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the Israeli data for the first, the second, the third and the fourth epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.04 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 1.95 \times 10^{-3}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.67 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 9.91 \times 10^{-3}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 5.74$ days	$f \geq \frac{N_\infty}{S_0} = 2.69 \times 10^{-2}$
Fourth epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.71 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 5.57 \times 10^{-2}$

S5.5. Japan

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Feb 20, 2020	Fixed
S_0	Number of susceptibles at time t_0	1.26×10^8	Fixed
E_0	Number of exposed at time t_0	2.61	Computed
I_0	Number of asymptomatic infectious at time t_0	5.45	Computed

Table S22. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the Japanese data for the first, the second, the third, the fourth and the fifth epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 8.18$ days	$f \geq \frac{N_\infty}{S_0} = 1.29 \times 10^{-4}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.34 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 4.77 \times 10^{-4}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 6.92$ days	$f \geq \frac{N_\infty}{S_0} = 7.22 \times 10^{-4}$
Fourth epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 7.17$ days	$f \geq \frac{N_\infty}{S_0} = 2.77 \times 10^{-3}$
Fifth epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.30 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 1.82 \times 10^{-3}$

S5.6. Peru

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Mar 20, 2020	Fixed
S_0	Number of susceptibles at time t_0	3.32×10^7	Fixed
E_0	Number of exposed at time t_0	1.64×10^2	Computed
I_0	Number of asymptomatic infectious at time t_0	3.85×10^2	Computed

Table S23. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the Peruvian data for the first, the second and the third epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 2.20 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 1.09 \times 10^{-2}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 3.47 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 2.32 \times 10^{-2}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 2.01$ days	$f \geq \frac{N_\infty}{S_0} = 2.11 \times 10^{-1}$

S5.7. Spain

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Feb 15, 2020	Fixed
S_0	Number of susceptibles at time t_0	3.95×10^7	Fixed
E_0	Number of exposed at time t_0	5.10	Computed
I_0	Number of asymptomatic infectious at time t_0	6.87	Computed

Table S24. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the Spanish data for the first, the second, the third and the fourth epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.05 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 5.87 \times 10^{-3}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 2.81 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 2.50 \times 10^{-2}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.58 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 2.49 \times 10^{-2}$
Fourth epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 9.84$ days	$f \geq \frac{N_\infty}{S_0} = 3.29 \times 10^{-2}$

S5.8. United Kingdom

Period	Interpretation	Parameters value	Method
t_0	Time at which we started the epidemic model	Feb 15, 2020	Fixed
S_0	Number of susceptibles at time t_0	6.81×10^7	Fixed
E_0	Number of exposed at time t_0	3.41	Computed
I_0	Number of asymptomatic infectious at time t_0	5.15	Computed

Table S25. In this table we list the values of the parameters of the epidemic model used for the simulations.

Compatibility condition between data and epidemic model

By using the data from Great Britain for the first, the second and the third epidemic waves, we get from Eqs (2.59) and (2.60) the following estimates for the average duration of the exposed and asymptomatic infectious periods and the fraction of reported cases

First epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 2.06 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 4.20 \times 10^{-3}$
Second epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 3.14 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 3.15 \times 10^{-2}$
Third epidemic wave	$\frac{1}{\alpha}$ and $\frac{1}{\nu} \leq \frac{1}{\chi\theta} = 1.08 \times 10^1$ days	$f \geq \frac{N_\infty}{S_0} = 3.56 \times 10^{-2}$

S6. Dependency with respect to the parameters for the French data

Influence of f on basic reproduction number:
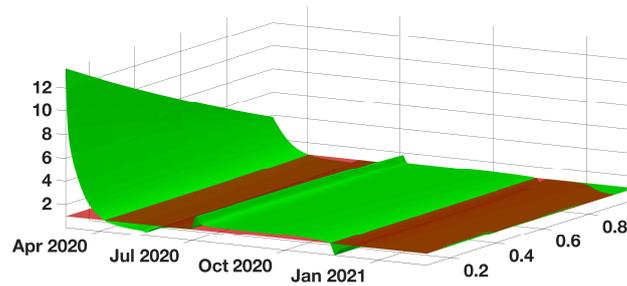
Figure S10. In this figure we plot $(t, f) \rightarrow R_e(t)$ when t varies from January 03 2020 to January 04 2021 and f varies from 0.1 to 1.
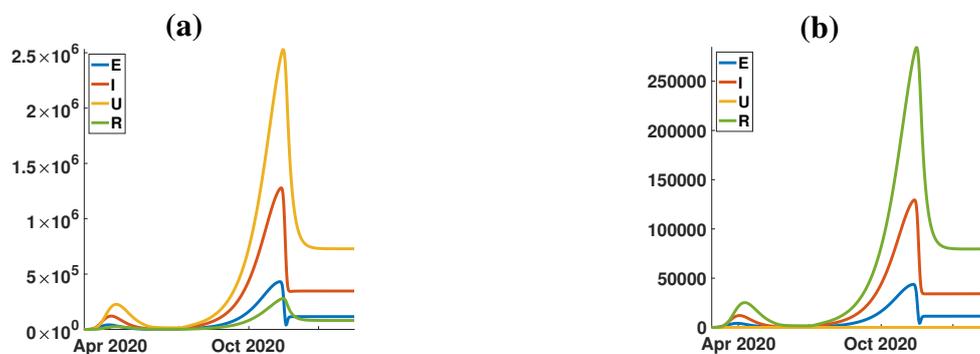
Figure S11. In this figure we explore the influence of the parameter f on the solution of model. The figure (a) corresponds to $f = 0.1$ and figure (b) corresponds to $f = 1$. The remaining parameters are unchanged.

Influence of κ on basic reproduction number:
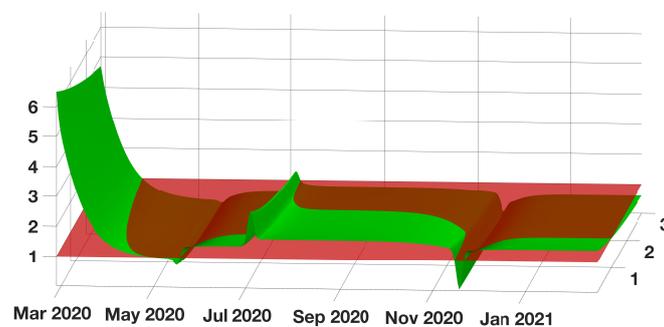
Figure S12. In this figure we plot $(t, \kappa) \rightarrow R_e(t)$ when t varies from January 03 2020 to January 04 2021 and κ varies from 0.1 to 3.
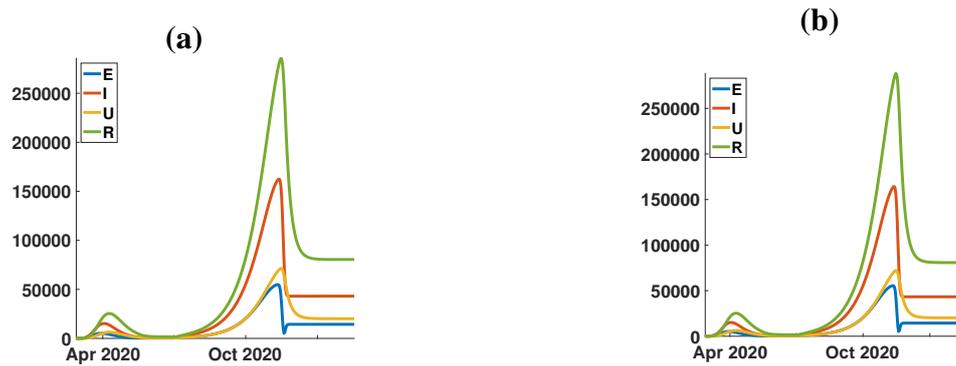
Figure S13. In this figure we explore the influence of the parameter f on the solution of model. The figure (a) corresponds to $\kappa = 0.1$ and figure (b) corresponds to $\kappa = 3$. The remaining parameters are unchanged.

Influence of ν on basic reproduction number:
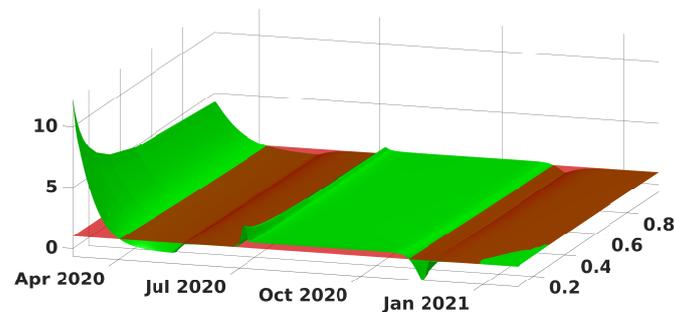
Figure S14. In this figure we plot $(t, \nu) \rightarrow R_e(t)$ when t varies from January 03 2020 to January 04 2021 and ν varies from 0.1 to 1 (or equivalently $1/\nu$ varies from 10 days to 1 day).
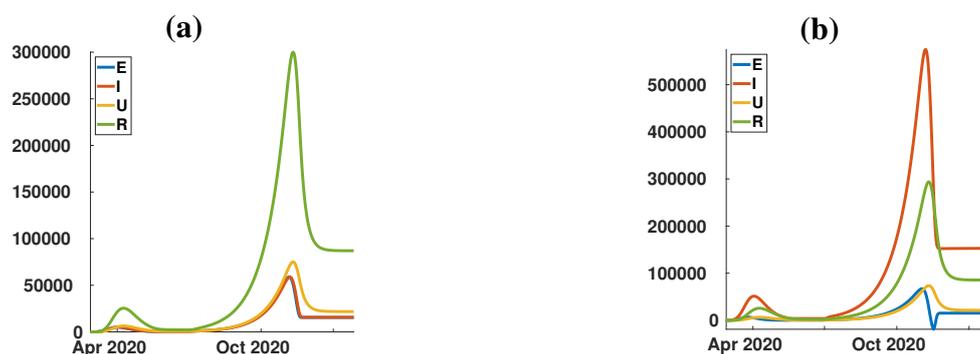
Figure S15. In this figure we explore the influence of the parameter $1/\nu$ on the solution of model. The figure (a) corresponds to $1/\nu = 1$ and figure (b) corresponds to $1/\nu = 10$. The remaining parameters are unchanged.

Influence of η on basic reproduction number:
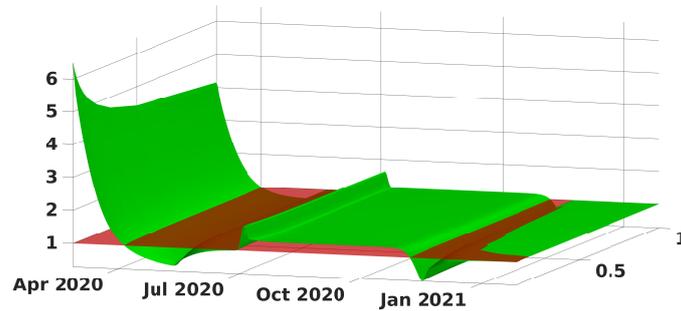
Figure S16. In this figure we plot $(t, \eta) \rightarrow R_e(t)$ when t varies from January 03 2020 to January 04 2021 and η varies from 0.1 to 1 (or equivalently $1/\eta$ varies from 10 days to 1 day).
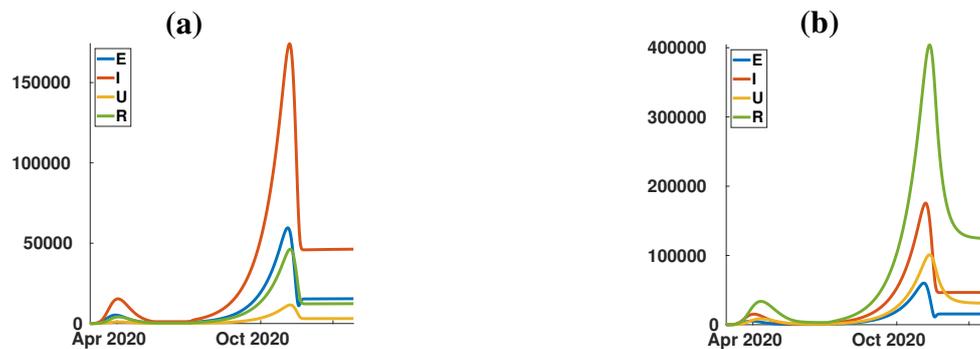
Figure S17. In this figure we explore the influence of the parameter f on the solution of model. The figure (a) corresponds to $1/\eta = 1$ days and figure (b) corresponds to $1/\eta = 10$ days. The remaining parameters are unchanged.

Influence of α on basic reproduction number:
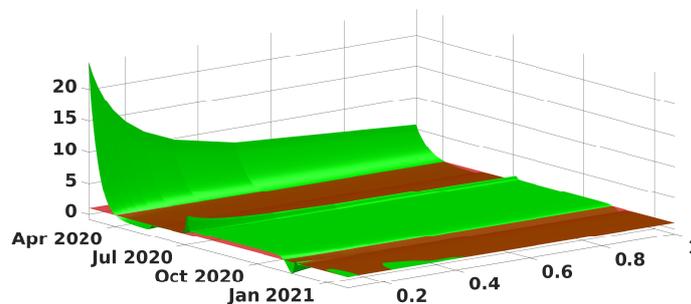
Figure S18. In this figure we plot $(t, \alpha) \rightarrow R_e(t)$ when t varies from January 03 2020 to January 04 2021 and α varies from 0.1 to 1 (or equivalently $1/\alpha$ varies from 10 days to 1 day).

(a)

(b)
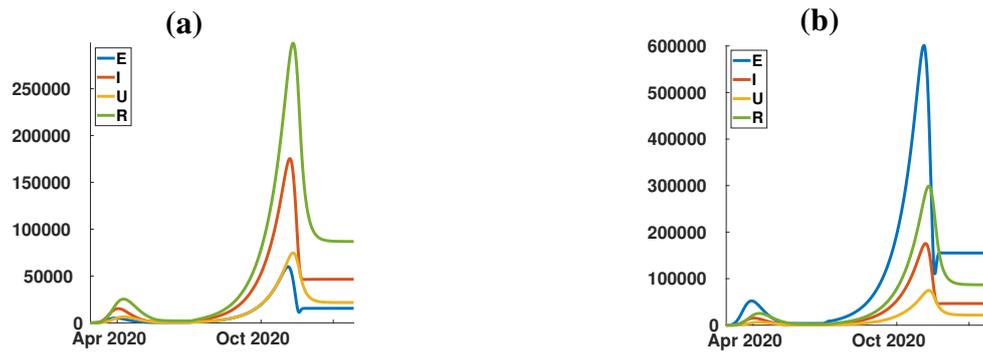
Figure S19. In this figure we explore the influence of the parameter f on the solution of model. The figure (a) corresponds to $1/\alpha = 1$ days and figure (b) corresponds to $1/\alpha = 10$ days. The remaining parameters are unchanged.

S7. Upper bound of the duration for the exposed period and the asymptomatic infectious period

Let us finally mention that for each country and each epidemic wave we evaluated the parameter $1/(\chi\,\theta)$. In Figure S20 we plot the histogram of its estimated value and obtain a median value be 15.61 days. Therefore the length of exposure and the length asymptomatic infectious period should smaller than 15.61 days.
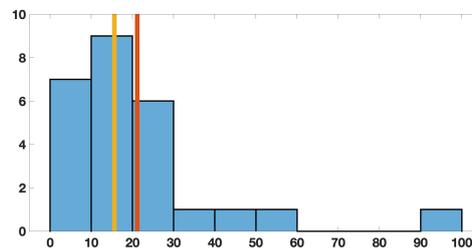
Figure S20. In this Figure we plot the histogram for the estimated values $1/(\chi\,\theta)$ (see Appendix E). The red vertical line is mean value which is equal to 21 days. The yellow vertical line is median value which is equal to 15.61 days.

In this section, we plot the estimated values of the parameter $1/(\chi\,\theta)$ for each epidemic period and each country consider in this study. The parameter corresponds to the upper bound of the length of the exposed period and asymptomatic infectious period. Indeed from the section devoted to the compatibility condition we know that the average duration of the exposed period should satisfy

$$1/\nu \leq 1/(\chi\,\theta),$$

and the average duration of the asymptomatic infectious period should should satisfy

$$1/\alpha \leq 1/(\chi\,\theta).$$
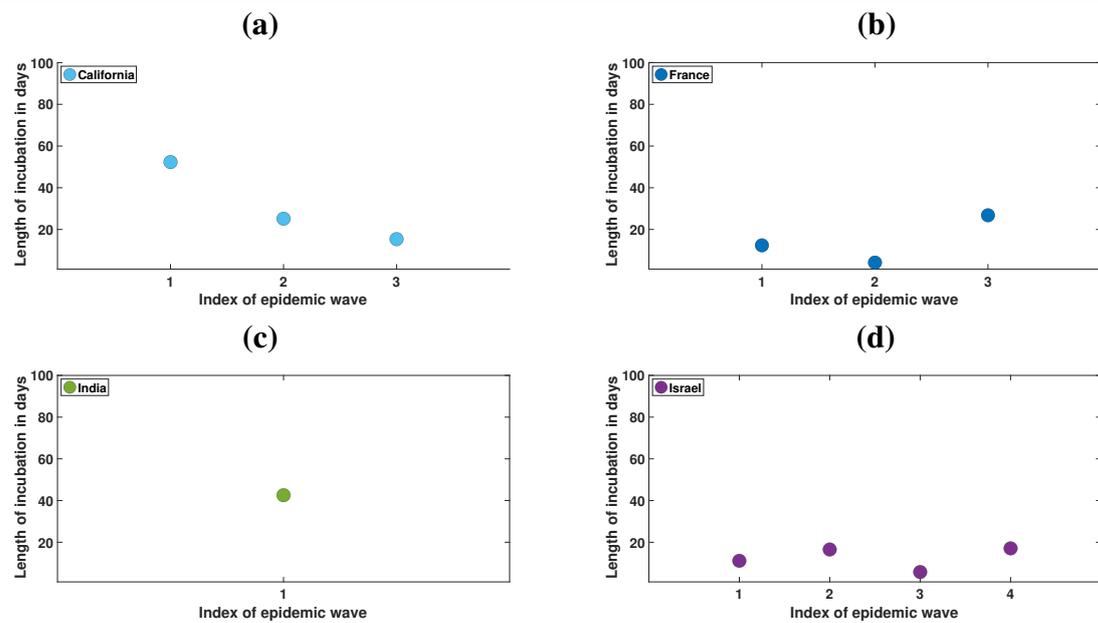
Figure S21. In this figure we plot the values of the parameter $1/(\chi\theta)$ estimated for each epidemic wave and for California (a), France (b), India (c), Israel (d). This parameter represents the maximal length of the incubation period. In each figure, we plot this parameter for each epidemic wave and for each country.
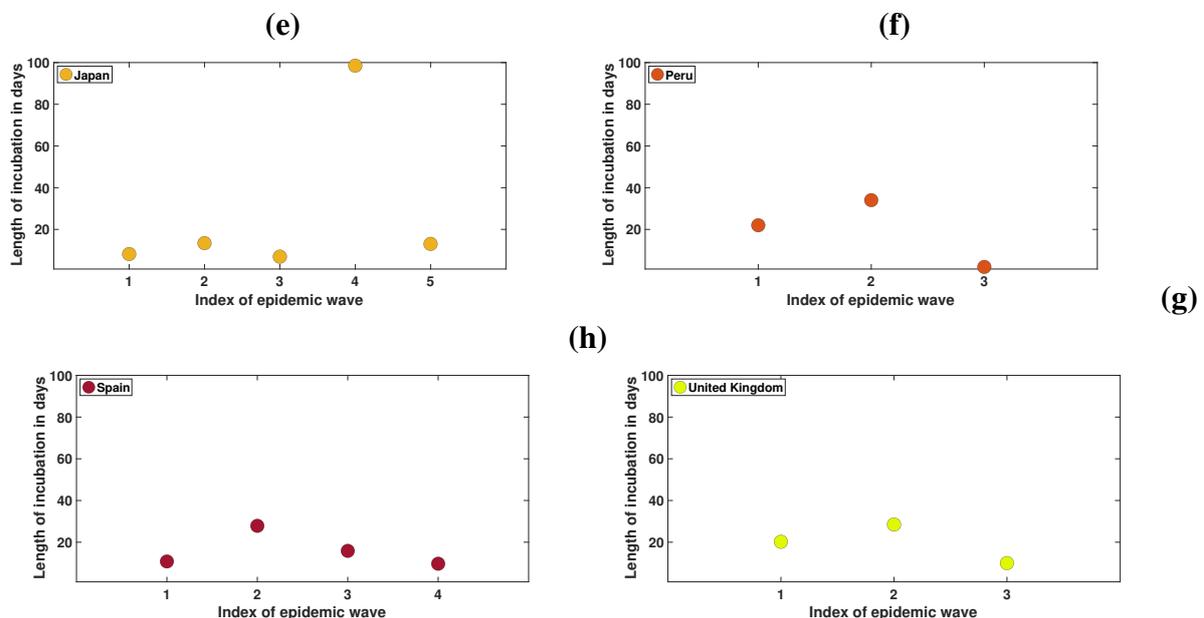
Figure S22. In this figure we plot the values of the parameter $1/(\chi\theta)$ estimated for each epidemic wave and for Japan (e), Peru (f), Spain (g) and United Kingdom (h). This parameter represents the maximal length of the incubation period. In each figure, we plot this parameter for each epidemic wave and for each country.

S8. Computing R_0

The basic reproduction number R_0 can be computed for the SEIUR model by the formula (see [71,72])

$$R_0 = \rho(FV^{-1}),$$

where F is the matrix containing new infections and V contains the rates of transfer between classes:

$$F := \begin{pmatrix} 0 & \tau S & \tau \kappa S & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad V := \begin{pmatrix} \alpha & 0 & 0 & 0 \\ -\alpha & v & 0 & 0 \\ 0 & -v(1-f) & \eta & 0 \\ 0 & v(1-f) & 0 & \eta \end{pmatrix},$$

see [71] and [72] for details. Therefore

$$V^{-1} = \begin{pmatrix} 1/\alpha & 0 & 0 & 0 \\ 1/v & 1/v & 0 & 0 \\ (1-f)/\eta & (1-f)/\eta & 1/\eta & 0 \\ f/\eta & f/\eta & 0 & 1/\eta \end{pmatrix},$$

$$FV^{-1} = \frac{\tau S}{\eta v} \begin{pmatrix} \eta + \kappa v(1-f) & \eta + \kappa v(1-f) & \kappa v & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

It follows that

$$R_0 = \frac{\tau S}{\eta v}(\eta + \kappa v(1-f)).$$

S9. Stochastic approach to effective reproductive ratio

In numerical applications, we also present the results obtained by applying the method described in the paper of Cori et al. [69]. Let us summarize the principle of the method. We consider that the incidence data (*i.e.,* the daily number of new reported cases) correspond to infection events that have occurred in the past. For each new reported case, we reconstruct the time the infectious period started by sampling a Gamma distribution (i.e. the time from the infection to the moment at which the individual is reported follows a Gamma distribution). The parameters of this Gamma distribution are computed to match the differential equation framework. In numerical application, $1/\mu = 10$ days, we took the average for the average of the Gamma distribution as well as its standard deviation. We denote I_t the resulting number of individuals that begin their infectious period on the day t. As described in [69], we use a smoothing window of τ days ($\tau = 14$ days in numerical applications). The resulting effective reproductive ratio R_t is then computed as

$$R_t = \frac{a + \sum_{s=t-\tau+1}^{t} I_s}{\frac{1}{b} + \sum_{s=t-\tau+1}^{t} \Lambda_s},$$

where a and b are a prior distribution on R_t (we took $a = 1$ and $b = 5$, as in [69]) and Λ_s is computed by the formula

$$\Lambda_s = \sum_{s=1}^{t} I_{t-s} w_s,$$

where w_s is the average infectiousness profile after time s. In numerical applications, and following [69, Web Appendix 11], we used the following formula for w_s

$$w_s = sF_{\Gamma,\alpha,\beta}(s) + (s - 2)F_{\Gamma,\alpha,\beta}(s - 2) - 2(s - 1)F_{\Gamma,\alpha,\beta}(s - 1)$$
$$+ \alpha\beta(2F_{\Gamma,\alpha+1,\beta}(s - 1) - F_{\Gamma,\alpha+1,\beta}(s - 2) - F_{\Gamma,\alpha+1,\beta}(s)),$$

where $F_{\Gamma,\alpha,\beta}(s)$ is the cumulative density of a Gamma distribution of parameters (α, β):

$$F_{\Gamma,\alpha,\beta}(t) = \int_0^t \frac{1}{\Gamma(\alpha)\beta^\alpha} s^{\alpha-1} e^{-\frac{s}{\beta}} \mathrm{d}s.$$

The parameters α and β are computed to match the Gamma distribution of the serial intervals which, in our case, have mean value $1/\mu = 10$ days and standard deviation as well of $1/\mu$, so that $\alpha = 1/\mu$ and $\beta = 1/\mu$.

Because of the sampling of random numbers involved in the computation of R_t, the procedure described above was repeated 100 times (each time drawing a new sequence of I_s from the daily number of new cases) and the final value of R_t presented in Figure 4 of the main text (green curves) is the average of the values obtained during these 100 simulations.