Combining and comparing regional epidemic dynamics in Italy: Bayesian meta-analysis of compartmental models and model assessment via Global Sensitivity Analysis

Giulia Cereda (✉ giulia.cereda@unifi.it)

University of Florence

Cecilia Viscardi

University of Florence

Michela Baccini

University of Florence

Combining and comparing regional epidemic dynamics in Italy: Bayesian meta-analysis of compartmental models and model assessment via Global Sensitivity Analysis

Giulia Cereda[1,2]**, Cecilia Viscardi**[1,2]**, and Michela Baccini**[1,2]

[1]Department of Statistics, Computer Science, Applications, University of Florence, Florence, 50134, Italy
[2]Florence Center for Data Science, University of Florence, Florence, 50134, Italy
[1]Correspondence to: michela.baccini@unifi.it,giulia.cereda@unifi.it,cecilia.viscardi@unifi.it

ABSTRACT

During autumn 2020, Italy faced a second important SARS-CoV-2 epidemic wave. We explored the time pattern of the instantaneous reproductive number, $R_0(t)$, and estimated the prevalence of infections by region from August to December calibrating SIRD models on COVID19-related deaths, fixing at values from literature Infection Fatality Rate (IFR) and infection duration. A Global Sensitivity Analysis (GSA) was performed on the regional SIRD models. Then, we used Bayesian meta-analysis and meta-regression to combine and compare the regional results and investigate their heterogeneity. The meta-analytic $R_0(t)$ curves were similar in the Northern and Central regions, while a less peaked curve was estimated for the South. The maximum $R_0(t)$ ranged from 2.61 (North) to 2.15 (South) with an increase following school reopening and a decline at the end of October. Average temperature, urbanization, characteristics of family medicine and health care system, economic dynamism, and use of public transport could partly explain the regional heterogeneity. The GSA indicated the robustness of the regional $R_0(t)$ curves to different assumptions on IFR. The infectious period turned out to have a key role in determining the model results, but without compromising between-region comparisons.

Introduction

After the first SARS-CoV-2 outbreak during spring 2020, Italy faced a second epidemic wave during the autumn. In order to reduce the rate of contagion and prevent the collapse of the health care system, the Italian government introduced regional-level measures of social distancing of different degrees, starting from November 6th, 2020. Among others, a national curfew from 10 pm to 5 am was implemented, and the regions were weekly classified as low risk, medium risk, and high risk zones (yellow, orange, and red zones, respectively), according to indicators centrally calculated by the Istituto Superiore di Sanità (ISS). The timing and the degree of the containment measures likely influenced the epidemic dynamics[1], but also socio-economic, demographic characteristics of the population, and environmental factors may have had a role in determining and moderating the level of contagion and its pattern over time.

Several studies investigated different aspects related to the epidemic dynamics during the first and/or the second waves in Italy by using compartmental models[2–5] or different approaches[6–8]. Some studies performed descriptive comparisons among regions (mostly during the first wave)[9–12] or between the first and the second epidemic wave[13]. Others explored possible determinants of the heterogeneity in COVID19 incidence and mortality across the country, focusing on the beginning of the emergency in the spring 2020[14–16].

The aim of our study was to describe the epidemic dynamics in Italy from August 1st, 2020 to the end of the same year, in order to obtain an overall picture of the second wave in the country with a special focus on the contagion spread, highlighting and investigating the heterogeneity among regions. To this end, we adopted a two-step procedure.

First, we estimated a compartmental model of SIRD-type for each region in order to investigate the trend of the contagions over time. Compartmental models, which take their name from the fundamental assumption that at each time during the epidemic the population is divided into homogeneous groups or "compartments", are widely used in the literature for forecasting and inference purposes, with examples also in COVID19 research[17]. When the interest is to make inference, the model parameters are estimated minimizing the distance between observed data and model predictions (calibration). In our study, we calibrated the regional SIRD models on the daily number of notified COVID19-related deaths, made publicly available by the Protezione Civile[18]. Compared with the number of notified cases, mortality data – although possibly subject to some

notification delays – can be considered as more reliable and likely less influenced by the capacity of the health system to detect infections. The calibration procedure allowed us to investigate the behavior of the contagion over time in terms of instantaneous reproductive number $R_0(t)$, which quantifies the average number of secondary infections caused by a single infected individual over time, as well as in terms of the number of infected individuals, which can exceed, even by far, the number of the notified ones.

At the second step, we combined and compared the regional results by using Bayesian multivariate and univariate meta-analytic techniques, and we investigated through meta-regressions the possible role of region-level variables in explaining between-region observed discrepancies in terms of contagion spread.

In the regional SIRD models, we set the infection fatality rate (IFR) and the average time from infection onset to infection resolution to plausible values arising from the literature, in order to assure parameter identifiability. Treating part of the parameters in the compartmental models as fixed is a common practice that poses the problem, usually not addressed, of evaluating the robustness of the results when different values are specified for these unestimated parameters[19]. In this paper, we performed a sensitivity analysis on the estimated $R_0(t)$ curve and prevalence of infections, by changing one at a time the values of IFR and infection duration in the SIRD models. Additionally, we implemented a Global Sensitivity Analysis (GSA) procedure[20] to quantify and characterize the uncertainty around the calibration results that propagated from the uncertainty around the values of those parameters that were not the object of the inference. Despite GSA is not yet widely used in epidemiology, it appears as one of the recommendations in the manifesto by Saltelli and colleagues, which offers a critical view of modeling in time of pandemic[21].

Finally, as an ancillary result, we obtained an evaluation of the submerged fraction of contagion and an indirect assessment of the admissible IFR values, by comparing the number of infections estimated by the regional SIRD models, which by definition includes both notified and not notified cases, with the observed number of notified positives reported by the Protezione Civile[18].

1 Data

For our analyses, we used the national database on the evolution of the COVID19 pandemic, made available on a daily basis by the Protezione Civile[18]. This database collects the numbers of notified positive, hospitalized, recovered and deceased subjects by region. In the estimation phase, we used the daily number of COVID19-related deaths from August 1st, 2020 to January 14th, 2021 for all Italian regions, and the number of notified circulating infections on July 31st. The daily number of new infections and the daily number of circulating infections were used in a descriptive way to indirectly assess the submerged fraction of contagion. Although provided separately, we merged data of the two autonomous provinces of Trentino Alto Adige (Bolzano-Alto Adige and Trento), in order to obtain a more stable estimate for the entire region. We removed from the death counts of Emilia-Romagna 154 cases that, although happened during spring 2020, have been added to the data set on August 15th. In the map of Supplementary Figure S1, we represented, for each Italian region, the population size and the total number of COVID19-related deaths notified during the study period (see also Supplementary Table S1).

We collected socio-demographic and economic indicators measured at the regional level for 2020 or for the last available year from the website of the Italian National Institute of Statistics (ISTAT)[22] (see Section 2.3 for further details). We calculated the regional average temperatures during the study period from the daily temperature measurements reported on the website www.ilmeteo.it.

2 Methods

2.1 Regional SIRD models

For each region, we adopted a compartmental model of SIRD type, described by the following system of differential equations:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta(t)\frac{S(t-1)}{S(0)}I(t-1) \\ \frac{dI(t)}{dt} = \beta(t)\frac{S(t-1)}{S(0)}I(t-1) - \alpha I(t-1) - \delta I(t-1) \\ \frac{dR(t)}{dt} = \alpha I(t-1) \\ \frac{dD(t)}{dt} = \delta I(t-1) \end{cases} \tag{1}$$

where $S(t), I(t), R(t)$ and $D(t)$ are the sizes of the Susceptible, Infected, Recovered and Deceased compartments at time t[23]. For each region, we fixed $I(0)$ to the number of notified circulating infections on July 31st (from now on denoted by i_0) as reported by the Protezione Civile[18]. We approximated the number of susceptible people at time 0 (August 1st) as $S(0) = N - I(0)$,

with N regional population size at January 1st 2020 (Permanent Census of Population[22]). We set $D(0) = 0$ and $R(0) = 0$, thus starting to count deaths and recoveries from August 1st. The parameters α and δ are the transition rates from the compartment of the infected individuals to the compartments of the recovered and deceased ones, respectively. They depend on the IFR, p, on the average times from infection to death, T_D, and from infection to recovery, T_R. Having set $T_D = T_R = T$, the following relationships hold: $\alpha = \frac{1-p}{T}$, $\delta = \frac{p}{T}$[24]. The infection rate $\beta(t)$ is related to the instantaneous reproductive number $R_0(t)$, modelled as time-dependent, as follows:

$$\beta(t) = R_0(t)(\alpha + \delta) = \frac{R_0(t)}{T}. \tag{2}$$

At the beginning of the epidemic, $R_0(t)$ corresponds to the basic reproductive number, defined as the number of secondary infections generated by the first infected individual in the population. $R_0(t)$ is also related to the effective reproductive number, $R_t = R_0(t)S(t)/S(0)$[25], that measures the actual transmission at a specific time accounting for the natural depletion of susceptible individuals as the contagion spreads. R_t departs from $R_0(t)$ only at an advanced stage of the epidemic . To get a flexible estimate of $R_0(t)$, we modelled it through a natural cubic regression spline[26], with 4 internal equi-spaced knots (6 degrees of freedom): $R_0(t) = s(t; \boldsymbol{\vartheta})$, where $\boldsymbol{\vartheta}$ is a vector of unknown coefficients, to be estimated.

We assured parameter identifiability by fixing in the model $T = 14$ days and $p = 1.15\%$. The value $p = 1.15\%$ is the IFR estimate reported for the upper-income countries by the Imperial College COVID19 response team[27]. It is also consistent with the value of 1.14% estimated for Italy by the Italian Institute for International Political Studies[28] and used in a previous paper by the authors[29]. Regarding T, the value of 14 days is in line with both the median time from symptoms onset to death reported by ISS for Italy (12 days)[30] and the estimated average time from infection onset to recovery of 13.4 days arisen from a recent meta-analysis[31]. We explored these two choices explored via sensitivity analysis and GSA.

In the estimation phase, we discretized the differential equations in (1) and evaluated the size of the compartments by considering unit time intervals (details in Supplementary Section S1). This allowed us to estimate the model minimizing over $\boldsymbol{\vartheta}$ the following sum of squares:

$$Q(\boldsymbol{\vartheta}) = \sum_{t=1}^{K} \left(D(t; \boldsymbol{\vartheta}) - D^{\text{obs}}(t) \right)^2, \tag{3}$$

where $t = 1$ corresponds to August 1st 2020, $t = K$ to January 14th 2021, and $D^{\text{obs}}(t)$ denotes the cumulative number of deaths observed from August 1st. Notice that deaths observed during the first two weeks of January provide information about the contagion at the end of the year, given that deaths are assumed to happen on average $T = 14$ days after the disease onset.

We performed the minimization of (3) through the *auglag* function of the `nloptr` package of R software (http://ab-initio.mit.edu/nlopt), constraining the function $R_0(t)$ to positive values. We ran the estimation algorithm 100 times, using different initial values sampled from a multivariate grid defined on the values of $\boldsymbol{\vartheta}$. Among the 100 parameters estimates thus obtained, we selected the estimate $\hat{\boldsymbol{\vartheta}}$ associated to the lowest value of $Q(\cdot)$.

We implemented a parametric bootstrap procedure, in order to quantify the sampling variability around the estimates. Following a consolidated procedure[32,33], we assumed a Negative Binomial distribution on the daily increments of the estimated time series $D(t; \hat{\boldsymbol{\vartheta}})$ and generated 500 bootstrap samples to be used as observed time series in as many calibrations. The 90% confidence intervals or bands for the quantities of interest have been calculated as the 5th and 95th percentiles of the bootstrap distributions (see Supplementary Section S2 for further details).

2.2 Sensitivity analyses

To investigate how changes in the values of p and T affected the estimates of $R_0(t)$ and the shapes of the epidemiological curves arising from the regional SIRD models, we repeated the analyses for $p = 1.79\%, 0.78\%$ and $T = 10, 18$. The values $p = 0.78\%$ and $p = 1.79\%$ are the 95% confidence interval bounds of the IFR estimate in Brazeau et al.[27]. We performed an additional analysis fixing $p = 0.5\%$, as an extreme lower bound for the IFR. The value $T = 10$ is consistent with the 95% lower bound of the estimated mean time of Byrne et al.[31], while $T = 18$ is the estimated mean duration of the maximal infectious period from the same study.

Then, we went beyond the previous one-factor-at-a-time sensitivity analysis by performing a GSA, calculating the Sobol's variance indexes[20]. Given a function that relates inputs to outputs, the GSA explores how the outputs vary as the inputs change, to determine the inputs most contributing to the outputs' behavior (factor prioritization), finding non-influential inputs (model simplification), and investigating interaction effects between inputs. This can be done by propagating the uncertainty around the inputs to the outputs via MC simulations, then using Sobol's decomposition of the variance of each output thus obtained, and apportioning it among the different inputs[34] (see Supplementary Section S3). The contribution of each input to the output variance can be quantified by computing first order indexes (and superior order indexes) and total variance indexes. In particular,

the first order index of a given input represents the proportion of the output variance which is due to the main effect of the input (i.e. the first-order effect), while the total effect index represents the proportion of the output variance which is due to the main effect of the input and all its interactions with the other inputs (the higher-order effects).

In our application, we considered as inputs the fixed parameters of the SIRD model (p, T_R, T_D, i_0) and as outputs the parameters estimated by calibration, as well as derived quantities, such as the maximum and minimum $R_0(t)$, the peak of infections, and the dates at which they occurred, together with the first date when $R_0(t)$ reached the threshold of 1 after the maximum infection peak. The model was the calibration algorithm, given the observed data.

We calculated the first order and total variance indexes of each input on each output, relying on the results of 5'040 MC simulations, that in our case corresponded to 5'040 calibrations of the SIRD model. Each calibration was performed under a different combination of inputs, obtained by sampling p from the continuous uniform distribution $\mathscr{U}[0.0078, 0.0179]$[27], and the transition times and $I(0)$ from the following discrete uniform distributions: $T_R \sim \mathscr{U}\{7, 21\}$, $T_D \sim \mathscr{U}\{7, 21\}$, $I(0) \sim \mathscr{U}\{i_0, 3i_0\}$. For each input, the aggregate total variance index on the vector $\boldsymbol{\vartheta}$ was calculated as a weighted average of the total variance indexes of the single spline coefficients[35].

Given the huge computational effort required by the GSA implementation, we performed it on a virtual machine with 16 vCPU only for one region (Tuscany). However, we expect that the results can be generalized to the other regions. The GSA was conducted by using the *soboljansen* and the *sobolMultOut* functions of the R package `sensitivity` (https://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf).

2.3 Bayesian meta-analyses

We used a multivariate Bayesian random effects meta-analysis model to combine the estimated region-specific $R_0(t)$ curves, accounting for the heterogeneity among regions, and to combine the regional estimates of the monthly average prevalence of infection from September to December.

Additionally, we conducted univariate Bayesian random effects meta-analyses on the following quantities obtained from the regional $R_0(t)$ curves:

- average value of $R_0(t)$ from October 1st to December 31st;

- maximum value of $R_0(t)$ during the study period;

- variation of $R_0(t)$ during the 4 weeks following November 6th, date of the introduction of social distancing measures by the central government, with a classification of the regions according to three risk levels;

- variation of $R_0(t)$ during the 4 weeks following the beginning of the school.

For the variation of $R_0(t)$ during the 4 weeks following November 6th, we performed also meta-analyses by intensity assigned in the first week of the region-specific measures. Univariate and multivariate meta-analyses were conducted on all regions and separately by geographical area: Southern regions (Basilicata, Calabria, Campania, Molise, Puglia, Sardegna, Sicilia), Central regions (Abruzzo, Lazio, Marche, Toscana, Umbria), Northern regions (Emilia Romagna, Friuli Venezia Giulia, Liguria, Lombardia, Piemonte, Trentino, Valle d'Aosta, Veneto).

Finally, Bayesian meta-regression analyses were performed on the average $R_0(t)$ from October 1st to December 31st in order to investigate possible sources of between-region heterogeneity. Specifically, we evaluated as possible effect modifiers the following variables measured at the regional level (for 2020 or for the last available year): percentage of people with at least two chronic diseases in the population, number of general practitioners per 10'000 residents, number of pediatricians per 10'000 children, number of hospitals per 1'000 residents, and percentage of public hospitals over the total number of hospitals, aging index (number of over 65 per 100 individuals younger than 15), mean age of the population, average size of households, percentage of people with an academic degree per 10'000 residents, schooling rate (percentage of the individuals aged 20-24 who have at least a high school diploma), percentage of children attending kindergartens, percentage of workers using public transports to go to work, percentage of people (18-34) using public transports to go to school, poverty index (percentage of people living in households below the poverty threshold), employment rate (percentage of employed persons 15-64), tourism rate (days of presence of tourists during the year per inhabitant), total energy consumption of industries and manufactures, percentage of residents living in high-urbanization areas, average temperature from October to December 2020 (see Supplementary Table S2 for details on meta-regressors and related references). We specified separate meta-regression models each of which included only one meta-regressor.

In all analyses, non-informative priors were assumed on the model hyperparameters. We get a sample from the joint posterior distribution of the hyperparameters via MCMC algorithms. A detailed description of the models and the software used for the analysis is reported in Supplementary Section S4.
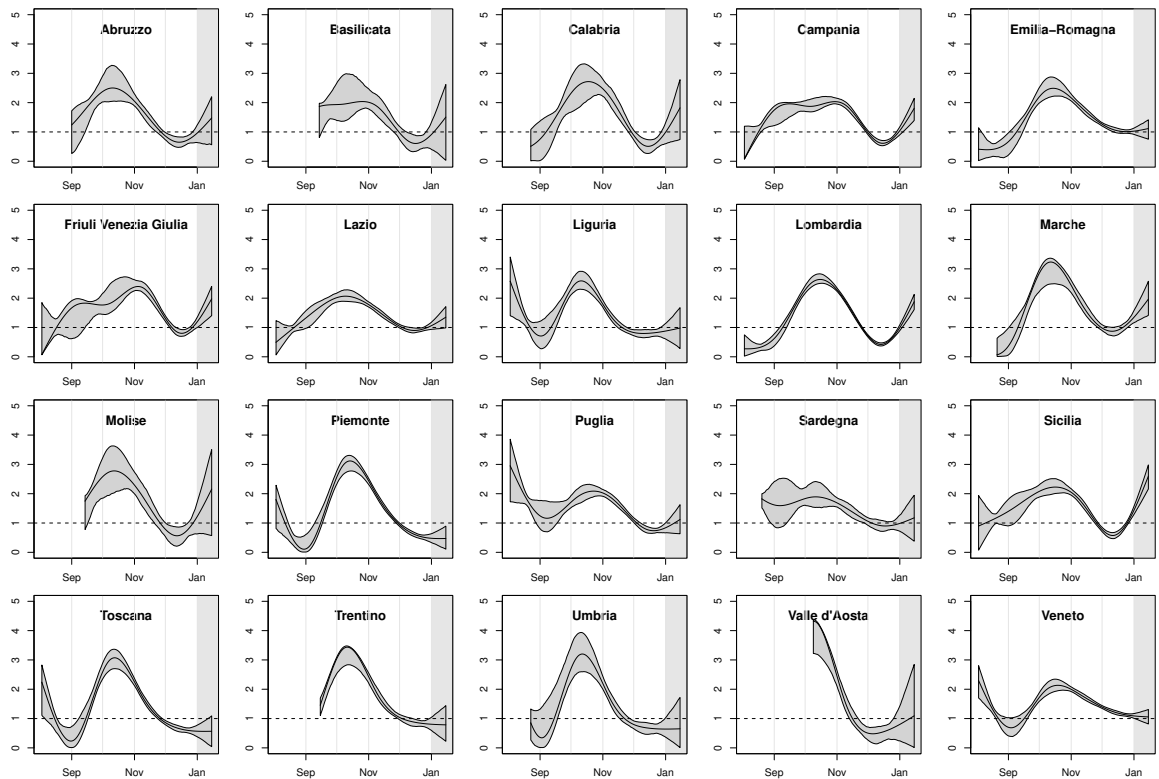
Code availability

R code used for calibration of the regional SIRD models and for sensitivity analysis and GSA are available from the authors upon request.

3 Results

3.1 Results of the main analysis ($p = 1.15\%$, $T = 14$)

The fit of the SIRD model was good for all regions, with the expected cumulative deaths close to the observed ones (Supplementary Figure S2).


1.pdf

Figure 1 Estimated $R_0(t)$ with point-wise 90% confidence bands, by region; $p = 1.15\%$, $T = 14$ days.

In Figure 1 we reported the estimated curves of $R_0(t)$ arising from the regional SIRD models with their 90% point-wise confidence bands. For the regions where the first COVID19-related death was observed after the second week of August, the curve was shown starting 14 days before the first death, because of the extremely poor information on the previous period. Moreover, we highlighted in grey the last 14 days of the study period, to stress that the estimates for that period were quite unreliable, having assumed an average time from infection onset to death of two weeks. This is also the reason why, in order to draw valid conclusions about $R_0(t)$ until the end of the year, we calibrated the model up to January 14th, 2021. The pattern of $R_0(t)$ was heterogeneous among regions, but, with few exceptions, a mid/late-October peak was visible as the culmination of growth started in early/mid September. In several regions, $R_0(t)$ declined and then increased again starting from mid December.

Figure 2 summarizes these regional results showing the overall meta-analytic Italian $R_0(t)$ curve and the meta-analytic curves by geographical area, obtained from Bayesian multivariate meta-analyses (pointwise posterior mean and 90% credible intervals). It is quite evident that during the second epidemic wave the shape of $R_0(t)$ was similar in the Northern and Central regions, with an initial increase and a clear peak around middle October. On the contrary, the overall curve for Southern regions was flatter, without the initial decrease. The combination of all the regional curves leads to an overall $R_0(t)$ curve (left plot) similar to the one obtained for Northern and Central regions, even though with tighter credible bands.

In Figure 3 (A), we reported the regional curves describing the prevalence of infections per 1'000 inhabitants over time, arising from the SIRD models (see also Supplementary Figures S3 and S4 and Supplementary Table S4). These estimates are inclusive of detected and undetected cases, thus in principle they should be an upper bound for the number of notified
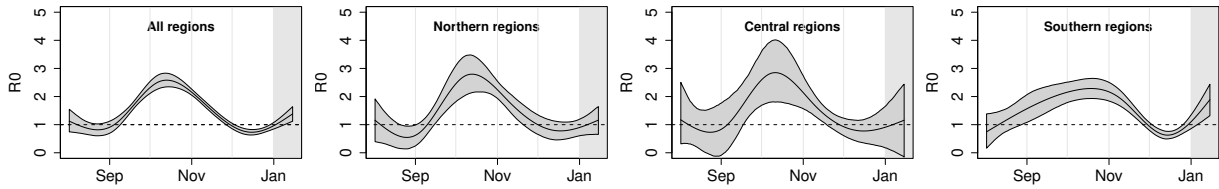
2.pdf

Figure 2   Posterior mean of the meta-analytic curves of $R_0(t)$ with point-wise 90% credible bands for the entire country and by geographical area; $p = 1.15\%$, $T = 14$ days.
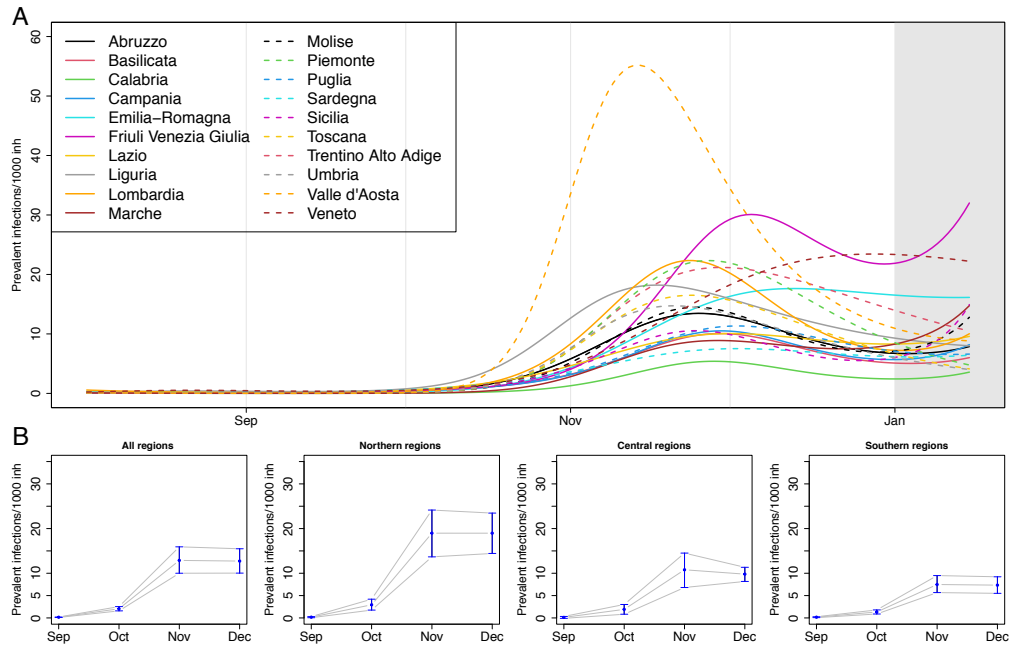
Figure 3   (A) Estimated prevalence of infections (number of circulating infections over 1'000 inhabitants) by region, and (B) posterior mean of the meta-analytic monthly-average prevalences, for the entire country and by geographical areas, with 90% credible intervals (lower panel); $p = 1.15\%$, $T = 14$ days.

cases provided by Protezione Civile (see Section 3.3 for a discussion about this point). Valle d'Aosta exhibited the largest peak of prevalence, reached in the first half of November, with more than 50 circulating infections every 1'000 inhabitants. It was followed - though with less than half the value of its prevalence - by Friuli Venezia Giulia, Veneto, Piemonte, Lombardia, Trentino Alto Adige and then by Liguria and Emilia Romagna. The lowest prevalence was estimated for Calabria. Liguria and Valle d'Aosta reached the peak of circulating infections before the other regions, while Veneto was the last one.

The overall estimates of the average monthly prevalence arising from the Bayesian multivariate meta-analysis (Figure 3, B) highlight how the prevalence of infections in the Northern regions was larger than in the Central and Southern ones, exceeding on average 15 cases every 1'000 inhabitants during the month of November.

3.1.1 Results of univariate meta-analyses and meta-regressions

|		Estimate	90% CrI		I^2	90% CrI	
Average $R_0(t)$ from October to December	All regions	1.59	1.53	1.65	90.2	83.0	94.7
	North	1.66	1.57	1.76	92.1	82.0	97.4
	Center	1.61	1.44	1.78	93.1	79.6	98.6
	South	1.48	1.39	1.56	80.2	56.5	94.0
Maximum $R_0(t)$ over the study period	All regions	2.61	2.39	2.83	91.6	84.8	95.7
	North	2.83	2.44	3.26	94.5	85.0	98.3
	Center	2.78	2.27	3.32	87.5	65.2	97.5
	South	2.15	1.97	2.35	55.3	20.3	88.9
4 week change in $R_0(t)$ after school re-opening	All regions	0.50	0.35	0.66	87.5	78.0	93.3
	North	0.74	0.54	0.94	74.4	32.5	93.0
	Center	0.58	0.15	1.03	89.7	69.9	98.0
	South	0.16	0.02	0.30	55.7	24.6	86.1
4 week change in $R_0(t)$ after November 6th	All regions	-0.75	-0.84	-0.66	88.4	80.9	93.7
	North	-0.72	-0.91	-0.54	95.8	90.4	98.7
	Center	-0.73	-0.90	-0.56	76.7	46.6	94.8
	South	-0.81	-0.97	-0.66	76.3	45.0	93.1

Table 1 Results of the Bayesian meta-analyses (posterior mean of the quantity of interest and posterior median of the I^2 index, with associated 90% credible intervals) conducted on: average value of $R_0(t)$ from October to December, maximum value of $R_0(t)$ over the study period, changes in $R_0(t)$ after school re-opening and after the introduction of national restrictions on November 6th. $p = 1.15\%$, $T = 14$ days.

Table 1 summarizes the results of the Bayesian meta-analyses on the quantities derived from the regional $R_0(t)$ curves: average value of $R_0(t)$ from October to December, maximum value of $R_0(t)$ over the study period, changes in $R_0(t)$ arising in four weeks after school re-opening and after the introduction of national restrictions on November 6th. For all quantities, the heterogeneity among regions was high, with the lowest I^2 estimated among the Southern regions. The average level of $R_0(t)$, as well as its maximum, was higher in the Central and Northern regions than in the Southern ones. An overall increase of $R_0(t)$ equal to 0.50 was estimated during the first 4 weeks after school re-opening in September (the dates of school re-opening in every region are reported in Supplementary Table S3). This increase was lower in the South than in the Central and Northern regions.

We found also evidence of an overall decline of $R_0(t)$ equal to 0.75 during the four weeks following the introduction of the restrictions on November 6th. This decline was similar across geographical areas (Table 1), but appeared to be associated with the strength of the measures adopted at a regional level, as shown by the posterior distributions reported in Figure 4 (A). In the four regions initially classified as at high risk (red regions: Calabria, Lombardia, Piemonte, Valle d'Aosta), where stronger restrictions were immediately adopted, the decline was steeper than in the two classified as at medium risk (orange regions: Puglia, Sicilia) and in the remaining 14 low risk regions (yellow), subject to lighter measures.

Figure 4 (B) shows the results of the meta-regressions on the average value of $R_0(t)$ from October to December, in terms of change in the average value of $R_0(t)$ associated with an interquartile range variation in the meta-regressor. The change was multiplied by 100 so that the reported values indicated the changes in the number of secondary infections derived from 100 infected individuals at time t. We considered as relevant a meta-regressor if the posterior probability that the change was larger than 5 or lower than -5 exceeded 50%. According to this criterion, it was evident a marginal positive association of the average $R_0(t)$ with employment rate, use of kindergartens, tourism rate and, to a less extent, with population in high density areas, schooling index, and use of public transport to go to school. A negative association was found with number of practitioners and pediatricians, poverty index, temperature, prevalence of people with at least two chronic diseases, and, to a less extent, with the number of hospitals and the proportion of the public ones. The results of the meta-regressions are reported also in Supplementary Table S5.
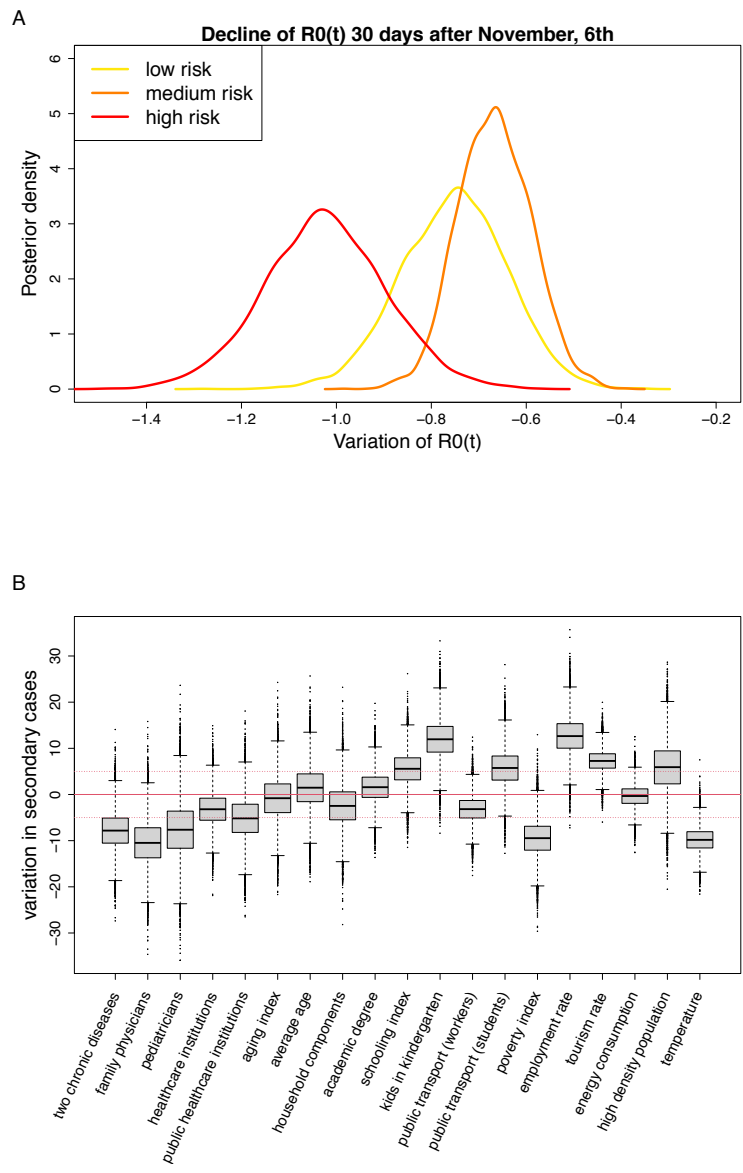
Figure 4 (A) Posterior distributions of the change in $R_0(t)$ during the first four weeks following the introduction of the restriction measures on November 6th, by level of alarm assigned to the region, and (B) results of the meta-regressions on the average value of $R_0(t)$ from October to December (posterior distribution of the variation in the average value of $R_0(t) \times 100$, associated to a change in the meta-regressor equal to its observed interquartile range). $p = 1.15\%$, $T = 14$ days.

3.2 Results of the sensitivity analyses
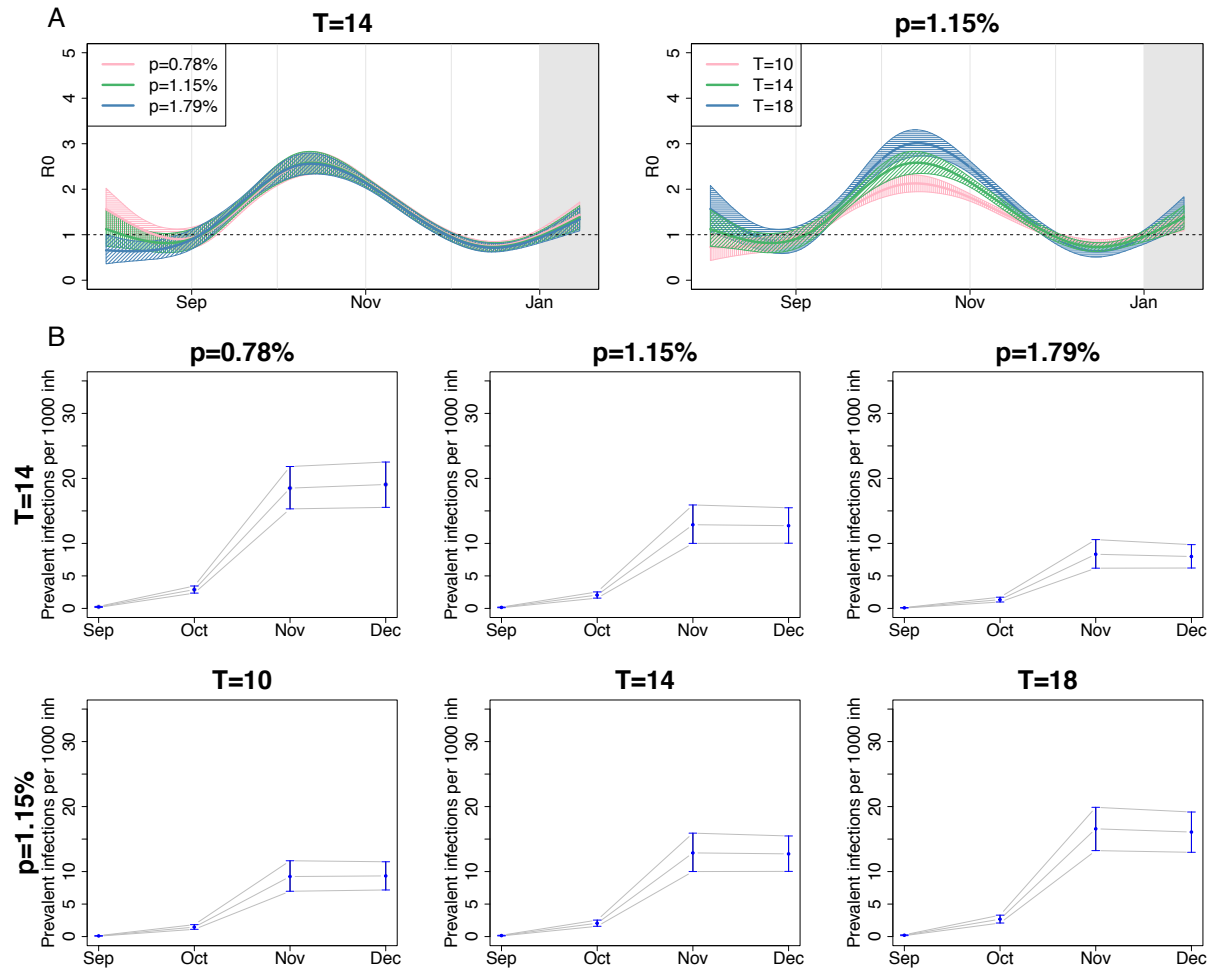
5b.pdf

Figure 5 (A) Posterior mean and 90% credible bands of the meta-analytic R_0 curves and (B) posterior mean and 90% credible intervals of the meta-analytic monthly-averaged prevalence per 1'000 inhabitants (second and third rows) for the entire country, when using different values for p, fixed $T = 14$ days, and when using different values for T, fixed $p = 1.15\%$.

The shape of the regional $R_0(t)$ curves appeared quite similar under different IFR scenarios when $T = 14$, as shown by the comparison of the overall meta-analytic curves obtained by fixing the IFR to different values in the regional SIRD models (Figure 5, A, left). Some discrepancy was observed only at the beginning of the study period, when larger values of $R_0(t)$ were obtained in correspondence of smaller values of p. The regional curves obtained changing p are shown in Supplementary Figures S5-S7. Figure 5 (A, right) compares the overall meta-analytic curve obtained by assuming different values for T, having fixed $p = 1.15\%$. A shorter infection duration corresponded to a less peaked curve, but the date when $R_0(t)$ was maximum and the date when it crossed the value of 1 after the introduction of containment measures were preserved.

In the lower panel of Figure 5, we reported the meta-analytic estimates of the monthly averaged prevalence of infections per 1'000 inhabitants from September to December, when changing p, fixed T (second row), and when changing T, fixed p (third row). The estimated prevalence decreased consistently with the value of p and increased with the values of T. Notice that the number of prevalent infections when $T = 10$ and $p = 1.15\%$ and when $T = 14$ and $p = 1.79\%$ resembled one another. The same happened with the pairs of parameters ($T = 14$, $p = 0.78\%$) and ($T = 18$, $p = 1.15\%$). This suggests that increasing (decreasing) p produces the same effect of decreasing (increasing) T, when the other parameter is fixed, and that a global evaluation of the impact of these quantities on the results is required.

Table 2 shows the total variance indexes, including the aggregate index for the vector $\boldsymbol{\vartheta}$, and the first order indexes obtained from the GSA procedure. As suggested by the magnitude of the total variance indexes, changes of the fixed parameters p, T_D, and i_0 had a very low impact on the coefficients vector $\boldsymbol{\vartheta}$. On the contrary, T_R exhibited the greatest aggregated total

			Spline coefficients						$R_0(t)$				Infections	
	ϑ_1	ϑ_2	ϑ_3	ϑ_4	ϑ_5	ϑ_6	Aggr. ϑ	Max	Max (date)	Min	Min (date)	Cross 1 (date)	Peak	Peak (date)
p	0.10	0.21	0.12	0.08	0.08	0.21	0.09	0.02	0.39	0.19	0.14	0.16	0.16	0.50
i_0	0.18	0.35	0.23	0.16	0.14	0.36	0.17	0.02	0.58	0.27	0.19	0.03	0.00	0.45
T_D	0.18	0.40	0.21	0.15	0.14	0.34	0.17	0.03	0.73	0.30	0.20	0.18	0.32	0.57
T_R	0.68	0.10	0.61	0.72	0.78	0.23	0.72	0.99	0.72	0.69	0.78	0.87	0.74	0.85

First order indexes

			Spline coefficients						$R_0(t)$				Infections	
	ϑ_1	ϑ_2	ϑ_3	ϑ_4	ϑ_5	ϑ_6		Max	Max (date)	Min	Min (date)	Cross 1 (date)	Peak	Peak (date)
p	0.07	0.20	0.09	0.07	0.05	0.12		-0.01	-0.06	0.10	0.03	0.01	0.04	0.14
i_0	0.12	0.34	0.16	0.11	0.08	0.32		0.00	0.00	0.13	0.02	-0.04	-0.04	0.05
T_D	0.11	0.32	0.13	0.09	0.07	0.21		-0.03	-0.26	0.14	0.06	0.00	0.11	0.02
T_R	0.53	0.03	0.43	0.60	0.64	0.09		0.95	-0.18	0.42	0.61	0.73	0.56	0.25

Table 2 Total variance indexes and first order indexes of each model input (by row) on the coefficients of the $R_0(t)$ regression spline, maximum and minimum values of $R_0(t)$ with the corresponding dates, date in which $R_0(t)$ crossed the value one, infection peak with the corresponding date (by column).

variance index (0.72), proving to be an input that contributes very much to the $R_0(t)$ curve as a whole. Regarding the impact on the derived quantities, T_R was still the most relevant input (the only relevant on the maximum values of $R_0(t)$), even if non-negligible indexes were found also for the other inputs, especially on the date of occurrence of the maximum $R_0(t)$, on the date when $R_0(t)$ crossed the threshold of 1 and on the date of the peak of infections. In order to correctly interpret these results, one should however consider that some outputs could vary within a small range of values as the inputs change. Hence, the corresponding total variance indexes, although high, could actually derive from the apportionment of a small total variance. This is in itself indicative of overall robustness of these outputs to inputs perturbations. For instance (Supplementary Figure S8 and Table S5), the date of the infection peak (coefficient of variation 0.002) was almost unaffected by variations of the model inputs as well as the date in which $R_0(t)$ crossed the value of 1 (coefficient of variation 0.01). Analogously ϑ_4 had the most dispersed distribution among the other coefficients, suggesting its greater sensitivity to changes of the inputs values. This larger dispersion was considered in the calculation of the aggregate total variance index, which averages the single total variance indexes with weights proportional to the ϑ_i variances[35].

While in the regional SIRD models we constrained $T_D = T_R = T$, the GSA was conducted allowing T_D to vary independently from T_R. This choice emphasized that T_D was less relevant than T_R on output variability. Therefore, we can conclude that, when T_R is properly set, misspecification of T_D has a negligible effect. In such a case, forcing T_D to be equal to T_R, as in our regional analysis, produces a simplification of the SIRD model without inducing relevant variations in the outputs.

Finally, being the total variance indexes for $\boldsymbol{\vartheta}$ only slightly higher than the corresponding first order indexes, it was evident that interactions among inputs were not relevant on the estimated $R_0(t)$ curve. On the contrary, interactions were a relevant source of variability in the derived quantities. The fact that the first order indexes on the derived quantities were sometimes negative, in most cases close to zero, was due to a poor MC approximation. Negative indexes are not unusual when the contribution of the inputs is negligible and can be avoided by increasing the number of MC simulations[20].

3.3 Submerged infections and role of the IFR
The number of new notified infections estimated by the SIRD model should be interpreted as inclusive of the undetected cases, thus one would expect that it was an upper bound for the observed number of new notified cases. However, when fixing $p = 1.15\%$, for some regions the observed number of new notified infections sometimes exceeded the number of new infections estimated by calibrating the SIRD model on the observed COVID19-related deaths (Supplementary Figure S3). This paradoxical result could be partly due to systematic errors (e.g. notifications concentrated on particular days of the week) and delays in the notification of cases, but also suggests that an IFR lower than 1.15% could be more consistent with the observed infection dynamics.

As an example, in Figure 6 we compared the predicted number of new infections for Campania and Liguria under different IFR scenarios. As the value assumed for the IFR increased, the number of new infections estimated by the model became smaller, to the extent of being, in the case of Campania, lower than the observed number of new notified infections when $p = 1.15\%$ and $p = 1.79\%$. On the contrary, for Liguria the estimated curves were overall consistent with the observed number of new notified infections, regardless of the IFR used in the analysis. The only exception was for $p = 1.79\%$, when the observed number of new cases notified during August and September slightly exceeded the model predictions.
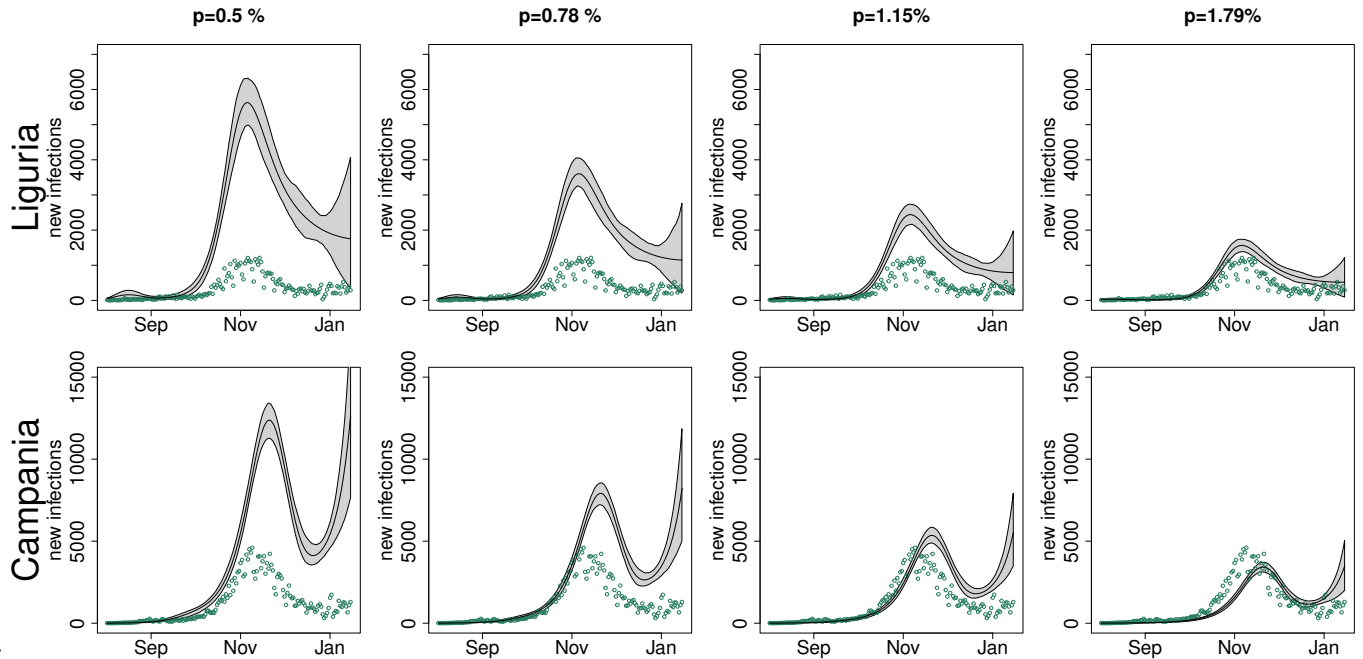
Figure 6 Estimated number of new infections (black line) in Liguria and Campania with pointwise 90% confidence bands, when setting in the SIRD model the infection fatality rate *p* to 0.5%, 0.78%, 1.15%, 1.79%, along with the observed number of new notified infections (green points), by region; $T = 14$ days.

4 Discussion

In this paper, we used data publicly available to study the SARS-CoV-2 epidemic dynamics in Italy and to investigate regional heterogeneity. We conducted our analyses on the time window corresponding to the second epidemic wave in Italy, which did not include the first months of 2021, when an extensive vaccination campaign began in the country. This permitted us to specify regional models that did not account for the immunity progressively acquired by part of the population.

A special focus was on the instantaneous reproductive number, $R_0(t)$, that we modeled through a regression spline in order to capture how the strength of contagion varied over time. $R_0(t)$ describes the speed of the contagion and depends directly or indirectly on countless factors, including virus infectiousness, socio-demographic and economic characteristics of the population, the efficacy of contact-tracing procedures, restriction policies. Being directly related to the number of contacts across the population, it is very sensitive to the introduction of social distancing measures. Alternative methods exist to estimate $R_0(t)$[25]; the main advantage of estimating $R_0(t)$ through compartmental models is that they allow quantifying also the number of incident and prevalent infections over time, and, more in general, the size of all defined compartments. Additionally, they permit to estimate $R_0(t)$ relying on the observed time series of daily COVID19-related deaths, which, in our context, was the most reliable quantity among those reported in the Protezione Civile database. Even though we cannot exclude a certain amount of mortality under-reporting, this was probably negligible during the second epidemic wave.

4.1 Results of the main analysis

Our results indicated that in Italy, during the second SARS-CoV-2 epidemic wave, the instantaneous reproductive number changed over time heterogeneously across regions, but with some important common elements. Among them, the increase of $R_0(t)$ during September, stronger in the Northern and Central regions than in Southern ones, which led to a dangerously high level of contagion in mid-October, followed by a decline of the contagion spread. We cannot exclude that this pattern was due to a phenomenon of seasonal variation typical of respiratory infections[36], but it was also suggestive of a possible role of school reopening in September in amplifying infection spreading, as argued also by other authors[37]. Although contagion within school has been sometimes considered as not higher than in other contexts[38,39], this issue is still debated[40]. In particular, reopening schools may have facilitated contagions through intensifying the number and the duration of interpersonal contacts within and outside schools, especially with the use of the public transport[1]. This seems to be confirmed by the meta-regression results: the regions where the percentage of students using public transport was higher tended to be characterized by higher average levels of $R_0(t)$. Also meteorological conditions (temperature, humidity and UV radiations) may have had an impact on

contagion spread through modulating SARS-CoV-2 infectiousness[36,41–43]. Additionally, with the arrival of the autumn season, recreational and sport activities moved, as usual, to closed places, with a consequent increase of the contagion risk[44,45]. The possible role of ambient temperature as a moderator of the contagion was suggested also by the meta-regression: higher average regional temperatures were associated with lower values of the average $R_0(t)$ from October to December. A factor that could have contributed to reach the maximum of the $R_0(t)$ curve in October was also the full recovery of many work activities after the summer holidays. Finally, at the end of September (20th-21st) regional elections took place in seven regions (Toscana, Marche, Campania, Puglia, Veneto, Liguria, Valle d'Aosta) and administrative elections took place in 1'184 municipalities around the country. The fact that, as per tradition, voting stations have been set up in school buildings during the weekend and that people moved within and between regions to reach their places of residence in order to vote should not be a priori ruled out as a possible source of contagion amplification, as documented elsewhere[46,47].

The decline of the $R_0(t)$ curves after the October peak was likely related to the restrictions progressively put in place by the central government. This decline tended to be steeper in the regions earlier classified as high-risk zones, where stronger restrictions were immediately introduced. However, it should be also noticed that, especially in the Northern and Central regions, the decline apparently started before the introduction of restrictions on November 6th. This could be indicative of the efficacy of local measures introduced in some regions before the national ones or reflect spontaneous changes in behavior of the population in the face of worrying levels of contagion.

At the end of November, the instantaneous reproductive number was below or very close to the threshold of 1 everywhere. Then, a new increase was observed during the second half of December. This was particularly evident in the Southern regions, probably due to travels within and between regions related to the Christmas holidays.

The "flat" $R_0(t)$ curve estimated for the Southern regions, with a maximum value slightly larger than 2, versus a maximum close to 3 in the rest of the country, was indicative of contagion dynamics in the South apparently different, possibly related to specific environmental, demographic and socio-economic features or a greater capacity of contagion control. Focusing on the average level of $R_0(t)$, we tried to explain through meta-regression part of the observed between-region heterogeneity. Our meta-regression results should not be used to draw conclusions about the existence of causal relationships linking the strength of contagion with the regional features, but suggest hypotheses that should be investigated through ad-hoc studies. The positive association of the average value of $R_0(t)$ with the employment rate and the schooling rate, and its negative association with the poverty index seem to indicate that population lifestyles typical of richer socio-economic contexts may have helped the virus spread. Also, population density, here measured in terms of percentage of people residing in high-density areas, was one of the possible predictors of a higher instantaneous reproductive number, as well as the tourism attractiveness of the region. However, in interpreting this latter result, one should consider that tourism flows have undergone considerable changes due to the COVID19 emergency and that in our meta-regression we included the tourism rate of the year 2018, the last available. As already discussed, our meta-regressions highlighted the potential role of school-related commuting in enhancing the contagion. We also found that the percentage of children attending kindergarten had a positive association with the $R_0(t)$ level. This result could simply reflect the fact that taking advantage of the kindergarten service is a proxy of the employment level in the region or could be indicative that kindergarten themselves had a role in spreading infections. Notice that kindergartens have been usually kept open during the second epidemic wave. Interestingly, we found that the average $R_0(t)$ was lower when the number of family physicians and pediatricians per inhabitants, as well as the proportion of public health care institutes, were higher. These results could indicate the important role of family medicine and public sanitary service in preventing contagion. In interpreting the meta-regression results, we should account that some of the meta-regressors followed a North-South gradient so that it was difficult to disentangle their role from that of possibly unobserved factors that varied with the latitude. Additionally, some observed associations could reflect a reverse causality phenomenon. Finally, we did not focus on environmental variables other than the regional average temperature in the study period. Other meteorological factors, e.g. solar radiation, humidity, and air pollution have been suggested as affecting the contagion rate as well as COVID19-related mortality or hospitalization, but exploring these complex relationships was outside the scope of our paper[48–50].

4.2 Sensitivity analyses

Compartmental models can be very complex when many compartments are defined and many transitions between them are allowed. Complexity goes hand in hand with an increase in the number of unknown parameters that cannot be estimated due to structural and practical identifiability problems. Most of the complicated compartmental models proposed in the literature rely on fixing the values of a large number of parameters without even studying the impact on the results of those arbitrarily chosen values[21,51].

In this paper, we performed both one-factor-at-time sensitivity analysis and a GSA. The two sensitivity analyses had a different interpretation and their results were not directly comparable. In fact, with the one-factor-at-time sensitivity analysis, we ignored how the inputs interacted and inspected the model outputs only for a few values of IFR and infection duration. On the contrary, with the GSA, we explored the whole space of the unestimated parameters of the SIRD modelThe GSA indicated

that there was interaction among the inputs and that there were no negligible inputs. The initial number of infected i_0 resulted to be the less influential fixed parameter of the SIRD model. Notice that we were interested in evaluating the relevance of i_0 since in the main analysis we forced it to be equal to the notified number of circulating infections on July 31st, thus assuming that the regional screening and tracing systems were initially able to detect all new infections. A second important result was related to the model simplification that we adopted in the regional SIRD models, assuming $T_D = T_R = T$: this simplification was admissible if the infection duration in the SIRD was set to plausible values for T_R. On the other hand, T_R resulted to be the most influential parameter, suggesting that an accurate knowledge of T_R would be needed for better predictions and understanding of the pandemic. Unfortunately, improving the empirical evidence on T_R is not trivial, because of its complex nature (T_R is both the average time from infection onset to recovery and the average time of infectiousness) and because of its dependence on undetected infections. Finally, the GSA confirmed the robustness of the estimated $R_0(t)$ curve to the value assigned to p, already documented elsewhere[29]. Under the reasonable assumption that the average infection time was homogeneous across the regions (the infection time is mainly related to virus and disease characteristics) and considering that changing the value of T the regional $R_0(t)$ curves inflated/deflated but preserved their overall shape, this robustness should assure that the conclusions drawn from the comparison of the regional $R_0(t)$ curves were not affected by specific choices of IFR and T.

4.3 Submerged infections and role of the IFR

Our model produced regional estimates of the number of circulating infections, inclusive of the submerged fraction of the contagion. This quantity, combined with $R_0(t)$, determines the number of new infections, and it is useful to assess the actual impact of the contagion on the health care system: if a large instantaneous reproductive number can be sustainable at the beginning of the epidemic when the number of cases is low, a low instantaneous reproductive number may not be sustainable when the number of circulating infected is large.

For identifiability reasons, we did not estimate the IFR, but our analysis provided indirectly indications about plausible values for this parameter through a comparison between the number of infections predicted by the SIRD model and the observed cases reported by Protezione Civile[18]. Under the assumption of an average waiting time from infection onset to infection resolution (death or recovery) equal to 14 days, our findings supported the hypothesis that during the study period the IFR varied over time and space. In particular, according to our results, the IFR was likely lower than 1.15% in the initial phase of the second epidemic wave, when the virus mainly circulated among younger people[52], as well as in many Southern regions, including Campania, which is, not by chance, the region with the lowest aging index in Italy[22]. Conversely, an IFR equal to 1.79%, the upper bound considered in our analysis, was not consistent with the observed infection dynamics in most regions, because it led to a predicted number of infections smaller than the observed number of notified cases.

Fixing $p = 0.5\%$, not far from the IFR estimate reported in the review by Ioannidis[53], we obtained quite high estimates of infection prevalence which could be reasonably considered as an extreme - even if not impossible - lower bound.

Notice that, in principle, the observed inconsistency between predicted infections and notified infections in the Southern regions could partly be solved by assuming for them a longer average waiting time T (e.g. $T = 18$ days), with $p = 1.15\%$. Indeed, as shown by the sensitivity analysis results, the parameters p and T jointly affect the epidemic curve generated by the SIRD. However, as already explained, this hypothesis seems less plausible than the hypothesis of a heterogeneous IFR.

4.4 Study limitations

Our study has some limitations. Regarding the regional analyses, the SIRD model relied on strong assumptions that could be partly unrealistic: the population was homogeneously mixed, with people making contact at random, and closed, with no contacts among individuals belonging to different regions or countries; transition parameters were constant across individuals who were present at the same time in the same compartment; there was not reinfection and no incubation period. Importantly, an individual becoming infected on the day t was supposed to be infectious starting from the day $t + 1$ until infection resolution. This was quite an unrealistic assumption for the infected individuals that were notified, thus isolated. Additionally, as usual in compartmental models in their simplest form, we implicitly assumed that the transition times were exponentially distributed[54], inducing a non-negligible probability that the infection duration was much longer than the average, as well as a high probability of very short waiting times. As already discussed in Section 4.3, also the assumption of an IFR which is constant over time was questionable, in the light of the comparison between the observed notified cases and the infections predicted by the SIRD model. An additional limitation concerns the fact that we calibrated the SIRD models only on the observed COVID19 related deaths, without exploiting the availability of other observed time series, like the ones of notified infected. Calibrating on the notified infected would have required the formulation of a more complex compartmental model with separate compartments for detected and undetected cases and the introduction of additional unknown transition parameters[2–4, 29]. On the other hand, this more complex model would have also allowed us to take into account the fact that the detected and isolated infected individuals are not able to spread the contagion as the undetected ones.

Additionally, the regional $R_0(t)$ curves could be quite sensitive to spline knots position. One could extend the procedure to use penalized splines[26].

Following a procedure which is frequently used in compartment models literature[32,55], we did not make assumptions about data distribution in the phase of estimation of the SIRD parameters, but we generated bootstrap samples for confidence intervals construction by assuming a Negative Binomial distribution on the daily increments of deaths. Alternative approaches rely on likelihood maximization or Bayesian inference, but compartmental models often exhibit complex likelihood requiring particle filtering methods to be maximized[56] or computationally intensive methods based on data augmentation procedures[57]. In the Bayesian framework, likelihood-free approaches have been also proposed[58].

The bootstrap percentile intervals could have lower coverage than the nominal one[59]. Bias-corrected and accelerated bootstrap confidence intervals could have better performance[60], but they are too computationally demanding to be applied in this context.

Finally, regarding meta-regression, a finer geographical detail, e.g. considering provinces or municipalities instead of regions, would have allowed us to better appreciate the sources of heterogeneity in contagion spread. However, estimating $R_0(t)$ at a provincial or at finer level would have led to a more unstable inference. Not to mention that the assumption of a close population would be, at this geographical detail, even more exceptionable.

5 Conclusions

Despite the difficulty of drawing information from poor and limited data, our approach allowed us to estimate $R_0(t)$ and get an evaluation of the prevalence of infections in Italy at a regional level. Besides the common elements – including a mid-October peak and a decline during the month of November, steeper in those regions where stronger restriction measures were earlier applied – the rate of contagion changed over time heterogeneously across regions. This proves that models treating the phenomenon mostly at a national level could ignore important characteristics, specific to certain areas. Meta-regression results show that heterogeneity can be partly explained by socioeconomic and demographic factors, like the grade of urbanization, family medicine and health care system, employment rate, use of public transport for school commuting. Higher temperatures were associated to lower $R_0(t)$ levels. These factors should be further explored with analyses at a finer geographical scale.

Although we fixed some parameters to make the model identifiable, the results of the sensitivity analyses reassured us that the overall shape of the estimated $R_0(t)$ curves and the conclusions drawn from comparisons among regional curves were robust to changes of those parameters. On the contrary, the estimate of the prevalence of infections was strongly related to changes in the IFR. Moreover, our findings seemed to support the hypothesis of a heterogeneous IFR over time and among regions, that could have been lower than 1% in some regions, especially at the beginning of the epidemic wave.

Regarding sensitivity analysis, we want to stress that GSA has been proved to be a powerful and promising tool and its use should be encouraged in this and other research contexts.

Data availability

The datasets used for the current study are available in the Protezione Civile github repository, https://github.com/pcm-dpc/COVID19. All methods were carried out in accordance with relevant guidelines and regulations.

References

1. Chirico, F., Sacco, A., Nucera, G. & Magnavita, N. Coronavirus disease 2019: the second wave in Italy. *J. Heal. Res.* **35**, 359–363 (2021).

2. Gaeta, G. A simple SIR model with a large set of asymptomatic infectives. *Math. Eng.* **3**, 1–39 (2021).

3. Romano, S., Fierro, A. & Liccardo, A. Beyond the peak: A deterministic compartment model for exploring the Covid-19 evolution in Italy. *PLoS ONE* **15**, 1–23 (2020).

4. Russo, L. *et al.* Tracing day-zero and forecasting the COVID19 outbreak in Lombardy, Italy: A compartmental modelling and numerical optimization approach. *PLoS ONE* **10**, e0240649 (2020).

5. Faranda, D. & Alberti, T. Modeling the second wave of COVID19 infections in France and Italy via a stochastic SEIR model. *Chaos: An Interdisc. J. Nonlinear Sci.* **30**, 111101 (2020).

6. Alaimo Di Loro, P. *et al.* Nowcasting COVID19 incidence indicators during the Italian first outbreak. *Stat. Med.* **40**, 3843-3864 (2021).

7. Farcomeni, A., Maruotti, A., Divino, F., Jona-Lasinio, G. & Lovisong, G. An ensemble approach to short-term forecast of COVID-19 intensive care occupancy in Italian regions. *Biom J.* **63**, 503-513 (2021).

8. Campi, G. *et al.* Metastable states in plateaus and multi-wave epidemic dynamics of COVID-19 spreading in Italy. *Sci Rep* **11**, 12412 (2021).

9. Quaranta, G., Formica, G., Machado, J. T., Lacarbonara, W. & Masri, S. F. Understanding COVID-19 nonlinear multi-scale dynamic spreading in Italy. *Nonlinear Dyn.* **10**, 1583–1619 (2020).

10. Ferrari, L. *et al.* Modeling provincial COVID-19 epidemic data using an adjusted time-dependent SIRD model. *Int. J. Environ. Res. Public Heal.* **18**, 6563 (2021).

11. Bertuzzo, E. *et al.* The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nat. Commun.* **11**, 4264 (2020).

12. Gatto, M. *et al.* Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *Proc. Natl. Acad. Sci.* **117**, 10484–10491 (2020).

13. Vinceti, M., Filippini, T., Rothman, K. J., Di Federico, S. & Orsini, N. SARS-CoV-2 infection incidence during the first and second COVID-19 waves in Italy. *Environ Res* **197**, 111097 (2021).

14. Coccia, M. Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID-19. *Sci Total. Environ* **729**, 138474 (2020).

15. Boschi, T., Di Iorio, J., Testa, L., Cremona, M. A. & Chiaromonte, F. Functional data analysis characterizes the shapes of the first COVID-19 epidemic wave in Italy. *Sci. Reports* **11**, 17054 (2021).

16. Rovetta, A. & Castaldo, L. Relationships between demographic, geographic, and environmental statistics and the spread of novel Coronavirus disease (COVID-19) in Italy. *Cureus* **12**, e11397 (2020).

17. Gnanvi, J. E., Salako, K. V., Kotanmi, G. B. & Glèlè Kakaï, R. On the reliability of predictions on COVID-19 dynamics: A systematic and critical review of modelling techniques. *Infect. Dis. Model.* **6**, 258–272 (2021).

18. Protezione Civile. Repository of COVID-19 outbreak data for Italy. https://github.com/pcm-dpc/COVID-19.

19. Wu, J., Dhingra, R., Gambhir, M. & Remais, J. V. Sensitivity analysis of infectious disease models: methods, advances and their application. *J R Soc Interface* **10**, 20121018 (2013).

20. Saltelli, A. *et al. Global Sensitivity Analysis: The Primer* (Wiley, 2008).

21. Saltelli, A., Bammer, G., Bruno, I. & et al. Five ways to ensure that models serve society: a manifesto. *Nature* **582**, 482–484 (2020).

22. ISTAT. Bilancio demografico nazionale. Anno 2019. https://www.istat.it/it/archivio/245466 (2020).

23. Lin, F., Muthuraman, K. & Lawley, M. An optimal control theory approach to non-pharmaceutical interventions. *BMC Infect Dis* **10**, 1–13 (2010).

24. Legrand, J., Grais, R. F., Boelle, P.-Y., Valleron, A.-J. & Flahault, A. Understanding the dynamics of Ebola epidemics. *Epidemiol. & Infect.* **135**, 610–621 (2007).

25. O'Driscol, M., Harry, C., Donnelli, C., Cori, A. & Dorigatti, I. A comparative analysis of statistical methods to estimate the reproduction number in emerging epidemics with implications for the current COVID-19 pandemic. *Clin Infect Dis* ciaa1599 (2020).

26. Ruppert, D., Wand, M. & Carroll, R. *Semiparametric Regression*. Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press, 2003).

27. Brazeau, N. F. *et al. COVID-19 infection fatality ratio: estimates from seroprevalence.* (Imperial College London (29-10-2020), 2020).

28. Villa, M. *Coronavirus: la letalità in Italia, tra apparenza e realtà.* (ISPI, Milano, 2020).

29. Baccini, M., Cereda, G. & Viscardi, C. The first wave of the SARS-CoV-2 epidemic in Tuscany (Italy): a SI^2R^2D compartmental model with uncertainty evaluation. *PLoS ONE* **16**, e0250029 (2021).

30. ISS, SARS-CoV-2 Surveillance Group. Characteristics of SARS-CoV-2 patients dying in Italy report based on available data on December 9th 2020. Tech. Rep., ISS (2020).

31. Byrne, A. W. *et al.* Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open* **10**, e039856 (2020).

32. Chowell, G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infect. Dis. Model.* **2**, 379–398 (2017).

33. Grenfell, B. T., Bjørnstad, O. N. & Finkenstädt, B. F. Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the TSIR model. *Ecol. Monogr.* **72**, 185–202 (2002).

34. Sobol, I. M. Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.* **1**, 407–414 (1993).

35. Lamboni, M., Monod, H. & Makowski, D. Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. *Reliab. Eng. Syst. Saf.* **96**, 450–459 (2011).

36. Choi, Y.-W., Tuel, A. & Eltahir, E. A. B. On the environmental determinants of Covid-19 seasonality. *GeoHealth* **5**, e2021GH000413 (2021).

37. Larosa, E. *et al.* Secondary transmission of COVID-19 in preschool and school settings in northern Italy after their reopening in September 2020: a population-based study. *Eurosurveillance* **25**, 2001911 (2020).

38. Buonsenso, D., De Rose, C., Moroni, R. & Valentini, P. SARS-CoV-2 infections in italian schools: Preliminary findings after 1 month of school opening during the second wave of the pandemic. *Front. Pediatr.* **8**, 909 (2021).

39. Gandini, S. *et al.* A cross-sectional and prospective cohort study of the role of schools in the SARS-CoV-2 second wave in Italy. *The Lancet Reg. Heal. – Eur.* **5** (2021).

40. Tupper, P. & Colijn, C. COVID-19 in schools: Mitigating classroom clusters in the context of variable transmission. *PLoS Comput. Biol.* **17**, 1–20 (2021).

41. Wu, Y. *et al.* Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Sci. The Total. Environ.* **729**, 139051 (2020).

42. Moriyama, M., Hugentobler, W. J. & Iwasaki, A. Seasonality of respiratory viral infections. *Annu. Rev. Virol.* **7**, 83–101 (2020).

43. Chin, A. W. H. *et al.* Stability of SARS-CoV-2 in different environmental conditions. *The Lancet Microbe* **1**, e10 (2020).

44. Li, Y. *et al.* Probable airborne transmission of SARS-CoV-2 in a poorly ventilated restaurant. *Build. Environ.* **196**, 107788 (2021).

45. Rencken, G. K., Rutherford, E. K., Ghanta, N., Kongoletos, J. & Glicksman, L. Patterns of SARS-CoV-2 aerosol spread in typical classrooms. *Build Environ* **204**, 108167 (2021).

46. Palguta, J., Levínský, R. & Škoda, S. Do elections accelerate the COVID-19 pandemic?: Evidence from a natural experiment. *J Popul Econ* **35**, 1–44 (2021).

47. Cotti, C., Engelhardt, B., Foster, J., Nesson, E. & Niekamp, P. The relationship between in-person voting and COVID-19: Evidence from the Wisconsin primary. *Contemp Econ Policy* **39**, 760-777 (2021).

48. Rahimi, N. R. *et al.* Bidirectional association between COVID-19 and the environment: A systematic review. *Environ. Res.* **194**, 110692 (2021).

49. Xu, R. *et al.* Weather, air pollution, and SARS-CoV-2 transmission: a global analysis. *Lancet Planet Heal.* **5**, e671–e680 (2021).

50. Santurtún, A. *et al.* Exposure to particulate matter: Direct and indirect role in the COVID-19 pandemic. *Environ Res* 112261 (2021).

51. James, L. P., Salomon, J. A., Buckee, C. O. & Menzies, N. A. The use and misuse of mathematical modeling for infectious disease policymaking: Lessons for the COVID-19 pandemic. *Med. Decis. Mak.* **41**, 379–385 (2021).

52. Gruppo di lavoro AIE. La progressione della seconda ondata in italia e nelle regioni per classi di età. *Scienza rete* (2020).

53. Ioannidis, J. The infection fatality rate of COVID-19 inferred from seroprevalence data *Bull. World Heal. Organ.* **99**, 19-33F (2021).

54. Brauer, F., van den Driessche, P. & Wu, J. *Mathematical Epidemiology*, chap. Compartmental models in epidemiology, 19–79 (Springer Berlin Heidelberg, 2008).

55. Roosa, K. & Chowell, G. Assessing parameter identifiability in compartmental dynamic models using a computational approach: application to infectious disease transmission models. *Theor. Biol. Med. Model.* **16**, 1–15 (2019).

56. Ionides, E. L., Bretó, C. & King, A. A. Inference for nonlinear dynamical systems. *Proc. Natl. Acad. Sci.* **103**, 18438–18443 (2006)

57. O'Neill, P. D. A tutorial introduction to Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods. *Math. Biosci.* **180**, 103–114 (2002).

58. Kypraios, T., Neal, P. & Prangle, D. A tutorial introduction to Bayesian inference for stochastic epidemic models using approximate Bayesian computation. *Math. Biosci.* **287**, 42 – 53 (2017).

59. Efron, B. & Tibshirani, R. *An Introduction to the Bootstrap* (CRC press, Boca Raton, 1994).

60. Kim, M., Kim, Y. & Schmidt, P. On the accuracy of bootstrap confidence intervals for efficiency levels in stochastic frontier models with panel data. *J. Prod. Analysis* **28**, 165–181 (2007).

Acknowledgements

Author contributions statement

All authors conceived the study, performed the statistical analyses, discussed the results and wrote the paper. M.B. supervised the project.

Competing interests

The authors have not competing interests.

Additional information

Requests for materials should be addressed to G.C.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- suplfilenuovo.pdf