Similarities and differences of the COVID-19 second surge in Europe and the Northeast United States

Jose M. G. Vilar[1,2,*] and Leonor Saiz[3,*]

[1] Biofisika Institute (CSIC, UPV/EHU), University of the Basque Country (UPV/EHU), P.O. Box 644, 48080 Bilbao, Spain

[2] IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain

[3] Department of Biomedical Engineering, University of California, 451 E. Health Sciences Drive, Davis, CA 95616, USA

[*]Correspondence to: j.vilar@ikerbasque.org or lsaiz@ucdavis.edu

ABSTRACT

We estimated the aggregate prevalence over time for Europe and the Northeast US to characterize the COVID-19 second surge in these regions. We find a starting date as early as July 3 (95 %CI: July 2 - July 5) for Europe (Schengen Area) and August 19 (95 %CI: August 16 - August 22) for the Northeast US; subsequent infectious populations that, as of December 7, have always increased or remained stagnant; and the resurgences being the collective effect of each overall region with no one country or state dominating the dynamics by itself.

INTRODUCTION

The Northeast United States (US) and Europe are progressing through a second surge of the COVID-19 outbreak after experiencing a similar initial outburst and subsequent regression. None of these resurgences was widely expected nor anticipated [1, 2]. Both areas display high mobility among their locations and broad independence among locations to enact measures to mitigate the propagation of the outbreak. In the case of Europe, Schengen Area countries allow for unrestricted border crossings among them. The characterization of the similarities and differences of the outbreak progression in these two regions is needed to provide insights into the effectiveness of the actions taken, to ascertain the extent their results can be extrapolated from one region to another, and to informedly mitigate the current and potentially forthcoming resurgences.

We computed the aggregate values of the infectious populations and the corresponding time-varying reproduction numbers for the Northeast US and Europe's Schengen Area from the individual values of the locations of each region [3]. Reproduction numbers were computed from growth rates considering a gamma-distributed generation interval with a mean of 6.5 days and a standard deviation of 4.2 days (Methods) [4]. Aggregate values were computed from infectious-population-weighted individual location values (Methods). The individual location data was downloaded on December 31, 2020 from https://github.com/Covid19Dynamics/trajectories.

RESULTS

The initial progression of the overall infectious populations for both regions consisted of exponential growth followed by an exponential decay (Figures 1A and 1B). Subsequently, there was a sharp transition to fast exponential growth in Europe on July 3 (95% Confidence Interval (CI): July 1–July 6) and a stagnant overall infectious population in the Northeast US, which started to grow slowly but with increasing speed on August 19 (95% CI: August 16 – August 23). Concomitantly, the time-varying reproduction numbers crossed above one on the resurgence dates less abruptly in the Northeast US than in Europe (Figure 1C), reaching maximum values of 1.50 (95% CI: 1.48–1.51) in Europe and 1.30 (95% CI: 1.27–1.34) in the Northeast US. The sharp resurgence to exponential growth in Europe is coincidental with lifting major nonpharmaceutical interventions that curved the outbreak [5], including the coordinated end of travel bans in Schengen Area's countries on July 1, 2020 [6]. As of December 7, the estimated infectious population is still in a growing phase but slowing down in the Northeast US (Figure 1A) and has remained in a stagnant state since early November in Europe (Figure 1B).

The aggregated values for each region provide precise evidence of sustained growth of the outbreaks already over the summer, despite the uncertainty and variability present in each of the locations independently (Figure 1D). There is also robust evidence that the resurgence was not driven by a unique location since any aggregate value of the starting date for each region leaving one of their locations out is within the confidence limits of that of the overall region (Methods, Figure 1D, and Supplementary Figure 1). Therefore, the resurgence was the collective effect of each overall region.

DISCUSSION

These results highlight the need to implement policies and surveillance approaches that include data at a supra-location level when there is high mobility among locations. In this regard, the Northeast US, as a region, is closely trailing Europe in the second surge of the outbreak, but with a markedly smaller growth and evidence of slowing down earlier in the growth phase than Europe. Key differences in the actions taken included more gradual lifting and swifter progressive reimplementation of measures in the Northeast US than in Europe [5]. The progression over time of the aggregate prevalence of Europe's Schengen Area countries shows, with high reliability, that Europe's initial acting upon the second surge in mid-late October [1] took place well after a three-month-long period of sustained growth of the COVID-19 infectious population in the overall region, which has resulted in a second surge deadlier than the first one [7]. Such a death toll has not taken place yet in the Northeast US and it is potentially avertable with the social restrictions currently in place [5] according to the observed trend in the progression of its estimated overall aggregate prevalence.

METHODS

The aggregate infectious population for a region is expressed as

$$n_I(t) = \sum_j n_{I,j}(t),$$

where $n_{I,j}(t)$ is the infectious population of the specific location with index j. The corresponding upper and lower confidence intervals are computed as

$$n_I^U(t) = n_I(t) + \sqrt{\sum_j \left(n_{I,j}^U(t) - n_{I,j}(t)\right)^2}$$

and

$$n_I^L(t) = n_I(t) - \sqrt{\sum_j \left(n_{I,j}(t) - n_{I,j}^L(t)\right)^2}$$

from the upper, $n_{I,j}^U(t)$, and lower, $n_{I,j}^L(t)$, confidence intervals for each location.

The overall time-dependent growth rate is given by

$$k_G(t) = \sum_j k_{G,j}(t) \frac{n_{I,j}(t)}{n_I(t)},$$

where $k_{G,j}(t)$ is the growth rate of the infectious population of the specific location with index j. This expression follows from

$$\frac{d}{dt} n_I(t) = \sum_j \frac{d}{dt} n_{I,j}(t) = \sum_j k_{G,j}(t)\, n_{I,j}(t) = k_G(t) n_I(t).$$

The corresponding upper and lower confidence intervals are computed as

$$k_G^U(t) = k_G(t) + \frac{1}{n_I(t)} \sqrt{\sum_j \left(k_{G,j}^U(t) - k_{G,j}(t)\right)^2 n_{I,j}(t)^2}$$

and

$$k_G^L(t) = k_G(t) - \frac{1}{n_I(t)} \sqrt{\sum_j \left(k_{G,j}(t) - k_{G,j}^L(t)\right)^2 n_{I,j}(t)^2}$$

from the upper, $k_{G,j}^{U}(t)$, and lower, $k_{G,j}^{L}(t)$, confidence intervals for each location.

The time-varying reproduction number R_t is obtained from the growth rate $k_G(t)$ through the Euler–Lotka equation,

$$R_t^{-1} = \int_0^{\infty} f_{GT}(\tau) e^{-k_G(t)\tau} d\tau,$$

where $f_{GT}(\tau)$ is the probability density function of the generation time. We consider the usual description of generation times through a gamma distribution

$$f_{GT}(\tau) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta\tau},$$

which leads to

$$R_t = \left(1 + \frac{k_G(t)}{\beta}\right)^{\alpha}$$

for $k_G(t) > -\beta$ and $R_t = 0$ for $k_G(t) \leq -\beta$. The values of the parameters are given by $\alpha = \tau_G^2/\sigma_G^2$ and $\beta = \tau_G/\sigma_G^2$, where τ_G and σ_G^2 are the mean and the variance of the generation time, respectively.

The starting date of the second surge is computed as the date the infectious population reached a minimum value after the maximum of the first surge. The lower bound of its confidence interval is computed as the last day before the minimum in which the upper bound of the confidence interval of the reproduction number is below 1. Analogously, the upper bound of the confidence interval is computed as the first day after reaching the minimum in which the lower bound of the confidence interval of the reproduction number is above 1.

The trajectories of the infectious populations, the growth rates, and the confidence intervals for each location were downloaded on December 31, 2020 from https://github.com/Covid19Dynamics/trajectories.

ACKNOWLEDGMENTS

J.M.G.V. acknowledges support from Ministerio de Ciencia e Innovación under grant PGC2018-101282-B-I00 (MCI/AEI/FEDER, UE). L.S. acknowledges support from the University of California, Davis.

REFERENCES

1. Looi, M.-K., *Covid-19: Is a second wave hitting Europe?* BMJ, 2020. **371**: p. m4113.
2. Kim, J., *Is a Second Wave Starting in N.Y.C.?* The New York Times 2020: p. www.nytimes.com/2020/11/10/nyregion/second-coronavirus-wave-.html.
3. Vilar, J.M.G. and L. Saiz, *The evolving worldwide dynamic state of the COVID-19 outbreak.* medRxiv, 2020: p. 2020.11.26.20239434.
4. Bi, Q., et al., *Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study.* The Lancet Infectious Diseases, 2020. **20**(8): p. 911-919.
5. Hale, T., et al. *Oxford COVID-19 Government Response Tracker, Blavatnik School of Government*. 2020. www.bsg.ox.ac.uk/covidtracker.
6. *Council of the European Union. Council Recommendation on the temporary restriction on non-essential travel into the EU and the possible lifting of such restriction*. 2020. https://data.consilium.europa.eu/doc/document/ST-9208-2020-INIT/en/pdf.
7. *JHU CSSE COVID-19 Data*. December 31, 2020]; Available from: https://github.com/CSSEGISandData/COVID-19.

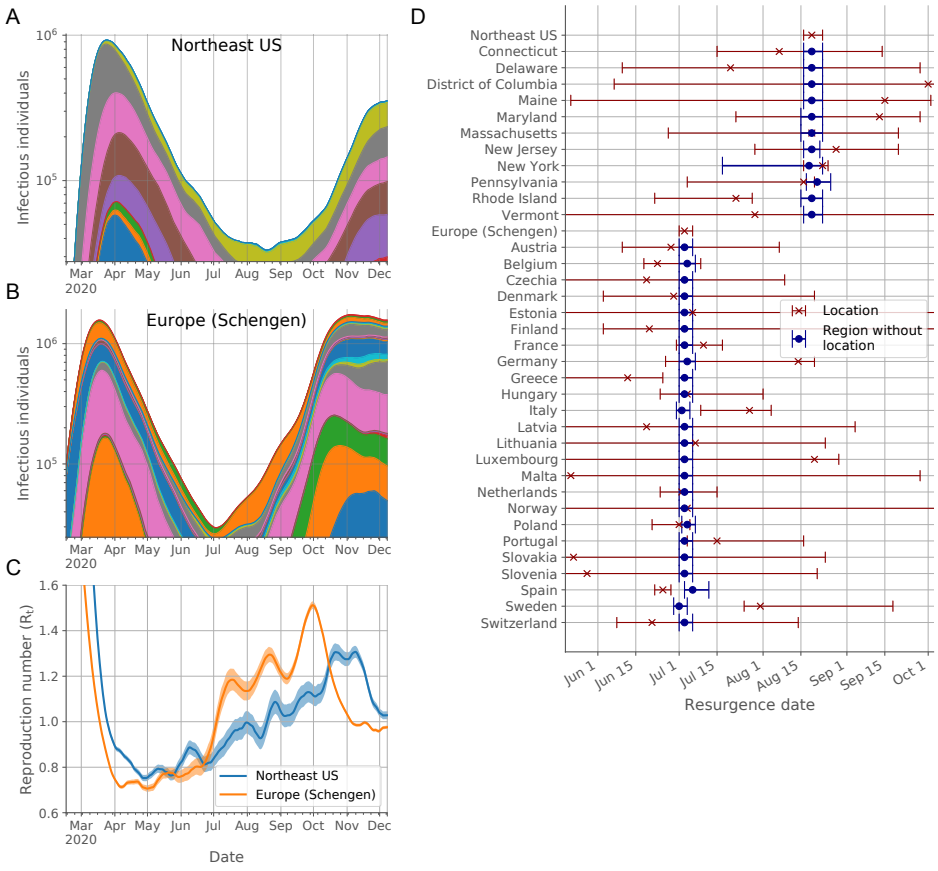


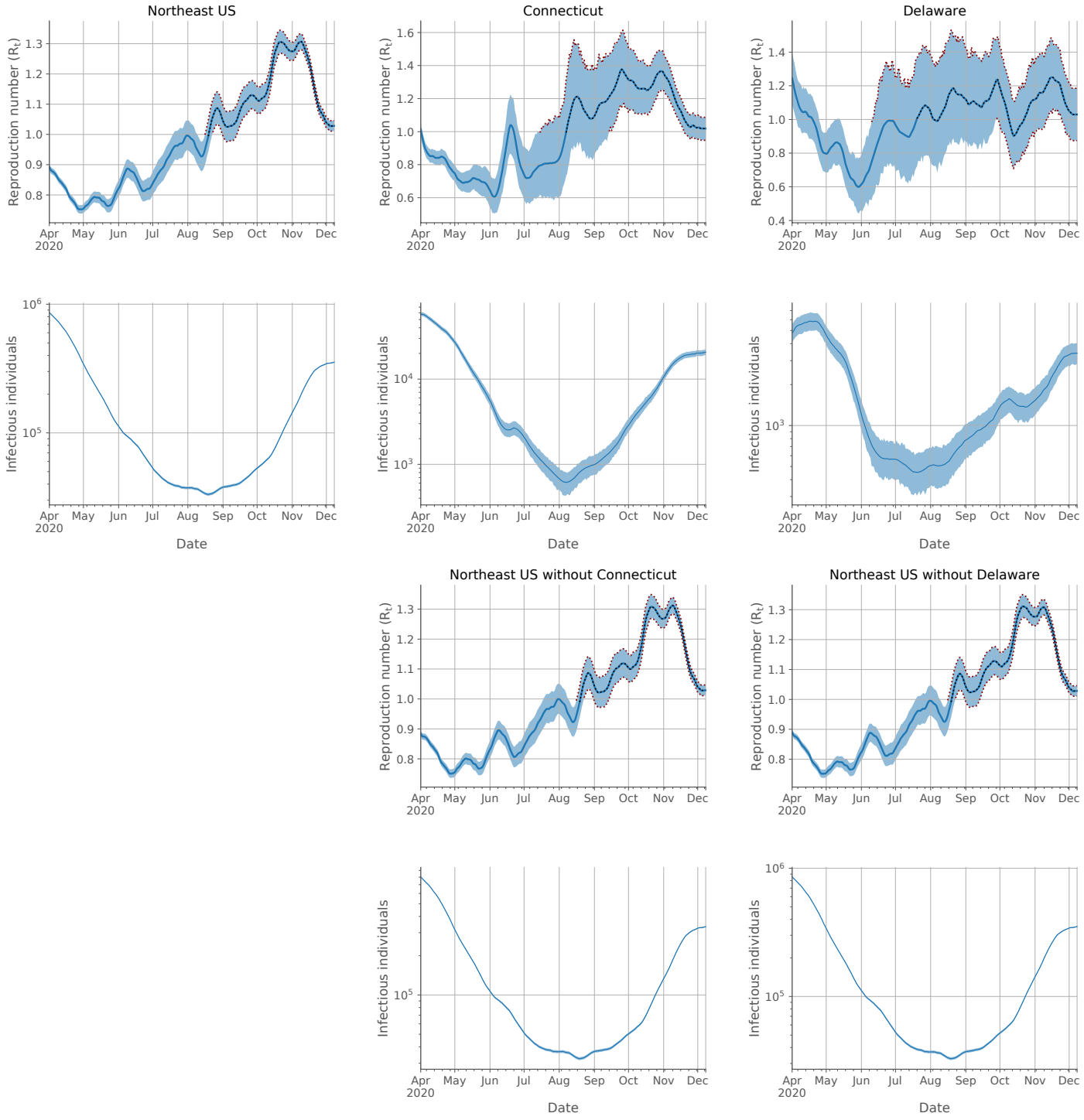
Figure 1. Temporal evolution of the COVID-19 outbreak in the Northeast US and Europe. Progression over time of the infectious population for locations in the Northeast US (A) and countries in Europe's Schengen Area (B), their corresponding reproduction numbers (C), and the timing of the resurgence in each location and the corresponding region without the location (D). Each colored section in the area plots represents the contribution of a state (A) or country (B) to the overall infectious populations. Countries and states are arranged in alphabetical order from bottom to top. The infectious populations are plotted on a logarithmic scale to highlight the triphasic behavior (growth-decay-growth) of the outbreak. The shaded regions in the reproduction number plots (C) represent the 95% CI. The

resurgence dates correspond to the minimum infectious population reached for each location and region (red crosses) and for the whole region without the population of the location indicated (blue circles) (D). The intervals represent the 95% CI.

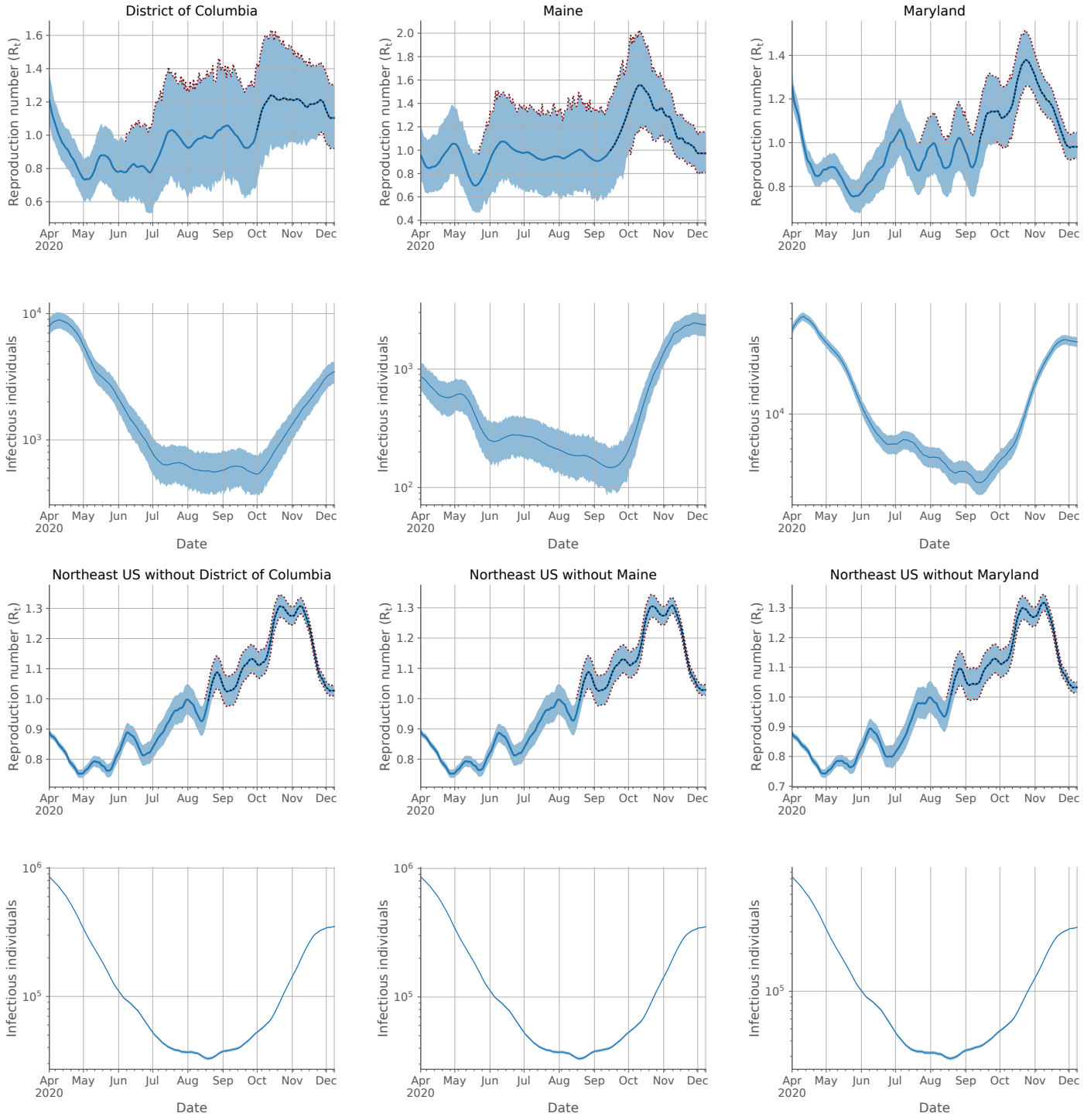
Supplementary Figure 1. Location-specific characterization of the temporal evolution of the COVID-19 outbreak for the states in the Northeast US and countries in Europe's Schengen Area.

(The 13-page long figure is appended after the legend)

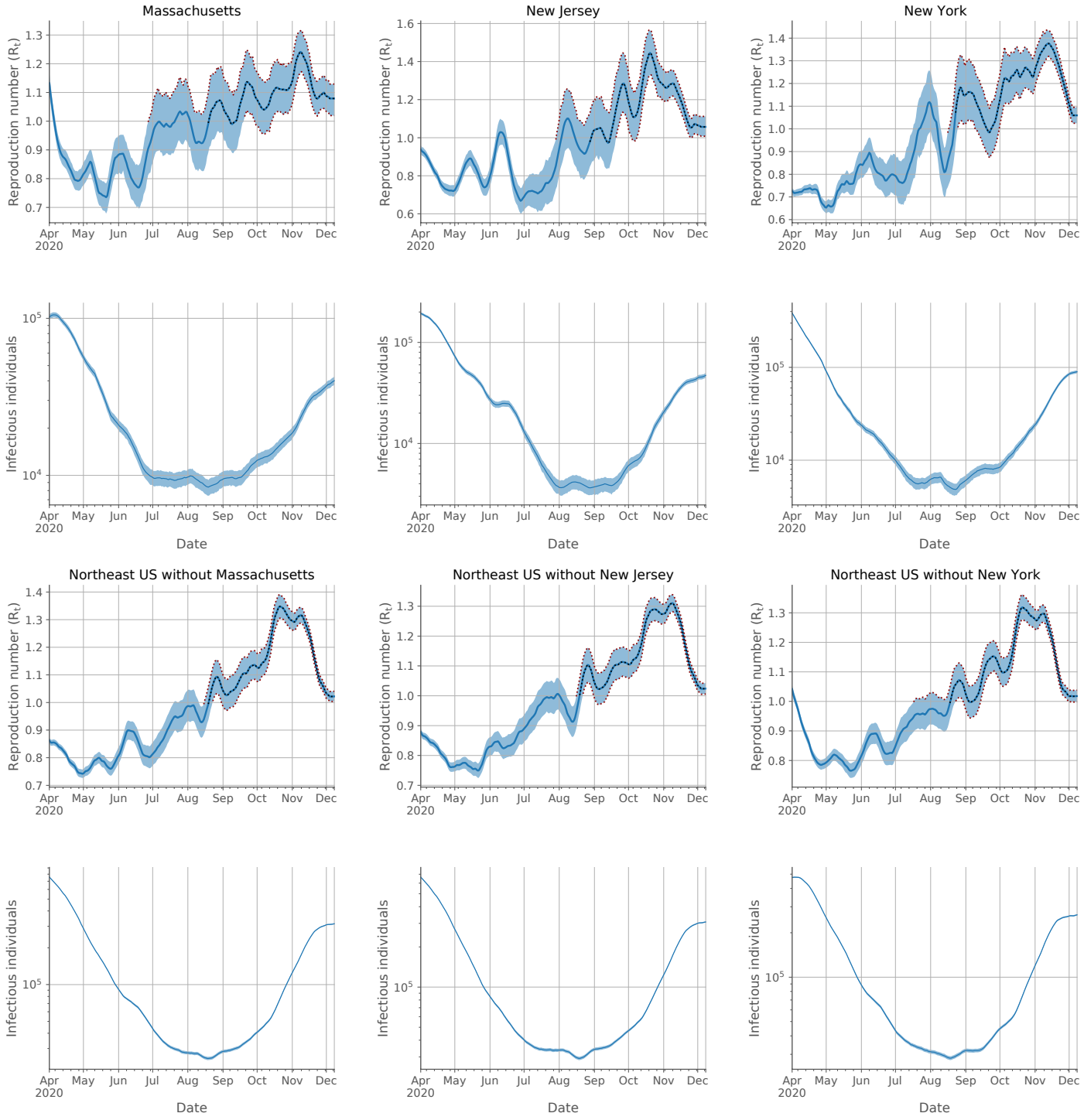
The data for each location and for the region without the location is shown in a column of 4 panels. The 1st and 3rd panels from the top show the temporal evolution of the reproduction number (blue line) with the shaded blue region indicating the 95% confidence intervals (CI). The dotted lines highlight the span of the second wave and its confidence interval over the reproduction number data. The 2nd and 4th panels from the top show the temporal evolution of the infectious population with the shaded blue region indicating the 95% CI. The 1st and 2nd panels correspond to the location and the 3rd and 4th, to the region without the location. Reproduction numbers were computed from growth rates considering a gamma-distributed generation interval with a mean of 6.5 days and a standard deviation of 4.2 days.



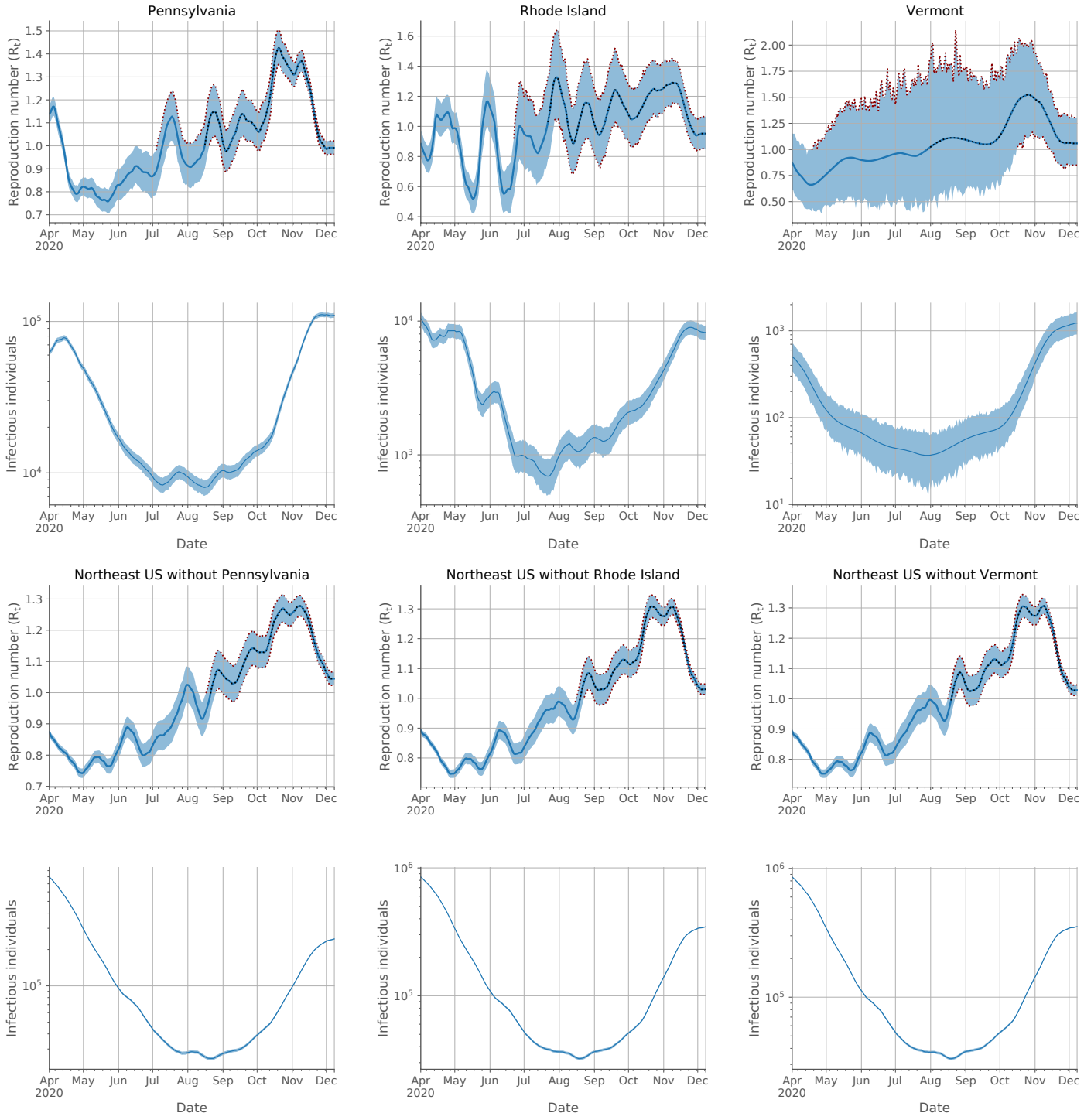
Supplementary Figure 1.1



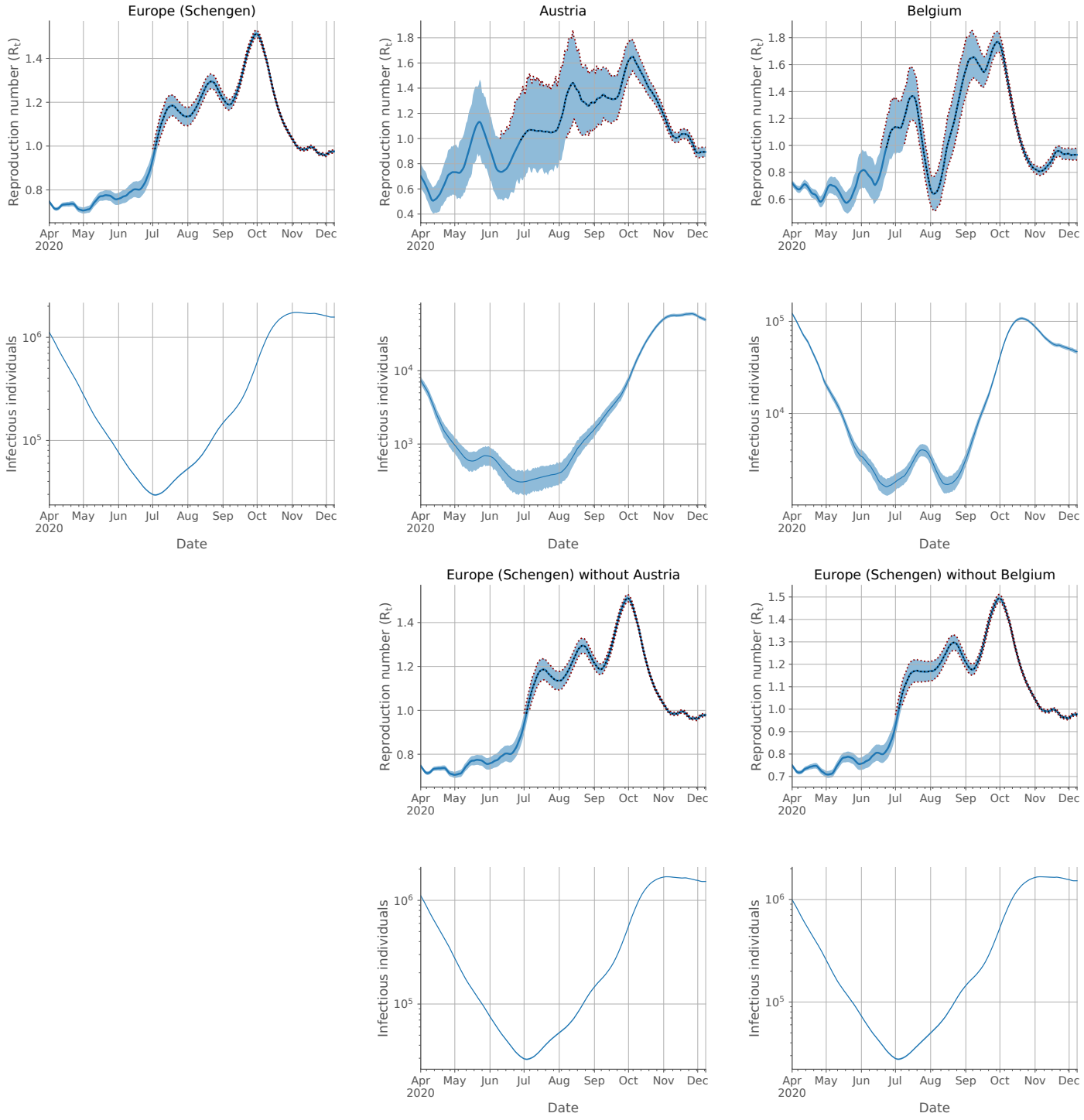
Supplementary Figure 1.2



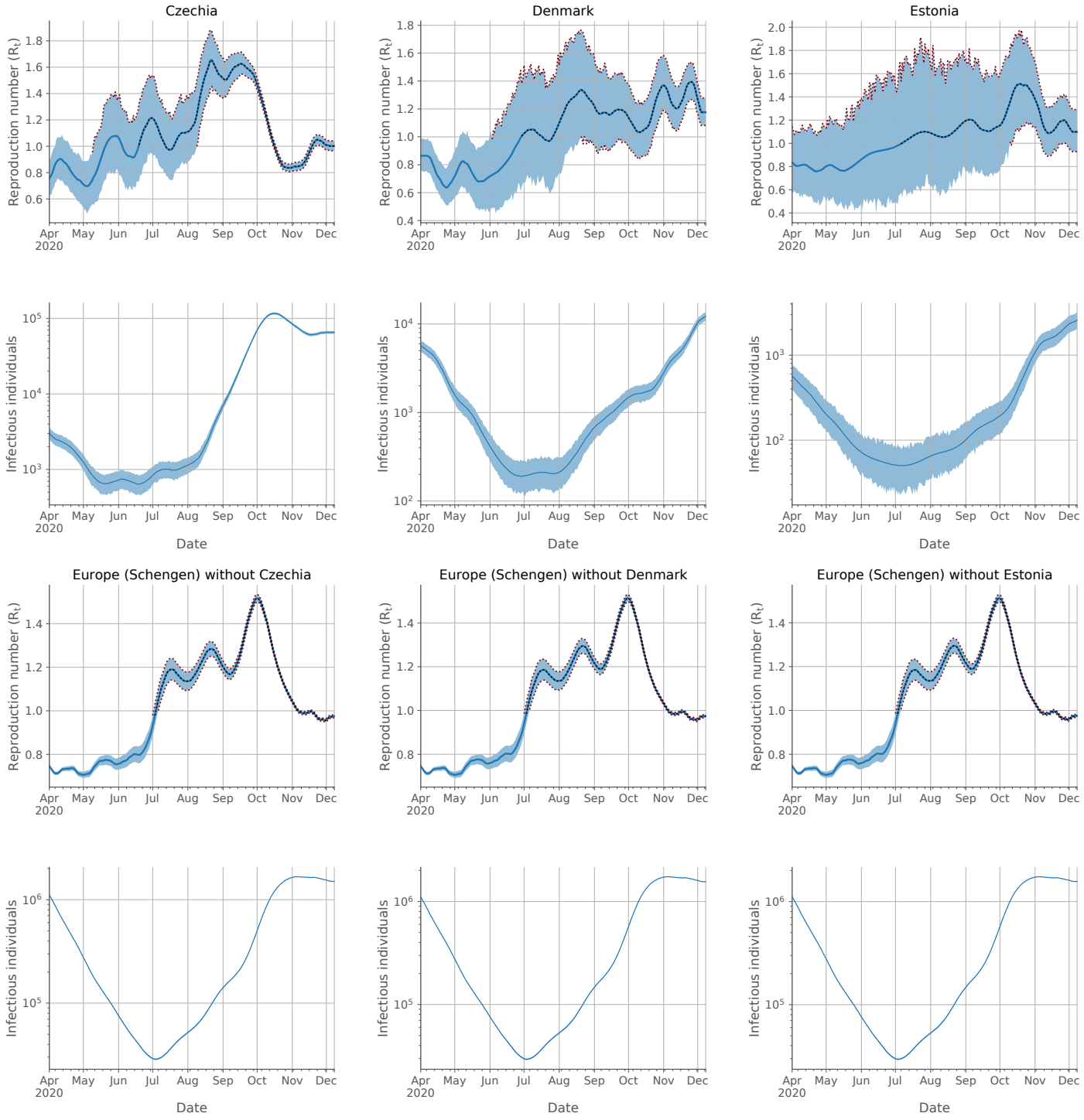
Supplementary Figure 1.3



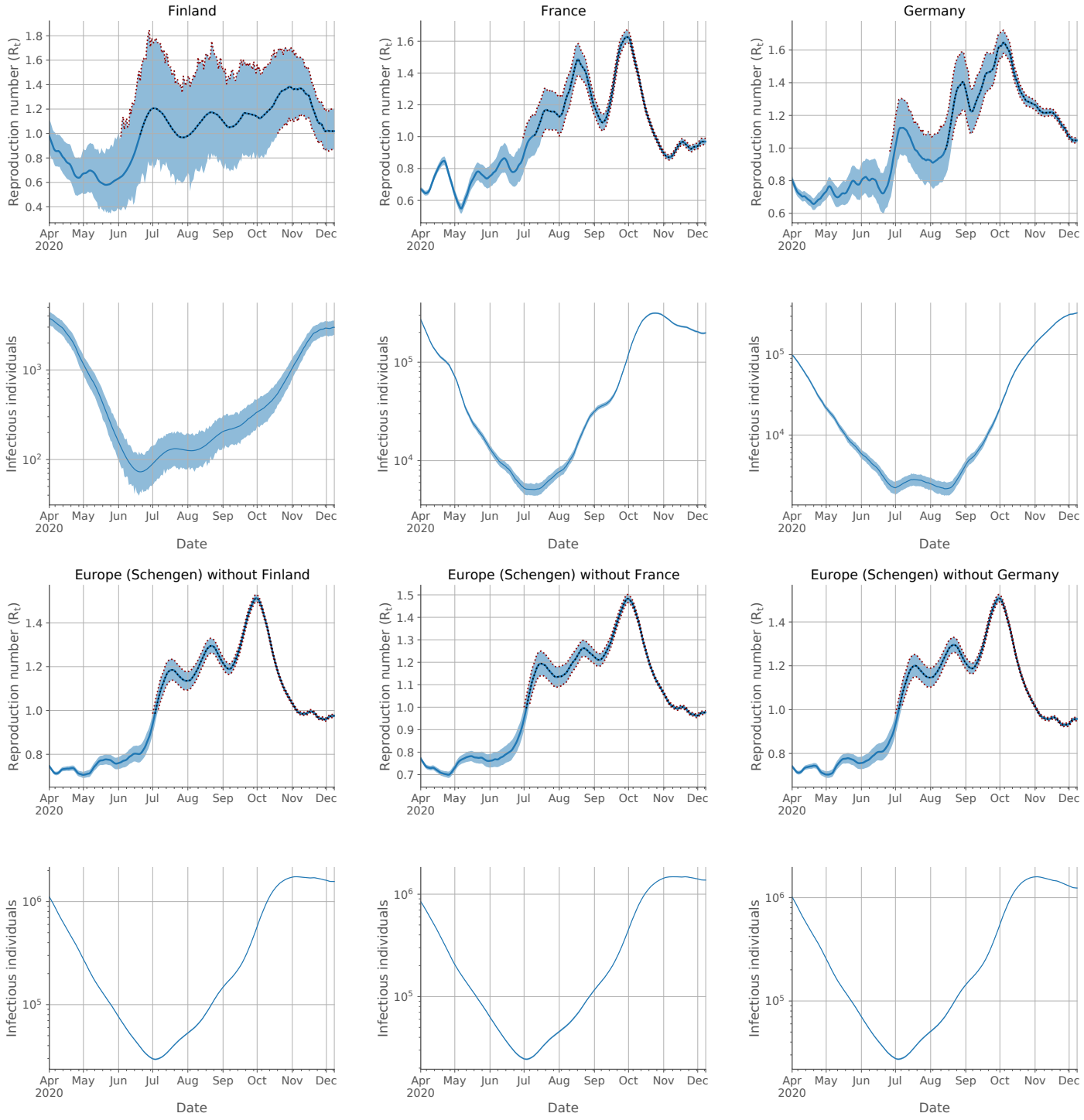
Supplementary Figure 1.4



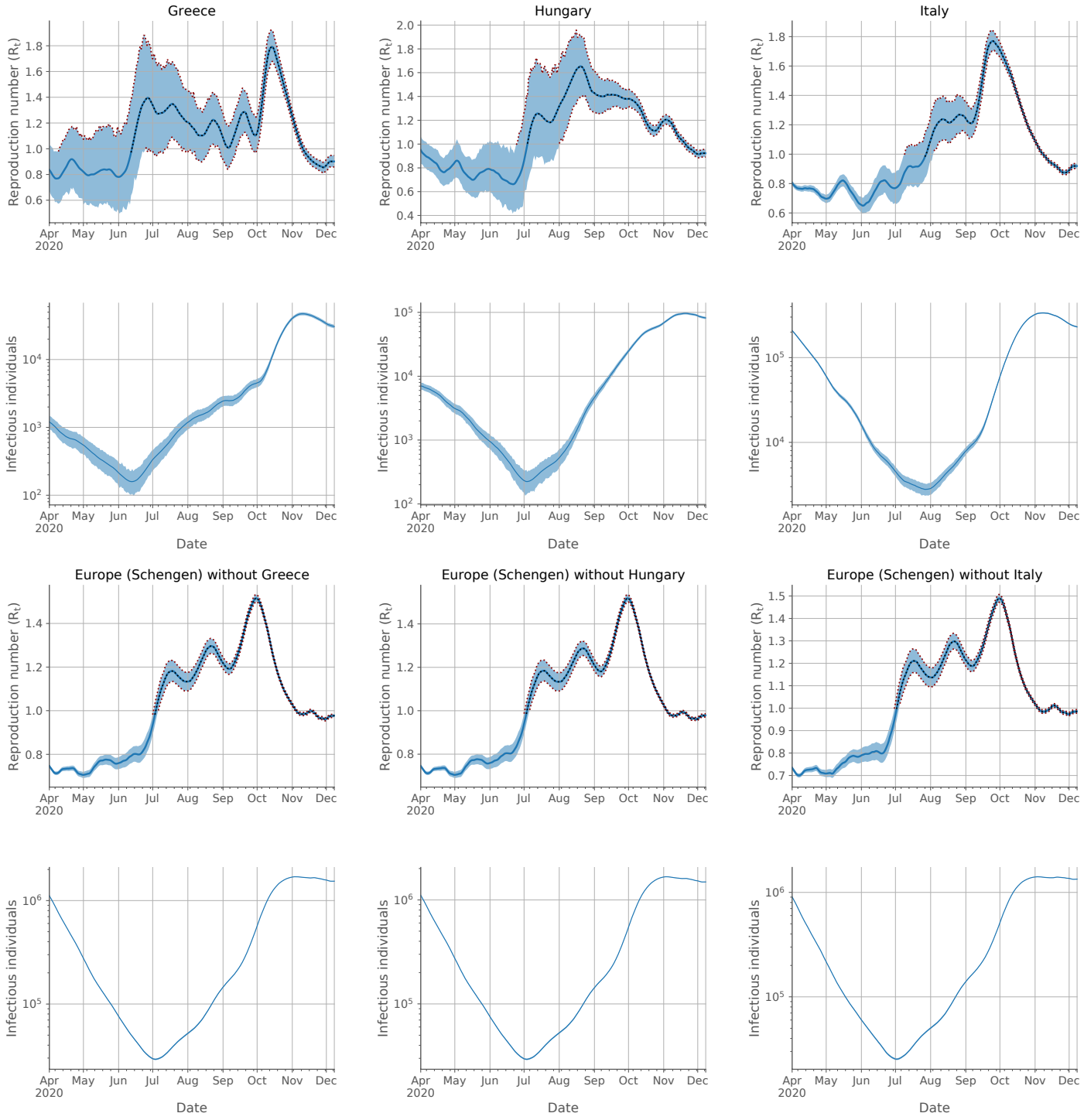
Supplementary Figure 1.5



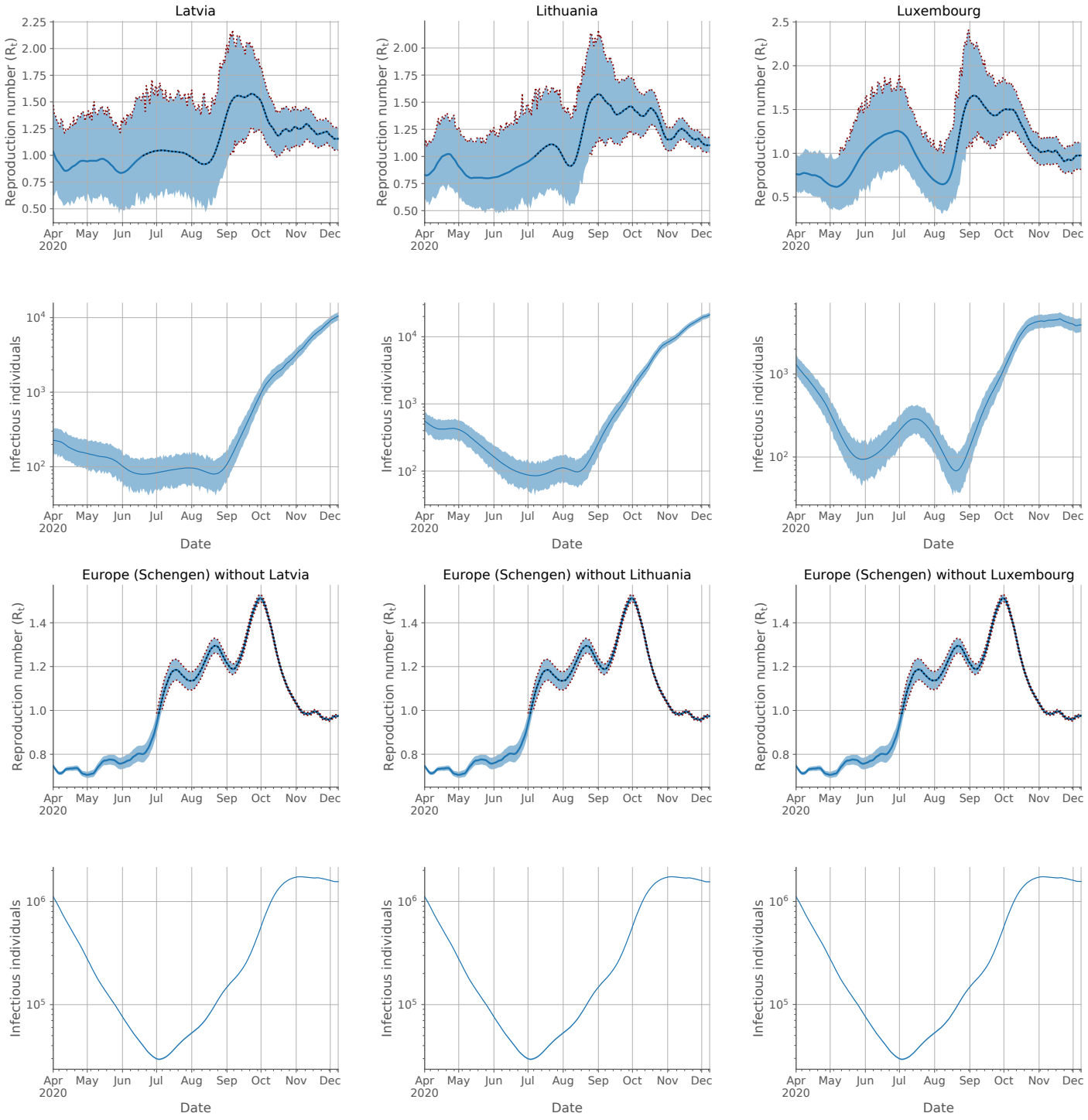
Supplementary Figure 1.6



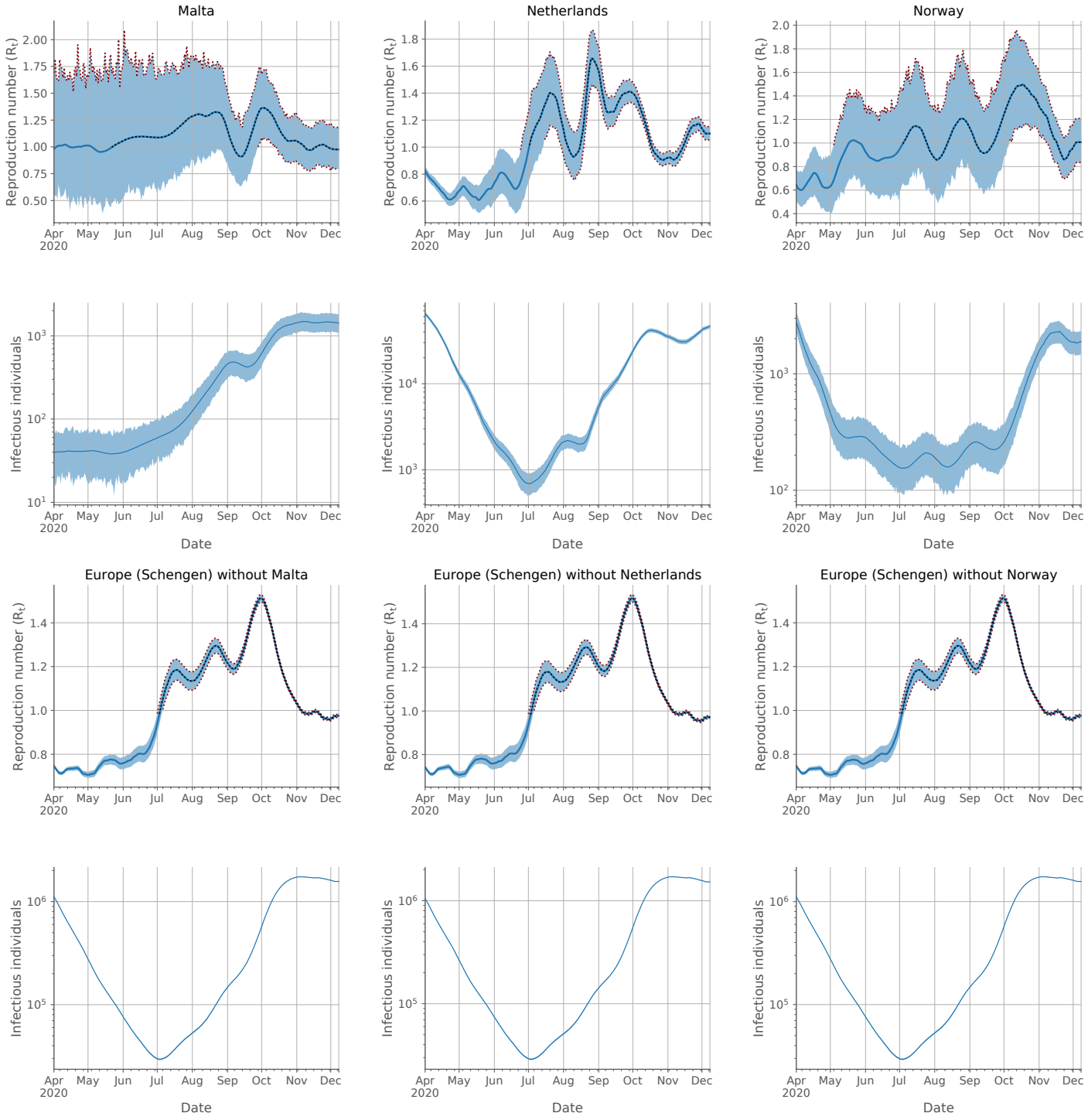
Supplementary Figure 1.7



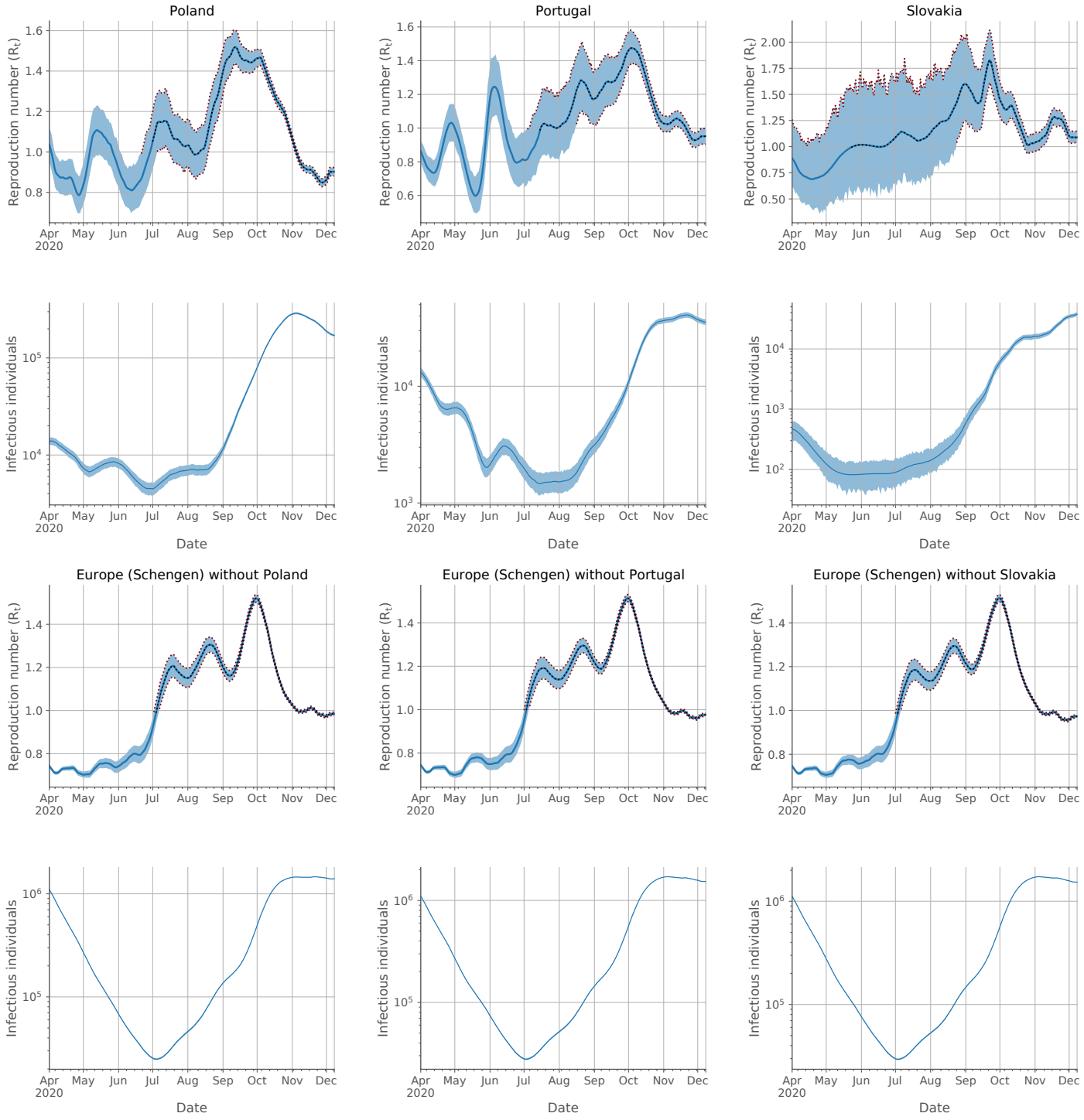
Supplementary Figure 1.8



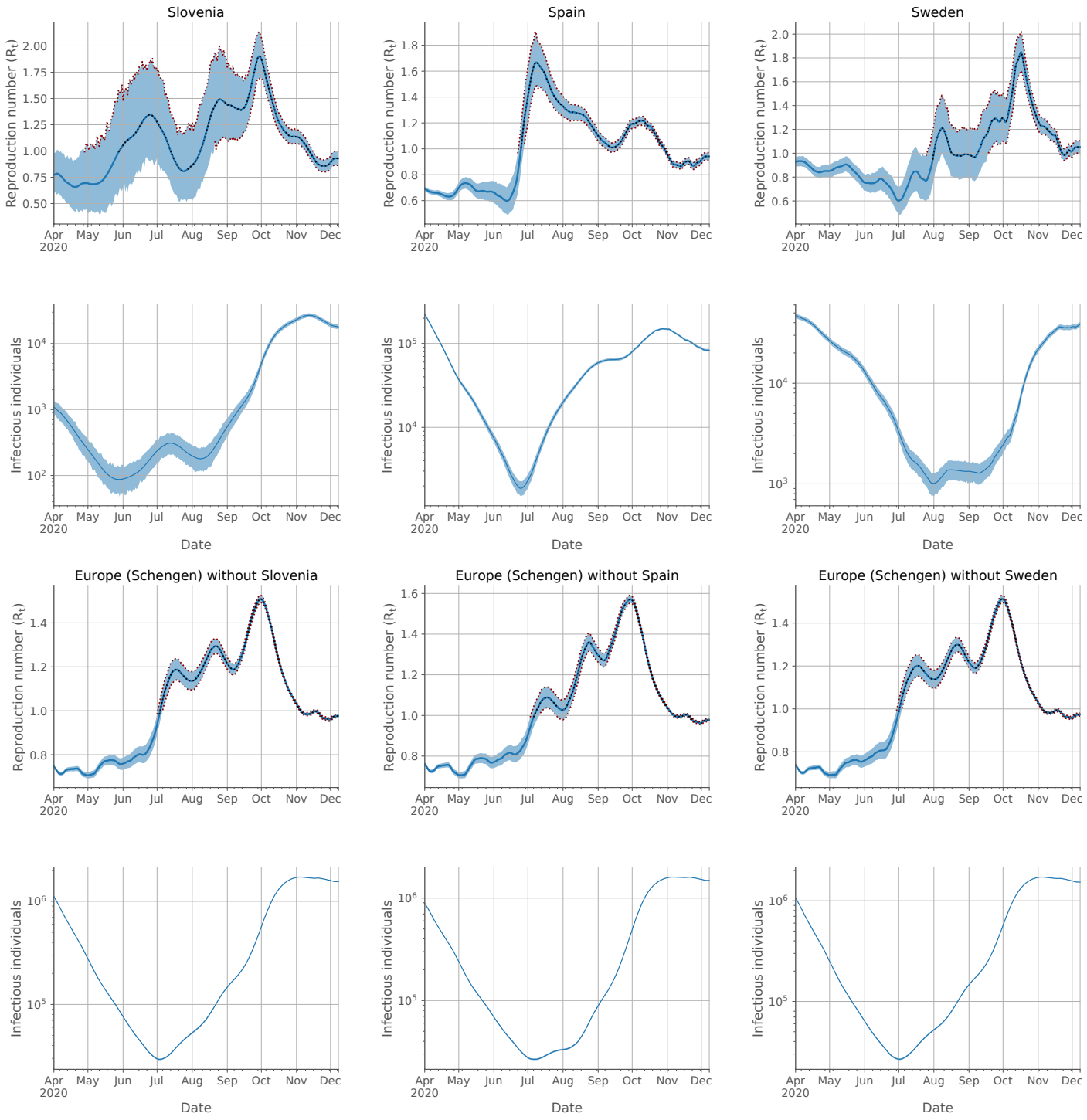
Supplementary Figure 1.9



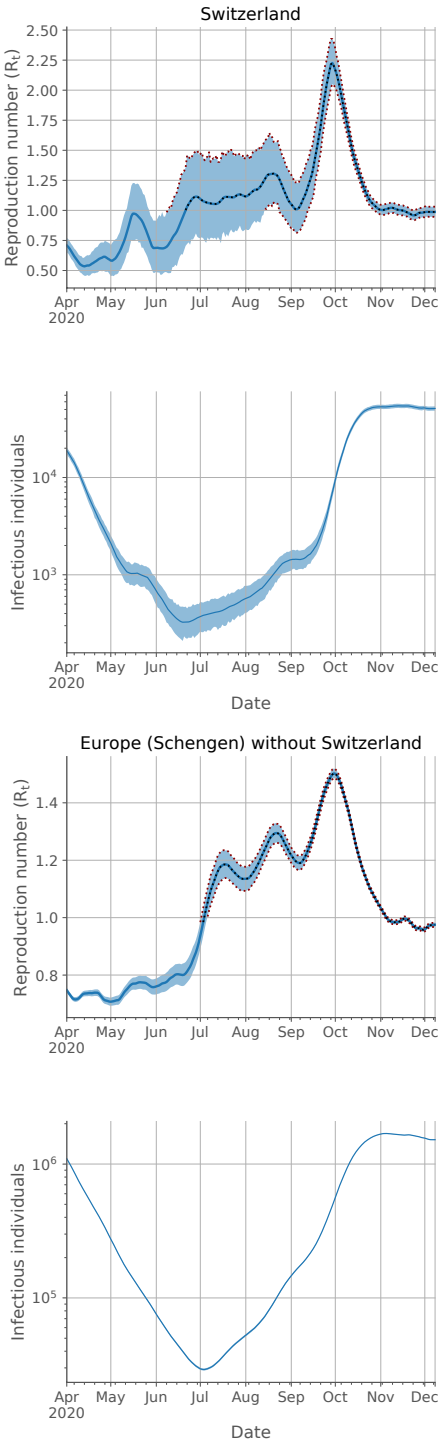
Supplementary Figure 1.10



Supplementary Figure 1.11



Supplementary Figure 1.12



Supplementary Figure 1.13