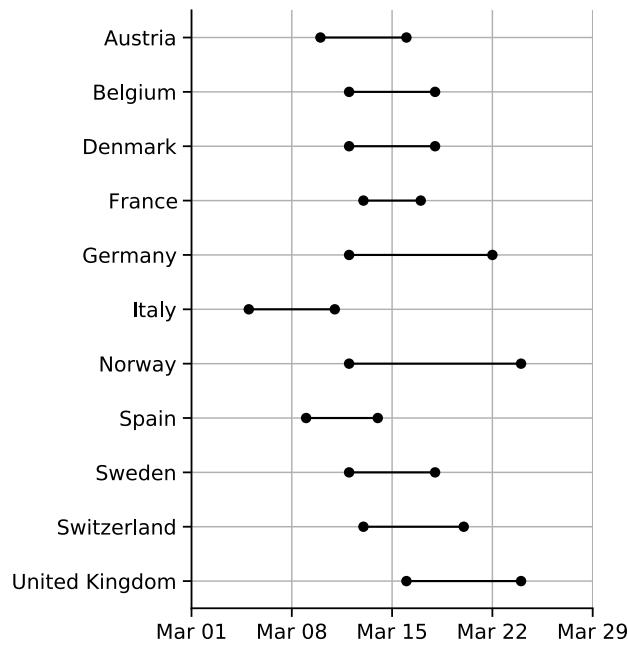
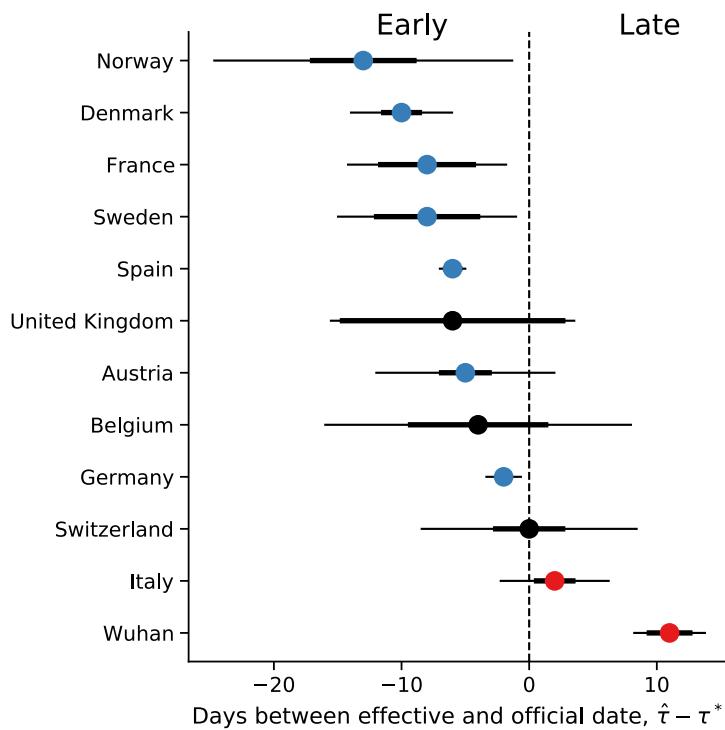


## Supplementary Material



**Figure S1: Official start of non-pharmaceutical interventions.** See Table 1 for more details. Wuhan, China is not shown.



**Figure S2: Official vs. effective start of non-pharmaceutical interventions estimated up to Mar 28.** The difference between  $\hat{\tau}$  the effective and  $\tau^*$  the official start of NPIs estimated from case data up to Mar 28, 2020, shown for different regions. Here,  $\hat{\tau}$  is the posterior median, see Table 2.  $\tau^*$  is the last NPI date (Table 1). Thin and bold lines show 95% and 75% credible intervals, respectively.

WAIC

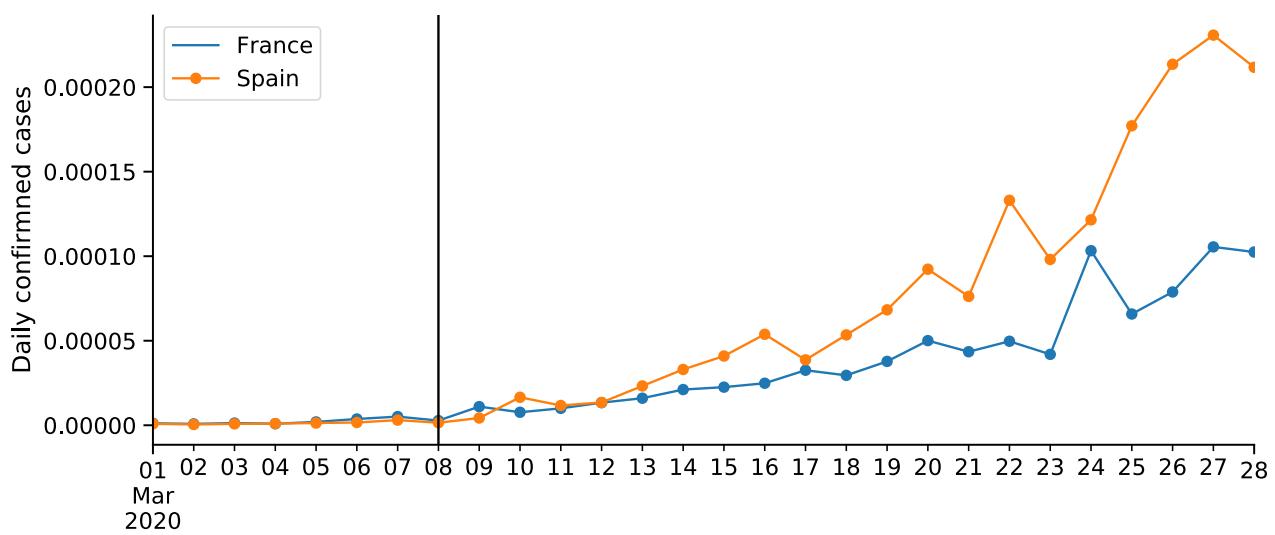
Country	No	Fixed	Free
Austria	219.49	95.06	<b>35.96</b>
Belgium	148.37	98.41	<b>49.16</b>
Denmark	44.36	<b>40.30</b>	43.11
France	581.59	255.14	<b>172.08</b>
Germany	1029.36	327.50	<b>174.90</b>
Italy	898452.34	5484.56	<b>80.18</b>
Norway	70.03	42.04	<b>39.79</b>
Spain	1476.46	647.34	<b>128.58</b>
Sweden	32.53	<b>30.06</b>	31.10
Switzerland	265.80	83.95	<b>63.89</b>
United Kingdom	258.18	117.54	<b>68.17</b>
Wuhan China	107.31	94.00	<b>73.75</b>

**Table S1: WAIC values for the different models.** WAIC (widely applicable information criterion; Eq. 10)<sup>11</sup> values for models with: no  $\tau$  at all, *No*;  $\tau$  fixed at the official last NPI date  $\tau^*$ , *Fixed*; and free parameter  $\tau$ , *Free*. WAIC values are scaled as a deviance measure: lower values imply higher predictive accuracy and a difference of 2 is a popular threshold for model comparison<sup>13</sup>. Bold values emphasize cases in which the *Free* model has the lowest WAIC.

RMSE

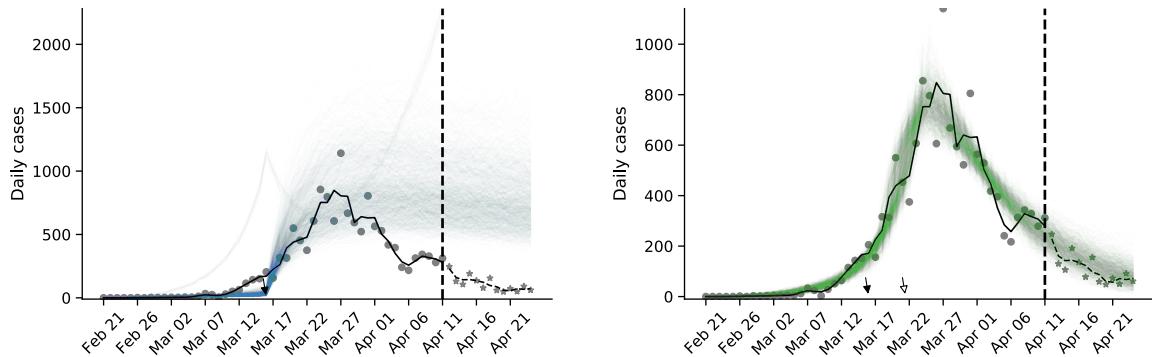
Country	No	Fixed	Free
Austria	926.7	740.7	<b>58.9</b>
Belgium	3286.0	2541.0	<b>902.9</b>
Denmark	434.8	1058.0	<b>400.0</b>
France	10530.0	8206.0	<b>1816.0</b>
Germany	14240.0	17760.0	<b>1952.0</b>
Italy	14160.0	9536.0	<b>1106.0</b>
Norway	292.2	578.2	<b>63.3</b>
Spain	17840.0	14410.0	<b>1704.0</b>
Sweden	775.9	1205.0	<b>596.4</b>
Switzerland	1839.0	1721.0	<b>214.0</b>
United Kingdom	14700.0	14780.0	<b>2735.0</b>

**Table S2: Posterior RMSE of out-of-sample predictions with the different models.** Expected posterior predictive RMSE (root mean squared error) for models with: no  $\tau$  at all, *No*;  $\tau$  fixed at the official last NPI date  $\tau^*$ , *Fixed*; and free parameter  $\tau$ , *Free*. In all cases, the model with free parameter  $\tau$  has the lowest RMSE. Models were fitted to case data up to Apr 11, 2020, and then used to generate 1,000 predictions up to Apr 24 by sampling model parameters from the posterior distribution. These predictions were then compared to the real data using RMSE, and the mean RMSE value is shown in the table for each country and model. Bold values emphasize cases in which the *Free* model has the lowest RMSE.

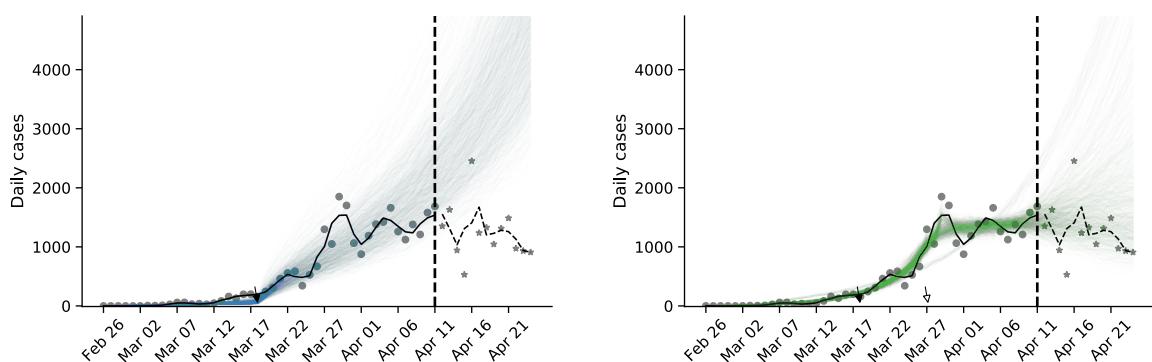


**Figure S3: COVID-19 confirmed cases in France and Spain.** Number of cases proportional to population size (as of 2018). Vertical line shows Mar 8, the effective start of NPIs  $\hat{\tau}$  in both countries.

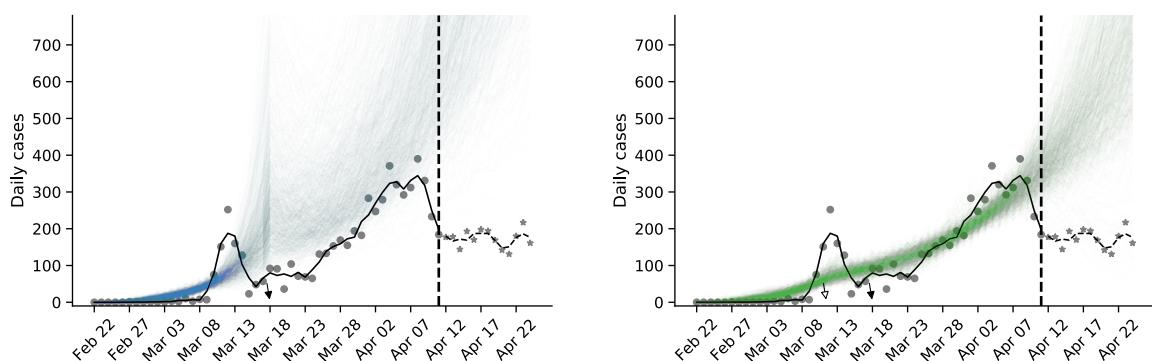
### Austria



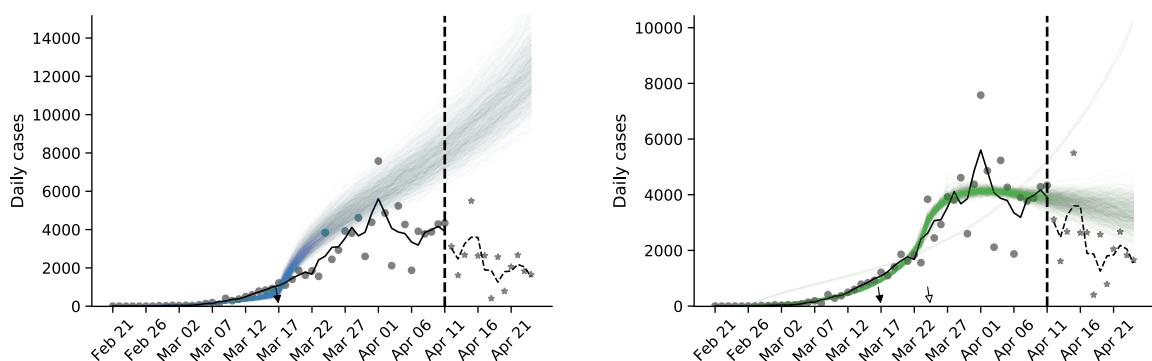
### Belgium



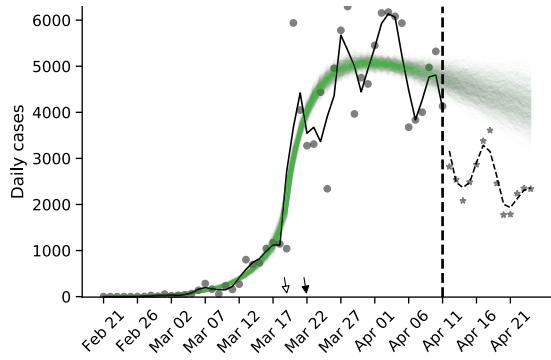
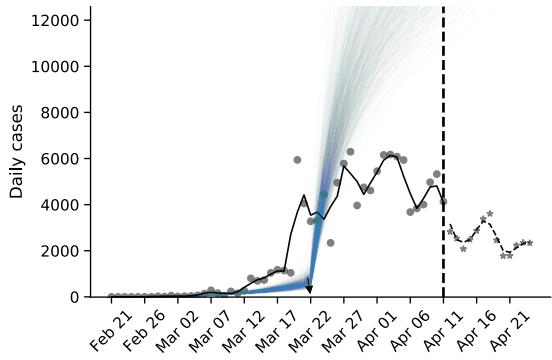
### Denmark



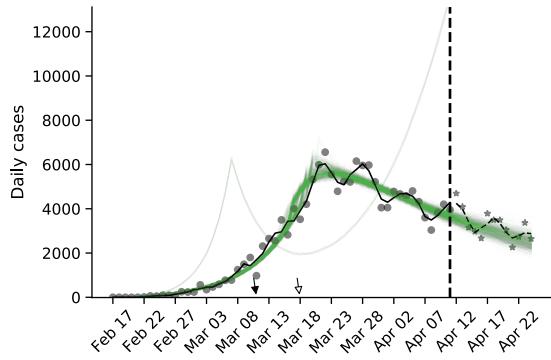
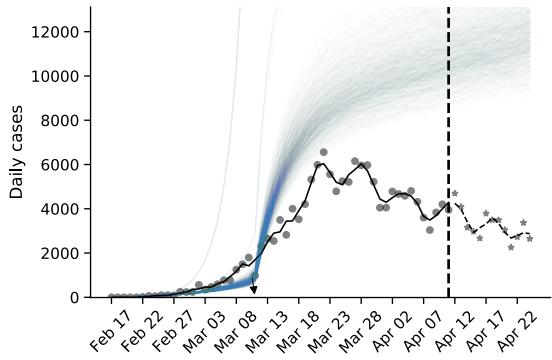
### France



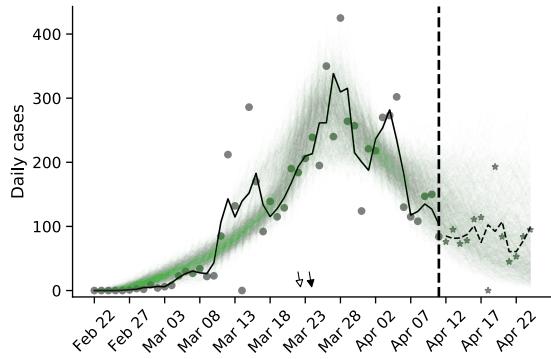
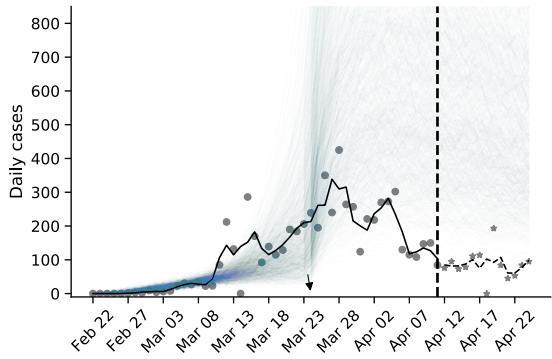
## Germany



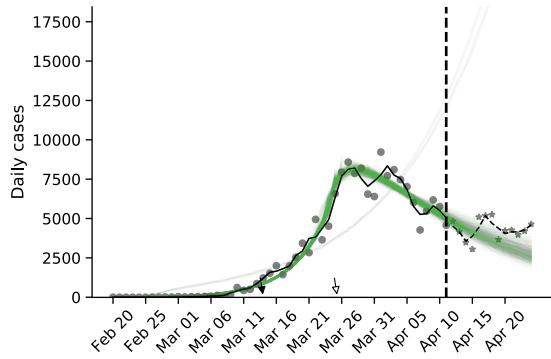
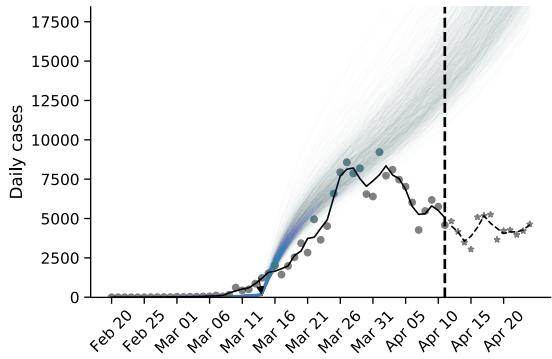
## Italy



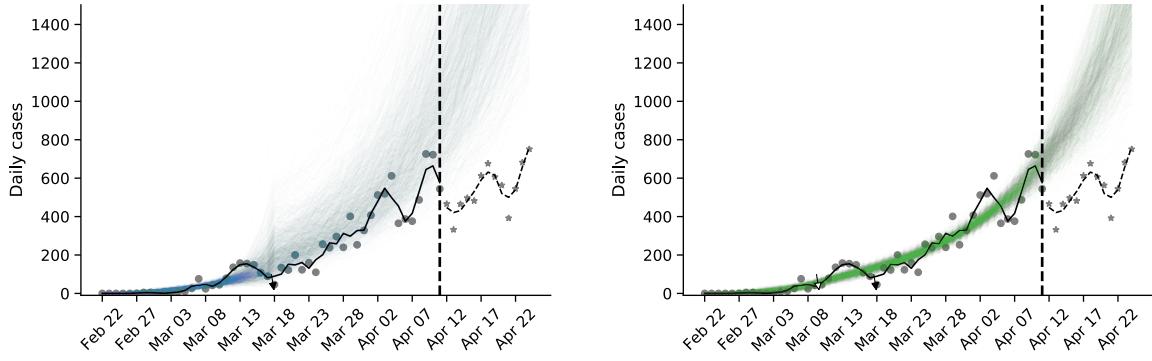
## Norway



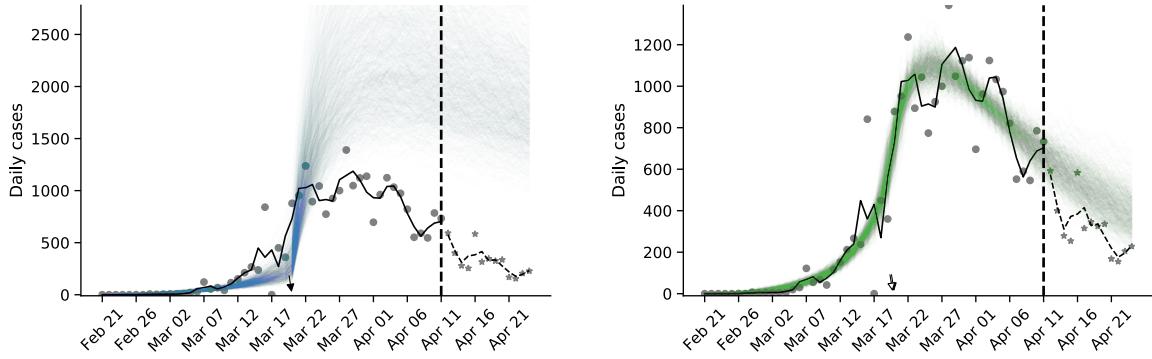
## Spain



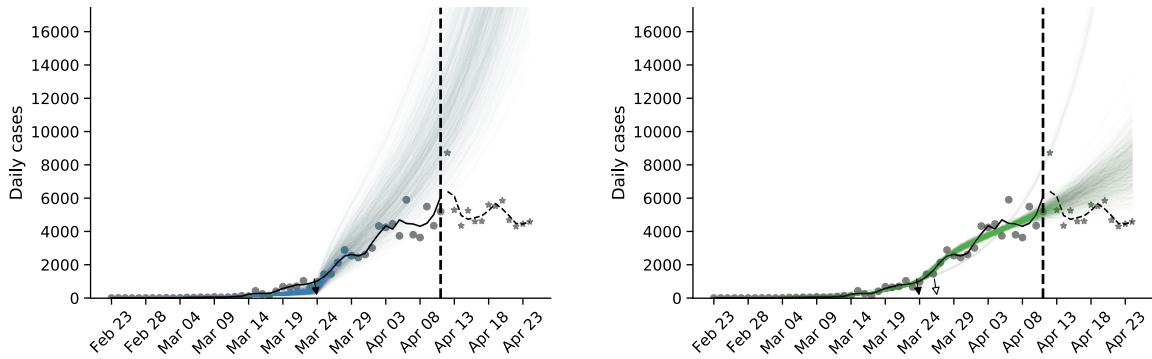
### Sweden



### Switzerland



### United Kingdom



**Figure S4. Posterior prediction plots for 11 European countries.** The vertical dashed line represents Apr 11, 2020. Circles and stars represent daily case data up to and after Apr 11, respectively. Black and white arrows denote the official  $\tau^*$  and effective  $\hat{\tau}$  start of NPIs, respectively. Black lines represent a smoothing of the data points using a Savitzky-Golay filter. Coloured lines represent posterior predictions from a model with fixed  $\tau$  (blue) and free  $\tau$  (green). Models were fitted with data up to Apr 11. The predictions are generated by drawing 1,000 parameter sets from the posterior distribution, and then generating a daily case count using the SEIR model up to Apr 24. Note the differences in the y-axis scale. Posterior predictions with the free  $\tau$  model predict the out-of-sample data well for all countries except Denmark and Sweden, but poorly for the fixed  $\tau$  model. The predictions of the model without  $\tau$  (not shown) are even worse.