Research Article Summary: Inferring COVID-19 spreading rates and potential change points for case number forecasts

Jonas Dehning,[1,2*] Johannes Zierenberg,[1*] F. Paul Spitzner,[1,2*]
Michael Wibral,[3] Joao Pinheiro Neto,[1] Michael Wilczek,[1*] Viola Priesemann[1*]

[1] Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany,
[2] Institute of Neuroscience and Medicine (INM-6), Research Center Jülich, Germany
[3] Campus Institute for Dynamics of Biological Networks, University Göttingen, Germany

* Authors contributed equally. Correspondence should be addressed to viola.priesemann@ds.mpg.de.

Introduction. When faced with the outbreak of a novel epidemic like COVID-19, rapid response measures are required by individuals as well as by society as a whole to mitigate the spread of the virus. During this initial, time-critical period, neither the central epidemiological parameters, nor the effectiveness of measures like cancellation of public events, school closings, and social distancing are known.

Rationale. As one of the key epidemiological parameters, we infer the spreading rate λ from confirmed COVID-19 case numbers at the example in Germany by combining Bayesian inference with an SIR (Susceptible-Infected-Recovered) model from compartmental epidemiology. Our analysis characterizes the temporal change of the spreading rate and, importantly, allows us to identify potential change points and to provide short-term forecast scenarios based on various degrees of social distancing. A detailed, educational description is provided in the accompanying paper, and the model, inference, and prediction are available on github. While we apply it to Germany, our approach can be readily adapted to any other country or region.

Results. In Germany, political interventions to contain the outbreak were implemented in three steps over three weeks: Around March 9, large public events like soccer matches were cancelled. On March 16, schools and other educational institutions as well as many non-essential stores were closed. One week later, on March 23, a far-reaching contact ban ("Kontaktsperre"), which includes the prohibition of even small public gatherings as well as the further closing of restaurants and non-essential stores, was imposed by the government authorities.

From the observed case numbers of COVID-19, we can quantify the impact of these measures on the spread (Fig. 1). As of April 10, we have evidence of the first change point in the spreading rate from $\lambda_0 = 0.40$ (95 % Confidence interval (CI: [0.33,0.49]) to $\lambda_1 = 0.24$ (CI: [0.20,0.28]), which occurred around March 8 (CI: March 5 to March 10). Moreover, we have evidence for a second change point to $\lambda_2 = 0.15$ (CI: [0.12,0.19]), which occurred around March 16 (CI: March 15 to March 18). Both changes in λ slowed the spread of the virus, but still imply exponential growth (Fig. 1, red and orange traces). To contain the disease spread, and turn from exponential growth to a decline of novel cases, a further decrease in λ is necessary. We have first indications that this transition has been reached by the third change-point around March 23 (CI: March 21 to March 25).

With the start of this third change point, λ takes approximately the critical value where the spreading rate λ balances the recovery rate μ, i.e. the effective growth rate $\lambda^* = \lambda - \mu \approx 0$ (Fig. 1, green traces). The case numbers in the coming week will provide more information on its precise value. Importantly, $\lambda^* = 0$ presents the watershed between exponential growth or decay. Together with the delay of approximately two weeks between infection and first inference of λ^*, any future intervention such as lifting restrictions therefore warrants careful consideration.

Our detailed analysis shows that, *in the current phase*, reliable short- and long-term forecasts are very difficult, if not impossible: In Fig. 1C,D already the three example scenarios quickly diverge from each other, and consequently span a huge range of future case numbers. Thus, any uncertainty on the magnitude of our social distancing in the past two weeks can have a major impact on the case numbers in the next two weeks. Beyond two weeks, the case numbers depend on our future behavior, for which we have to make explicit assumptions. We illustrate how the precise magnitude and timing of potential change points impact the forecast of case numbers (see Fig. 2, main paper).

Conclusions. We developed a Bayesian framework to infer the spreading rate λ and the timing and magnitude of change points. Thereby, the efficiency of political and individual measures for social distancing and containment can be assessed in a timely manner. We find first evidence for a successive decrease of the spreading rate in Germany around March 9 and around March 16, which significantly reduced the magnitude of exponential growth, but was not sufficient to turn growth into decay. The development in the coming week will reveal the efficiency of the contact ban initiated on March 23. In general, our analysis code may help to infer the efficiency of measures taken in other countries and inform policy makers about tightening, loosening and selecting appropriate rules for containment.

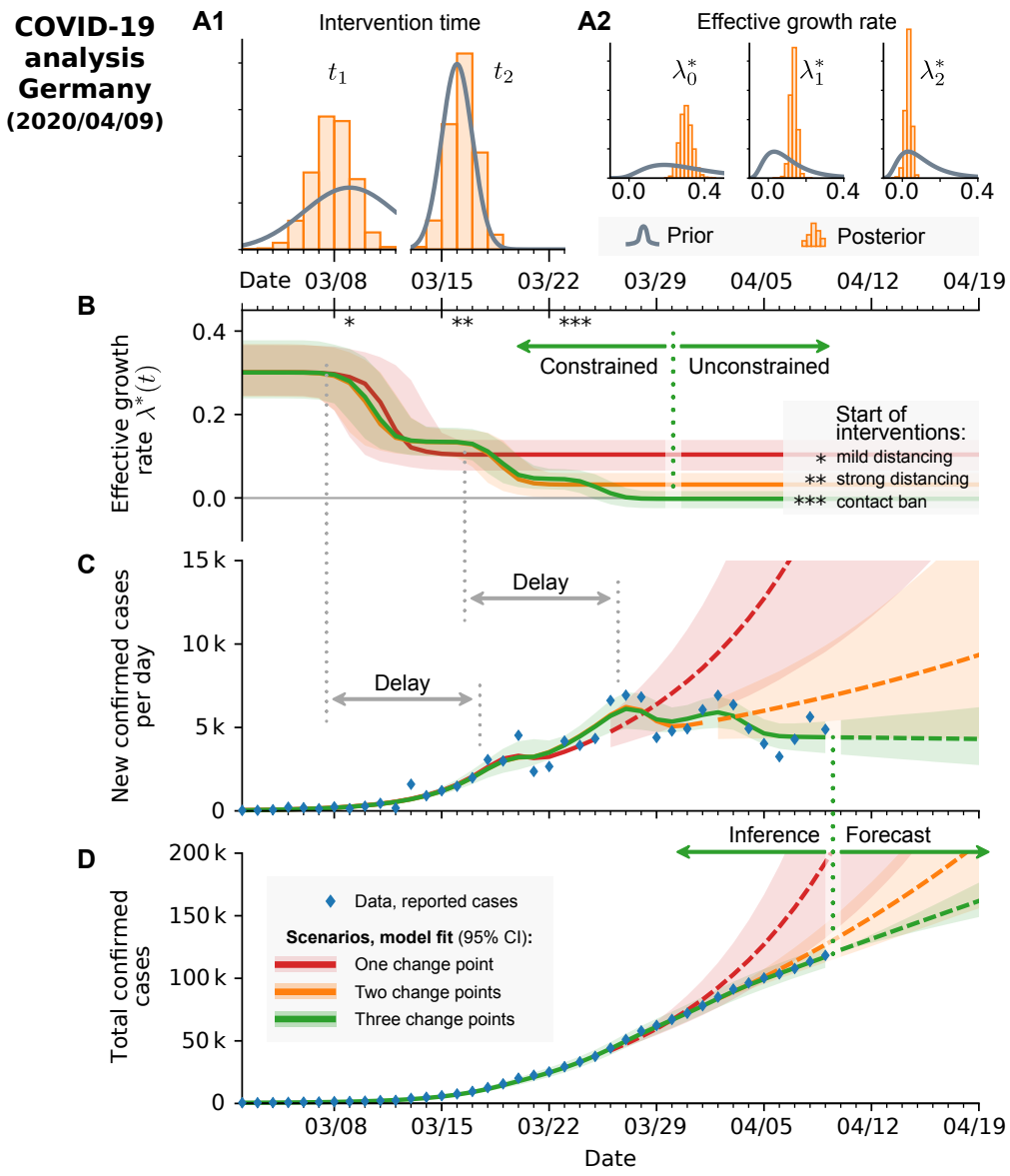

FIG. 1. [Our paper uses data available as of April 10. On our github repository you find the current figure versions.] Inference of change points in the spreading rate λ from confirmed COVID-19 cases in Germany. **A:** Prior (blue) and posterior (orange) distributions for five of the central parameters of an SIR model with two change points (at time t_1 and t_2), where the spreading rate changes from $\lambda_0 \to \lambda_1 \to \lambda_2$. **B:** The inferred growth rate λ^*, i.e. the difference between spreading and recovery rate ($\lambda^* = \lambda - \mu$) for an SIR model that assumes scenarios with one, two or three change points (red, orange, green; fitted to case reports until March 25, April 1 and April 9, respectively). The timing of the inferred change points corresponds well to the timing of the governmental interventions in Germany (depicted as $*$). **C,D:** The model-fit of the new confirmed cases and (cumulative) total confirmed cases is depicted for the models with one, two or three change points. The three scenarios depend strongly on whether one includes the third change point or not: the number of new confirmed cases grow exponentially (red, orange) or are approximately constant (green). This illustrates that the future development depends strongly on our distancing behavior. **B,C:** Note the delay D between change point (i.e. change in spreading behavior) and observation of confirmed cases of almost two weeks.

Inferring COVID-19 spreading rates and potential change points for case number forecasts

Jonas Dehning[1,2*], Johannes Zierenberg[1*], F. Paul Spitzner[1,2*], Michael Wibral[3], Joao Pinheiro Neto[1], Michael Wilczek[1*], and Viola Priesemann[1*]
[1] *Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany*
[2] *Institute of Neuroscience and Medicine (INM-6), Research Center Jülich, Germany and*
[3] *Campus Institute for Dynamics of Biological Networks, University of Göttingen, Germany*
(Dated: April 11, 2020)

As COVID-19 is rapidly spreading across the globe, short-term modeling forecasts provide time-critical information for decisions on containment and mitigation strategies. A main challenge for short-term forecasts is the assessment of key epidemiological parameters and how they change as first governmental intervention measures are showing an effect. By combining an established epidemiological model with Bayesian inference, we analyze the time dependence of the effective growth rate of new infections. For the case of COVID-19 spreading in Germany, we detect change points in the effective growth rate that correlate well with the times of publicly announced interventions. Thereby, we can (a) quantify the effects of recent governmental measures to mitigating the disease spread, and (b) incorporate the corresponding change points to forecast future scenarios and case numbers. Our code is freely available and can be readily adapted to any country or region.

f

INTRODUCTION

During the initial outbreak of an epidemic, reliable short-term forecasts are key to estimate required medical capacities, and to inform and advice the public and decision makers [1]. During this initial phase, three tasks are of particular importance to provide time-critical information for crisis mitigation: (1) establishing central epidemiological parameters such as the basic reproduction number that can be used for short-term forecasting; (2) simulating the effects of different possible interventions aimed at the mitigation of the outbreak; (3) estimating the actual effects of the measures taken – to rapidly adjust them and to adapt short-term forecasts. Tackling these tasks is challenging due to the large statistical and systematic errors that are present during the initial stages of an epidemic with its low case numbers. This is further complicated by the fact that mitigation measures are taken rapidly, while the outbreak unfolds, but they take an effect only after an a priori unknown delay. To obtain reasonable parameter estimates for short-term forecasting and policy evaluation despite these complications, any prior knowledge available needs to be integrated into modeling efforts to reduce uncertainties. This includes knowledge about basic mechanisms of disease transmission, recovery, as well as preliminary estimates of epidemiological parameters from other countries, or from closely related pathogens. The integration of prior knowledge, the quantitative assessment of the remaining uncertainties about epidemiological parameters, and the principled propagation of these uncertainties into forecasts is the domain of Bayesian modeling and inference [2, 3].

Here, we draw on an established class of models for epidemic outbreaks: The Susceptible-Infected-Recovered (SIR) model [4–7] specifies the rates with which population compartments change over time, i.e., with which susceptible people become infectious, or infectious people recover. This simple model can be formulated in terms of coupled ordinary differential equations (in mean field), which enable analytical treatment [8, 9] or fast evaluation (ideally suited for Bayesian inference). Accordingly, SIR-like models have been used to model epidemic spreads, from Bayesian Markov-Chain Monte Carlo (MCMC) parameter estimation [10–12] to detailed scenario discussions [13–16]. Recently, this family of models also played a dominant role in the analyses of the global corona virus (SARS-CoV-2) outbreak, from inference [17–19] to scenario forecast [20–27] to control strategies [28, 29].

We combine the SIR model with Bayesian parameter inference and augment the model by a time-dependent spreading rate. The time dependence is implemented via potential change points reflecting changes in the spreading rate driven by governmental interventions. Based on three distinct measures taken in Germany, we detect three corresponding change points from reported COVID-19 case numbers. Already on April 1 we had reported evidence for the first two change points, and predicted the third one [30]. Now, with data until April 9, we have evidence for all three change points. First, the spreading rate decreased from 0.40 (CI [0.33,0.49]) to 0.24 (CI [0.20,0.28]), with this decrease initiated around March 8 (CI [March 6, March 10]). This matches the cancellation of large public events such as trade fairs and soccer matches. Second, the spreading rate decreased further to 0.15 (CI [0.12,0.19]) initiated around March

16 (CI [March 15, March 18]). This matches the closing of schools, child-care facilities, and non-essential stores. Third, the spreading rate decreased further to 0.10 (CI [0.07,0.13]) initiated around March 23 (CI [March 21, March 25]). This corresponds well to the strict contact ban, which was announced on March 22. While the first two change points were not sufficient to switch from growth of novel cases to a decline, the third change point probably brought this crucial reversal.

Our framework is designed to infer the effectiveness of past measures and to explore potential future scenarios with propagating the respective uncertainties. In the following we demonstrate the potential impact of timing and magnitude of change points, and report our inference about the three past governmental interventions in Germany. Our framework can be readily adapted to any other country or region. The code (already including data sources from many other countries), as well as the figures are all available on Github [31].

BACKGROUND: INFERENCE OF CENTRAL EPIDEMIOLOGICAL PARAMETERS AND THE EFFECTS OF INTERVENTIONS

In order to simulate the general effect of different possible interventions on the spread of COVID-19 in Germany, we first focus on the initial phase of the outbreak where no serious mitigation measures were implemented. In the absence of interventions, an epidemic outbreak can be described by SIR models with constant spreading rate (Methods). In Germany, first serious interventions occurred around March 9 and affected the case reports with an observation delay, a combination of incubation period with median $5-6$ days [32] and a test delay (time until doctor is visited plus test-evaluation time) that we assume to be about 2 days. Hence, we consider as initial phase the time period from March 1 to March 15 in order to infer central epidemiological parameters. We then model the effects of interventions as change points in the spreading rate (Methods) in order to simulate the effect of different possible interventions.

Bayesian inference for central epidemiological parameters during the initial phase of the outbreak

We perform Bayesian inference for the central epidemiological parameters of an SIR model using Markov-Chain Monte Carlo (MCMC) sampling (Fig. 1). The central parameters are the spreading rate λ, a recovery rate μ, a reporting delay D, and the number of initially infected people I_0. We chose informative priors based on available knowledge for λ, μ, and D, and uninformative priors for the remaining parameters (Methods). We intentionally kept also the informative priors as broad as possible such that the data would constrain the parameters (Fig. 1).

As median estimates, we obtain for the spreading rate $\lambda = 0.41$, $\mu = 0.12$, $D = 8.6$, and $I_0 = 19$ (see Fig. 1D–H for the posterior distributions and the CIs). Converted to the basic reproduction number $R_0 = \lambda/\mu$, we find a median $R_0 = 3.4$ (CI [2.4, 4.7]), which is consistent with previous reports that find median values between 2.3 and 3.3 [18, 33, 34]. Overall, the model shows good agreement for both new cases C_t (Fig. 1 A) and the cumulative cases $\sum_{t'=0}^{t} C_{t'}$ (Fig. 1 B) with the expected exponential growth (linear in lin-log plot). The absolute deviation between data and model (Fig. 1 C) is well captured by the case-number-dependent width of our likelihood (Methods) motivated by demographic noise in mean-field models of spreading processes [35, 36]. The observed data are clearly informative about λ, I_0 and σ (indicated by the difference between the priors (gray line) and posteriors (histograms) in Fig. 1 D,I,H). However, μ and D are largely determined by our prior choice of parameters (histograms match gray line in Fig. 1 E,G). This is to be expected for the initial phase of an epidemic outbreak, which is dominated by exponential growth.

In order to quantify the impact of possible interventions, we concentrate on the effective growth of active infections before and after the intervention. As long as the number of infections and recoveries are small compared to the population size, the number of active infections per day can be approximated by an exponential growth (Fig. 1A,B) with effective growth rate $\lambda^* = \lambda - \mu$ (see Methods). As a consequence, λ and μ cannot be estimated independently. This is further supported by a systematic scan of the model's log-likelihood in the λ–μ space that shows an equipotential line for the maximum likelihood (Fig. 1 J). This strongly suggests that the growth rate λ^* is the relevant free parameter with a median $\lambda^* = 28\%$ (Fig. 1 I). The control parameter of the dynamics in the exponential phase is thus the (effective) growth rate: If the growth rate is larger than zero ($\lambda > \mu$), case numbers grow exponentially; if the growth rate is smaller than zero ($\lambda < \mu$), the recovery dominates and the new confirmed cases decrease. The two different dynamics (supercritical and subcritical, respectively) are separated by a critical point at $\lambda^* = 0$ ($\lambda = \mu$) [36].

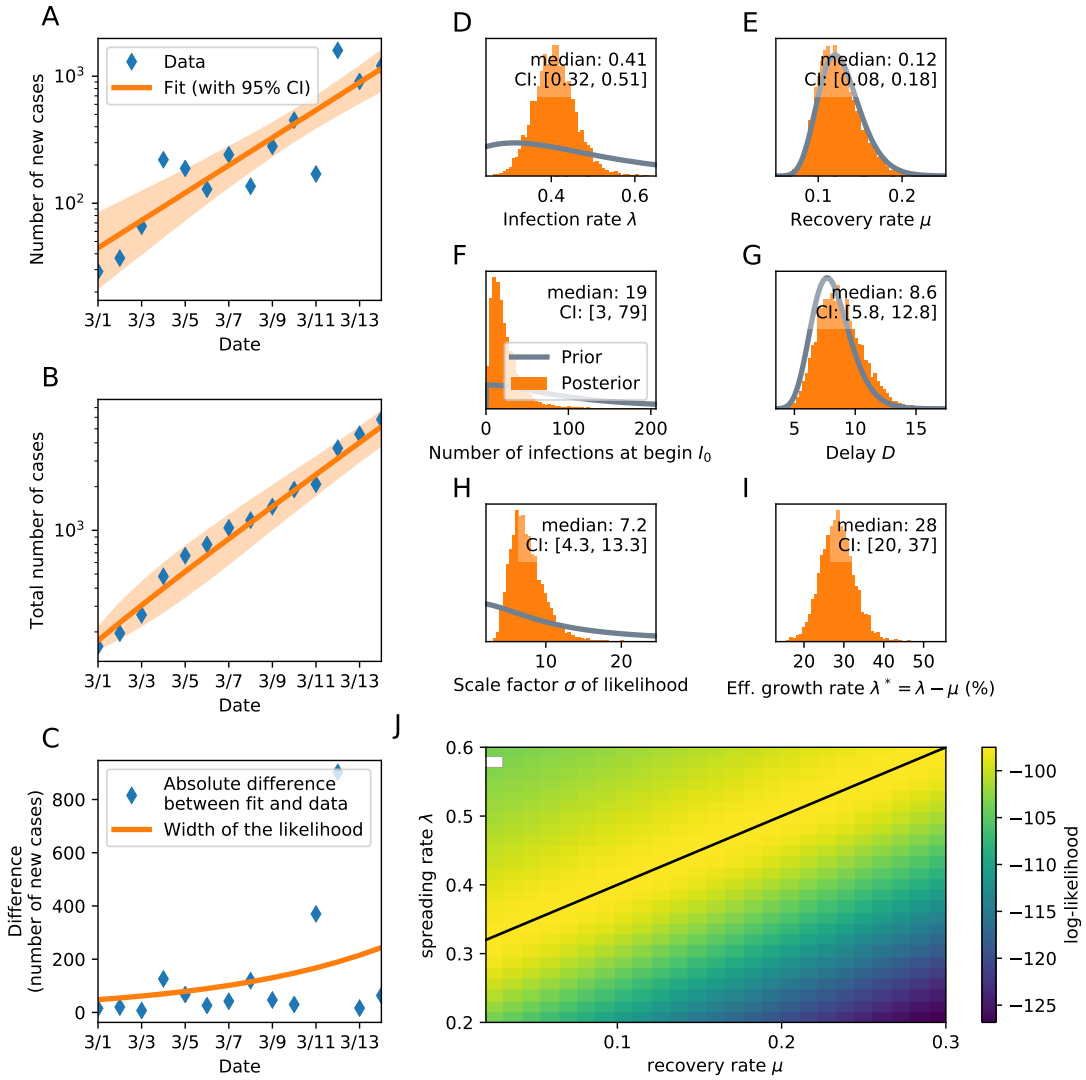

FIG. 1. Inference of central epidemiological parameters of the SIR model during the initial onset period, March 1–15. **A:** Number of new cases over time, and **B:** total number of cases over time (cumulative) increase exponentially. **C:** Absolute difference between model and data is captured by the width of our likelihood, scaling with the square root of new cases $\sigma\sqrt{C_t}$ (orange line). **D-I:** Prior (gray) and posterior (orange) distributions for all model parameters: estimated spreading rate λ, recovery rate μ, reporting delay D between infection date and reporting date, number of cases I_0 at the start of the simulation, scale-factor σ of the width of the likelihood distribution, and the effective growth rate $\lambda^* = \lambda - \mu$. **J:** Log-likelihood distribution for different combinations of λ and μ. A linear combination of λ and μ yield the same maximal likelihood (black line). White dot: Inference did not converge.

Magnitude and timing of interventions matter for the mitigation of the outbreak

We simulate different, hypothetical interventions following the initial phase in order to show that both, the amount of change in behavior (leading to a change in spreading rate λ, Fig. 2 A,B) and the exact timing of the change (Fig. 2 C,D) determine the future development. Hypothetical interventions build on the inferred parameters from the initial phase (Fig. 1, in particular median $\lambda_0 = 0.41$ and median $\mu = 0.12$) and were implemented as change points in the spreading rate from the inferred λ_0 to a new value λ_1. With such a change point, we model three potential scenarios of public behavior: **(I) No social distancing;** Public behavior is unaltered and the spread continues with the inferred rate ($\lambda_1 = \lambda_0$ with median $\lambda_1 = 0.41 > \mu$). **(II) Mild social distancing;** The spreading rate decreases to 50%, ($\lambda_1 = \lambda_0 / 2$ with median $\lambda_1 = 0.21 > \mu$). Although people effectively reduce the number of contacts by a factor of two in this second scenario, the total number of reported cases continues to grow alongside scenario (I) for the time period of the reporting delay D (median $D = 8.6$ from initial phase, see below for a more constrained estimation).

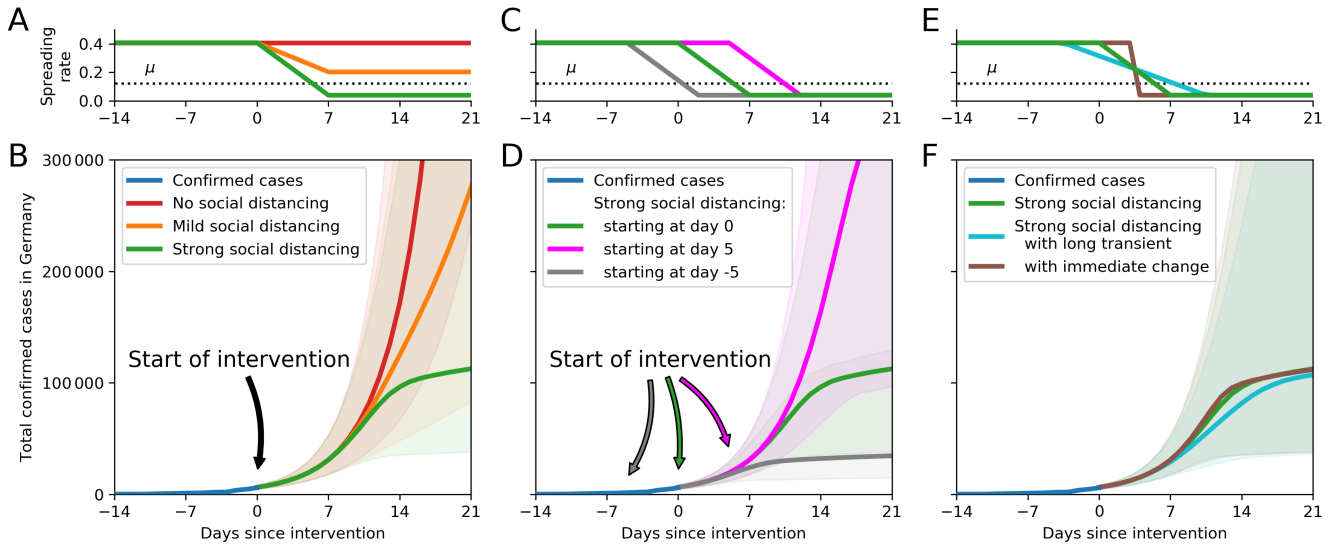

FIG. 2. The timing and effectiveness of interventions strongly impact future COVID-19 cases. **A, B:** We assume three different scenarios for interventions starting on March 16: (I) no social distancing – red, (II) mild social distancing – orange, or (III) strict social distancing – green. **C, D:** We also analyzed how a delayed restriction impacts case numbers: Strict restrictions starting on March 16 (green), or five days later or earlier. A delay (or advance) of five days in implementing restrictions has a major impact on the expected case numbers. **E, F:** Comparison of the time span over which interventions ramp up to full effect. For all ramps that are *centered around the same day*, the resulting case numbers are fairly similar.

Also, we still observe an exponential rise of new infections after the intervention becomes effective, because the growth rate remains positive, $\lambda_1^* = \lambda_1 - \mu > 0$. **(III) Strong social distancing;** Here, the spreading rate decreases to 10%, ($\lambda_1 = \lambda_0 / 10$ with median $\lambda_1 = 0.04 < \mu$). The assumptions here are that contacts are severely limited, but even when people stay at home as much as possible, *some* contacts are still unavoidable. Even under such drastic policy changes, no effect is visible until the reporting delay D is over. Thereafter, a quick decrease in daily new infections manifests within two weeks (delay plus change point duration), and the total number of cases reaches a stable plateau. Only in this last scenario a plateau is reached, because here the growth rate becomes negative, $\lambda^* < 0$, which leads to decreasing numbers of new infections.

Further more, the timing of an intervention matters: Apart from the strength of an intervention, its *onset* time has great impact on the total case number (Fig. 2 C,D). For example, focusing on the strong intervention (III) — where a stable plateau is reached — the effect of advancing or delaying the change point by just five days leads to more than a three-fold difference in cumulative cases. While we find that the timing of an intervention has great effect on case numbers, the duration over which the change takes place has only minor effect — if the intervals of change are centered around the same date. In Fig. 2 E,F we illustrate the adjustment of $\lambda_0 \to \lambda_1$ having durations of 14, 7 and 1 day(s). Note that the change point duration is a simple way to incorporate variability in individual behavior, and is not linked to the reporting delay D.

RESULTS

We incorporate the effect of governmental interventions into our model by introducing flexible change points in the spreading rate (see Methods). During the COVID-19 outbreak in Germany, governmental interventions occurred in three stages from (i) the cancellation of large events with more than 1000 participants (around March 9), through (ii) closing of schools, childcare centers and the majority of stores (in effect March 16), to (iii) the contact ban and closing of all non-essential stores (in effect March 23). The aim of all these interventions was to reduce the (effective) growth rate $\lambda^* = \lambda - \mu$. As soon as the growth rate becomes negative ($\lambda^* < 0$), the number of new confirmed infections

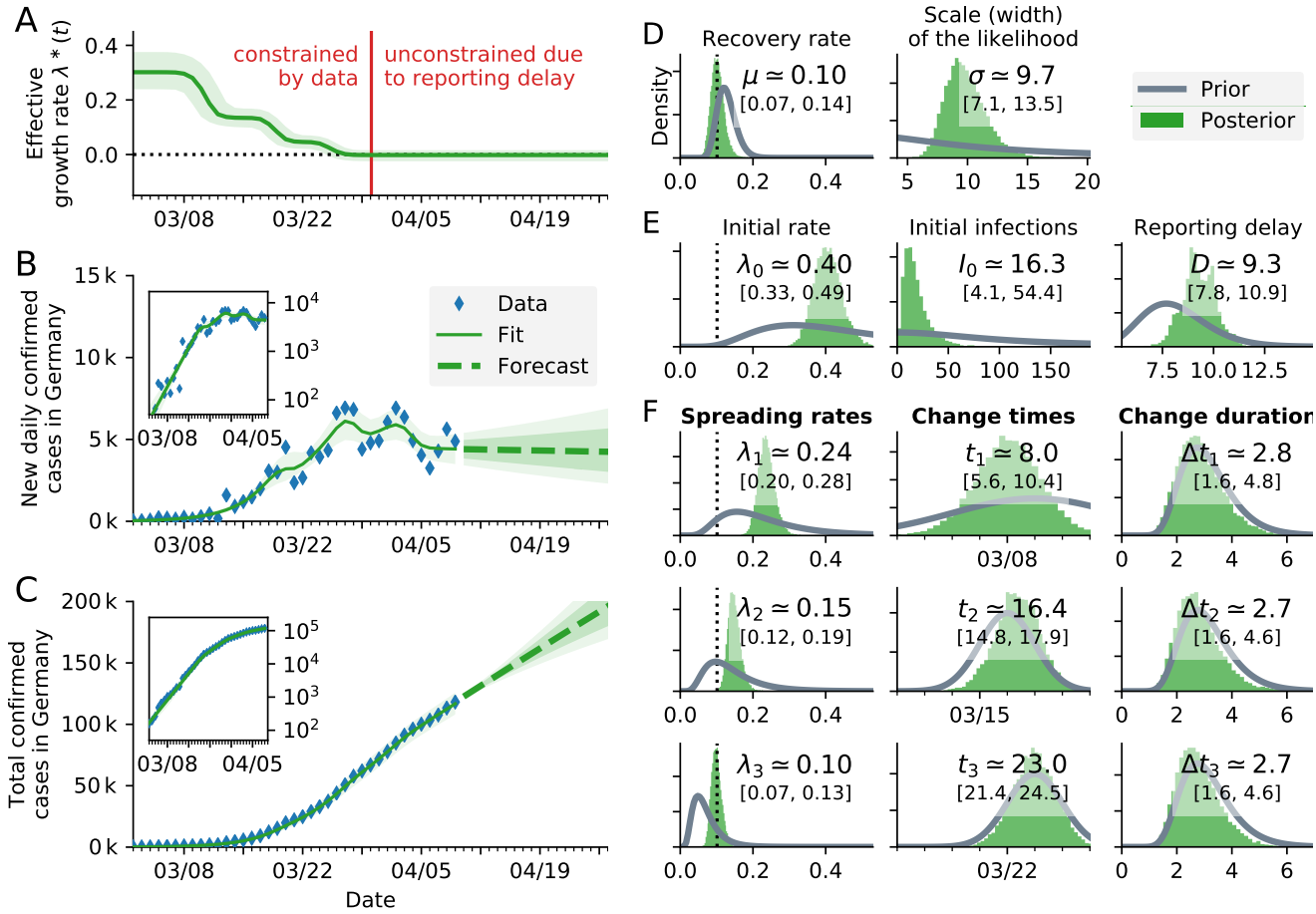

FIG. 3. The German COVID-19 data (blue diamonds) until April 9 indicate three change points, consistent with three major governmental interventions. **A:** Time-dependent model estimate of the effective spreading rate $\lambda^*(t)$. **B:** Comparison of daily recorded new cases and the model (green solid line for median fit with 95% confidence intervals, green dashed line for median forecast with 75% and 95% confidence intervals); **inset:** same data in log scale. **C:** Comparison of total recorded cases and the model (same representation as in B); **inset:** same data in log scale. **D-F:** Priors (gray lines) and posteriors (green histograms) of all model parameters; inset values indicate the median and 95% confidence intervals of the posteriors. For the same model with one or two change points, please see the corresponding figures in the SI (Fig. S1 and S2).

decreases after the respective reporting delay.

Detecting change points in the spreading rate — and quantifying the amount of change as quickly as possible — becomes a central modeling challenge when short-term forecasts are required. To address this challenge, we assume an initial spreading rate λ_0 (the exponential growth phase, cf. Fig. 1) and up to three potential change points motivated by the German governmental interventions: The first change point ($\lambda_0 \to \lambda_1$) is expected around March 9 (t_1) as a result from the official recommendations to cancel large events. A second change point ($\lambda_1 \to \lambda_2$) is expected around March 16 (t_2), when schools and many stores were closed. A third change point ($\lambda_2 \to \lambda_3$) is expected around March 23 (t_3), when all non-essential stores were closed, and a contact ban was enacted. We expect the behavioral changes introduced at these change points to unfold over a few days Δt_i, however, changes in duration can be partly compensated by changes in the onset time t_i (see Fig. 2 E,F, scenarios). We chose priors for all parameters based on the information available to us up to March 28 (see Methods). On that date, the data were already informative about the first change point, and thereby helped to inform our forecast scenarios.

The inferred parameters for the models with change points are consistent with the inferred parameters from the exponential onset phase (Fig. 1, Fig. 3 & Figs. S1- S2), wherever parameters were comparable. In particular, all estimated λ_0 from models with multiple change points are compatible with the spreading rate during the exponential onset phase ($\lambda = 0.41$, CI $[0.32, 0.51]$, assuming a stationary λ until March 15, Fig. 1 A). Also the scale factor σ and the number of initial infections I_0 for the models with change points are fully consistent with the initial model inference during the exponential onset phase (Fig. 1 F,H).

TABLE I. Model comparison using leave-one-out (LOO) cross-validation. Lower LOO-scores represent a better match between model and data.

Model	LOO-score	Effective number of parameters (pLOO)
zero change points	683 ± 12	4.09
one change point	629 ± 15	9.98
two change points	608 ± 16	7.84
three change points	602 ± 16	9.87

The models with two or three change points fit the observed data better than those with less change points

The model with three change points describes the data better than all other models, as indicated by the leave-one-out (LOO) cross-validation-based Bayesian model comparison [37] (lowest LOO-score in Table I). However, the model with two change points still has a comparable LOO-score within one standard deviation, which is expected since the present data (as of April 9) just barely entered the regime of the third change point – based on the known timing of the third intervention (March 23) and the estimate for the delay D (posterior median: 9.3 days). As expected, the models with none or a single change point have LOO-scores that are at least about one standard deviation higher (worse) than those of the best models and may be discarded.

When comparing our inference based on three change points to the number of confirmed cases, we find them to largely match (Fig. 3 B,C). Note that the tell-tale transient decrease in the data, which are expected independently of the weekend-related effects on sampling, are more evident in the raw number of newly confirmed cases (Fig. 3 B) than in the cumulative report (Fig. 3 C). Such a transient decrease of new cases, before increasing again, originates from changing an exponential growth rate over small time-interval in the model. It is consistent with the observed temporary drop in newly confirmed cases and suggests a rapid implementation of mitigation measures by the public. However, we also observed a spread in the data points that was somewhat larger than expected by the model. We assign this to the fact that in the main model we did not incorporate an additional prior describing uncertainty and noise that is introduced by fluctuations in reporting (less reports on weekends, availability of test kits, etc.) – however, we verified that our results are consistent when we extend our main model to account for the week-related alterations (Fig. S3 and S5). Given these and other additional sources of noise, we consider the match of model and data convincing.

Change point detection quantifies the effect of governmental interventions on the outbreak of COVID-19 in Germany

Ideally, detected changes can be related to specific mitigation measures, so that one gains an understanding about the effectiveness of different measures (Fig. 3). Indeed, we found clear evidence for three change points in the posterior distributions of the model parameters: First, $\lambda(t)$ decreased from $\lambda_0 = 0.40$ (with 95% confidence interval, CI $[0.33, 0.49]$) to $\lambda_1 = 0.24$ (CI [0.20, 0.28]). The date of the change point was inferred to be March 8 (CI [March 6, March 10])]; this inferred date matches the timing of the first governmental intervention including cancellations of large events, as well as increased awareness. After this first intervention, the (effective) growth rate $\lambda^*(t) = \lambda(t) - \mu$ decreased by more than a factor 2, from median $\lambda_0 - \mu = 0.3$ to median $\lambda_1 - \mu = 0.14$, given that the recovery rate was inferred as $\mu = 0.10$ (CI [0.07, 0.14]). Second, $\lambda(t)$ decreased from $\lambda_1 = 0.24$ to $\lambda_2 = 0.15$ (CI [0.12, 0.19]), which is larger than our prior assumption. The date of the change point was inferred to be March 16 (95% CI [15, 18])]; this inferred date matches the timing of the second governmental intervention including closing schools and some stores. After this second intervention, the median growth rate became $\lambda^*(t) = \lambda_2 - \mu = 0.05 \approx 0$ and is thus in the vicinity of the critical point, yet still positive. The first two interventions in Germany thereby mitigated the spread by drastically reducing the growth rate, but the spread of COVID-19 remained exponential. Third, $\lambda(t)$ decreased from $\lambda_2 = 0.15$ to $\lambda_3 = 0.10$ (CI [0.07, 0.13]). The date of the change point was inferred to be March 23 (CI [21, 25])]; this inferred date matches the timing of the third governmental intervention including contact ban and closing of all non-essential shops. Only after this third intervention, the median (effective) growth rate, $\lambda^*(t) = \lambda_3 - \mu = -0.009 < 0$ (CI $[-0.039, 0.015]$)], finally became minimally negative, pointing at a possible decrease in the number of new infections. We can thus clearly relate the change points to the governmental interventions and quantify their mitigation effect.

DISCUSSION

We presented a Bayesian approach for online-monitoring the effect of governmental interventions on the spread of an epidemic outbreak. This approach enabled a timely inference of the central epidemiological parameters for Germany – as well as three change points in the spreading rate – of an SIR model from the number of reported cases during the COVID-19 outbreak. We showed that change points in the spreading rate affect the confirmed case numbers with a delay of about two weeks (median reporting delay of $D = 9.3$ days plus a median duration of changes of 3 days). Thereby, we were able to relate the inferred change points to the three major governmental interventions in Germany: We found a clear reduction of the spreading rate related to each governmental intervention (Fig. 3), (i) the cancellation of large events with more than 1000 participants (around March 9), (ii) the closing of schools, childcare centers and the majority of stores (in effect March 16), and (iii) the contact ban and closing of all non-essential stores (in effect March 23).

In absolute terms, the first two governmental interventions brought a reduction of the initial growth rate from $\lambda_0^* = 30\%$ to 14% and then down to 5%. However, these numbers still implied exponential growth. Only with the third intervention, i.e. the contact ban, we found first evidence that we have reached the transition in new case numbers from growth to decay. The decay rate of about -1% (CI $[-4\%, 2\%]$) is close to zero. This still implies that already a minor increase in the spreading rate may again switch the dynamics to the unstable regime with exponential growth.

We used a formal Bayesian model comparison in order to validate the presence of change points. Our model comparison ruled out models with less than two change points (Tab. I,V). While this may seem trivial, it has important consequences for making the short-term forecasts that decision makers rely on. Demonstrating and quantifying the effect of change points in the past enables us to project the effects of recent change points, that are not apparent in the observed case numbers yet, into future forecasts. Hence, it is important to look out for and identify potential change points as early as possible, and incorporate them appropriately into forecasts.

The detection of change points and their interpretation depend crucially on an accurate estimate of the reporting delay D. Therefore, the validity of its estimate should be evaluated. In our model D contains at least three distinct factors: the biological incubation period (median 5-6 days) [32], an additional delay from first symptoms to symptoms motivating a test (1-2 days) and possible delay before a testing results come in (1-2 days). The sum of these delays seems compatible with our inferred median delay of $D = 9.3$ days, especially given the wide range of reported incubation periods.

We chose to keep our main model comparatively simple, because of the small number of data points available during an epidemic outbreak. With such a low number of data points only a limited number of parameters can be effectively constrained. Hence, we chose to approximate a time-dependent spreading rate $\lambda(t)$ by episodes of constant spreading rates λ_i that are separated by three change points where a transition occurs. Our results show that this main model is currently sufficient for Germany: While we introduced fairly broad priors on the spreading rates, we obtained fairly narrow posterior distributions for each spreading rate λ_i (Fig. 3). We additionally evaluated obvious extensions of our main model with three change points by excluding weekend data that may suffer from under-reporting (Fig. S3), and by using the more involved model that includes the four classes: susceptible, exposed, infected, recovered (Fig. S4). All of these models yield consistent results for the three change points, and all have LOO scores within one standard error of each other. Thus, we consider our main model to be sufficient for case numbers in Germany at present.

Our framework can be easily adapted to other countries and enables one to incorporate future developments. For other countries, or for forecasts within smaller communities (e.g. federal states or cities), additional details may become important, such as explicit modeling of incubation time distributions [17, 38] (i.e. as done in Fig. S4) , spatial heterogeneity [17, 21], isolation effects [20, 38], subsampling effects hiding undetected cases even beyond the reporting delay [39, 40], or the age and contact structure of the population [26]. In countries where drastic changes in test coverage are expected this will have to be included as well. The methodology presented here is capable in principle of incorporating such details. It also lends itself to modeling of continuous drifts in the spreading rate, e.g. reflecting reactions of the public to news coverage of a catastrophic situation, or people growing tired of mitigation measures. Such further adaptations, however, can only be performed on a per-country basis by experts with an intimate knowledge of the local situation. Our code provides a solid and extensible base for this. For Germany, several developments in the near future may have to be included in the model. First, people may transiently change their behavior over the Easter holidays; second, we expect a series of change points, as well as continuous drifts, with governments trying to ease and calibrate mitigation measures.

Effective growth rates in Germany are still around zero and warrant careful consideration of future interventions. At present, with all mitigation measures in place, effective growth rates are very close to zero – the watershed between exponential growth or decay. Together with the delay of approximately two weeks between infection and case report

this warrants utmost caution in easing restrictions for two reasons: First, easing restrictions too much will quickly lead to renewed exponential growth; second, we would be effectively blind to this worsened situation for nearly two weeks in which it will develop uninhibited. This may result in unwanted growth in case numbers beyond the level that the health system can cope with – especially when the active cases have not gone down close to zero before lifting restrictions, thus re-initiating growth from a high base level. Therefore, it is important to consider lifting restriction only when the number of active cases are so low that a two-week increase will not pose a serious threat.

In conclusion, the presented Bayesian approach allows to detect and quantify the effect of recent governmental interventions and – combined with potential subsequent interventions – to forecast future case number scenarios. Our analysis highlights the importance of precise timing and magnitude of interventions for future case numbers. It also stresses the importance of including the reporting delay D between the date of infection and the date of the confirmed case in the model . The delay D, together with the time required to implement interventions causes a total delay between an intervention and its visibility in the case numbers of about two weeks for COVID-19 in Germany. This means that changes in our behavior today can only be detected in confirmed cases in two weeks. Combined with the current spreading rate that is still around zero, the inferred spreading and observation dynamics warrant an extremely careful planning of future measures.

MATERIALS AND METHODS

As a basis for our forecast scenarios, we use the differential equations of the well-established SIR (Susceptible-Infected-Recovered) model. Case data comes from the COVID-19 data repository maintained by the Johns Hopkins University Center for Systems Science and Engineering [41]. While the model dynamics is well understood in general, here our main challenge is to estimate model parameters specifically for the COVID-19 outbreak. To that end, we combined a Bayesian approach — to incorporate prior knowledge — with Markov Chain Monte Carlo (MCMC) sampling — to explore the parameters. Put simply, we first estimate the parameter distribution that best describes the observed situation, and then we use many samples from this parameter distribution to evolve the model equations and thus forecast future developments.

Simple model: SIR model with stationary spreading rate

We consider a time-discrete version of the standard SIR model. In short, we assume that the disease spreads at rate λ from the infected population stock (I) to the susceptible population stock (S), and that the infected stock recovers (R) at rate μ. This well-established model for disease spreading can be described by the following set of (deterministic) ordinary differential equations (see, e.g., Refs [5, 6, 20]). Within a population of size N,

$$\begin{aligned} \tfrac{\mathrm{d}S}{\mathrm{d}t} &= -\lambda \tfrac{SI}{N} \\ \tfrac{\mathrm{d}I}{\mathrm{d}t} &= \lambda \tfrac{SI}{N} - \mu I \\ \tfrac{\mathrm{d}R}{\mathrm{d}t} &= \mu I\,. \end{aligned} \tag{1}$$

As a remark, during the onset phase of an epidemic only a very small fraction of the population is infected (I) or recovered (R), and thus $S \approx N \gg I$ such that $S/N \approx 1$. Therefore, the differential equation for the infected reduces to a simple linear equation, exhibiting an exponential growth

$$\frac{dI}{dt} = (\lambda - \mu) I \quad \text{solved by} \quad \mathrm{I(t) = I(0)\ e^{(\lambda-\mu)t}}\,. \tag{2}$$

Because our data set is discrete in time (Δt =1 day), we solve the above differential equations with a discrete time step ($\mathrm{d}I/\mathrm{d}t \approx \Delta I/\Delta t$), such that

$$\begin{aligned} S_t - S_{t-1} &= -\lambda \Delta t \tfrac{S_{t-1}}{N} I_{t-1} &&=: -I_t^{\mathrm{new}} \\ R_t - R_{t-1} &= \mu \Delta t I_{t-1} &&=: \quad R_t^{\mathrm{new}} \\ I_t - I_{t-1} &= \left(\lambda \tfrac{S_{t-1}}{N} - \mu\right) \Delta t I_{t-1} = I_t^{\mathrm{new}} - R_t^{\mathrm{new}}\,. \end{aligned} \tag{3}$$

Importantly, I_t models the number of all (currently) active infected people, while I_t^{new} is the number of new infections that will eventually be reported according to standard WHO convention. Importantly, we explicitly include a reporting

delay D between new infections I_t^{new} and newly reported cases (C_t) as

$$C_t = I_{t-D}^{\text{new}}. \tag{4}$$

We begin our simulations at time $t = 0$ with I_0 infected cases and start including real-word data of reported cases $\hat{C}_t$ from day $t > D$ (see below for a parameterization).

Full model: SIR model with change points in spreading rate

Our change point detection builds on a generalization of the simple SIR model with stationary spreading rate. Instead, we now assume that the spreading rate λ_i, $i = 1, ..., n$, may change at certain time points t_i from λ_{i-1} to λ_i, linearly over a time window of Δt_i days. Thereby, we account for policy changes to reduce λ, which were implemented in Germany step by step. Thus, the parameters t_i, Δt_i, and λ_i are added to the parameter set of the simple model above, and the differential equations are augmented by the time-varying λ_i.

TABLE II. Overview of model parameters.

Variable	Parameter
$\theta = \{\lambda_i, t_i, \mu, \sigma, I_0\}$	Set of model parameters that are optimized
λ	Spreading rate
μ	Recovery rate
$\lambda^* = \lambda - \mu$	Effective spreading rate
λ_i	Spreading rate after i-th intervention
t_i	Time of i-th intervention
σ	Scale factor of the width of Student's t-distribution
N	Population size (83.700.000)
S_t	Susceptible at time t
I_t	Infected at time t
R_t	Recovered at time t
Δt	Time step
$R_t^{\text{new}} = \mu \Delta t I_{t-1}$	New recoveries at time t
$I_t^{\text{new}} = \lambda \Delta t \frac{S_{t-1}}{N} I_{t-1}$	New infections at time t
$C_t = I_{t-D}^{\text{new}}$	New reported cases at time t
D	Delay of case detection

Estimating model parameters with Bayesian MCMC

We estimate the set of model parameters $\theta = \{\lambda_i, t_i, \mu, D, \sigma, I_0\}$ using Bayesian inference with Markov-chain Monte-Carlo (MCMC). The parameter σ is the scale factor for the width of the likelihood $P\big(\hat{C}_t \big| \theta\big)$ between observed data and model (see below). Our implementation relies on the python package pymc3 [42] with NUTS (No-U-Turn Sampling) [43]. The structure of our approach is the following:

Choose random initial parameters and evolve according to model equations. Initially, we choose parameters θ from prior distributions that we explicitly specify below. Then, time integration of the model equations generates a (fully deterministic) time series of new infected cases $C(\theta) = \{C_t(\theta)\}$ of the same length as the observed real-world data $\hat{C} = \left\{\hat{C}_t\right\}$.

Iteratively update the parameters using MCMC. The drawing of new candidate parameters and the time integration of the SIR model is repeated in every MCMC step. The idea is to probabilistically draw parameter updates and to accept them such that the deviation between the model outcome and the available real-world time-series is likely to reduce. We quantify the inverse deviation (which needs to be maximized) between the model outcome for one time point t, $C_t(\theta)$ and the corresponding real-world data point $\hat{C}_t$ with the local likelihood

$$p\big(\hat{C}_t \big| \theta\big) \sim \text{StudentT}_{\nu=4}\left(\text{mean} = \text{C}_\text{t}(\theta),\, \text{width} = \sigma\sqrt{\text{C}_\text{t}(\theta)}\right).$$

We chose the Student's t-distribution because it resembles a Gaussian distribution around the mean but features heavy tails, which make the MCMC more robust with respect to outliers [44]. The case-number-dependent width is motivated

by the demographic noise of typical mean-field solutions for epidemic spreading, e.g., $\dot{\rho}(t) = a\rho(t) - b\rho^2(t) + \sqrt{\rho(t)}\eta(t)$, where ρ is the activity and $\eta(t)$ is Gaussian white noise [35, 36]. This choice is consistent with our data (Fig. 1 A-C). The overall deviation is then simply the product of local likelihoods over all time points.

For each MCMC step, the new parameters are drawn so that a set of parameters that minimizes the previous deviation is more likely to be chosen. In our case, this is done with an advanced gradient-based method (NUTS [43]). To reiterate, every time integration that is performed has its own set of parameters and yields one complete model time series. If the time integration describes the data well the parameter set is accepted, and this yields one Monte Carlo sample of the model parameters for the posterior distribution; the MCMC step is then repeated to create more samples from the posterior. Eventually, the majority of accepted parameter samples will describe the real-world data well, so that consistent forecasts are possible in the forecast phase.

Forecast using Monte Carlo samples. For the forecast, we take all samples from the MCMC step and continue time integration according to different forecast scenarios explained below. Note that the overall procedure yields an ensemble of forecasts — as opposed to a single forecast that would be solely based on one set of (previously optimized) parameters.

Priors that constrain model parameters

As short-term forecasts are time-critical at the onset of an epidemic, the available real-world data is typically not informative enough to identify all free parameters, or to empirically find their underlying distributions. We therefore chose informative priors on initial model parameters where possible and complemented them with uninformative priors otherwise. Our choices are summarized in Tab. III for the simple model, SIR model with stationary spreading rate for the exponential onset phase, and in Tab. IV for the full model with change points, and justified in the following.

TABLE III. Priors on the model parameters for the SIR model with stationary spreading rate.

Parameter	Variable	Prior distribution
Spreading rate	λ	$\text{LogNormal}(\log(0.4), 0.5)$
Recovery rate	μ	$\text{LogNormal}(\log(1/8), 0.2)$
Reporting delay	D	$\text{LogNormal}(\log(8), 0.2)$
Initially infected	I_0	$\text{HalfCauchy}(100)$
Scale factor	σ	$\text{HalfCauchy}(10)$

Priors for the simple model (Table III): In order to constrain our simple model, an SIR model with stationary spreading rate for the exponential onset phase, we chose the following informative priors. Because of the ambiguity between the spreading and recovery rate in the exponential onset phase (see description of simple model), we chose a narrow log-normal prior for the recovery rate $\mu \sim \text{LogNormal}(\log(1/8), 0.2)$ with median recovery time of 8 days [20]. Note that, our implementation of μ accounts for the recovery of infected people and isolation measures because it describes the duration during which a person can infect others. For the spreading rate, we assume a broad log-normal prior distribution $\lambda \sim \text{LogNormal}(\log(0.4), 0.5)$ with median 0.4. This way, the prior for $\lambda - \mu$ has median 0.275 and the prior for the base reproduction number ($R_0 = \lambda/\mu$) has median 3.2, consistent with the broad range of previous estimates [18, 33, 34]. In addition, we chose a log-normal prior for the reporting delay $D \sim \text{LogNormal}(\log(8), 0.2)$ to incorporate both the incubation time between 1–14 days with median 5 [32] plus a delay from infected people waiting to contact the doctor and get tested.

The remaining model parameters are constrained by uninformative priors, in practice the Half-Cauchy distribution [45]. The half-Cauchy distribution $\text{HalfCauchy}(\mathrm{x}, \beta) = 2/\pi\beta[1 + (\mathrm{x}/\beta)^2]$ is essentially a flat prior from zero to $O(\beta)$ with heavy tails beyond. Thereby, β merely sets the order of magnitude that should not be exceeded for a given parameter. We chose for the number of initially infected people in the model (16 days before first data point) $I_0 \sim \text{HalfCauchy}(100)$ assuming an order of magnitude $O(100)$ and below. In addition, we chose of the scale factor of the width of the likelihood function $\sigma \sim \text{HalfCauchy}(10)$, which is further multiplied to the number of new cases.

Priors for the full model (Table IV): In order to constrain our full model, an SIR model with change points in the spreading rate, we chose the same priors as for the simple model but added the required priors associated with the change points. In general, we assume that each set of governmental interventions (together with the increasing awareness) leads to a reduction (and not an increase) of λ at a change point. As we cannot know yet the precise reduction factor, we adhere to assume a reduction by 50%, but always with a fairly wide uncertainty, so that in principle even an increase at the change point would be possible. We model the time dependence of λ as change points, and not as continuous changes, because the policy changes were implemented in three discrete steps, which were

TABLE IV. Priors on the model parameters for the SIR model with change points.

Parameter	Variable	Prior distribution
Change points	t_1	Normal$(2020/03/09, 3)$
	t_2	Normal$(2020/03/16, 1)$
	t_3	Normal$(2020/03/23, 1)$
Change duration	Δt_i	LogNormal$(\log(3), 0.3)$
Spreading rates	λ_0	LogNormal$(\log(0.4), 0.5)$
	λ_1	LogNormal$(\log(0.2), 0.5)$
	λ_2	LogNormal$(\log(1/8), 0.2)$
	λ_3	LogNormal$(\log(1/16), 0.2)$
Recovery rate	μ	LogNormal$(\log(1/8), 0.2)$
Reporting delay	D	LogNormal$(\log(8), 0.2)$
Initially infected	I_0	HalfCauchy(100)
Scale factor	σ	HalfCauchy(10)

presumably followed by the public in a timely fashion. Continuous changes, originating e.g. from increased awareness of the population can be accounted for by the discrete steps as well, within the precision of reported cases we have.

For the spreading rates, we chose log-normal distributed priors as in the simple model. In particular, we chose for the initial spreading rate the same prior as in the simple model, $\lambda_0 \sim \text{LogNormal}(\log(0.4), 0.5)$; after the first change point $\lambda_1 \sim \text{LogNormal}(\log(0.2), 0.5)$, assuming the first intervention to reduce the spreading rate by 50% from our initial estimates ($\lambda_0 \approx 0.4$) with a broad prior distribution; after the second change point $\lambda_2 \sim \text{LogNormal}(\log(1/8), 0.2)$, assuming the second intervention to reduce the spreading rate to the level of the recovery rate, which would lead to a stationary number of new infections. This corresponds approximately to a reduction of λ at the change point by 50%; and after the third change point $\lambda_3 \sim \text{LogNormal}(\log(1/16), 0.2)$, assuming the third intervention to reduce the spreading rate again by 50%. With that intervention, λ_3 is smaller than the recovery rate μ, causing a decrease in new case numbers and a saturation of the cumulative number of infections.

For the timing of change points, we chose normally distributed priors. In particular, we chose $t_1 \sim \text{Normal}(2020/03/09, 3)$ for the first change point because on the weekend of March 8, large public events, like visits to soccer matches or fairs, were cancelled. For the second change point, we chose $t_2 \sim \text{Normal}(2020/03/16, 1)$, because on March 15, the closing of schools and other educational institutions along with the closing of non-essential stores were announced and implemented on the following day. Restaurants were allowed to stay open until 6 pm. For the third change point, we chose $t_3 \sim \text{Normal}(2020/03/23, 1)$, because on March 23, a far-reaching contact ban ("Kontaktsperre"), which includes the prohibition of even small public gatherings as well as complete closing of restaurants and non-essential stores was imposed by the government authorities. Further policy changes may occur in future; however, for now, we do not include more change points.

The change points take effect over a certain time period Δt_i for which we choose $\Delta t_i \sim \text{LogNormal}(\log(3), 0.3)$ with a median of 3 days over which the spreading rate changes continuously as interventions have to become effective. The precise duration of the transition has hardly any affect on the cumulative number of cases (Fig. 2 E-F). We assumed a duration of three days, because some policies were not announced at the same day for all states within Germany; moreover, the smooth transition also can absorb continuous changes in behavior.

Model comparison

Since change point detection entails evaluating models with different numbers of parameters, some form of fair model comparison needs to be performed. Here, we compared the models with different numbers of change points by their pointwise out-of-sample prediction accuracy using the log-likelihood evaluated at the posterior simulations of the parameter values obtained from the fitted models. Out-of-sample accuracy was approximated using Leave-one-out cross-validation (LOO) [37].

* viola.priesemann@ds.mpg.de; Authors contributed equally

[1] M. Enserink, K. Kupferschmidt, *Science* **367**, 1414 (2020).
[2] E. T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge University Press, 2003).
[3] A. Gelman, *et al.*, *Bayesian Data Analysis, Third Edition* (CRC Press, 2013).

[4] W. O. Kermack, A. G. McKendrick, G. T. Walker, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **115**, 700 (1927).
[5] H. Hethcote, *SIAM Rev.* **42**, 599 (2000).
[6] J. Anderson, I. Lampl, I. Reichova, M. Carandini, D. Ferster, *Nat. Neurosci.* **3**, 617 (2000).
[7] N. C. Grassly, C. Fraser, *Nat Rev Microbiol* **6**, 477 (2008).
[8] R. Parshani, S. Carmi, S. Havlin, *Phys. Rev. Lett.* **104**, 258701 (2010).
[9] T. Harko, F. S. N. Lobo, M. K. Mak, *Applied Mathematics and Computation* **236**, 184 (2014).
[10] T. Britton, P. D. O'Neill, *Scand J Stat* **29**, 375 (2002).
[11] J. Lourenço, *et al.*, *eLife* **6**, e29820 (2017).
[12] N. R. Faria, *et al.*, *Sci Rep* **7**, 1 (2017).
[13] B. Shulgin, L. Stone, Z. Agur, *Bull. Math. Biol.* **60**, 1123 (1998).
[14] O. N. Bjørnstad, B. F. Finkenstädt, B. T. Grenfell, *Ecol. Monogr.* **72**, 169 (2002).
[15] L. Hufnagel, D. Brockmann, T. Geisel, *Proc. Natl. Acad. Sci. USA* **101**, 15124 (2004).
[16] A. Pandey, *et al.*, *Science* **346**, 991 (2014).
[17] R. Li, *et al.*, *Science* (2020).
[18] A. J. Kucharski, *et al.*, *The Lancet Infectious Diseases* (2020).
[19] J. Lourenco, *et al.*, *medRxiv* p. 2020.03.24.20042291 (2020).
[20] B. F. Maier, D. Brockmann, *arXiv:2002.07572* (2020).
[21] P. Bittihn, R. Golestanian, *arXiv:2003.08784* (2020).
[22] R. M. Anderson, H. Heesterbeek, D. Klinkenberg, T. D. Hollingsworth, *The Lancet* **395**, 931 (2020).
[23] J. R. Fauver, *et al.*, *medRxiv* p. 2020.03.25.20043828 (2020).
[24] A. Arenas, *et al.*, *medRxiv* p. 2020.03.21.20040022 (2020).
[25] O. Mitjà, *et al.*, *The Lancet* (2020).
[26] S. L. Chang, N. Harding, C. Zachreson, O. M. Cliff, M. Prokopenko, *arXiv:2003.10218* (2020).
[27] W. Bock, *et al.*, *medRxiv* p. 2020.03.25.20043109 (2020).
[28] C. Gros, R. Valenti, K. Valenti, D. Gros, *arXiv:2004.00493* (2020).
[29] V. Zlatić, I. Barjašić, A. Kadović, H. Štefančić, A. Gabrielli, *arXiv:2003.08479* (2020).
[30] J. Dehning, *et al.*, *ArXiv200401105 Version 1* (2020).
[31] https://github.com/Priesemann-Group/covid19_inference_forecast.
[32] S. A. Lauer, *et al.*, *Ann Intern Med* (2020).
[33] J. Zhang, *et al.*, *medRxiv* p. 2020.03.19.20039107 (2020).
[34] Y. Liu, A. A. Gayle, A. Wilder-Smith, J. Rocklöv, *J Travel Med* **27** (2020).
[35] S. di Santo, P. Villegas, R. Burioni, M. A. Muñoz, *Phys. Rev. E* **95** (2017).
[36] M. A. Muñoz, *Rev. Mod. Phys.* **90**, 031001 (2018).
[37] A. Vehtari, A. Gelman, J. Gabry, *Stat Comput* **27**, 1413 (2017).
[38] L. Peng, W. Yang, D. Zhang, C. Zhuge, L. Hong, *arXiv:2002.06563* (2020).
[39] J. Wilting, V. Priesemann, *Nat. Commun.* **9**, 2325 (2018).
[40] Y.-C. Chen, P.-E. Lu, C.-S. Chang, T.-H. Liu, *arXiv:2003.00122* (2020).
[41] https://github.com/CSSEGISandData/COVID-19.
[42] J. Salvatier, T. V. Wiecki, C. Fonnesbeck, *PeerJ Comput. Sci.* **2**, e55 (2016).
[43] M. D. Hoffman, A. Gelman, *J. Mach. Learn. Res.* **15**, 1593 (2014).
[44] K. L. Lange, R. J. A. Little, J. M. G. Taylor, *J. Am. Stat. Assoc.* **84**, 881 (1989).
[45] A. Gelman, *Bayesian Anal.* **1**, 515 (2006).

ACKNOWLEDGMENTS

We thank Tim Friede, Theo Geisel, Knut Heidemann, Moritz Linkmann, Matthias Loidolt, and Vladimir Zykov for carefully and promptly reviewing our work internally. We thank the Priesemann group - especially Daniel Gonzalez Marx, Fabian Mikulasch, Lucas Rudelt & Andreas Schneider - for exciting discussions and for their valuable comments. We thank the colleagues of the Göttingen Campus, with whom we were discussing the project and the COVID-19 case forecast in the past weeks very intensively: Heike Bickeböller, Eberhard Bodenschatz, Wolfgang Brück, Alexander Ecker, Andreas Leha, Ramin Golestanian, Helmut Grubmüller, Stephan Herminghaus, Reinhard Jahn, Norbert Lossau & Simone Scheithauer. We thank Nils Bertschinger for stimulating discussion on Bayesian modeling and model comparison. All authors received support from the Max-Planck-Society. JD and PS acknowledge funding by SMARTSTART, the joint training program in computational neuroscience by the VolkswagenStiftung and the Bernstein Network. JZ received finanicial support from the Joachim Herz Stiftung. MW is employed at the Campus Institute for Dynamics of Biological Networks funded by the VolkswagenStiftung.

SUPPLEMENTARY MATERIAL

TABLE V. Using leave-one-out (LOO) cross-validation, we compare advanced models (SIR with an adapted reporting rate during weekends, SEIR with explicit incubation time) with the model of the main paper (SIR main). For more details on the model variants, see the figure captions of the respective models in the SI. Lower LOO-scores represent a better match between model and data.

Model	# c-pts.	LOO-score	eff. # parameters (pLOO)
SIR main	0	683 ± 12	4.09
SIR main	1	629 ± 15	9.98
SIR main	2	608 ± 16	7.84
SIR main	3	602 ± 16	9.87
SIR w/ abs-sine mod. weekend	2	596 ± 16	12.21
SIR w/ abs-sine mod. weekend	3	588 ± 17	12.24
SEIR	3	613 ± 15	8.94

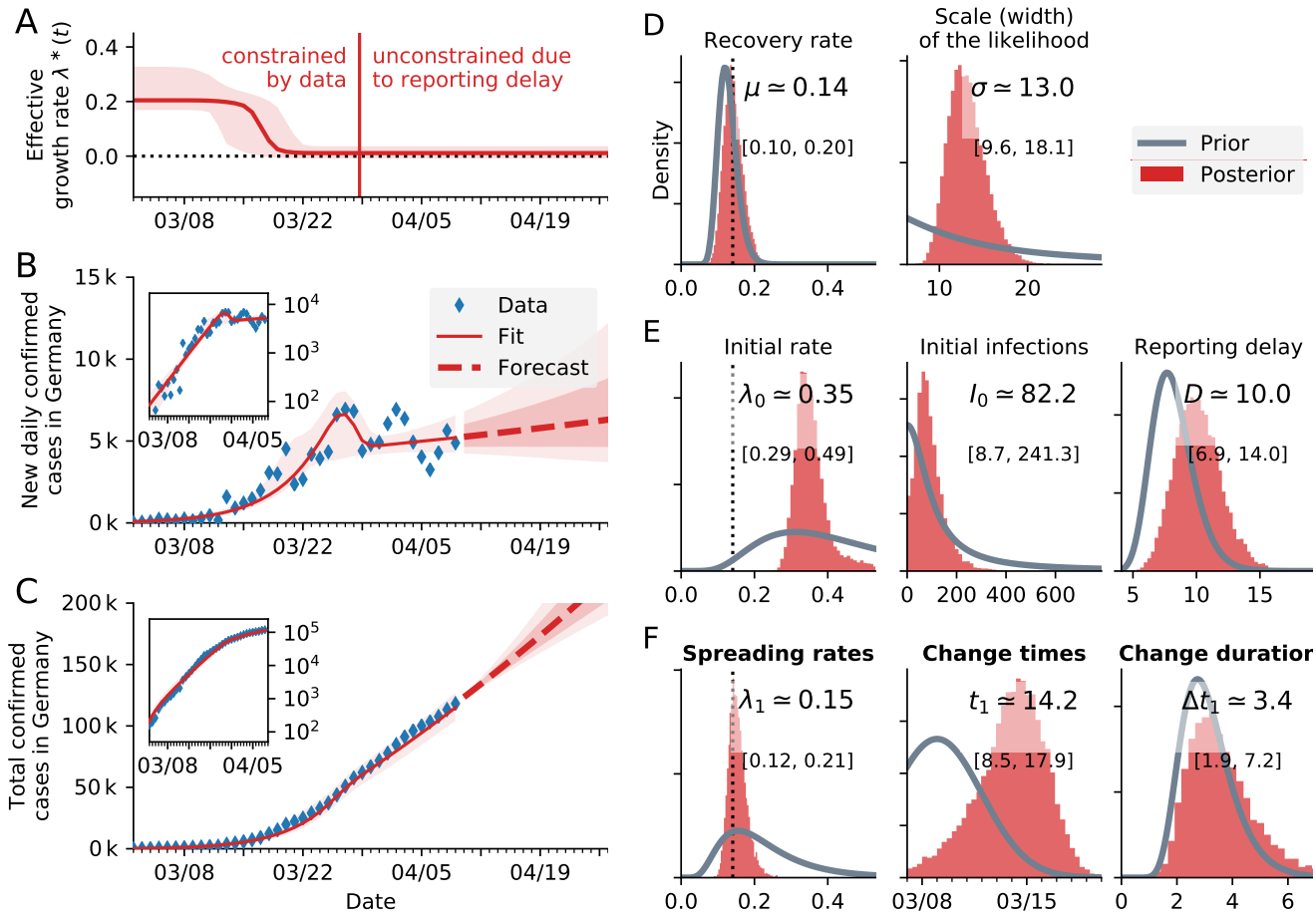

FIG. S1. Inferring (the time of) the change point in spreading rate λ, **for the model with only one change point**. Please refer to Fig. 3 for the case with three change points. **A:** Time-dependent model estimate of the effective growth rate $\lambda^*(t)$. **B:** Forecast and comparison of daily recorded new cases with the model fit. With only one change point, the model cannot describe the data well after April 1, **inset:** Same on log log scale. **C:** Same as B but for cumulative (total) cases. **D–F:** Posterior distributions from the change point detection (red) compared to prior distributions (gray). Please refer to Fig. 1 for a more detailed description of the distributions.

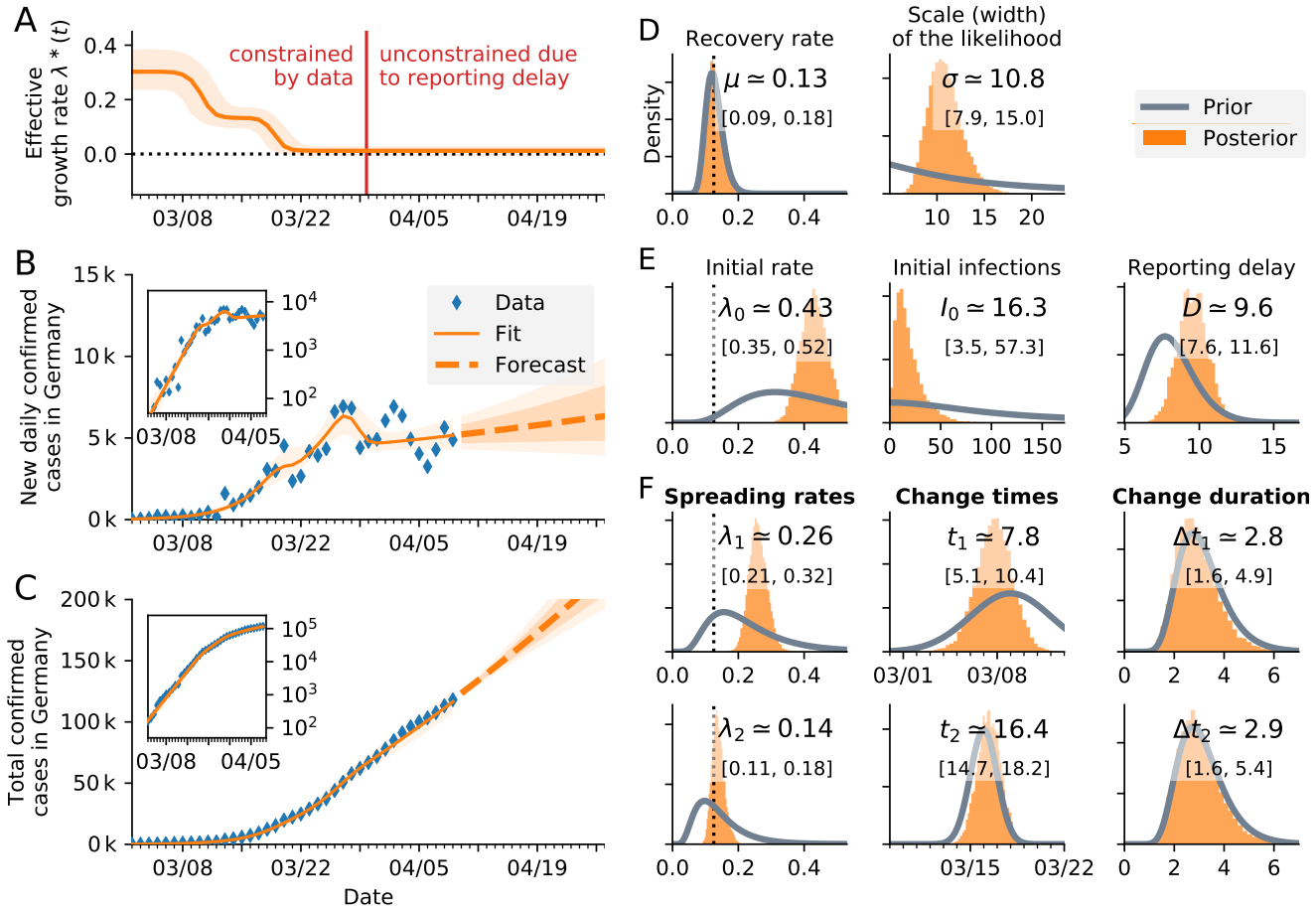

FIG. S2. Inferring (the time of) change points in spreading rate λ, **for the model with two change points**. Please refer to Fig. 3 for the case with three change points. **A:** Time-dependent model estimate of the effective growth rate $\lambda^*(t)$. Two change points are clearly visible. **B:** Forecast and comparison of daily recorded new cases with the model fit, **inset:** Same on log log scale. **C:** Same as B but for cumulative (total) cases. **D–F:** Posterior distributions from the change point detection (orange) compared to prior distributions (gray). Please refer to Fig. 1 for a more detailed description of the distributions.

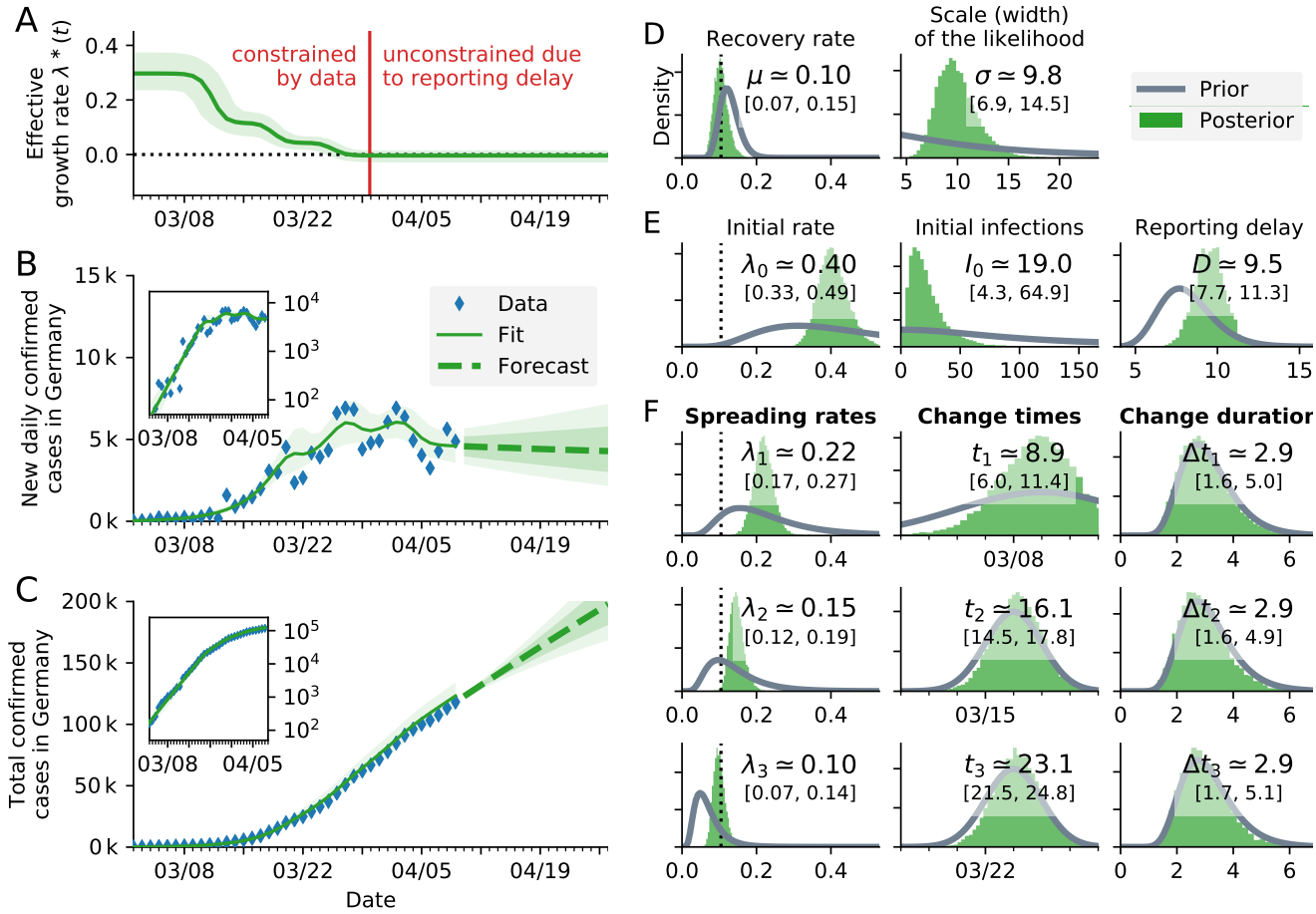

FIG. S3. As a check of our main model, we compare with a more complex model: here, we still consider three change points, but **here the fitting does not take into account the case reports obtained on weekends**, because these tend to be lower than during the week. Weekends were simply taken out of the likelihood. **A:** Time-dependent model estimate of the spreading rate λ_t. **B:** Model forecast of new cases, based on the inferred λ_t, linear scale (inset: log scale). **C:** Model forecast of total cases, based on the inferred λ_t, linear scale (inset: log-scale). **D–F:** Comparison of estimated (posterior) parameter distributions with initially assumed (prior) distributions.

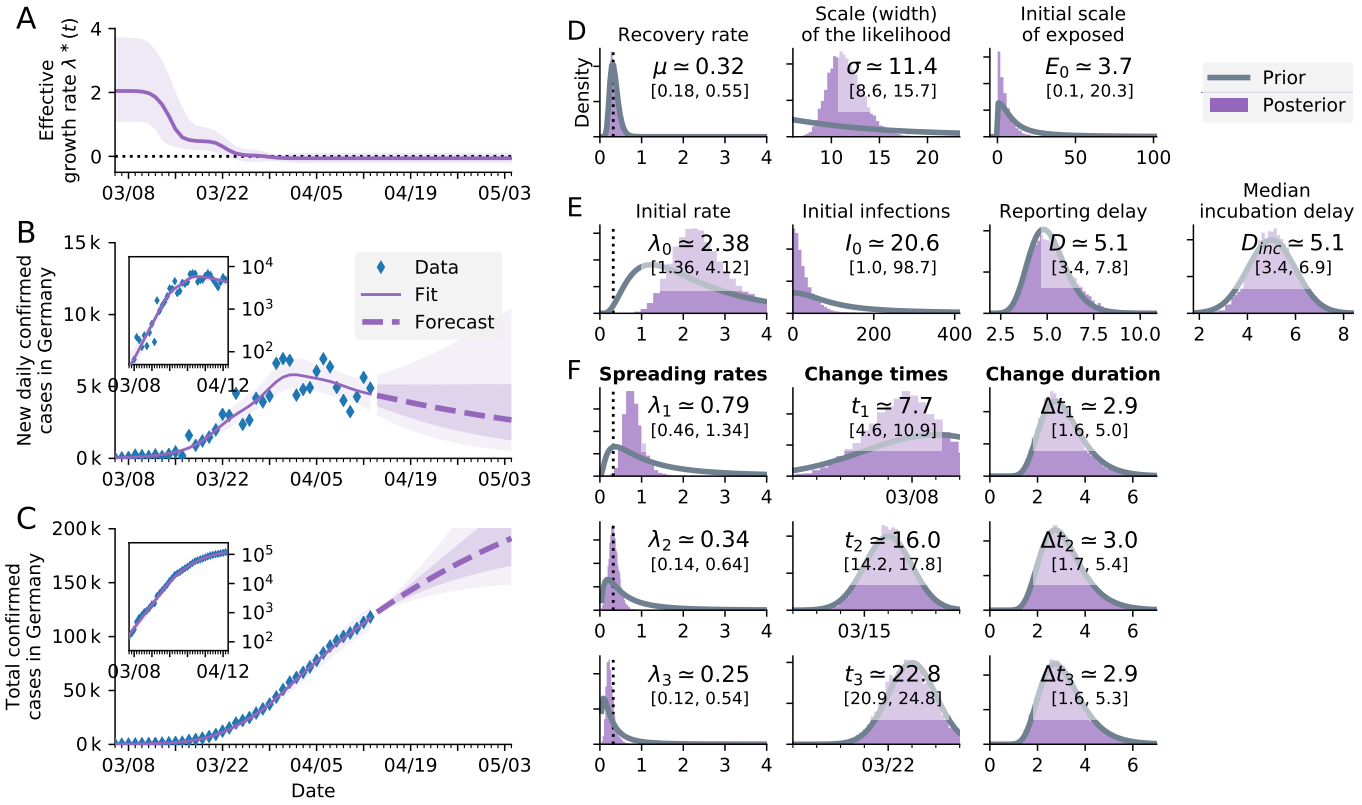

FIG. S4. Inference with a more involved **SEIR-*like*** model. Two additions were made to the SIR-model: **(1)** It includes an **explicit incubation period** during which infected people are not infectious, in the spirit of a SEIR model. In contrast to usual SEIR models, the length of incubation period is not exponentially distributed but has a lognormal distribution to match the characteristic incubation time of COVID-19. The incubation period has as prior a normal distribution with median Normal$(5, 1)$ (days) and a scale parameter σ of 0.418 [32]. **(2)** People that are **infectious are observed with a delay that is now lognormal distributed**. In the prior SIR model we assumed a fixed delay between infection and observation. The delay has a scale parameter σ of 0.3 and as median a LogNormal$(5, 0.2)$ (days) prior, to match approximately the total delay between infection and observation of the previous model. We changed the prior for the recovery rate μ to a median of $1/3$, which is similar to other SEIR simulation studies [17]. The priors for λ_0 to λ_3 were increased to 2,1, 0.5 and 0.25 respectively and a scale parameter of 1. **A:** Time-dependent model estimate of the effective growth rate $\lambda^*(t)$. Note that $\lambda^*(t)$ in the SEIR-like model is not directly comparable to $\lambda^*(t)$ in SIR models, because of the lognormal-distributed incubation period, decreasing the effective growth rate. **B:** Comparison of daily recorded new cases and the model, linear scale (inset: log scale). **C:** Comparison of total recorded cases and the model, linear scale (inset: log-scale). **D–F**: Prior and posterior distributions of all free parameters.

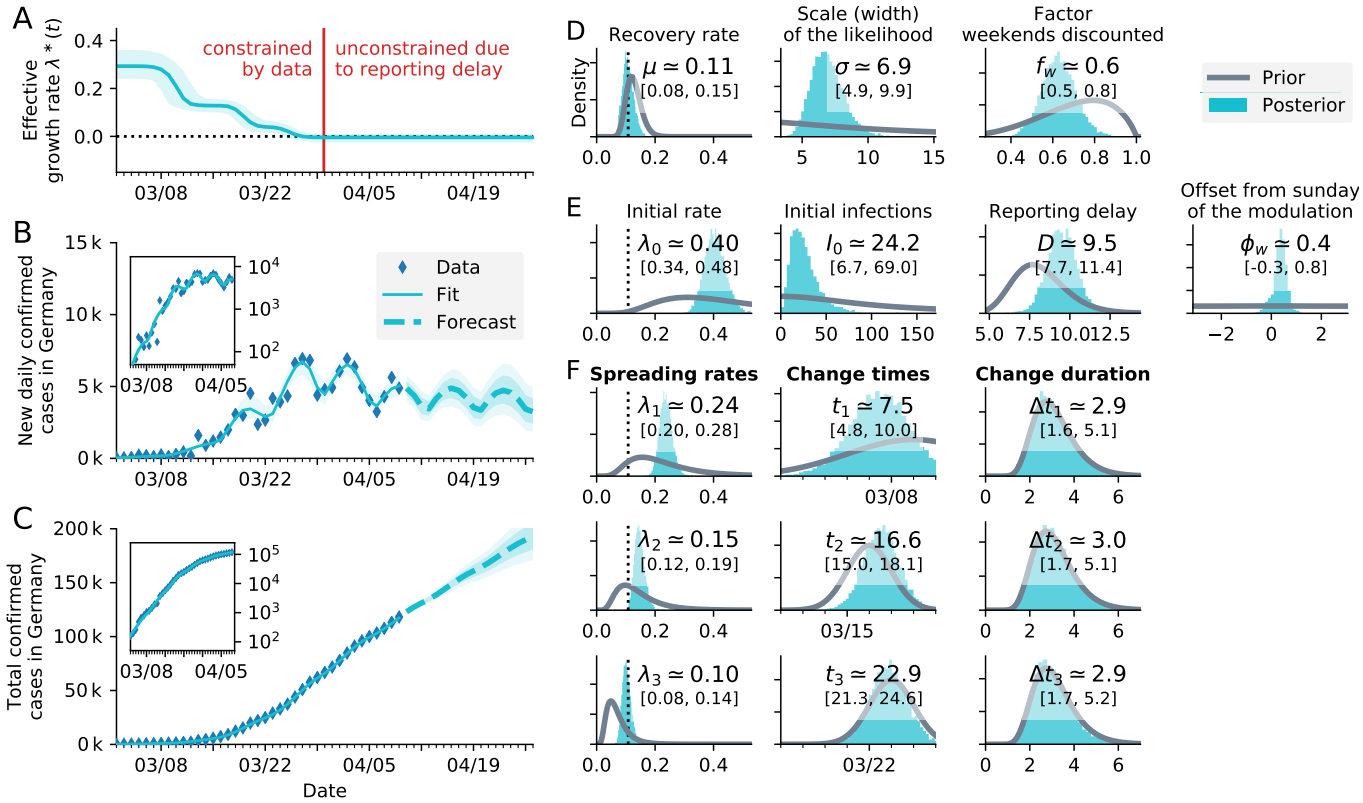

FIG. S5. Inference with the SIR model, as in the main text, but **with a weekly modulation of the number of observed cases**, to take into account the fewer number of tests performed on weekends. This was modeled by a multiplication of the inferred cases by the absolute value of a sine function (one week period). The offset of the sine function was given a flat prior. The magnitude of the sine function was given a beta-distribution as prior, with mean 0.7 and standard deviation 0.17. **A:** Time-dependent model estimate of the effective growth rate $\lambda^*(t)$. **B:** Comparison of daily recorded new cases and the model. Comparison of total recorded cases and the model. **C:** Model forecast of total cases, based on the inferred $\lambda^*(t)$. **D–F**: Prior and posterior distributions of all free parameters.