CESifo WORKING PAPERS

8446
2020

July 2020

Right and Yet Wrong: A Spatio-Temporal Evaluation of Germany's COVID-19 Containment Policy

Michael Berlemann, Erik Haustein

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8446*

Right and Yet Wrong: A Spatio-Temporal Evaluation of Germany's COVID-19 Containment Policy

Abstract

In order to get the COVID-19 pandemic under control, most governments around the globe have adopted some sort of containment policies. In the light of the enormous costs of these policies, in many countries highly controversial discussions on the adequacy of the chosen policies evolved. We contribute to this discussion by evaluating three waves of containment measures adopted by the German government. Based on a spatio-temporal endemic-epidemic model we show that in retrospective, only the first wave of containment measures clearly contributed to flattening the curve of new infections. However, a real-time analysis using the same empirical model reveals that based on the then available information, the adoption of additional containment measures was warranted. Moreover our spatio-temporal analysis shows that a one-size-fits-all policy, as it was adopted in Germany on the early stages of the epidemic, is not optimal.

JEL-Codes: I120, I180.

Keywords: containment measures, policy uncertainty, Covid-19, SIR model, infections, spatio-temporal modeling.

*Michael Berlemann**
Faculty of Economics & Social Sciences
Helmut-Schmidt-University
Holstenhofweg 85
Germany – 22043 Hamburg
Michael.Berlemann@hsu-hh.de

Erik Haustein
Faculty of Economics & Social Sciences
Helmut-Schmidt-University
Holstenhofweg 85
Germany – 22043 Hamburg
Erik.Haustein@hsu-hh.de

*corresponding author

1. Introduction

When SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2 - SARS-CoV-2), causing the respiratory disease COVID-19 (Coronavirus Disease 2019), spread first quickly within China (especially in the Hubei region) and then developed into a pandemic early 2020, the vast majority of affected countries adopted measures against the further (local) spread of virus. While a few countries at least temporarily considered a herd immunity strategy with only very mild containment measures, the majority of countries adopted strategies aiming at "flattening the curve". This strategy aims at slowing the spread of the epidemic so that the peak number of people requiring care at a time is reduced and the health care system does not exceed its capacity. However, governments differed substantially in the measures they adopted and how quickly they adopted them (Petherick et al. (2020)). The chosen containment measures range from public information campaigns, international travel restrictions, closings of educational institutions, workplaces, public transport and leisure and retail facilities, the cancellation of public and private events, restrictions on internal movement and obligations

to wear face masks to stay-at-home requirements.

Especially in those countries which already passed the (preliminary) peak of new infections and in which the health systems did not reach their capacity limits, recently a debate on the adequacy of the chosen containment measures unfolded. As the (expected) costs of the adopted strategies are often enormous, this is not too surprising. The critics of containment policies typically argue that at least some of the chosen policies were unnecessary as finally the pandemic proved to be much less severe than the overly pessimistic prophecies have initially told. The supporters of restrictive containment policies argue that this point of view is the result of a self-defeating prophecy, as the final reason for the success of containment policies is the existence of "prophecies of doom".

Germany is a prominent example for these lively discussions. When Germany experienced strongly rising infection numbers in early March 2020, quickly spreading all over Germany, the Federal Government initiated three waves of containment measures.[1] Soon after new infections reached their peak and started decaying, a discussion on the adequacy of the German containment policy unfolded in both, the public and among scientists from various disciplines. Since early April, numerous theoretical and empirical studies of the adequacy of the German containment measures evolved. Interestingly enough, they come to heavily differing results.[2]

This paper contributes to the literature by delivering new empirical evidence for the adequacy of various waves of containment measures adopted in Germany. Our empirical analysis is based on forecasts derived from an endemic-epidemic model which has proved to perform well in describing and forecasting other infectious diseases. As COVID-19 has a strong epidemic component, we employ a spatio-temporal model variant and conduct our analysis on the county-level. Doing so allows us constructing meaningful and consistent forecasts on various levels of spatial aggregation and to exploit all information in the available raw data (Giuliani et al. (2020)).

Our major finding is that the final judgment of the necessity of the adopted containment measures depends strongly on the available information set. When basing our analysis on all data which was available when this paper was written in the mid of June 2020, we find that only the first

[1] We describe these measures in detail in Section 4.2.

[2] We review this literature briefly in Section 2 of this paper.

wave of containment measures clearly contributed to "flattening the curve". However, when we take a real-time perspective and use only data which was available when the decisions on the adoption of the containment measures were made, all three waves of containment measures appear to be justified, at least in principle. Our results also indicate that regionally differing containment policies are strongly superior in comparison to one-size-fits-all policies.

The paper is structured as follows. Section 2 discusses the related literature. Section 3 describes the employed data. Section 4 outlines the empirical strategy, explains the considered containment measures, introduces the employed empirical model and presents the main empirical results. Section 5 delivers the results for the real-time perspective. Section 6 discusses the adequacy of the containment measures on the disaggregated spatial level. Section 6 concludes.

2. Related Literature

Although the COVID-19 pandemic is still evolving in many parts of the world, there is already quite some literature which is concerned with evaluating different containment policies. In the following we briefly discuss this literature, however, with a focus on Germany. Broadly, the literature can be divided into two strands: papers basing on calibrated theoretical models and econometric approaches.

Especially the model-based literature often bases upon the SIR model (Britton (2010)), which goes back to early work by Kermack et al. (1927). It assumes the population can be subdivided into at least three compartments: susceptible (S), infectious (I) and recovered individuals (I). Whenever central parameters such as the likelihood of susceptible individuals to become infectious, the time infectious individuals remain infectious and the time until recovery are known, the SIR model can be formulated as a system of differential equations and, after calibration, can be used for forecasting or simulation purposes.

The first strand of the literature employs calibrated (variants of) SIR models to study the effect of containment measures on the development of the COVID-19 epidemic.[3] To the best of our knowledge, the first study for Germany was delivered by an der Heiden and Buchholz (2020) and bases on

[3]See e.g. Maier and Brockmann (2020) for China or the multi-country-studies by Flaxman et al. (2020) and Gros et al. (2020).

an SEIR model, which extends the standard SIR model for a latent state of being exposed before becoming infectious. The model is mostly calibrated with data from China and concludes that without containment measures Germany would have reached quickly a critical level of infections exceeding the health system's capacity. The authors argue that under most scenarios a combination of various containment measures is necessary to prevent the health system from collapsing. A subsequent study by Donsimoni et al. (2020) relies on a very similar model, but calibrates it with more recent data from Germany. The authors show that public interventions can lead to more or less severe outcomes of the epidemic, depending on their timing and the employed outcome measures. Even the most recent calibration study for Germany by Dehning et al. (2020) relies on an SIR model. Here, the authors use Bayesian inference on Markov-Chain Monte-Carlo sampling to calibrate their model. The authors find that a model variant including three change points on March 6, March 15 and March 23 explain the data best and that even the third wave of containment measures was necessary to leave the path of exponential growth of new infections.

The second strand of the literature uses econometric methods to study whether and how containment measures affected newly occurring infections, either on the country or the regional level. To the best of our knowledge, the earliest econometric study for Germany on the country-level was conducted by Hartl et al. (2020). The study is based on data collected by Johns-Hopkins-University and analyzes the effect of the policy package adopted on March 13, which included the decision to close educational institutions. Employing a simple linear trend model for the logarithm of confirmed cases the authors find a structural break on March 20. Assuming a time-lag of 7-8 days, they attribute the structural break to the measures adopted on March 13. Homburg (2020) follows a similar approach, based on the same (but more recent) data. He basically argues that the "lockdown" on March 23 and even the closure of educational institutions 7 days earlier was unnecessary as the the peak of new infections was already reached on March 29. Assuming a time-lag of the data of 17 days the new infections would already have started to decrease well before these measures were adopted by the German government.

Other econometric studies have exploited regional data to examine the

effect of containment measures.[4] All regional studies for Germany base on data on new infections on the county level, published by the RKI. Mense and Michelsen (2020) cumulate the data on the week level and study the overall effectiveness of the German containment measures adopted in the 12th and 13th week 2020. In order to do so they regress new infections on past infections and a spatial lag of past infections within a two-way fixed effects panel setting. They find systematically lower coefficients for the spatial effect after the implementation of the containment measures and argue that, as a package, the adopted measures were effective. Glogowsky et al. (2020) employ an event-study framework for their empirical analysis. The authors find that the implemented containment measures reduced mobility and also significantly decreased new infections. Wieland (2020) employs the RKI data in daily frequency. However, before using them in his empirical approach he infers missing data points on the reference date from the available observations via auxiliary regressions and assumes an incubation time of 5 days. Based on the corrected data, he estimates logistic growth models to determine the local infection points, e.g. the days when the growth rate of new infections started decreasing. For Germany as a whole he estimates the infection point to lie in between March 17 and March 20 and thus well before the third round of containment policies became effective. For as many as 255 out of 412 county observations, the infection point is estimated before March 23. Felbermayr et al. (2020) primarily aim at identifying the main "superspreader-event" which led to the subsequent spread of COVID-19 within Germany. In order to do so they conduct a (repeated) cross-section analysis of new infection counts using a negative binomial model and find a significant effect of the road distance to Ischgl, a skiing area in Austria which was heavily visited by German tourists. The authors interpret their finding that the distance-to-Ischgl-coefficient turned out to be significant even after the containment policies in Germany were adopted as indication that the containment policy was successful in limiting infections over county borders.

[4]See e.g. the studies for China by Kraemer et al. (2020), for Spain by Orea and Álvarez (2020) and for the United States by Abouk and Heydari (2020), Chernozhukov et al. (2020) and Courtemanche et al. (2020). A multi-country study based on regional data was conducted by Hsiang et al. (2020).

3. Data

The data we employ for our empirical analysis comes from the Robert Koch Institute (RKI). RKI is the German government's central scientific institution in the field of biomedicine. A major task of RKI is monitoring infectious diseases such as COVID-19. To fulfill this task, RKI collects data on all detected COVID-19 cases in Germany.

According to the German Infection Protection Act (Infektionsschutzgesetz, IfSG), physicians and laboratories detecting active COVID-19 cases have to report these cases within 24 hours to the local public health department (Gesundheitsamt).[5] COVID-19 cases meeting the definition of the RKI are transmitted electronically by the local health department to the state government which then forwards this information to the RKI at the latest on the next working day. Most of the involved health authorities transmit the data earlier and more frequently than required by law, usually daily and also at weekends. Nevertheless, there is typically a delay of several days in the transmission of cases. The data transmitted to RKI always contains information on gender and age (age groups) of the infected individuals, the place of living (only county information) and the day, when the local public health department acquired knowledge on the case ("reporting date"). In roughly two-thirds of the cases the data also comes with information on the day, when the first symptoms occurred ("reference date"). Whenever the reference date is unknown, the reference date is set to the reporting date.

For our subsequent empirical analysis we need data on the date of infection. As this date is not included in the RKI data, we construct this information in a two-step procedure. In the first step we estimate the reference date for those observations in the RKI data, for which only the reporting date is available. In order to do so we employ those observations, for which both reporting and reference date are available. We then regress the difference between reference date and reporting date on age, gender and week-day and, in addition, use commune-fixed effects. We then use the estimated coefficients to impute the missing reference dates in the dataset. In the second step we calculate the most likely day of infection by assuming an incubation

[5] Note that COVID-19 often occurs without any or with only mild symptoms (see e.g. Streeck et al. (2020)). Thus, the factual number of infections is likely larger than the one reported in the RKI data. However, as we explain later, underreported data is not a problem in our empirical approach.

period of 5.8 days (see the meta-study by McAloon et al. (2020)).

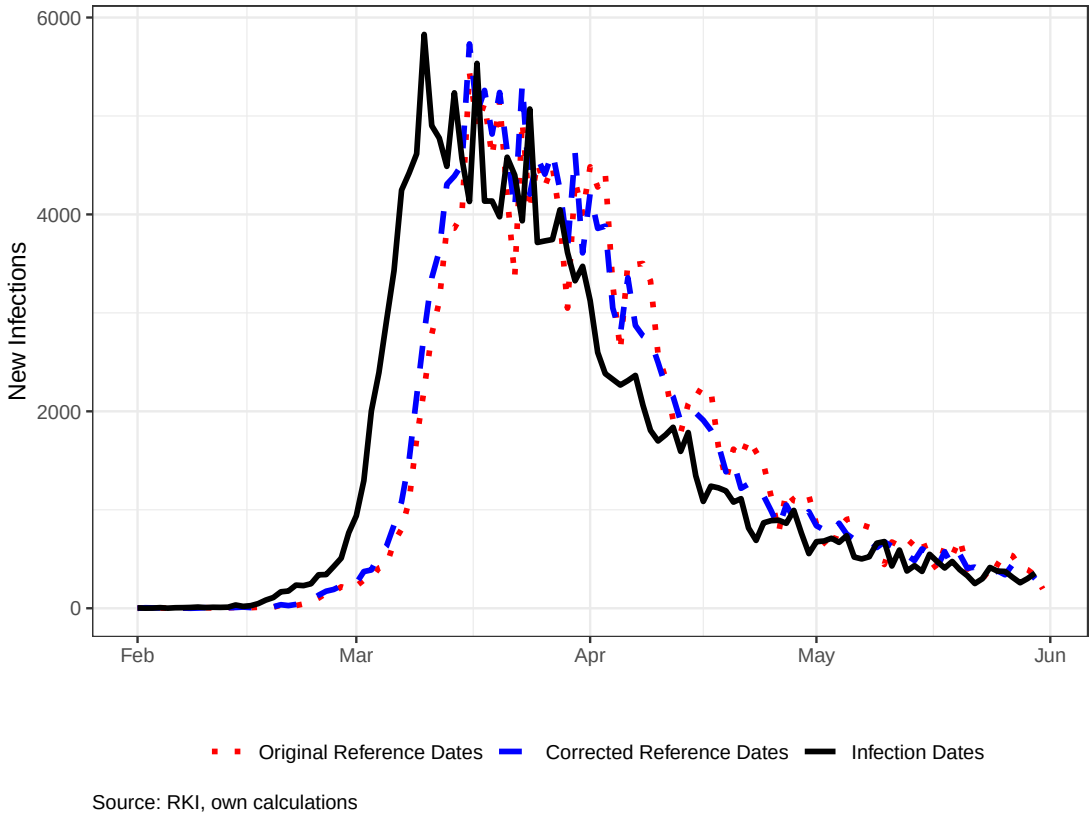

Figure 1: New Infections Based on Original Reference Dates, Corrected Reference Dates and Inferred Infection Dates

In Figure 1 we show the development of new infections when referring to (i) the reference dates reported in the original RKI data, (ii) the corrected reference dates and (iii) the likely infection date. By construction, both types of corrections result in more cases earlier on the time axis. Obviously, the effect of the correction of missing reference dates is comparatively small until the mid of March 2020 and increases slightly in size over the rest of the sample period. In general, the correction for the incubation time has a much more pronounced effect on the resulting data of new infections.

4. Were the German Containment Measures Necessary?

4.1. Empirical Strategy

Our aim is to evaluate three waves of containment measures, initiated by the German Federal Government in March 2020. In order to judge whether these measures were necessary to reach the central goal of preventing a collapse of the health system we proceed in two steps: In a first step we define the three (groups of) containment measures we investigate in this paper; we thereby also have to define when exactly the measures became effective. This step is important, as Germany is a federal state where many of the containment measures become effective not before the referring local governments implemented them formally.[6] In the second step of our analysis we employ the data presented in Section 3 to estimate a spatio-temporal model using only data before a certain measure was adopted. We then use the estimated model to predict the likely development of new infections for the subsequent period. The likely effect of the containment measure is then the difference between factual new infections and the predicted values.

4.2. Definition of Containment Measures

The first containment measure, we study in the following, is a ban on mass events. The ban was announced by German Minister of Health Jens Spahn who recommended to cancel all events with more than 1000 participants. While a number of events was already cancelled earlier (such as e.g. the international tourist fare ITB in Berlin on February 28 or the Leipzig book fare on March 4), the official recommendation was made on the afternoon of March 8. We assume that the ban on mass events became factually effective two days later, on March 10, when already numerous German states formally adopted the recommendation.

In the evening of March 16, a second round of containment measures was announced by Chancellor Angela Merkel. These containment measures included the closure of educational institutions (nursery schools, schools and universities), leisure facilities (e.g. gyms, playgrounds, bars and clubs) and retail facilities (with the exception of pharmacies, drugstores and groceries) as well as the introduction of national and international traveling restrictions.

[6] As the three groups of measures we study in the following were discussed and approved in meetings of the federal government and the German state's prime ministers, it is justified to assume that the measures became effective at roughly the same point in time.

We assume that these measures became effective two days later, i.e. on March 18.

The third group of containment measures, we investigate in this paper, was announced late on March 22, again by Chancellor Merkel. The population was asked to minimize social contacts as much as possible ("social distancing"). Firms were advised to allow home-office wherever possible and to guarantee a minimization of social contact at the working place. While it was allowed to leave home for work, visiting the doctor, buying food or having a walk, a physical distance to non-family members of at least 1.5 meters had to be kept. Finally, even restaurants and hairdressers had to close. Again we assume that these measures became effective two days later, on March 24.

4.3. Prediction Model

The model we use to predict the onset of the COVID-19 epidemic in Germany is based on the earlier mentioned SIR model and focuses on describing the transmission from the state of susceptible individuals to infectious individuals, as reported in the earlier described RKI data. Our empirical implementation follows the basic idea of Held et al. (2005) to model our panel of areal count time series as Poisson branching process with immigration. In line with Meyer et al. (2017) we assume that the regional count of newly infected individuals $Y_{r,t}$ is determined by an endemic and two epidemic components.[7] More precisely, we assume that this process follows a negative binomial distribution (with overdispersion parameter $\psi > 0$) and has the conditional mean

$$\mu_{r,t}^{Y} = e_r \cdot \nu_t + \lambda_r \cdot \sum_{d=1}^{D} u_d \cdot Y_{r,t-d} + \phi_r \cdot \sum_{s \neq r} \sum_{d=1}^{D} w_{r,s} \cdot u_d \cdot Y_{s,t-d}. \tag{1}$$

The endemic component, i.e. the share of the population in region r at time t which is newly infected, regardless of a county's infection history and regardless of the infection histories of its neighbours, is modeled as

$$\ln(\nu_t) = \alpha_0 + \eta \cdot t + \gamma \cdot \sin(\omega \cdot t) + \delta \cdot cos(\omega \cdot t) \tag{2}$$

[7] This approach has recently also been used to model the spatio-temporal spread of the COVID-19 in Italian provinces, see Giuliani et al. (2020).

with α_0 being a constant, $\eta \cdot t$ being a time trend[8] and $\gamma \cdot \sin(\omega \cdot t) + \delta \cdot cos(\omega \cdot t)$ capturing possible seasonal variation of the endemic component as it is typical for many virus diseases. In order to receive the mean of the endemic component in region r, we further have to multiply ν_t by the size of the local population e_r.[9]

The epidemic components of the infection process consist of an autoregressive and a spatial part. The autoregressive epidemic component $\lambda_r \cdot \sum_{d=1}^{D} u_d \cdot Y_{r,t-d}$ accounts for the reproduction of COVID-19 within the same region. In line with Bracher and Held (2020a) we allow for more than one autoregressive lag (with the weighting factors u_d fulfilling $\sum_{d=1}^{D} u_d = 1$) to better capture the time-series properties of new infections. This component is modeled as

$$\ln(\lambda_r) = b_r, \tag{3}$$

with b_r being a region-specific random effect (with $b_r \sim N(0, \sigma_\lambda^2)$) that accounts for random differences between regions.

The spatial autoregressive component $\phi_r \cdot \sum_{s \neq r} \sum_{d=1}^{D} w_{r,s} \cdot u_d \cdot Y_{s,t-d}$ accounts for the transmission of COVID-19 between regions. The spatial weights $w_{r,s}$ describe the flow of infections from region s to region r.[10] As for the autoregressive part, we allow for more than one spatial lag in our estimation approach. Similar as the autoregressive component we model the spatial component as

$$\ln(\phi_r) = c_r, \tag{4}$$

with c_r being a region-specific random effect (with $c_r \sim N(0, \sigma_\phi^2)$).

As our subsequent empirical analysis is partly based on relatively short panel data, we opt for two lags ($D = 2$) in the epidemic components of our model. We use exponentially decaying weights $u_d = \gamma \cdot (1 - \gamma)^{d-1}$ with $\gamma = 0.6$.[11]

[8] The time trend also corrects for potential changes in the testing intensity.

[9] Note that in principle, e can vary over time; however, as local population counts are not available in high frequency on the county level, we employ the newest population counts, which relate to the end of 2019 and which are available in the RKI data.

[10] The weights were derived from a row-normalized contiguity matrix of order one and type queen.

[11] In order to find the optimal value for γ we evaluated each estimated model for alternative values of γ ($\gamma \in \{0.6, 0.7, 0.8, 0.9\}$) using one-step-ahead in-sample forecasts and compared the models based on the logarithmic score, the ranked probability score and the Dawid-Sebastiani score (see e.g. Gneiting and Katzfuss (2014)). As the result of this

The described model can be estimated using penalized maximum likelihood procedures, as described in Paul and Held (2011) and Meyer and Held (2014). These techniques are implemented in the R package "surveillance"[12]. As we allow for more than one autoregressive term in our specification, we in addition use the R package "HHH4addon"[13] for the subsequent empirical analysis. Note that explicitly accounting for underreporting in the RKI data has little benefit in our application as we use the model primarily for forecasting and not for identifing parameters (see Bracher and Held (2020b)).

4.4. Empirical Results

We start out with a discussion of the results for the first wave of containment measures, announced on the afternoon of March 8, which are shown in the upper part of Figure 2. The black line illustrates new infection counts on the referring day. The forecast model is fitted over the period of February 1 to March 9. In the left part of the plot we show the fitted values of the model, disaggregated in the endemic, the autoregressive epidemic and the spatial epidemic part. We then use the model to predict the values of infections over the seven subsequent days[14] and show the resulting forecast interval.[15] It is easy to see that the model predicts strongly increasing new infections per day which almost double from around 4,600 to 8,600 new infections over the forecast horizon of one week. Over the same period, factual new infections slightly decreased and thus remained well below the projection. We take this as an indication that the first wave of containment measures contributed significantly to flattening the curve of new infections.

The results for the second wave of containment measures, announced on the evening of March 16, are shown in the middle part of Figure 2. The epidemic model was estimated for the period of February 1 to March 17. Again we use the estimated model to generate one-week-ahead forecasts[16] While

procedure $\gamma = 0.6$ was chosen for all subsequent empirical specifications.

[12]Meyer et al. (2017).

[13]This package is available at https://github.com/jbracher/hhh4addon. See Bracher and Held (2020a) for an application to dengue fever in Puerto Rico and viral gastroenteritis in Berlin.

[14]We refrain from further extending the forecast horizon as the next wave of containment measures was announced as early as on March 16.

[15]The forecast intervals were constructed via 10.000 Monte Carlo simulations.

[16]Again this is due to the fact that the third wave of containment measures was already announced a week later, on the evening of March 22.

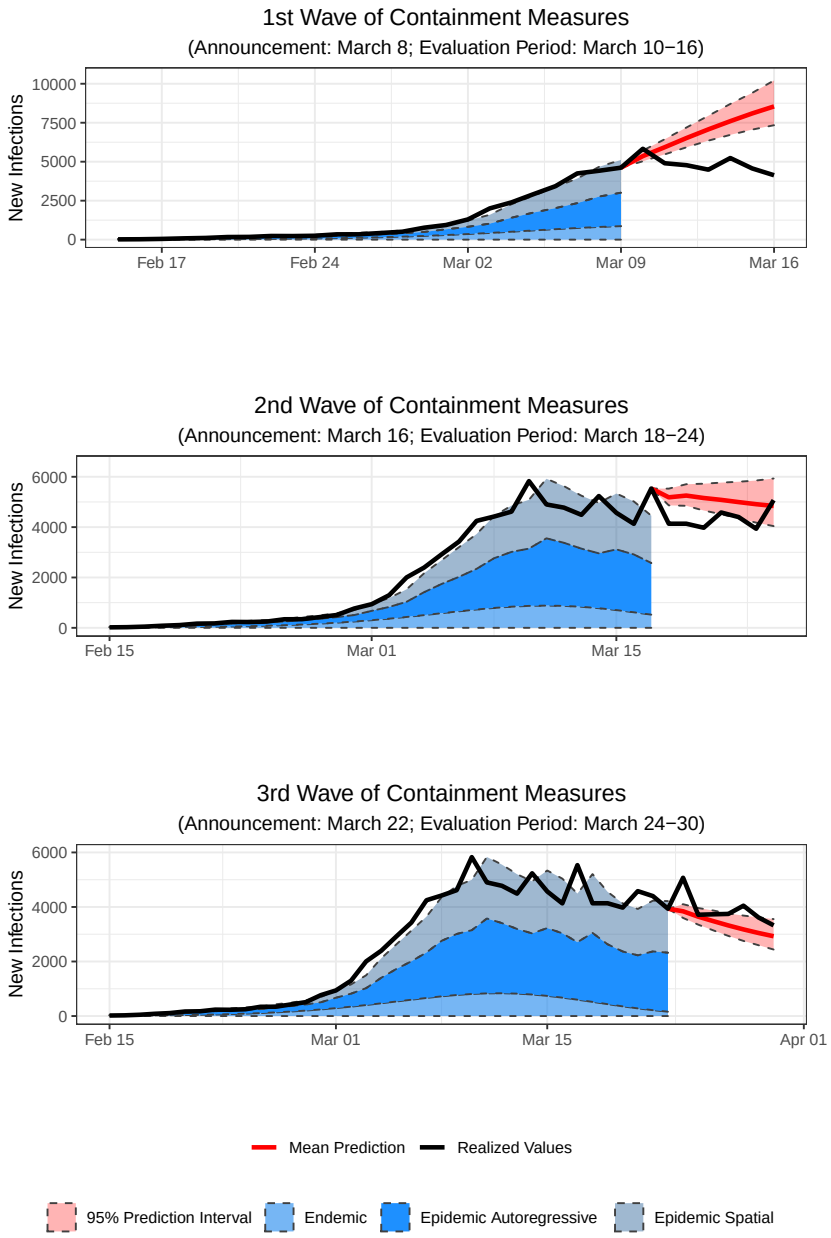

Figure 2: Evaluation of Alternative German Containment Measures

the factual new infections are mostly below the mean values predicted by the epidemic model, the difference between the mean prediction and the realized new infections differ significantly only in the first few days. Thereafter, the 95-percent prediction interval includes the realized values. Moreover, the mean prediction has already a downward slope, indicating that even without the second wave of containment measures the number of new infections could be expected to decrease. While this does not imply that the second wave of containment measures was without effect, one might at least question whether they in fact were necessary to enforce decreasing new infections and to prevent the health system from collapsing. However, one should also take into account that the 95-percent prediction interval widens quickly and is partly consistent with rising new infections, so that we have to interpret this finding with some caution.

In the lower part of Figure 2 we show the results for the third wave of containment measures, which were announced on the evening of March 22. The epidemic forecast model was fitted over the period of February 1 to March 23. For the third wave of containment measures we do not observe a systematic difference between the predicted and the realized new infections. As the 95-percent prediction interval is strongly downward sloping, the empirical evidence points into the direction that the third wave of containment measures was not necessary to prevent a collapse of the health system.

Note that the reported empirical results do not imply that the containment measures adopted in the second and the third wave such as the closure of educational institutions or enforced social distancing in general have no effect on new infections at all. First, note that we focused here on the question, whether the adopted measures were necessary to prevent a collapse of the German health system. Even when a measure is thus classified as unnecessary to contribute to this goal, it nevertheless might have contributed to lowering the subsequent number of new infections. Second, in our empirical setting the estimated effect of the adopted measures is contingent on which measures were adopted before. Thus, a different sequencing of the measures could result in a different evaluation of the different containment measures. Third, it should be considered that the German government and the Robert-Koch-Institute conducted a continuous media campaign, explaining possible consequences of COVID-19 and delivering information on how individuals could contribute to avoiding a further spread of the SARS-CoV-2 virus. Part of this information campaign were the recommendations to keep physical distance to others, to wash hands regularly and not to shake hands or to cough

openly. These recommendations coincided to quite some extent with the social distancing measures, formally adopted on March 23. It is well possible (if not likely) that these recommendations were already sufficient to induce behavioral changes in the population and to reach the goal of preventing a collapse of the German health system.

5. The Real-Time Perspective of the Acting Politicians

We conducted our previous analysis of the adequacy of the chosen containment policies in Germany on all data which was available in the mid of June. Doing so allowed us correcting the original data published by RKI to account for cases with missing reference date and the incubation period of 5.8 days. We also employed all data before the adoption of a certain containment measure to predict the likely future development of new infections. While doing so is adequate to judge the necessity of the containment measures in retrospect, parts of this information was unavailable to the acting politicians when they had to decide on the implementation of containment policies. Thus, while certain containment policies might be judged as unnecessary in retrospective, they might have looked reasonable at the time when they were adopted. In order to study this issue, we repeat our analysis under quasi real-time conditions, i.e. under the premise that only the data published on the day of announcement of a containment policy was available.

In the following we illustrate our procedure at the example of the first wave of containment measures, announced on March 8. We applied the same procedure to the later two waves of containment measures.

In the first step of our real-time analysis, we again infer the missing reference dates from those cases, where the difference between reporting and reference date is known via a regression model (as outlined in Section 3). However, we now use only data which was available on March 8, i.e. data with reporting dates until March 7. Moreover, we have to take into account that the most recent data is highly incomplete due to the fact that in most cases there is a delay between the reporting and the reference day. When we would use all observations until March 7 for the correction, the correction would be strongly biased towards too short corrections. In order to reduce this bias, we use only observations with reference dates before March 1 (e.g. 7 days earlier), as we then can expect to have at least 75 percent of all observations in our sample. After running the auxiliary regressions we use their results to infer the missing reference dates. We then correct for the

mean incubation period of 5.8 days to end up with a corrected panel of infection counts.

In the second step we have to determine, which is the last reliable observation of new infection counts in our sample. As the mean incubation time amounts to 5.8 days and it takes another 7 days until at least 75 percent of all reference dates became part of the data, the last reliable observation of our newly constructed new infections variable is February 23. Thus, when estimating the parameters of our spatio-temporal model we use only infection data until February 23.

In the third and final step of our analysis we use the estimated model parameters to nowcast new infections until March 8. In the same manner we then construct projections for new infections over the subsequent week. We argue that this projection is describing what a well-informed politician should have expected for the near future.

Figure 3 shows the results we receive when applying this procedure to all three waves of containment measures.

In the upper part of Figure 3 we show the situation on March 8, when the German Minister of Health Jens Spahn announced the ban on mass events. The now- and forecasted subperiods are separated by dashed vertical lines. The model nowcasts slightly less than 500 new infections per day for March 8. According to the model's prediction, the new infection count is expected to double over the two subsequent weeks to almost 1000. The quickly widening and highly asymmetric 95-percent prediction interval[17] indicates a high degree of uncertainty on future infection counts and also includes exponential infection growth paths. Based on this projection, one might hardly classify the adoption of the first wave of containment measures as unnecessary.

In the middle part of Figure 3 we display the forecast derived from data available on March 15, when the second wave of containment measures was announced. Here, the model predicts exponential growth of new infections reaching values of more than 80.000 daily new infections on March 24. All infection paths consistent with the 95-percent prediction interval must be judged as a severe threat to the German health system. Thus, there is little doubt that based on the available information additional containment measures appeared as necessary to flatten the curve of new infections.

[17]Note that we skipped parts of the upper part of the prediction interval due to visibility of the mean prediciton.

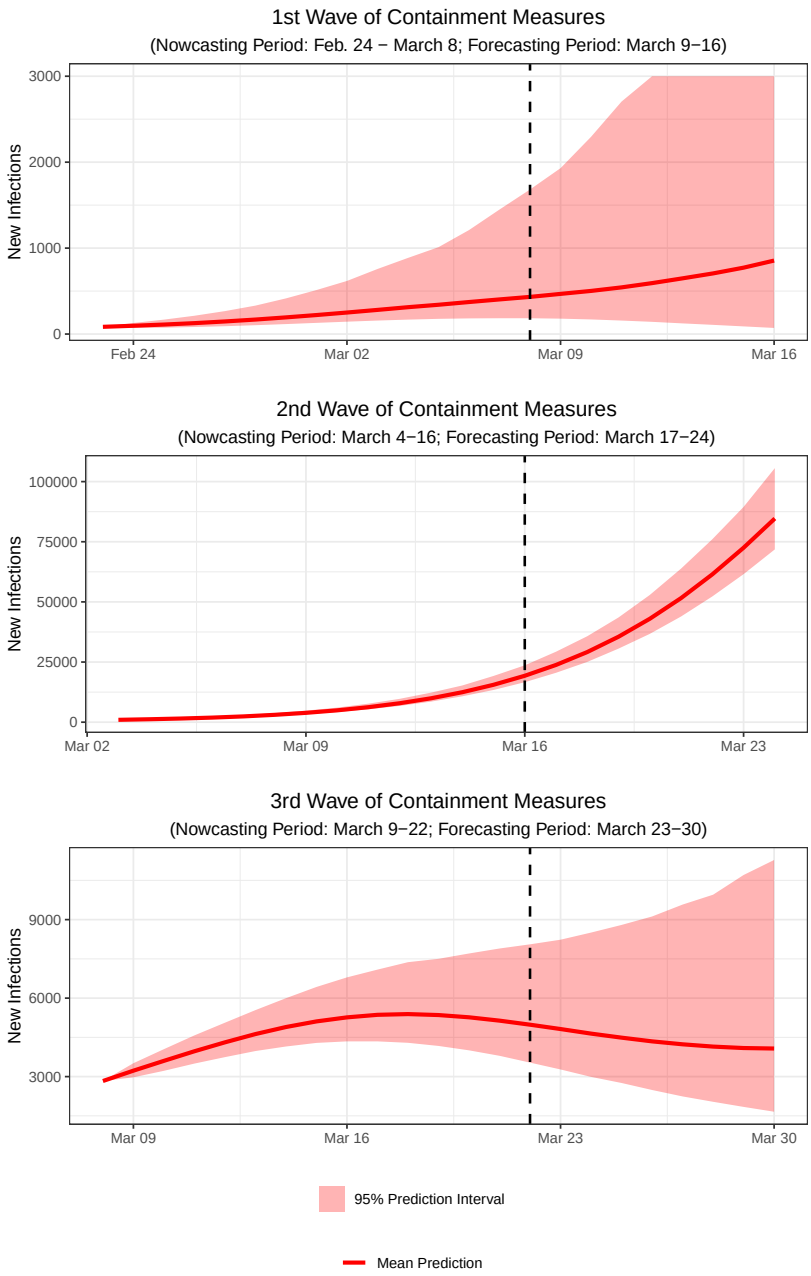

Figure 3: Real-time Forecasts without Further Containment Measures

Finally, in the lower part of Figure 3, we show the model prediction for the third wave of containment measures, announced on March 22. Here, the situation is somewhat ambiguous. According to the mean nowcast for March 22, new infections are still on a comparatively high level, but already started to slightly decrease. A further slight decrease is expected over the period until March 30. One might argue that based on the information available on March 22, additional containment measures were unnecessary to reach the goal of preventing a collapse of the German health system. However, according to the mean model prediction the number of new infections remains on a comparatively high level. Moreover, the 95-percent prediction interval for the mean forecast turns out to be large and highly asymmetric towards higher new infections. Thus, there was still a significant probability of further rising new infections over the subsequent week(s). Thus, even when the mean forecast pointed already into the direction of a slight relaxation of the situation, a comparatively mild degree of risk aversion would render the decision to adopt further measures correct. One might therefore conclude that the decision to adopt further containment measures on March 22 was warranted, given the then available information.

6. Did One Size Fit All?

In the early phase of the COVID-19 epidemic, the German containment policies followed a one-size-fits-all strategy. All adopted measures were discussed on regular joint meetings of the federal and the state governments. On these meetings the involved politicians, after intense and sometimes controversial discussions, agreed on the measures to be adopted. After the announcement of the measures by the heads of the German federal coalition government, the state governments independently implemented these measures. Although there was a mild variety in the exact timing and even the implementation of the measures, at least throughout March 2020 it was the declared will of the acting politicians to realize a joint and highly coordinated containment policy.[18]

While one might argue that the implementation of the same measures in all parts of Germany might have at least initially contributed to a higher

[18]Note that this does not hold true for later implemented containment measures such as the obligation to wear face masks. Even the (much later) initiated relaxation of some of the containment measures differed enormously in both the time and the spatial dimension.

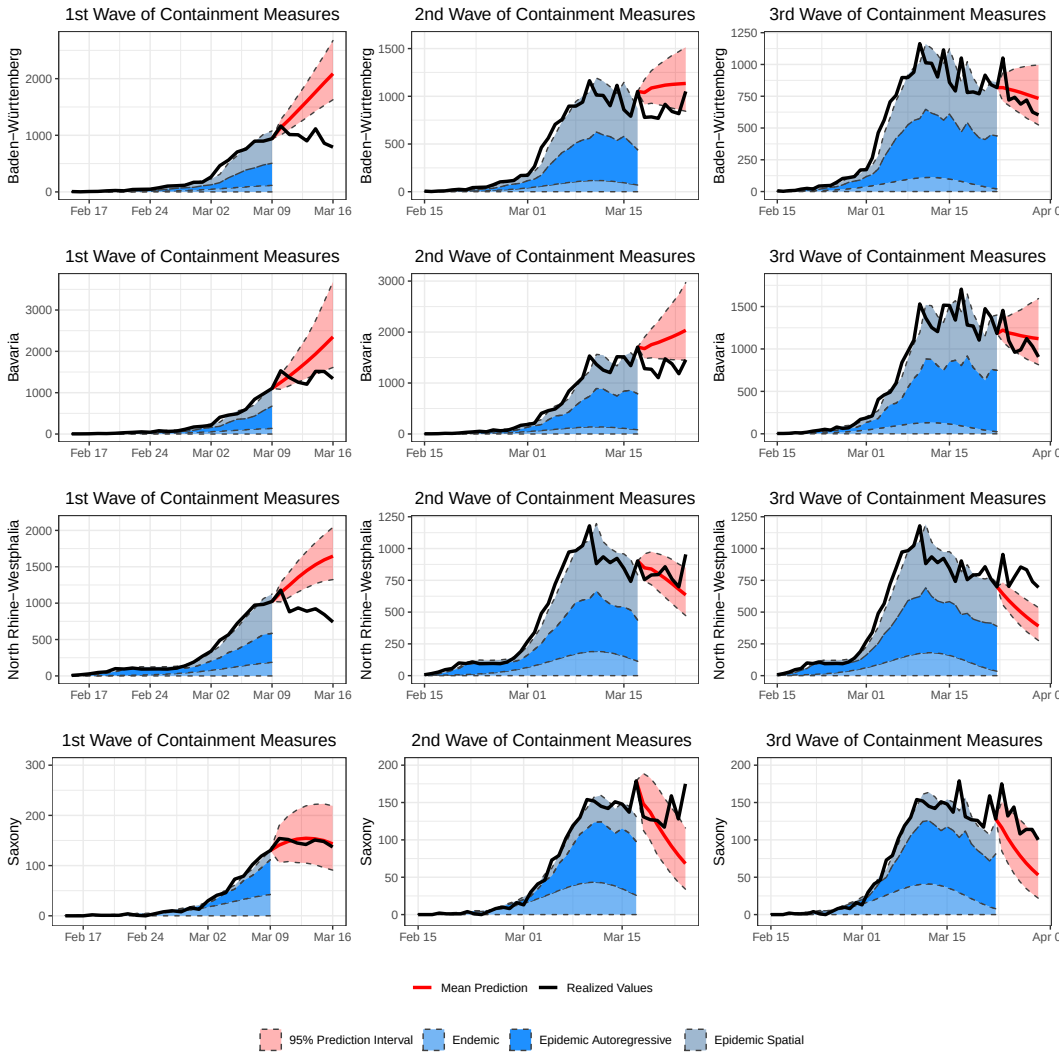

Figure 4: Adequacy of Alternative Containment Measures for Selected German States

degree of acceptance of the containment policy in the population, one might question whether such a one-size-fits-all policy was the best choice for a country like Germany with many quite diverse regions and organized as a federal state. In order to shed some light on this question, we use the spatial dimension of the empirical model, we estimated in Section 4. As the model delivers results on the county-level, we can aggregate the results even on lower

administrative levels such as the state level. In Figure 4 we show the results for four different states. We opted for the three most affected states Baden-Württemberg (first row), Bavaria (second row) and North Rhine-Westphalia (third row) as well as the most populated East German state Saxony (fourth row). In the columns we show the results for the three waves of containment measures.

While the first wave of containment measures turns out to be adequate for Baden-Württemberg, Bavaria and North Rhine-Westphalia, this hardly holds true for Saxony, where the mean forecast almost perfectly coincides with the factual development of new infections. Especially for the states with initially low infection counts in East and North Germany[19] already the first round of containment measures turns out to be somewhat questionable.

Similarly, even for the second wave of containment measures we find remarkable differences between the four states. While additional measures seemed to be adequate for Baden-Württemberg and Bavaria, for North Rhine-Westphalia the mean forecast was already indicating strongly decreasing new infections. The same applies to Saxony.

For the third wave of containment measures the mean predictions turn out to be downward sloping for all four states, thereby questioning the necessity of an additional round of containment measures. However, for Baden-Württemberg and Bavaria the 95-percent prediction interval includes also paths implying increasing new infections, rendering the decision to have an additional round of containment measures more rational than for North Rhine-Westphalia and Saxony. Interestingly enough, it was in fact Bavaria's prime minister Markus Söder who insisted on additional containment measures whereas Armin Laschet, prime minister of North Rhine-Westphalia early advocated for milder containment policies and relaxations.

The disaggregated data also reveal that there is quite some variety in the relative importance of the endemic, the autoregressive epidemic and the spatial endemic components. As Figure 4 reveals, the spatial endemic component played only a minor role in the case of Saxony whereas spatial spillovers contributed much to the development of new infections in North

[19] An exception is Hamburg which had high infection counts in the early phase of the epidemic. This is mostly due to the fact that Hamburg is the only German state with skiing holidays in the first two weeks of March. At that time many tourists got infected in the skiing areas like Ischgl in Austria (see Felbermayr et al. (2020)) and then returned to their home regions.

Rhine-Westphalia. These regional differences might be taken as an indication that different sorts of containment measures are adequate in these regions.

7. Summary and Conclusions

When the COVID-19 epidemic reached Germany in the first quarter of 2020, the German government adopted various waves of country-wide containment measures. Employing a spatio-temporal endemic-epidemic model, which is estimated for reference-date- and incubation-time-corrected RKI data, we showed that the second and especially the third wave of containment measures was likely not necessary to prevent a collapse of the German health system. However, based on a quasi real-time analysis we also show that, based on the available information, the decisions to adopt additional measures can hardly be judged as wrong or even irrational. However, the depicted discrepancy between the ex-post and the ex-ante perspective indicates that the payoff of better and earlier available data on unfolding epidemics might be large, especially in the light of the enormous costs of many containment measures. Investments in the collection of reliable raw data in medical practices and laboratories and a quicker transmission of this data to the relevant policymakers and researchers might help to reduce the follow-up costs of unnecessary containment measures.

Our study also questions one-size-fits-all containment policies, as they were initially adopted in Germany and many other countries. While initially a common containment policy might be helpful in organizing the necessary public support as all citizens are exposed to the same measures, this comes at the price that the adopted measures might be too strict for less affected areas. In consequence, the total costs of the containment policy are unnecessarily large. A regional differentiation of containment policies, dependent on the local infection situation, seems to be preferable, at least in countries where the federal institutions are capable of conducting and supervising locally differing policies. The German states seem to have realized this in early May, when first Thuringia (May 4) and Bavaria (May 5) departed from the country-wide strategy and relaxed various measures. Only a few days later, on May 7, Chancellor Merkel announced the end of the regular coordination meetings and declared the states to be responsible for further containment strategies.

References

Abouk, R. and B. Heydari (2020, April). The Immediate Effect of COVID-19 Policies on Social Distancing Behavior in the United States. *medRxiv*, 2020.04.07.20057356. Publisher: Cold Spring Harbor Laboratory Press.

an der Heiden, M. and U. Buchholz (2020, March). RKI - Coronavirus SARS-CoV-2 - Modellierung von Beispielszenarien der SARS-CoV-2-Epidemie 2020 in Deutschland (20.3.2020).

Bracher, J. and L. Held (2020a, March). Endemic-epidemic models with discrete-time serial interval distributions for infectious disease prediction. *arXiv:1901.03090 [stat]*. arXiv: 1901.03090.

Bracher, J. and L. Held (2020b, March). A marginal moment matching approach for fitting endemic-epidemic models to underreported disease surveillance counts. *arXiv:2003.05885 [stat]*. arXiv: 2003.05885.

Britton, T. (2010, May). Stochastic epidemic models: A survey. *Mathematical Biosciences 225*(1), 24–35.

Chernozhukov, V., H. Kasaha, and P. Schrimpf (2020, May). Causal Impact of Masks, Policies, Behavior on Early Covid-19 Pandemic in the U.S. *arXiv:2005.14168 [econ, stat]*. arXiv: 2005.14168.

Courtemanche, C., J. Garuccio, A. Le, J. Pinkston, and A. Yelowitz (2020, May). Strong Social Distancing Measures In The United States Reduced The COVID-19 Growth Rate. *Health Affairs*, 10.1377/hlthaff.2020.00608. Publisher: Health Affairs.

Dehning, J., J. Zierenberg, F. P. Spitzner, M. Wibral, J. P. Neto, M. Wilczek, and V. Priesemann (2020, May). Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science*. Publisher: American Association for the Advancement of Science Section: Research Article.

Donsimoni, J. R., R. Glawion, B. Plachter, and K. Wälde (2020, June). Projecting the spread of COVID-19 for Germany. *German Economic Review 21*(2), 181–216. Publisher: De Gruyter Section: German Economic Review.

Felbermayr, G., S. Chowdry, and J. Hinz (2020). The Spread of Coronavirus from Ischgl through Germany. *Covid Economics* (22), 177–204. Library Catalog: www.ifw-kiel.de.

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, M. Monod, A. C. Ghani, C. A. Donnelly, S. M. Riley, M. A. C. Vollmer, N. M. Ferguson, L. C. Okell, and S. Bhatt (2020, June). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 1–8. Publisher: Nature Publishing Group.

Giuliani, D., M. M. Dickson, G. Espa, and F. Santi (2020, March). Modelling and Predicting the Spatio-Temporal Spread of Coronavirus Disease 2019 (COVID-19) in Italy. SSRN Scholarly Paper ID 3559569, Social Science Research Network, Rochester, NY.

Glogowsky, U., E. Hansen, and S. Schächtele (2020, June). How Effective are Social Distancing Policies? Evidence on the Fight Against COVID-19 from Germany. SSRN Scholarly Paper ID 3619845, Social Science Research Network, Rochester, NY.

Gneiting, T. and M. Katzfuss (2014). Probabilistic Forecasting. *Annual Review of Statistics and Its Application 1*(1), 125–151. _eprint: https://doi.org/10.1146/annurev-statistics-062713-085831.

Gros, C., R. Valenti, L. Schneider, K. Valenti, and D. Gros (2020, April). Containment efficiency and control strategies for the Corona pandemic costs. *arXiv:2004.00493 [physics, q-bio, q-fin]*. arXiv: 2004.00493.

Hartl, T., K. Wälde, and E. Weber (2020, April). Measuring the impact of the German public shutdown on the spread of Covid-19. *Covid Economics 1*, 25–32.

Held, L., M. Höhle, and M. Hofmann (2005, October). A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling 5*(3), 187–199. Publisher: SAGE Publications India.

Homburg, S. (2020, April). Evidenz zur Coronainfektion und der Wirkung des Lockdown. Technical Report 670, Leibniz-University of Hannover.

Hsiang, S., D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, P. Lau, J. Lee, E. Rolf, J. Tseng, and T. Wu (2020, June). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 1–9. Publisher: Nature Publishing Group.

Kermack, W. O., A. G. McKendrick, and G. T. Walker (1927, August). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 115*(772), 700–721. Publisher: Royal Society.

Kraemer, M. U. G., C.-H. Yang, B. Gutierrez, C.-H. Wu, B. Klein, D. M. Pigott, O. C.-. D. W. Group†, L. d. Plessis, N. R. Faria, R. Li, W. P. Hanage, J. S. Brownstein, M. Layan, A. Vespignani, H. Tian, C. Dye, O. G. Pybus, and S. V. Scarpino (2020, May). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science 368*(6490), 493–497. Publisher: American Association for the Advancement of Science Section: Research Article.

Maier, B. F. and D. Brockmann (2020, May). Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China. *Science 368*(6492), 742–746. Publisher: American Association for the Advancement of Science Section: Research Article.

McAloon, C. G., A. Collins, K. Hunt, A. Barber, A. Byrne, F. Butler, M. Casey, J. M. Griffin, E. Lane, D. McEvoy, P. Wall, M. J. Green, L. O'Grady, and S. J. More (2020, April). The incubation period of COVID-19: A rapid systematic review and meta-analysis of observational research. *medRxiv*, 2020.04.24.20073957. Publisher: Cold Spring Harbor Laboratory Press.

Mense, A. and C. Michelsen (2020). Räumliche Ausbreitung von COVID-19 durch interregionale Verflechtungen. *Wirtschaftsdienst 2020*(6), 416–421.

Meyer, S. and L. Held (2014, September). Power-law models for infectious disease spread. *Annals of Applied Statistics 8*(3), 1612–1639. Publisher: Institute of Mathematical Statistics.

Meyer, S., L. Held, and M. Höhle (2017, May). Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package surveillance. *Journal of Statistical Software 77*(1), 1–55. Number: 1.

Orea, L. and I. C. Álvarez (2020, April). How effective has been the Spanish lockdown to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces. Technical Report 2020-03, FEDEA. Publication Title: Working Papers.

Paul, M. and L. Held (2011). Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. *Statistics in Medicine 30*(10), 1118–1136. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.4177.

Petherick, A., N. Angrist, T. Phillips, S. Webster, and B. Kira (2020, May). Variation in government responses to COVID-19. *University of Oxford, Blavatnik School working paper 2020/032*(Version 6.0). Library Catalog: www.bsg.ox.ac.uk.

Streeck, H., B. Schulte, B. Kuemmerer, E. Richter, T. Hoeller, C. Fuhrmann, E. Bartok, R. Dolscheid, M. Berger, L. Wessendorf, M. Eschbach-Bludau, A. Kellings, A. Schwaiger, M. Coenen, P. Hoffmann, M. Noethen, A.-M. Eis-Huebinger, M. Exner, R. Schmithausen, M. Schmid, and G. Hartmann (2020, June). Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. *medRxiv*, 2020.05.04.20090076. Publisher: Cold Spring Harbor Laboratory Press.

Wieland, T. (2020, May). Flatten the Curve! Modeling SARS-CoV-2/COVID-19 Growth in Germany on the County Level. *medRxiv*, 2020.05.14.20101667. Publisher: Cold Spring Harbor Laboratory Press.