# A Bayesian approach to analyze the efficacy of government SARS-CoV-2 policies in European countries

## Abstract

Even when the SARS-CoV-2 pandemic recedes, evidence based researches regarding the effectiveness of pharmaceutical and non-pharmaceutical governments interventions (NPIs) are ongoing. We utilize a Bayesian approach to examine the impact of government policies on containing the virus spread based on the longitudinal data of 30 European countries from early 2020 up to the middle of the year 2022. Thereby, we analyze the observations of the SARS-CoV-2 pandemic growth rates with relation to a comprehensive set of the pandemic influential factors comprising each country’s NPIs per day, each country’s pre and post vaccination period, each country’s per day dominant virus variants as well as the seasonality impacts on the spread of the virus in each country. The assessments regarding the impact of the influential factors reveal similar patterns over the studied countries and lend themselves to inform public health policy and contribute to reasonable selection of prevention policies in the subsequent or future epidemics.

## Introduction

While governments’ vaccination and non-pharmaceutical interventions (NPI) are generally acknowledged to control the spread of the virus during the recent SARS-CoV-2 pandemic, evidence-based research in order to convey analysis regarding the effectiveness of the governments’ policies – are proceeding. Our research aims at running a comprehensive longitudinal analysis capable to draw inferences from the data of European countries’ policies between the early 2020 up to the middle of the year 2022.

A number of preceding studies have estimated the effectiveness of vaccination policies and NPIs across different spatial and temporal scopes of the globe relying on different input variables and analytical approaches. The outcomes of the existing literature are partially overlapping but simultaneously are not entirely consistent with each other. Hence, exploring more insights with regard to control the growth of epidemic needs further analytical approaches.

Among other studies, a longitudinal analysis with regard to the relationship between pandemic growth rates and various pandemic influential factors by means of the data from European countries is detected by us as a relevant research gap. A Bayesian assessment of the growth rates of the pandemic with incorporation of a comprehensive set of pandemic influential factors comprising each country’s NPIs as well as the pre and post vaccination periods, dominant various virus variants and seasonality in each country is realized in this paper.

The core model of our study compares two distinct conditions for each element of the set of pandemic influential factors in each European country, in which it explores the (predictive) distribution of pandemic growth rates when a selected influential factor (as an explanatory variable) has been active and it explores the (predictive) distribution of pandemic growth rates when the selected explanatory variable has not been active and conveys the probability of whether and to which degree, activation of the selected factor in comparison to non-activation of it, is expected to comprise *less* values of growth rates. While, the efficiency of each factor is perused independent from other factors through the analysis, yet, pooling the experiences through analyzing the data from 30 countries and over a relative extensive range of the time horizon can convey a reliable global view with regard to the effectiveness of each explored explanatory variable.

The remainder of the paper is as follows. In section 2, the literature background is explained. Section 3 comprehends how the data regarding the set of the influential factors are extracted, cleaned and merged into a unique data table, how the applied Bayesian model is constructed and utilized, and how the inputs (explanatory variables) and the output (the dependent variable) of the model are defined. Section 4 serves to present the drawn inferences regarding the significance and overall impact of each of pandemic influential variables in each country and over all countries, respectively. Further discussion and concluding remarks are highlighted in section 5.

## Background

Various Various studies within the preceding literature among others Flaxman et. al. (2020), Li et. al., (2022), Huy et. al., (2022), Ge et. al., (2022), Leech et. al., (2022), Bahnholzer et. a., (2022), Liu et. al., (2022), Stockes et. al. (2022), Ge et. al., (2022), Liu et. al., (2022), Zhou et. al., (2022) and Lawson & Rotejanaprasert (2023) have estimated the effectiveness of government SARS-CoV-2 policies across multiple countries and regions by means of different methodologies and criteria. Table 1 expresses a summary of the recent studies within this context.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Study | Method | Evaluation criterion | Scope | Key results |
| [Li et. al., (2022)](https://pubsonline.informs.org/doi/full/10.1287/opre.2022.2306) | Epidemiological compartment model | Number of the new cases | Pre-vaccination era of the pandemic in large-scale epidemics areas, e.g. the United States, United Kingdom, and Russia | Mass gathering restrictions and school closings are associated with the largest average reductions in infection rates |
| [Liu Y et. al., (2022)](https://aricjournal.biomedcentral.com/articles/10.1186/s13756-021-01039-x) | Lasso regression | Reproduction numbers | European countries within the year 2020 | The mobility index generated by personal behavior in prevention and control may be more important than wearing a mask |
| [Stockes et. al. (2022)](https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-022-13546-6) | Multivariate regression | Mortality rate | 130 countries up to June 2020 | Earlier and stricter school and workplace closures are associated with lower mortality rates |
| [Huy et. al., (2022)](https://www.mdpi.com/1660-4601/19/3/1139) | Longitudinal correlation analysis | Growth rate of infection | 30 Asian countries over the 20 weeks of the pre- and post-vaccination period | The facial covering policy in the pre-vaccination period and restrictions on gatherings and public transport closure in the post-vaccination period, are the most effective intervention |
| [Ge et. al., (2022](https://www.nature.com/articles/s41467-022-30897-1)) | Bayesian regression | Reproduction numbers | European countries from August 2020 to October 2021 | NPIs might remain an important complementary to vaccination in reducing COVID-19 transmission before herd immunity has been reached |
| [Leech et. al., (2022)](https://www.pnas.org/doi/10.1073/pnas.2119266119) | Bayesian regression | Reproduction numbers | 56 countries between May and September 2020 | Mask wearing in community settings reduces SARS-CoV-2 transmission |
| [Ge et. al., (2022)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8666325/) | Bayesian regression | Growth rate of infection | 133 countries within 2020 | Gathering restrictions and facial coverings played significant roles in epidemic mitigation before the vaccine rollout |
| Liu D et. al.,(2022) | Deep neural network | Number of the new cases | All counties of the United States from January 1 to July 10, 2020 | Social distancing is essential to control the spread of COVID-19, as well as the duration and degree of implementing social distancing interventions |
| Zhou et. al.,(2022) | Deep neural network | Number of the new cases | 3 states of the United States from April 1st, 2020 to March 31st, 2021 | School closing, canceling the public events and closing public transport are the most successful containment and closure policies |

Table 1: Key features from the preceding studies of vaccination and NPIs’ effectiveness

Li et. al., (2022) introduce a compartmental epidemiological model which predicts the detected cases and deaths in large scale epidemic areas. The model is then used to quantify the impact of government interventions on the pandemic’s spread. Liu et. al., (2022) use a lasso regression model to compute reproduction numbers associated with government control and prevention policies. Stockes et. al. (2022) use a multivariate regression model to understand the relative impact of NPIs on mortality rates. Huy et. al., (2022) apply Longitudinal correlation analysis to examine the effectiveness of NPIs on decreasing the epidemic growth of SARS-COV-2 between, before and after the vaccine roll-outs period. Ge et. al., 2022, Leech et. al., (2022), Bahnholzer et. a., (2022), Ge et. al., (2022) and Lawson & Rotejanaprasert (2023) utilize Bayesian inference models with different criteria applied in various geographical areas. Liu et. al., (2022) and Zhou et. al., (2022) are among the rare studies, which apply deep learning methods to analyze the efficacy of government SARS-COV-2 intervention policies. The majority of deep learning approaches within the recent pandemic context are comprising time series analysis focusing on forecasting the spread of the virus (e.g. case prediction) *n* days ahead of the time based on the previous times statistic reports as input (see e.g. Xu et al., 2022; Verma et al., 2022; Alassafi et al., 2022; Sinha et al., 2022).

Not all the above mentioned studies focus on research regarding the effect of the *individual* sets of government policies. Bahnholzer et. al., (2022) estimate country-specific differences of the effectiveness of overall NPIs. Lawson & Rotejanaprasert (2023) analyze the spread of SARS-COV-2 to assess the possible counter-factual realization of disease spread in an incidence that would have been evidenced with no NPI. Ge et. al., 2022 analyze the effect of NPIs’ *stringency* and vaccination together to reduce the transmission of the pandemic.

Those studies, which look into the effect of individual NPIs, are applied within various regions and come out to some extent with various results. Li et. al., (2022) find that mass gathering restrictions and school closings are associated with the largest average reductions in infection rates in large scale epidemic areas. While studying the effects of government policies from the empirical changes of weekly growth rates within different pandemic waves in 133 countries before vaccine roll-outs, Ge et. al., (2022) find that gathering restrictions and facial coverings play significant roles in epidemic mitigation.

Among Asian countries, the facial covering policy in the pre-vaccination period and restrictions on gatherings and public transport closure in the post-vaccination period, are the most effective interventions, according to Huy et. al., (2022).

Liu et. al., (2022) (using lasso regression model) conclude that the rate of masks used in individual prevention does not seem to be related to cumulative mortality or morbidity in European countries, but that the mobility index generated by personal behavior might be the more important prevention policy.

Leech et. al., (2022) aim at declaring the inconsistency regarding the effects of mask wearing in the literature. They focus on how masks mandates represent a poor proxy for mask-wearing effectiveness. They study 56 countries between May and September 2020 and conclude that mask wearing reduces the reproduction rate of SARS-CoV-2 by 19 percent.

Evidence-based research into the effectiveness of government SARS-CoV-2 policies is still under progress and can be traced through the recent literature e.g. in the works of Chirwa et. al. (2023), Pozo-Martin et. al. (2023), Damette et. al. (2023), Nguyen et. al. (203), Kamineni et. al. (2023), Novak et. al. (2023) and Pung et. al. (2023). This paper departs from the preceding context methodologically and through its scope of research. We apply a Bayesian approach to draw predictive inferences with regard to the pandemic growth rates and various government policies as well as the effects of seasonality, dominant virus variants and pre and post vaccination period in 30 European countries over a relative long period of the SARS-CoV-2 pandemic.

## Analysis basis

### 3.1 Basic response variable

In line with preceding literature analysis we presume *reproduction rate* and the *growth rate*  of the pandemic as the main response variable to a number of explanatory variables.

The calculation of the point estimate of reproduction number for a given day *i* is defined as the simple quotient of the number of new cases for that day divided by the number of new cases 7 days *before* ([statista](https://de.statista.com/statistik/daten/studie/1117478/umfrage/reproduktionszahl-des-coronavirus-covid-19-in-deutschland/" \o "https://de.statista.com/)). Hence, if the number of observed SARS-CoV-2 positive cases in a day *i* is equal to *I(i)* and the average number of new cases from the day *i-n-1* on until the day *i* (7-days backward looking) is articulated by the incidence presented in the equation (1):

, n=7 (1)

Then, the formal definition of the reproduction number *R* is expressed in the term through equation (2)[[1]](#footnote-1)

(2)

As we are going to have a predictive criterion to be related to the the local effect of each NPI at each arbitrary day *i* in hand, we replace the numerator term of the equation 2 by a weekly *forward* looking average as is articulated by the incidence presented in the equation (3):

, n=7 (3)

And reformulate the equation (2) in equation (4):

(4)

In addition, we work with a simple transformed version of equation 4, which expresses the *growth rate*  of the pandemic as a relative indicator for *changing* the reported infection numbers, as described in equation (5):

*=* (5)

represents a kind of *symmetric* predictive criteria to indicate the change in the spread of the virus between 7 days’ average backward and 7 days’ average forward at any arbitrary day *i*.

### 3.2 Data

In this paper we utilize 4 data source files (per country per day government measures NPIs, per country per day virus variant distributions, per country per day per dose new vaccinations and per country per day new reported infections) from the information provided by [European Centre for Disease Prevention and Control](https://www.ecdc.europa.eu/en)[[2]](#footnote-2). We studied the countries within the time spans as it is presented in the table 2.[[3]](#footnote-3)

|  |  |  |
| --- | --- | --- |
| Country | from | until |
| Luxembourg | 2020-03-09 | 2022-06-29 |
| Hungary | '2020-03-11 | 2021-06-11 |
| Germany | 2020-02-14 | 2022-06-10 |
| Latvia | 2020-03-12 | 2022-05-14 |
| Spain | 2020-03-09 | 2022-04-19 |
| Norway | 2020-03-10 | 2022-02-12 |
| Romania | 2020-03-16 | 2022-03-08 |
| Liechtenstein | 2020-10-01 | 2022-03-31 |
| Portugal | 2020-02-27 | 2022-06-30 |
| Belgium | 2020-03-10 | 2022-05-22 |
| Bulgaria | 2020-02-24 | 2022-07-17 |
| France | 2020-02-29 | 2022-07-31 |
| Italy | 2020-03-04 | 2022-06-14 |
| Slovakia | 2020-03-10 | 2022-04-20 |
| Ireland | 2020-01-27 | 2022-03-15 |
| Denmark | 2020-03-11 | 2022-04-27 |
| Iceland | 2020-07-23 | 2021-12-03 |
| Cyprus | 2020-03-10 | 2022-04-14 |
| Greece | 2020-02-27 | 2022-04-30 |
| Netherlands | 2020-03-12 | 2022-05-31 |
| Czech | 2020-03-10 | 2022-05-05 |
| Lithuania | 2020-02-24 | 2022-04-30 |
| Austria | 2020-03-10 | 2022-05-15 |
| Poland | 2020-03-08 | 2022-03-27 |
| Slovenia | 2020-03-10 | 2022-05-30 |
| Estonia | 2020-01-28 | 2022-07-20 |
| Malta | 2020-03-13 | 2022-07-24 |
| Croatia | 2020-03-03 | 2022-04-30 |
| Finland | 2020-03-12 | 2022-06-30 |
| Sweden | 2020-03-12 | 2022-03-31 |

Table 2: data time spans by countries

The day-to-day based data of the accumulative vaccination doses, distribution of the types of virus variants per day in each country, implemented each European country’s government measures NPIs per day as well as the data regarding the number of reported new infections (per day)[[4]](#footnote-4) in each of studied 30 countries are first extracted from the 4 data source files. The data sets are merged then into a unique pandas table named *data\_encoded*. The code for constructing the data table *data\_encoded* through merging the primary data sets as well as relevant data files have been appended to the supplementary material of this paper.

Each of government’s policies NPIs is set equal to 1 when it is implemented within a day in a country and is set to be zero when not implemented. A detailed explanation of NPIs is presented in Lorenzo et. al. (2022). A variable dictionary related to the data on non-pharmaceutical interventions NPIs can be downloaded from the [ecdc](https://www.ecdc.europa.eu/sites/default/files/documents/Variable_Dictionary_and_Disclaimer_non-pharmacautical_measures_v3.pdf) website. The set of studied NPIs are represented in alphabetical order in table 3.

|  |  |  |  |
| --- | --- | --- | --- |
| Government NPIs[[5]](#footnote-5) |  |  |  |
| 'AdaptationOfWorkplace' | 'IndoorOver100' | 'MassGather50' | 'RegionalStayHomeOrder' |
| 'AdaptationOfWorkplacePartial' | 'IndoorOver1000' | 'MassGather50Partial' | 'RegionalStayHomeOrderPartial' |
| 'BanOnAllEvents' | 'IndoorOver50' | 'MassGatherAll' | 'RestaurantsCafes' |
| 'BanOnAllEventsPartial' | 'IndoorOver500' | 'MassGatherAllPartial' | 'RestaurantsCafesPartial' |
| 'ClosDaycare' | 'MasksMandatoryAllSpaces' | 'NonEssentialShops' | 'SocialCircle' |
| 'ClosDaycarePartial' | 'MasksMandatoryAllSpacesPartial' | 'NonEssentialShopsPartial' | 'SocialCirclePartial' |
| 'ClosHigh' | 'MasksMandatoryClosedSpaces' | 'OutdoorOver100' | 'StayHomeGen' |
| 'ClosHighPartial' | 'MasksMandatoryClosedSpacesPartial' | 'OutdoorOver1000' | 'StayHomeGenPartial' |
| 'ClosPrim' | 'MasksVoluntaryAllSpaces' | 'OutdoorOver50' | 'StayHomeOrder' |
| 'ClosPrimPartial' | 'MasksVoluntaryAllSpacesPartial' | 'OutdoorOver500' | 'StayHomeOrderPartial' |
| 'ClosPubAny' | 'MasksVoluntaryClosedSpaces' | 'PlaceOfWorship' | 'StayHomeRiskG' |
| ,'ClosPubAnyPartial' | 'MasksVoluntaryClosedSpacesPartial' | 'PlaceOfWorshipPartial', | 'StayHomeRiskGPartial' |
| 'ClosSec' |  | 'PrivateGatheringRestrictions' | 'Teleworking' |
| 'ClosSecPartial' |  | 'PrivateGatheringRestrictionsPartial', | 'TeleworkingPartial' |
| 'ClosureOfPublicTransport' |  | 'QuarantineForInternationalTravellers' | 'WorkplaceClosures' |
| 'ClosureOfPublicTransportPartial' |  | 'QuarantineForInternationalTravellersPartial' | 'WorkplaceClosuresPartial' |
| 'EntertainmentVenues' |  |  |  |
| 'EntertainmentVenuesPartial' |  |  |  |
| 'GymsSportsCentres' |  |  |  |
| 'GymsSportsCentresPartial' |  |  |  |
| 'HotelsOtherAccommodation' |  |  |  |
| 'HotelsOtherAccommodationPartial' |  |  |  |
|  |  |  |  |

Table 3: NPIs

The vaccination data is elaborated through computing the accumulated percentage numbers of each of the first 5 vaccine doses received by the population of each corresponded country in each day. The pre vaccination time period (where the value for the first dose is equal to zero) and the post vaccination time period (where the value for the first dose is larger than zero) are distinguished for each country’s data in a column named *vaccination\_modus*. Each virus variant is presented within a day and a country through the percentage it is sequenced. Hence, the sum of percentage values of all virus variants in a country within each day is equal to 1.[[6]](#footnote-6) The dominant virus variant in each country and each day is defined as the virus variant with the highest sequenced percentage and owns its own column named *dominant\_virus* within the *data\_encoded*. The most frequently observed dominant virus types in the table are *B.1.617.2*, *B.1.1.7* and *BA.2*.

Furthermore, 7 days’ average backward and 7 days’ average forward incidences described in section 3.1 at any arbitrary day *i* for each country are integrated into the *data\_encoded*.

### 3.3 Basic Bayes model

In our study, the distribution of the growth rate conditioned to a distinct SARS-CoV-2 explanatory factor *E* is computed in accordance with the Bayes theorem explained in Bishop (2006) as expressed in the equation (4):

(4)

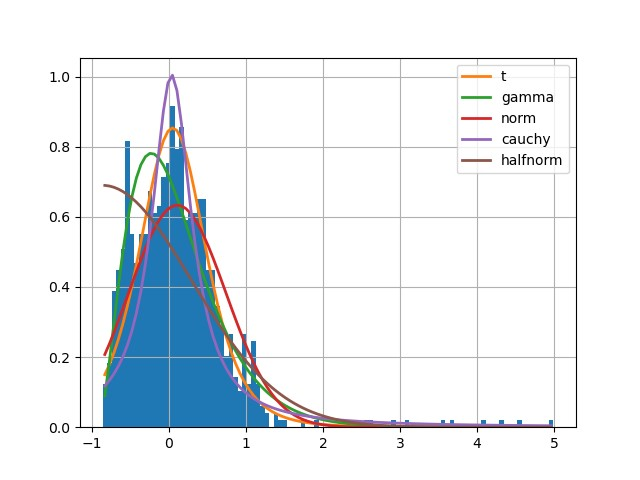
The equation 4 is composed of a prior probability π() function (, which expresses the overall distribution of the pandemic growth rates without the data being observed), a likelihood function (, which expresses the conditional distribution of the to be active or not active upon the prior distribution parametrized by ), and a posterior function (, which is defined proportional to the product of the likelihood and prior and represents the distribution of after incorporating the evidence provided by the prior probability and likelihood).

The Bayes theorem is applied in our study to infer conditioned to through observing the distribution of pandemic growth rates within the days where the has been active as well as observation the days, where the has *not* been active.

### 3.4 Basic distribution

In order to model the distribution of probabilistic parameters in our model and based on the fitting data to the studied countries, we use a Student-T distribution to predict the distribution of s in our analysis[[7]](#footnote-7). The reason of choosing a Student-T is its fitness to the shape of the s in the studied countries. Figure 2 and Table 2 depict the histogram of the overall s numbers by the example of Germany during the pandemic between 03.02.2020 and 07.07. Basically, the t-distribution gives more probability to observations in the tails of the distribution than a standard normal distribution, which suits well the distribution of in our study.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | sumsquare\_error | aic | bic | kl\_div | ks\_statistic | ks\_pvalue |
| t | 0.555638 | 958.8544 | -6112.84 | inf | 0.052344 | 0.019538 |
| gamma | 0.601352 | 915.6697 | -6046.59 | inf | 0.052242 | 0.01989 |
| norm | 0.700098 | 1780.438 | -5925.91 | inf | 0.07842 | 6.27E-05 |
| cauchy | 1.195052 | 662.0179 | -5477.81 | inf | 0.10869 | 4.47E-09 |
| halfnorm | 1.71739 | 932.5817 | -5173.94 | inf | 0.125505 | 5.78E-12 |



Tabl2 2 & Figure 1 (high resolution image): Student-T distribution fits optimal to describe the likelihood function of growth rate values in Germany.

### 3.5 Applied Bayesian approach

To obtain the posterior probability of with regard to presence and not-presence of a selected factor, we analyze the data based on a *hierarchical* Bayesian model ([Congdon, 2019](https://www.taylorfrancis.com/books/mono/10.1201/9780429113352/bayesian-hierarchical-models-peter-congdon); [Johnson et. al., 2022](https://www.bayesrulesbook.com/)). A hierarchical Bayesian model considers a hyper parameter at the top level of its analysis, as well as specific parameters in its lower level. The top hierarchical level takes the overall distribution of growth rates in a country into account (through pooling all data from a country regardless of the condition whether a certain factor *E* has been active or not). The lower level takes the situation-specific developments into account i.e. whether the selected influential factor is implemented or not. The model used in this paper is illustrated in the figure 2. The top level of the model in figure 2 depicts the hyper mu parameter, which is a normal distribution function consisting of a mean and a standard deviation, which both are resulting from pooling all data from a related country. The left hand side of the model (consisting of variables with *positive* sign at the end) devotes itself to the effect of *activation of* a certain factor *E* on the epidemic growth rates. The right hand side of the model (consisting of variables with *negative* sign at the end) devotes itself to the effect of *non-activation* of the selected factor on the epidemic growth rates. The lower level of the model is inferring posterior predictive values in our analysis with regard to the positive signed and negative signed parameters, respectively. Thereby, it uses the Student-T distribution as it is justified at 3.4. The parameter *nu* is the *degree of freedom* part of the subsequent Student-T distribution, which usually ranges from 0 to 30 and becomes shared between the left hand side and the right hand side of the model.

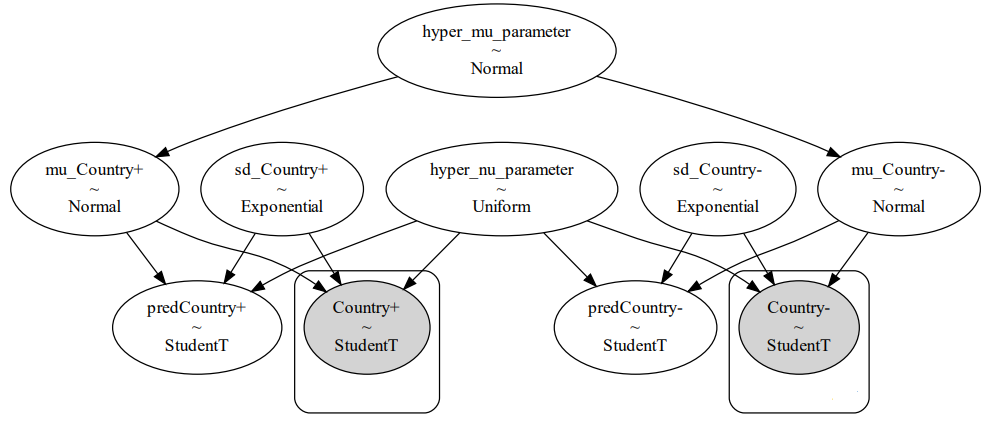


Figure 2 (high resolution image): The hierarchical Bayesian model for analyzing a distinct government NPI in a distinct country.

The presence of the *sd\_Country* parameters specified in the both right hand side and left hand side of the model figure provides a weekly informative prior which permits the lower level of the model to sample its parameters (either corresponded to the activation of a factor on the left hand side or corresponded to the non-activation of the selected factor in the right hand side) without bias (towards the overall higher parameters of the model i.e. the hyper parameter) with some standard deviation. Note that, while basically Normal distribution provides a reasonable prior for the *mean* parameter, the *Exponential* distributions provides reasonable priors for the *standard deviation* i.e. for the distribution of the abovementioned sd\_Country terms (Johnson et. al, 2022).

### 3.5 Technical setting

Obtaining posterior distributions through describing steps at previous sections, are often analytically intractable. An accurate approximation of the posterior distributions is accomplished via using NUTS sampler implemented in PyMC3 by setting a sample size *k* in order to draw *k* samples from the posterior[[8]](#footnote-8). PyMC3 is an open source framework of probability distribution sampling with a syntax close to the natural syntax statisticians ([Salvatier et al., 2016](https://peerj.com/articles/cs-55/)). It incorporates the Markov Chain Monte Carlo (MCMC) algorithm designed for probability distribution sampling as the No-U-Turn Sampler (NUTS). The NUTS method is a self-tuning variant of Hamiltonian Monte Carlo (HMC) ([Duane et al., 1987](https://ui.adsabs.harvard.edu/abs/1987PhLB..195..216D/abstract)). As the HMC and NUTS apply the gradient information from the likelihood, they can achieve a fast convergence rate, which avoids the random walk behavior by taking a series of steps informed by first-order gradient information ([Hoffman & Gelman, 2014](https://www.jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf)). It is worth noting that PyMC3 takes the advantage of Theano ([Bergstra et al., 2010](https://www.iro.umontreal.ca/~lisa/pointeurs/theano_scipy2010.pdf)) as backend to transparently transcode models to a C code and compile them to machine code.

The results of the posterior MCMC sampling draws get stored in a PyMC3 data object called *trace*. The values of a trace are not only representing the obtained posterior values but also might indicate the reliability of parameter space exploration by the employed MCMC sampling algorithm. The model trace convergence is evaluated by means of the Gelman-Rubin *R\_hat* statistic ([Brooks and Gelman, 1998](https://www.tandfonline.com/doi/abs/10.1080/10618600.1998.10474787)). *R\_hat* Values not close to one indicate that trace values move wildly or get stuck and that sampling failed to cover the parameter space effectively. Hence, in practice we are looking for R\_hat values close to one. The resulting trace plots regarding the predictive distributions of by investigation of the Gyms and sport centers closings in Germany are shown in figure 3.

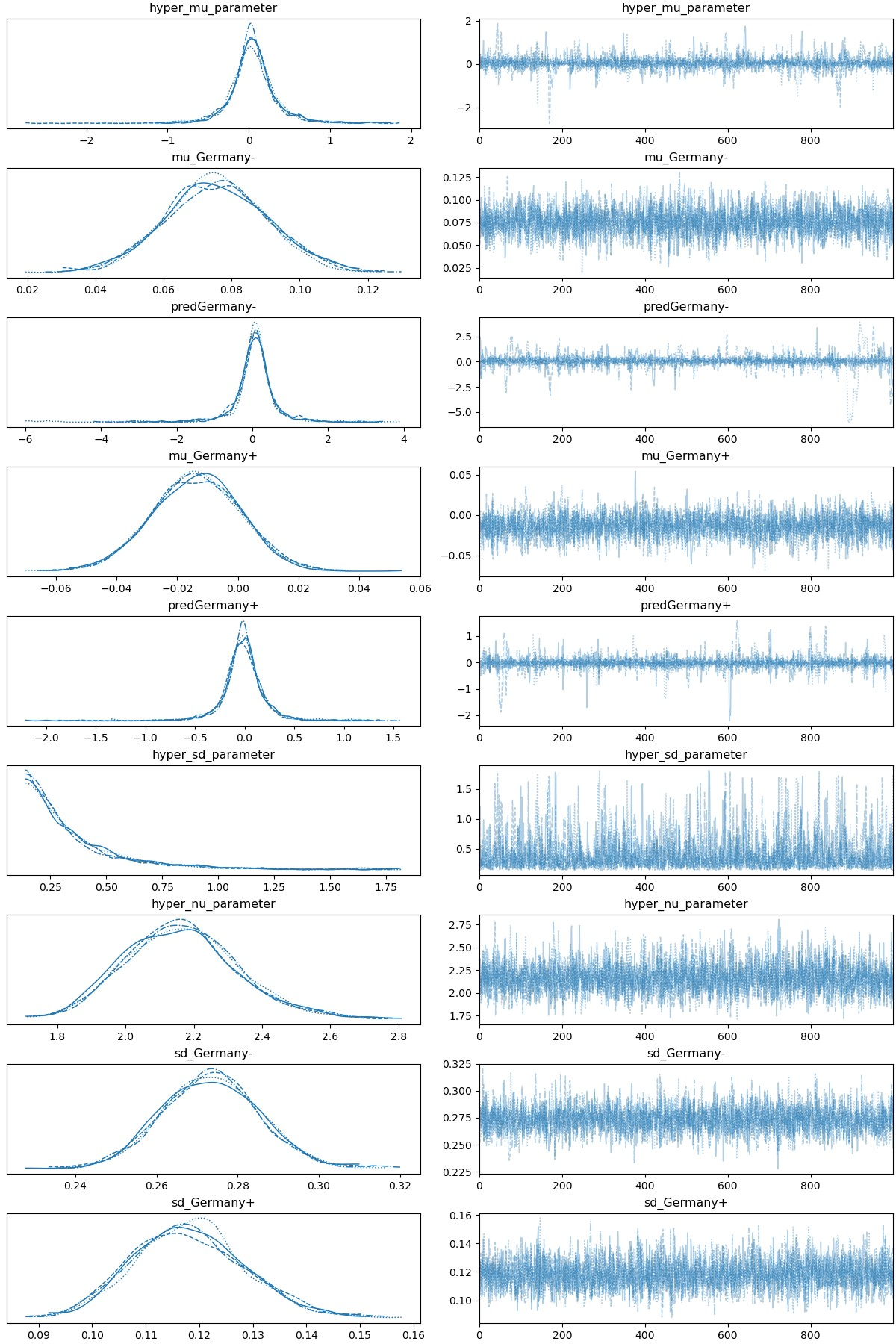


Figure 3 (high resolution image): The resulting trace regarding the effect of gyms and sport center closings in Germany.

### 3.6 Explanatory factors’ efficiencies

In this section we define the concept of *efficiency* of a factor to be the probability that presence (activation) of a factor within an arbitrary day, grants less value of growth rate in comparison to the absence (non-activation) of that factor. Before formalizing the definition in equation 5, one can look at the figures 5 and 6, which illuminate the inference procedure from the observed data (left column of the figure 5) to obtaining posterior predictive distributions of growth rates (middle and the right hand column of the figure 5) and to computing the efficiencies (figure 6) by means of the example of gyms and sport centers’ closures in 4 most populated European countries.

In order to compute each NPI’s efficiency in each country, we first distinguish between the days, where the selected factor (gyms and sport centers’ closures) has been active (blue bars at the left hand side of the figure 5) within each selected country and those days, where the selected has not been active (orange bars at the left hand side of the figure 5) in that country. A posterior predictive analysis succeeds by sampling posteriors from parameters existing in the figure 3. The model code is attached to the supplementary material of the paper. The results of the sampling inference from posterior distributions are depicted in the middle column of the figure 5 for each country. The right hand side column of the figure 5 represents probability density functions, which are corresponded to the probability bars for each country in the middle column as well as to the model trace parameters related to the pred\_Country+ and pred\_Country- in figure 3 (and as it is shown by the example of pred\_Germany in figure 4).

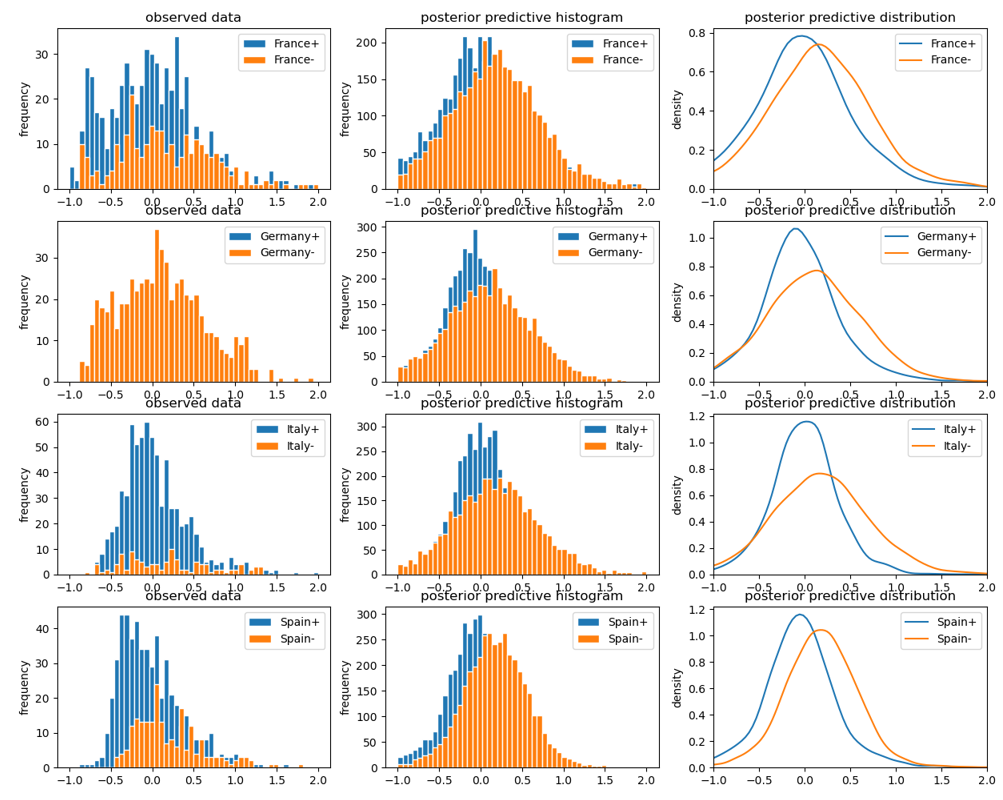


Figure 5 (high resolution image): Posterior predictive results regarding implementation (positive signed) versus non implementation (negative signed) of gyms and sport centers in 4 most populated European countries.

The distinguished (positive-signed and negative-signed) probability bars as well as density curves at the middle and right hand side columns of the figure 5 can be considered to compare the *efficiency* of activation versus non-activation of the selected NPI as the explanatory factor. Though, a precise statistically comparison between the outcome of each blue (positive signed) distribution and each orange (negative signed) distribution, can be carried out by 1000 times sampling 1000 random draws from both blue and orange distributions and looking whether the probability of s is *smaller* in the predictive (blue) distribution (representing the activation of the selected NPI) compared to s in the predictive (orange) distribution (representing the non-activation of the selected NPI). In other word, each time (from 1000), 1000 random draws from the positive signed trace[[9]](#footnote-9) distribution will generate a list of s named , which is of length 1000. Likewise, each time (from 1000) we do 1000 random draws from the negative signed trace[[10]](#footnote-10) distribution and generate a list of s named , which is of length 1000. Then, the efficiency ratio expressed in equation 5, becomes resulted from the comparison of each positional element *i* of the above mentioned two generated lists , which becomes 1 if true and 0 if false in equation 5.

(5)

It is furthermore worth noting that based on the definition, the explanatory variables with efficiency values larger than 50 percent can be labeled as *effective* with regard to the pandemic *control*. Analogously, the efficiency values smaller than 50 percent, however can be labeled as *effective* with regard to the pandemic *progression*. The histogram of the obtained 1000 computed efficiencies ‘s for the countries of figure 5 is depicted in figure 6.

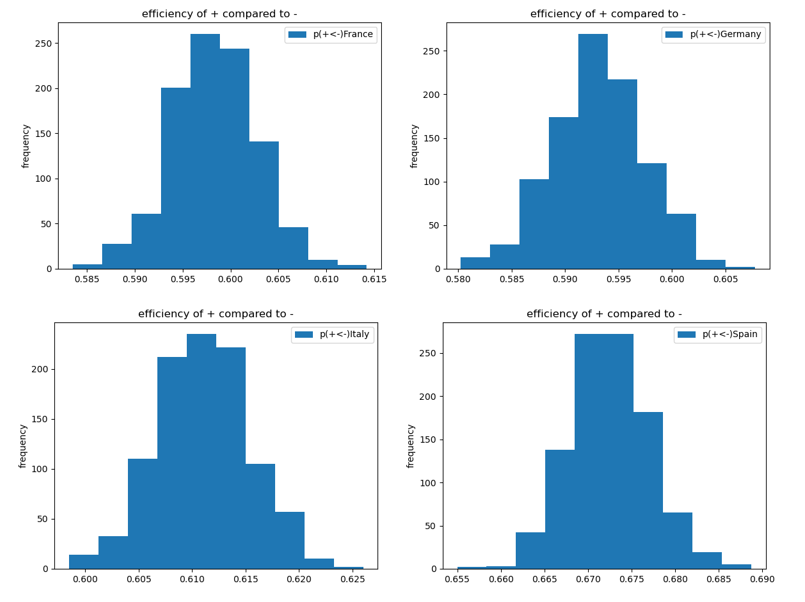


Figure 6 (high resolution image): Efficiency of implementation of the selected NPI (gyms and sport centers closures) in comparison to non-implementation of it within the selected subset of countries.

## Results

In this paper we analyze the set of explanatory factors containing a comprehensive set of NPIs as well as the most frequently sequenced virus variants, pre and post vaccination periods and seasonality impacts on the spread of the virus in 30 European countries.[[11]](#footnote-11) The detailed results regarding the summary of government NPI efficiencies and the related corresponded statistics of the traces are attached to the supplementary material of this paper. The *r-hat* statistic in all obtaining results are precisely enough close to 1.00 for all parameters. This indicates no problems during sampling. The average values corresponded to the each explanatory factor efficiencies is presented in figure 7.[[12]](#footnote-12)

From the first 3 rows of the figure 7 one can infer that within the days of the presence of the *B.1.1.7(Coronavirus Alpha variant*) in comparison to the days of non-activation of it, enormous considerable amount of reduction in the growth rates of the pandemic could have been achieved. Likewise, one can infer that how the activation of the explanatory variable *B.1.617.2 (Coronavirus Delta variant)* has been harmful. In contrast, the *BA.2 (Coronavirus Omicron variant) activation* might have been beneficial in terms of amelioration of the virus spread. Regarding the seasonality effects it is apparent that large probable growth rates of the virus is achieved by activation of the season Autumn (within September, October and November) and the smallest probabilities with regard to the pandemic growth is predicted to be in the season Spring (within March, April and May). The gained efficiencies in terms of less values of growth rates in Winter (within December, January and February) are higher than the anticipated efficiencies of Summer (June, July and August). With reference to the predecessor literature, the role of seasonal trends has been explored in Merow and Urban (2020) and Wiemken et. al. (2023). Merow and Urban (2020) developed statistical models that predict the maximum potential of COVID-19 worldwide and throughout the year and predicted that COVID-19 will decrease temporarily during summer, rebound by autumn, and peak next winter. In a more recent study, Wiemken et. al. (2023) use time series decomposition to extract the annual seasonal component of COVID-19 case, hospitalization, and mortality rates from March 2020 through December 2022 for the United States and Europe and identify seasonal spikes in COVID-19 from approximately November through April for all outcomes and in all countries. The results of our model *partially* overlap with the aforementioned studies when it comes to the rebounding the virus spread in the time span around October and November in Autumn and that within the time span around April a receding phase begins. It gets from the figure 7 obvious that in the majority of countries the efficiency of the post vaccination period is relatively higher that the pre vaccination.

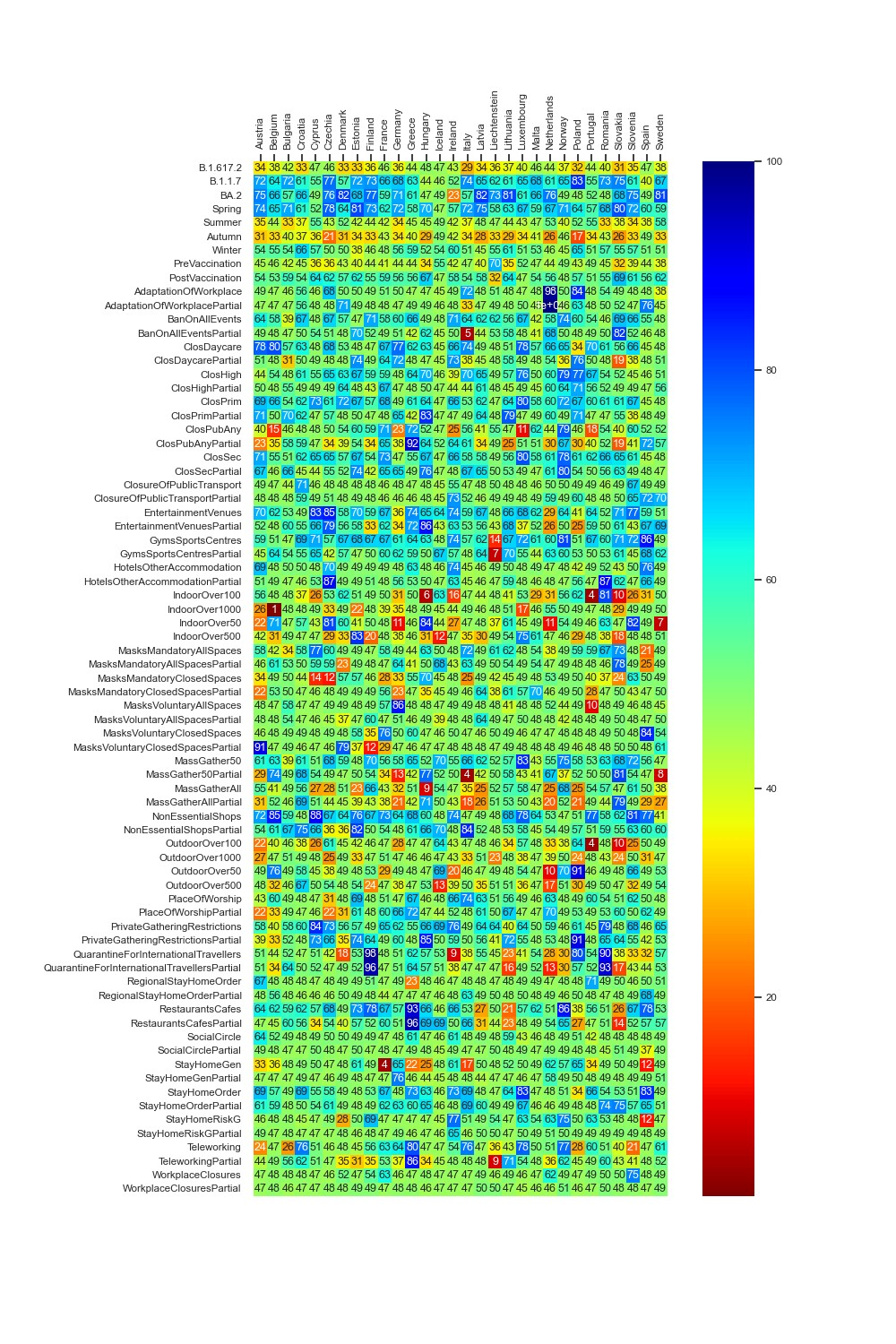
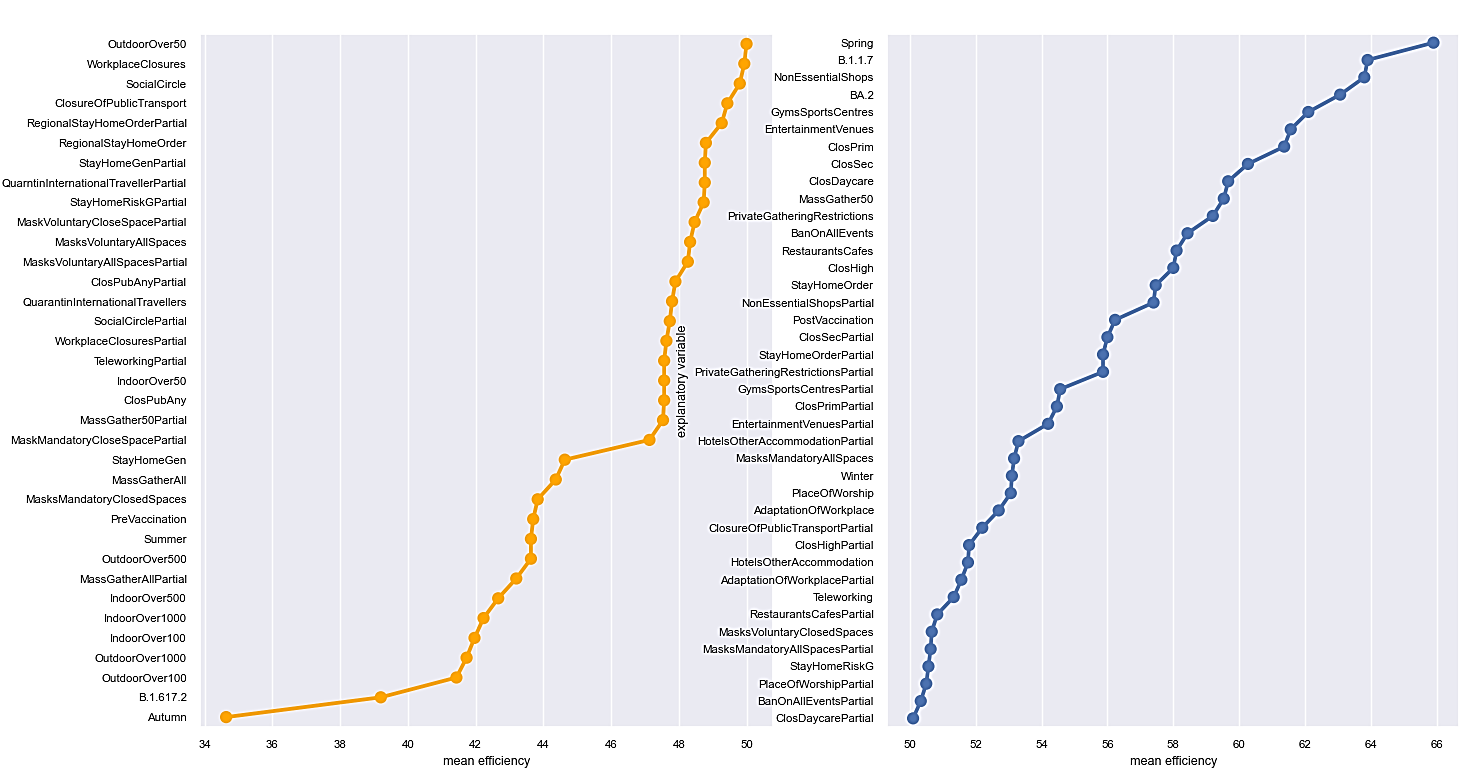


Figure 7 (high resolution image): the summary of explanatory factors average efficiencies in percentage values.

Within the NPIs[[13]](#footnote-13), it becomes evident from the experience of the European countries that among the government non-pharmaceutical interventions[[14]](#footnote-14), closing the primary schools, closures of non-essential shops (as a partially relaxed measure), restrictions on private gathering, stay-at-home orders for the general population (these are enforced and also referred to as ‘*lockdown*’) as a partially relaxed measure, closures of non-essential shops, mass gathering interventions i.e. interventions to limit mass/public gatherings (any interventions on mass gatherings up to 1000 participants included) and the closure of gyms and sport centers (as a partially relaxed measure) are the most impacting factors in containing the pandemic growth followed by a range of other NPIs which can be read from the depicted efficiency values depicted for each country in figure 7.

The next relative efficient intervention, which to some extent stands out is *MasksMandatoryAllSpaces,* which is encompassing the protective mask usage in all public spaces on mandatory basis (enforced by law). In contrast it is peculiar that the factor MasksMandatoryClosedSpaces representing protective mask use in closed public spaces/transport on mandatory basis (enforced by law) is just revealing efficiencies under 50 percent efficiency values in the most of the countries.  
The relative critical efficacy of mask mandates in *closed* space might offer further researches to elaborate more on the role of mask mandates to control the pandemic. While the effects of government mask mandates in the literature are to some extent inconsistent (Leech et. al., 2022), the *partially* inconstant result regarding mask mandates in *closed* spaces in figure 7 might be further examined. Thereby, there exists some hypothesis that enforcing masks works if and only when entire or at least a certain mass of the society becomes committed to it and it becomes significantly inefficient if it is prescribed or obeyed partially (Kai et.al., 2020) or if the observed levels of mask wearing in closed spaces are critical (Leech et. al., 2022). Some other hypothesis might raise the question of effectiveness of mask mandates in regions where commitment to face masks lacks cultural background (Barceló & Sheen, 2020). Despite that conjectures, examining the predicted non efficiency of mask mandates in *closed spaces* despite the high efficiency of *all spaces* mask mandates might not be explained completely based on the above mentioned hypothesis nor within the framework of our analysis method. This might need more illuminations.

In order to summarize the experiences of the entire 30 studied countries the average of the average efficiencies for each explanatory variables are computed and demonstrated in the figure 8. Due to the high length of the factor numbers, the resulted *averaged efficiency* terms in figure 8 are presented in two side by side panels. The right panel is indeed the continuation of the left hand side one. Thereby, the factors within the left hand side panel are those who are characterized by under 50 percent efficiency values and can be labeled as *effective* with regard to the pandemic *progression*. In contrast, the factors within the right hand side panel are those who are characterized by larger than 50 percent efficiency values and can be labeled as *effective* with regard to the pandemic *control*. Hence, the factors at the button of the left hand side panel (e.g. the season Autumn or the omicron variant) can be interpreted as the ones with the largest pandemic progression efficacy and the factors at the top levels of the right hand side panel e.g. the season Spring and the alpha variant) can be interpreted as the ones with the largest pandemic control efficacy.

 Figure 8 (high resolution image): The sorted average efficiencies of the pandemic influential factors over all studied countries.

## Conclusion

The research objective of our study was set out to run a probabilistic Bayesian correlation analysis based on the data of European countries and through a wide time span to gain an extended insight regarding the efficiency of governments policies to contain the pandemic growth rates. We measured the efficiency of a comprehensive set of pandemic influential factors comprising each country’s NPIs as well as the pre and post vaccination periods, dominant various virus variants and seasonality in a realistic manner. We compared two distinct circumstances for each corresponded influential factor, in which we explored the (predictive) distribution of pandemic growth rates when the selected factor was active and when the selected factor was not active, respectively and computed the probability of whether and to which degree, activating the selected factor in comparison to non-activating of it, is expected to trigger less values of growth rates. This approach might not be interpreted as a complete causal analysis. However, it can inform policy setting about similar patterns in containing the spread of the disease across 30 European countries over a relative long span of the observation. These similar insights might provide reasonable viewpoint to evaluate the governments NPIs effectiveness as well as other influential factors e.g. seasonal impacts. If a specific NPI shows poorly related predictive effectiveness when applied in several countries, this might be inviting to further research and more deliberate at policy setting level. Our analysis might not be capable to further elaborate regarding the partial inefficiencies when it comes to some NPIs e.g. mask mandates. The analysis, moreover does not incorporate the aggregate effects of the explanatory variables to shed light into the possible causal interrelations. is requires to be mixed with other analytical and qualitative methods to come out with more consistent policy relevant explanations.

## Acknowledgement

We would like to acknowledge the funding source of the Covid&AI project, which is a collaboration between four research groups at the University of Koblenz and is founded by German Science Foundation DFG.

## Supplementary material

You can find all supplementary material related to this paper consisting of code, data, results and figures in the [GitLab account](https://gitlab.uni-koblenz.de/hamedkhalili/covid_ai_project/-/tree/main/p_1) corresponded to this paper, which is provided by University of Koblenz.

## References

A. Riccardi, J. Gemignani, F. Fernández-Navarro and A. Heffernan, "Optimisation of Non-Pharmaceutical Measures in COVID-19 Growth via Neural Networks," in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 5, no. 1, pp. 79-91, Feb. 2021, doi: 10.1109/TETCI.2020.3046012.

Binggui Zhou, Guanghua Yang, Zheng Shi, Shaodan Ma, Interpretable Temporal Attention Network for COVID-19 forecasting, Applied Soft Computing, Volume 120, 2022, 108691, ISSN 1568-4946, <https://doi.org/10.1016/j.asoc.2022.108691>.

Banholzer, N., Feuerriegel, S. & Vach, W. Estimating and explaining cross-country variation in the effectiveness of non-pharmaceutical interventions during COVID-19. Sci Rep 12, 7526 (2022). <https://doi.org/10.1038/s41598-022-11362-x>

Bishop, Christopher M. ( 2006). Pattern recognition and machine learning. New York :Springer

Barceló J, Sheen GC. Voluntary adoption of social welfare-enhancing behavior: Mask-wearing in Spain during the COVID-19 outbreak. PLoS One. 2020 Dec 1;15(12):e0242764. doi: 10.1371/journal.pone.0242764. PMID: 33259531; PMCID: PMC7707551.

ByAlicia A. Johnson, Miles Q. Ott, Mine Dogucu. Bayes Rules! An Introduction to Applied Bayesian Modeling <https://doi.org/10.1201/9780429288340>

Congdon, P.D. (2019). Bayesian Hierarchical Models: With Applications Using R, Second Edition (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429113352>

Cory Merow and Mark C. 2020. Urban, Seasonality and uncertainty in global COVID-19 growth rates, Proceedings of the National Academy of Sciences, 117(44): 27456-27464. doi: 10.1073/pnas.2008590117

Duane, Simon ; Kennedy, A. D. ; Pendleton, Brian J. ; Roweth, Duncan. Hybrid Monte Carlo. 1987. Physics Letters B, Volume 195, Issue 2, p. 216-222. DOI:10.1016/0370-2693(87)91197-X

De Kai, Guy-Philippe Goldstein, Alexey Morgunov, Vishal Nangalia, Anna Rotkirch. Universal Masking is Urgent in the COVID-19 Pandemic: SEIR and Agent Based Models, Empirical Validation, Policy Recommendations. arXiv:2004.13553

Dixizi Liu, Weiping Ding, Zhijie Sasha Dong, Witold Pedrycz, Optimizing deep neural networks to predict the effect of social distancing on COVID-19 spread,Computers & Industrial Engineering, Volume 166, 2022, 107970, ISSN 0360-8352, <https://doi.org/10.1016/j.cie.2022.107970>

Ge, Y., Zhang, WB., Wu, X. et al. Untangling the changing impact of non-pharmaceutical interventions and vaccination on European COVID-19 trajectories. Nat Commun 13, 3106 (2022). <https://doi.org/10.1038/s41467-022-30897-1>

Gavin Leech, Charlie Rogers-Smith, Joshua Teperowski Monrad, Jonas B. Sandbrink, Benedict Snodin, Robert Zinkov, Benjamin, John S. Brownstein, Yarin Gal, Samir Bhatt, Mrinank Sharma, Sören Mindermann, Jan M. Brauner, and Laurence Aitchison. 2022. Mask wearing in community settings reduces SARS-CoV-2 transmission <https://www.pnas.org/doi/full/10.1073/pnas.2119266119>

Ge Y, Zhang WB, Liu H, Ruktanonchai CW, Hu M, Wu X, Song Y, Ruktanonchai NW, Yan W, Cleary E, Feng L, Li Z, Yang W, Liu M, Tatem AJ, Wang JF, Lai S. Impacts of worldwide individual non-pharmaceutical interventions on COVID-19 transmission across waves and space. Int J Appl Earth Obs Geoinf. 2022 Feb;106:102649. doi: 10.1016/j.jag.2021.102649. PMCID: PMC8666325.

Huy, Le Duc, Nhi Thi Hong Nguyen, Phan Thanh Phuc, and Chung-Chien Huang. 2022. "The Effects of Non-Pharmaceutical Interventions on COVID-19 Epidemic Growth Rate during Pre- and Post-Vaccination Period in Asian Countries" International Journal of Environmental Research and Public Health 19, no. 3: 1139. <https://doi.org/10.3390/ijerph19031139>.

James Bergstra, Olivier Breuleux, Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume Desjardins, Joseph Turian, David Warde-Farley, Yoshua Bengio. Theano: A CPU and GPU Math Compiler in Python. PROC. OF THE 9th PYTHON IN SCIENCE CONF. (SCIPY 2010)

Kamineni, Meghana and Engø-Monsen, Kenth and Midtbø, Jørgen E and Forland, Frode and de Blasio, Birgitte Freiesleben and Frigessi, Arnoldo and Engebretsen, Solveig, Effects of non-compulsory and mandatory COVID-19 interventions on travel distance and time away from home, Norway, 2021 Eurosurveillance, 28, 2200382 (2023), <https://doi.org/10.2807/1560-7917.ES.2023.28.17.2200382>

Liu, Y., Yu, Q., Wen, H. et al. What matters: non-pharmaceutical interventions for COVID-19 in Europe. Antimicrob Resist Infect Control 11, 3 (2022). <https://doi.org/10.1186/s13756-021-01039-x>

Lawson, Andrew, and Chawarat Rotejanaprasert. 2023. "Bayesian Spatio-Temporal Prediction and Counterfactual Generation: An Application in Non-Pharmaceutical Interventions in COVID-19" Viruses 15, no. 2: 325. <https://doi.org/10.3390/v15020325>

Lionello, Lorenzo and Stranges, Debora and Karki, Tommi and Wiltshire, Emma and Proietti, Chiara and Annunziato, Alessandro and Jansa, Josep and Severi, Ettore and on behalf of the ECDC–JRC Response Measures Database working group, Non-pharmaceutical interventions in response to the COVID-19 pandemic in 30 European countries: the ECDC–JRC Response Measures Database Eurosurveillance, 27, 2101190 (2022), <https://doi.org/10.2807/1560-7917.ES.2022.27.41.2101190>

Matthew D. Hoffman, Andrew Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research 15 (2014) 1593-1623.

Michael Lingzhi Li, Hamza Tazi Bouardi, Omar Skali Lami, Thomas A. Trikalinos, Nikolaos Trichakis, Dimitris Bertsimas (2022) Forecasting COVID-19 and Analyzing the Effect of Government Interventions. Operations Research 71(1):184-201. <https://doi.org/10.1287/opre.2022.2306>

Mozaffer, F., Cherian, P., Krishna, S., Wahl, B., Menon, G.  2023. EFFECT OF HYBRID IMMUNITY, SCHOOL REOPENING, AND THE OMICRON VARIANT ON TRAJECTORY OF COVID-19 EPIDEMIC IN INDIA: A MODELLING STUDY, International Journal of Infectious Diseases. <https://doi.org/10.1016/j.ijid.2023.04.062>

Nguyen, M.H., Nguyen, T.H.T., Molenberghs, G. et al. The impact of national and international travel on spatio-temporal transmission of SARS-CoV-2 in Belgium in 2021. BMC Infect Dis 23, 428 (2023). <https://doi.org/10.1186/s12879-023-08368-9>

Pozo-Martin, F., Beltran Sanchez, M.A., Müller, S.A. et al. Comparative effectiveness of contact tracing interventions in the context of the COVID-19 pandemic: a systematic review. Eur J Epidemiol 38, 243–266 (2023). <https://doi.org/10.1007/s10654-023-00963-z>

Pung, R., Clapham, H.E., Russell, T.W. et al. Relative role of border restrictions, case finding and contact tracing in controlling SARS-CoV-2 in the presence of undetected transmission: a mathematical modelling study. BMC Med 21, 97 (2023).<https://doi.org/10.1186/s12916-023-02802-0>

Stokes, J., Turner, A.J., Anselmi, L. *et al.* The relative effects of non-pharmaceutical interventions on wave one Covid-19 mortality: natural experiment in 130 countries. *BMC Public Health***22**, 1113 (2022). <https://doi.org/10.1186/s12889-022-13546-6>

Salvatier J., Wiecki T., Fonnesbeck C. Probabilistic programming in Python using pymc3. PeerJ Comput. Sci., 2 (2016), Article e55

Stephen P. Brooks & Andrew Gelman (1998) General Methods for Monitoring Convergence of Iterative Simulations, Journal of Computational and Graphical Statistics, 7:4, 434-455, DOI: 10.1080/10618600.1998.10474787

Wiemken, T.L., Khan, F., Puzniak, L. et al. Seasonal trends in COVID-19 cases, hospitalizations, and mortality in the United States and Europe. Sci Rep 13, 3886 (2023). <https://doi.org/10.1038/s41598-023-31057-1>

1. The calculation of the point estimate of reproduction number for a given day *i* is equal as the simple quotient of the number of new cases for that day divided by the number of new cases 7 days *before* ([statista](https://de.statista.com/" \o "https://de.statista.com/)) . In our study however, we use the number of new cases *n* days *after* the day *i* as described in equation 3 to replace a forward looking sight instead of the conventional backward looking sight. Furthermore, we generalize the 7 days’ default to the arbitrary parameter *n*, which is set to 7, 14 and 21 in this paper. [↑](#footnote-ref-1)
2. https://www.ecdc.europa.eu/en [↑](#footnote-ref-2)
3. Thereby we decided to confine the time scope of the study regrading some countries to the dates mentioned in table 2 e.g. just until 2021 data, even though more data for a wider range were also available. The reason is that in the case of Hungary and Iceland, within the recent phases of the pandemic, only data on 1 or 2 days per week are non-zero numbers followed by numbers far away from zero, which might generate outlier ranges of reproduction rates of infections, corresponded to some NPIs. In the case of Liechtenstein, reported numbers of infections until October 2020 are multiple blocks of zero together with rare non zero numbers. [↑](#footnote-ref-3)
4. Before incorporating the data regarding the number of new reported cases over all countries (containing 28729 rows), the data table of new cases is cleaned especially in two respects. First, in 1275 rows the reported zero daily cases of positive infections are modified by us to be 1. This is done in order to enable the program to compute the monthly reproduction rates without being troubled through a number divided by zero. Second, in 93 rows, the reported number of positive cases are missing (NaN) values or negative values. We replaced the missing values by using the interpolate function in Python. [↑](#footnote-ref-4)
5. A detailed explanation of NPIs is presented in Lorenzo et. al. (2022). A variable dictionary related to the data on non-pharmaceutical interventions NPIs can be downloaded from the [ecdc](https://www.ecdc.europa.eu/sites/default/files/documents/Variable_Dictionary_and_Disclaimer_non-pharmacautical_measures_v3.pdf) website. [↑](#footnote-ref-5)
6. Note that a parameter named Not\_sequenced is applied to those days where no information regarding the proportion of various virus mutants in a country were available. Ina addition, there exists a virus variant named ‘Others’ in the list of virus variants, which represents a collection of other not labeled virus variants in the data. But these 2 types are not included in our study. [↑](#footnote-ref-6)
7. We use the [fitter](https://github.com/cokelaer/fitter) package in python to explore aggregate and individual data across the studied countries. [↑](#footnote-ref-7)
8. To provide a minimum level of convergence in our model parameters, we set the sampling size by default as 1000 in our study. [↑](#footnote-ref-8)
9. E.g. pred\_Germany+ in figure 4 [↑](#footnote-ref-9)
10. E.g. pred\_Germany- in figure 4 [↑](#footnote-ref-10)
11. Due to relative long time needed to be available to the MCMC sampling process the studied factors in all Countries, we run the PYMC3 model training on a virtual machine on the Linux server at the university of . [↑](#footnote-ref-11)
12. The max and min values of s differ in the range of maximal 2 percent from the depicted mean values and are represented in supplementary material. [↑](#footnote-ref-12)
13. Note that when analyzing the days where a completely implemented NPI, *X* has been active (positive signed days) and where the *X* has not been active (negative signed days), we subtracted the days containing the *partially* *X* from the negative signed days, so in order to constitute a meaningful contrast between the positive signed and the negative signed days. In the same way, when analyzing the days where a partially implemented NPI, *X* has been active (positive signed days) and where the *X* has not been active (negative signed days), we subtracted the days containing the completely activation of *X* from the negative signed days. [↑](#footnote-ref-13)
14. A glossary for definition of abbreviated NPIs is provided by the European center for disease prevention and control and is added to the supplementary material of this paper. [↑](#footnote-ref-14)