A country-wise comparative analysis of the role of containment policies, vaccination strategies, and virus variants in the COVID-19 pandemic in Europe

Simone Rancati, Giovanna Nicora, Enea Parimbelli, Marco Salemi, Riccardo Bellazzi, Daniele Pala

Submitted to: JMIR Public Health and Surveillance
on: November 13, 2024

Table of Contents

A country-wise comparative analysis of the role of containment policies, vaccination strategies, and virus variants in the COVID-19 pandemic in Europe

Simone Rancati[1]; Giovanna Nicora[1]; Enea Parimbelli[1]; Marco Salemi[2]; Riccardo Bellazzi[1]; Daniele Pala[1]

[1]Department of Computer, Electrical and Biomedical Engineering University of Pavia Pavia IT
[2]Department of Pathology, Immunology and Laboratory Medicine College of Medicine University of Florida Gainesville US

Corresponding Author:
Daniele Pala
Department of Computer, Electrical and Biomedical Engineering
University of Pavia
Via Ferrata 5
Pavia
IT

Abstract

Background: The COVID-19 pandemic has been a catastrophic event caused by the SARS-CoV-2 virus, which has endangered the entire world population for about three years, upsetting several aspects of public health, economics, and society, with some effects still visible today. The virus rapid spread was initially countered with containment policies aimed at reducing personal interactions and increasing health resources. Recent research on which policies were the most effective to optimize the prevention strategies and improve preparedness for possible future epidemics showed, although with some common findings, that the relationship between containment policies and epidemiological indicators can vary depending on the geographic area considered.

Objective: In this study, we aim to assess the complex interplay between containment policies, epidemiological variables, vaccines and variants throughout the entire pandemic, providing a framework that enlightens the complex dynamics that characterized the link between epidemiology and public health in several European countries. This framework can also serve as a tool for public health policy management in case of future epidemics.

Methods: To analyze the complex interplay between containment policies, epidemiological variables, vaccines and variants, we used a vector autoregressive mixed-effect model on data collected throughout the entire pandemic time span in nine European countries: Spain, France, The Netherlands, Latvia, Slovenia, Greece, Ireland, Cyprus, Estonia. We apply this model on weekly data to create a global description of the pandemic in each one of the three years of its duration, providing both a global and a country-wise analyses.

Results: The interplay between the epidemiological situation and the policy stringency changed significantly over the course of the pandemic. The number of hospitalizations was the most associated variable for policy making at first, but it gradually left the spotlight to the number of new cases as vaccinations increased and new variants, characterized by higher infectivity but lower hospitalization rates, emerged. All vaccine doses are associated with a reduction of new cases and hospitalizations. However, we found only the second vaccine dose to be associated with policy stringency; finally, no virus variant was found associated with policy stringency. Small differences could be noticed across different countries.

Conclusions: Through the application of our model we extract insights on the complex inter-dependencies that were generated across different variables during the COVID-19 pandemic. We provide a framework that can be applied to improve the policy making process in case of future epidemics, considering both the global and the local dimension.

(JMIR Preprints 13/11/2024:68774)
DOI: https://doi.org/10.2196/preprints.68774

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

A country-wise comparative analysis of the role of containment policies, vaccination strategies, and virus variants in the COVID-19 pandemic in Europe

Simone Rancati[1], Giovanna Nicora[1], Enea Parimbelli[1], Marco Salemi[2,3], Riccardo Bellazzi[1], Daniele Pala*[1]

[1] Department of Computer, Electrical and Biomedical Engineering, University of Pavia, Pavia, Italy.
[2] Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA
[3] Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL, USA

13. novembar 2024.

Abstract

Background

The COVID-19 pandemic has been a catastrophic event caused by the SARS-CoV-2 virus, which has endangered the entire world population for about three years, upsetting several aspects of public health, economics, and society, with some effects still visible today. The virus rapid spread was initially countered with containment policies aimed at reducing personal interactions and increasing health resources. Recent research on which policies were the most effective to optimize the prevention strategies and improve prepa- redness for possible future epidemics showed, although with some common findings, that the relationship between containment policies and epidemiological indicators can vary depending on the geographic area considered.

Objectives

In this study, we aim to assess the complex interplay between containment policies, epidemiological variables, vaccines and variants throughout the entire pandemic, providing a

framework that enlightens the complex dynamics that characterized the link between epidemiology and public health in several European countries. This framework can also serve as a tool for public health policy management in case of future epidemics.

Methods

To analyze the complex interplay between containment policies, epidemiological variables, vaccines and variants, we used a vector autoregressive mixed-effect model on data collected throughout the entire pandemic time span in nine European countries: Spain, France, The Netherlands, Latvia, Slovenia, Greece, Ireland, Cyprus, Estonia. We apply this model on weekly data to create a global description of the pandemic in each one of the three years of its duration, providing both a global

and a country-wise analyses.

*Correspondence to daniele.pala@unipv.it

Results

The interplay between the epidemiological situation and the policy stringency changed significantly over the course of the pandemic. The number of hospitalizations was the most associated variable for policy making at first, but it gradually left the spotlight to the number of new cases as vaccinations increased and new variants, characterized by higher infectivity but lower hospitalization rates, emerged. All vaccine doses are associated with a reduction of new cases and hospitalizations. However, we found only the second vaccine dose to be associated with policy stringency; finally, no virus variant was found associated with policy stringency. Small differences could be noticed across different countries.

Conclusions

Through the application of our model we extract insights on the complex inter-dependencies that were generated across different variables during the COVID-19 pandemic. We provide a framework that can be applied to improve the policy making process in case of future epidemics, considering both the global and the local dimension.

1 Introduction

The COVID-19 pandemic has been one of the most significant global events of the 21st century, as it caused a series of disruptions in public health, economics and politics that affected the entire planet with effects still visible after more than one year. The pandemic started in December 2019 and was declared finished in May 2023 [1]. In this period, COVID-19 caused 760.79 million registered infections and 6.90 million deaths in the World, with an average of 71,916 cases and 1,498 death every week [2], [3]. It was caused by the Sars-CoV-2 virus, a new pathogen belonging to the Coronavirus family that possess the ability to spread very quickly and lead, in the initial phases of the pandemic, to extremely worrying hospitalization and mortality rates due to the inability of the human immune system to recognize and properly fight the virus [4]. The SARS- CoV-2 virus is characterized with a high mutation rate, leading to the emergence of different variants such as Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), or the currently circulating Omicron (B.1.1.529) and their (sub)lineages like XBB.1.5. This situation represented a public health emergency, with the first few outbreaks quickly escalating to the overwhelming of most healthcare systems, as countries did not have the hospital beds and resources to treat an unprecedented number of people with cardio-respiratory complications [5]. In order to prevent further spreading of the disease, national and local administrations started emanating containment policies that had the aim of reducing the probability of contacts within the population, for example mandating social distance and use of face masks, prohibiting leisure travel, canceling public events, closing food services. These policies had inevitable socioeconomic side effects and a tremendous impact also on the populations' mental health [6]. This was probably partly due to the fact that the initial response to the spreading of the virus was unorganized, as the novelty of the emergency made it impossible to create data-driven responses, therefore the public health interventions were not optimized to face the pandemic. Several studies aimed at optimizing containment policies, but the validity of the results was constantly undermined by the fast changes in the epidemiological situation, continuously occurring with the introduction of new variables such as the vaccination rates and the new variants of the virus. The PERISCOPE project, funded by the European Commission, created a multi-disciplinary consortium and undertake research on the COVID-19 pandemic in order to help policy makers in managing the current pandemic and be prepared for the new ones. One of the main aims of PERISCOPE was the creation a European COVID-19 Atlas [7], summarizing data and statistics about the epidemiological, economic, and socio-political indicators during the pandemic. Other scopes of the project included the development of descriptive and predictive models to assess the relationships between different

aspects of the pandemic impact, including the interplay between containment policies and epidemiological variables.

This study presents an overall description of the various phases of the pandemic in several European countries, with a particular focus on the reaction of those countries in terms of containment policies and their association with the epidemiological outcomes. Our aim is to provide a global overview of how the public health policy makers reacted to the spreading of the virus, both at the initial stage of the pandemic and in the subsequent phases, characterized by the introduction of large vaccination campaigns and the birth of new,

often more infectious, variants. We focused on eight countries, that were selected based on the completeness and availability of their data: Cyprus, Estonia, France, Netherlands, Latvia, Ireland, Slovenia and Spain. We used a multilevel vector auto-regressive model (mlVAR), considering each country as a member of a population, in order to take into account both the general relationships among the different variables in the entire population and the local effects for each individual countries.

1.1 Previous Work

This work has been conceived as the continuation of a research that started during the first year of the pandemic as part of the PERISCOPE project, and that has been published in 2023 in the journal Scientific Reports [8]. In this paper, we used an mlVAR model to study the two-way relationship between non-pharmaceutical interventions and epidemiological manifestations of COVID-19 in nine selected countries in Europe. The results of this study provided important information on the complex interplay between health policies, assessed using the CoronaNet database [9], and epidemiological surveillance, but also showed how the effectiveness of specific policies was rather country-specific, although with some common ground given by the fact that policies aimed at increasing health monitoring and resources were the most effective in all countries. The study, however, was performed only on 2020 data, as data about the vaccination rates and the advent of new variants were not yet available in sufficient quantity to assess their mutual influence with the governments' policies and the consequent impact on infection, hospitalization and death rates of the disease. The novel study described in this paper presents an extension of the analysis that includes newer data from the years 2021 and 2022, when the introduction of vaccinations and new variants forced the policy makers to continuously change their strategies. With these new results, we are able to provide better insights both on the mechanisms that influenced the strengthening or release of containment measures, and vice versa on those that led from public health policies to a change in the epidemiological situation of the pandemic,
taking into account all the different variables that characterized each phase.

2 Methods

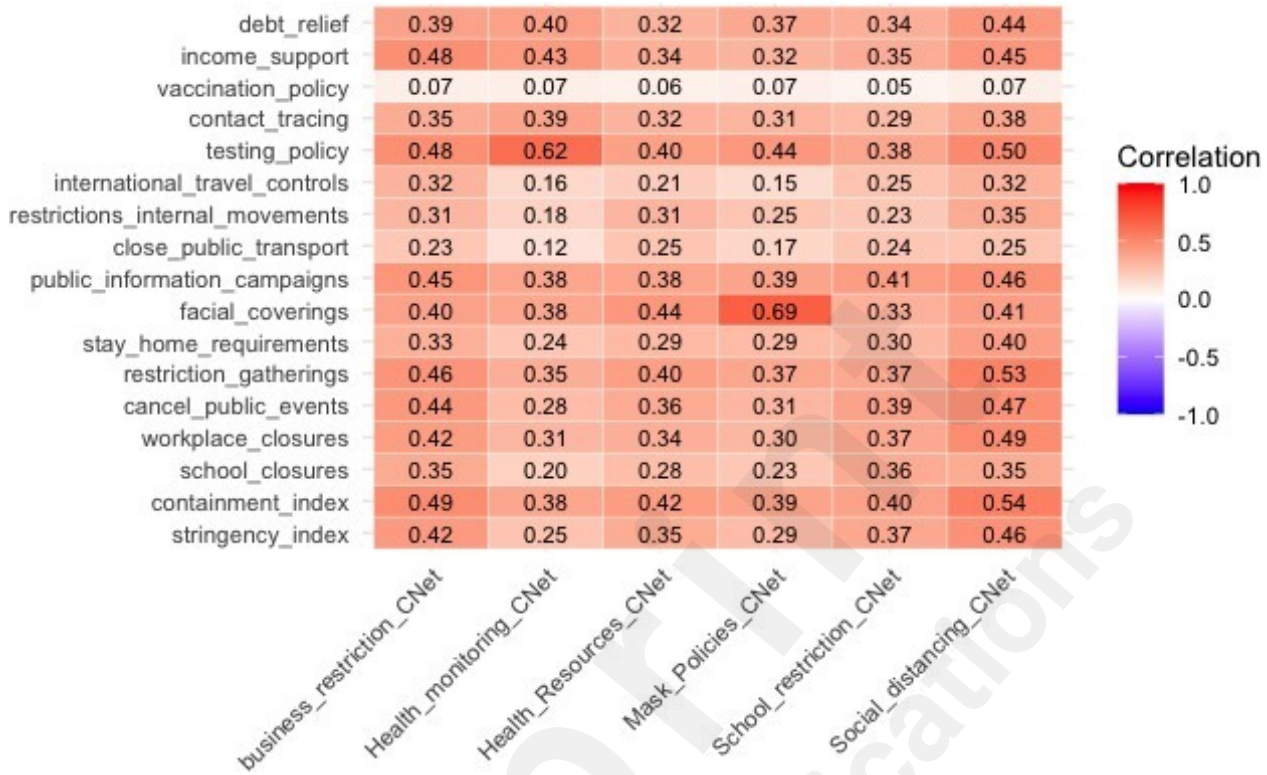
2.1 Data Sources and Preprocessing

Three main data sources were used for our analysis: OxCGRT [10], ECDC [11] and GISAID [12]. Data preprocessing was performed using both RStudio and PyCharm. The first preprocessing step, which is common across all datasets, involved applying a filter to retain only the data generated between January 2020 and December 2022. Additionally, each dataset was split based on the country of interest: Spain, France, the Netherlands, Latvia, Slovenia, Greece, Ireland, Cyprus, and Estonia.

2.1.1 Oxford COVID-19 Government Response Tracker

The analysis is based on the Stringency Index (SI), a composite measure derived from the Oxford COVID- 19 Government Response Tracker (OxCGRT) dataset [10]. This index is calculated starting from nine key metrics: school closures, workplace closures, cancellation of public events, restrictions on public gatherings, public transport closures, stay-at-home requirements, public information campaigns, restrictions on internal movements, and international travel limitations. The SI is computed as the average score of these nine metrics, each ranging from 0 to 100, with a higher score indicating a stricter government response. When policies differ at the subnational level, the index reflects the strictest region. It is important to note that the SI measures the

strictness of policies starting from January 2020, but does not assess their effectiveness or appropriateness, so a higher score does not necessarily imply a better response. In addition to the SI, other measures, such as the Policy Intensity Indices (PIIs) [9], have been developed, notably by Kubinec in the CoronaNet Project (CNet) and used in prior research, including our previous study [7]. The choice to use the SI from OxCGRT instead of the PIIs in this analysis is driven by the availability of the SI for the years 2022 and 2023, unlike the PIIs. Furthermore, the SI consolidates all relevant metrics into a single measure, offering ease of interpretation when analyzing results, whereas the PIIs are characterized by six distinct indices (Social Distancing, School Restrictions, Business Restrictions, Health Monitoring, Health Resources, and Mask Policies) without a single combined index. An exploratory analysis was conducted



Slika 1: Correlation Matrix between OxCGRT Metrics and CoronaNet Project Policy Intensity Indices (PIIs). This heatmap displays the Pearson correlation coefficients between various OxCGRT metrics, including the Stringency Index, and the six Policy Intensity Indices (PIIs) developed in the CoronaNet Project (CNet). The Stringency Index shows moderate correlations with all PIIs, particularly with the Business Restrictions Index (0.42), Social Distancing Index (0.46), Health Monitoring Index (0.25), Mask Policies Index (0.29) and School Restriction Index (0.37).

to assess the relationship between OxCGRT indices, including the SI, and the PIIs. Figure 1 presents the Pearson correlation between the OxCGRT indices and the PIIs, with a particular focus on the correlation between the Stringency Index (SI) and all PIIs. The results demonstrate that the SI serves as an effective summary measure, capturing the overall trends represented by the individual PIIs.

2.1.2 European Centre for Disease Prevention and Control

Our second key data source for our analysis is the European Centre for Disease Prevention and Control (ECDC) repository [11]. The ECDC has been monitoring the progression of COVID-19 across Europe since the initial outbreaks in 2020. It provides weekly reports that include various metrics such as the number of cases, case-to-test ratios, cumulative cases, incidence rates, hospitalizations, and mortality rates, both for the entire European region and for individual countries and vaccination data. Each report details the changes observed in the

previous week for these parameters and summarizes the current situation in each country. While the SI is recorded daily, ECDC data are available on a weekly basis. As a result, for our integrated dataset, we chose calendar weeks as the unit of analysis, aggregating the daily values of the Stringency Indices into 53 weeks for 2020, 52 weeks for 2021, and 52 for 2022. Using weekly data helps mitigate irregularities present in daily records, which are prone to errors and fluctuations due to factors such as inconsistent testing rates and methodologies. However, most of the 29 countries analyzed lacked complete data on hospital admissions and ICU bed occupancy. The selection of the nine nations included in this research was based on prior analyses, which showed that these countries had minimal missing data for the ECDC epidemiological and genetic variables relevant to our study. For our analysis, out of the ECDC

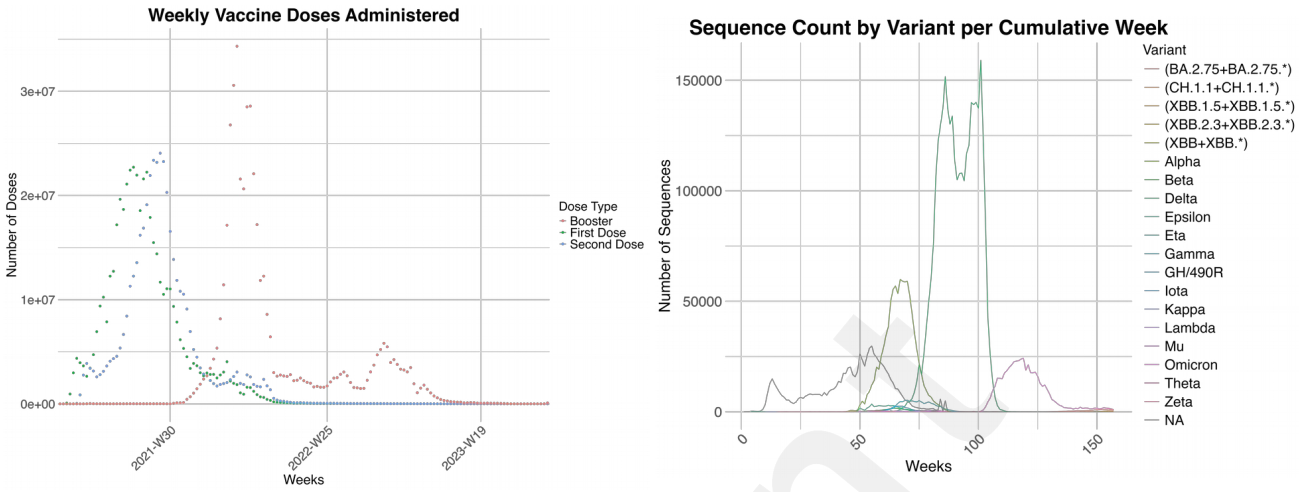
datasets' wide range of features, we selected the following key variables: New weekly case incidence (New cases), Weekly hospital admissions (admHosp), Weekly ICU admissions (admICU), Weekly death incidence (Deaths), First dose vaccination (FirstDose), Second dose vaccination (SecondDose), and three additional doses (DoseAdditional1, DoseAdditional2, DoseAdditional3). The additional doses were aggregated into a single variable, named Booster. Each variable was then rescaled by dividing it by the population of the respective country and normalized per 100,000 individuals to allow for meaningful comparisons. For missing data, the imputation method involved replacing missing values (NA) with the last valid entry from the same column, as long as the previous value was neither missing nor zero. If no valid previous value existed, or if the last value was zero, the missing value was replaced with zero. This approach ensures the dataset is complete, reducing the potential bias from incomplete data. At the end of ECDC Preprocessing, we produced nine datasets for each year, each corresponding to a specific country of interest.

2.1.3 GISAID

In addition the epidemiological surveillance variables, we included also the role that variants such as Alpha, Delta, and Omicron played in the pandemic and their impact on the SI. To this end, we conducted an analysis of frequency of Spike protein sequences in our third key dataset: GISAID, a global initiative that provides open access to genomic data of influenza and the Sars-CoV-2 Viruses. The Spike protein is a key component of a virus, responsible for allowing it to enter human cells [3]. This protein mutates rapidly, leading to the emergence of new variants with varying degrees of infectivity and severity [13]. During the pandemic, various SARS-CoV-2 variants emerged as a result of the high mutability of the virus's Spike protein. As shown in Figure 2B, the original Wuhan strain (NA) was the predominant and unique variant at the beginning of 2020. However, as the virus continued to evolve, new variants began to emerge over time. The Alpha variant (lineage B.1.1.7) was first identified in the United Kingdom in September 2020 [14], followed by the Beta variant (lineage B.1.351) in South Africa in May 2020 [15], and the Gamma variant (lineage P.1) originating in Brazil in November 2020 [16]. The Delta variant (lineage B.1.617.2) was first reported in India in October 2020 [**?**], while the Omicron variant (lineage B.1.1.529) was detected in South Africa on November 24, 2021, and has since become predominant across Europe [3]. Many other variants (Epsilon, Zeta, Kappa etc.) have been identified in the course of the pandemic, but they had less impact. The Omicron variant, including its sublineages such as XBB.1.5 [17], is currently the most widespread variant globally (Figure 2B).

For our analysis, we downloaded 16,187,950 Spike protein sequences from GISAID, spanning from De- cember 24, 2019, to November 8, 2023, along with their corresponding Pango Lineages. To enhance data quality, we implemented several filtering criteria. Firstly, we retained sequences only from specific countries and excluded any sequences that contained missing or unrecognized amino acids. Additionally, we selected sequences whose lengths were within ± 30 amino acids of the median length for their lineage, aligning with the expected Spike protein length of 1241 to 1301 amino acids [3] resulting a dataset with 4,506,983 Spike protein. Finally, we focused on sequences and metadata whose collection dates were included in the period of our analysis. Subsequent filtering was conducted to focus on sequences from the nation of interest in this research. For each pair of country and variant we calculated the variant frequency as a fraction of the total sequences uploaded to GISAID for the variant on a weekly basis (using calendar weeks) during the specified year. This process resulted in the creation of nine distinct datasets—one for each country for each year of the simulation—each containing the weekly percentage of observed variants. In 2020, the main COVID-19 variants circulating worldwide were Alpha, Beta, Delta, and Zeta. However, these variants did not appe- ar uniformly across all countries, due to bio-genetic differences or limited sequencing capacity. To ensure consistency in the data, any missing variant information for each year was filled with zero, maintaining the same set of features across the model. This method was also applied in 2021 and 2022. In 2021, the number of circulating variants increased as the virus adapted, with new variants such as Theta, Gamma, Epsilon, Eta, Kappa, Iota, Mu, Lambda,

GH490R, and Omicron joining Alpha, Beta, Delta, and Zeta. By 2022, the spread of vaccines and the virus's natural evolution reduced the number of dominant variants to Omicron, Delta, and Alpha. Despite this, the World Health Organization (WHO) defined more aggressive sublineages of Omicron as distinct variants, including BA.2.75, XBB, XBB.1.5, and CH.1.1 [18], [17].



(a) A (b) B

Slika 2: A)Number of sequences uploaded to the GISAID database for each COVID-19 variant from 2020 to 2022. Each colored line represents a different variant, showing the volume of protein sequences submitted over time.B)Weekly distribution of COVID-19 vaccine doses administered over time, categorized by first dose, second dose, and booster doses (third and fourth doses). Each colored dot corresponds to a different dose type, highlighting the trends in vaccine uptake and the progression of the vaccination campaign over time from 2020 to 2022.

2.2 Data Analysis

2.2.1 Correlation Analysis

To assess potential relationships between variables, the correlation between the SI and each epidemiologi- cal and genetic variables has been assessed for each year of the simulation using Pearson's linear correlation coefficient [8]. The results are presented in Table 1, and in tables S1, S2 in the supplementary material.

2.2.2 Population Modelling Approach

From our correlation analysis, we were able to discern the relationships between the SI index and various epidemiological and genetic variables. We implemented a multi-country, autoregressive predictive model for each simulation year[19], [8] on both policy and epidemiological data, as outlined in Equation 2, where I denotes the Stringency Index and t represents the weekly temporal granularity. Autoregressive models inherently require the incorporation of time lags to effectively capture the temporal dependencies within the data, allowing us to understand how past values influence current trends. To determine the optimal temporal granularity, we tested the model with various time lags and selected the best one based on the Akaike Information Criterion (AIC) values, whose formula is shown in Equation 1,where k indicates the number of parameters estimated by the model instead L is the maximum likelihood of the model.

$$\mathrm{AIC} = 2k - 2\ln(L) \tag{1}$$

In particular, for each year of simulation, we used Multilevel Vector Autoregression (mlVAR) to create an integrated model that could be applied across all the included countries. This approach is a statistical mode- ling technique that focuses on a population of related individuals, where the

variability between individuals is represented by unique model parameters for each one of them. More in detail, we treated each of the nine selected European countries as an individual within a larger population. Each country operates under its own model, which is a slight variation of a central population model depicted through Fixed Effect (FE) models. In this framework, fixed effects represent the parameters consistently applicable to all the countries. Equation 2 shows an example of the fixed effects calculation with the stringency index as dependent varia- ble. The complete model is made of the same framework applied to all other variables (See Supplementary

Equation 1 for the complete model).

$$\begin{aligned}\mathbf{I}(t) = \beta_{01} + \beta_{11} \cdot I(t-1) + \beta_{21} \cdot NewCases(t-1) + \beta_{31} \cdot admHosp(t-1) + \beta_{41} \cdot admICU\,(t-1) \\ +\beta_{51} \cdot death(t-1) + \beta_{61} \cdot FirstDose(t-1) + \beta_{71} \cdot SecondDose(t-1) + \beta_{81} \cdot BoosterDose(t-1) \\ +\beta_{91} \cdot Alpha(t-1) + \beta_{101} \cdot Beta(t-1) + \beta_{111} \cdot Delta(t-1) + \beta_{121} \cdot Zeta(t-1),\end{aligned} \quad (2)$$

Conversely, Random Effects (RE) are specific to each country and supplement the common population model. The core concept here is that the model parameters for each country derive from a common Gaussian distribution, where the mean is the estimated average for the population, and the standard deviation is calculated based on contributions from each country (Eq. 3).

$$\beta_{ij} \sim N\,(\beta_{j,pop}, \sigma_{\beta_{j,pop}}) \quad (3)$$

Eq. 2, also fits the definition of a multivariate autoregressive AR model, where all the variables at time *t* only depend from the ones at the previous timestamp *t* 1. In statistics an AR model is a representation of a random process in which the output variable depends linearly on its own previous values plus a stochastic term (an imperfectly predictable term); thus the model is in the form of a stochastic difference equation (or recurrence relation).

The multilevel vector autoregression has been implemented using *RStudio*, with the dedicated *R* packages

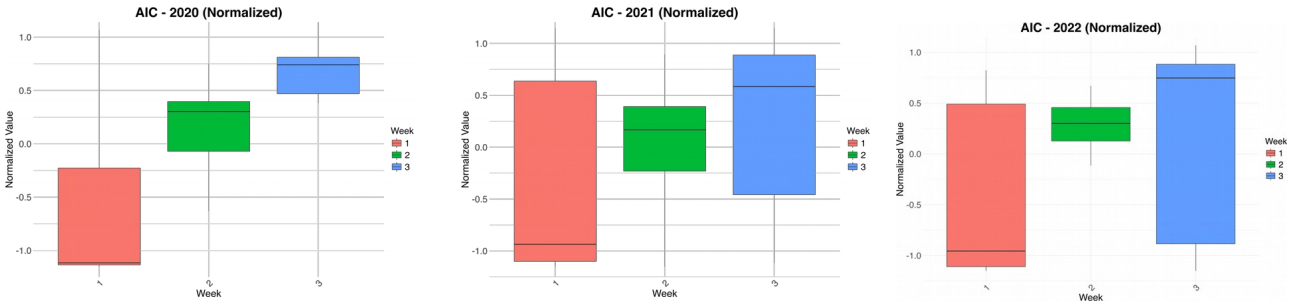
mlVAR and *lme4*.

The input for the model parameter estimation phase consists of a data table that includes SI, four selected epidemiological parameters (*new cases (nC), hospital admissions (aH), ICU admissions (aICU), deaths (d)*), three vaccination variables (*First Dose (fD), Second Dose (sD), Booster Dose (bD)*), and a column detailing the frequency of each variant for the simulation year. Specifically, for 2020, the variant columns include four (*Alfa (Al), Beta (Be), Delta (De), Zeta (Ze)*); for 2021, there are fourteen (*Alpha (Al), Beta (Be), Delta (De), Zeta (Ze), Theta (Th), Gamma (Ga), Epsilon (Ep), Eta (Et), Kappa (Ka), Iota (Io), Mu (Mu), Lambda (La), GH490R (GH), Omicron (Om)*); and for 2022, there are eight (*Delta (De), GH490R (GH), Omicron (Om), Alpha (Al), BA.2.75, CH.1.1, XBB, XBB.1.5*). Each table also includes a final co- lumn specifying the ID of the country. Thus, for each year of simulation, nine datasets are generated, one for each country of interest.

The output of the model is the estimate of the regression parameters of the autoregressive model. Equation 4 represents an example for SI in 2020 where the conditional part of the formula refers to the Random Effects: the contribute is conditional to the ID of the country.

$$\begin{aligned}I(t) \sim\ & I(t-1) + NewCases(t-1) + admHosp(t-1) + admICU\,(t-1) + death(t-1) + \\ & newCases(t) + admHosp(t) + admICU\,(t) + death(t) + \\ & FirstDose(t-1) + SecondDose(t-1) + BoosterDose(t-1) + \\ & Alpha(t-1) + Beta(t-1) + Delta(t-1) + Zeta(t-1) + \\ & NewCases(t) + admHosp(t) + admICU\,(t) + death(t) + \\ & FirstDose(t) + SecondDose(t) + BoosterDose(t) + \\ & Alpha(t) + Beta(t) + Delta(t) + Zeta(t) + \\ & ((1|\boldsymbol{ID}) + (\boldsymbol{I}(t-1)|\boldsymbol{ID}) + (NewCases(t-1)|\boldsymbol{ID}) + \\ & (admHosp(t-1)|\boldsymbol{ID}) + (admICU\,(t-1)|\boldsymbol{ID}) + (death(t-1)| \\ & \boldsymbol{ID}) + (\boldsymbol{FirstDose}(t-1)|\boldsymbol{ID}) + (SecondDose(t-1)|\boldsymbol{ID}) + \\ & (\boldsymbol{BoosterDose}(t-1)|\boldsymbol{ID}) + (Alpha(t-1)|\boldsymbol{ID}) + \end{aligned} \quad (\boldsymbol{Be}\ ta(t - 1)|$$

$$ID) + (Delta(t-1)|ID) + (Zeta(t-1)|ID)) \quad (4)$$



Slika 3: Normalized AIC for the model applied to all three years with a lag of one, two and three weeks. It can be seen that the model performs better with a lag of one week in 2020 (Kruskall-Wallis P-Value = 0.008). No significant differences are visible for 2021 and 2022, although the AIC is generally still lower with a one-week delay.

3 Results

3.1 Temporal Granularity Selection

Like most viral infections, COVID-19 has an incubation time between infection and symptoms manife- station. Several studies have demonstrated that the incubation time of this disease during the first waves was higher compared to most influenza viruses, ranging from 2 to 14 days [20], however the incubation time decreased with the advent of new variants in 2021 and 2022 [21]. One of the consequences of the presence of an incubation time is a time lag between the infections and the effects on the epidemiological variables, and also a delay in observing the effects of a containment policy after its establishment. Furthermore, hospita- lization rates may present further delay, as hospitalization generally occurs several days after infection. For these reasons, it is necessary to determine the best time lag for the mlVAR analysis, in order to optimize its effectiveness. To this aim, we ran the mlVAR model for all variables using lags of one, two and three weeks, and calculated the AIC for each one of the models. Results are reported in Figure 3 and show that the AIC is generally lower for a one-week delay, although the difference appears statistically significant only in 2020. Therefore, we decided to keep a lag of one week for all three years for continuity in the analysis both of this paper and the previous work presented in [7].

3.2 Fixed Effects

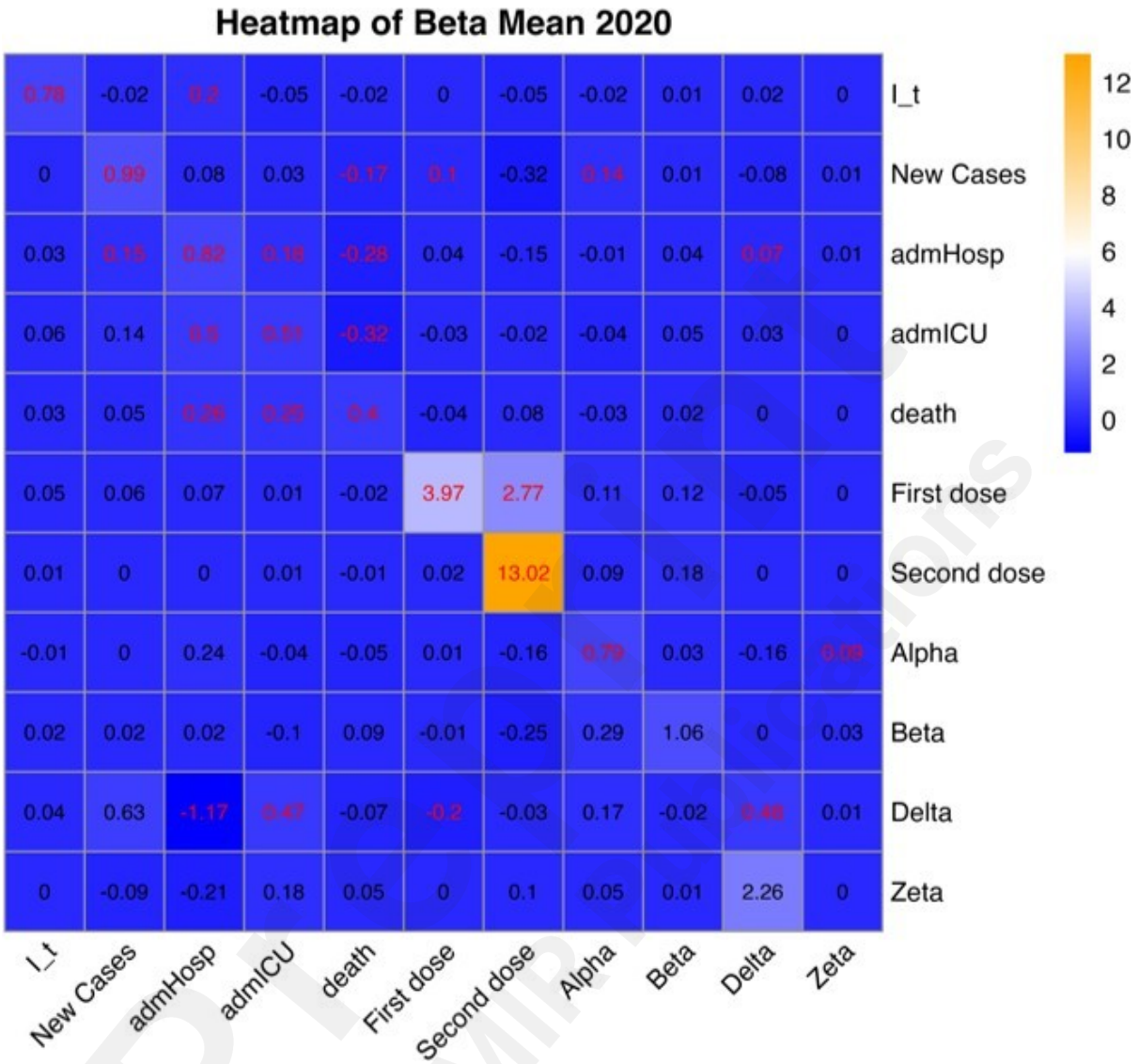
Fixed effects in the mlVAR model represent the relations that each variable has with all the other variables considering the entire population, i.e. all countries. We performed the analysis for each year of pandemic, obtaining different fixed effects, as the conditions changed in time with the introduction of vaccinations and new variants. Detailed results for each year are reported in the following sections.

3.2.1 2020

When the virus spread in Europe for the first time in 2020, vaccinations were not available, and the original strain and the Alpha variant were causing most of the infections. Towards the end of the

year, new variants arrived and the first vaccinations were made. Figure 4 shows a heatmap of the Beta coefficients for all the relations between all combinations of variables and a time lag of one year.

From this figure, looking at the significant relations, it can be noticed that the Stringency index in each week of 2020 strongly depends on its own value in the previous week, and is also positively influenced by the rate of hospitalizations. The epidemiological indicators (cases, hospitalizations and deaths) also have have some influence on each other in time. Vaccination data also shows some significant coefficients, however their relevance is probably low as vaccinations started at the end of the year, therefore the data points are not sufficient to show a significant trend.



Slika 4: Heatmap of the fixed effects' coefficients for the year 2020. The significant ones (P-Value $<$ 0.05) are highlighted in red. The rows indicate the variables at time t, i.e. the outcomes, whereas the columns indicate the variables at time $t - 1$, i.e. the covariates.

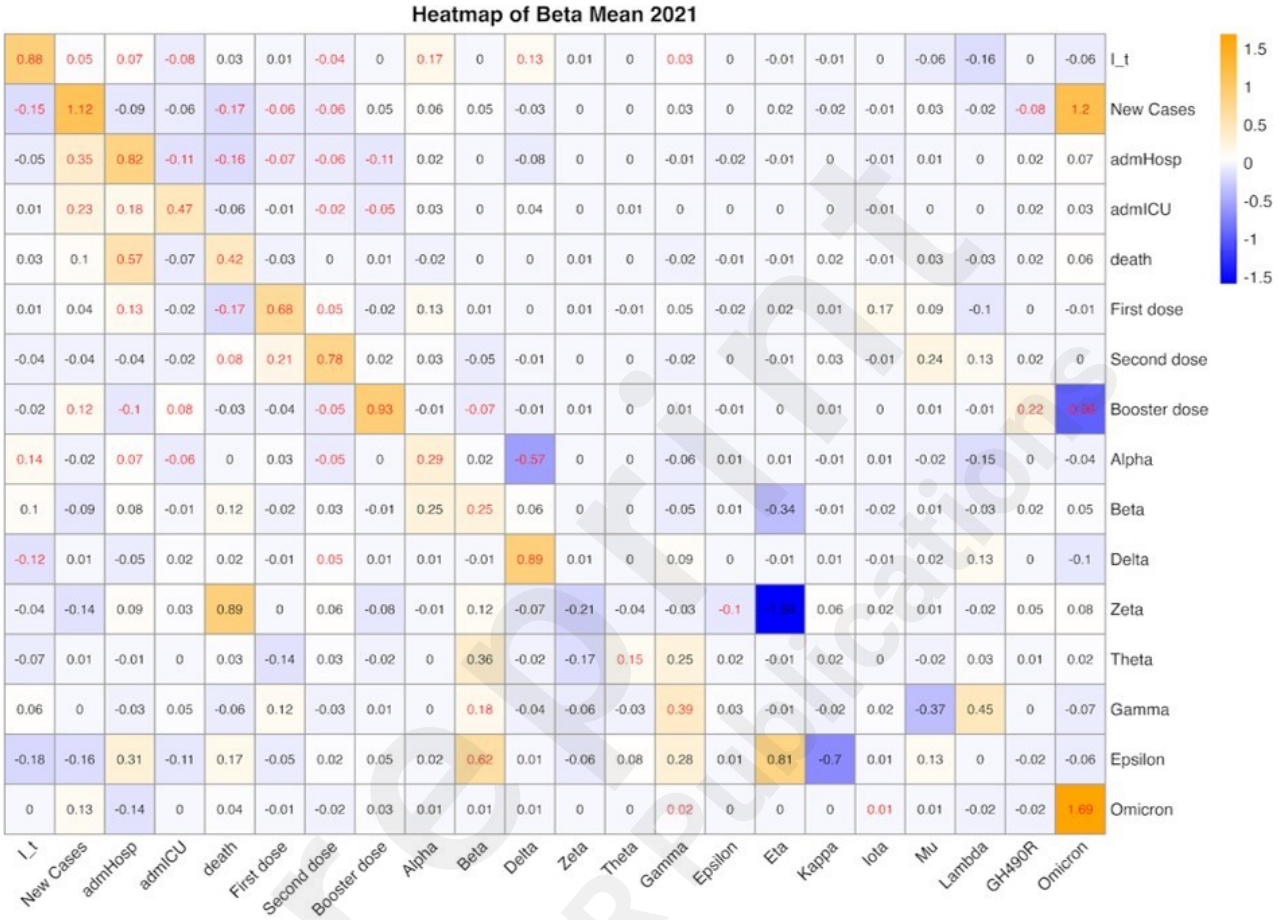
3.2.2 2021

The central year of the pandemic was the most eventful one, as several waves of infections and new variants came in a period in which vaccinations were becoming gradually more diffuse. Figure 5 shows the heatmap of the mlVAR fixed effects for the year 2021, considering all new variants together with the stringency index, the vaccination data including the booster dose, and all the epidemiological variables. Looking at the significant coefficients, highlighted in red, several interesting relations can be noticed. First of all, the stringency index appears to increase following an increase in new cases and hospitalizations, excluding the ICU ones. A decrease can be seen following an increase of second doses of vaccine. Vaccinations also have an impact on epidemiological variables, as all hospitalizations decrease with an increase of any vaccination dose. Also, positive significant relations are present between first doses and hospitalizations, second doses

and deaths, and boosters and ICU admissions.

Some interesting relationships can be noticed also concerning the variants: the Alpha variant, prevalent between the end of 2020 and the beginning of 2021, is negatively influenced by the increase of the Delta

variant, that rapidly became prevalent in 2021, replacing Alpha in a relatively short time. Plus, the Omicron variant seems to be related to a notable increase in new cases, with a positive, significant and relatively high coefficient.



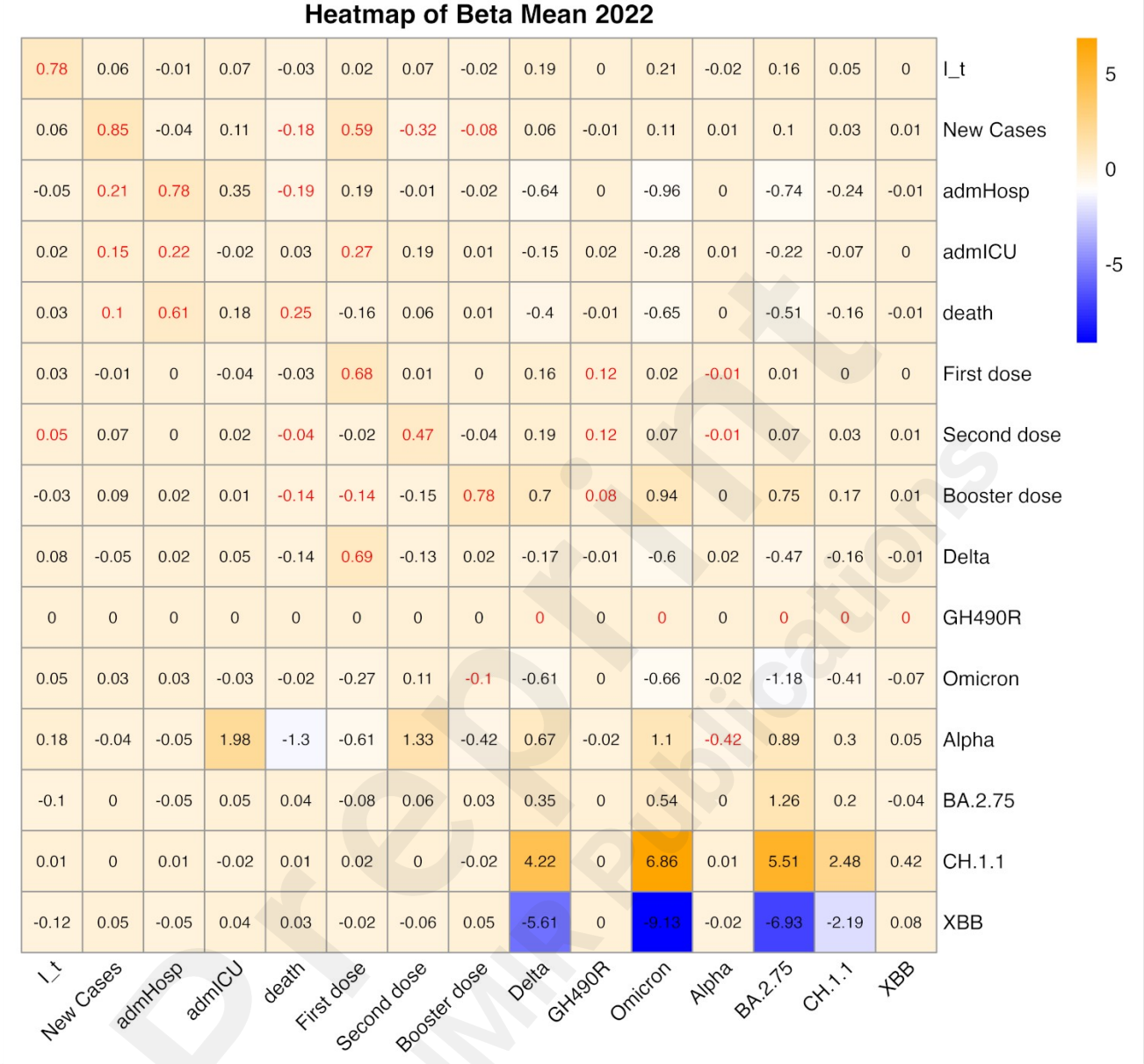
Slika 5: Heatmap of the fixed effects' coefficients for the year 2021. The significant ones (P-Value $<$ 0.05) are highlighted in red. The rows indicate the variables at time t, i.e. the outcomes, whereas the columns indicate the variables at time $t-1$, i.e. the covariates.

3.2.3 2022

Figure 6 shows the fixed effects of mlVAR applied to the year 2022, i.e. the final full year of pandemic. In this time frame, restrictions around the world were gradually lifted. Thanks to the increasing immunity of the population due to the combined effects of vaccinations and previous exposures, and to the diffusion of the Omicron variant that seemed to be less severe although highly contagious [22], hospitalizations and deaths showed a decreasing trend. As a consequence, the stringency index in 2022 appears not to be related with any epidemiological variable. Some temporal dependencies are still present for the epidemiological variables, and some significant coefficients can be noticed also concerning the variants, however considering that the variant variables depend on the percentage of genetic sequences identified, we do not believe that they represent meaningful results.

3.3 Random Effects

While fixed effects describe the general relations between variables in the entire population, random effects describe the individual variations of the model outcome for each of the countries and allow to take into account

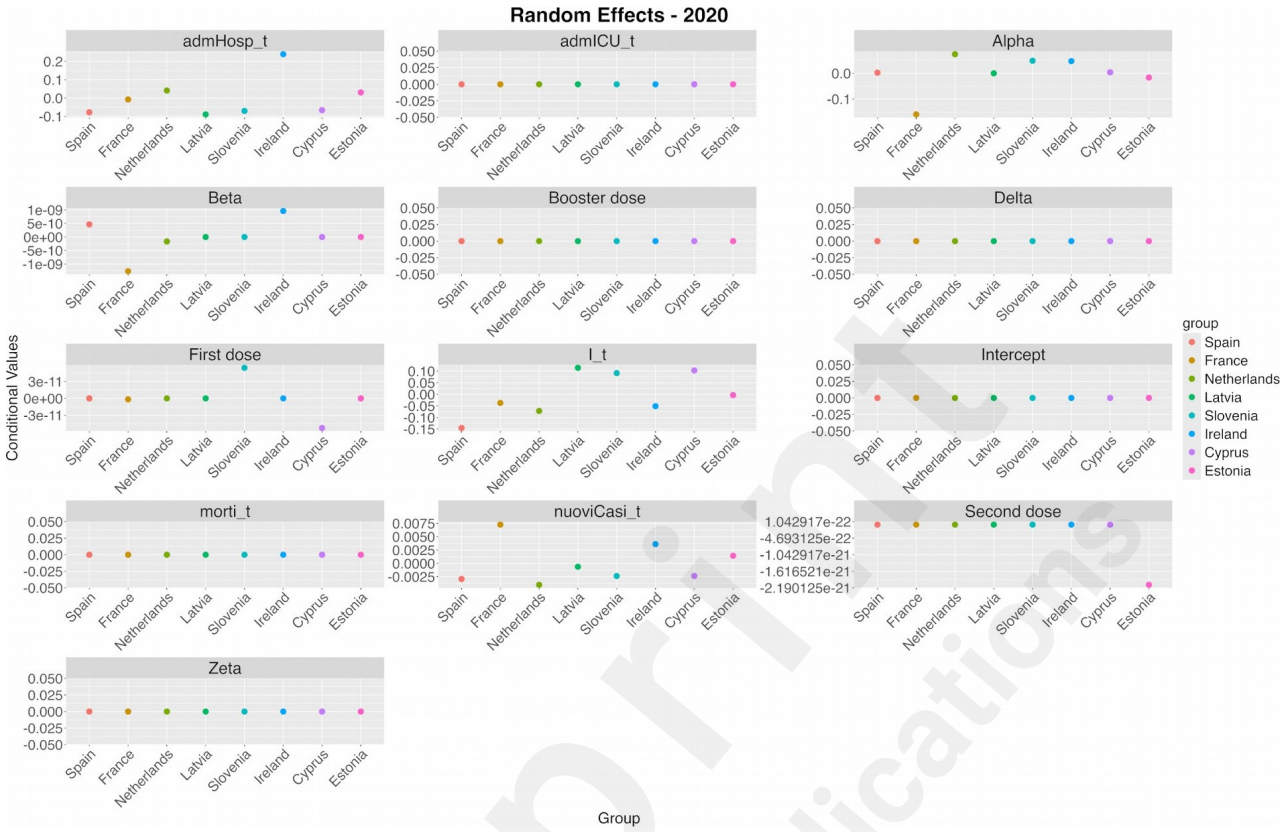


Slika 6: Heatmap of the fixed effects' coefficients for the year 2022. The significant ones (P-Value $<$ 0.05) are highlighted in red. The rows indicate the variables at time t, i.e. the outcomes, whereas the columns indicate the variables at time $t - 1$, i.e. the covariates.

the variability due to each country's own socio-demographic features. Figure 7 shows the random effects of the 2020 model for each variable, considering the stringency index as the dependent variable. Random effects of the 2021 and 2022 models are reported in figures A.1 and A.2 in the appendix section. Some interesting phenomena can be seen analyzing these figures together with the fixed effects results. For example, according to the fixed effects in 2020 the stringency index was significantly influenced only by its previous value and the increase of admissions to hospitals, however the random effects show that this influence was higher in some countries such as Ireland, but slightly lower in Spain, Slovenia, Latvia and Cyprus.

In 2021, on the contrary, the effect of hospitalizations on the index is constant for all countries, whereas some other variables show higher variation, for example new cases seem to have a greater impact on the stringency index in the Netherlands than in Spain or Ireland. Some country-specific

differences are present also in 2022, however the only significant fixed effect is the one related to the previous stringency index, which appears to be the highest in Latvia, Slovenia and Cyprus and the lowest in Spain.



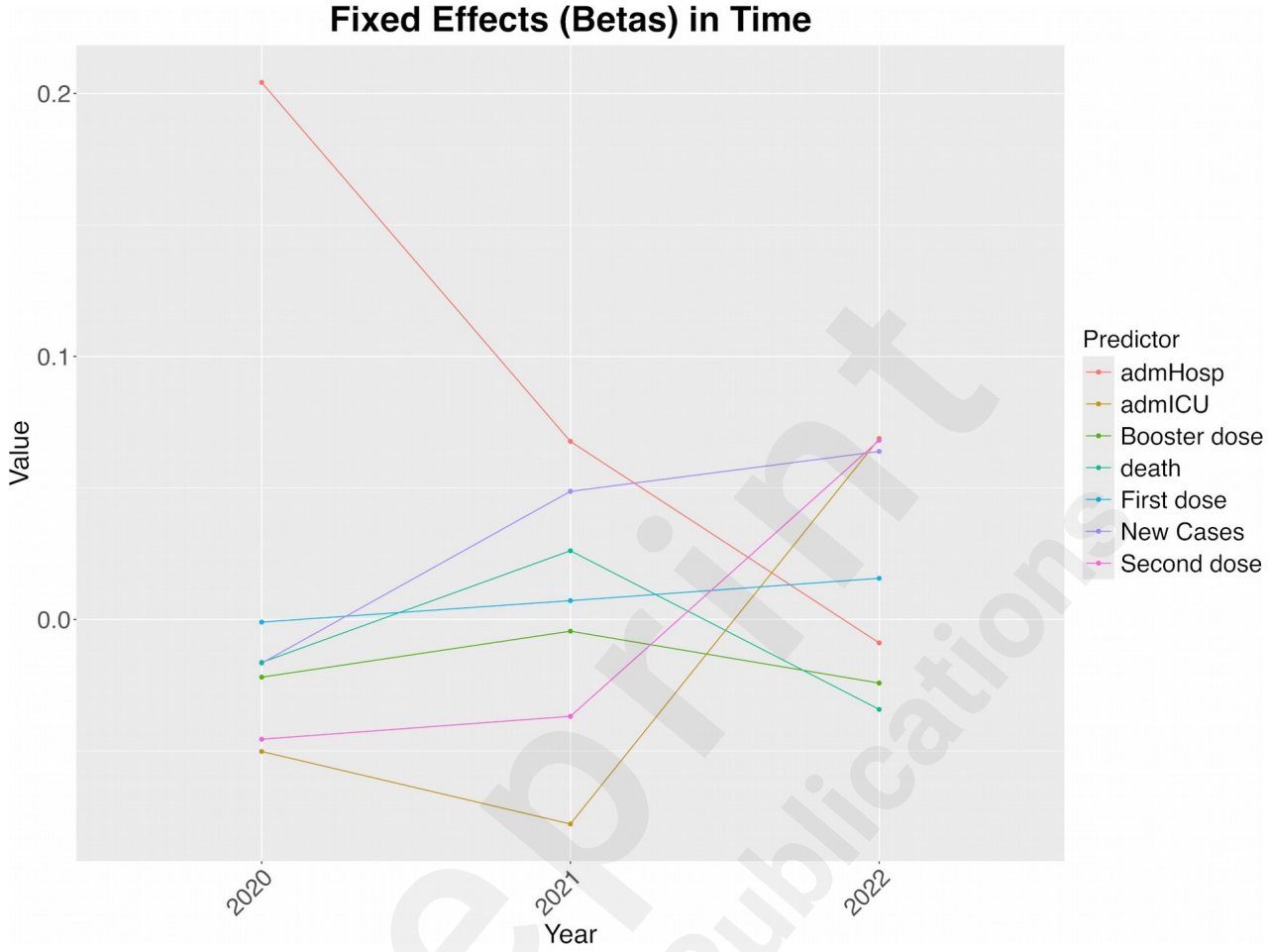
Slika 7: Random effects (country-wise variation on the model intercept) for all predictors with stringency index as outcome in 2020.

3.4 Temporal progression of the coefficients

A useful representation of how the general impact of the epidemiological variables on the policies, and vice versa, varied in all the considered countries can be obtained plotting the value progressions of the fixed effects coefficients in the three years of pandemic. Figure 8 shows this variation setting the stringency index as dependent variable. Observing this figure, it can be noticed that hospital admissions is the variable with the highest influence on the stringency index in the first two years of pandemic, especially in the first phase in 2020, when the deadliest variants were causing most of the infections and vaccines were not yet available. In 2021, with the increasing vaccinated population and new variants characterized by high infection power but relatively lower hospitalization rates, hospitalization rates tend to have less influence on the creation of stringent policies, whereas new cases gain importance. Another interesting observation can be made looking at the Seecond Dose ˇcoefficient, that is negative in 2020 and 2021, meaning that in general the increase in second dose administrations is followed by a relaxation of the policies. The coefficient increases in 2022, however from the heatmap it can be seen that this number is not significant, as the number of people who had received the second dose in 2022 is high and constant.

Considering the stringency index as covariate instead of outcome, interesting inter-dependencies can be seen as well. Figure A.3 shows the betas variation in time setting the number of new cases as dependent variable. Interestingly, an increase in the stringency index does not appear to reduce new cases in 2020, as shown also in the heatmaps, but it seems to become more effective in 2021, whereas the 2022 value appears to be not significant. Interesting insights can be noticed looking at the vaccination data: vaccines appear to significantly reduce new cases only when the percentage of vaccinated population is relatively high, in fact it can be seen that the coefficients are negative and significant in 2021 for first and second dose, and in 2022 for the booster dose. The booster dose,

however, significantly reduced hospitalizations also in 2021, as shown in Figure 5. It should be noted that the relatively high coefficient for the ”first dose” variable in 2022 is not reliable, as in 2022 the rate of population who had received the first dose was high and constant.



Slika 8: Variations of the fixed effects coefficients indicating the effects of the epidemiological variables on the stringency index in the three considered years.

4 Discussion

Being caused by a highly contagious and previously unknown virus, the COVID-19 pandemic quickly became a world crisis that lasted for three years. Finding themselves suddenly burdened by a high load of hospitalizations, governments and local political entities decided to establish containment policies aiming at reducing social gatherings, national and international travel, and business activities that could be the ground for new outbreaks. As some of these policies might have been designed as a reaction to the current public health situation, rather than as the result of a data-driven analysis, it is important to conduct such analysis. Identifying the best policies will help creating unified protocols, aiming at facing future pandemic threats via optimized interventions, and maximize effectiveness while limiting the negative social and economic impacts. Previous studies [23], including ours [7], focus only on the initial phase of the pandemic, and they do not take into account the presence of different virus variants, nor how the vaccinations influenced the choice of the policies. Past studies generally showed that the policy effectiveness has been rather heterogeneous depending on the geographic regions considered, although, not surprisingly, the most effective ones on average seem to be those related to increasing health monitoring and resources. In this work, we aimed at providing a framework to analyze the entire course of the pandemic, with a series of models showing the delicate dependencies among the epidemiological

and sociopolitical aspects of the pandemic, considering that those two aspects are enclosed in a mutual exchange of causal factors and are characterized by a spatial dimension that works on different levels. To the best of our knowledge, this is the first work to provide a complete analysis that includes both the population-level effects and the individual ones for the entire duration of the

pandemic, which includes several transition phases where health policies were combined with vaccination strategies to contrast the virus, and at the same time were countered by the advent of new variants.

The general results of our analysis show notable differences in the variable relations in time, but a certain agreement among the different countries considered. In particular, it seems that in 2020, when the virus was new and most lethal, the hospitalizations rate was the indicator more associated with stringent policies, which were probably a reaction to the increasing pressure on the health systems. This happened more in Ireland than in other countries, whereas the impact was relatively lower in Slovenia, Latvia, Spain and Cyprus. Increasing hospitalization rates were followed by more and more stringent countermeasures. However, the stringency increase did not lead to a significant decrease in hospitalizations after one week, however this relation would be different considering a larger time frame. In 2021, with a combination of increased health monitoring that led to faster case reports, and new variants with a higher infection power but less serious symptoms, ”new cases” became the epidemiological variable that is most associated with higher policy stringency. In this case, the coefficient is the lowest for Ireland and Spain, whereas the relation is higher in the other countries, especially the Netherlands. Differently from 2020, the reverse relation is significant, i.e. an increase in stringent policies is associated with a significant decrease in new cases. This is not true for 2020, although we would argue that new cases were not properly tracked, and this might have strongly contributed to the lack of statistical significance. All vaccination doses are significantly associated with a reduction of cases and hospitalizations when a larger part of population is vaccinated (first and second dose are most effective in 2021, second and booster dose in 2022). However, only an increase of second doses in 2021 is associated with a decrease in policy stringency in all countries. Concerning the sequenced variants, the only variant that is significantly associated with policy making was Delta in 2021, as a higher number of sequences is followed by higher stringency. The rising of the Omicron variant is related to an impressive spike in new cases in 2022, but an increase in hospitalizations or policy stringency does not seem to follow in any of the considered countries.

Arguably, this study presents several limitations, mainly consequential of the implicit limitations of the epidemiological data, i.e., the collected data does not necessarily represent the real situation. Collected data depend on the quantity and frequency of tests and their collection methods, as well as variant sequencing that were performed by a specific country. Even though the hierarchical nature of the model is a point of strength, as it considers both the entire set of countries and the individual ones, the geographical model granularity is still very low, as local regions and provinces could have separate epidemiological and sociopolitical situations. This could lead also to incidental correlations, some of which are clearly observed in this study (e.g. first doses and new cases in 2022). Despite these limitations, this study shows a thorough analysis of the interplay between virus spread, restriction policies, vaccination strategies, and viral variants. The model has the capacity to provide a multi-level, complete analysis including both the population-level (all countries) effects and the individual-level ones (a single country). This tool is easily scalable to other geographical entities and can be potentially applied to new settings, in view of optimizing the policy making process in preparation to new epidemics.

5 Conclusion

The COVID-19 pandemic has been the largest public health crisis of the last century so far, with govern- ments, countries, and public health authorities establishing a variety of containment policies in the attempt of reducing the virus circulation. In a time span of three years, these policies came to a complex interplay with the emergence of new virus variants, a continuous change of the epidemiological landscape, and an increase of vaccination coverage. In this study, we propose a framework to assess such complex interplay, using a vector autoregressive mixed-effect model on

data collected throughout the entire pandemic time span in eight European countries: Spain, France, Netherlands, Latvia, Slovenia, Greece, Ireland, Cyprus, Estonia. Results show that the mutual association of the epidemiological variables and the policies' stringency varied significantly over the course of the pandemic, with policy stingency strongest association with the increase in hospitalizations during the first year, followed by the number of new cases during the following years, with the increase of vaccinations and the advent of new variants characterized by higher infectivity but lower hospitalization rates. All vaccine doses seemed to be effective in reducing new cases and hospitalizations, although the second dose was the only one that significantly associates with a reduction of policy stringen-

cy. Finally, new variants had a string association with several epidemiological aspects, but not with policy stringency. On small differences can be noticed across different countries. Although this approach has some limitations due to the nature of the collected data and the low geographical granularity, this model provides deeper insights on the complex inter-dependencies that were generated across different entities during the COVID-19 pandemic, and represents an easily generalizable framework that can be applied to improve the policy making process in case of future epidemics, considering both the global and the local dimensions.

6 Acknowledgments

6.1 Funding

This study has been carried out as part of the PERISCOPE Project, funded by the European Commission under Grant Agreement N° 101016233, and has been funded in part by the NIH grant NIAID R01 AI170187.

6.2 Authors' Contribution

All authors significantly contributed to the design of this study and the creation of the manuscript. SR, DP, EP and GN designed the study, SR and DP wrote the first draft, that was revised by EP, GN, MS and RB. SR performed the analysis under the supervision of DP. Funding acquisition was carried out by MS and RB. All authors read and approved the manuscript in its current version.

7 Conflicts of Interest

All authors have no conflicts of interest to declare.

8 Availability of Data and Material

All data and codes used for the analysis reported in this manuscript is available in the following GitHub: COVID19-Europe-Policy-Vaccine-Variants-Analysis

The GISAID datasets used in this research are available at the following links: GISAID Dataset

References

[1] World Health Organization. WHO Homepage; 2024. [Accessed: 7-October-2024]. Available from: https://www.who.int.

[2] Mathieu E, Ritchie H, Rod´es-Guirao L, et al.. Coronavirus Pandemic (COVID-19); 2020. [Accessed: 7-October-2024]. Available from: https://ourworldindata.org/coronavirus.

[3] Rancati S, Nicora G, Prosperi M, Bellazzi R, Marini S, Salemi M. Forecasting dominance of SARS-CoV-2 lineages by anomaly detection using deep AutoEncoders. bioRxiv. 2023 Oct:2023.10.24.563721.

[4] Schultze JL, Aschenbrenner AC. COVID-19 and the human innate immune system. Cell. 2021 Apr;184(7):1671-92.

[5] Akande OW, Akande TM. COVID-19 pandemic: A global health burden. Niger Postgrad Med J. 2020 Jul;27(3):147-55.

[6] Ahmed N, Barnett P, Greenburgh A, Pemovska T, Stefanidou T, Lyons N, et al. Mental health in Europe during the COVID-19 pandemic: a systematic review. Lancet Psychiatry. 2023 Jul;10(7):537-56.

[7] Pala D, Parimbelli E, Larizza C, Cheng C, Ottaviano M, Pogliaghi A, et al. A new interactive tool to visualize and analyze COVID-19 data: The PERISCOPE Atlas. Int J Environ Res Public Health. 2022 Jul;19(15):9136.

[8] Podrecca C, Parimbelli E, Pala D, Cheng C, Messerschmidt L, Bu¨the T, et al. A comparative analysis of the effects of containment policies on the epidemiological manifestation of the COVID-19 pandemic across nine European countries. Sci Rep. 2023 Jul;13(1):11631.

[9] CoronaNet Research Project;. Accessed: 2024-10-04. https://www.coronanet-project.org/.

[10] OXFORD COVID-19 Governemnt Response Tracker;. Accessed: 2024-10-04. https://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker.

[11] European Centre for Disease Prevention and Control;. Accessed: 2024-10-04. https://www.ecdc.europa.eu/en/about-ecdc.

[12] Khare S, et al. GISAID's Role in Pandemic Response. China CDC Weekly. 2021;3(49):1049-51.

[13] Solis O, Beccari AR, Iaconis D, Talarico C, Ruiz-Bedoya CA, Nwachukwu JC, et al. The SARS- CoV-2 spike protein binds and modulates estrogen receptors. bioRxiv. 2022 May:2022.05.21.492920. Update in: Sci Adv. 2022 Dec 2;8(48):eadd4150. doi: 10.1126/sciadv.add4150. PMID: 35665018; PMCID: PMC9164441.

[14] Kumar N, Quadri S, AlAwadhi AI, AlQahtani M. COVID-19 Recovery Patterns Across Alpha (B.1.1.7) and Delta (B.1.617.2) Variants of SARS-CoV-2. Front Immunol. 2022 Feb;13:812606.

[15] Xu K, Gao P, Liu S, Lu S, Lei W, Zheng T, et al. Protective prototype-Beta and Delta-Omicron chimeric RBD-dimer vaccines against SARS-CoV-2. Cell. 2022 Jun;185(13):2265-78.e14. Epub 2022 Apr 27.

[16] Fumagalli MJ, Castro-Jorge LA, Fraga-Silva TFC, de Azevedo PO, Capato CF, Rattis BAC, et al. Pro- tective Immunity against Gamma and Zeta Variants after Inactivated SARS-CoV-2 Virus Immunization. Viruses. 2021 Dec;13(12):2440.

[17] Hu Y, Zou J, Kurhade C, Deng X, Chang HC, Kim DK, et al. Less neutralization evasion of SARS- CoV-2 BA.2.86 than XBB sublineages and CH.1.1. Emerg Microbes Infect. 2023 Dec;12(2):2271089. Epub 2023 Oct 26.

[18] Wang Q, Li Z, Guo Y, Mellis IA, Iketani S, Liu M, et al. Evolving antibody evasion and receptor affinity of the Omicron BA.2.75 sublineage of SARS-CoV-2. iScience. 2023 Oct;26(11):108254.

[19] Ferrari D, Violante T, Novelli M, Starlinger PP, Smoot RL, Reisenauer JS, et al. The death of laparoscopy. Surg Endosc. 2024 May;38(5):2677-88. Epub 2024 Mar 22.

[20] Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of Coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. Ann Intern Med. 2020 May;172(9):577-82.

[21] Park SW, Sun K, Abbott S, Sender R, Bar-On YM, Weitz JS, et al. Inferring the differences in incubation-period and generation-interval distributions of the Delta and Omicron variants of SARS-CoV-2. Proc Natl Acad Sci U S A. 2023 May;120(22):e2221887120.

[22] Arabi M, Al-Najjar Y, Mhaimeed N, Salameh MA, Paul P, AlAnni J, et al. Severity of the Omicron SARS-CoV-2 variant compared with the previous lineages: A systematic review. J Cell Mol Med. 2023 Jun;27(11):1443-64.

[23] Antonini M, Fouda A, Hinwood M, Melia A, Paolucci F. The interplay between global health policy and vaccination strategies in the shift towards COVID-19 endemicity. Health Policy and Technology. 2024;13(1):100854. The interplay between global health policy and vaccination strategies in the shift towards COVID-19 endemicity. Available from: https://www.sciencedirect.com/science/article/pii/S2211883724000170.

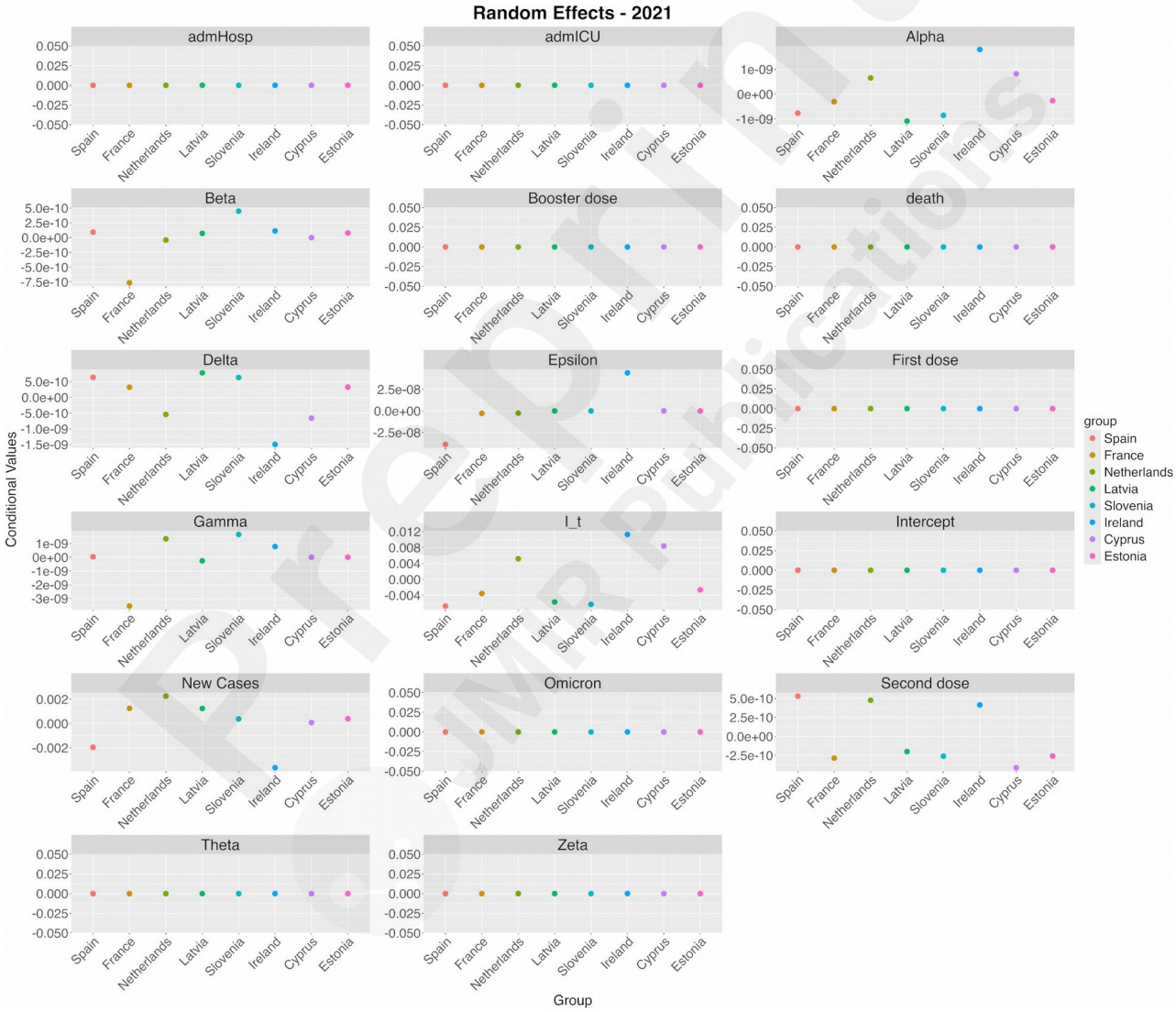
	New Cases	admHosp	admICU	death	First Dose	Second Dose	Booster	Alpha	Beta	Delta	Zeta
SI	**0.33**	**0.41**	**0.42**	**0.48**	**0.11**	7e-5	0.08	**0.27**	0.07	0.08	0.03

Tabela 1: The contemporaneous bivariate correlations between each index and the weekly

epidemiological and genetic variables for 2020 were analyzed across all countries. Significant results ($p < 0.05$) are highlighted in bold. Correlations were calculated by combining data from all available countries into a single vector for each variable.

9 Tables

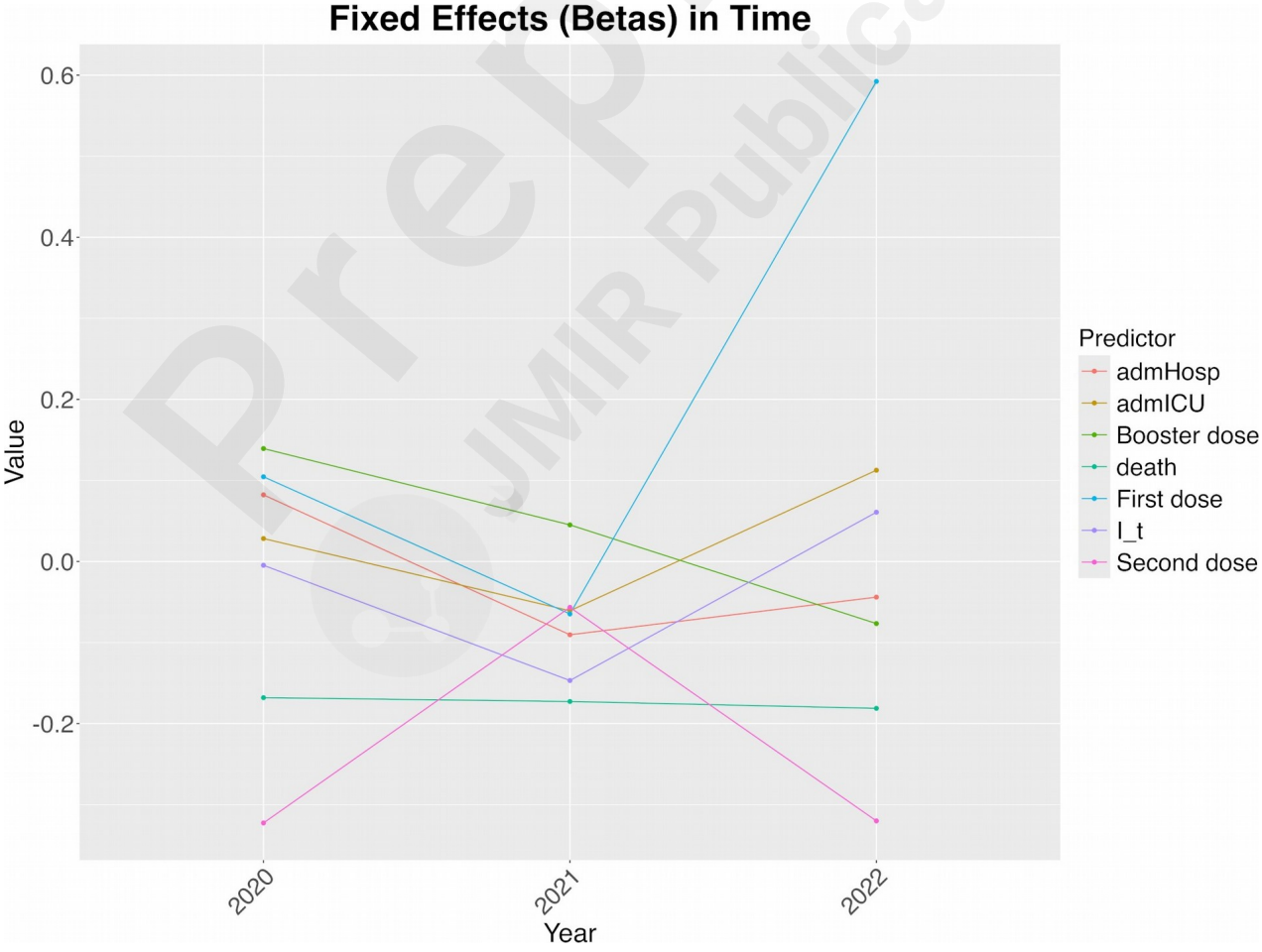
A Appendix



Slika A.1: Random effects (country-wise variation on the model intercept) for all predictors with stringency index as outcome in 2021.

Slika A.2: Random effects (country-wise variation on the model intercept) for all predictors with

stringency index as outcome in 2022.



Slika A.3: Variations of the fixed effects coefficients indicating the effects on new cases of the other

epidemi- ological variables and the stringency index in the three considered years.

Supplementary Files