

# Bayesian time-varying autoregressive models of COVID-19 epidemics

Paolo Giudici<sup>1</sup>  | Barbara Tarantino<sup>1</sup> | Arkaprava Roy<sup>2</sup>

<sup>1</sup>Department of Economics and Management, University of Pavia, Pavia, Italy

<sup>2</sup>Department of Biostatistics, University of Florida, Gainesville, FL, USA

## Correspondence

Paolo Giudici, Department of Economics and Management, University of Pavia, Pavia, PV 27100, Italy.  
Email: [giudici@unipv.it](mailto:giudici@unipv.it)

## Funding information

European Commission, Grant/Award Number: Periscope contract number 101016233



This article has earned an open data badge “**Reproducible Research**” for making publicly available the code necessary to reproduce the reported results. The results reported in this article could fully be reproduced.

## Abstract

The COVID-19 pandemic has highlighted the importance of reliable statistical models which, based on the available data, can provide accurate forecasts and impact analysis of alternative policy measures. Here we propose Bayesian time-dependent Poisson autoregressive models that include time-varying coefficients to estimate the effect of policy covariates on disease counts. The model is applied to the observed series of new positive cases in Italy and in the United States. The results suggest that our proposed models are capable of capturing nonlinear growth of disease counts. We also find that policy measures and, in particular, closure policies and the distribution of vaccines, lead to a significant reduction in disease counts in both countries.

## KEYWORDS

Bayesian models, B-splines, COVID-19 disease, NPI covariates, PARX models

## 1 | INTRODUCTION

Time series of observed counts arise in a variety of contexts in which a number of events are recorded over time, including studies of incidences of a particular disease (Davis, 2003; Zeger, 1988). The reported occurrences are often expressed with daily, weekly, or monthly frequency. Disease modeling has significant practical importance for governments since it provides information that can improve decisionmaking. In addition, studying the correlation between counts and policy variables allows to better gauge the impact of control measures. This paper provides a Bayesian analysis of a time series of counts with the goal of providing reliable short-term predictions of disease counts and to assess their correlation with explanatory variables. The time series of interest is the incidence of COVID-19 infectious disease in Italy and United States, two paradigmatic countries, and the explanatory variables are the varying Non-pharmaceutical interventions (NPIs) implemented by most countries to stop the virus propagation.

Since COVID-19 first emerged in early 2020, it has infected nearly every country and claimed the lives of several thousands of people worldwide. So far, many countries have been locked-down and stringent social distancing measures have

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Biometrical Journal* published by Wiley-VCH GmbH.

been implemented to contrast the high incidence of the infection over time and its high economic consequences. Previous studies, mostly focused on the first wave of the pandemic (before July 2020), have shown the effectiveness of NPIs in pandemic mitigation (Cowling et al., 2020; Hao et al., 2020). They have also shown that the impact of NPIs may differ among countries with various country characteristics, like health capacity, population density, air temperature, etc. (Esteve et al., 2020; Marcon et al., 2020). Existing studies also indicate that vaccination is the most promising strategy for overcoming the epidemic. However, the unequal distribution and allocation of vaccines among nations and demographic groups may obstruct the path to herd immunity. It is crucial to understand how vaccination combined with NPIs reduces COVID-19 transmission (Giordano et al., 2021).

The Poisson AutoRegressive (PAR) model is one of the most commonly used models for analyzing disease time series data, as it can well capture time dependence among observations. It can be seen as a special case of the generalized linear model (GLM) methodology (Fahrmeir et al., 2013; Kedem & Fokianos, 2002; Nelder & Wedderburn, 1972) where the link function is a linear function of the covariates, with coefficients that do not change over time. The advantage of this specification is the ease of model implementation and interpretation. However, when the impact of the epidemic changes over time, especially under the effect of policy measures, this type of model may be restrictive, and time-varying coefficients may be needed.

A flexible way to introduce time-varying parameters in a GLM model is through spline basis functions, as in the seminal work of Cleveland and Loader (1996). Spline basis functions have also been proposed to specify time-varying coefficients in AutoRegressive (AR) time series models (Amorim et al., 2008; Cai et al., 1999; Huang & Shen, 2004). For this type of models, Zeger (1988) has suggested a generalized estimating equation methodology, which still is the most commonly used estimation method. Within the Bayesian framework, Chan and Ledolter (1995) proposed a Monte Carlo expectation maximization (MCEM) technique. More recently, Hastie et al. (2001) compared different estimation procedures. Moving to the more complex autoregressive moving average (ARMA) models, the only work that considers time-varying parameters is that of Andrade et al. (2015).

None of these papers focuses on the time-varying character of the coefficients. We contribute to fill the gap with a model that extends Fokianos et al. (2008) within a time-varying coefficient framework. In doing so, we extend the time-dependent contagion models recently proposed by Agosto et al. (2021), Geir et al. (2022), and Roy and Karmakar (2021) with the inclusion of time-varying policy covariates. And we also extend the causal contagion models recently proposed by Bo et al. (2021) and Islam et al. (2020) with the inclusion of time-varying covariate effects, which also allows to compare alternative time lags of policy measures. Note also that our proposed model allows to account for overdispersion of counts, as in the Markov-switching Poisson GARCH model proposed by Roberts et al. (1998).

More specifically, We consider two classes of time series generalized linear models. The first is a class of AR models, in which the conditional mean of the Poisson process is linked to the past observations and to the potential covariate effects. The second one adds the dependence of the past values on the conditional mean. Since the mean of a Poisson model is the same as the variance, the latter can be seen as an Integer-Valued Generalized Autoregressive Conditional Heteroskedasticity (INGARCH) model. Originally introduced by Ferland et al. (2006), these models have been thoroughly analyzed in Zhu (2012a, 2012b, 2012c) and Roberts et al. (1998).

We thus compare alternative models along three dimensions: AR versus INGARCH structures; time-constant versus time-varying coefficient, and models with or without covariates while testing for different time lags.

The rest of the paper is organized as follows. Section 2 describes the proposed models, Section 3 describes the Bayesian specification and inference. Section 4 describes the data. Section 5 presents the main empirical findings. Section 6 summarizes the highlights of the paper and suggests future research directions.

## 2 | MODEL SPECIFICATION

The time dependence of COVID-19 transmission represents a critical factor to understand the epidemic dynamics. The reference Susceptible-infected-recovery (SIR) model assumes an exponential growth model, which assumes that the logarithm of the counts follows a linear trend, summarized by  $R_0$ , a function of the ratio between the new cases originating on consecutive days. However, the evolution of the epidemics is characterized by time heterogeneity, which leads to a nonlinear growth. This derives, for example, from time-varying testing and measurement procedures, among others.

From the previous consideration, it appears that the exponential growth assumption needs to be relaxed in favor of a time-dependent modeling framework. Recently, Agosto et al. (2021) proposed a PAR model for epidemic counts that can take into account both short-term and long-term dependence on past contagion shocks. The model is specified by a

Poisson conditional distribution for the COVID-19 counts, along with a logarithmic link function which allows to model the effect of covariates without constraining them to assume positive values, as in a linear link model.

While the model has been quite useful to monitor contagion in the first wave of the pandemic, it has become less performing during the following times since, even if there is a nonlinear growth assumption, it is based on coefficients that do not change over time and, therefore, cannot easily adapt to regime switches.

In this paper, we aim to extend the fixed coefficients PAR model of Agosto et al. (2021) into a time-varying coefficient framework, more flexible in model specification. Our modeling strategy is motivated to adapt to the rapid changes of the COVID spread. Hence, we propose a nonstationary time-varying model, which is suitable for our analysis. The constraints of the parameter space in Equations (7) and (18) are motivated by the stationary conditions for the time-constant case. This is fairly common in the time-varying AR and INGARCH model literature (Dahlhaus & Rao, 2006; Davis & Mikosch, 2009; Fryzlewicz et al., 2008; Ferreira et al., 2017; Rohan & Ramanathan, 2013).

In detail, we consider two alternative specifications: one based on INGARCH models for counts, equivalent to the PAR models and other one based on the more parsimonious AR models.

## 2.1 | AR models

The time series of counts  $Y_t$  reported at time (day)  $t$  is assumed to follow a Poisson distribution, conditional on the information up to  $t - 1$ :

$$Y_t | \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t), \quad (1)$$

where  $\mathcal{F}_{t-1}$  is the  $\sigma$ -field that contains all relevant information up to time  $t$ . Let  $X_t = (X_{t,1}, \dots, X_{t,r})$  be a time-varying  $r$ -dimensional covariate vector and let  $v_t = \log(\lambda_t)$ . A Poisson autoregressive model can then be specified, following Agosto et al. (2021), by the following link function:

$$v_t = \mu + \sum_{i=1}^p \alpha_i \log(1 + Y_{t-i}) + \eta^T X_t, \quad (2)$$

where  $\mu \in \mathbb{R}$ ,  $\alpha_i \in \mathbb{R}$  with  $i = 1, \dots, p$ , and  $\eta^T = (\eta_1, \dots, \eta_r) \in \mathbb{R}^r$  are unknown parameters to be estimated from the data. From an interpretational viewpoint,  $\mu$  is the intercept,  $\alpha_i$  describes the dependence of the expected count of cases on past observations and  $\eta^T$  describes a dependence of the expected counts on  $r$  explanatory covariates at time point  $t$ .

The above model, as discussed in Agosto et al. (2021), allows to take into account time dependency of COVID-19 counts, thereby reducing overdispersion. However, it remains the issue of a strong time heterogeneity of the policy measures taken by different countries, which can lead to an incorrect measurement of the effects of the covariates. To solve this issue, we propose to extend the AR model into a time-varying coefficient version, in which the link function for  $Y_t$  becomes:

$$v_t = \mu(t/T) + \sum_{i=1}^p \alpha_i(t/T) \log(1 + Y_{t-i}) + \eta^T(t/T) X_t, \quad (3)$$

where  $\alpha_i(t/T)$  and  $\eta^T(t/T)$  represent, for  $t = 1, \dots, T$ , time-varying splines. We recall that a spline is a continuous function that coincides with a polynomial on every subinterval in which it is defined. The coefficients of the polynomial differ from interval to interval, but the order of the polynomial is the same. We consider, for parsimony, cubic B-splines.

The introduction of splines allows the effect of covariates to vary along different time intervals, trying to better capture the effect of time-varying policy measures.

## 2.2 | INGARCH models

The PAR model of Agosto et al. (2021) is an INGARCH model which allows to specify the logarithmic link not only as a function of the past observations but also as a function of the past conditional mean of the process.

More formally, the INGARCH model can be specified as follows:

$$Y_t | \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t), \quad (4)$$

$$v_t = \log(\lambda_t) = \mu + \sum_{i=1}^p \alpha_i \log(1 + Y_{t-i}) + \sum_{j=1}^q \beta_j v_{t-j} + \eta^T X_t, \quad (5)$$

where the additional parameter  $\beta_j \in \mathbb{R}$ ,  $j = 1, \dots, q$  represents the dependence of the logarithmic link on the conditional mean of the Poisson process.

As done for AR models, we propose to consider a time-varying extension of INGARCH models, described by the following:

$$v_t = \mu(t/T) + \sum_{i=1}^p \alpha_i(t/T) \log(1 + Y_{t-i}) + \sum_{j=1}^q \beta_j(t/T) v_{t-j} + \eta^T(t/T) X_t, \quad (6)$$

where, in addition to the parameters in Equation (3),  $\beta_j(t/T)$  represents, for  $t = 1, \dots, T$ , time-varying parameters specified as functions of cubic B-splines.

### 3 | BAYESIAN INFERENCE

To better take model uncertainty into account, we consider a Bayesian specification for both AR (Section 3.1) and INGARCH models (Section 3.2). We specify the priors according to the parameter constraints in the time-constant framework, putting some transformations on parameters to improve sampling efficiency. At the end, the posterior distribution is specified.

Equations refer to the time-varying framework. However, for the time-constant case, they can be easily recovered.

#### 3.1 | tvBAR models

Let  $\theta = (\mu, \alpha_i, \eta_k)$  indicate the vector of parameters of the AR model. Their parameter space is given by

$$\Theta = \{\theta \in \mathbb{R}^{p+r+1} : |\alpha_1(x)|, \dots, |\alpha_p(x)| \leq 1, \sup_x \left| \sum_k \alpha_k(x) \right| \leq 1\}. \quad (7)$$

The constraints introduced in Equation (7) on  $\alpha_i(x)$ s are motivated by the time constant models where these conditions lead to stationarity and ergodicity. In the time-varying AR literature, they are not uncommon (Dahlhaus & Rao, 2006; Fryzlewicz et al., 2008).

The prior distributions for the time-varying Bayesian AutoRegressive (tvBAR) models can be specified, for each B-spline basis expansion as follows:

$$\mu(x) = \sum_{s=1}^{K_1} m_s B_s(x), \quad (8)$$

$$\alpha_i(x) = \sum_{s=1}^{K_2} g^{-1}(a_{i,s}) B_s(x), \quad (9)$$

$$\eta_k(x) = \sum_{s=1}^{K_3} n_{k,s} B_s(x), \quad (10)$$

$$m_s \sim N(0, 100), 1 \leq s \leq K_1, \quad (11)$$

$$a_{i,s} \sim N(0, 100), 1 \leq i \leq p, 1 \leq s \leq K_2, \quad (12)$$

$$n_{k,s} \sim N(0, 100), 1 \leq k \leq r, 1 \leq s \leq K_3, \quad (13)$$

where  $B_s$  are B-spline basis functions. The subscripts  $i$  and  $k$  refer to the time-constant dimension of the parameters (i.e.,  $i = 1, \dots, p$  and  $k = 1, \dots, r$ ), while  $s$  refers to the dimension of the B-splines functions.

For the unconstrained parameters  $m_s$  and  $n_{k,s}$  (which respectively refer to the intercept and the effect of explanatory covariates) we assume independent Gaussian priors in Equations (11) and (12) with variance equal to 100 to make the prior weakly informative. With reference to the constrained autoregressive parameter, we may specify a uniform prior in order to satisfy the stationarity conditions of time-constant models.

However, from a computational viewpoint, this implies that the autoregressive parameters ( $\gamma_i$ ) are to be sampled in a bounded space between  $-1$  and  $1$ , reducing the efficiency of the Markov chain Monte Carlo (MCMC) posterior approximation.

$$\gamma_i \sim U(-1, 1), \quad (14)$$

where  $1 \leq i \leq p$ . We therefore propose a parameter transformation, mapping the bounded parameter vector  $\Theta$  to an unbounded parameter vector  $u$  to improve sampling efficiency. The proposed mapping function is the following:

$$g(x) = \log\left(\frac{1-x}{1+x}\right), \quad (15)$$

where  $x \sim U(-1, 1)$ . In detail, the final autoregressive parameter ( $a_{i,s}$ ) to be sampled in Langevin Monte Carlo (LMC) algorithm is obtained as follows:

$$a_{i,s} = g(\gamma_i), a_{i,s} \sim N(0, 100), \quad (16)$$

where  $1 \leq i \leq p$  and  $1 \leq s \leq K_2$ . In this way, we put normal priors on the transformed B-spline coefficients to be sampled within the LMC algorithm and then transformed back the parameters for the likelihood function ( $g^{-1}(a_{i,s})$ , where  $g^{-1}$  is the inverse of function  $g$ ). This method results in better mixing of the posterior samples.

The log-likelihood function of the model, combined with the proposed prior distributions, gives rise to the following log posterior distribution ( $\mathcal{L}_1$ ) for the tvBAR model:

$$\begin{aligned} \mathcal{L}_1(\theta) \propto & \sum_{t=1}^T \left[ Y_t \left\{ \mu\left(\frac{t}{T}\right) + \sum_{i=1}^p \alpha_i\left(\frac{t}{T}\right) \log(1 + Y_{t-i}) + \eta^T X_t \right\} - \exp \left\{ \mu\left(\frac{t}{T}\right) + \sum_{i=1}^p \alpha_i\left(\frac{t}{T}\right) \log(1 + Y_{t-i}) + \eta^T X_t \right\} \right] \\ & - \sum_{s=1}^{K_1} \frac{m_s^2}{2 \cdot 100} - \sum_{s=1}^{K_3} \frac{n_s^2}{2 \cdot 100}. \end{aligned} \quad (17)$$

### 3.2 | tvBINGARCH models

Let  $\theta = (\mu, \alpha_i, \beta_j, \eta_k)$ , the vector of parameters of the INGARCH model. Their parameter space is given by

$$\Theta = \{\theta \in \mathbb{R}^{p+q+r+1} : |\alpha_1(x)|, \dots, |\alpha_p(x)| \leq 1, |\beta_1(x)|, \dots, |\beta_q(x)| \leq 1, \quad (18)$$

$$\sup_x \left| \sum_{k,l} \alpha_k(x) + \beta_l(x) \right| \leq 1\}.$$

The constraints in Equation (18) are inspired from the stationary condition from time-constant literature on INGARCH models (Geir et al., 2022).

For each B-spline basis expansion, prior distributions can be specified as follows:

$$\mu(x) = \sum_{s=1}^{K_1} m_s B_s(x), \quad (19)$$

$$\alpha_i(x) = \sum_{s=1}^{K_2} g^{-1}(a_{i,s}) B_s(x), \quad (20)$$

$$\beta_j(x) = \sum_{s=1}^{K_3} g^{-1}(b_{j,s}) B_s(x), \quad (21)$$

$$\eta_k(x) = \sum_{s=1}^{K_4} n_{k,s} B_s(x), \quad (22)$$

$$m_s \sim N(0, 100), 1 \leq s \leq K_1, \quad (23)$$

$$a_{i,s} \sim N(0, 100), 1 \leq i \leq p, 1 \leq s \leq K_2, \quad (24)$$

$$b_{j,s} \sim N(0, 100), 1 \leq j \leq q, 1 \leq s \leq K_3, \quad (25)$$

$$n_{k,s} \sim N(0, 100), 1 \leq k \leq r, 1 \leq s \leq K_4, \quad (26)$$

$$\lambda_0 \sim N(0, 100), \quad (27)$$

where  $B_s$  are B-spline basis functions. The subscripts  $i, j$ , and  $k$  refer to the time-constant dimension of the parameters (i.e.,  $i = 1, \dots, p$ ,  $j = 1, \dots, q$ , and  $k = 1, \dots, r$ ), while  $s$  refers to the dimension of the B-splines functions.

Priors for  $\mu$  and  $\eta_k$  can be recovered from Equations (11) and (13). Moreover, we need to additionally estimate the parameter  $\lambda_0$  to start the recursion of the conditional mean (Equation 27).

Similar to what is done in Section 3.1, we propose a transformation of the time-dependent parameters to ensure the condition about the absolute value of their sum (Equations 24 and 25), taking the logarithmic function (15) to map into an unbounded parameter space for sampling. The starting values for time-dependent parameters ( $r_i, k_j$ ), together with their transformed version ( $a_{i,s}, b_{j,s}$ ) can be summarized as follows:

$$r_i \sim U(-1, 1), \quad (28)$$

$$k_j \sim U(-1, 1), \quad (29)$$

$$a_{i,s} = g\left(\frac{r_i + k_j}{2}\right), a_{i,s} \sim N(0, 100), \quad (30)$$

$$b_{j,s} = g\left(\frac{r_i - k_j}{2}\right), b_{j,s} \sim N(0, 100), \quad (31)$$

where  $1 \leq i \leq p$ ,  $1 \leq j \leq q$ , and  $1 \leq s \leq K$  with  $K = (K_2, K_3)$ .

Specifically,  $m_s$ ,  $\alpha_{i,s}$ ,  $b_{j,s}$ ,  $n_{k,s}$ , and  $\lambda_0$  are the parameters to be sampled in MCMC iterations. After sampling, the parameters  $\alpha_{i,s}$  and  $b_{j,s}$  are transformed back ( $g^{-1}(\alpha_{i,s})$ ,  $g^{-1}(b_{j,s})$ ) for the likelihood function. The posterior distribution  $\mathcal{L}_2$  for the Bayesian INGARCH (BINGARCH) model can then be obtained as follows for the time-varying case:

$$\begin{aligned} \mathcal{L}_2(\theta) \propto & \sum_{t=1}^T \left[ Y_t \left\{ \mu \left( \frac{t}{T} \right) + \sum_{i=1}^p \alpha_i \left( \frac{t}{T} \right) \log(1 + Y_{t-i}) + \sum_{j=1}^q \beta_j \left( \frac{t}{T} \right) v_{t-j} + \eta^T X_t \right\} \right. \\ & \left. - \exp \left\{ \mu \left( \frac{t}{T} \right) + \sum_{i=1}^p \alpha_i \left( \frac{t}{T} \right) \log(1 + Y_{t-i}) + \sum_{j=1}^q \beta_j \left( \frac{t}{T} \right) v_{t-j} + \eta^T X_t \right\} \right] \\ & - \sum_{j=1}^{K_1} \frac{m_j^2}{2 \cdot 100} - \sum_{j=1}^{K_4} \frac{n_j^2}{2 \cdot 100} - \frac{\lambda_0^2}{2 \cdot 100}. \end{aligned} \quad (32)$$

To calculate the posterior distribution from the previous expressions, we employ the gradient-based. The Hamiltonian Monte Carlo (HMC) algorithm is similar to Roy and Karmakar (2021). Tuning of HMC parameters is necessary to obtain a good acceptance rate. HMC has two parameters, the number of leapfrog steps and step length. Since gradient computation is expensive, we consider setting the number of leapfrog steps to one. In this way, HMC becomes LMC (Neal, 2011). According to Roberts et al. (1998), the optimal acceptance rate for LMC is around 0.57. We thus maintain an acceptance rate between 0.45 and 0.7 by adaptively tuning the step-length parameter. Based on our results, this produces excellent mixing of the posterior samples (see Figure 3 for reference).

We want to remark that, due to the increasing complexity of the parameter space for the INGARCH model,  $\alpha_{i,s}$ ,  $b_{j,s}$ , and  $\lambda_0$  are sampled together.

## 4 | DATA

The proposed models have been applied to two types of data sources: COVID-19 impact data (taken from the European Centre for Disease Prevention and Control (ECDC): [www.ecdc.europa.eu/en/covid-19/data](http://www.ecdc.europa.eu/en/covid-19/data)) and policy intervention data (taken from Oxford Covid-19 Government Response Tracker (OxCGRT): [www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker](http://www.bsg.ox.ac.uk/research/research-projects/covid-19-government-response-tracker)).

From ECDC, we collected daily frequency COVID-19 count data from two representative countries: Italy and the United States. We consider, as target response variable, the *change in total positive cases*, defined as the difference between the total counts of positives for the current day and those the day before. The starting point of the time series was chosen as the date when the country recorded more than 100 recorded cases.

We also obtained data on the taken policy measures from OxCGRT (Hale et al., 2020), a project from Oxford's Blavatnik School of Government, in which each government response to the pandemic is measured according to an ordinal scale that ranges from no measures to the most severe form of measure. Our primary interventions of interest are the measures defined as *C—containment and closure policies* as well as those defined as *H—health system policies*.

We remark that the epidemic and the policy measures taken to tackle the epidemics have been quite heterogeneous across countries, with NPI measures taken at the local authority level, varying across regions and across time. This heterogeneity suggests that these covariates have an intrinsic measurement error.

Tables 1 and 2 show a preliminary analysis of the OxCGRT variables of interest. Table 1 indicates sparse frequency data, mostly in correspondence with health system policies, which have less frequent changes over time, compared to containment and closure policies. Table 2 indicates high collinearity between policy measures, which can be explained by the fact that NPIs have not been implemented independently of each other but, rather, in a multivariate *policy mix* approach.

We have addressed these issues by grouping NPIs into homogeneous categories, in line with Cowling et al. (2020) and Davies et al. (2020) who concluded that combined interventions have been the most effective strategy. Bo et al. (2021) also showed that social distancing in combination with other NPIs has led to a significant decrease in COVID-19 transmission.

More specifically, following Li et al. (2020), the original policy variables have been transformed from the ordinal to the binary scale, taking a value of “0” in case of *no intervention* or *recommended intervention* and a value of “1” in case of *required intervention*. We also grouped policies into five variables, following Islam et al. (2020): closure policies (*X1\_closures*), restrictions on mass gatherings and events (*X2\_events*), lockdown-type policies (*X3\_lockdown*), testing,



**TABLE 1** Frequency table of NPI variables obtained from OxCGRT. Variables have been coded on an ordinal scale according to the severity of government response. NPIs with reference to Italy and United States are shown

NPI	Country	0	1	2	3	4	5
<i>C1_School.closing</i>	Italy	0	22	161	246	0	0
	United States	0	0	98	331	0	0
<i>C2_Workplace.closing</i>	Italy	0	33	214	182	0	0
	United States	10	50	265	104	0	0
<i>C3_Cancel.public.events</i>	Italy	0	0	429	0	0	0
	United States	0	53	376	0	0	0
<i>C4_Restrictions.on.gatherings</i>	Italy	28	0	113	17	271	0
	United States	2	1	9	65	352	0
<i>C5_Close.public.transport</i>	Italy	174	233	22	0	0	0
	United States	8	421	0	0	0	0
<i>C6_Stay.at.home.requirements</i>	Italy	88	84	236	21	0	0
	United States	6	175	248	0	0	0
<i>C7_Restrictions.on.internal.movement</i>	Italy	142	0	287	0	0	0
	United States	5	49	375	0	0	0
<i>H2_Testing.policy</i>	Italy	0	0	429	0	0	0
	United States	0	0	5	424	0	0
<i>H3_Contact.tracing</i>	Italy	0	0	429	0	0	0
	United States	0	429	0	0	0	0
<i>H6_Facial.Coverings</i>	Italy	26	0	0	0	403	0
	United States	1	25	2	86	315	0
<i>H7_Vaccination.policy</i>	Italy	293	0	110	26	0	0
	United States	280	75	31	20	16	7

**TABLE 2** Most significant partial correlations between NPI covariates obtained from OxCGRT. A p-value less than 0.001 is flagged with \*\*\*. Results with reference to Italy and United States are shown

Country	NPI(1)	NPI(2)	Pcorr
Italy	<i>C1_School.closing</i>	<i>C3_Cancel.public.events</i>	-0.623***
	<i>C3_Cancel.public.events</i>	<i>C4_Restrictions.on.gatherings</i>	-0.955***
	<i>C4_Restrictions.on.gatherings</i>	<i>C5_Close.public.transport</i>	0.464***
	<i>C3_Cancel.public.events</i>	<i>C6_Stay.at.home.requirements</i>	-0.637***
	<i>C6_Stay.at.home.requirements</i>	<i>C7_Restrictions.on.internal.movement</i>	0.58***
United States	<i>C3_Cancel.public.events</i>	<i>H6_Facial.Coverings</i>	-0.673***
	<i>C5_Close.public.transport</i>	<i>C7_Restrictions.on.internal.movement</i>	0.591***
	<i>C2_Workplace.closing</i>	<i>H2_Testing.policy</i>	0.458***
	<i>C1_School.closing</i>	<i>H6_Facial.Coverings</i>	-0.483***
	<i>C1_School.closing</i>	<i>H7_Vaccination.policy</i>	-0.641***

tracing, and masks (*X4\_masks*), and vaccination policy (*X5\_vaccines*). The precise mapping between the OxCGRT variables and ours is shown in Table 3.

## 5 | EMPIRICAL FINDINGS

We first implemented the frequentist AR and INGARCH models using the `tsglm` function from the R package `tscount` (Liboschik et al., 2017). As shown in Table 4, the  $p$  order of the INGARCH structure has been selected according to the



**TABLE 3** Transformation of OxCGRT NPIs. Variables have been transformed from ordinal to binary scale taking the value of 0 in case of “no intervention” or “recommended intervention” and 1 in case of “required intervention.” Policy variables have then been grouped based on the thematic area of reference

NPI	Description	OxCGRT	Country	Value	Freq
<i>X1_closures</i>	Require closing of schools (or work from home) for some or all levels	c1; c2	Italy	0	247
			Italy	1	192
			United States	0	325
			United States	1	104
<i>X2_events</i>	Require canceling of events and restrictions on gatherings between 11-100 people or less	c3; c4	Italy	0	141
			Italy	1	298
			United States	0	62
			United States	1	367
<i>X3_lockdown</i>	Require closing of public transport, require not leaving the house with some exceptions and internal movement restrictions in place	c5; c6; c7	Italy	0	172
			Italy	1	267
			United States	0	52
			United States	1	377
<i>X4_masks</i>	Open public testing, comprehensive contact tracing and facial coverings required outside the home	h2; h3; h6	Italy	0	36
			Italy	1	403
			United States	0	28
			United States	1	401
<i>X5_vaccines</i>	Availability of vaccines for two or more groups of people	h7	Italy	0	303
			Italy	1	136
			United States	0	355
			United States	1	74

**TABLE 4** Selection of the order of autoregressive dependence on past counts ( $p$ ) for INGARCH structure based on MAPE

Country	Model	$p = 1$	$p = 7$	$p = 14$
Italy	AR	15%	24%	46%
	INGARCH	15%	24%	46%
United States	AR	11%	15%	29%
	INGARCH	12%	16%	28%

absolute error of the median predictions with respect to the true values (median absolute percentage error (MAPE))). Based on the mean incubation period of the disease,  $p$  equal to 1, 7, and 14 have been tested. Results show that  $p$  equal to 1 reports the lowest MAPE in both Italy and United States. Then, the time-dependent model structure has been extended adding exogenous covariates for policy effect.

Additionally, we have implemented the same model structures within the Bayesian framework for which we have developed specific *R* functions. Since we extend the frequentist framework models by introducing B-splines, we discuss about the selection of the number of knots for the B-splines (Table 7) before presenting parameter estimates for both time-constant and time-varying coefficients.

Finally, we compared and discussed the results obtained in terms of predictive performance, using the last week of data as a test set. Point forecasts have been evaluated in terms of MAPE, whereas the whole distribution has been evaluated using scoring rules (Gneiting & Raftery, 2007), implemented in the *R* package `scoringutils` (Bosse et al., 2020).

The source code to reproduce the results is available as [Supporting Information](#) on the journal’s web page.

**TABLE 5** Parameter estimates and standard errors for frequentist Poisson AR/INGARCH models for COVID-19 data. When the  $p$ -value against the null hypothesis of a zero coefficient is less than 0.001 the result is flagged with \*\*\*

Parameter	Country	Model AR	Model INGARCH
$\mu$	Italy	0.367*** (0.006)	0.428*** (0.007)
	United States	0.526*** (0.003)	0.169*** (0.001)
$\alpha$	Italy	0.962*** (0.001)	1*** (0.003)
	United States	0.954*** (0)	0.466*** (0.001)
$\beta$	Italy		-0.044*** (0.003)
	United States		0.52*** (0.001)

**TABLE 6** Parameter estimates and standard errors of NPI covariates for the frequentist Poisson ARX/INGARCHX models applied to the COVID-19 data. When the  $p$ -value against the null hypothesis of a zero coefficient is less than 0.001 the result is flagged with \*\*\*

Parameter	Country	Model ARX	Model INGARCHX
$X1\_closure$	Italy	-0.03*** (0.001)	-0.028*** (0.001)
	United States	-0.037*** (0.001)	-0.013*** (0)
$X2\_events$	Italy	0.189*** (0.005)	0.224*** (0.005)
	United States	-0.043*** (0.002)	-0.058*** (0.001)
$X3\_lockdown$	Italy	0.026*** (0.003)	0.05*** (0.003)
	United States	0.055*** (0.002)	0.059*** (0.001)
$X4\_masks$	Italy	0.068*** (0.004)	0.102*** (0.004)
	United States	-0.04*** (0.002)	-0.124*** (0.001)
$X5\_vaccines$	Italy	-0.035*** (0.001)	-0.042*** (0.001)
	United States	-0.048*** (0.001)	-0.02*** (0)

## 5.1 | Frequentist analysis

The results from the frequentist Poisson AR/INGARCH models, grouped by country, are shown in Table 5.

The AR estimates in Table 5 show a strong positive  $\alpha_1$  for both countries, which indicates a positive dependence of counts over time. The INGARCH estimates in the same table show that, for Italy,  $\alpha_1$  is much greater than  $\beta_1$ . This indicates that short-term dependence on past observed disease counts prevails over long-term dependence: a result that could be explained by the containment and closure measures introduced very promptly in response to the COVID-19 spread. Differently, in the United States, the estimates for  $\alpha_1$  and  $\beta_1$  have a similar magnitude, revealing the persistence of long memory effects.

The impacts of policy measures are summarized in Table 6.

Table 6 shows that, for Italy, closure policies ( $X1\_closures$ ) and vaccines ( $X5\_vaccines$ ) have significant negative coefficients, meaning that these policy measures have significantly reduced infection counts. In contrast, restrictions on gatherings and events ( $X2\_events$ ), lockdown measures ( $X3\_lockdown$ ), and testing, tracing, and masks ( $X4\_masks$ ) reveal a significant positive impact on counts. This could be explained by the fact that these policy measures are typically imposed when the number of cases is high, leading to a “reverse effect” that leads to stronger policies when counts are higher.

The results for the United States also indicate negative significant effects for closure policies ( $X1\_closures$ ) and vaccines ( $X5\_vaccines$ ), but also for restrictions on gatherings and events ( $X2\_events$ ) and for testing, tracing, and masks policies ( $X4\_masks$ ). Only lockdown-type policies ( $X3\_lockdown$ ) have a positive coefficient, which can be attributed to the “reverse effect” found for Italy. These results are in line with Hsiang et al. (2020) who have found effectiveness of most anticontagion policies in slowing down the exponential growth of the pandemic in the United States.

## 5.2 | Bayesian analysis

To implement the proposed Bayesian time-varying models, a preliminary selection of the number of spline knots is necessary. A large number of knots can cause overfitting, whereas a small number can result in underfitting. We select

TABLE 7 Selection of the number of knots for time-varying Bayesian models based on the MAPE

Country	Model	Knots = 2	Knots = 4	Knots = 6	Knots = 8	Knots = 10
Italy	tvBAR	15%	14%	44%	21%	17%
	tvBINGARCH	100%	15%	100%	100%	18%
United States	tvBAR	39%	14%	10%	18%	12%
	tvBINGARCH	100%	11%	100%	100%	19%

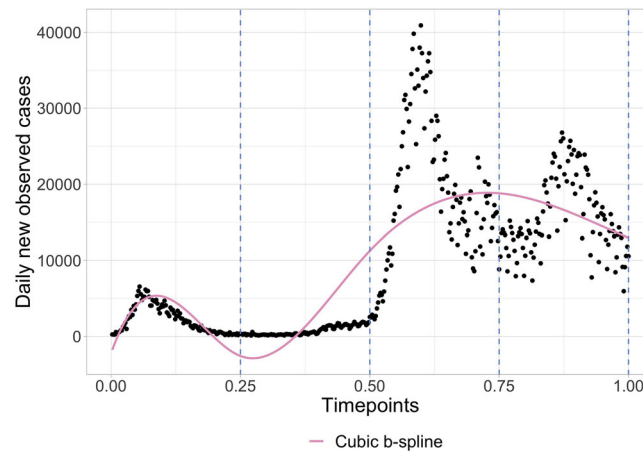


FIGURE 1 Cubic spline fit of daily new observed cases of COVID-19 in linear scale for Italy

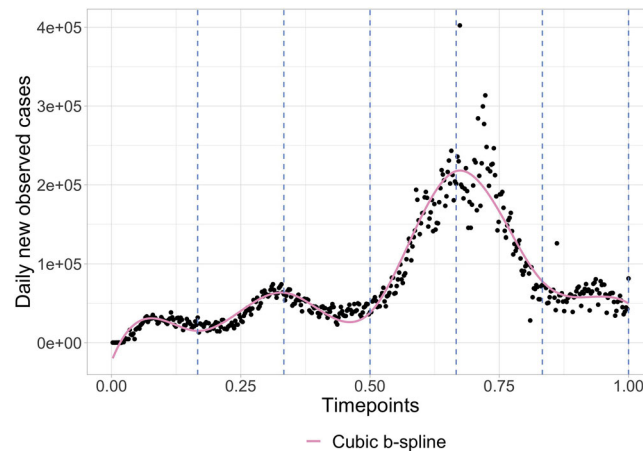


FIGURE 2 Cubic spline fit of daily new observed cases of COVID-19 in linear scale for the United States

the number of knots on the basis of the results from the MAPE of the tvBAR and time-varying Bayesian INGARCH (tvBINGARCH) models, applied to the training data, as shown in Table 7.

From Table 7, the best number of knots, which minimizes the MAPE for the tvBAR model is 4 for Italy and 6 for the United States; for the tvBINGARCH model, it is 4 for both countries. Figures 1 and 2 show the fitted cubic B-splines of the daily new observed cases of COVID-19 for Italy and for the United States, applying tvBAR models with, respectively, four and six knots.

Figure 2 shows that the peaks and the evolution of COVID-19 data are well captured by a spline with six knots for the United States, which identifies three main peaks. Figure 1 shows that a simpler four knots spline is a good approximation for Italy, which identifies two main peaks, smoothing the two observed peaks in the second wave of the pandemic.

To derive the posterior distributions, the LMC algorithm has been run for 10,000 iterations, with the first 5000 discarded as burn-in. All computations have been validated with convergence diagnostic assessments. For illustration, we report here the trace plots of the average likelihood function deviations along the MCMC chain for the coefficient functions ( $\theta$ ),

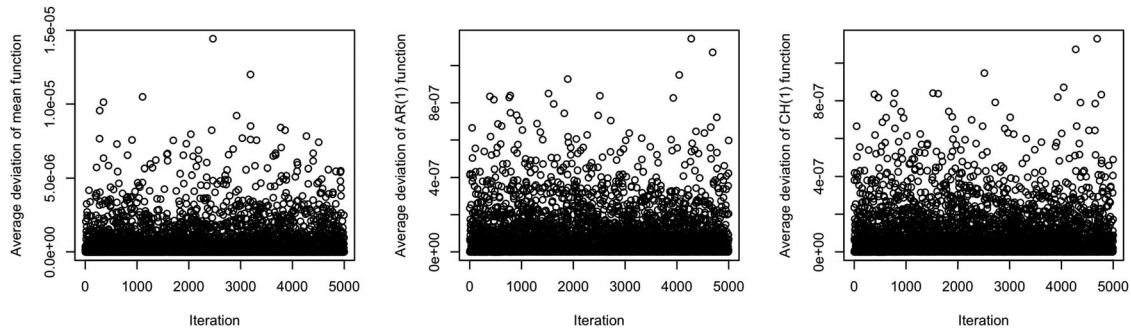


FIGURE 3 Trace plots of average likelihood function deviations across the MCMC chain for the BINGARCH model

TABLE 8 Posterior mean and 5%, 25%, 75%, and 95% posterior quantiles using time-constant BAR/BINGARCH models for Italy

Parameter	Model	Posterior mean	5%	25%	75%	95%
$\mu$	BAR	0.3837	0.3773	0.3807	0.3863	0.3903
	BINGARCH	0.4356	0.4259	0.4324	0.4399	0.4422
$\alpha$	BAR	0.9608	0.9601	0.9605	0.9611	0.9614
	BINGARCH	0.4776	0.4773	0.4774	0.4778	0.4781
$\beta$	BINGARCH	0.5873	0.5824	0.5849	0.5895	0.5921

TABLE 9 Posterior mean and 5%, 25%, 75%, and 95% posterior quantiles using time-constant BAR/BINGARCH models for the United States

Parameter	Model	Posterior mean	5%	25%	75%	95%
$\mu$	BAR	0.5326	0.531	0.532	0.5329	0.5342
	BINGARCH	0.18	0.1784	0.18	0.1803	0.1806
$\alpha$	BAR	0.954	0.9539	0.954	0.9541	0.9542
	BINGARCH	0.4666	0.4664	0.4665	0.4667	0.4668
$\beta$	BINGARCH	0.5183	0.5182	0.5183	0.5184	0.5185

calculated with the following formula:

$$\frac{1}{T} \left[ \sum_{j=1}^T (\theta(j/T)^t - \theta(j/T)^{t+1})^2 \right]^{1/2}, \quad (33)$$

where  $\theta(j/T)^t$  represents the  $t$ th postburn samples with  $T = 5000$  post burn iterations. For explanatory purposes, the trace plots for  $\theta = \{\mu, \alpha_i, \beta_j, \eta\}$  with regard to the BINGARCH model are reported in Figure 3. They exhibit efficient mixing of the posterior samples.

The posterior distributions of the time constant parameters are given in Tables 8 and 9 for both countries.

From Tables 8 and 9, note that, for the United States, the frequentist findings in Section 5.1 are confirmed by the Bayesian estimates: both short-term and long-term dependence of counts are present. The same result is found for Italy, which is different from what occurred with the frequentist estimates, and consistent with the similarity between the two countries.

We now examine the posterior distributions of the time-varying coefficients. Figures 4 and 5 depict the time evolution of the estimated  $\mu$  and  $\alpha_1$  from the Bayesian AutoRegressive (BAR) models for Italy and the United States. Figure 4 suggests that, for Italy, the estimated autoregressive coefficient of order 1 has a high impact on disease counts at the beginning of the period of analysis: the first wave of the pandemic. The impact starts to decrease when the number of cases decrease and increases in correspondence to the second wave of the pandemic. An almost opposite behavior is shown by the mean trend, which depends on the stock of cases, rather than on their flows. In April 2021, decreasing autoregressive coefficient estimates emerge, together with a low mean trend, resulting in a decreasing magnitude of the pandemic.

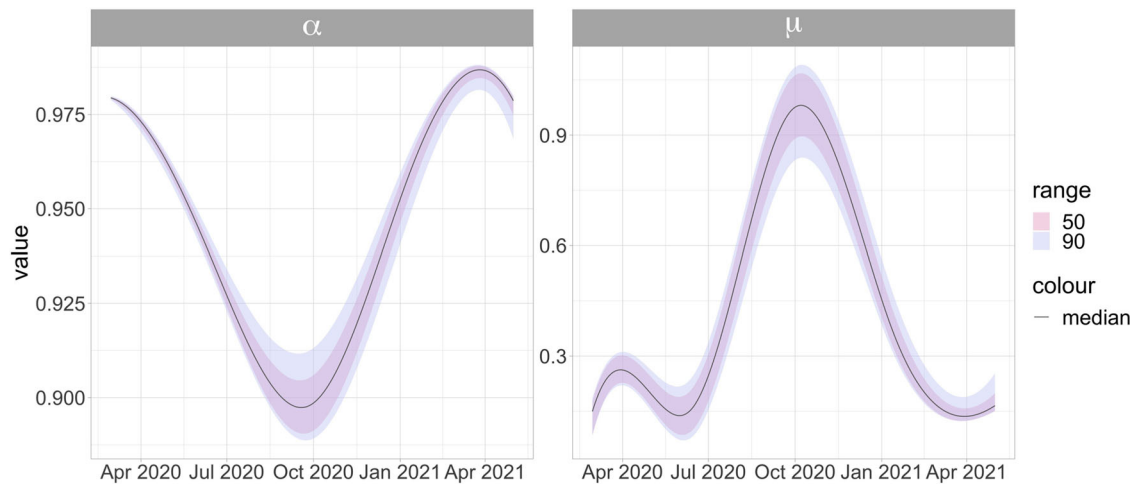


FIGURE 4 Time-varying posterior coefficient distribution (tvBAR model) for Italy

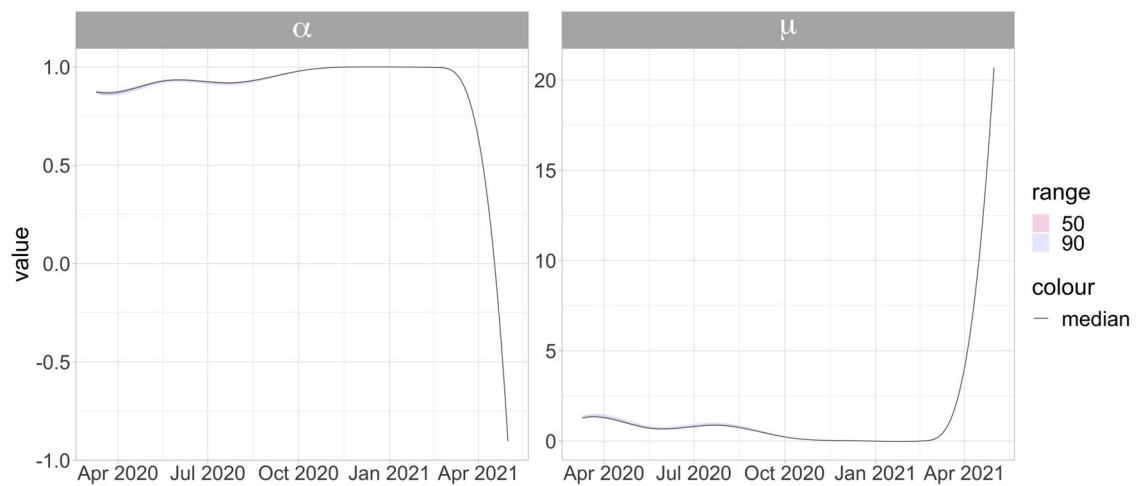


FIGURE 5 Time-varying posterior coefficient distribution (tvBAR model) for the United States

Figure 5 suggests that, for the United States, no evident turning points in the trend of estimated functions appear; the autoregressive coefficient is almost always close to 1 with a high mean trend. In the last period of analysis, the autoregressive component decreases and the mean trend becomes the prevailing factor.

Figures 6 and 7 report the time evolution of the  $\mu$ ,  $\alpha_1$ , and  $\beta_1$  estimated parameters from the BINGARCH models.

For Italy, Figure 6 confirms the results from the BAR model, also showing an increasing contribution of the conditional mean in the initial period of analysis, reaching the peak in the summer, where the contagion was decreasing and quite stable.

For the United States, Figure 7 reports a higher contribution of  $\alpha_1$  over  $\beta_1$  during the first wave of the pandemic and the opposite behavior in the second wave. A growing mean trend is still present in April 2021.

The further step of the analysis involves adding NPIs covariates within the proposed modeling framework, comparing the results from the BAR and BINGARCH models. Tables 10 and 11 show the results, when time constant parameters are considered.

The frequentist results in Table 6 are confirmed by the Bayesian posterior estimates in Tables 10 and 11. Table 10 suggests that for Italy the closures of schools and workplaces ( $X1\_closures$ ) and vaccination policy ( $X5\_vaccines$ ) are both effective in reducing COVID-19 counts; whereas other variables have a positive sign. Table 11 confirms that, in the United States, most policies (with the exception of lockdown measures) reduce COVID-19 counts.

When time-varying parameters are considered, the median point estimates in Tables 10 and 11 are replaced by a range of estimates, corresponding to different time points. Figures 8 and 9 show the results for Italy and the United

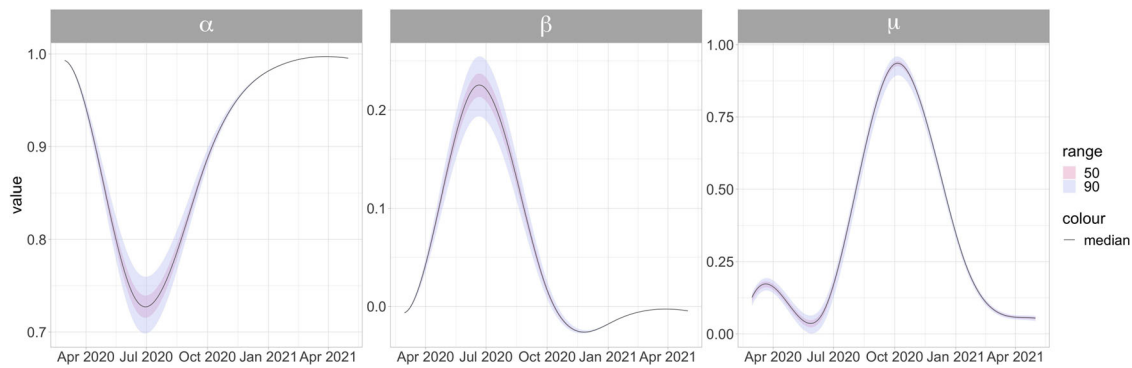


FIGURE 6 Time-varying posterior coefficient distribution (tvBINGARCH model) for Italy

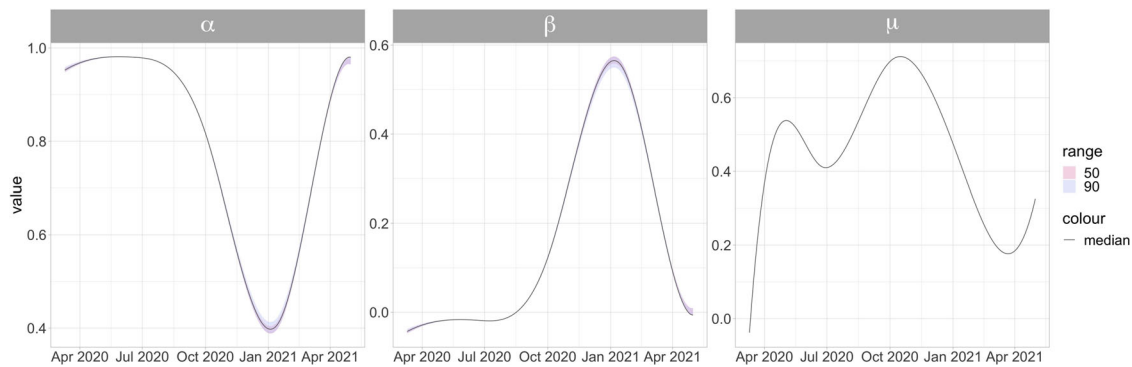


FIGURE 7 Time-varying posterior coefficient distribution (tvBINGARCH model) for the United States

TABLE 10 Posterior mean and 5%, 25%, 75%, and 95% posterior quantiles of NPI covariates using time-constant BARX/BINGARCHX models for Italy

Parameter	Model	Posterior mean	5%	25%	75%	95%
<i>X1_closures</i>	BARX	-0.0243	-0.0268	-0.0259	-0.0229	-0.0218
	BINGARCHX	-0.0223	-0.0237	-0.0229	-0.0217	-0.0207
<i>X2_events</i>	BARX	0.1916	0.189	0.1903	0.193	0.1942
	BINGARCHX	0.2327	0.2294	0.2312	0.2343	0.2352
<i>X3_lockdown</i>	BARX	0.0307	0.0271	0.0296	0.032	0.0329
	BINGARCHX	0.0478	0.0457	0.0469	0.0487	0.0503
<i>X4_masks</i>	BARX	0.0731	0.0703	0.0722	0.0743	0.0752
	BINGARCHX	0.1088	0.1064	0.1076	0.1101	0.1124
<i>X5_vaccines</i>	BARX	-0.0301	-0.0319	-0.0308	-0.0293	-0.0282
	BINGARCHX	-0.0378	-0.0399	-0.0388	-0.0369	-0.0354

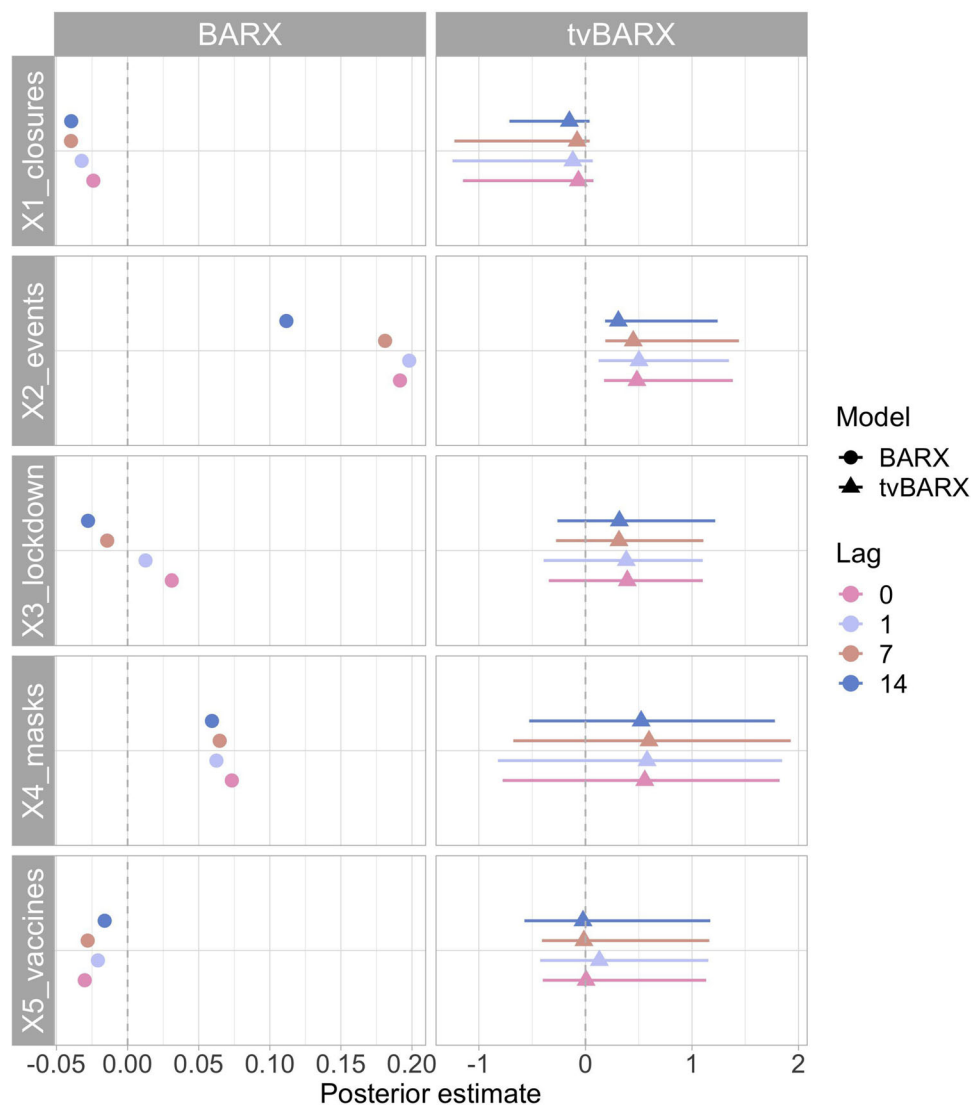
States, comparing the median estimates of the Bayesian AutoRegressive eXogenous (BARX) models with the estimates of the time-varying Bayesian AutoRegressive eXogenous (tvBARX) models. In addition, we consider different lag length windows of 1, 7, and 14 days to identify a possible time lag between policy interventions and their epidemiological effect.

From Figure 8, we can gauge that a time lag of about 1–2 weeks seems to increase the impact of covariates representing containment and closure policies (*X1\_closure*), confirming that a large number of days are needed between the activation of this type of policy and the realization of its epidemiological effect. Figure 9 suggests a more immediate time impact of policies for the United States, as they lead to more negative estimates without time lag or with a delay of at most 1 day.

Both Figures 8 and 9 show a large variation in policy effects within the considered period. In particular, the effects of lockdowns (*X3\_lockdowns*) and testing, tracing, and masks (*X4\_masks*) vary much more than any of the other policy

**TABLE 11** Posterior median and 5%, 25%, 75%, and 95% posterior quantiles of NPI covariates using time-constant BARX/BINGARCHX models for the United States

Parameter	Model	Posterior mean	5%	25%	75%	95%
X1_closures	BARX	-0.0369	-0.0378	-0.0374	-0.0365	-0.036
	BINGARCHX	-0.0129	-0.0132	-0.0131	-0.0128	-0.0126
X2_events	BARX	-0.0432	-0.0469	-0.0443	-0.0417	-0.0405
	BINGARCHX	-0.056	-0.0565	-0.0561	-0.0559	-0.0557
X3_lockdown	BARX	0.04	0.0366	0.0384	0.0424	0.0432
	BINGARCHX	0.0517	0.0513	0.0515	0.0517	0.0525
X4_masks	BARX	-0.0421	-0.045	-0.044	-0.0405	-0.0386
	BINGARCHX	-0.1234	-0.1239	-0.1237	-0.1231	-0.1219
X5_vaccines	BARX	-0.0476	-0.0492	-0.0482	-0.047	-0.0463
	BINGARCHX	-0.0202	-0.0207	-0.0203	-0.02	-0.0197



**FIGURE 8** Time-constant (BARX) versus time-varying (tvBARX) posterior coefficient estimates for NPI covariates (Italy). The range of variation has been computed from the min-max range of the posterior median over the time interval of interest



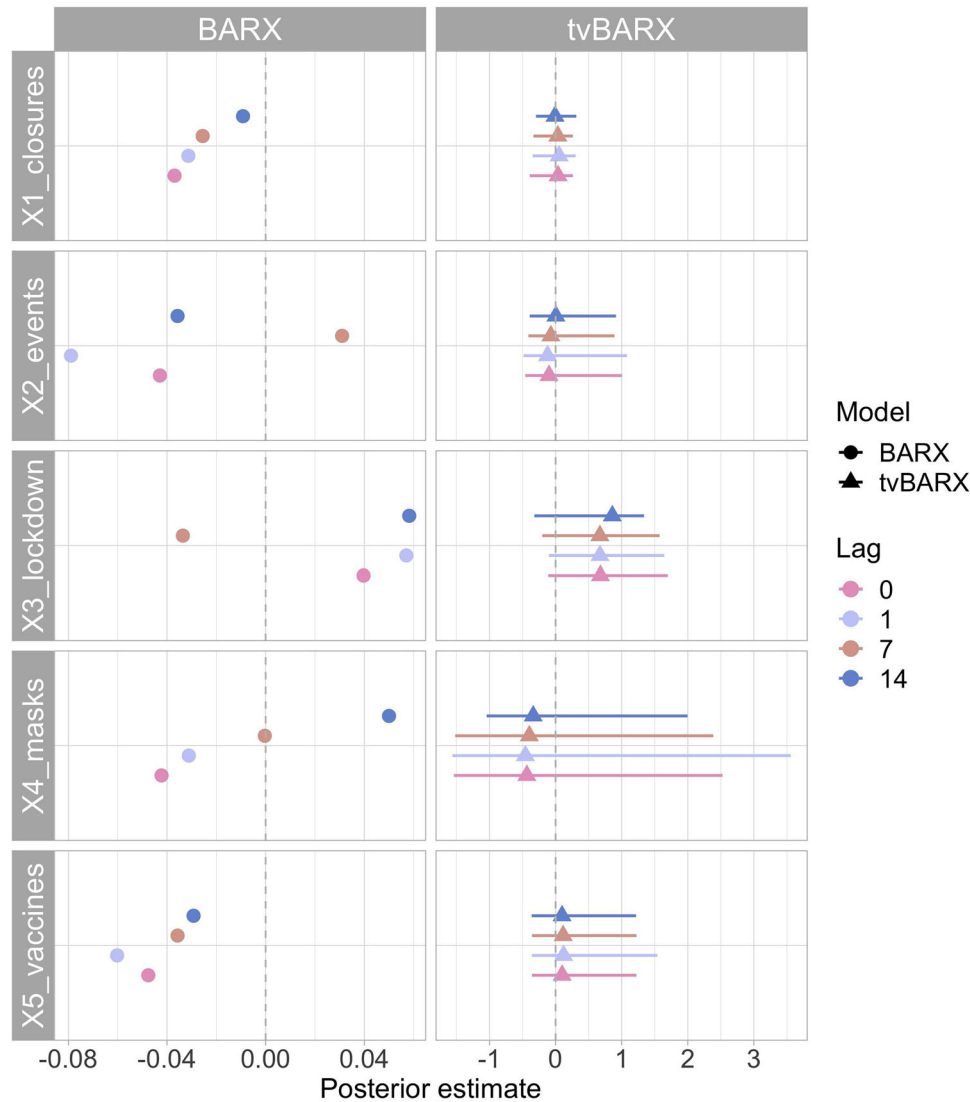


FIGURE 9 Time-constant (BARX) versus time-varying (tvBARX) posterior coefficient estimates for NPI covariates (United States). The range of variation has been computed from the min-max range of the posterior median over the time interval of interest

measures. This can be attributed to the presence of reverse causality effects, as policies have been implemented in response to the rise in pandemic counts.

Overall, it can be concluded that Bayesian estimates are in line with frequentist ones since the model structure is the same in the time constant case, but, additionally, incorporate greater uncertainty by means of the posterior distribution and the time-varying coefficient estimates.

### 5.3 | Model comparison

All models have been assessed in terms of the absolute error of the median forecasts, calculated as the absolute value of the difference between the true values and median posterior forecasts obtained from MCMC iterations. A high MAPE means that the median forecasts tend to be far away from the true values.

The accuracy of model prediction intervals has also been evaluated in terms of overprediction, underprediction, and bias, calculated from the scoring rules applied to the predictive Monte Carlo samples. They are defined with respect to the upper bound (overprediction) or lower bound (underprediction) of the prediction interval. The bias is a measure between  $-1$  and  $1$  that expresses the tendency to underpredict ( $-1$ ) or overpredict ( $1$ ) the true values. All posterior predictions

TABLE 12 Model comparison for Italy with the Bayesian AR models

Model	Set	MAPE	Under	Over	Bias
BAR	Train	15%	672	672	0.13
	Test	36%	0	3576	1
BARX	Train	15%	646	647	0.053
	Test	28%	0	2934	1
tvBAR	Train	15%	362	369	-0.019
	Test	10%	10	950	0.357
tvBARX	Train	12%	128	155	0.002
	Test	15%	11	128	0.057

TABLE 13 Model comparison for Italy with the Bayesian INGARCH models. Note that the BINGARCH model specification has not been reported due to low-performance metrics

Model	Set	MAPE	Under	Over	Bias
BINGARCHX	Train	15%	446	463	0.073
	Test	28%	0	1671	0.686
tvBINGARCH	Train	15%	572	579	0.004
	Test	9%	27	1718	0.554
tvBINGARCHX	Train	17%	360	399	-0.118
	Test	21%	440	479	-0.074

TABLE 14 Model comparison for the United States with the Bayesian AR models

Model	Set	MAPE	Under	Over	Bias
BAR	Train	11%	5352	5354	0.046
	Test	50%	1520	18098	0.714
BARX	Train	11%	5088	5088	0.047
	Test	31%	3003	12170	0.714
tvBAR	Train	10%	4426	4472	-0.034
	Test	32%	3455	6700	0.443
tvBARX	Train	11%	3942	3977	-0.037
	Test	29%	4410	7816	0.436

have been evaluated both on the train and on the test set (one week ahead). Tables 12 and 13 give the results of model comparison for Italy, whereas Tables 14 and 15 show the results for the United States.

The results show that most of the proposed models fit the data quite well. For Italy, the tvBARX model seems to be the best, with a small MAPE (equal to 12%) on the training set and a bias almost equal to 0. The model also performs quite well on test data, with a 15% MAPE and a bias equal to 0.057. Similar conclusions are obtained for the United States, although the bias on the test set is higher (0.436), consistent with the higher number of cases with respect to Italy.

To further demonstrate the accuracy of our proposed models, Figures 10 and 11 plot the posterior median of the fitted and predicted values, together with the number of observed cases of COVID-19 for Italy and the United States. The interval range (50–90%) has been added to both figures.

From Figures 10 and 11 both models appear to perform quite well.

We finally remark that our proposed TvBINGARCHX model was employed within the ECDC European Forecast Hub initiative and compared with many other models that attempted to predict counts of positive cases from 1 to 4 weeks ahead. The latter model was chosen since allows to account for both short-term and long-term dependence together with

TABLE 15 Model comparison summary for the United States with the Bayesian INGARCH models

Model	Set	MAPE	Under	Over	Bias
BINGARCH	Train	12%	5110	5116	-0.05
	Test	33%	2684	12016	0.714
BINGARCHX	Train	12%	4998	5022	-0.013
	Test	26%	3706	8168	0.671
tvBINGARCH	Train	11%	2304	2219	-0.033
	Test	60%	131	4976	0.414
tvBINGARCHX	Train	10%	247	219	-0.017
	Test	35%	1353	9421	0.717

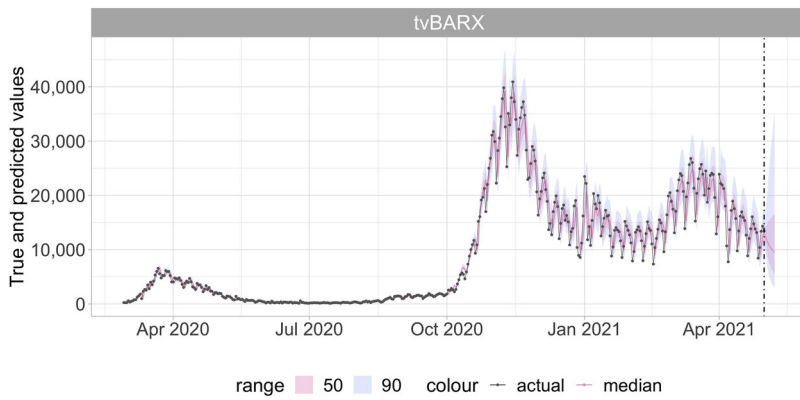


FIGURE 10 Predictions versus observed values over train and test data, with daily frequency (tvBARX model for Italy)

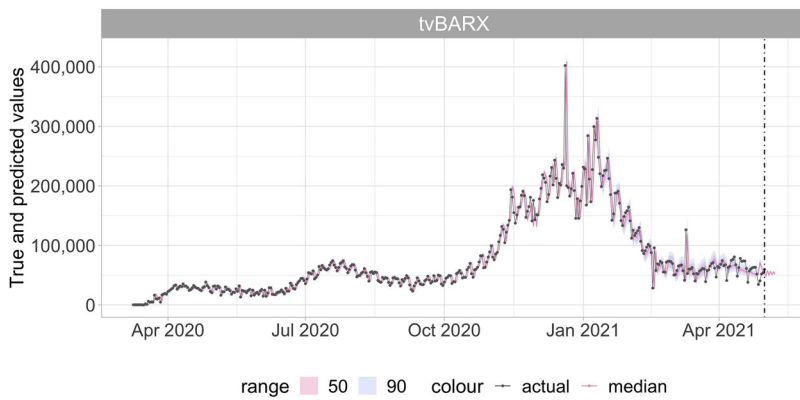


FIGURE 11 Predictions versus observed values over train and test data, with daily frequency (tvBARX model for the United States)

policy effects. The countries of interest were France, Germany, and Italy. Table 16 summarizes the evaluation metrics for the submission period that goes from March 2021 to November 2021, showing the best performance for Germany, with a bias equal to 0, slightly higher for Italy (0.6), and France (0.7). Real-time forecasts of cases are really challenging, given data issues in COVID-19 databases and intrinsic model error together with changing policies, population behavior, and testing procedures.

TABLE 16 Evaluation metrics of weekly forecasts made with the tvBINGARCHX model for Germany, France, and Italy for the European Forecast Hub project summarized throughout all the submission period (from 8 March 2021 to 29 November 2021)

Location	Under	Over	Bias
DE	2	0	0
FR	0	1858	0.7
IT	0	3124	0.6

## 6 | DISCUSSION

Motivated by the need to deliver reliable estimates of the COVID-19 contagion dynamics, in this paper we have provided novel Bayesian Poisson autoregressive models, with time-varying coefficients. While time-varying coefficients allow to better take into account of nonlinear growth patterns, the Bayesian framework allows to incorporate the uncertainty of model outputs. The addition of policy intervention variables allows to better understand the timing, sign, and magnitude of the policy measures undertaken by the governments to face the pandemic.

The application of the model to two paradigmatic cases as Italy and the United States reveals that some combinations of policy interventions together with the first phase of the vaccination campaign are associated on average with a reduction in the incidence of COVID-19. In detail, findings suggest beneficial effects of the closure of schools and workplaces for both countries. We also found that earlier implementation of restrictions results in a greater reduction in the incidence of COVID-19 for Italy. The same conclusion seems not to be shared by the United States.

Accuracy evaluation of our model prediction intervals shows good in-sample and out-of-sample performance. With respect to the latter, since our tvBINGARCHX model has also been employed within the ECDC European Forecast Hub initiative, low error with respect to 1–4 weeks ahead forecasts confirm the previous results.

### 6.1 | Literature comparison

Agosto et al. (2021) extended the log-linear Poisson autoregression (Geir et al., 2022) to epidemiologic contagion. The frequentist framework of our paper (Section 5.1) refers to this model, being the starting point of our analysis and the model against which comparisons have been made with the Bayesian framework. Unlike Agosto et al. (2021), policy covariates have been introduced.

The Bayesian framework instead introduces the time-varying models of Roy and Karmakar (2021) and evaluates their performances against the extended structure with time-varying policy covariates. Time constant coefficient models have also been added to have a direct comparison with Agosto et al. (2021) in the Bayesian framework. So, the model structure has been recovered and extended from these two contributions in terms of policy covariates.

On the other hand, in terms of policy evaluation, Bo et al. (2021) and Islam et al. (2020) have been recalled as a starting point for policy definition and evaluation. The effectiveness of NPIs was evaluated by Bo et al. (2021) in terms of time-varying reproduction number ( $R_t$ ) while Islam et al. (2020) evaluated policies in terms of counts, more similar to our approach. We confirm results of Bo et al. (2021) and Islam et al. (2020) on the effectiveness of distancing policies but we extend their work by introducing and quantifying the effectiveness of vaccines. Similar Islam et al. (2020), we introduce lag effects and a combination of policies, resulting in a better evaluation of policy effects.

### 6.2 | Strengths and limitations of this study

The strength of our study relies on both methodological extension and epidemic monitoring. The structure of our proposed methods shows both extensions of previous works in terms of time-dependence structure (i.e., time-varying coefficients) and exogenous components of the model (i.e., policy covariates). Our model formulation is more flexible since it allows to consider both time dependence structure in terms of autocorrelation with previous counts but also with exogenous covariates. Alternative time lags allow to capture the timing of the effect of public interventions on the trajectory of the pandemic. Furthermore, the time-varying coefficient framework allows both to model rapid changes of COVID-19 spread and address overdispersion issues.

At the same time, our study is also subject to several limitations. First, data quality could be an issue for forecasting models, especially in terms of the measurement of policy covariates. We relied on OxCGRT variables to represent government interventions, but it is really challenging to collect information on the exact date, nature, and extent of policies undertaken by the different countries. Measurement with error could result in higher model bias. The discussion could be extended in the same way with regard to pandemic counts since day-to-day reports are usually made up of errors. In the end, real-time forecasts of cases are really challenging, given data issues in COVID-19 databases and intrinsic model error together with changing policies, population behavior, and testing procedures.

### 6.3 | Implications for policy-makers

Quantifying the impact of NPIs is a critical step in ensuring a strong public health response to COVID-19 in the first phase of the vaccination campaign.

Risk factors for cluster formation are expected to be similar across countries, including closed and poorly ventilated locations together with crowded places where people have limited possibility for physical distancing. High-risk environments include venues, events, and activities with certain environmental conditions. As a result, the findings suggest beneficial effects of the closure of schools and workplaces for both countries, the latter representing a possible option for reducing transmission but with a large-scale economic and social impact.

Mobility plays an important role in human-to-human transmission. However, given the difference in terms of population size, limiting people's movement and gathering by requiring public transport closures in a combination of stay-at-home orders, and internal movement restrictions allows to decrease municipal (and regional) mobility fluxes and, consequently, the infection rate in Italy. On the other hand, requiring canceling gatherings and events at the stricter level (between 11 and 110 people or less) is effective in large countries such as the United States.

We also found that earlier implementation of restrictions results in a greater reduction in the incidence of COVID-19 in Italy. The impact on the trajectory of infection changes for the United States, being more immediate. However, more factors need to be accounted for to obtain more accurate conclusions for both countries.

While NPIs reduce transmission rates, vaccination primarily reduces symptomatic disease, hospitalization, and death from infection, as well as infection rates to a lesser amount. Our findings show that vaccination is effective on epidemic evolution but in combination with some NPIs, advocating for the need to keep NPIs in place during the first phase of the vaccination campaign.

However, our results regarding several health system interventions need to be cautiously interpreted. The noneffectiveness of some NPIs can be attributed to several factors. For example, if the outcome of the study is related to the number of cases, as case detection improves through more effective testing and tracing, the number of cases reported will increase without representing a real rise of cases. Finally, for both countries, we have found a reverse effect of some policies, consistent with the fact that policy measures are triggered by the increase in COVID-19 counts.

### 6.4 | Further studies

Further extensions of the work in the paper should consider the inclusion of finer policy measurements, as new databases become available; and, possibly, the inclusion of spatial dependence in the model, especially for large countries such as the United States, as longer time series become available. For all countries, including Italy and the United States, it would also be important to include policy and contagion data at the subnational and regional level, if available, to better capture spatial variation.

#### ACKNOWLEDGMENTS

This research has received funding from the European Union's Horizon 2020 research and innovation program "PERISCOPE: Pan European Response to the ImpactS of COvid-19 and future Pandemics and Epidemics" [101016233, H2020-SC1-PHE\_CORONAVIRUS-2020-2-RTD]. The project is the result of a close collaboration among the three authors; in addition, Paolo Giudici's main role has been in ideating and supervising the work; coding and model specification, and data analysis is mainly due to Barbara Tarantino together with Arkaprava Roy with regard to coding and revision of the paper. The Authors thank the European COVID-19 Forecast Hub of the European Center for Disease Prevention and Control, to which the forecasts based on this paper are sent every week, for useful comments, feedback, and discussions.

Open Access Funding provided by Università degli Studi di Pavia within the CRUI-CARE Agreement.


#### CONFLICT OF INTEREST

The authors have declared no conflict of interest.

#### DATA AVAILABILITY STATEMENT

The data from the case study are available as [Supporting Information](#) on the journal's web page.

## OPEN RESEARCH BADGES

 This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available in the [Supporting Information](#) section.

This article has earned an open data badge **“Reproducible Research”** for making publicly available the code necessary to reproduce the reported results. The results reported in this article could fully be reproduced.

## ORCID

Paolo Giudici  <https://orcid.org/0000-0002-4198-0127>

## REFERENCES

- Agosto, A., Campmas, A., Giudici, P., & Renda, A. (2021). Monitoring COVID-19 contagion growth. *Statistics in Medicine*, *40*(18), 4150–4160.
- Amorim, L., Cai, J., Zeng, D., & Barreto, M. (2008). Regression splines in the time-dependent coefficient rates model for recurrent event data. *Statistics in Medicine*, *27*, 890–906.
- Andrade, B. S., Andrade, M. G., & Ricardo, S. E. (2015). Bayesian GARMA models for count data. *Communications in Statistics: Case Studies, Data Analysis and Applications*, *1*, 192–205.
- Bo, Y., Guo, C., Lin, C., Zeng, Y., Li, H. B., Zhang, Y., Hossain, M. S., Chan, J., Yeung, D. W., Kwok, K. O., Wong, S., Lau, A., & Lao, X. Q. (2021). Effectiveness of non-pharmaceutical interventions on COVID-19 transmission in 190 countries from 23 January to 13 April 2020. *International Journal of Infectious Diseases*, *102*, 247–253.
- Bosse, N. I., Abbott, S., EpiForecasts, & Funk, S. (2020). scoringutils: Utilities for scoring and assessing predictions, <https://cran.r-project.org/web/packages/scoringutils/index.html>
- Cai, Z., Fan, J., & Yao, Q. (1999). Functional-coefficient regression models for nonlinear time series. *Journal of the American Statistical Association*, *95*, 941–956.
- Chan, K., & Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving counts. *Journal of The American Statistical Association*, *90*, 242–252.
- Cleveland, W. S., & Loader, C. (1996). Smoothing by local regression: Principles and methods. In *Statistical Theory and Computational Aspects of Smoothing* (pp. 10–49).
- Cowling, B. J., Ali, S. T., Ng, T., Tsang, T. K., Li, J., Fong, M. W., Liao, Q., Kwan, M. Y., Lee, S. L., Chiu, S. S., Wu, J. T., Wu, P., & Leung, G. M. (2020). Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: An observational study. *The Lancet Public Health*, *5*, 279–288.
- Dahlhaus, R., & Rao, S. (2006). Statistical inference for time-varying ARCH processes. *The Annals of Statistics*, *34*, 1075–1114.
- Davies, N. G., Kucharski, A. J., Eggo, R. M., Gimma, A., Edmunds, W. J., & Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group. (2020). Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: A modelling study. *The Lancet Public Health*, *5*, 375–385.
- Davis, R. A. (2003). Observation-driven models for Poisson counts. *Biometrika*, *90*, 777–790.
- Davis, R. A., & Mikosch, T. (2009). Extreme value theory for GARCH processes. *Handbook of Financial Time Series*, *90*, 187–200.
- Esteve, A., Permanyer, I., Boertien, D., & Vaupel, J. W. (2020). National age and coresidence patterns shape COVID-19 vulnerability. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 16118–16120.
- Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). Generalized linear models. In *Regression: Models methods and applications* (pp. 269–324). Springer Berlin Heidelberg.
- Ferland, R., Latour, A., & Oraichi, D. (2006). Integer-Valued GARCH Process. *Journal of Time Series Analysis*, *27*, 923–942.
- Ferreira, G., Navarrete, J. P., Rodríguez-Cortés, F. J., & Mateu, J. (2017). Estimation and prediction of time-varying GARCH models through a state-space representation: A computational approach. *Journal of Statistical Computation and Simulation*, *87*, 2430–2449.
- Fokianos, K., Rahbek, A., & Tjøstheim, D. (2008). Poisson Autoregression. Discussion Papers 08–35, University of Copenhagen. Department of Economics.
- Fryzlewicz, P., Sapatinas, T., & Rao, S. (2008). Normalized least-squares estimation in time-varying ARCH models. *Annals of Statistics*, *36*, 742–786.
- Geir, D., Berentsen, J. B., Maruotti, A., & Bård, S. (2022). Modelling clusters of corporate defaults: Regime-switching models significantly reduce the contagion source. *Journal of The Royal Statistical Society Series, C-Applied Statistics*, *71*(3), 698–722.
- Giordano, G., Colaneri, M., Di Filippo, A., Blanchini, F., Bolzern, P., De Nicolao, G., Sacchi, P., Colaneri, P., & Bruno, R. (2021). Modeling vaccination rollouts, SARS-CoV-2 variants and the requirement for non-pharmaceutical interventions in Italy. *Nature Medicine*, *27*, 993–998.
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, *102*, 359–378.
- Hale, T., Petherick, A., Phillips, T., & Webster, S. (2020). Variation in government responses to COVID-19. Blavatnik School of Government Working Paper, Blavatnik School of Government.
- Hao, X., Cheng, S., Wu, D., Wu, T., Lin, X., & Wang, C. (2020). Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature*, *584*, 420–424.
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning*. Springer Series in Statistics, Springer New York Inc.



- Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L. Y., Hultgren, A., Krasovich, E., Lau, P., Lee, J., Rolf, E., Tseng, J., & Wu, T. (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, *584*, 262–267.
- Huang, J., & Shen, H. (2004). Functional-Coefficient regression models for nonlinear time series: A polynomial spline approach. *Scandinavian Journal of Statistics*, *31*, 515–534.
- Islam, N., Sharp, S. J., Chowell, G., Shabnam, S., Kawachi, I., Lacey, B., Massaro, J. M., D'Agostino, R. B., Sr, & White, M. (2020). Physical distancing interventions and incidence of coronavirus disease 2019: Natural experiment in 149 countries. *BMJ: British Medical Journal*, *370*, m2743.
- Kedem, B., & Fokianos, K. (2002). *Regression models for time series analysis*, Wiley, London.
- Li, Y., Campbell, H., Kulkarni, D., Harpur, A., Nundy, M., Wang, X., Nair, H., & Usher Network for COVID-19 Evidence Reviews (UNCOVER) group. (2021). The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: A modelling study across 131 countries. *The Lancet Infectious Diseases*, *21*, 193–202.
- Liboschik, T., Fokianos, K., & Fried, R. (2017). tscount: An R package for analysis of count time series following generalized linear models. *Journal of Statistical Software*, *82*, 1–51.
- Marcon, G., Tettamanti, M., Capacci, G., Fontanel, G., Spanò, M., Nobili, A., Forloni, G., & Franceschi, C. (2020). COVID-19 mortality in Lombardy: the vulnerability of the oldest old and the resilience of male centenarians. *Aging*, *12*, 15186–15195.
- Neal, R. (2011). MCMC Using Hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, *2*, 2.
- Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Series A, General*, *135*, 370–384.
- Roberts, G. O., & Rosenthal, J. S. (1998). Optimal scaling of discrete approximations to Langevin diffusions. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, *60*, 255–268.
- Rohan, N., & Ramanathan, T. (2013). Nonparametric estimation of a time-varying GARCH model. *Journal of Nonparametric Statistics*, *25*, 33–52.
- Roy, A., & Karmakar, S. (2021). Time-varying auto-regressive models for count time-series. *Electronic Journal of Statistics*, *15*, 2905–2938.
- Zeger, Scott L. (1988). A regression model for time series of counts. *Biometrika*, *75*, 621–629.
- Zhu, F. (2012). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *Journal of Mathematical Analysis and Applications*, *389*, 58–71.
- Zhu, F. (2012). Modeling time series of counts with COM-Poisson INGARCH models. *Mathematical and Computer Modelling*, *56*, 191–203.
- Zhu, F. (2012). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Statistical Planning and Inference*, *142*, 826–839.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Giudici, P., Tarantino, B., & Roy, A. (2022). Bayesian time-varying autoregressive models of COVID-19 epidemics. *Biometrical Journal*, 1–23. <https://doi.org/10.1002/bimj.202200054>

## APPENDIX: ABBREVIATIONS

AR	AutoRegressive
ARX	AutoRegressive eXogenous
ARMA	autoregressive moving average
BAR	Bayesian AutoRegressive
BARX	Bayesian AutoRegressive eXogenous
BINGARCH	Bayesian INGARCH
BINGARCHX	Bayesian INGARCH eXogenous
ECDC	European Centre for Disease Prevention and Control
GLM	generalized linear model
HMC	Hamiltonian Monte Carlo
INGARCH	Integer-Valued Generalized Autoregressive Conditional Heteroskedasticity
INGARCHX	Integer-Valued Generalized Autoregressive Conditional Heteroskedasticity eXogenous
LMC	Langevin Monte Carlo
MAPE	median absolute percentage error
MCEM	Monte Carlo expectation maximization
MCMC	Markov chain Monte Carlo



---

NPIs	Non-pharmaceutical interventions
OxCGRT	Oxford Covid-19 Government Response Tracker
PAR	Poisson AutoRegressive
SIR	Susceptible-infected-recovery
tvBAR	time-varying Bayesian AutoRegressive
tvBARX	time-varying Bayesian AutoRegressive eXogenous
tvBINGARCH	time-varying Bayesian INGARCH
tvBINGARCHX	time-varying Bayesian INGARCH eXogenous