

# Application of Markov-Switching-Dynamic-Regression Model for COVID-19 Reproduction Rate: A Case study of Zimbabwe

C. Shoko\*, K. Makatjane and G. Madisa-Maseko

Department of Statistics, Faculty of Social Sciences, University of Botswana, Gaborone, Botswana

Received: 12 May 2025, Revised: 22 Jul. 2025, Accepted: 16 Aug. 2025.

Published online: 1 Sep. 2025.

**Abstract:** In this study, we develop a Markov Switching Dynamic Regression (MS(k)-DR(p)) model for the regime switches based on the reproduction rate in real-time. The effects of the stringency index, daily COVID-19 deaths and total vaccinated individuals on the regimes are also analysed. Before the fitting of the MS(k)-DR(p) model, this study analysed the cointegration of the response variable and the explanatory variable using the Johansen procedure. Results from analysis show that the peaks and reproduction rate can best be described by shifts between two regimes, MS(2)-DR(1). The first regime is defined by high reproduction number and the second regime is identified by a low reproduction rate. The MS(2)-DR(1) model for the reproduction rate could only be explained by the stringency index and the daily COVID-19 deaths. An increase in stringency index results in a decrease in the reproduction rate for regime 2 and the opposite is true for regime 1. The developed model closely tracks regime changes caused by changes in the stringency index that include lock-down, mandatory mask-wearing, social distancing, etc. Thus, the MS-DR model is a useful policy tool for monitoring interventions by the public health sector in controlling epidemics that have the same behaviour as COVID-19.

**Keywords:** Reproduction rate, COVID-19, MS(k)-DR(p) model, Cointegration, Zimbabwe.

## 1 Introduction

The epidemiologist, Klaus Dietz (Dietz, 1975), was the first to give an explicit definition of  $R$ . He defined it as “the expected number of infections or secondary cases generated by a typical infected individual.” However,  $R$  can be defined in two different ways depending on the stage of the infection. In a scenario where the whole population is susceptible, this number is called the

basic reproduction number, or  $R_0$ , and it shows how many secondary infections resulted from a single case at the start of the epidemic. However, at later stages of the pandemic,  $R$  is referred to as the effective reproductive number and is denoted as  $R(t)$ . The reproduction number  $R(t)$ , can be a cohort  $R$  also known the instantaneous  $R$ , as it is computed for each time  $t$

since the start of the epidemic (Gostic et al., 2020).  $R(t)$  represents a disease's capacity to spread and as such it is a key diagnostic tool for monitoring the dynamics of COVID-19 (Luis Rosero-Bixby and Tim Miller, 2022) and also to assess the effects of intervention strategies that include a stringent index, vaccines, quarantine, among others.

The COVID-19 virus's reproduction number,  $(R(t))$ , which is the average number of new infections produced from a population made up of exposed/immune and naïve/susceptible individuals, is a measure of a virus's transmissibility. The reproduction number is the ratio of new infections at the current time  $t$ , over the number of new infections in the interval  $\Delta t$  and it represents the instantaneous proliferation rate of an infection. A reproduction number greater than 1 (Rinaldo et al., 2020), implies that the number of infected subjects is increasing leading to an epidemic. Whereas a reproduction number below 1, is an indication that the transmission is likely to reach extinction, and when the reproduction number is 1, the disease is said to be endemic. The basic reproduction number, which indicates an infectious agent's likelihood of causing an epidemic, is a fundamental concept in infectious disease epidemiology (Liu et al., 2020). Thus,  $R(t)$ , together with other indicators, form the basis for characterizing unfolding outbreaks (Pasetto et al., 2020).

To control the spread of COVID-19, stringency measures were put in place. This was measured based on the stringency index (SI) ranging from 0 to 100. The SI suggests that lockdown procedures are rigorous, primarily to limit human conduct. Schools, public spaces, and workplaces were all be closed as part of this SI. Public events were also be canceled, gatherings and international travel were restricted, and stay-at-home rules were enforced. A SI of 0 means no restrictions or lockdowns and a

\*Corresponding author e-mail: [shokoc@ub.ac.bw](mailto:shokoc@ub.ac.bw)

SI of 100 means harsh restrictions and lockdown policies (Shah et al., 2021). According to Makki (2020), the assumption is that increasing the SI implies limited contact between the infected and susceptible populations and this results in the reduction of the reproduction rate. In addition to the stringency measures were the development of the vaccines that help in strengthening the immune system to fight COVID-19 infections. The assumption is that as the population of the vaccinated increases, the reproduction rate is likely to decrease. On the other hand, the death of individuals from the infected population reduces the size of the infected population leading to a reduction in the reproduction rate (Mezencev and Klement 2021).

Based on the ups and downs nature of the COVID-19 pandemic, Haimerl and Tobias (2023) developed a regime-switching approach to model the trend in COVID-19. For each day of the observation period, their suggested model divided the trend into two regimes: an infection that was either rising or falling. However, Oliveira et al (2021) discovered that the occurrence of a new more contagious wave of COVID-19 for the year 2021 spread faster than the first outbreak in spring 2020 resulting in populations in member states around the WHO European Region beginning to exhibit signs of pandemic weariness. Given the protracted nature of this catastrophe and the resulting annoyance and misery, pandemic weariness is a normal and expected response. It presents a significant risk to attempts to stop the virus's spread; for a policy framework for reviving the populace to stop the pandemic, see WHO (2020).

In a similar vein, evidence presented by Maragakis (2020) indicates that after the initial wave of illness, people's commitment to the non-pharmacological COVID-19 prevention methods of mask wearing, hand washing, and social distancing decreased. Additionally, Shafer et al. (2021) confirmed that easing social distancing to contact levels that are 50% of what they were before COVID-19 may result in over 35% of the population being infected at the same time, while McGrail et al. (2020) showed that social distancing can lead to an estimated 65% reduction in new COVID-19 cases. As an example, they cited the circumstances in Manitoba, Canada. These findings corroborate the claim by WHO (2020) on the beginning of the spread of pandemic fatigue.

The dynamics in the daily case numbers of COVID-19 have been studied using a range of econometric techniques; however, unobserved components (UC) models have been particularly effective in predicting the previously described features of COVID-19-related data. The UC models were applied early in the pandemic to detect structural breakdowns and fit linear deterministic trends to COVID-19 case data (Hartl et al. 2020; Lee et al. 2021; Liu et al. 2021). While Navas Thorakkattile et al. (2022) and Xie (2022), among others, have included (stationary) seasonal components, Moosa (2020) and Doornik et al. (2022) have studied UC models with stochastic trends as the number of available data grows. It is still difficult to correctly account for the COVID-19 case numbers trend's alternating peaks and troughs.

By specifically modeling the up and down pattern of the COVID-19 reproduction numbers—which has become one of the defining characteristics of the COVID-19 pandemic—we add to the UC literature with these data. Hence, we introduce a Markov-switching dynamic regression (MS-DR) that is adept at handling situations where the underlying data-generating process changes over time. This is particularly relevant for COVID-19, where the pandemic has led to significant shifts in people's behaviour, policy responses, and public health outcomes. By allowing for different regimes (or states) in the data, our approach can accurately reflect periods of high and low transmission rates or varying levels of public compliance with health measures.

The ability to switch between different states according to Rossouw et al (2021), enables better predictions of future trends based on historical data. For instance, MS-DR can capture the volatility and abrupt changes in infection rates or economic indicators associated with COVID-19 waves, leading to more reliable forecasts than traditional static models (Bouteska et al 2023). This is crucial for policymakers who need to make informed decisions based on expected future scenarios. Markov switching dynamic regression models provide a powerful tool for analyzing COVID-19 data by accommodating regime shifts, improving forecasting accuracy, offering insights into transition dynamics, capturing volatility, and being applicable across diverse fields of study. In the past, regime-switching and mixture models have been thoroughly investigated. For a summary, see Frühwirth-Schnatter (2006); Kim and Nelson (2017). According to Kim (1994); Kim and Nelson (2017), the Kim filter is an extension of the Kalman filter to regime-switching models, and we use it to estimate the trend, seasonal, and cyclical components. By using Hamilton's (1989) recursions to estimate the regime probabilities between the Kalman filter's prediction and updating steps, this enables regime switching in a state-space context.

Our second contribution is the application of Johansen Cointegration and Granger causality test. In this way, we can identify whether influential variables and total number of COVID-19 cases share a long-term equilibrium relationship. This is crucial, as many variables tend to move together over time despite short-term fluctuations. Recognizing these long-term relationships allows researchers to model and forecast more accurately, avoiding spurious correlations that can arise from non-stationary data where we can establish if the number of COVID-19 cases and the reproduction rate are both non-stationary but cointegrated. This suggests that there is a long-term equilibrium relationship between these two variables and implies that as public health measures change over time, the infection rates will adjust to maintain this relationship.

While Granger causality, helps us to determine whether different policy interventions can predict the daily reproduction rates.

In the context of this study, if we find that past values of vaccination rates cause future infection rates, this suggests that changes in vaccination strategies directly impact on controlling the spread of COVID-19. Finally, causality tests require establishing that the cause precedes the effect. In health data analysis, this is crucial for identifying effective interventions. For instance, if increased testing leads to reduced infection rates after some time, it supports the hypothesis that testing is an effective public health intervention. By incorporating cointegration and causality tests into the analysis of COVID-19 health data enables a deeper understanding of the complex interactions among various determinants. This approach not only enhance theoretical insights but also provides practical guidance for effective public health strategies.

## 2 Materials and Methods

In this study, we use a Markov-Switching Dynamic Regressive (MS-DR) model to assume the behavior of stochastic parameters. This section of the study discusses the techniques and methodologies. The findings in this part will be used as a guide to determine the kind of model to estimate and the sort of data to use.

### 2.1 Markov Switching Dynamic Regression Model

Decomposing nonlinear time series into a limited sequence of discrete stochastic processes, states, or regimes is the basic concept of Markov Switching Models (MSM). This allows the parameters to take on different values based on the state or regime that is in effect at a given time  $t$ . The transitions between the states or regimes are caused by an unobservable regime variable denoted by  $S_t$ , which is believed to evolve according to a Markov Chain. Given the time series say,  $\{Y_t : t = 1, 2, 3, \dots, n\}$ , the dynamic regression is therefore represented as follows

$$Y_t = \mu + X_t' + \varepsilon_t \quad (1.1)$$

where,  $\varepsilon \sim N[0, \sigma^2]$  and  $X_t \in \mathcal{R}^k$  is vector of explanatory (exogenous or lagged endogenous) variables. The numbering of the regime is arbitrary; hence, the mean of Model (1.1) can be written as a function of  $S_t$  is expressed as the output of the identity link function  $\eta(\cdot)$  which is now given as  $\mu_{ijs_t} = \eta(\cdot)$ . If it is assumed that the Markov process operates over the set of  $k$  states  $[1, 2, \dots, k]$ , then, it is expressed as

$$\mu_{ij} = \eta(\cdot) = \beta_{ij}, S_t = i, j, i = 0, j = 1 \quad (1.2)$$

In this case, the transition probabilities between the  $k$  states is given by

$$p_{i|j} = P[S_{t+1} = i | S_t = j], i, j = 0, \dots, k-1 \quad (1.3)$$

and as a result, the likelihood of transitioning from stage  $i$  in one period to stage  $j$  in the next is solely determined by the prior condition. Because the system must be in one of the  $i, j$  states we now show that

$$\sum_{i=0}^{k-1} p_{i|j} = 1 \quad (1.4)$$

where the full matrix of transition probabilities  $P$  is

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

such that  $p_{11} + p_{21} = 1$  and  $p_{12} + p_{22} = 1$  (Xaba et al 2019), and if the value of  $p_{ij}$  is small, the system will stay longer in state  $i$  than in state  $j$ . The duration of this state is expected to be  $\frac{1}{p_{ij}}$ . Nonetheless, Geraldi et al. (2021) emphasised that the number of regimes can be  $r \geq 2$  and for our case, we utilise the simple and effective MS approach, which eliminates a great

deal of empirical estimation issues. Our proposed MS-DR specification follows the dynamic regression model in the specification of the dynamics that

$$Y_t = v(S_t) + \alpha X_t' + X_t' \beta + \varepsilon_t, S_t = i, j, i = 0, j = 1 \quad (1.5)$$

We finally present a Markov-switching dynamic regression as follows:

$$Y_t = \begin{cases} \alpha_0 + \sum_{i=1}^p \gamma_{0,i} X_{t-i} + \varepsilon_{1,i}, S_t = 0 \\ \alpha_1 + \sum_{i=1}^p \gamma_{1,i} X_{t-i} + \varepsilon_{1,i}, S_t = 1 \end{cases} \quad (1.6)$$

where,  $\varepsilon_t$  is independently and identically distributed (i.i.d) with mean 0 and variance  $\sigma_t^2$ . The state variable is governed by a first-order Markov chain.

## 2.2 Unit root test

Moroke (2014) confirmed the non-stationarity of several macroeconomic data. Therefore, to make these series stationary, they must be rendered stationary and this is one of the prerequisites for cointegration analysis. It is necessary to prove that the variables have the same integration order. Kennedy (1996) states that if a variable need to be differenced  $d$  times in order to become stable, it is integrated of order This study uses the Augmented Dickey-Fuller (ADF), Kwiatkowski-Phillips-Schmidt-Shin (KPSS), and Box Ljung tests to first examine the series stationarity properties.

**Augmented-Dick-Fuller (ADF) test:** This test is used to check for stationarity of variables and is based on the null hypothesis that assumes non-stationarity. As suggested by Dickey and Fuller (1981), the following regressions are estimated

$$\Delta Y_t = \rho Y_{t-1} + \sum_{i=1}^p \delta \Delta Y_{t-i} + \varepsilon_t \quad (1.7)$$

where the model with intercept is estimated as follows

$$\Delta Y_t = \beta_0 + \rho Y_{t-1} + \sum_{i=1}^p \delta \Delta Y_{t-i} + \varepsilon_t \quad (1.8)$$

and finally, the model with intercept together with trend is given by

$$\Delta Y_t = \beta_0 + \beta_1 t + \rho Y_{t-1} + \sum_{i=1}^p \delta \Delta Y_{t-i} + \varepsilon_t \quad (1.9)$$

In Model (1.7) to Model (1.9),  $\Delta$  is a differencing operator,  $t$  is time drift;  $\rho$  denotes the selected maximum lag based on the minimum criteria such as Akaike's information criteria (AIC), Schwartz Bayesian criteria (SBC) or Hannan-Quin criteria (HQC) values and  $\varepsilon_t$  is the error term. Makatjane and Moroke (2016) disclosed that  $\beta$ 's and  $\delta$ 's are model bounds. Depending on the findings, the intercept, and intercept plus trend may be included in the model. Finally, the ADF test is then defined as

$$\hat{\tau} = \frac{\hat{\gamma} - \gamma_0}{se(\hat{\gamma})} \sim t_{\alpha, n-p} \quad (1.10)$$

where,  $\hat{\gamma}$  is the process root coefficient. If the observed probability value is less than the calculated probability value, we fail to reject the null hypothesis of non-stationary time series.

**Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test:** The test is based on the null hypothesis that assumes a stationary series. The KPSS test statistic is given as

$$KPSS = n^{-2} \sum_{t=1}^n \frac{S_t}{\hat{\sigma}^2} \quad (1.11)$$

where  $S_t = \sum_{i=1}^t e_i$  and  $\hat{\sigma}^2$  is an estimate of the residuals' long-run variance. If the KPSS is greater than the critical value,

indicating that the series deviates from its mean, the research rejects the null hypothesis. In order to make the variables stationary, first order differencing is used. Since the data are gathered daily, a lag of up to 7 is incorporated to account for autocorrelation. The stationarity requirement is met if the variables are both stationary and integrated of I(1).

**Box Ljung test:** This test statistic looks at the presence of auto-correlation. Correlation analysis assesses the strength and direction of a relationship between two indices. Risk analysts try to find if absolute or squared returns are highly autocorrelated. This objective can be achieved by using partial autocorrelation functions (PACF) and autocorrelation functions (ACF) (Montgomery et al, 2015). Jonathan and Kung-Sik (2008) describe the ACF as

$$\rho_k = \frac{\rho_k}{\rho_0} = \frac{\text{cov}(X_t, X_{t-k})}{\text{var}(X_t)}, k = 1, \dots \quad (1.12)$$

Where  $\rho_0$  is the autocovariance computed as  $\text{Cov}(X_t, X_t) = \text{Var}(X_t)$ . Normally, a correlogram is used and it is a graph of  $\rho_k$  against  $k$ . A PACF is further defined as follows

$$a_{kk} = \begin{cases} \rho_1, k = 1 \\ \frac{\rho_k - \sum_{j=1}^{k-1} a_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} a_{k-1,j} \rho_j}, k = 2, \dots, \rho \\ 0, k > 1 \end{cases} \quad (1.13)$$

For a given sample of a time series is  $\{r_t\}_{t=1}^T$ , let  $\bar{r}$  be a sample mean, then a lag-1 sample autocorrelation of  $r_t$  is

$$\hat{\rho}_1 = \frac{\sum_{t=2}^T (r_t - \bar{r})(r_{t-1} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2} \quad (1.14)$$

Under some general conditions, Model (1.14) is a consistent estimator for  $\rho_1$ . However, if  $r_t \sim \text{I.i.d}$  sequence, and  $E(r_t^2) < \infty$ , then Tsay (2014) suggested that  $\rho_1$  is asymptotically normal with mean zero and variance  $\frac{1}{T}$ . This result in practice tests the following hypothesis

$$\begin{aligned} H_0: \rho_1 &= 0 \\ H_1: \rho_1 &\neq 0 \end{aligned}$$

A test statistic is the usual  $t$  ratio, which is  $\sqrt{T\hat{\rho}_1}$ , and it is asymptotically a standard normal distribution. In general, the lag -  $\ell$  sample autocorrelation of  $r_t$  is

$$\hat{\rho}_\ell = \frac{\sum_{t=\ell+1}^T (r_t - \bar{r})(r_{t-\ell} - \bar{r})}{\sum_{t=1}^T (r_t - \bar{r})^2}, 0 \leq \ell < T-1 \quad (1.15)$$

In finite samples,  $\hat{\rho}_\ell$  is a biased estimator of  $\rho_\ell$ . For a small sample size, that is  $T < 30$ ,  $\frac{1}{T}$  causes this bias and this can be overcome by making sure that  $T > 30$ . Therefore, in practice,  $\hat{\rho}_1$  is plotted against time, and Shumway and Stoffer (2017) revealed that if the spikes of the plotted ACF and PACF fall beyond the control bands of the plots, this indicates highly correlated returns at different lags, and Tsay (2014) exhibited that there is no correlation if  $\hat{\rho}_\ell = 0, \ell > 0$ .

Therefore, Ljung and Box (1978) propose the following Portmanteau statistic

$$Q^*(m) = T \sum_{\ell=1}^m \hat{\rho}_\ell^2 \quad (1.16)$$

as a test statistic for the following hypothesis

$$\begin{aligned} H_0: \rho_1, \dots, \rho_m &= 0 \\ H_1: \rho_1, \dots, \rho_m &\neq 0 \end{aligned}$$

Under the assumption that the series is i.i.d sequence, with certain moment conditions. Tsay (2010) showed that model (1.17) is asymptotically a chi-squared random variable with  $m$  degrees of freedom that increase the power of a test statistic in Model (3.17). Because of finite samples, Ljung and Box (1978) modified model (1.17) to

$$Q(m) = T(T+2) \sum_{\ell=1}^m \frac{\hat{\rho}_{\ell}^2}{T-\ell} \quad (1.17)$$

The null hypothesis is rejected if the observed probability value is lower than the conventional probability value and conclude that the returns series are highly correlated.

### 2.3 Cointegration Test

The cointegration test is used to confirm whether the variables move together or not in the long run. The test requires all variables to be integrated in the same order. In this study, we used the Johansen procedure. Four variables are considered, the cointegration of the explanatory variables: stringency index (SI), total vaccinated (TV), and number of deaths (ND) on the response variable reproduction rate (RR).

The maximum Eigenvalue and the Trace test are used to determine these vectors. These test statistics test the null hypothesis of  $r$  cointegrating relationships

$$H_0 : \lambda_i = 0 \quad \text{for } i = r+1, \dots, k \quad \text{VS} \quad H_a : \lambda_i \neq 0$$

for  $r = 0, 1, \dots, n-1$ . According to Fountis and Dickey (1989), the four steps below are suggested to cater for multivariate analysis:

**Step 1:** Fit linear multivariate time series as

$$X_t = \Phi_1 X_{t-1} + \dots + \Phi_{\rho} X_{t-\rho} \quad (1.18)$$

**Step 2:** Compute the largest eigenvalue,  $\lambda_{\max}$ , based on the following characteristic equation

$$\left| \lambda^{\rho} I - \Phi_1 \lambda^{\rho-1} - \dots - \Phi_{\rho} \right| = 0 \quad (1.19)$$

where,  $I = \rho \times \rho$  matrix.

**Step 3:** Test of the null hypothesis of unit root is based on the following test statistic

$$\hat{\tau}_{mfd} = N[\lambda_{\max} - 1] \quad (1.20)$$

where  $\lambda_{\max}$  is computed using model (1.20).

Step 4: At the given significance level, particularly 5% significance level, obtain the critical value from the table. Reject the null hypothesis if the calculated test statistic is greater than the tabulated value or alternatively, if the calculated probability is less than the level of significance. With respect to Moroke (2014), the following are the Johansen trace and maximum eigenvalues formulae also calculated through these four steps

$$J_{\text{Trace}} = -N \sum_{i=r+1}^n \ln(\hat{\lambda}_{\max}) \quad (1.21)$$

$$J_{\max} = -N \ln(1 - \hat{\lambda}_{\max}) \quad (1.22)$$

The critical values are listed in the Johansen and Juselius (1990) tables. The null hypothesis for these statistics are rejected when the observed values exceed the critical thresholds. This will imply that the variables are cointegrated and that there is a long-run link, as proposed by Sjö (2008). Osterwald-Lenum (1992) provides detailed tables containing these crucial values. It is crucial to note that Alexander (2001) said that these two tests might provide different outcomes. However, after the cointegration link has been validated, the study proceeds to construct a VECM, which is described below.

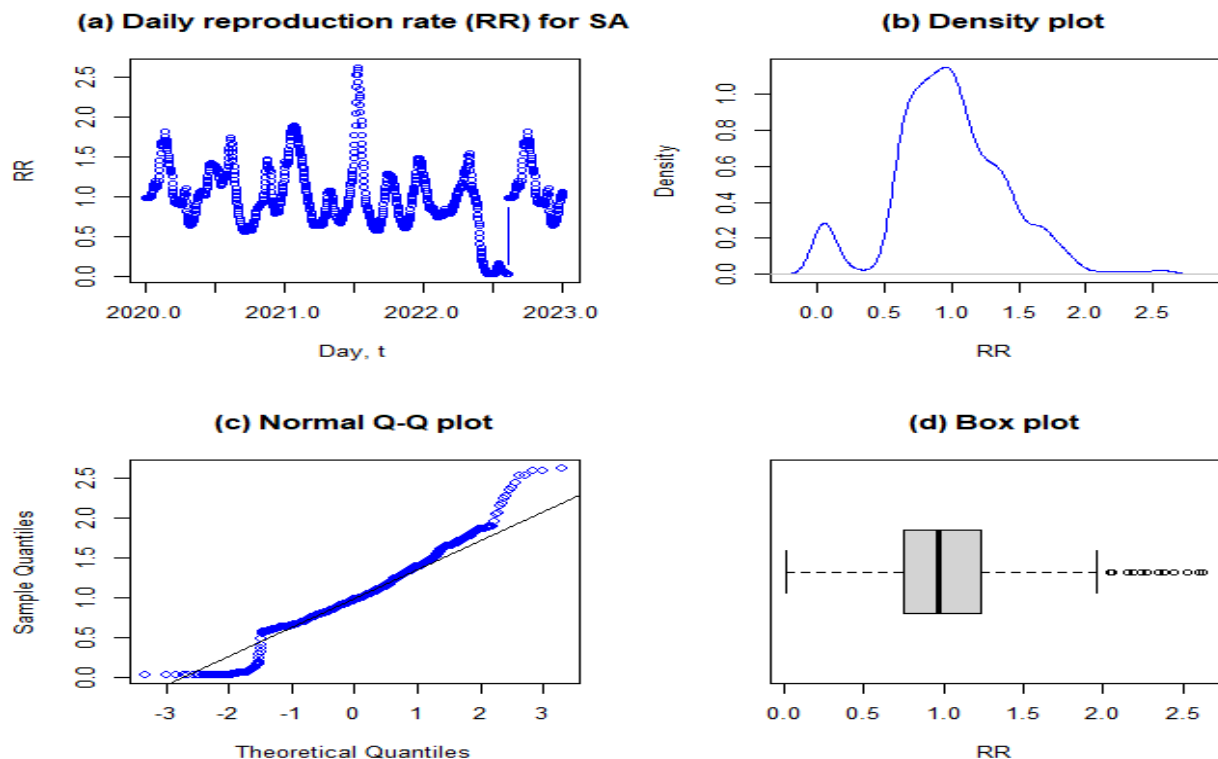


### 3 Results and Discussion

This section presents empirical data analyses using daily-confirmed reported cases of corona virus in Zimbabwe for the period of 28 May 2020 to 31 December 2022. We cast—off a Markov-Switching Dynamic Regressive (MS-DR) to capture regime shifts of COVID-19 reproduction rate in Zimbabwe.

#### 3.1 Exploratory data analysis

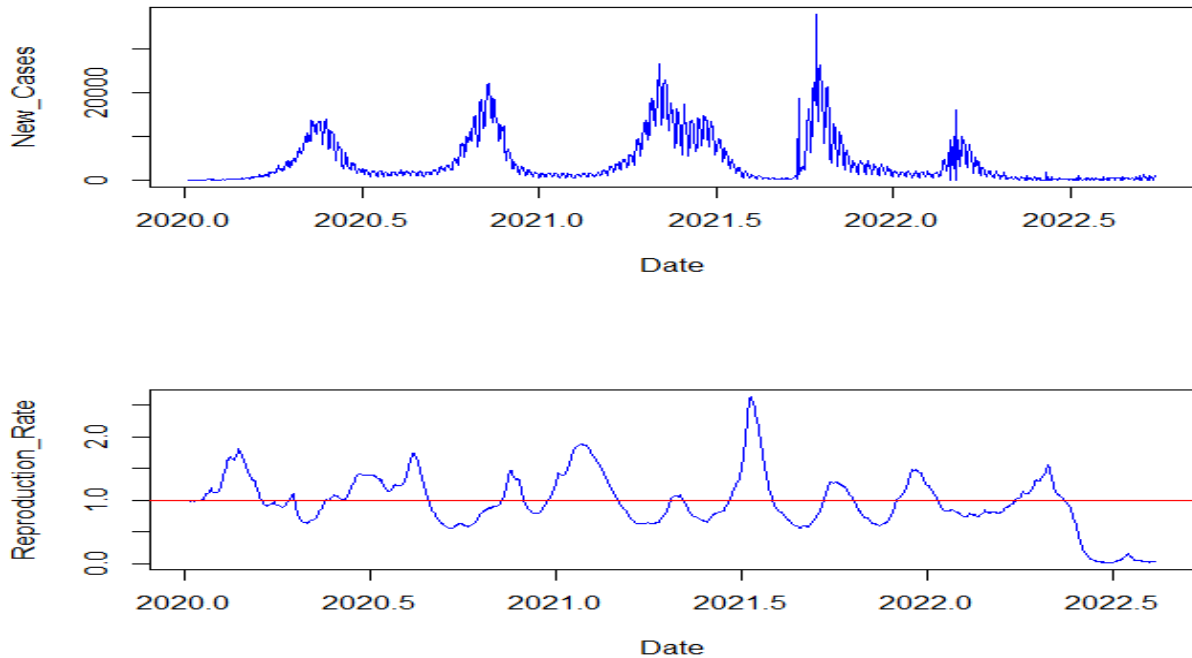
Based on Figure 1, a time series plot in the top-left panel shows both upward and downward trends in conjunction with seasonality; with the highest peak around June 2021. This is a clear indication that the series is not stationary. The quantile-quantile (Q-Q) plot on the bottom-left panel of Figure 1 also confirms these results. The Q-Q plot indicates that the distribution of the daily reproduction rate is deviating from normal distribution indicating that the series follows a fat-tailed distribution. The fat-tail conclusion is further confirmed by the density plot (top-right plot) which clearly shows a non-normal distribution. In addition, the Box and Whiskers plot shows that the distribution of the reproduction rate is positively skewed with some outliers which shows existence of extremely high values. The same results of leptokurtic coronavirus spread are also found by Wong and Collins (2020) and Shoko and Sigauke (2023) where the authors were therefore interested in determining whether the empirically observed distribution of  $Z$  for coronavirus spread exhibited an exponential tail. Rather, Wong and Collins (2020) used three complementing techniques: 1) a Zipf plot; 2) a meplot; and 3) statistical estimators of the tail index to find that the tail is consistent with fat-tail behavior. In their investigation on the extreme quantiles method to COVID-19 dissemination in South Africa, Shoko and Sigauke (2024) also discovered fat-tail behavior utilising the Box and Whiskers plot, normal Q-Q plot, and density plot.



**Fig. 1:** Time series and Quantile—quantile plot for Confirmed Daily corona virus Cases

In Figure 2 the time series plots for the new COVID-19 cases and the reproduction rate are presented. The reproduction rate also follows an up-and-down trend. During the period of study, the reproduction rate was mostly above 1 with the highest reproduction rate of more than 2.5. Relating the two graphs shows that the peaks in daily COVID-19 cases follow after an increase in the reproduction rate and vice-versa. This implies that when the reproduction rate ( $R$ ) rises above 1, it signals that each infected person is, on average, transmitting the virus to more than one other person, leading to a potential surge in cases. This rise in cases strained healthcare systems, especially when protective measures were relaxed too quickly. High case counts may lead to resource shortages, impacting patient care for both COVID-19 and non-COVID conditions.

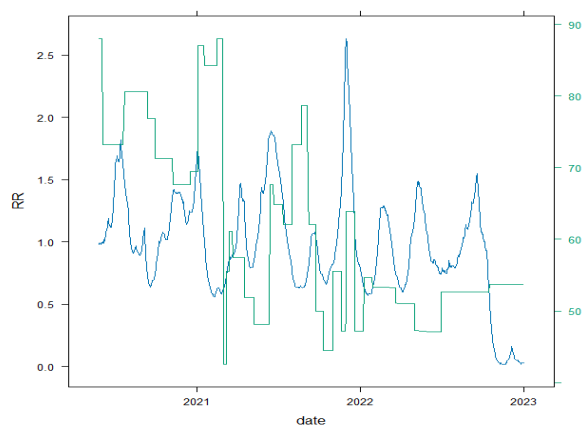
Furthermore, the cyclic nature of COVID-19 can prompt authorities to implement adaptive strategies, such as reintroducing mask mandates, social distancing, or travel restrictions during peak periods. This "ebb and flow" approach helps to prevent overwhelming the healthcare system but may also cause public fatigue or resistance, complicating long-term compliance. Finally, these fluctuations also affect mental health and economic stability. Repeated cycles of rising cases can lead to prolonged psychological stress and anxiety within the community, especially among high-risk groups. Economically, ongoing waves of infection can disrupt business continuity, employment, and global supply chains. Cutler (2023) further emphasised that Health systems may need to scale up contact tracing, testing, and isolation measures during periods when the reproduction rate and case numbers rise. This is resource-intensive and can be challenging to maintain, particularly in lower-resourced areas, leading to potential lapses in containment efforts



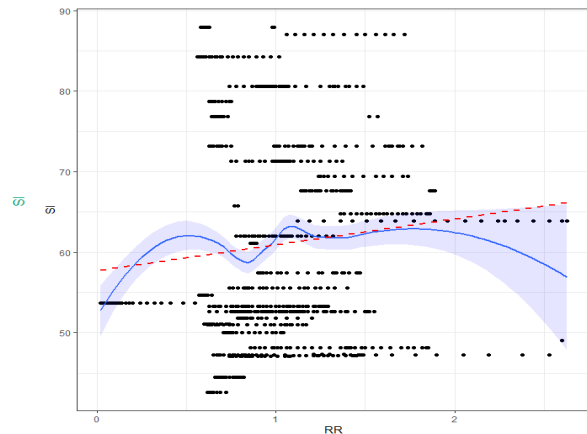
**Fig. 2:** Time series plots for the new COVID-19 cases and the reproduction rate.

We present plots of two time series, reproduction rate, and each one of the explanatory variables, stringency index (SI), new deaths (ND), and total vaccines (TV), in one. The correlation plot between each of the two variables is also presented.

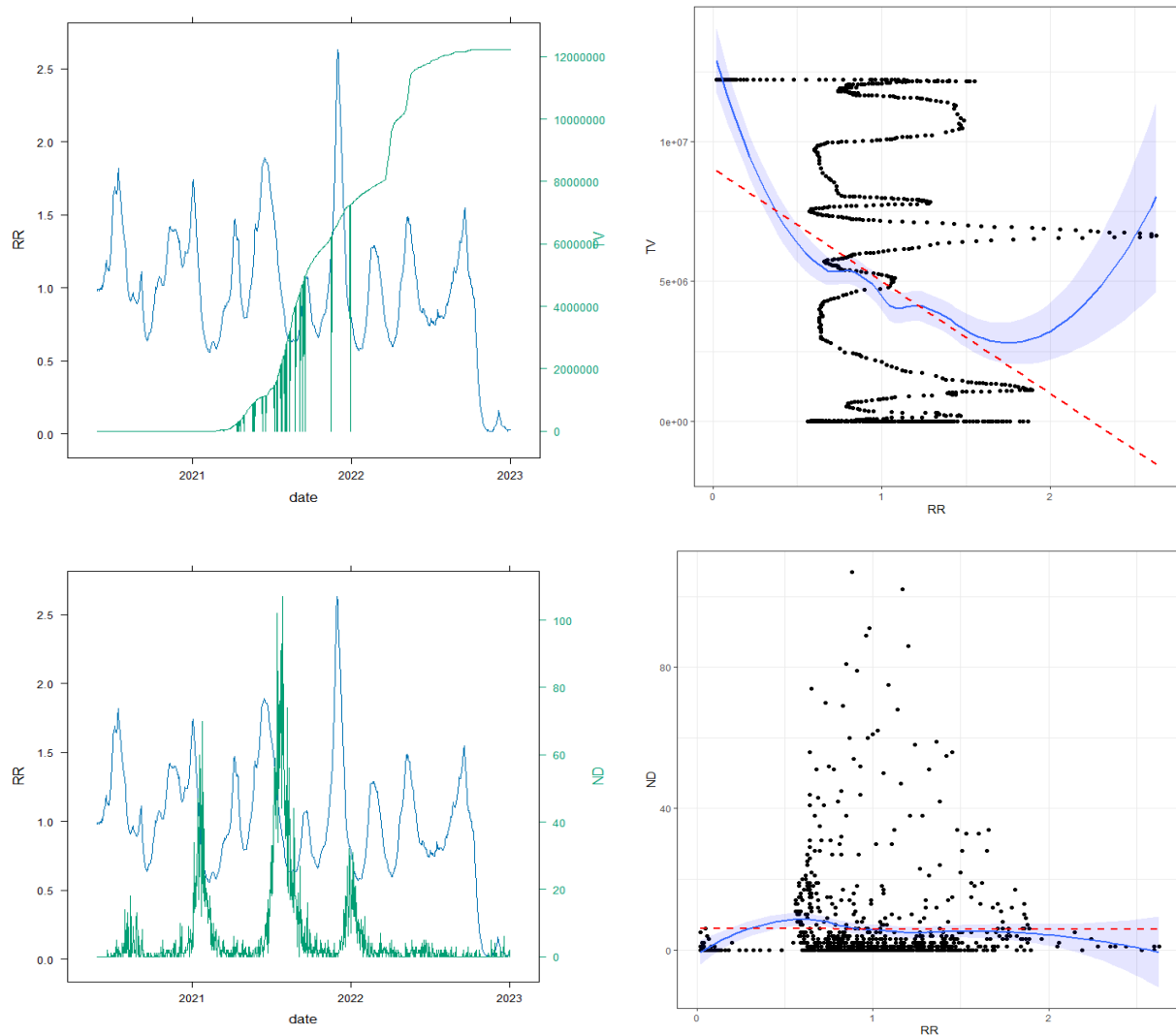
a. RR vs each of the explanatory variables



b. Correlation: RR & covariates







**Fig. 3:** a. Time series plots for the reproduction number and each of the variables of interest (SI, TV, ND, respectively), b. Correlation between RR and each of the covariates.

Results from Figure 3 show a long-term relationship between the reproduction number and the stringency index. High reproduction number is characterised by the low stringency index. The corresponding correlation plot show a positive correlation between the two variables. In the middle we have the time series plot for the total vaccines and the reproduction number. The plots show that in the initial stages, which has fewer number of the vaccinated population, the reproduction rate was relatively although it shows fluctuations due the tightening and loosening of the stringency index. As the size of the vaccinated population increased a reduction in the reproduction rate was noted from over 2.5 towards the end of year 2022 to 0 in the year 2023. The correlation plot show a negative correlation between the size of the vaccinated population and the reproduction rate. On the other hand, daily COVID-19 deaths follow after a peak in the infectious rate and this trend is consistent throughout the observation period.

Before we proceed with the main analysis, the unit root tests are applied. This is to ensure that the assumptions of the Johansen test are met. This is done using the ADF test, Elliot-Rothenberg-Stock test, KPSS test, and the Box Ljung test. The targeted variables are reproduction rate (RR) and the explanatory variables are SI, ND, and TV and the results are reported in Table 1. The p-values of the ADF test are less than 5% for RR and ND rejecting the null hypothesis of the existence of unit root and concluding that both variables are stationary at zero differencing. However, for ND and SI, the p-values are greater than 5% we therefore, fail to reject the null hypothesis of unit root and conclude that the series are non-stationary hence differencing is recommended. It is worth noting that with the KPSS all the variables are stationary at 5% level of significance. For RR, TV, ND, and SI, the Box-Ljung statistics are all extremely high with very low p-values ( $< 2.2e-16$ ), suggesting that all series exhibit strong serial correlation.

**Table 1:** Unit root test for the RR, TV, and ND time series

|    | ADF-test at lag 10 | KPSS-test at lag 7 | Box-Ljung-test     |
|----|--------------------|--------------------|--------------------|
| RR | -4.9814(0.01)      | 0.1164(0.1)        | 2146.9(< 2.2e-16)  |
| TV | -0.5507 (0.9795)   | 1.7356(0.01)       | 2132.3 (< 2.2e-16) |
| ND | -3.9024 (0.01387)  | 0.5428 (0.01)      | 1549.9 (< 2.2e-16) |
| SI | -2.4529 (0.1028)   | 1.7992 (0.01)      | 2077.1 (< 2.2e-16) |

**Note:** Values in () are the Probability values for all the tests

In the next step, this study applies the Johansen procedure to test for cointegration. The Johansen test can detect multiple cointegration vectors and is more appropriate for multivariate analysis than other approaches (Zikidou and Hadjidema, 2020). Wassell and Saunders, (2008) further argued that the Johansen procedure is a desirable property in that it treats every test variable as endogenous variable. The outcome from the Johansen test cointegration test is reported in Table 2. The findings reveal that

the null hypothesis of no cointegration  $H_0 : r = 0$  can be rejected, implying a long-run relationship between reproduction rate (RRts) and daily reported deaths (NDts), stringency index (SIts), and total vaccines (TVts) in Zimbabwe.

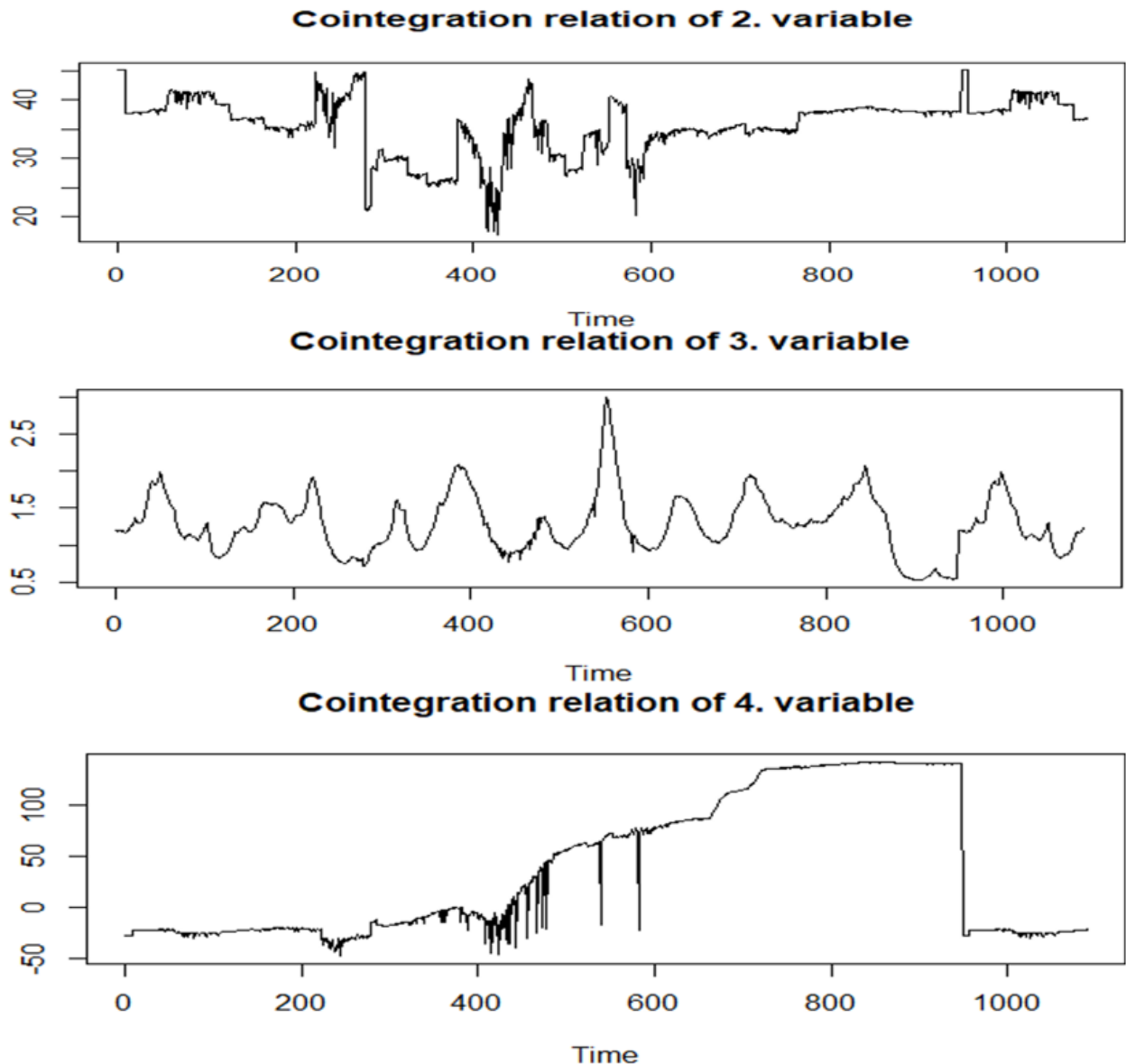
The test statistics and critical values in Table 2 indicate the results of the trace test for different ranks ( $r$ ), which helps identify the number of cointegrating vectors. For  $r \leq 3$ , the test statistic is 3.44, which is below the 10%, 5%, and 1% critical values. This suggests that we do not reject the null hypothesis of at most three cointegrating relationships. But with  $r \leq 2$  the test statistic is 20, which exceeds the critical values at all significance levels, meaning that we reject the null hypothesis and conclude that there are more than two cointegrating relationships at the 1% level. But when  $r \leq 1$ , the test statistic is 48.94, which exceeds the critical values, so we reject the null hypothesis of at most one cointegrating relationship, suggesting more than one. Finally, when  $r \leq 0$  the test statistic is 114.4, which far exceeds the critical values, indicating we reject the null hypothesis of no cointegrating relationships. Based on the test statistics, there are likely two cointegrating relationships (since we reject the null for  $r \leq 2$  but fail to reject for  $r \leq 3$ ). The model suggests that there are at least two long-run equilibrium relationships among the RRts, NDts, SIts, and TVts series. The loading matrix (also called the adjustment matrix) shows how each variable adjusts toward the long-run equilibrium after a short-term shock. In this case, RR adjusts with coefficients (-0.00193, -0.00039, -0.00785, 0.00001), meaning RR will adjust to deviations from the long-run equilibrium; while ND adjusts with larger coefficients, especially in response to SI and TV.

**Table 2:** Cointegration Test using the Johansen Procedure

| Eigenvalues (lambda): 0.0584 0.0262 0.015 0.0032                                                                |             |            |            |           |
|-----------------------------------------------------------------------------------------------------------------|-------------|------------|------------|-----------|
| Values of test statistic and critical values of test:                                                           |             |            |            |           |
|                                                                                                                 | test        | 10%        | 5%         | 1%        |
| $r \leq 3$                                                                                                      | 3.44***     | 6.5        | 8.18       | 11.65     |
| $r \leq 2$                                                                                                      | 20***       | 15.66      | 17.95      | 23.52     |
| $r \leq 1$                                                                                                      | 48.94       | 28.71      | 31.52      | 37.22     |
| $r \leq 0$                                                                                                      | 114.4       | 45.23      | 48.28      | 55.43     |
| Note: *, **, and *** denote statistically significant at the 10%, 5%, and 1% significance levels, respectively. |             |            |            |           |
| (These are the cointegration relations)                                                                         |             |            |            |           |
|                                                                                                                 | RRts.l2     | NDts.l2    | SIts.l2    | TVts.l2   |
| RRts.l2                                                                                                         | 1.0000      | 1.0000     | 1.0000     | 1.0000    |
| NDts.l2                                                                                                         | 0.1906      | -0.1599    | -0.0009    | -0.2985   |
| SIts.l2                                                                                                         | 0.1097      | 0.5016     | 0.0024     | -0.3156   |
| TVts.l2                                                                                                         | 0.0000004   | 0.0000009  | 0.0000     | 0.00001   |
| (This is the loading matrix)                                                                                    |             |            |            |           |
|                                                                                                                 | RRts.l2     | NDts.l2    | SIts.l2    | TVts.l2   |
| RRts.d                                                                                                          | -0.0019     | -0.0004    | -0.0079    | 0.00001   |
| NDts.d                                                                                                          | -0.4717     | 0.0896     | 0.9226     | 0.0006    |
| SIts.d                                                                                                          | -0.0063     | -0.0530    | 0.2185     | 0.0014    |
| TVts.d                                                                                                          | -11470.2100 | -9430.0890 | 60800.3300 | -478.7100 |

In Figure 4, we present the plots from the Johansen cointegration test. Based on the results from Figure 4, we can draw some

preliminary conclusions about the dynamics of reproduction rate and new deaths, total vaccines and the stringency index over our time period: There appears to be some foundation for the co-movement of the reproduction number and each one of the variables, new deaths stringency index, and total vaccines.



**Fig. 4:** Cointegration plots for the reproduction number with number of deaths (2. Variable), stringency index (3. Variable) and total vaccinated (4. Variable).

### 3.2 Empirical Analysis

Since our study captured a long-run linkage among the time series variables, we move to the next step where we explore the possible effect of the new deaths, total vaccines and stringency index on the COVID-19 reproduction rate in Zimbabwe. To address this issue, a Markov-switching dynamic regression is utilised to mitigate the autocorrelation problem because the MS-DR captures the possibility that the time series behaves differently in distinct regimes (e.g., high-volatility vs. low-volatility states), which helps explain the autocorrelation observed in a given data. By allowing the parameters of the regression model to change between regimes, it can adjust to shifts in trends or cycles that might not be captured in a standard linear regression model. Finally, the regime-switching feature can capture patterns of serial correlation by modelling transitions between different states where autocorrelation properties differ.

The first step in determining if the two-regime switching model may be employed is to perform a likelihood ratio (LR) test and

the results of this test confirm that the null hypothesis of no regime switching is rejected in support of the presence of two-regimes because the chi-square statistic's reported p-value is less than 5% significant level. A similar set of findings was published by Wasin and Bandi (2011), Yarmohammadi et al. (2012), Saji (2017), and Psaradakis et al. (2009). This was done by first fitting a linear regression model for the relationship between the reproduction rate and each of the covariates, stringency index, total vaccines, and total vaccines. The covariates had a significant linear relationship with the reproduction number ( $p$ -value $<0.05$ ) except for the daily new vaccines. The results are presented in Table 3 below.

**Table 3:** Estimated parameters from the fitted linear regression model

|                  | Estimate     | Std. Error  | t value  | Pr(> t )     |
|------------------|--------------|-------------|----------|--------------|
| (Intercept)      | 1.6920       | 0.0946      | 17.8950  | < 0.0001 *** |
| SI <sub>ts</sub> | -0.0078      | 0.0013      | -5.9190  | <0.0001 ***  |
| TV <sub>ts</sub> | -0.000000044 | 0.000000003 | -13.1130 | < 0.0001 *** |
| ND <sub>ts</sub> | -0.0022      | 0.0009      | -2.3300  | 0.0200 *     |

**Note:** \*\*\*significant at 1%, \* significant at 5%

Since the null hypothesis of no regime shift is rejected in favor of at least 2 regimes available, we then fit our proposed a Markov Switching-Dynamic regression (MS-DR) to the data using various orders of the autoregressive part, starting with  $p=1$  for different regimes from  $k=1$ . The selection of the best MS-DR model was based on convergence, the log-likelihood, residual standard error, and the multiple R-squared ( $R^2$ ). A forward selection of the variables is used, a procedure that starts with regressing the reproduction rate variable with each of the covariates and eliminating the variables from which the MS-DR model does not converge. In Table 4 we present the fitted MS-DR models.

**Table 4:** Selection of the best MS-DR model

| MS(k)-DR(p)       | Loglik          | Regime 1 |                 |            | Regime 2 |                 |            | Regime 3 |                 |       |
|-------------------|-----------------|----------|-----------------|------------|----------|-----------------|------------|----------|-----------------|-------|
|                   |                 | RSE      | MR <sup>2</sup> | t-val      | RSE      | MR <sup>2</sup> | t-val      | RSE      | MR <sup>2</sup> | t-val |
| M1:SI(k=2,p=1)    | 2251.8          | 0.0196   | 0.9985          | Yes        | 0.0215   | 0.9964          | Yes        | -        | -               | -     |
| M2:SI(k=2,p=2)    | <b>2671</b>     | 0.0105   | 0.9993          | NaN        | 0.0290   | 0.9959          | Yes        | -        | -               | -     |
| M3:SI(k=3,p=3)    | 2742.305        | 0.0204   | 0.9983          | Yes        | 0.0088   | 0.9996          | NaN        | 0.0084   | 0.9996          | Yes   |
| M4:SI(k=3,p=1)    | 2252.323        | 0.0192   | 0.9979          | Yes        | 0.0186   | 0.9973          | Yes        | 0.0110   | 0.9994          | Yes   |
| M5:NV(k=2,p=1)    | 2247.763        | 0.0204   | 0.9983          | NaN        | 0.0209   | 0.9964          | NaN        |          |                 |       |
| M6:ND(k=2,p=1)    | 2251.64         | 0.0221   | 0.998           | Yes        | 0.0180   | 0.9973          | Yes        |          |                 |       |
| M7:SI ND(k=2,p=1) | <b>2253.104</b> | 0.0187   | 0.997           | <b>Yes</b> | 0.0215   | <b>0.9982</b>   | <b>Yes</b> |          |                 |       |

Results from Table 4 show that MS(3)-DR(3) from the regression of the reproduction rate on the stringency index has the highest log-likelihood. However, the model fails to estimate the t-value for the stringency index variable. Therefore, the model cannot be considered for further analysis. Based on this argument, model M7 which can best be described as MS(2)-DR(1) is the best model for describing regime-switching of the reproduction rate dynamics. Another study by Paul, and Hartl (2023), although it did not consider the effects of other variables, also proposes a two-state MSM of either an infection up- or down-turning regime for the daily observational period. The estimated model considers the effects of the stringency index and new deaths and allows the variance to oscillate between regime 1 and regime 2. Table 5 shows the estimated parameters from the fitted model.

**Table 5:** Estimated parameters for the fitted MS-DR

|                                                              | Estimate | Std. Error | t value | Pr(> t )     |
|--------------------------------------------------------------|----------|------------|---------|--------------|
| <b>Regime 1 (RSE= 0.0782; Multiple R<sup>2</sup>=0.9650)</b> |          |            |         |              |
| (Intercept)(S)                                               | 0.0811   | 0.0292     | 2.7774  | 0.0055 ***   |
| SI <sub>ts</sub> (S)                                         | -0.0007  | 0.0004     | -1.7500 | 0.08012 *    |
| ND <sub>ts</sub> (S)                                         | -0.0020  | 0.0009     | -2.2222 | 0.0263 **    |
| RR <sub>ts</sub> 1(S)                                        | 0.9805   | 0.0119     | 82.3950 | < 0.0001 *** |
| <b>Regime 2 (RSE=0.0215; Multiple R<sup>2</sup>=0.9982)</b>  |          |            |         |              |
| (Intercept)(S)                                               | -0.0086  | 0.0038     | -2.2632 | 0.0236 **    |
| SI <sub>ts</sub> (S)                                         | 0.0002   | 0.0001     | 2       | 0.0455 **    |
| ND <sub>ts</sub> (S)                                         | -0.0005  | 0.0070     | -1      | < 0.0001 *** |
| RR <sub>ts</sub> 1(S)                                        | 0.9957   | 0.0020     | 497.85  | < 0.0001 *** |
| <b>Transition Probabilities</b>                              |          |            |         |              |
|                                                              | Regime 1 | Regime 2   |         |              |
| Regime 1                                                     | 0.9116   | 0.0287     |         |              |
| Regime 2                                                     | 0.0884   | 0.9713     |         |              |

**Note:** \*\*\*significant at 1%, \* significant at 5%

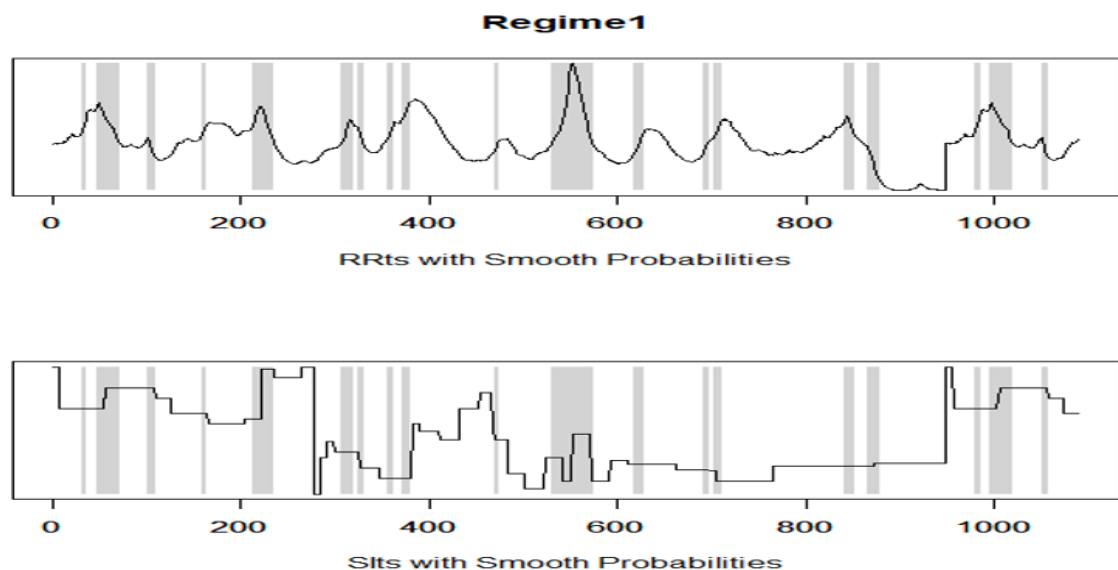
From Table 5 we find that the estimated regime switching trends estimates are all negative for both the stringency index and

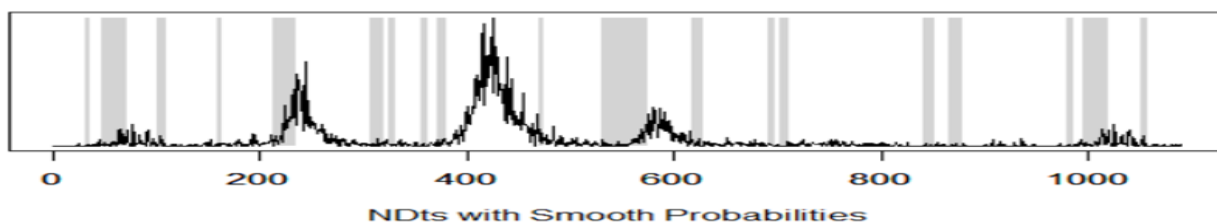
new deaths. This means that a high stringency index results in a low reproduction rate. Miladinov (2021) argued that the reproduction rate is a function of COVID-19 cases and deaths such that an increase in the number of deaths results in a reduction in the reproduction rate. This is because deaths result in the reduction of the number of infected subjects in the population. Another study by Sha et al. (2021) on the Markov switching regression modelling of renewable electricity production shows a negative relationship between the stringency index and renewable energy production. The p values are highly significant suggesting a significant effect, indicating that as the stringency index increases, the reproduction rate of COVID-19 by 0.0007 units in regime 1 and by 0.0002 in regime 2. The transition probability represented as

$$\begin{bmatrix} p(S_t = 1 | S_{t-1} = 1) = 0.911624 \\ P(S_t = 2 | S_{t-1} = 2) = 0.971261 \end{bmatrix}$$

suggesting that regime 1 is less persistent. When the process is in regime 1 (Low COVID-19), there is a high probability that it switches to regime 2 (High COVID-19) and the probability of switching to high COVID-19  $P(S_t = 2 | S_{t-1} = 1) = 0.0883$ . The average duration supports this conclusion because when the system is in regime 1, it takes approximately 11 days. This means that, on average, once the daily cases enter a period of lower infection rates, this period tends to last for about 11 days. While relatively short, it might indicate a temporary stability or decline in cases, possibly due to recent health measures, improved immunity, or changes in public behaviour. Regime 2 on the other hand has an average duration of 34.84 days. This suggests that, on average, when cases surge and enter a period of higher daily infections, it tends to persist for approximately 35 days. This prolonged high-case period could reflect the time it takes for public health interventions (e.g., lockdowns, vaccination drives) to slow down the rate of infections, or it could represent the duration of a typical wave of infections. The system tends to stay longer in the high-case phase (Regime 2), which could indicate that once a wave of infections begins, it takes more time for cases to decline. Periods of low cases (Regime 1), though beneficial, tend to last for shorter periods, suggesting that even when infections drop, vigilance is needed to prevent rapid rebounds. See for instance Abimbola et al (2022) for more readings on regime persistence.

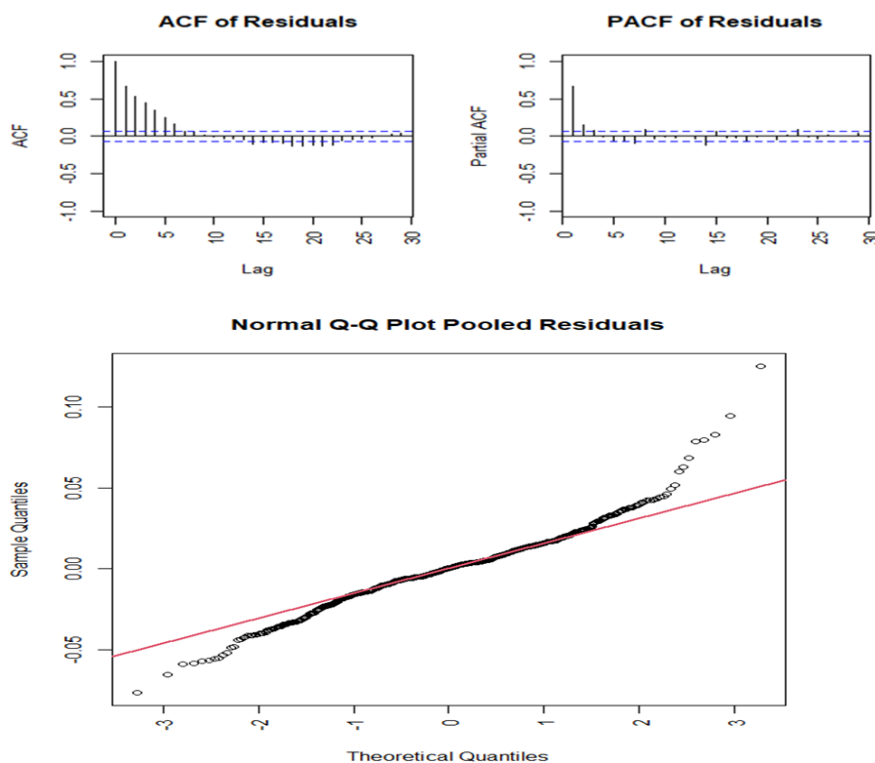
To have a visual appeal of how the three variables, reproduction rate, stringency index, and new deaths, locate the two regimes, we present Figure 5. Figure 5 shows the classification of the regimes based on the reproduction rate and how the two variables, stringency index and new deaths relate to these regimes. Whereas the reproduction rate captures regime 1 (the grey region) as states associated with a high reproduction rate, the stringency index variable associates regime 1 with the general reduction in the stringency index and vice versa. These stringent measures include lockdowns, mask-wearing, social distancing, etc. Lowering (relaxing) these measures increases the spread of COVID-19, thus, the reproduction rate is expected to increase. A rise in deaths due to the COVID-19 pandemic follows immediately after a rise in the reproduction number. The proposed model closely tracks regime shifts as a result of viral mutations, policy interventions, and public behaviour.





**Fig. 5:** Relationship between the stringency index, new deaths, and the reproduction rate in locating the two regimes

We further diagnose the residuals of the fitted model and the results are presented in Figure 6. The residuals look like white noise and they fit into the Normal Distribution. Moreover, the autocorrelation has disappeared.



**Fig. 6:** Normal Probability plot and Autocorrelation Function of the residuals from the fitted Markov switching model.

### Granger Causality test

This study used the Granger causality test to examine the causal relationship between the daily confirmed deaths (ND) and the strict index (SI) and the daily reproduction rate (RR) in Zimbabwe at various lags. The test indicates the causality's direction. Table 6 shows the results of the Granger causality test. At lag order 2 and lag order 5, we find evidence of causation between SI and RR, suggesting that the stringency index is a significant predictor of Zimbabwe's daily reproduction rate. Furthermore, daily COVID-19 fatalities, particularly at lag 2, can also forecast a sizable fluctuation in the reproduction rate. At lag 5, we are unable to identify any meaningful causal relationship between the daily COVID-19 deaths and the reproduction rate.

**Table 6:** The Wald test statistic from the Granger causality test and corresponding p-values

| Direction of Causality | Order 2               | Order 5               |
|------------------------|-----------------------|-----------------------|
| SI→RR                  | 19.5360 (<0.0001 ***) | 7.4866 (<0.0001 ***)  |
| ND→RR                  | 6.1466 (0.0022 **)    | No causality          |
| RR→ND                  | No causality          | 3.3819 (0.004893 **)  |
| TV→RR                  | 17.0920 (<0.0001 ***) | 13.7320 (<0.0001 ***) |

Note: The path of causality is represented by →. 10%, 5%, and 1% levels of significance are illustrated by \*, \*\* & \*\*\*, correspondingly



## 4 Conclusion and Recommendations

This study proposes a Markov-Switching dynamic regression (MS-DR) model to model the dynamics of the COVID-19 reproduction number using data from Zimbabwe. The MS-DR model is adept at handling the spread of COVID-19 which is characterised by significant shifts in behaviours. Our results show that the MS-DR model is capable of capturing the characteristics of the COVID-19 infection rate. The two regimes are classified as high reproduction rate (regime 1) and a low reproduction rate (regime 2) and both states are persistent. Using the model, we also investigated the association between the regimes and other factors: the stringency index, the total number of the vaccinated population, and the daily number of reported deaths. This study is unique in that before fitting and coming up with the best MS-DR model, it assesses the cointegration of the COVID-19 infectious rate with the factors and further assesses the direction of causality. Throughout the pandemic, deaths due to COVID-19 were significantly high after an increase in the reproduction rate. As more infected individuals die the reproduction rate makes a transition to regime 2. This makes mathematical sense since the deaths reduce the number of infected individuals and consequently lower the transmission rates. High reproduction rates follow after the lowering of the stringency index and as the stringency index is increased, the reproduction rate decreases to regime 2. This shows the significance of policy responses as an intervention to the COVID-19 pandemic. Thus, the MS-DR model is the best model for assessing structural changes such as the loosening and tightening of the stringency index. Hence, the MS-DR is an appropriate policy tool for monitoring the state of the pandemic. The proposed model can be used to examine effectiveness of interventions for future epidemics with similar characteristics as COVID-19. However, more research on MS-DR modelling of COVID-19 and application in other countries need to be conducted in order to generalise the findings from this study.

## Acknowledgement

The authors sincerely thank the anonymous reviewers for their helpful comments and suggestions on this paper.

## References

- [1] Cutler, D. M. (2023). Health System Change in the Wake of COVID-19. In JAMA Health Forum. 4(10): 234355-234355. American Medical Association.
- [2] Khan, M., Aslam, F., & Ferreira, P. (2021). Extreme Value Theory and COVID-19 Pandemic: Evidence from India. *Economic Research Guardian*, 11(1), 2-10.
- [3] Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27(2). doi:10.1093/jtm/taaa021
- [4] Makatjane, K., & Moroke, N. (2022). Examining stylized facts and trends of FTSE/JSE TOP40: a parametric and Non-Parametric approach. *Data Science in Finance and Economics*, 2(3), 294-320. doi:<https://doi.org/10.3934/dsfe.2022015>.
- [5] Makki, F., Sedas, P. S., Kontar, J., Saleh, N., and Krpan, D. (2020). Compliance and stringency measures in response to COVID-19: a regional study. *Journal of Behavioral Economics for Policy*, 4(S2), 15-24.
- [6] Wong, F., & Collins, J. J. (2020). Evidence that coronavirus superspreading is fat-tailed. *Proceedings of the National Academy of Sciences*, 117(47), 29416-29418. doi:<https://doi.org/10.1073/pnas.2018490117>.
- [7] Rossouw S, Greyling T, Adhikari T (2021) The evolution of happiness pre and peri-COVID-19: A Markov Switching Dynamic Regression Model. *PLoS ONE* 16(12): e0259579. <https://doi.org/10.1371/journal.pone.0259579>
- [8] Haimerl, Paul, and Tobias Hartl. 2023. Modeling COVID-19 Infection Rates by Regime-Switching Unobserved Components Models. *Econometrics* 11: 10. <https://doi.org/10.3390/econometrics11020010>
- [9] Damiano Pasetto a,\* , Joseph C. Lemaitre b, Enrico Bertuzzo a, Marino Gatto c, Andrea Rinaldo b,d. Range of reproduction number estimates for COVID-19 spread. *Biochemical and Biophysical Research Communications* 538 (2021) 253e258. <https://doi.org/10.1016/j.bbrc.2020.12.003>
- [10] Geraldi, M. S., Bavaresco, M. V., Triana, M. A., Melo, A. P., & Lamberts, R. (2021). 539 Addressing the impact of COVID-19 lockdown on energy use in municipal buildings: a case study 540 in Florianópolis, Brazil. *Sustainable Cities and Society*, 102823
- [11] Rabeya Anzum, Md. Zahidul Islam Mathematical Modeling of Coronavirus Reproduction Rate with Policy and Behavioral Effects RedRxiv 2020.06.16.20133330; doi: <https://doi.org/10.1101/2020.06.16.20133330>
- [12] George Shirreff, Jean-Ralph Zahar, Simon Cauchemez, Laura Temime,1 Lulla Opatowski,1 EMEA-MESuRS Working Group on the Nosocomial Modelling of SARS-CoV-2. Measuring Basic Reproduction Number to Assess Effects of Nonpharmaceutical Interventions on Nosocomial SARS-CoV-2 Transmission. *Emerging Infectious Diseases* •

www.cdc.gov/eid • Vol. 28, No. 7, July 2022. DOI: <https://doi.org/10.3201/eid2807.212339>

- [13] Claris Shoko, Caston Sigauke. Estimation of extreme quantiles of confirmed COVID-19 cases using South African data. *Stat., Optim. Inf. Comput.*, Vol. x, Month 2024, DOI: [10.19139/soic-2310-5070-2079](https://doi.org/10.19139/soic-2310-5070-2079).
- [14] Claris Shoko, Caston Sigauke, Short-term forecasting of COVID-19 using support vector regression: An application using Zimbabwean data, *AJIC: American Journal of Infection Control* (2023), doi: <https://doi.org/10.1016/j.ajic.2023.03.010>.
- [15] Stavroula Zikidou\* and Stamatina Hadjidema. Households Health Expenditure in Interannual Correlation With Public Health Expenditure in Greece. *Front Public Health*. 2020; 8: 448. Published online 2020 Oct 2. doi: [10.3389/fpubh.2020.00448](https://doi.org/10.3389/fpubh.2020.00448).
- [16] Wassell, C and Saunders, P. Time series evidence on social security and private saving: The issue Revisited. (2008). Available online at: <http://www.cwu.edu/business/sites/cts.cwu.edu.business/files/Soc%20Sec%20Final%20Draft.pdf> in Stephanie Glen. "Johansen's Test: Simple Definition" From [StatisticsHowTo.com](https://www.statisticshowto.com): Elementary Statistics for the rest of us! <https://www.statisticshowto.com/johansens-test/> (accessed October 10, 2024).
- [17] Rinaldo, M. Gatto, I. Rodriguez-Iturbe, River Networks as Ecological Corridors. Species, Populations, Pathogens, Cambridge University Press, Cambridge, 2020.
- [18] Luis Rosero-Bixby<sup>1</sup> and Tim Miller<sup>2</sup>. The mathematics of the reproduction number R for Covid-19: A primer for demographers. *Vienna Yearbook of Population Research* 2022 (Vol. 20), pp. 143–166
- [19] Miladinov G. The mechanism between mortality, population growth and ageing of the population in the European lower and upper middle income countries. *PLoS One*. 2021 Oct 29;16(10):e0259169. doi: 10.1371/journal.pone.0259169. PMID: 34714830; PMCID: PMC8555844.
- [20] Muhammad Ibrahim Shah, Dervis Kirikkaleli, Festus Fatai Adedoyin, Regime switching effect of COVID-19 pandemic on renewable electricity generation in Denmark, *Renewable Energy*, Volume 175, 2021, Pages 797-806, ISSN 0960-1481, <https://doi.org/10.1016/j.renene.2021.05.028>.
- [21] Sarkodie, S. A. (2020). Causal effect of environmental factors, economic indicators and 599 domestic material consumption using frequency domain causality test. *Science of The Total Environment*, 139602.
- [22] Breitung, J., & Candelon, B. (2006). Testing for short-and long-run causality: A frequency domain approach. *Journal of econometrics*, 132(2), 363-378.
- [23] Mezencev, R., & Klement, C. (2021). Stringency of the containment measures in response to COVID-19 inversely correlates with the overall disease occurrence over the epidemic wave. *medRxiv*, 2021-01.
- [24] Oliveira, A.M., Binner, J.M., Mandal, A (2021).. Using GAM functions and Markov-Switching models in an evaluation framework to assess countries' performance in controlling the COVID-19 pandemic. *BMC Public Health* 21, 2173 <https://doi.org/10.1186/s12889-021-11891-6>
- [25] WHO (2020). Pandemic Fatigue: Reinvigorating the Public to Prevent Pandemic Fatigue. <https://apps.who.int/iris/bitstream/handle/10665/335820/WHO-EURO-2020-1160-40906-55390-eng.pdf>. Accessed 17 October 2024.
- [26] Xaba, D., Moroke, N. D., and Rapoo, I. (2019). Modeling stock market returns of BRICS with a Markov-switching dynamic regression model. *Journal of Economics and Behavioral Studies*, 11(3 (J)), 10-22.
- [27] Bouteska, A., Sharif, T., and Abedin, M. Z. (2023). COVID-19 and stock returns: Evidence from the Markov switching dependence approach. *Research in International Business and Finance*, 64, 101882.
- [28] Rossouw, S., Greyling, T., & Adhikari, T. (2021). The evolution of happiness pre and peri-COVID-19: A Markov Switching Dynamic Regression Model. *Plos one*, 16(12), e0259579.
- [29] Maragakis LL (2020). Coronavirus Second Wave? Why Cases Increase. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/coronavirus/first-and-second-waves-of-coronavirus>. Accessed 17 October 2024
- [30] McGrail DJ, Dai J, McAndrews KM, Kalluri R (2020). Enacting national social distancing policies corresponds with dramatic reduction in COVID19 infection rates. *PLoS ONE*. 5(7):0236619. <https://doi.org/10.1371/journal.pone.0236619>.
- [31] Shafer LA, Nesca M, Balshaw R (2021). Relaxation of social distancing restrictions: Model estimated impact on COVID-19 epidemic in Manitoba, Canada. *PLoS ONE*. 16(1):0244537. <https://doi.org/10.1371/journal.pone.0244537>.

- [32] Abimbola, O.; Ogbekile, C.; Oseghale, E. (2022). Impact of COVID-19 on Local Markets: An Application of Garch and Markov-Switching Dynamic Regression. Preprints, 2022030281. <https://doi.org/10.20944/preprints202203.0281.v1>.
- [33] Moroke, N. D. (2014). Household debts-and macroeconomic factors nexus in the United States: A cointegration and vector error correction approach. *Journal of Economics and Behavioral Studies*, 6(6), 452-465.
- [34] Kennedy, P. (1996). *A Guide to Econometrics*. 3rd Edition. Massachusetts, Blackwell Publishers Inc.
- [35] Tsay, R. S. (2010b), *Financial Time Series and Their Characteristics*, third edn, Analysis of Financial Time Series.
- [36] Tsay, R. S. (2014), *Financial Time Series*, American Cancer Society, Chapte5, pp. 200-205.
- [37] Shumway, R. H. and Stoffer, D. S. (2017), *Time series analysis and its applications: with R examples*, Springer.
- [38] Montgomery, D. C., Jennings, C. L. and Kulahci, M. (2015), *Introduction to time series analysis and forecasting*, John Wiley & Sons. USA.
- [39] Ljung, G. M. and Box, G. E. (1978), 'On a measure of lack of fit in time series models', *Biometrika* 65(2), 297-303.
- [40] Jonathan, D. C. & Kung-Sik, C. (2008), 'Time series analysis with applications in r', SpringerLink, Springer eBooks.