To Be or Not to Be :
A Study on Government Vaccination Policy on Covid Infection Rate

Yige Luo[1,2], Poorbita Kundu[2], Yutian Yang[2]

ABSTRACT

The primary goal of this project is to explore the effect of government vaccination policies on covid infection rate based on data from several countries across different continental territories. It also aims to study the variation of this effect across the countries. Both the response and the covariate being time series, ARIMAX model is used to fit the data. The proposed final model is an ARIMAX fitted to the suitably transformed response (after regressing out the effect of potential confounders including testing rate and continent, weekly aggregates are taken to remove seasonality) with only government policy as a categorical predictor. Although the exact order of the model varies across countries, in general the proposed full ARIMAX model illustrates a good predictive performance on the test data. The ARIMAX model with covariate significantly outperforms the corresponding null model for the tail point forecasts. Consequently, the prediction MSE for the test data is substantially lowered for all countries. Of course as expected, the regression parameter estimates differ from one country to another. Nevertheless, the above result holds true in general. The very last part of the report acknowledges the shortcomings of the final model and enlists few considerations for future analysis.

INTRODUCTION

As the Covid-19 pandemic continues, it remains a critical task for governments across the world how to respond to such a public health crisis. Protective measurements such as Covid vaccinations are usually helpful at the individual level, yet it is not always clear at the government level what actions to take and when to take them. In many places of the US, for example, people are reported to spend hours in the vaccination queues soon after the vaccination requirement came into effect (Schumaker, 2020), which increased the physical contact and therefore could work against the policy's purpose in slowing the spread of Covid.

In this project, we are planning to assess the effect of vaccination policy on the spread of Covid, and determine if policy effectiveness varies across countries. The vaccination policy is coded as dummies, with increasing stringencies ranging from no policies to all groups. The response variable is a time series of Covid growth rate, measured by daily number of new cases per million roughly from October 2020 to October 2021 across a few selected countries in the northern hemisphere. To examine the association between vaccination policy and Covid growth, we first remove the confounding effects of Covid testing rate and continent, and then 1) build an intercept-only ARIMAX model (a.k.a. time series regression with ARIMA errors) as the null

[1] Department of Evolution and Ecology, University of California-Davis, Davis, CA, USA
[2] Department of Statistics, University of California-Davis, Davis, CA, USA

model; 2) include the vaccination policy in the ARIMAX models; and 3) examine the gain (if any) of prediction power by introducing the vaccination policy. Finally, we use the gain in prediction power as a proxy to evaluate the policy effectiveness in that country and compare it across all countries based on the coefficient estimates.

The performance of the null ARIMAX model and the full ARIMAX model will be examined by internal and external validation. Note that for both models, we used a subset of historical data instead of the full range data (January 2020 - November 2021). This allows us to evaluate the model external prediction accuracies by comparing the model "forecasts" and the realized data in October 2021, where many other potentially contributing factors such as face covering, travel restrictions, weather conditions (Hossain Ahmed & Uddin 2021) are also at play.

The significance of our aims lie in our abilities to make suggestions to policy makers in countries where they have not publicized vaccination guidelines. If the effectiveness of vaccination policy (measured by the gain in variance explained) is uniformly strong, there are good reasons to believe the policy is generally helpful and the government should follow the existing practices. On the other hand, if substantial variations exist in the policy effectiveness, the government should probably think twice before leap, and hopefully incorporate its nation-specific reality (such as the public response to vaccination policies and schemes used for administration of vaccines) during decision making.

METHOD

- **Data selection & filtering**

  Our data were retrieved from Our World In Data (OWID) on November 4, 2021. The time series data for our analysis comprised 1) daily infection rate, measured by the number of new confirmed cases per million and 2) daily testing rate, measured by the number of new tests per million, as well as 3) vaccination policy responses of the government, coded as dummies ranging from 0 to 5 with increasing stringency (Table 1).

Table 1. Factor levels of vaccination policy and corresponding explanations.

Vaccination Policy Code	Explanation	Detailed Explanation
0	None	No availability
1	One group	Availability for ONE of following: key workers/ clinically vulnerable groups / elderly groups
2	Two group	Availability for TWO of following: key workers/ clinically vulnerable groups / elderly groups
3	All vulnerable groups	Availability for ALL of following: key workers/ clinically vulnerable groups / elderly groups
4	Vulnerable + some others	Availability for all three plus partial additional availability (select broad groups/ages)
5	Universal	Universal availability

As the climate cycle of the Southern Hemisphere could be quite different, we limited the scope of our study to the Northern Hemisphere, including all European & North American countries and some African & Asian countries. Out of 186 available countries, we first filtered out countries with less than 90 days of historical data before any vaccination policy was in effect. To ensure sufficient time span of data, we kept countries which have enforced vaccination policies before March 1, 2021. To ensure countries under consideration have enough capacity for Covid testing, we excluded countries with cumulative tests below ten per thousand. Countries with more than 5% missing data (missing daily infection rate, missing daily testing rate, missing date, missing vaccination policy) were also excluded. For each continent, we hand-picked four representative countries with higher rankings of regional GDP (Wikipedia, 2021), leaving 16 selected countries with historical dates (mostly year 2021) for our study.

To control for potential confounding factors (e.g. daily testing rate, continental origin of country), multiple linear regression was first applied on the untransformed response variable over confounders. Further analyses on time series modeling were performed on the adjusted daily infection rate, i.e. residuals of the linear model where effects of the confounders have been regressed out.

As will be discussed in detail in the descriptive analysis section, most of the daily records showed noticeable weekly patterns. This was likely the result of partial operation of testing centers during weekends, which led to aggregation of unrecorded cases to the very next working day (Hale, 2020). To address this issue, two independent methods were considered: 1) weekly contrast and 2) weekly aggregate. The first method took the difference of time series data at lag 7, whereas the second one summed up daily records into weekly aggregates.

$$\text{weekly contrast: } Y_t' = (1 - B^7)Y_t$$
$$\text{weekly aggregate: } Y_t^* = \sum_{i=1}^{7} Y_{t+i}$$

Besides this seasonality component, we also removed the trend component in our data before fitting an ARMA or ARIMA model .

- **ARIMA model**

A process ($X_t : t \in Z$) is called autoregressive-integrated moving average of order (p, d, q), abbreviated by ARIMA(p, d, q), if it satisfies the difference equations (Hyndman, 2018):

$$\nabla^d \phi(B) Y_t = \theta(B) Z_t\,, \text{ where } Z_t \sim WN(0, \sigma^2)$$

Here p denotes the order of MA component, q denotes the order of AR component and d denotes the order of differencing. Consequently, an ARIMA(p, d, q) process is stationary if and only if d = 0, in which case it is ARMA. For any d ≥ 1, the above difference equations will still be satisfied if X_t is replaced with $X_t + \mu_t$, where μ_t is a polynomial of degree d − 1. Therefore, ARIMA processes can potentially play an important role in representing data with trend. Hence we use ARIMA model for our analysis with covid data. In fact, in the model building step we shall try several values of d, ranging from 0 to 3.

- **ARMAX model**

Analogous to the extension of ARMA to ARIMA, ARIMAX may be viewed as a simple extension of the ARMAX model discussed below. Suppose that the time series of interest can be modelled suitably by ARMA(p,q). Now if we further want to study the effect of one or more exogenous variables (covariates) on our response time series, then this leads us to the ARMAX(p,q) model.

The most intuitive extension of ARMA(p,q) model to ARMAX(p,q) model results from the addition of a covariate effect term to the right hand side of the ARMA(p,q) difference equation as follows,

$$Y_t = \beta X_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$
$$\text{where } X_t \text{ denotes covariate at time } t \text{ and } \beta \text{ denotes regression coefficient}$$

However the above model encounters a setback in terms of interpretation of the regression coefficient **β**. Note that **β** can no longer be interpreted as the effect on the response when the covariate value is increased by one unit. The presence of lagged values of the response on the right hand side means that **β** can only be interpreted as the concurrent effect of the covariate on the response conditional on the past values of the response. And this clearly is not intuitive. In fact, writing the above model using a backshift operator, we see that the AR coefficients get mixed up with both the covariates and the error term. All these issues create a need to consider other classes of ARMAX models, of which the two most popular are, 1. Regression with ARMA errors and 2. Transfer Function Models. We shall primarily focus on the first one in our report owing to its more intuitive approach.

A. Regression with ARMA errors

This is probably the most popularly used ARMAX model due to its simplicity. It comprises the conventional regression model setup with the error term being modelled as ARMA(p,q).

$$Y_t = \beta X_t + \eta_t$$
$$\eta_t = \phi_1 \eta_{t-1} + \cdots + \phi_p \eta_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

The conventional model setup ensures the usual and intuitive interpretation of the regression coefficient **β**. On the other hand, the ARMA component in the response time series is incorporated in the model via the error term being modelled as ARMA process. Using backshift operator, the final model can be written in the form,

$$Y_t = \beta X_t + \frac{\theta(B)}{\phi(B)} Z_t$$

B. Transfer Function Models

Transfer function models belong to a general class of models, called dynamic function models. These models were popularized by Box and Jenkins. Although they are less intuitive in approach, they are more suitable for some data due to the greater flexibility of their model components.

$$Y_t = \frac{\beta(B)}{v(B)} X_t + \frac{\theta(B)}{\phi(B)} Z_t$$

$\beta(B)$ allows for lagged effects of the covariate
$v(B)$ allows for decaying effects of the covariate

- **ARIMAX model**

ARIMAX model is a very simple extension of the aforementioned ARMAX model (Hyndman 2010, Hyndman 2018). It uses the same conventional regression model setup but now models the error term as ARIMA(p,d,q). The resulting ARIMAX(p,d,q) model thus takes the form,

$$Y_t = \beta X_t + \eta_t = \beta X_t + \frac{\theta(B)}{\nabla^d \phi(B)} Z_t$$

Clearly ARIMAX(p,d,q) is equivalent to ARMAX(p,q) model fitted to $\nabla^d Y_t$ and $\nabla^d X_t$

Extending the above to the case with k covariates, we have

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + \frac{\theta(B)}{\nabla^d \phi(B)} Z_t$$

RESULT

Descriptive Analysis

Data for this study was retrieved from OWID, which includes but is not limited to daily updated Covid dynamics and policy responses. Our analyses were conducted on data collected from 16 selected countries out of 186 countries worldwide, based on the criteria described in the method. As interests lie on the vaccines that only became available after December 2020, we also limited the date range mostly to the year 2021. To control for differences of population size across countries, the response variable Y_t was defined as daily confirmed Covid cases per million, whose time series progression was shown in Figure 1.

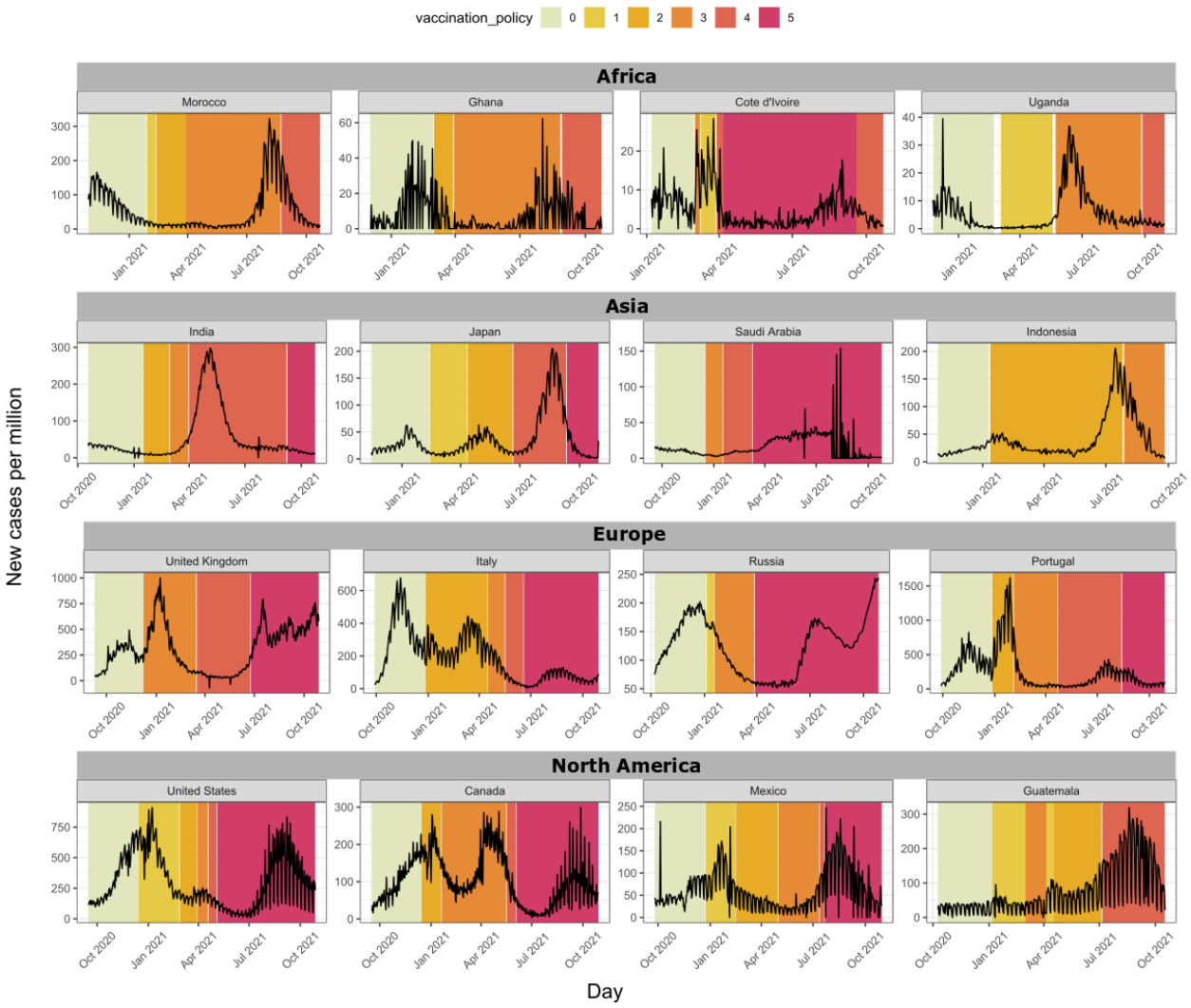


Figure 1: Time series progression of untransformed Covid infection rate across studied countries. Each row represents each continent under study, where country-specific data are displayed inside each panel. Vaccination policies are colored coded as shaded rectangles, and stringency increases with intensity of colors.

The magnitude of covid infection rate seemed to vary across countries and continents (Figure 1), with African countries being less affected than European countries. Asian countries seemed to have comparable levels of infection rate to countries in North America, with the exception that the Covid infection in the US resembled European countries more. Many countries showed more than one episode of Covid outbreak, i.e. peaks of high infection rates, with the first episode during the 2021 winter and a recent rebounce around the 2021 summer. Visual inspection of the coincidence between the vaccination stringency and infection rate did not seem to suggest a negative correlation: countries including the UK, Russia, the US and Mexico showed an infection spike despite that vaccinations became universally enforceable. Yet in Portugal, higher policy stringency seemed to correlate with lower infection rate on an overall basis.

- ***Strong weekly periodicities exist on untransformed Covid infection rate***

  After applying the best ARIMA fit on the raw response variable, the ACF of residuals of many countries show strong lags at multiple of 7 (Figure 2), suggesting the presence of weekly periodicities. Exceptional cases were found in the U.S. and Canada, which were consistent with their noticeable weekly patterns in Figure 1. The lack of confirmed cases during weekends is likely an artifact as some of the testing centers aggregated weekend testing results to the following business day (Hale, 2020). Nevertheless, the prevalence of seasonality on the raw infection rate was a clear violation of stationarity assumption of time series data, and thus entailed proper transformation procedure on the response variable. Two transformation methods were proposed. The first approach to handle such periodicity is to take weekly contrasts of the infection rates, i.e. define a new response variable as $Y_t' = (1 - B^7)Y_t$. An alternative approach of better interpretability is to aggregate data within the week, i.e., define a new response variable as $Y_t^* = \sum_{i=1}^{7} Y_{t+i}$. Results were shown later in the inferential analysis section.

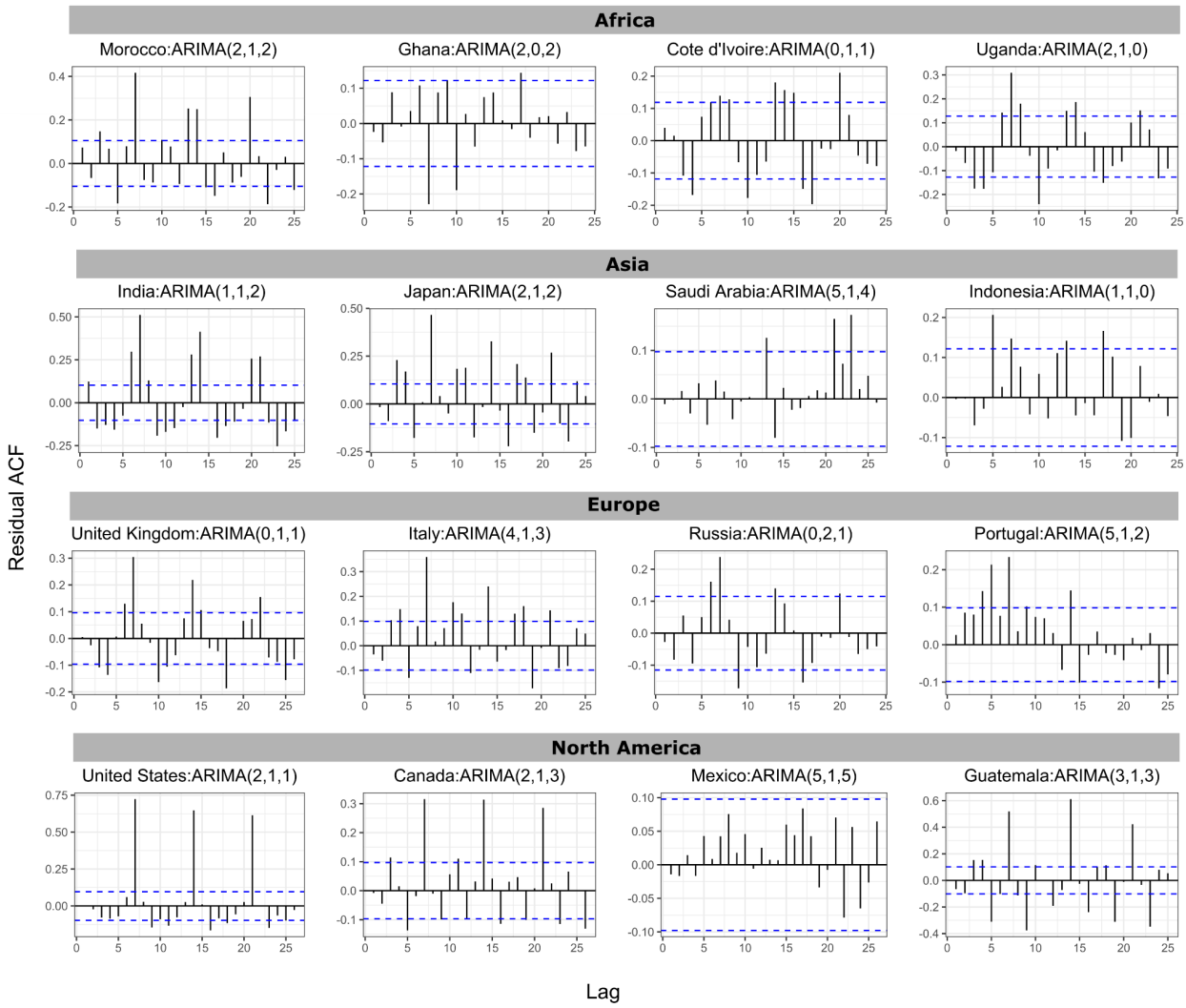


Figure 2: Residual sample ACF of the best ARIMA(p, q, d) fit on the untransformed response variable for each country, showing up to 25 lags. 95% pointwise confidence level of population ACF are indicated by blue dotted lines.

- ***Infection rates associated with testing rates and continental origin of countries***

  Moderate to strong linear relationships were found between infection rates and testing rates in African, Asian and North American countries in our study (Figure 3). On the other hand, however, the trend for European countries seemed to be less clear. For example, testing rates were almost uninformative of infection rate ($R^2 = 3.6 \times 10^{-4}$) in the United Kingdom, whereas in Russia the prediction power of testing rates was as good as other non-European countries. The slight concave shape of smoothing curves could suggest that Covid detections were saturated. In theory, when the number of tests given reaches beyond a certain threshold (several thousands), diminishing returns in positive tests were generally expected (Chiu & Ndeffo-Mbah, 2021). Similar situations

could have occurred in countries including the UK, Portugal and Italy, masking the strong associations between infection and testing when tests were rare, and therefore led to an apparent Simpson's Paradox.

Although we limited our attention to countries of the northern hemisphere, spatial heterogeneities such as geoclimate, population density and socio-economic disparities were still likely to be present and might well confound our analyses. Indeed, continents where our selected countries resided explained 5.3% variances and found to be significantly associated with infection rate (Table 2), despite being less pronounced as the effects of testing (Table 1), which explained 28.0% variances. Overall, these results necessitated the removal of confounding effects of testing rates and continental origin of countries.

Table 2: ANOVA Table showing significant associations between Covid infection rate and the confounding factors.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% variance explained
new_tests_per_million	1	41369748	41369747.7	2295.3	0	28.0
continent	3	7786113	2595371.1	144.0	0	5.3
Residuals	5482	98803994	18023.3	-	-	66.8

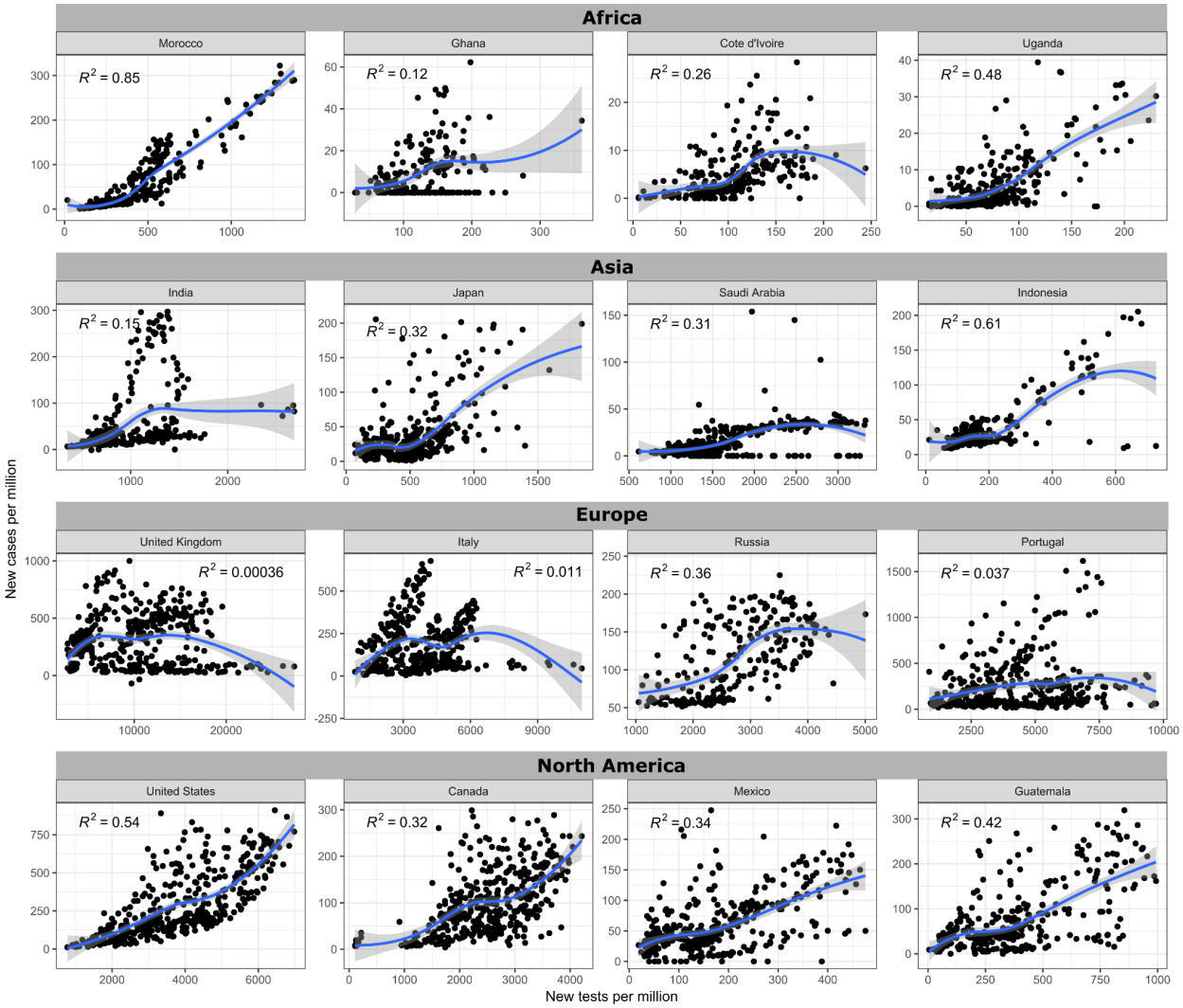
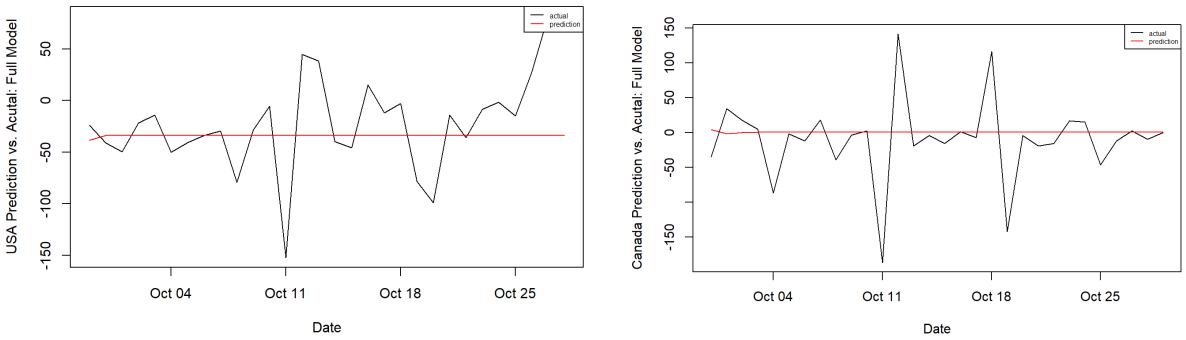


Figure 3: Scatter plot of Covid infection rate vs. testing rate across countries, with loess smoothing (solid blue curve) and 95% confidence interval (light gray ribbon). Each panel also displays the coefficient of determination (R^2) of the best linear fit.

Inferential Analysis

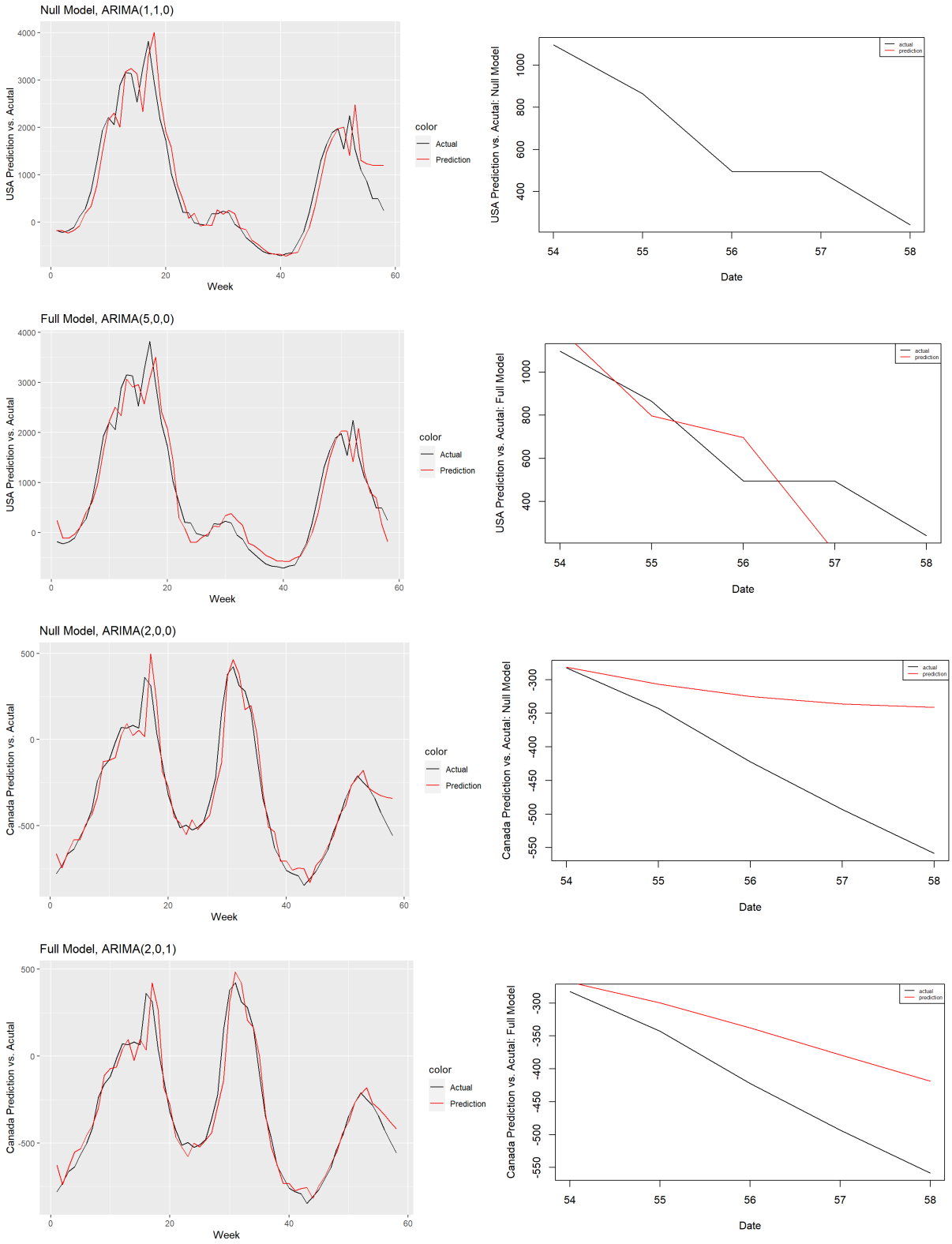
To test the hypothesis that the vaccination policy played a significant role in the evolution of Covid infection, two ARIMAX models were considered. For each of 16 selected countries, we built 1) the null ARIMAX model with only intercept and 2) the full ARIMAX model including vaccination policy as the predictor. Both aforementioned approaches, i.e. Method 1: weekly contrast and Method 2: weekly aggregate, were effective to mitigate the weekly periodicities (Figure S1 and Figure S2). Also as expected, adjusted Covid infection rates were no longer associated with testing rates or continents (Table S1 and Table S2), therefore justifying the use of ARIMAX models to investigate effects of vaccination policy on transformed infection rates.



Method 1: Prediction Plot of USA and Canada in terms of Differences

We then decided to not reversely transform the data, and interpret the data in terms of the difference between Yt and Yt-7. However, since the transformed data are weakly stationary, and the residuals from the linear regression with the new confirmed cases per million and vaccination policy are likely white noises, the prediction after a few lags tends to be the mean of our data. (*Find more details in Supplementary Materials section*)

Next, we decide to aggregate the daily data to weekly data by summing up the values of every 7 days in each column. Since the seasonality of the data has a period of 7, aggregating the daily data to weekly data will also eliminate the seasonality. The drawback of this approach is that we have sharply decreased the sample size to 1/7 of the original data, which leaves us with even less data to train our model since we are eventually splitting the data set into training and testing groups.



Method 2: goodness of fit and prediction plot for USA and Canada

From the plots we are able to observe that (since we are not transforming the data) the range of the fitted values are pretty similar to the original volume, the predicted values follow the same trend and are relatively close to the actual data. Therefore, we decide to use the weekly aggregated data to perform the ARIMAX model analysis.

We choose the USA as an example to show how we determined whether vaccination policy has a significant impact on the new COVID-19 infection rate.

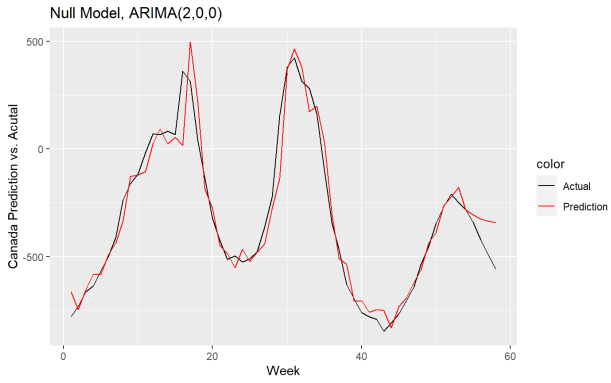
Note : Although our ARIMAX model is similar to linear regression model, since the anova test would require some normality and independence assumption, we cannot apply them on our time series data.

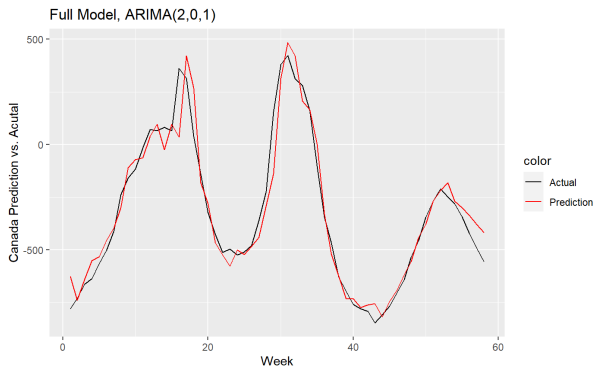
Goodness of fit and prediction for USA: null model vs. full model

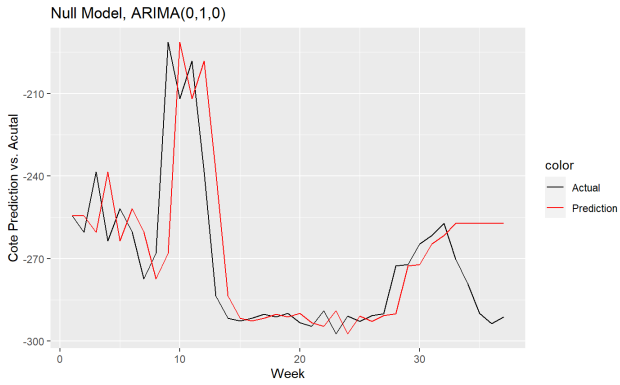
MSE for USA, Null	MSE for USA, Full
Train: 134,373.4 Test: 416,172	Train: 94,950.39 Test: 67,544.61

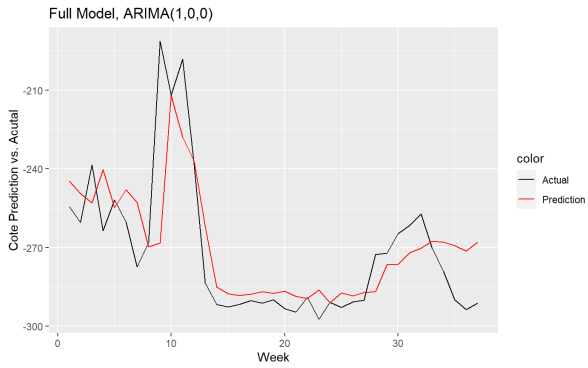
As we can see from the above plots, the prediction for the testing group (after week 50) of the full model has improved with respect to the null model in the sense of lowering the mean squared error. Both the training MSE and the testing MSE of the full model have been significantly lowered, suggesting that having the extra information of vaccination policy does help the model to make better predictions. However we do not observe much of an improvement in pattern matching except for at the tail end. It is this better fit for the tail values that actually lower the test MSE to such a great extent as observed above.

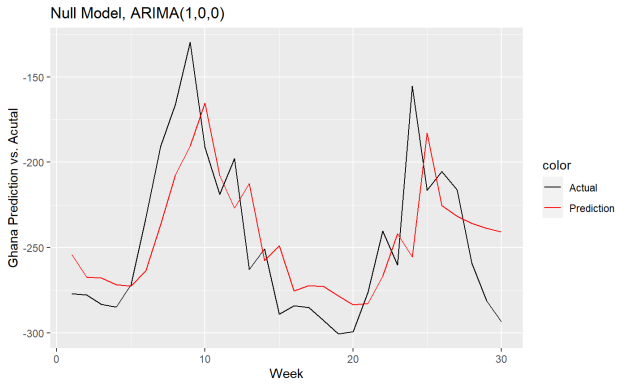
The plots below describe the null vs full model comparison for 15 other selected countries, besides the USA.

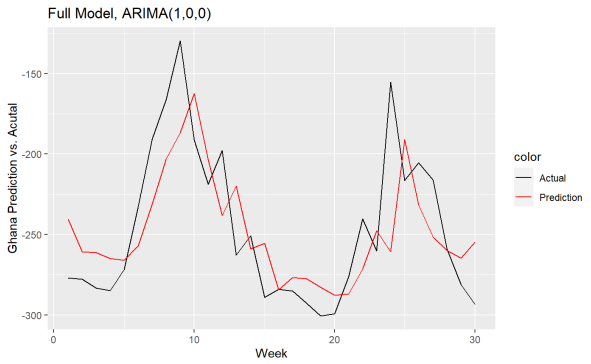

Null Model, ARIMA(2,0,0)
Canada Prediction vs. Acutal
Week
color
Actual
Prediction

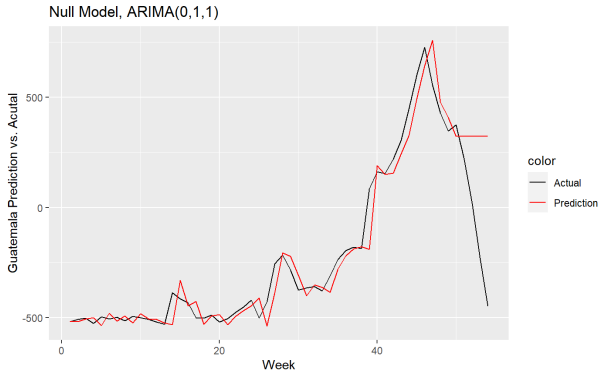

Full Model, ARIMA(2,0,1)
Canada Prediction vs. Acutal
Week
color
Actual
Prediction

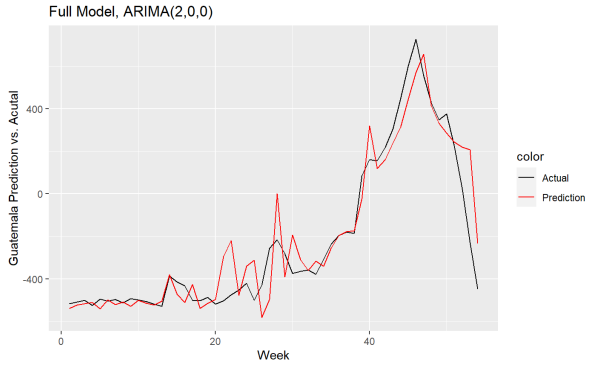

Null Model, ARIMA(0,1,0)
Cote Prediction vs. Acutal
Week
color
Actual
Prediction

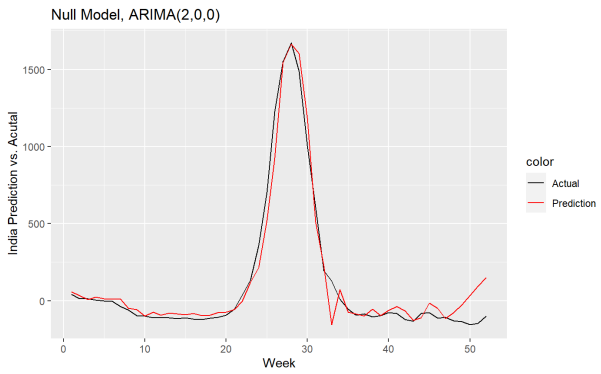

Full Model, ARIMA(1,0,0)
Cote Prediction vs. Acutal
Week
color
Actual
Prediction

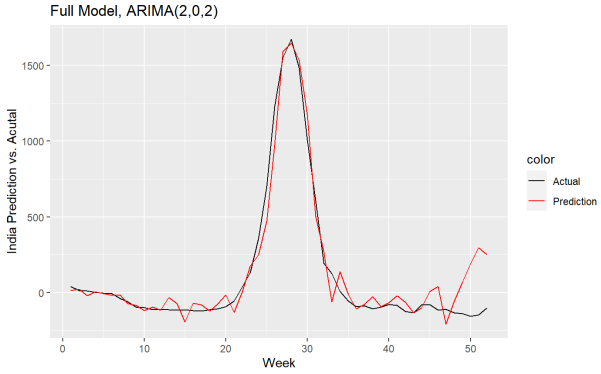

Null Model, ARIMA(1,0,0)
Ghana Prediction vs. Acutal
Week
color
Actual
Prediction

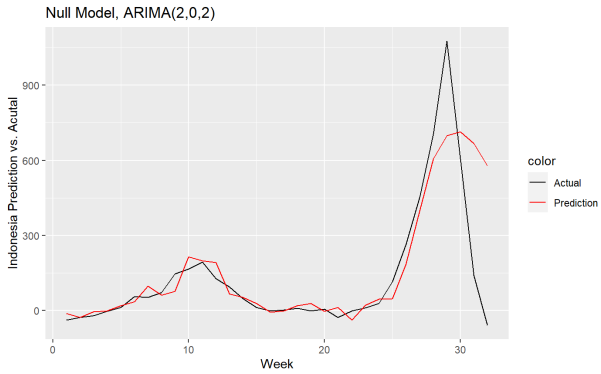

Full Model, ARIMA(1,0,0)
Ghana Prediction vs. Acutal
Week
color
Actual
Prediction

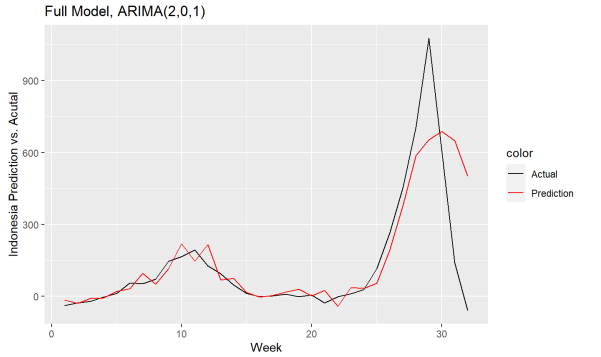

Null Model, ARIMA(0,1,1)
Guatemala Prediction vs. Acutal
Week
color
Actual
Prediction

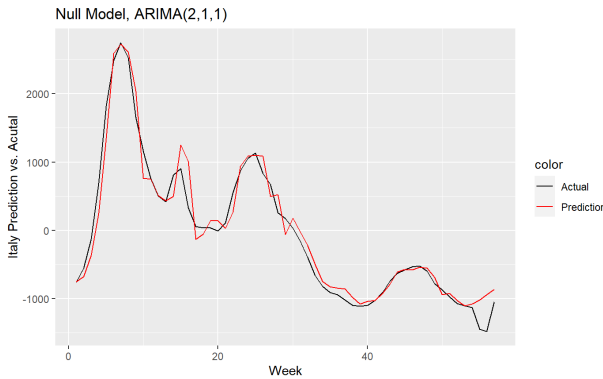

Full Model, ARIMA(2,0,0)
Guatemala Prediction vs. Acutal
Week
color
Actual
Prediction

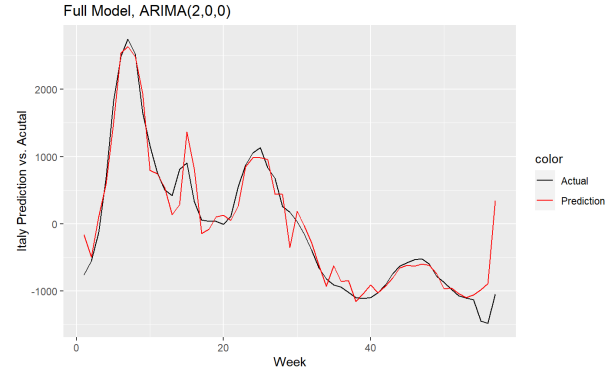

Null Model, ARIMA(2,0,0)
India Prediction vs. Acutal
Week
color
Actual
Prediction

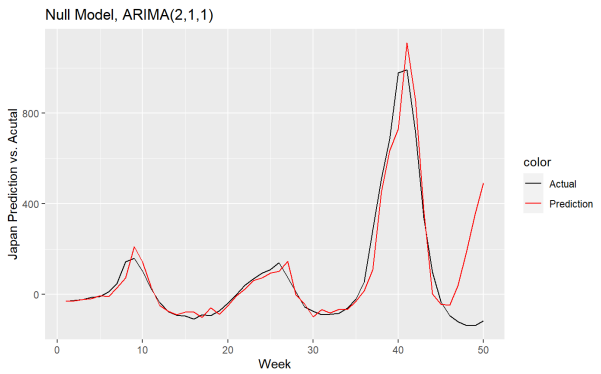

Full Model, ARIMA(2,0,2)
India Prediction vs. Acutal
Week
color
Actual
Prediction

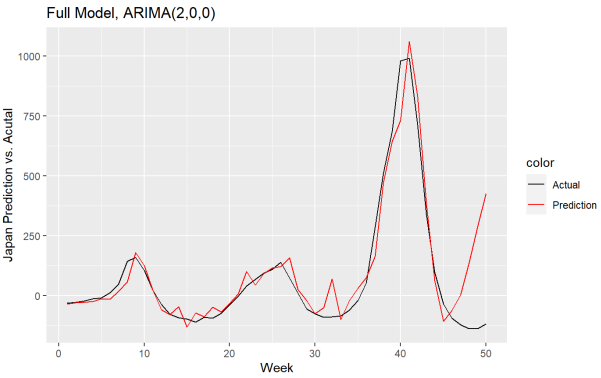

Null Model, ARIMA(2,0,2)
Indonesia Prediction vs. Acutal
Week
color
Actual
Prediction

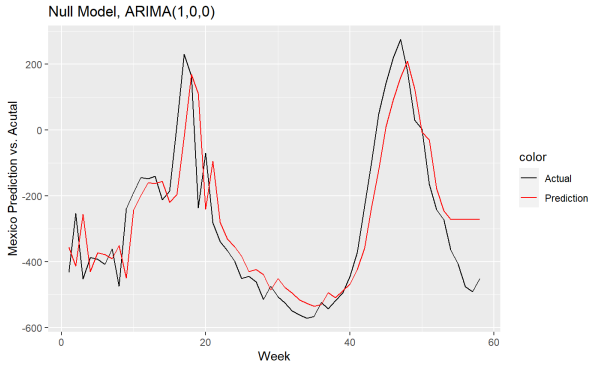

Full Model, ARIMA(2,0,1)
Indonesia Prediction vs. Acutal
Week
color
Actual
Prediction

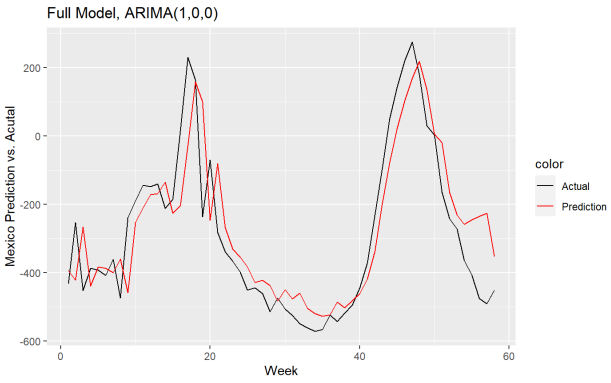

Null Model, ARIMA(2,1,1)
Italy Prediction vs. Acutal
Week
color
Actual
Prediction

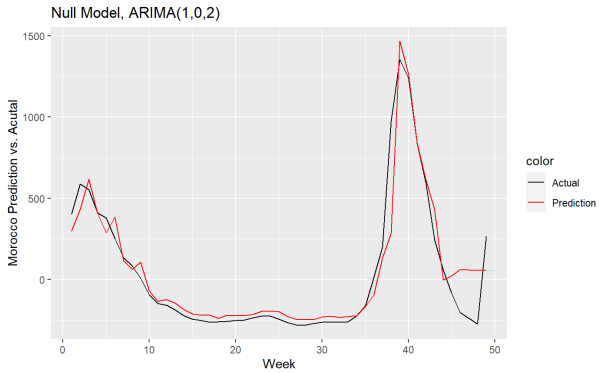

Full Model, ARIMA(2,0,0)
Italy Prediction vs. Acutal
Week
color
Actual
Prediction

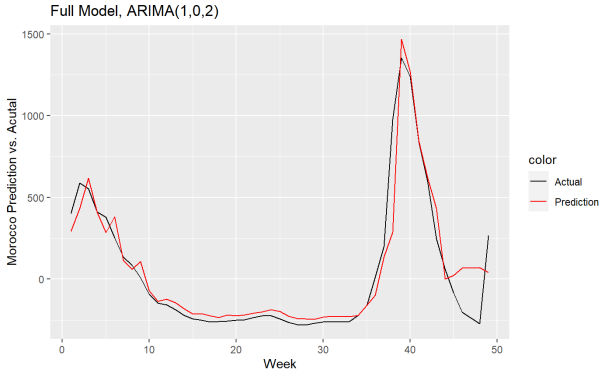

Null Model, ARIMA(2,1,1)
Japan Prediction vs. Acutal
Week
color
Actual
Prediction

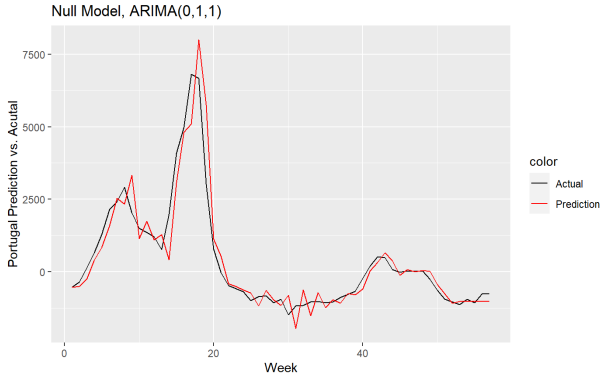

Full Model, ARIMA(2,0,0)
Japan Prediction vs. Acutal
Week
color
Actual
Prediction

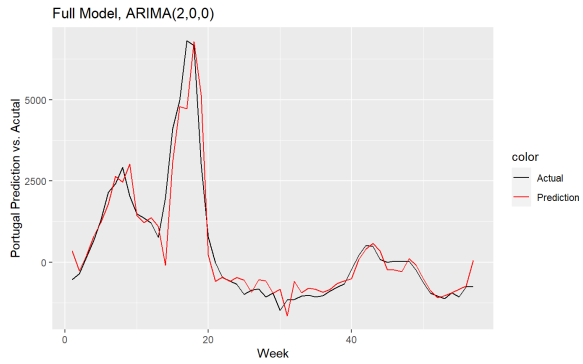

Null Model, ARIMA(1,0,0)
Mexico Prediction vs. Acutal
Week
color
Actual
Prediction

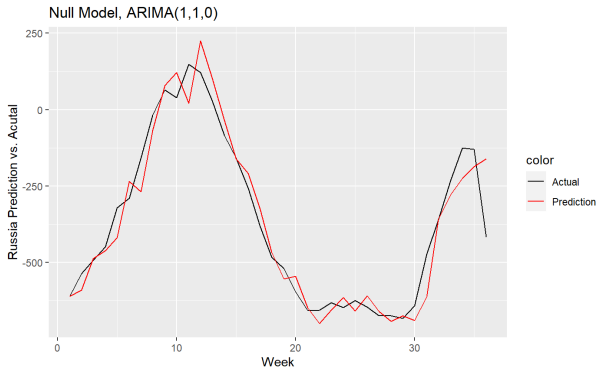

Full Model, ARIMA(1,0,0)
Mexico Prediction vs. Acutal
Week
color
Actual
Prediction

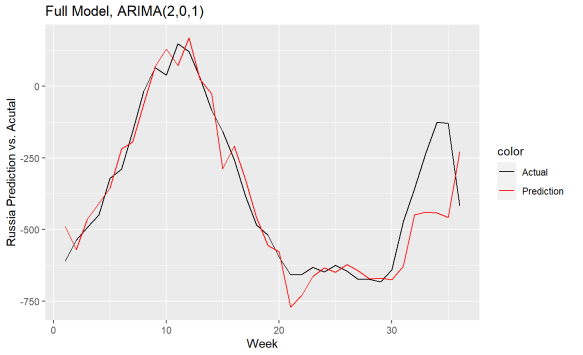

Null Model, ARIMA(1,0,2)
Morocco Prediction vs. Acutal
Week
color
Actual
Prediction

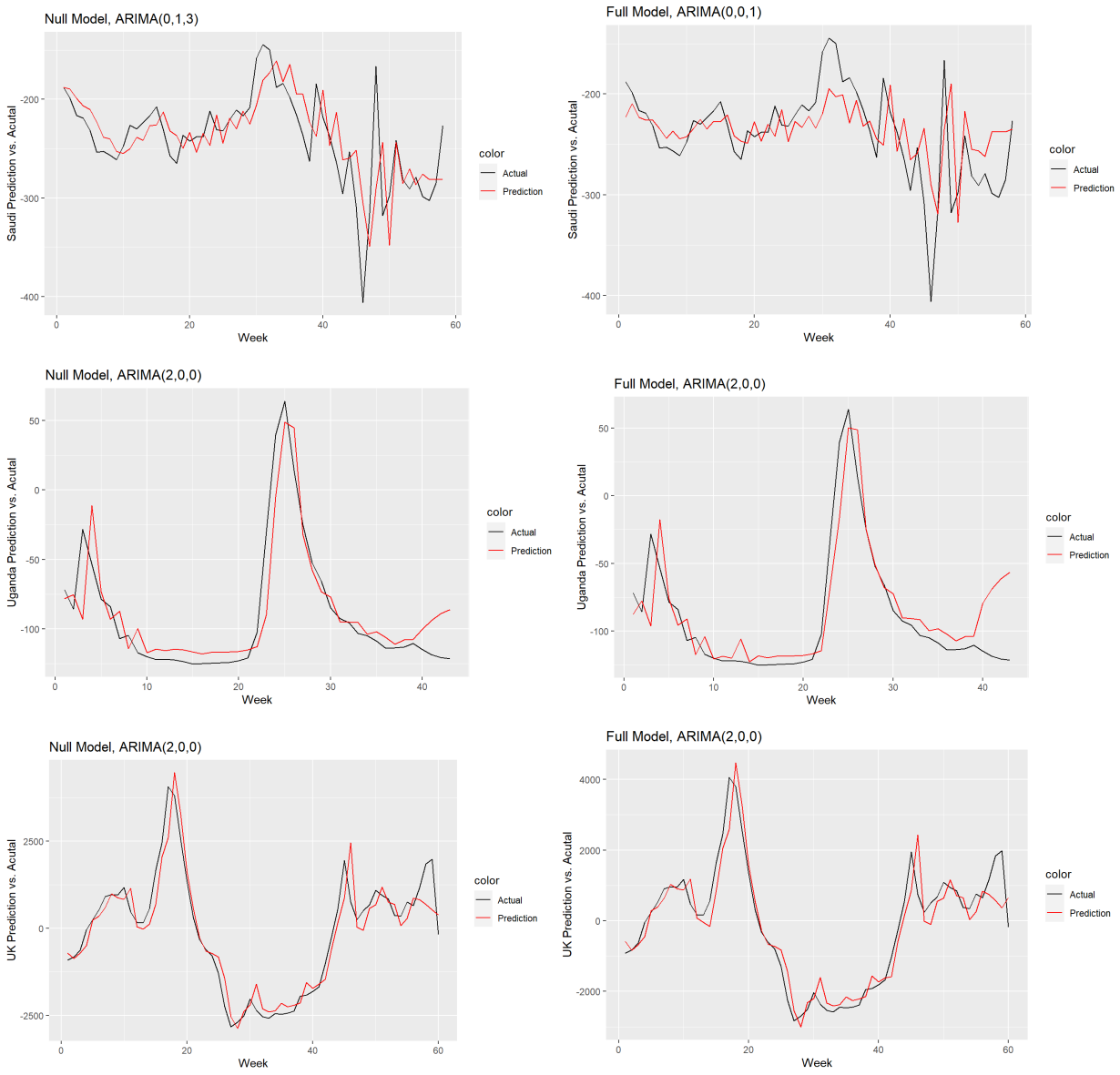

Full Model, ARIMA(1,0,2)
Morocco Prediction vs. Acutal
Week
color
Actual
Prediction


Null Model, ARIMA(0,1,1)
Portugal Prediction vs. Acutal
Week
color
Actual
Prediction


Full Model, ARIMA(2,0,0)
Portugal Prediction vs. Acutal
Week
color
Actual
Prediction


Null Model, ARIMA(1,1,0)
Russia Prediction vs. Acutal
Week
color
Actual
Prediction


Full Model, ARIMA(2,0,1)
Russia Prediction vs. Acutal
Week
color
Actual
Prediction



Comparison of Goodness of Fit and Prediction for rest Fifteen Countries Between Null and Full Model

The above plots illustrate a noteworthy feature. We observe that, not only in the USA, but in most of the other countries, that the prediction for the testing group (after week 50) of the full model has improved with respect to the null model in the sense of lowering both the training MSE and the testing MSE. However again, we do not observe much of an improvement in pattern matching except for at the right tail values which comprise a large proportion of the test data.

By comparing MSE across the 16 countries (in the table below), we can gain a general understanding of how the effects of vaccination policy vary in these countries.

Train and Test MSE for Null and Full model across 16 selected countries:

USA, Null	USA, Full	Canada, Null	Canada, Full
Train: 134,373.4 Test: 416,172	Train: 94,950.39 Test: 67,544.61	Train: 7,709.992 Test: 16,462.48	Train: 7,503.692 Test: 8,346.138
Cote, Null	Cote, Full	Ghana, Null	Ghana, Full
Train: 389.9936 Test: 842.3394	Train: 311.4124 Test: 318.1598	Train: 1,158.84 Test: 1,150.946	Train: 1,156.252 Test: 745.3843
Guatemala, Null	Guatemala, Full	India, Null	India, Full
Train: 5,339.966 Test: 200,357.1	Train: 9,748.447 Test: 56,498.24	Train: 6,728.735 Test: 34,547.26	Train: 6,906.291 Test: 98,352.42
Indonesia, Null	Indonesia, Full	Italy, Null	Italy, Full
Train: 1,265.082 Test: 170,099.7	Train: 1,470.991 Test 155,289.7	Train: 39,815.22 Test: 102,746.2	Train: 45,912.11 Test: 500,346.4
Japan, Null	Japan, Full	Mexico, Null	Mexico, Full
Train: 3,711.901 Test: 149,276.6	Train: 3,753.749 Test: 113,307.6	Train: 11,653.22 Test: 29,623.64	Train: 11,597.87 Test: 34,939.01
Morocco, Null	Morocco, Full	Portugal, Null	Portugal, Full
Train:14,658.97 Test: 64,730.88	Train: 14,610.02 Test: 69,294.38	Train: 422,511.2 Test: 32,079.86	Train: 374,263.6 Test: 141,228.8
Russia, Null	Russia, Full	Saudi, Null	Saudi, Full
Train: 6,538,372 Test: 676,692.2	Train: 6,553,271 Test: 441,397.7	Train: 1,209.52 Test: 801.4738	Train: 1,333.291 Test: 2,095.487
Uganda, Null	Uganda, Full	UK, Null	UK, Full
Train: 391.0393 Test: 614.8072	Train: 372.425 Test: 2288.198	Train: 239,348.9 Test: 781,661	Train: 243,867.4 Test: 1,021,245

From the comparison table above, we derive that the MSE for the training group has either maintained the same level or has decreased sharply. In fact, we tend to have a loose notion that adding vaccination policy will likely explain the data better for countries that are more developed, and it may not be very significant for developing countries but it will at least maintain the same level as without it.

Finally we move on to the second part of our inferential analysis, which focuses on how the effects of vaccination policy vary across the different countries. We directly compare the regression coefficient of the full ARIMAX model in different countries because the interpretation is most intuitive and straightforward. Our intuitive understanding and immediate scientific reasoning, a more strict vaccination policy should imply less new infection rate. Naturally, we expect to see a negative coefficient estimate in our fitted model.

Table below presents the coefficient estimates obtained for the different countries under study:

Coefficients of Full ARIMAX model:

USA: -31.67	Canada: -7.375	Cote: -1.1313	Ghana: -1.6251
Guatemala: 27.17	India: 12.32	Indonesia: 4.676	Italy: -58.110
Japan: 10.654	Mexico: 5.459	Morocco: 1.213	Portugal: -71.70
Russia: -12.811	Saudi: -0.2053	Uganda: 13.072	UK: -19.22

Quite surprisingly, we notice from the table above that not all countries have negative regression coefficients. We also note that some of the more developed countries have larger negative coefficients than the rest. For example, countries in North America like the USA and Canada, and countries in Europe have much larger negative coefficients, which indicates that vaccination policies have expected effects in these countries. Countries in Africa and Asia generally have coefficients around 0 or even positive coefficients, which indicates that vaccination policy seems to have insignificant effects in these countries.

DISCUSSION

Before we proceed to summarise the results of our data analysis, we must note a few typical features of our data that have largely contributed to the observed consequences.

Firstly, we have data for each of the different countries, wherein each such country comprises many different states. These states vary vastly with respect to their covid infection rate and reaction to vaccination policies. Some of these states have high population density while others are sparsely inhabited. Some states have citizens strongly following covid restrictions even without government interference while some others have the majority protesting against many of the covid protocols. The implementation of the covid vaccination policies also differ from state to state within the same country depending on several factors like population, health care facilities, covid awareness. Hence, looking at each country as a single unit, ignoring the unimaginably

vast heterogeneity within a country, may be a plausible reason for the overall lack of significant relationship.

Secondly, our detrended and deseasonalized data hardly shows any pattern and behaves mostly like white noise apparently showing random small range fluctuations around a mean value. Lack of significant relationship between the response and predictor, especially for some of the developing countries, naturally makes the regression method grow a tendency to predict the (unconditional) mean as the forecast. Clearly, if the relationship between covid infection rate and government policies is very weak, it shall not take time for the effect to die out. Hence after the first few forecast values, the prediction simply corresponds to the mean value (around which the training data fluctuates closely) under Method 1.

For each of the countries under consideration, we have tried ARIMAX with d = 0,1,2,3 values but the above observations persisted even for the best choice of d. Although the countries showed vast fluctuations among themselves following our intuitive logic, the aforementioned features persisted for all of them.

Usually data like this are appropriately modelled by (Pattern 2/3/4) Intervention Analysis, accounting for the effect of an external force lasting for a specific period of time. However in our case, our data did not exhibit any apparent shift in mean level and in fact no distinct pattern in particular. Hence given the nature of our data, intervention analysis also did not seem to be a good choice.

We do acknowledge the fact that factors like testing rate and continent may play the role of confounders in our analysis. So we regress them out in the very first place. Eventually, after trying all the aforementioned analyses, we propose our final model.

The final model comprises an ARIMAX model fitted on the weekly aggregates with number of new covid cases (per million) as response and government policy as the only covariate. The optimal order (p,d,q) of the ARIMAX model varies from one country to another.

In general, our final model performs quite well in terms of predicting the test data. Infact, by dint of the inherent nature of our data, the null ARIMAX model without covariates also performs appreciably good for almost all the cases. This kind of spells out that, in general, our covariate does not have a significant effect on the response time series. This implication is further strengthened by the regression parameter estimates fitted by the model.

Thus in the light of the given data, we tend to draw a general conclusion that the government policy does not have a significant effect on the daily new number of covid cases (per million). However, the results show an interesting phenomenon. While predicting the test data, the ARIMAX model shows much lesser tail departure than the corresponding null model in each case. In the domain of our analysis, the tail represents the counterfactual forecast and our covariate seems to play an interesting role in predicting the same. This in turn matches our intuitive belief that the effect of government policies on covid infection rate is lagged rather than concurrent. Lagged covariate analysis may be one of the future considerations for this data.

However, before believing in these results entirely, we must acknowledge an important shortcoming of our model. We did not rule out every factor that may cause the COVID-19 infection rate to vary. Such factors are many in the real world, for instance say: the policy exists, but the governments are not enforcing it efficiently enough or say, people are not responding affirmatively, or say, the population density and climatic conditions of a region catalyse or hinder the infection rate. These are some of the innumerable accountable factors and there lies an even larger pool of unaccountable ones. All these could result in a coefficient around 0 or even positive coefficient in our ARIMAX model. In our future study, we should try to take out the effect of some of these factors, and then the ARIMAX model may provide more dependable results.

APPENDIX

Github link for code is provided to reproduce visualizations and statistical analysis from the study, in addition an HTML is provided with pre-computed results:

AUTHOR CONTRIBUTION

Yige Luo: Conceptualized the project ideas, cleaned up data, conducted exploratory analysis, wrote Introduction & descriptive analysis of the Result section, and maintained the GitHub repository.

Poorbita Kundu: Literature review on time series regression, wrote Abstract, Method & Discussion sections, helped in interpretation of data analysis results.

Yutian Yang: Conducted and wrote inferential analysis of the Result section.

All authors discussed, edited, proofread and agreed upon the submitted manuscript.

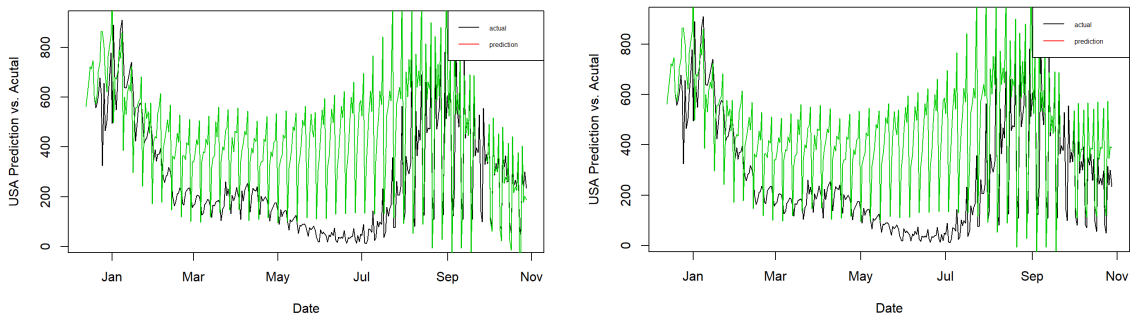
ACKNOWLEDGEMENT

We are grateful to Prof. Alexander Aue and Rui Hu for course materials, instructions and suggestions for data analysis methodology. We also thank Dongjie Chen for exchanging thoughts on using the ARIMAX model.

REFERENCE

- Schumaker, E. (2020, December 31). *Long lines for COVID-19 vaccines build in Florida, Tennessee, Puerto Rico*. ABC News. https://abcnews.go.com/US/long-lines-covid-19-vaccines-build-florida-tennessee/story?id=74988405
- Chiu, W. A., & Ndeffo-Mbah, M. L. (2021). Using test positivity and reported case rates to estimate state-level COVID-19 prevalence and seroprevalence in the United States. *PLOS Computational Biology*, *17*(9), e1009374. https://doi.org/10.1371/journal.pcbi.1009374
- Hale, L. (2020, December 19). Why are COVID-19 numbers lower on the weekends? *Dear Pandemic*. https://dearpandemic.org/covid-19-numbers-lower-on-the-weekends/
- Hossain, Md. S., Ahmed, S., & Uddin, Md. J. (2021). Impact of weather on COVID-19 transmission in south Asian countries: An application of the ARIMAX model. The Science of the Total Environment, 761, 143315. https://doi.org/10.1016/j.scitotenv.2020.143315
- Hyndman, R. J. (2010, October 4). The ARIMAX model muddle | R-bloggers. Retrieved November 6, 2021, from https://www.r-bloggers.com/2010/10/the-arimax-model-muddle
- Hyndman, R. J. & Athanasopoulos, G. (2018) 8.7 ARIMA modelling in R | Forecasting: Principles and Practice (2nd ed). Retrieved November 14, 2021, from https://Otexts.com/fpp2/
- Hyndman, R. J. & Athanasopoulos, G. (2018) 9.2 Regression with ARIMA errors in R | Forecasting: Principles and Practice (2nd ed). Retrieved November 14, 2021, from https://Otexts.com/fpp2/
- List of African countries by GDP (PPP). (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=List_of_African_countries_by_GDP_(PPP)&oldid=1053048796
- List of Asian and Pacific countries by GDP (PPP). (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=List_of_Asian_and_Pacific_countries_by_GDP_(PPP)&oldid=1058211886
- List of North American countries by GDP (nominal). (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=List_of_North_American_countries_by_GDP_(nominal)&oldid=1056068894
- List of sovereign states in Europe by GDP (nominal). (2021). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=List_of_sovereign_states_in_Europe_by_GDP_(nominal)&oldid=1057805830

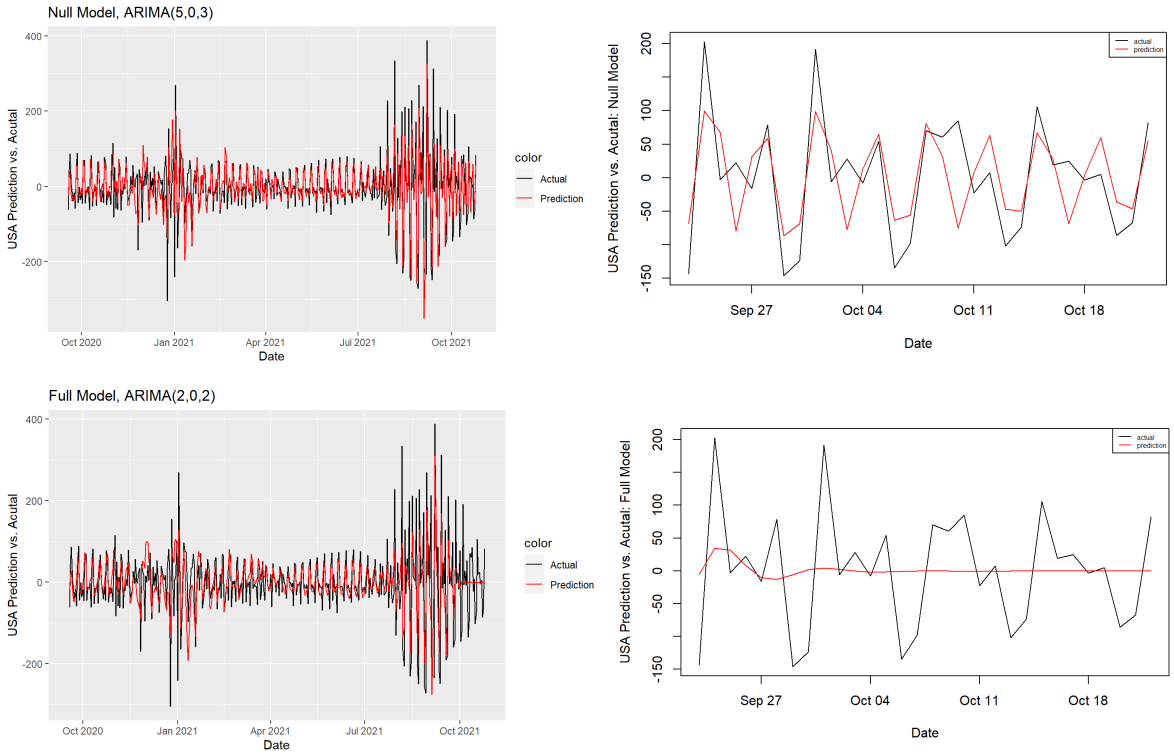
SUPPLEMENTARY MATERIAL

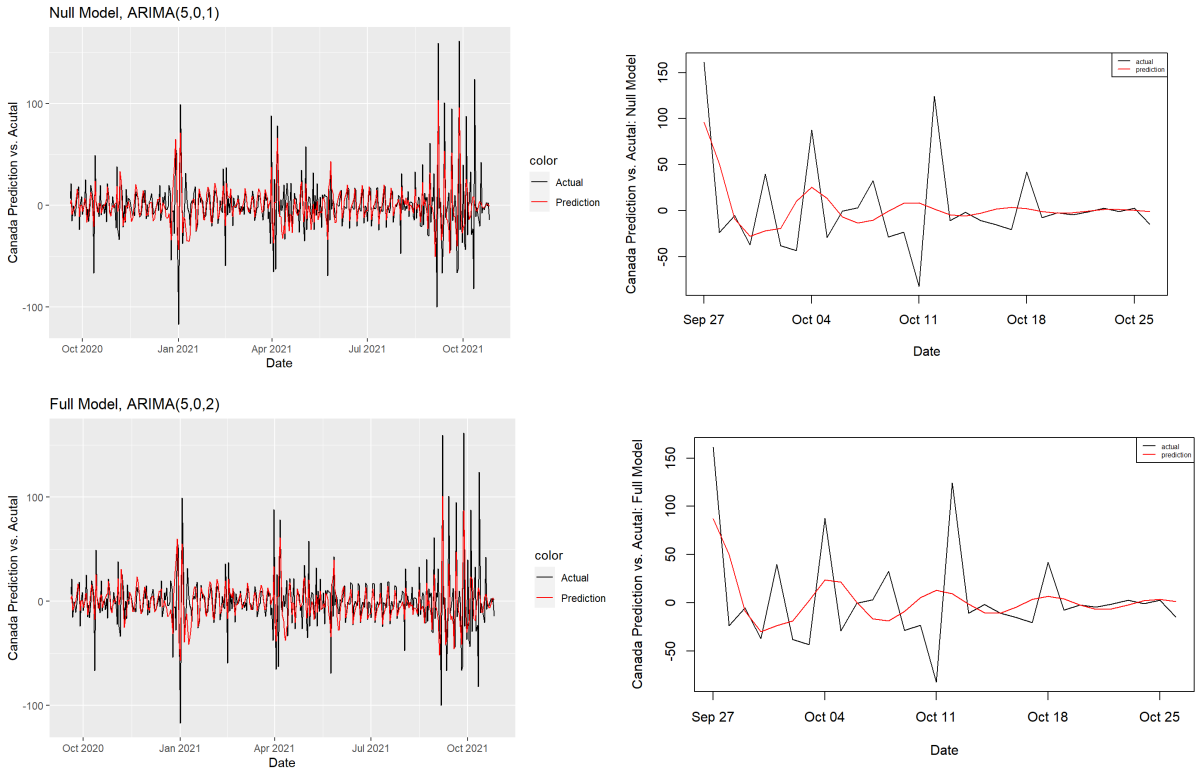


.

Method 1: Goodness of fit plot for Null Model and Full Model of United States

Although the prediction from this model is reasonable, as we were trying to interpret the result, we try to reverse the transformed data but we are unable to keep the data in the same range as the original data. This is because we are trying to predict the difference between Yt and Yt-7, and the predicted difference will have a slight error compared to the original data. When we are doing the reverse transformation process, we are actually enlarging the scale of the error, which results in a much larger range than the original data.





Method 3: Goodness of fit and prediction plots for USA and Canada

Another approach we have attempted is to use Method 3: decompose function from the forecast package in R to manually remove trend and seasonality from the response variable. This approach gives us a more intuitive interpretation and we can still obtain the stationary data to fit the ARIMAX model. However, the same problem occurred. The residuals of the linear regression model still behaved like white noise, so the prediction is just predicting the mean of the data after a few lags. (Find figure above)

Table S1: ANOVA Table showing non-significant associations between adjusted weekly contrast of Covid infection rate and the confounding factors

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% variance explained
New tests per million	1	4180.9	4180.9	1.3	0.3	0
Continent	3	662.0	220.7	0.1	1.0	0
Residuals	5230	16711977.2	3195.4	-	-	100

Table S2: ANOVA Table showing non-significant associations between adjusted weekly Covid infection rate and the confounding factors.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% variance explained
Weekly tests per million	1	309985.1	309985.1	0.4	0.5	0.1
continent	3	201621.0	67207.0	0.1	1.0	0.0
Residuals	797	611637080.5	767424.2	-	-	99.9

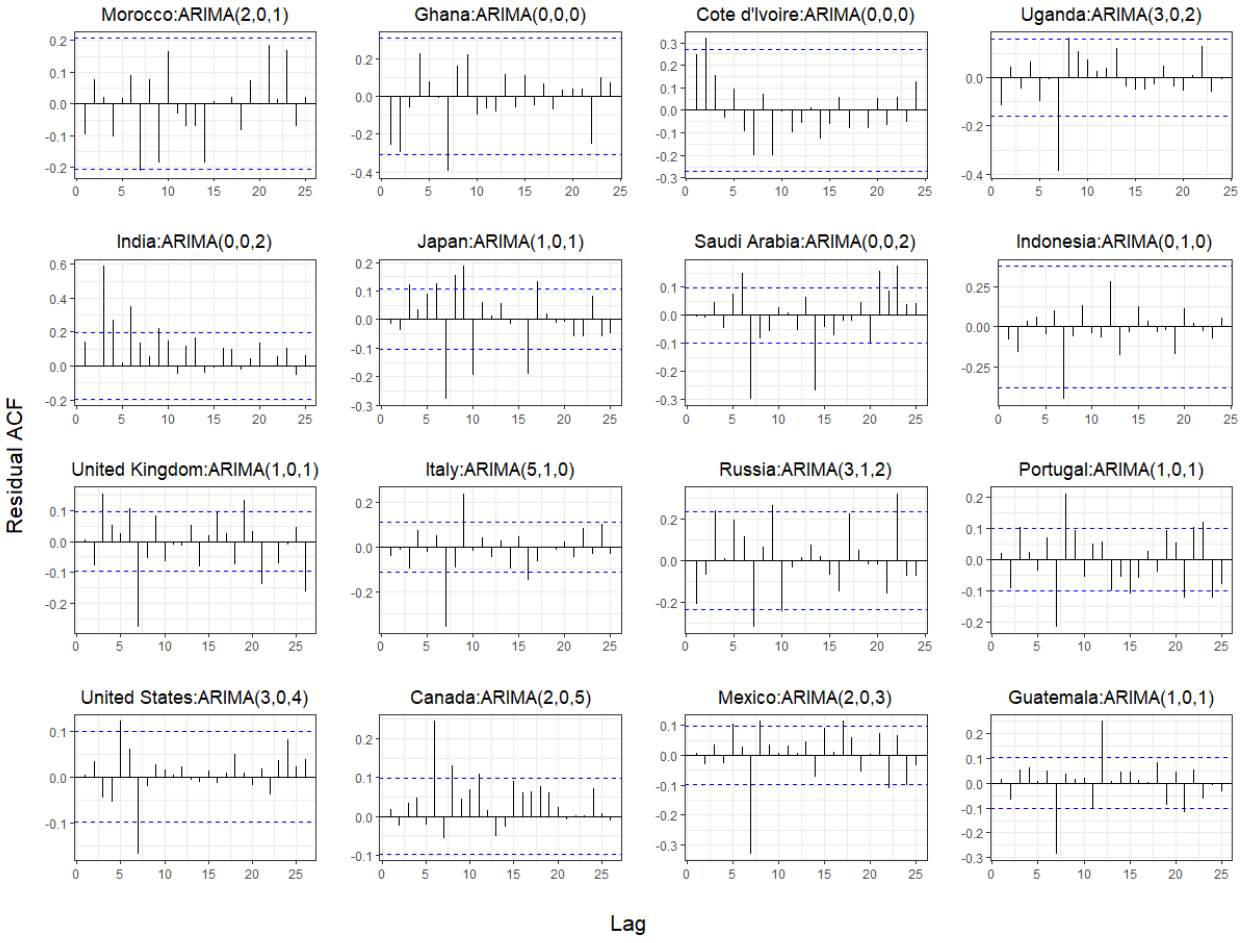


Figure S1: Residual sample ACF of the best ARIMA(p, q, d) fit on the confounder-adjusted weekly contrast of infection rate for each country, showing up to 25 lags. 95% pointwise confidence level of population ACF are indicated by blue dotted lines. Note that effects of weekly periodicities are mitigated.

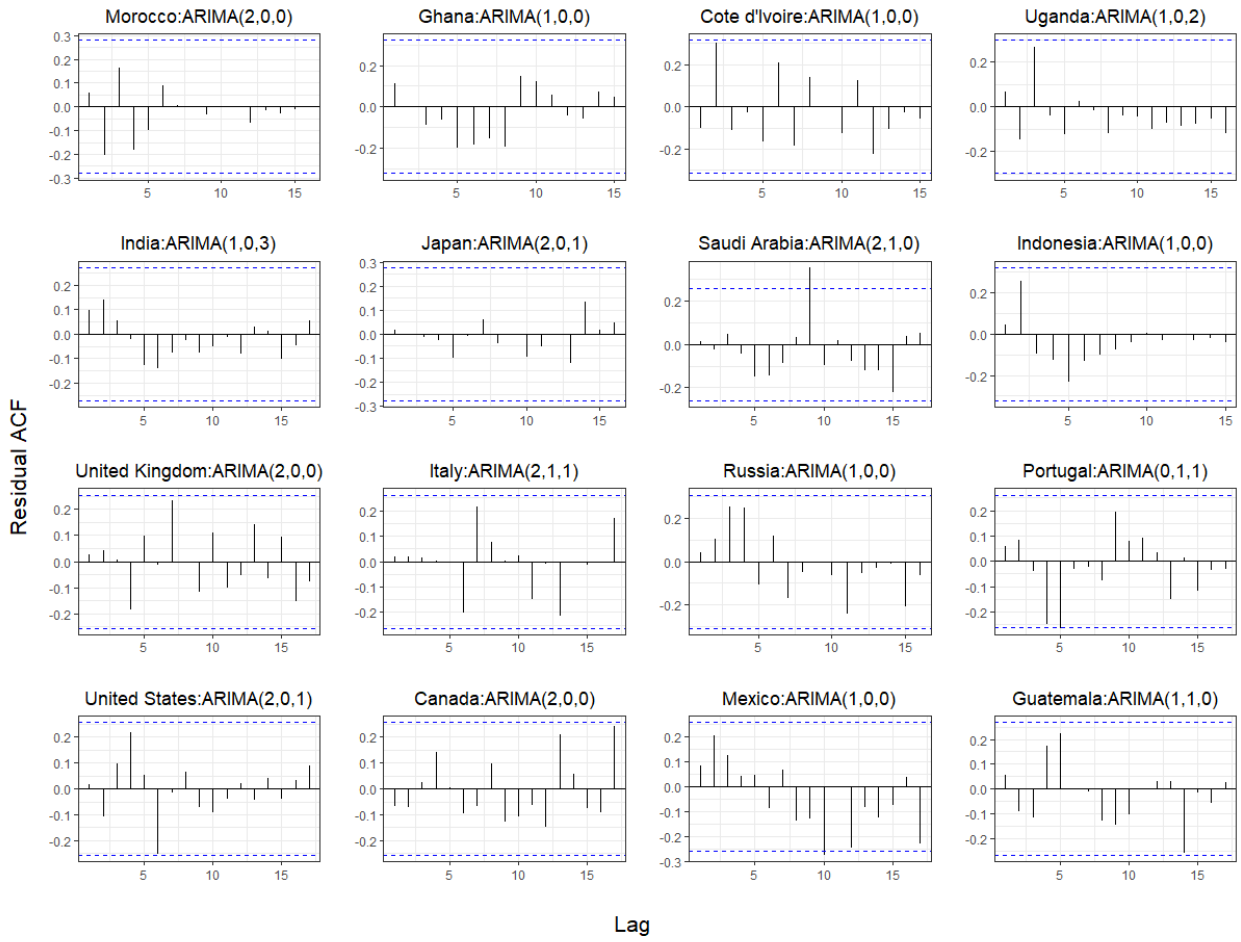


Figure S2: Residual sample ACF of the best ARIMA(p, q, d) fit on the confounder-adjusted weekly infection rate for each country, showing up to 25 lags. 95% pointwise confidence level of population ACF are indicated by blue dotted lines. Note that effects of weekly periodicities are successfully removed.