Shubha Changappa Palachanda
Student ID: 2129848
DATA 512- Project Report

Impact Analysis of human mobility and government interventions effects on the spread of COVID-19

County Assigned: Jefferson County (Kentucky, US)

Introduction

The COVID-19 pandemic has greatly affected people's lives, world economies, and the public health threat it represents is the most serious seen in a respiratory virus since the 1918 H1N1 influenza pandemic. In the absence of a vaccine or an effective treatment, the rapid spread of the disease elicited a wide range of responses from different governments across the globe to contain the spread of the pandemic.

It is well known that the US government had also adopted several interventions apart from the mask mandate policy like travel restrictions, ban on large gatherings, stay-at-home orders, etc., to determine the set of restrictions that will effectively help remediate the spread of COVID-19 without excessively limiting the economic activity. However, these policies did have severe public repercussions. These non-traditional restrictions caused social belonging, quality of education, human needs of safety, and financial security to be threatened. Hence, it's essential to understand which government policies which include pharmaceutical and non-pharmaceutical inventions, had a positive turnaround in containing the spread and were truly worth implementing. This analysis can help the policymakers get more insights on what measures can be quickly implemented/ adjusted in the future if a similar situation were to arise, so as to ensure that the general public is not subjected to restrictions that perhaps don't yield the intended results.

Human mobility also plays an important role in the dynamics of infectious disease spread. By analyzing social mobility data, one can evaluate if changing social mobility in outdoor or indoor settings along with masking policies influenced COVID-19 infection spread in real-time as well. One of the motivations for this exercise is to understand if tracking human mobility both inclusive and exclusive of periods where mask mandates policies were in effect, through publicly available data sources can be used as an effective measure in identifying patterns in the COVID-19 spread.

The Kentucky state government (like other state governments) has implemented several policies at different time periods to help reduce the spread. However, it can be possible that the population responds differently to government interventions and that these differences can cause different human mobility patterns which leads the epidemic development to change. Hence, human mobility is a good additional factor that can also be used to evaluate the effectiveness of pharmaceutical and non-pharmaceutical inventions when analyzing the spread of COVID-19.

The main motivation of this analysis is to truly understand the effects of government interventions and human mobility on the spread of COVID-19 and to identify how these interventions and human mobility, in general, had a significant impact in reducing the COVID-19 spread in Jefferson County, Kentucky, US.

Background/Related Work

There are many publications that have studied the relationship between mobility and the COVID-19 pandemic in different regions of the world. The paper "Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries" by Patrick Bryant and Arne Elofsson [1] is one such study that I went through for my analysis. Through the paper, I learned how a Bayesian model was utilized to determine that the changes in mobility have a considerable overlap with the introduction of governmental non-pharmaceutical interventions using a Bayesian model that estimates the number of deaths on a given day dependent on changes in the basic reproductive number. The shift in mobility (of different types, like visits to parks, grocery, transit stations, etc.,) shows high correlations with the death rates 1 month later. Reduction of movement within the grocery and pharmacy sector resulted in a reduction in death rates overall. This paper gave me several insights into my impact analysis of mobility and daily case spread in Jefferson County. I learned that the inferences between mobility and case rate information can be easily derived through multi-variate time series regression analysis techniques like Vector Autoregression models, as it displays any bi-directional association between the two time-series information and highlights the lagged impact between them.

The paper "Effects of government policies on the spread of COVID-19 worldwide" [2] published by Nature, also helped me understand how to go about analyzing the effectiveness of implemented government policies in the alleviation of COVID-19. Here the Poison regression models were employed using the generalized estimation equations approach in the analysis of COVID-19 daily confirmed cases against policies. However, poison regression was used here also the policy information at their disposal was numerical in nature. Since the government response dataset from the University of Oxford was categorical in nature, I learned that a multi-linear regression model can be used for the analysis but encoding the government mandates as dummy variables on the basis of a stringency index which is a composite measure based on nine response indicators like including school closures, workplace closures, travel bans, etc., rescaled to a numerical value for model fitting.

Postulated research questions:

All the questions listed before are pertaining to Jefferson County, KY

1) How has mobility changed with the government restrictions and policies in place?
   Hypothesis:

Null: The mean difference in the percentage change in mobility prior to and post-mask mandates is not different from 0.
Alternative: The mean difference in the percentage change in mobility prior to and post-mask mandates is different from 0.

2) Which of the different government responses/policies have influenced mobility the most?
Hypothesis: Mobility (of a particular type, say mobility to drug stores or pharmacies) reduced by X% post a varied combination of government policy implementation on the travel ban, masking regulations, and vaccination mandates

3) Which of the different government responses/policies have influenced the COVID-19 spread the most?
Hypothesis: COVID-19 cases reduced by X% post a varied combination of government policy implementation on the travel ban, masking regulations, and vaccination mandates

4) How have different mobility habits influenced the spread of the COVID-19 pandemic? Has reduction/increase in mobility had a significant impact on case reduction/increase?
Hypothesis: A time lag correlation between the different types of human mobility and the covid-19 cases, identifies the mobility contributions towards the spread.

Methodology

All of the methodological methods discussed below are human-centered in nature. To begin, the proposed study design is oriented toward individuals. It uses human mobility behaviors to make decisions, which are subsequently utilized to promote policy health care, and government mandates improvements for the public. It employs data created by individuals for people. It also employs a participatory design in which data from users' inquiries is blended into the created answer in real-time. This immediate, quick, and low-cost feedback will help in the proactive updating of policy suggestions based on people's activities. Certain ethical issues are also included in this analytical process. It is based on data from Jefferson County inhabitants and is free of any other demographic biases. To protect people's privacy, no personal information or particular search searches are used in the study. The approach and analysis developed are also replicable and can be used for any data set containing people's mobility patterns and COVID-19 region responses.

- Change point detection [3]:

For Common Analysis

Changes in time series or signals can take different forms. Roughly speaking, a change point is an abrupt change in a time series, meaning a change in the underlying trends, frequencies, or probability distributions. It is critical to determine whether or not a change has happened, or whether many changes have occurred, and to pinpoint the timing of such changes. Change point detection is a method used in time series analysis to detect signal changes. The Auto-detect number of change points (PELT) option uses the Pruned Exact Linear Time algorithm to estimate the number and location of change points. The intuition behind PELT is that for a time step to be detected as a change point, it must reduce the segmentation cost by more than the penalty value that is added. If the cost reduction is less than the added penalty, the penalized cost will increase, and the time step will not be detected as a change point.

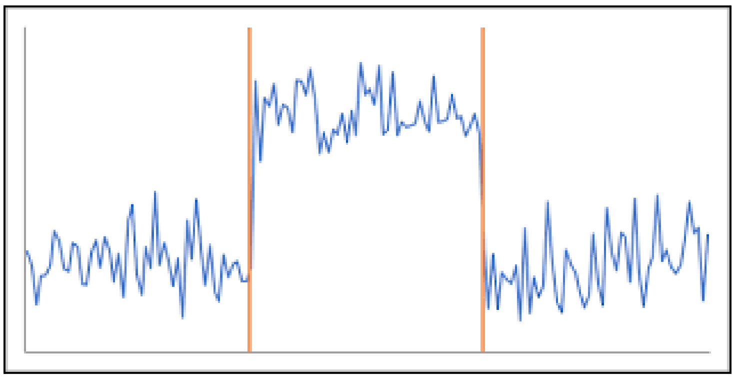


Fig. Example of change points in time series plots

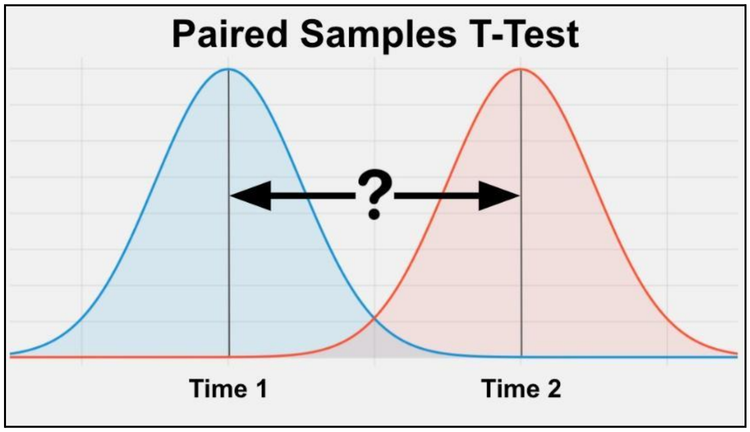
As the PELT method generally produces quick and consistent results when detecting change points through the minimization of costs, this algorithm is used in my common analysis for detecting change points in infection, daily case rate while incorporating the mask mandate policy to better understand the impact (if any) on the COVID-19 infections information for Kentucky County. I also leveraged this algorithm for the changepoint detection in different time series information related to different types of human mobility.

- Dependent/Paired t-test technique [4]:

To address the research question: 1) How has mobility changed with the government restrictions and policies in place?

The dependent t-test (also called the paired t-test or paired-samples t-test) compares the means of two related groups to determine whether there is a statistically significant difference between these means. A dependent t-test is an example of a "within-subjects" or "repeated-measures" statistical test. This indicates that the same participants are tested more than once. Thus, in the dependent t-test, "related groups" indicates that the same participants are present in both groups. The reason that it is possible to have the same participants in each group is that each participant has been measured on two occasions on the same dependent variable.

Since the mobility data is pulled from the same population (i.e., the Population of Jefferson County, Kentucky, US), we utilize a dependent t-test to answer the question of how mobility has changed with mask mandate government restriction and to learn if there is a change in mobility prior to and post mask mandate restrictions in Jefferson County.



Basic Hypotheses: Null: The mean difference in the percentage change in mobility prior to and post-mask mandates is not different from 0. Alternative: The mean difference in the percentage change in mobility prior to and post mask mandates is different from 0.

- Multiple-linear Regression Analysis [5]:

To address the research questions: 2) Which of the different government responses/policies have influenced mobility the most? and 3) Which of the different government responses/policies have influenced the COVID-19 spread the most?

Multiple-linear Regression analysis is an important statistical method that allows us to examine the relationship between two or more variables in the dataset. Multivariate regression allows one to have a different view of the relationship between various variables from all possible angles. It helps us to predict the behavior of the response variables depending on how the predictor variables move. This analysis is a way of mathematically differentiating variables that have an impact. It answers the questions: what are the important variables? Which of them can be ignored? How do they interact with each other? And most important is how certain we are about these variables. Hence for analysis of understanding which of the government restrictions impact human mobility and the spread of Covid-19 in Jefferson County we use Multiple-linear Regression. The model here is fit with a dependent variable: the main factor that we are trying to understand or predict. And then we have independent variables: the factors we believe have an impact on the dependent variable.

To understand which of the different government responses/policies have influenced mobility the most, the dependent variable is the percentage change in human mobility (of different types, such as visits to grocery and pharmacy, visits to retail and recreation stores, visits to workspaces, visits to residential areas, and visits to transit hubs). The independent variables are the different types of government restrictions mandated in Jefferson county to control the spread of Covid-19, namely, facial coverings, school closing, workplace closing, canceling public events, restrictions on gatherings, public transport closing, and stay-at-home requirements. These variables are categorical in nature and hence they are encoded with values based on an index that associates the level of stringency when the government restrictions were mandates before fitting the model.

To understand which of the different government responses/policies have influenced cases the most, the dependent variable is the daily case count in Jefferson County and the independent variables are again the different types of government restrictions mandated in Jefferson county to control the spread of Covid-19, namely, facial coverings, school closing, workplace closing, cancel public events, restrictions on gatherings, public transport closing and stay at home requirements. These independent variables are again encoded as mentioned for the previous model.

- Vector Auto Regression (VAR) [6]:

To address the research question: 4) How have different mobility habits influenced the spread of the COVID-19 pandemic? Has reduction/increase in mobility had a significant impact on case reduction/increase?

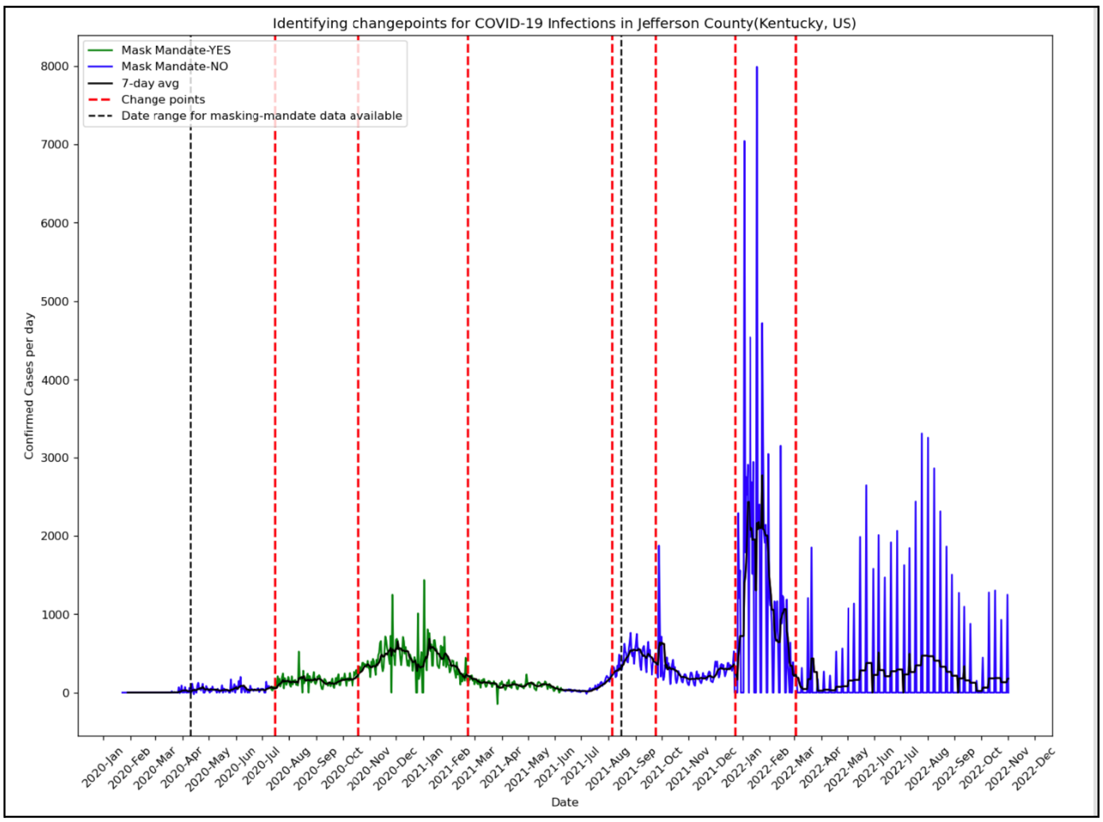
VAR models are an essential component of multivariate time series modeling. The vector autoregression (VAR) model is one of the most successful, flexible, and easy-to-use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to dynamic multivariate time series. It is a workhouse multivariate time series model that relates current observations of a variable with past observations of itself and past observations of other variables in the system. The structure of VAR models enables one to explain the values of endogenous variables from their past observed values.

The VAR models are used for my analysis here, to understand how mobility habits of different types, such as visits to grocery and pharmacies, visits to retail and recreation stores, visits to workspaces, visits to residential areas, and visits to transit hubs, which is times series information influences the daily case counts (also time series data) in Jefferson County. Here, Augmented Dickey Fuller test (ADF Test) which is a common statistical to check whether the two series information being fitted in the VAR model is stationary or not.

Findings

- Common analysis

A time series plot has been created to visualize the variations in the COVID-19 daily infection rate in Jefferson County (state of Kentucky) during and beyond the periods when mask mandate policies were established in Jefferson County. The mask mandate policies data w.r.t Jefferson County is also incorporated into this visual. The changepoints (represented by the red dotted lines in the graph) indicating the abrupt variations in the rate of infections i.e., timestamps where the change in inflected cases is significant from the confirmed cases dataset made available for Jefferson County are calculated using Pruned Exact Linear Time (PELT) Test as mentioned in methodology.

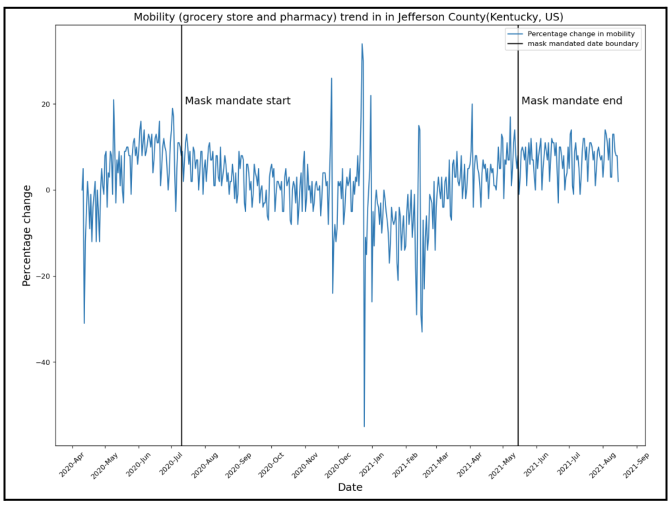


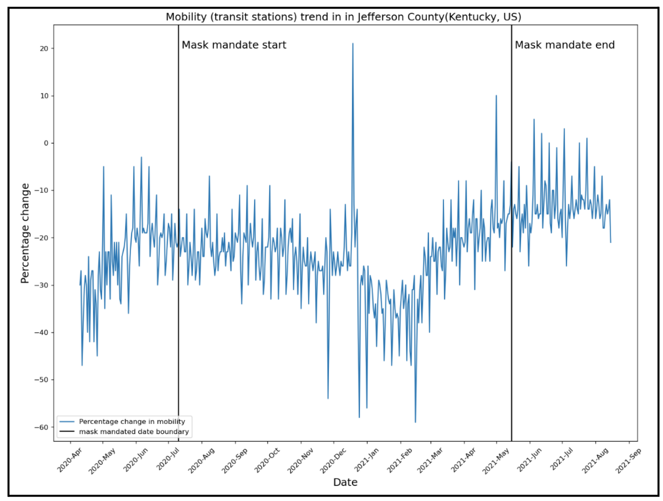
This visual shows that masks were declared necessary in Jefferson County from mid-June 2020 through mid-July 2021 (7/10/2020 to 6/10/2021). It is important to note that before and throughout the commencement of the mask mandate, the cases per day spread was somewhat controlled, however during the center time of the mask mandate (October 2020 to February 2021), the spread appears to be moving towards an uncontrolled trend. The same cannot be said for the elimination of the mask mandate. During the beginning and conclusion of the mask mandate period, we witness a controlled spread (i.e., from Mar

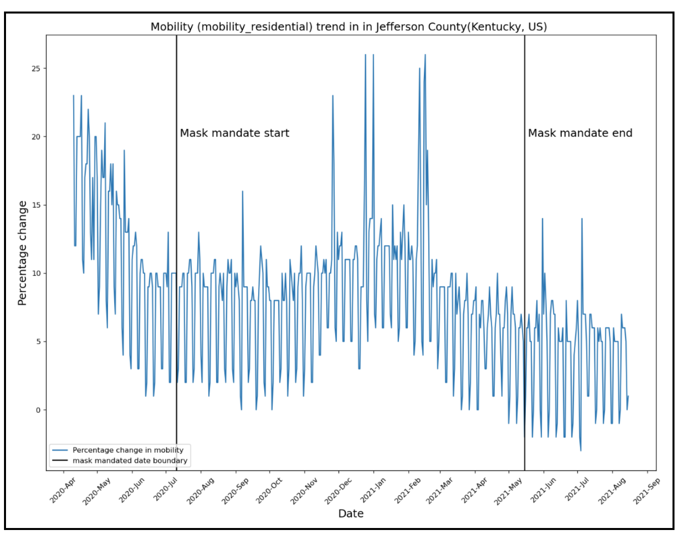
2021 to June 2021). If we assume the masking requirement to be the primary cause of the trend period, we may claim that the impacts of eliminating the mask mandate can be seen in the next 30 days (i.e., beginning in August 2021), when the instances per day begin to spread uncontrollably again.

According to the voluntary masking study, 60.2% of residents in this county always wore masks, and 91.5% used masks more than sometimes (including sometimes survey points). The CDC has countrywide instructions for wearing masks during this time period (July 2 - July 14). So, theoretically, we assume that when there was a state mandate, everyone in the county followed CDC standards rigorously. As a result of these non-uniform fluctuations in case rates throughout mask mandate periods, we can conclude that mask mandates are not the only factor influencing case rate variations.

- How has mobility changed with the government restrictions and policies in place?







Above are the time series plots for changes in mobility trends of different types from the Jefferson County population. Through just visualizing the plots, we can't infer whether mask mandates truly influenced changes in mobility. Hence, we look at the results of a dependence sample test to access the influence of mask mandates on mobility changes.

The dependent/ paired t-test was conducted using the ttest_rel module from scipy package to test the null hypothesis mentioned in methodology on the human mobility information of Jefferson County prior to and post mask mandates (Null hypothesis: The mean difference in the percentage change in mobility prior to and post mask mandates is not different from 0)

Results from the test

Mobility Type	**Tests statistic and p-value**	**Hypothesis action**
Visits to retail and recreation point	statistic=-5.87, pvalue=7.1e-08	Reject null hypothesis
Visits to grocery and pharmacy	statistic=1.72, pvalue=0.088	Don't reject null hypothesis
Visits to transit stations	statistic=-2.17, pvalue=0.032	Reject null hypothesis
Visits to workplace	statistic=-8.12, pvalue=2.2e-012	Reject null hypothesis
Visits to residential areas	statistic=10.49, pvalue=2.69e-17	Reject null hypothesis

From the above results, we see that the p-value is less than the significance level of 0.05 for all mobility types except for mobility corresponding to visits to grocery and pharmacies. This indicated that there is a significant difference in the percentage change in mobility of most types of post-introduction of mask mandates. To look at how different the mobility changes are when other government restrictions are in place we look to the results of the multi-linear regression model next.

- Which of the different government responses/policies have influenced mobility the most?

The multi-linear regression model has been fitted with the following variables data to analyze the impact of government policies on mobility.

Independent variables: 'mask_required',' facial_coverings ', school_closing',
workplace_closing ', 'cancel_public_restrictions_on_gatherings ', 'public_transport',
'stay_at_home_requirements

Response variable: mobility change (for different mobility types: such as visits to grocery and pharmacy, visits to retail and recreation stores, visits to workspaces, visits to residential areas, and visits to transit stations)

```
                            OLS Regression Results
==========================================================================================
Dep. Variable:     mobility_grocery_and_pharmacy   R-squared:                       0.263
Model:                                       OLS   Adj. R-squared:                  0.245
Method:                            Least Squares   F-statistic:                     14.99
Date:                           Mon, 05 Dec 2022   Prob (F-statistic):           8.54e-19
Time:                                   21:31:19   Log-Likelihood:                -1174.6
No. Observations:                            345   AIC:                             2367.
Df Residuals:                                336   BIC:                             2402.
Df Model:                                      8
Covariance Type:                       nonrobust
================================================================================================
                                coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------------
const                         9.9444      5.838      1.703      0.089      -1.539      21.428
mask_required                -1.8088      1.904     -0.950      0.043      -5.554       1.936
facial_coverings             -2.5189      1.627     -1.548      0.022      -5.719       0.681
school_closing                5.0846      1.229      4.137      0.000       2.667       7.502
workplace_closing             0.5888      1.905      0.309      0.057      -3.159       4.337
cancel_public_events         -1.3920      3.655     -0.381      0.004      -8.582       5.798
restrictions_on_gatherings   -0.3464      0.709     -0.488      0.026      -1.742       1.049
public_transport_closing     -0.5619      1.361     -0.413      0.680      -3.239       2.115
stay_at_home_requirements    -8.8001      2.147     -4.098      0.000     -13.024      -4.576
==============================================================================
Omnibus:                      153.883   Durbin-Watson:                   2.064
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1863.531
Skew:                          -1.514   Prob(JB):                         0.00
Kurtosis:                      13.976   Cond. No.                         91.6
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

From the results of the model fitted for mobility type – visit to grocery and pharmacy we see that government restrictions (apart from mask mandates) like stay at home requirements, canceling of public events and public transport close and restriction of gatherings have the most impact in reducing human mobility. For example, we see that when stay-at-home orders are in place, keeping all other variables constant, we see that there is an 8.8% decrease in mobility to grocery and pharmacy stores.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     mobility_retail_and_recreation   R-squared:                       0.466
Model:                                        OLS   Adj. R-squared:                  0.453
Method:                             Least Squares   F-statistic:                     36.58
Date:                            Sun, 11 Dec 2022   Prob (F-statistic):           1.65e-41
Time:                                    22:39:17   Log-Likelihood:                -1173.6
No. Observations:                             345   AIC:                             2365.
Df Residuals:                                 336   BIC:                             2400.
Df Model:                                       8
Covariance Type:                        nonrobust
===============================================================================================
                                coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------------
const                        -8.4171      5.821     -1.446      0.149     -19.867       3.033
mask_required                -3.4251      1.898     -1.804      0.072      -7.159       0.309
facial_coverings             -0.4250      1.622     -0.262      0.793      -3.616       2.766
school_closing                5.3932      1.225      4.401      0.000       2.983       7.804
workplace_closing            -1.7120      1.900     -0.901      0.368      -5.449       2.025
cancel_public_events         -1.9218      3.645     -0.527      0.008      -9.091       5.247
restrictions_on_gatherings    0.1046      0.707      0.148      0.083      -1.287       1.496
public_transport_closing     -0.1267      1.357     -0.093      0.026      -2.796       2.542
stay_at_home_requirements    -1.1449      2.141     -7.541      0.000     -20.356     -11.934
==============================================================================
Omnibus:                      279.339   Durbin-Watson:                   1.968
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             7128.337
Skew:                          -3.139   Prob(JB):                         0.00
Kurtosis:                      24.365   Cond. No.                         91.6
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

From the results of the model fitted for mobility type – visits to retail and recreation points, we see that government restrictions (apart from mask mandates) like stay-at-home requirements, canceling of public events and public transport close, and restriction of gatherings again have the most impact in reducing human mobility. For example, we see that when stay-at-home orders are in place, keeping all other variables constant, we see that there is a 1.1% decrease in mobility to retail and recreation points.

On fitting similar models to the rest of the mobility types we see that stay-at-home orders and the canceling of public events orders are the most common influential government restrictions for reducing the mobility of all types.

- Which of the different government responses/policies have influenced the COVID-19 spread the most?

To understand the impact of government policies on the spread of Covid-19 in Jefferson county we fit another multi-linear regression model with the following variables.

Independent variables: 'mask_required',' facial_coverings ', school_closing’, workplace_closing ', 'cancel_public_restrictions_on_gatherings ’, 'public_transport’, ‘stay_at_home_requirements

Response variable: daily cases count (moved 7-day averaged count)

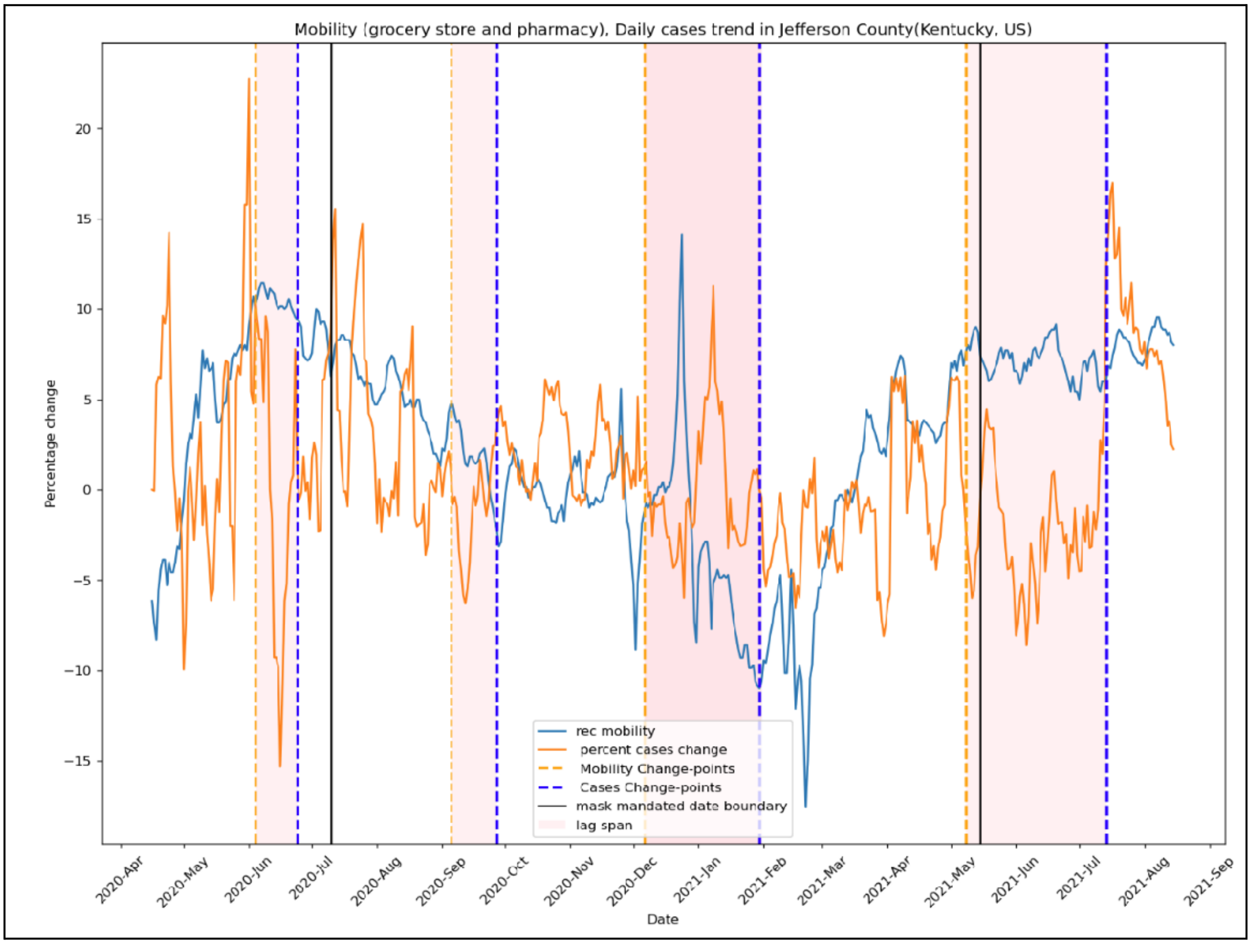
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.063
Model:                            OLS   Adj. R-squared:                  0.040
Method:                 Least Squares   F-statistic:                     2.801
Date:                Sun, 11 Dec 2022   Prob (F-statistic):            0.00514
Time:                        22:43:50   Log-Likelihood:                -1317.3
No. Observations:                 345   AIC:                             2653.
Df Residuals:                     336   BIC:                             2687.
Df Model:                           8
Covariance Type:            nonrobust
==============================================================================================
                                 coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------------------
const                          0.8274      8.829      0.094      0.925     -16.540      18.195
mask_required                 -5.0660      2.880     -1.759      0.079     -10.730       0.598
facial_coverings               0.8828      2.460      0.359      0.020      -3.957       5.722
school_closing                 5.2531      1.859      2.826      0.005       1.597       8.910
workplace_closing             -0.6596      2.882     -0.229      0.019      -6.328       5.009
cancel_public_events          -2.1848      5.528     -0.395      0.023     -13.059       8.690
restrictions_on_gatherings    -0.8355      1.073     -0.779      0.037      -2.946       1.275
public_transport_closing      -1.2796      2.058      0.622      0.034      -2.768       5.328
stay_at_home_requirements     -6.2762      3.247     -1.933      0.054     -12.664       0.112
==============================================================================
Omnibus:                      189.928   Durbin-Watson:                   2.314
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             3344.683
Skew:                           1.874   Prob(JB):                         0.00
Kurtosis:                      17.786   Cond. No.                         91.6
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

From the results of the model fitted we see that most government restrictions have an impact on reducing daily case count. Predominantly, we see that when the stay-at-home restriction is mandated, holding all other variables constant we see a 6.2% reduction in daily case counts. Similarly, we see a 2.1% and 1.27% reduction in daily cases (holding all other variables constant) when the cancellation of public events and public transport closing mandates respectively are put in place.

- How have different mobility habits influenced the spread of the COVID-19 pandemic? Has reduction/increase in mobility had a significant impact on case reduction/increase?

To access the mobility impact on the spread of COVID-19 in Jefferson County, we first plot the time series graphs of the percentage change in mobility (mobility of type- visits to grocery and pharmacy is displayed in the plot below) and change in daily case rate and their corresponding change points calculated through change point detection. The shaded red blocks in the graph show that there is a lag gap in the changes between mobility and the daily case rate. Note: Similar lags have been observed in other mobility types as well. Hence, we look to the results of the VAR model to infer how much of a lagged influence mobility change has on the COVID-19 spread.



Variables for the VAR models: percentage mobility change (mobility of different types, such as visits to grocery and pharmacies, visits to retail and recreation stores, visits to workspaces, visits to residential areas, and visits to transit hubs)) and percent daily case rate change.

```
 Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Mon, 05, Dec, 2022
Time:                     18:23:46
--------------------------------------------------------------------
No. of Equations:         2.00000    BIC:                    12.7888
Nobs:                     69.0000    HQIC:                   10.4058
Log likelihood:          -378.746    FPE:                    53027.4
AIC:                      8.83861    Det(Omega_mle):         14938.7
--------------------------------------------------------------------
Results for equation cases
=====================================================================================================
                                       coefficient       std. error           t-stat            prob
-----------------------------------------------------------------------------------------------------
const                                   164.863164        76.147955            2.165           0.030
L1.cases                                 -0.195984         0.349145           -0.561           0.575
L1.mobility_grocery_and_pharmacy          4.589214         2.564030            1.790           0.073
L2.cases                                 -0.040124         0.401087           -0.100           0.920
L2.mobility_grocery_and_pharmacy         -2.985386         3.328994           -0.897           0.370
L3.cases                                 -0.197872         0.246353           -0.803           0.422
L3.mobility_grocery_and_pharmacy          3.284529         3.297114            0.996           0.319
L4.cases                                 -0.411564         0.294900           -1.396           0.163
L4.mobility_grocery_and_pharmacy         -0.327309         2.560254           -0.128           0.898
L5.cases                                 -0.126979         0.303716           -0.418           0.676
L5.mobility_grocery_and_pharmacy         -1.097784         2.644493           -0.415           0.678
L6.cases                                  0.119933         0.276227            0.434           0.664
L6.mobility_grocery_and_pharmacy          1.022795         1.983395            0.516           0.606
L7.cases                                 -0.034247         0.188143           -0.182           0.856
L7.mobility_grocery_and_pharmacy          4.575773         2.365945            1.934           0.053
L8.cases                                 -0.392406         0.185877           -2.111           0.035
L8.mobility_grocery_and_pharmacy          0.635195         2.069717            0.307           0.039
L9.cases                                  0.121790         0.253371            0.481           0.631
L9.mobility_grocery_and_pharmacy         -0.609680         2.067425           -0.295           0.068
L10.cases                                -0.208361         0.280461           -0.743           0.458
L10.mobility_grocery_and_pharmacy         3.184462         1.884824            1.690           0.041
```

From the results of the fitted model summary above, we see that mobility to grocery and pharmacy places has a significant 7-day lagged impact on daily case rates. That is, reducing the mobility to grocery and pharmacy by 1% reduced the daily case rate by 4.5% after a duration of 7 days (holding all other variables constant)

```
--------------------------------------------------------------------
Results for equation cases
===================================================================================================
===================================================================================================
                                     coefficient        std. error          t-stat            prob
---------------------------------------------------------------------------------------------------

L7.mobility_transit_stations            0.849025          0.356713           2.380           0.017
L10.mobility_retail_and_recreation      0.596583          2.169144           1.197           0.031
```

We see similar results obtained when the VAR model is fitted with other types of mobility. For example, we see that mobility to transit stations also has a significant 7-day lagged impact on daily case rates. That is, reducing the mobility to transit stations by 1% reduced the daily case rate by 0.84% after a duration of 7 days (holding all other variables constant). Mobility to retail and recreation spots has a significant 10-day lagged impact on daily case rates. That is, reducing the mobility to retail and recreation spots by 1% reduced the daily case rate by 0.59% after a duration of 7 days (holding all other variables constant). Overall we see an average lag impact of 8.5 days on the case rates when all types of mobility are considered.

Discussion/Implications

This analysis is mainly aimed at mitigative measures, which focuses on slowing but not necessarily stopping the epidemic spread, reducing peak healthcare demand while protecting those most at risk of severe disease from infections in the future, and suppressive measures, which aims to reverse epidemic growth, reducing case numbers to low levels and maintaining that situation indefinitely. Maintenance can be carried out but controlling human mobility patterns and government interventions during pandemic-like situations. Through the analysis we see that common policies like school closures, travel restrictions, bans on public gatherings, stay-at-home orders, closure of public transportation, emergency investments in the healthcare system and other government policies help in limiting social mobility and in turn control COVID-19 spread. Therefore, further targeted research can be performed in the future for the suppression of social contact by limiting mobility in workplaces, schools, and other public spheres with the aim of quickly reducing the transmission rate.

Through the analysis, the effectiveness of different implemented government policies in the alleviation of the COVID-19 pandemic has been demonstrated. A deterministic stage-structured multi-linear regression model showed the positive effects of extended workplace distancing, reduction in mixing in the community, and school closure in the control of the pandemic situation in Jefferson County, Kentucky. A novel machine learning model like this can be employed to future examine the role of selected socioeconomic factors in mediating local and cross-city transmission of coronavirus. All these studies help to show that implemented government policies have a positive effect on reducing the spread of COVID-19. However, a lot of the researchers projected the effects of these fast-changing policies by the comparison of the number of daily confirmed cases under when these policies were implemented and when not, to show the impact of these policies in the suppression and mitigation of COVID-19. The direct benefits of these policies cannot be observed but are currently only inferred from the fitted model summary. Therefore, a direct measure to observe if these policies had a positive impact on reducing the number of daily confirmed cases has not been done. Also, this analysis has mainly focused on the effects of government actions and mobility changes in Jefferson County only. Hence, a similar systematic evaluation of these policies has to be made on a global scale in order to mitigate pandemic-like situations in the future.

Limitations

The COVID-19 Mobility dataset demonstrates how visits to sites like grocery shops and parks have altered (during the pandemic) from a pre-pandemic baseline value. The baseline is the median value on the relevant weekday from January 3 to February 6, 2020. As time passes and we move further away from the baseline period, Jefferson County populations might vary due to relocation or new regional and remote working options. Google's understanding of categorized places might also have changed. For example, the same value today and in April 2020 might not indicate the same behavior or adherence; it might be that Google has updated information about shops and restaurants in the region or that fewer people live there now. These differences could shift the values up or down over long time periods. Here, I would like to highlight that this potential variation in the population over time which in turn affects the changes in the mobility change rates is not considered during my analysis due to a lack of information on new baselines at different periods during the pandemic.

The Government response dataset from the University of Oxford provides data about the government restriction that can at most be filtered down to the state level. Here, this dataset is filtered out for Kentucky state, and data is inherited into our analysis at County (Jefferson) level. Also, some of the government mandates and policies namely, international support, public information campaigns, and emergency investment in healthcare, have been removed from the dataset during the data cleaning phase of analysis as these divisions comprised a lot of missing information. Using complete information about all the government restrictions information in the future will help with deriving more holistic inferences w.r.t the said interventions in the future.

For the statistical test and the regression models used in the analysis, we see that all the data assumptions have been met. However, for the multi-linear regression models fitted with government response and mobility information, we see that the R-squared error of the models is not very high (not close to 0.9). As it means that the model is not the best-fit version of itself, the model results have to be taken into account with a pinch of salt as there is no universal rule on how to incorporate this statistical measure in assessing the pre-defined model.

It is crucial to highlight that the inclusion of Google data in the computation of mobility changes is dependent on user choices, connection and if it fulfills Google's privacy limits. As a result, when the data does not match Google's quality and privacy requirements, there may be empty entries for specific locations and dates. It's also worth noting that this data is based on users who have enabled Location History in their Google Accounts, thus it reflects a sample group of users from any given region. As with all samples, this may or may not accurately represent the behavior of a larger population.

Conclusion

In summary, this analysis helps to understand the effects of government interventions and human mobility on the spread of COVID-19 and also helps identify how these interventions and human mobility, in general, had a significant impact in reducing the COVID-19 spread in Jefferson County, Kentucky, US. Starting with the change point detections in the infection rates, we see the variations in the spread of the pandemic, before, during, and after mask mandates. Through the dependent t-tests, we learn that the social mobility changes are different prior to and post-mask mandates in Jefferson County. We also learn that human mobility of different types did reduce by significantly varying percentages post the introductions of varied combinations of government policy implementations like canceling of public events, stay-at-home orders, and travel restrictions. We also learn that there's a significant reduction in the percentage daily case rate again with the introduction of varied combinations of government policy implementation on the travel ban, canceling public events, and stay-at-home orders. Through the times-series regression analysis, we also see that there is a lagged impact (on an average of 7 to 10 days) between case rates and mobility, showing clear indications that a reduction in mobility overall leads to a reduction of the spread. Overall, we learn that understanding how well mobility information reflects the population infection rates and whether that relationship changes with government measures are important for tracking the trajectory of the pandemic in Jefferson County and assessing the effectiveness of the previous issues control measures.

Analysis of mobility data from Google has undoubtedly been beneficial to better understand how changes in mobility in outdoor and indoor settings may impact COVID-19 infection spread. Moreover, assessing the effectiveness of government interventions has also aided in gaining insights into policies that can be adjusted in the future if a similar pandemic situation were to arise. In general, this analysis helps with the internal auditing of similar data by stakeholders involved in data mining and analysis for validating the quality of data to ensure that we do not misinform policymakers or the public. Importantly, the design decisions in this analysis respect the privacy of individuals in Jefferson County as it uses their powerful data on mobility, government response, and daily case rate to inform and modulate public policy decisions.

References

[1] Estimating the impact of mobility patterns on COVID-19 infection rates in 11 European countries

[2] Effects of government policies on the spread of COVID-19 worldwide

[3] Detecting the Change Points in a Time Series

[4] Paired Samples T-Test

[5] Multiple Linear Regression Using Python and Scikit-learn

[6] A Multivariate Time Series Guide to Forecasting and Modeling (with Python codes)

[7] Utilization of Mobility Data in the Fight Against COVID-19

[8] Exploring the relationship between mobility and COVID− 19 infection rates for the second peak in the United States using phase-wise association

[9] How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study

[10] Association Between Population Mobility Reductions and New COVID-19 Diagnoses in the United States Along the Urban–Rural Gradient, February–April 2020

[11] Public mobility data enables COVID-19 forecasting and management at local and global scales

[12] Reduction in mobility and COVID-19 transmission

[13] Mobility network models of COVID-19 explain inequities and inform reopening

[14] The relationship between mobility and COVID-19 pandemic: Daily evidence from an emerging country by causality analysis

Data Sources

Data Source 1: COVID-19 Data from John Hopkins University

This dataset consists of the cumulative confirmed case counts were gathered from the Kaggle repository of the John Hopkins University; Raw United States confirmed COVID-19 cases dataset.
Link-
https://www.kaggle.com/datasets/antgoldbloom/covid19-data-from-john-hopkins-university?select=RAW_us_confirmed_cases.csv

Data Source 2: U.S. State and Territorial Public Mask Mandate from April 10, 2020, through August 15, 2021, by County by Day

This dataset consists of the data for masking mandates was sourced from the CDC dataset of masking mandates by county.
Link -
https://github.com/aaliyahfiala42/data-512-a7#us-state-and-territorial-public-mask-mandates-from-april-10-2020-through-august-15-2021-by-county-by-day-mask-use-by-countycsv

https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i

Data Source 3: COVID-19 Mobility dataset from Google

The dataset shows how visits to places, such as grocery stores, public transport hubs, and parks, to name a few, have changed since Feb 2020 compared to a baseline. The baseline is the median value, for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020.

Link-
https://github.com/GoogleCloudPlatform/covid-19-open-data/blob/main/docs/tablemobility.Md

Data Source 4: Covid-19 Government Response dataset from the University of Oxford

This dataset consists of the government's response to the pandemic through different policies like school closures, restrictions on travel/ internal movement, and workplace closures, to name a few, including a stringency index associated with each government response.
Link-
https://github.com/GoogleCloudPlatform/covid-19-open-data/blob/main/docs/tablegovernment-response.md