Will my government provide income support during the COVID-19 pandemic? A study on the factors for governments to make these decisions

Jeffray Tsai

December 20, 2021

1 Introduction

The COVID-19 pandemic has impacted many different facets of livelihoods, from health systems to workplaces. With early shutdowns due to the uncertainties the coronavirus could bring and what the disease would do, lockdowns were imposed all around the world, leading to a slowdown in many parts of people's lives. The economy was fully tanked during the shutdown, but has slowly been improving again due to vaccine access and more knowledge of the coronavirus, except the uncertainties of the variants that happen to come along during the pandemic.

A lot of people lost jobs or had their salaries cut, and governments around the world have tried to provide some income support to alleviate some of the stress caused by the pandemic. There are some determining factors on when governments should provide support, and also some correlations on the kind of responses governments have already made. In this study, we want to see how governments provide income support to citizens during the COVID-19 pandemic.

The independent variables are COVID-19 cases, vaccination rates, government response, and public health response. The dependent variable is income support provided by governments. We will be sampling all countries/territories around the world.

The main hypothesis is countries and territories that have better control of cases and vaccination rates, and have better government and public health responses will generally have provided economic relief. Doing this research will give us an idea of different country's economic situation during the pandemic. We will be running regression models to understand the relationships between the variables.

2 Description of Dataset and Variables

Our World in Data provides datasets that relate to the COVID-19 pandemic. It is updated on a regular basis, thus, our data ranges from the beginning of the pandemic in January 2020 to the submission of this paper in mid-December 2021.

2.1 Dependent Variable

Our World in Data has datasets that can help us understand government's economic responses during the COVID-19 pandemic. The dependent variable will come from the "Income support" dataset. This is a good predictor on how the government was able to support citizens in terms of economic relief packages.

- Income_support: This categorical variable has three categories, no income support (0), covers up to 50% of lost salary (1), and covers more than 50% of lost salary (2). This is determined on a daily basis.

2.2 Independent Variables

Our World in Data provides many COVID-19 related datasets. The datasets that will be used for independent variables include "COVID-19 dataset," "Stringency Index," and "Daily number of COVID-19 vaccinations administered per 100 people". These datasets will help us understand the correlations of government and public health related variables with the economic response dependent variables. [1]

- New_cases: This variable is a daily count of new COVID-19 cases for each day.

- New_deaths: This variable is a daily count of deaths related to COVID-19 each day.

- Stringency_index: This variable is calculated as an index based on school closures, workplace closures, public event cancellations, public gathering restrictions, public transport closures, stay at home implementations, public information, restrictions of domestic movement, and border control. An index rate is given on a daily basis. 0 is the lowest score, meaning low stringency, while 100 is the highest score, meaning highest stringency.

- Containment_index: This variable is based on testing policy, contact tracing, face covering policy, and vaccine roll out policies. This is also calculated amongst the same metrics as the stringency index. An index rate is given on a daily basis. 0 is the lowest score, meaning low stringency, while 100 is the highest score, meaning highest stringency.

- New_vaccinations_smoothed_per_million: This variable gives us vaccination rates per million people on a daily basis from when the country started administering vaccines.

3 Descriptive Statistics

Within this study, we have different types of variables, including categorical and continuous variables. As such, we get a summary statistic of each variable that will be used to understand how each of the variables work. For the data, we filled NaN values with 0 and filtered for datapoints to be greater than 0 for new_cases and new_deaths since there were negative numbers which does not work with counting daily cases and deaths of COVID-19.

income_support	counts
0 (no support)	70,685
1 (up to 50% support)	46,896
2 (more than 50% support)	27,469

Table 1: Counts of income_support.

As seen from Table 3, there are 70,685 counts of where governments did not provide any income support, 46,896 counts of where governments provided up to 50% of income support, and 27,469 counts of where governments provided more than 50% of support, all on a daily basis.

We set income_support as binary, where 0 is no government income support and 1 is government income support provided, no matter how much. As seen from Table 4, there are 70,685 counts of

income_support	counts
0 (no support)	70685
1 (yes support)	74365

Table 2: Counts of income_support as binary.

where governments did not provide any income support each day. There are 74,365 counts of where governments provided income support each day.

Summary	Min	1st Q	Median	Mean	3rd Q	Max
new_cases	0.0	1.0	67.0	7864	899	908289
new_vaccinations_smoothed_per_million	0.0	0.0	0.0	0.1618	0.2040	11.7497
new_deaths	0.0	0.0	1.0	152.8	14.0	18007.0
containment_index	0.0	30.36	52.38	45.05	64.64	93.95
stringency_index	0.0	23.15	49.07	45.27	67.59	100.00

Table 3: Summary statistics of independent variables.

Table 5 shows the summary statistics of the independent variables. The highest containment_index for governments is 93.45, while the lowest is 0.0. The highest stringency_index is 100.00, while the lowest is 0.0. The highest vaccination rate per million people is 11.75, while the lowest is 0. The highest new cases per day is 908,289, while the lowest is 0. The highest new deaths per day is 18,007, while the lowest is 0.

4 Initial Models

We run the an initial model in the form of a linear regression model $y_i = \beta_0 + \beta_1 X_i + \epsilon_i$. Our equation will be:

$$\begin{aligned} income\_support = \beta_0 + \beta_1 containment\_index + \beta_2 stringency\_index + \\ \beta_3 new\_vaccinations\_smoothed\_per\_million + \beta_4 new\_cases + \beta_5 new\_deaths + \epsilon \quad (1) \end{aligned}$$

From this model, we wee the model summary in Figure 1. For every increase in income support, on average, there is a slight increase in public health responses, government restrictions, and deaths.

```
Call:
lm(formula = income_support ~ containment_index + stringency_index +
    new_vaccinations_smoothed_per_million + new_cases + new_deaths,
    data = data)

Residuals:
     Min       1Q   Median       3Q      Max
-1.28374 -0.61215 -0.08982  0.31199  1.90509

Coefficients:
                                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                            9.491e-02  4.227e-03  22.452  < 2e-16 ***
containment_index                      1.065e-02  7.856e-05 135.590  < 2e-16 ***
stringency_index                       2.844e-03  7.191e-05  39.550  < 2e-16 ***
new_vaccinations_smoothed_per_million -5.242e-03  5.732e-03  -0.914     0.36
new_cases                             -9.921e-07  1.093e-07  -9.077  < 2e-16 ***
new_deaths                             4.721e-05  5.919e-06   7.976 1.52e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6998 on 145044 degrees of freedom
Multiple R-squared:  0.167,     Adjusted R-squared:  0.167
F-statistic:  5815 on 5 and 145044 DF,  p-value: < 2.2e-16
```

Figure 1: Summary of linear regression model.

The vaccination rates and number of new cases decreases slightly. All of the variables are significant except for vaccination rates. We note that the R-squared value, both multiple and adjusted, is 0.167, which is really low. This means there is high variability around the regression line.

Our main hypothesis is therefore slightly correct, as there is better government and public health responses. Vaccination rates slightly increases and new cases slightly decrease, however, the only variable that does not follow the main hypothesis is new deaths, as it does slightly increase.

Although it does seem like this linear regression model works to predict the outcome of the predictor variables, we also want to test out when we set the predictor variables into binary. A similar equation is used when the predictor variables are set to binary. We will also see if the R-squared value will be higher.

$$income = \beta_0 + \beta_1 containment\_index + \beta_2 stringency\_index + \\ \beta_3 new\_vaccinations\_smoothed\_per\_million + \beta_4 new\_cases + \beta_5 new\_deaths + \epsilon \quad (2)$$

In Figure 2, the model with the dependent variable set to binary generated similar results as the model with the original dependent variable. For every increase in income support from governments, there is a slight increase in government restrictions and public health responses. Daily deaths slightly increase, while new cases and vaccination rates slightly decrease on average. All variables in this model are significant, and vaccination rates is more significant in this model compared to the previous model. The main hypothesis is therefore not entirely correct, since daily deaths did still increase and

```
Call:
lm(formula = income ~ containment_index + stringency_index +
    new_vaccinations_smoothed_per_million + new_cases + new_deaths,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9336 -0.4472  0.1662  0.3706  0.9291

Coefficients:
                                       Estimate Std. Error t value Pr(>|t|)
(Intercept)                           7.440e-02  2.686e-03  27.701  < 2e-16 ***
containment_index                     7.932e-03  4.991e-05 158.925  < 2e-16 ***
stringency_index                      1.831e-03  4.569e-05  40.070  < 2e-16 ***
new_vaccinations_smoothed_per_million -8.309e-03  3.642e-03  -2.282   0.0225 *
new_cases                            -6.218e-07  6.944e-08  -8.954  < 2e-16 ***
new_deaths                            2.829e-05  3.760e-06   7.523 5.37e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4446 on 145044 degrees of freedom
Multiple R-squared:  0.2087,    Adjusted R-squared:  0.2087
F-statistic:  7653 on 5 and 145044 DF,  p-value: < 2.2e-16
```

Figure 2: Summary of linear regression model with dependent variable set to binary.

vaccination rates slightly decrease for this model. The R-squared value is slightly higher than the previous model, with both multiple and adjusted being 0.2087. This means that there is still high variability around the regression line.

Although when the model is set to having the dependent variable being binary, a linear regression model does not work the best for this study. The logistic regression model works better in this instance because we have set income support into binary: whether governments have provided income support or not. We get the form of the logit model $log(\frac{P(X=1)}{P(X=0)}) = \alpha + \beta X_1$.

We have the logit model of

$$log(\frac{P(income=1)}{P(income=0)} = \alpha + \beta_1 containment\_index + \beta_2 stringency\_index + \beta_3 new\_vaccinations\_smoothed\_per\_million + \beta_4 new\_cases + \beta_5 new\_deaths \quad (3)$$

From Figure 3, we observe slightly different results, especially because a couple of the coefficients have changed. For every increase in income support from governments, there is a slight increase in containment measures, government restrictions, and daily deaths. However, there is also a slight decrease in vaccination rates and new cases. All of the variables are significant except for vaccinations. The main hypothesis is not entirely correct, with similar results as the linear regression model with income support set to binary.

In Table 4, we get the odds ratio of each of the independent variables. The odds increase by 3.9% for containment, 1% for stringency, -2.3% for vaccinations, about 0% for new cases, and about 0% for

```
Call:
glm(formula = income ~ containment_index + stringency_index +
    new_vaccinations_smoothed_per_million + new_cases + new_deaths,
    family = binomial, data = data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0948  -1.0437   0.6129   0.9518   2.1416

Coefficients:
                                       Estimate Std. Error  z value Pr(>|z|)
(Intercept)                          -2.169e+00  1.688e-02 -128.512  < 2e-16 ***
containment_index                     3.827e-02  2.803e-04  136.517  < 2e-16 ***
stringency_index                      9.934e-03  2.321e-04   42.810  < 2e-16 ***
new_vaccinations_smoothed_per_million -2.309e-02  1.778e-02   -1.299    0.194
new_cases                            -3.926e-06  4.022e-07   -9.761  < 2e-16 ***
new_deaths                            1.498e-04  2.161e-05    6.932 4.14e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 200989  on 145049  degrees of freedom
Residual deviance: 167714  on 145044  degrees of freedom
AIC: 167726

Number of Fisher Scoring iterations: 4
```

Figure 3: Summary of logistic regression model with dependent variable set to binary.

new deaths. Since the odds of income support are greater than 1 for containment and stringency, they have higher odds of outcome.

variable	odds ratio
intercept	0.1142781
containment_index	1.0390096
stringency_index	1.0099838
new_vaccinations_smoothed_per_million	0.9771699
new_cases	0.9999961
new_deaths	1.0001498

Table 4: Odds of independent variables.

In order to make the variables slightly more significant, we use interaction terms for a couple of the independent variables. In this case, we will make vaccinations and deaths as the interactions.

$$log(\frac{P(income = 1)}{P(income = 0)} = \alpha + \beta_1 containment\_index + \beta_2 stringency\_index + \beta_3 new\_vaccinations\_smoothed\_per\_million * \beta_5 new\_deaths + \beta_4 new\_cases \quad (4)$$

From Figure 4, we observe that this will help make the vaccinations variable slightly more significant than the previous models. For every increase in income support, on average, there is a slight increase in containment measures, government restrictions, and daily deaths. There is a slight decrease in vaccination rates and new cases. The main hypothesis is still rejected because of the increase in daily deaths and decrease in vaccination rates and new cases.

```
Call:
glm(formula = income ~ containment_index + stringency_index +
    new_vaccinations_smoothed_per_million * new_deaths + new_cases,
    family = binomial, data = data)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-2.093  -1.043   0.613   0.952   2.142

Coefficients:
                                                   Estimate Std. Error  z value Pr(>|z|)
(Intercept)                                      -2.167e+00  1.691e-02 -128.159  < 2e-16 ***
containment_index                                 3.826e-02  2.803e-04  136.479  < 2e-16 ***
stringency_index                                  9.939e-03  2.321e-04   42.826  < 2e-16 ***
new_vaccinations_smoothed_per_million            -3.490e-02  1.900e-02   -1.836   0.0663 .
new_deaths                                        1.497e-04  2.169e-05    6.900 5.19e-12 ***
new_cases                                        -4.054e-06  4.096e-07   -9.897  < 2e-16 ***
new_vaccinations_smoothed_per_million:new_deaths  6.075e-05  3.487e-05    1.742   0.0815 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 200989  on 145049  degrees of freedom
Residual deviance: 167711  on 145043  degrees of freedom
AIC: 167725

Number of Fisher Scoring iterations: 4
```

Figure 4: Summary of logistic regression model with interaction terms.

In Table 5, we get a similar odds ratio of each of the independent variables. The odds increase by 3.9% for containment, 1% for stringency, -3.4% for vaccinations, and about 0% for daily deaths, daily cases, and the interactive variables. Since the odds of income support are greater than 1 for containment and stringency, they have higher odds of outcome.

variable	odds ratio
intercept	0.1145006
containment_index	1.0390032
stringency_index	1.0099886
new_vaccinations_smoothed_per_million	0.9657022
new_deaths	1.0001497
new_cases	0.9999959
new_vaccines_smoothed_per_million * new_deaths	1.0000608

Table 5: Odds of independent variables with interaction terms.

5 Final Models

Initially, it looked like the linear regression model without setting the dependent income support variable into binary, was going to be a good option, as all of the variables were significant except for vaccination rates. In order to improve that, we set income support into binary. Both of these linear regression models resulted in low R-squared values.

A linear regression model would be insufficient since it does not model the probability of binary event. Hence, we use a logistic regression model in order for the study to be more accurate. In this case, the logistic regression model did work, but the vaccination rates variable was not statistically significant. We then made vaccination rates and daily deaths as interactive variables, which helped the statistical significance of vaccination rates.

We run the ANOVA function to get the Analysis of Deviance table for the two logit models, and determine which model is more statistically significant. From the Analysis of Deviance table in Figure 5, we conclude that the logit model with the interactive terms is the best model to use for this study, as it is more statistically significant.

```
Analysis of Deviance Table

Model 1: income ~ containment_index + stringency_index + new_vaccinations_smoothed_per_million +
    new_cases + new_deaths
Model 2: income ~ containment_index + stringency_index + new_vaccinations_smoothed_per_million *
    new_deaths + new_cases
  Resid. Df Resid. Dev Df Deviance Pr(>Chi)
1    145044     167714
2    145043     167711  1   3.2711  0.07051 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 5: Analysis of Deviance.

Th best model for this study is the logistic regression model with interactive terms:

$$log(\frac{P(income=1)}{P(income=0)} = \alpha + \beta_1 containment\_index + \beta_2 stringency\_index + \beta_3 new\_vaccinations\_smoothed\_per\_million * \beta_5 new\_deaths + \beta_4 new\_cases \quad (5)$$

6 Conclusion

Different governments have different ways of responding on the COVID-19 pandemic. This includes restricting movement, international borders, mask policies, and economic relief. From this study, our independent variables generally were daily cases and deaths of COVID-19, vaccination rates, government stringency, and public health response. The dependent variable was income support provided by governments around the world. All of the countries and territories around the world were sampled for the study.

Initially, our main hypothesis was that countries and territories around the world that have better control of cases and vaccination rates, while also having better government and public health responses, will have better economic relief in the form of income support. From the final logit model, for every increase in income support, on average, there is a slight increase in containment, stringency, daily deaths, and vaccination rates, while there is a slight decrease on daily cases and the interactive terms of vaccination rates and daily deaths. This slightly differs from the main hypothesis, as daily deaths did slightly increase and vaccination rates and new cases slightly did slightly decrease.

From this study, we are able to understand that government stringency in terms of lockdowns, restrictions, and domestic movement, and also public health responses, are signs of where governments are taking the pandemic seriously and also acknowledging the fact that their citizens may need some income support for lost jobs or decrease in salary. It can be also understood from daily COVID-19 cases and deaths that governments may have implemented income support when there were more cases, hence why the main hypothesis is partially incorrect.

In future studies, we need to take into account the contexts where governments may not provide as much assistance when it is not needed, in this case being providing income support during lower number of deaths and cases. For the thesis, it would be interesting to rephrase what the hypothesis is while layer in different variables as well. There are also other related economic relief packages, such as debt relief, that we can study on.

References

[1] Lucas Rodés-Guirao Cameron Appel Charlie Giattino Esteban Ortiz-Ospina Joe Hasell Bobbie Macdonald Diana Beltekian Hannah Ritchie, Edouard Mathieu and Max Roser. Coronavirus pandemic (covid-19). *Our World in Data*, 2020.