The Impact of a Country's Health Expenditure on COVID-19 Cases

Sachin Kadam

Abstract

The coronavirus pandemic has upended modern society and forced governments in countries around the world to reevaluate their current healthcare systems. This study attempts to understand the relationship between a country's healthcare capabilities (as measured by current health expenditure) and their response to the coronavirus pandemic (as measured by COVID-19 cases). Other explanatory variables in this study are HDI indexes, life expectancy, GDP per capita and COVID-19 deaths will be used to further understand the relationship. This study uses simple and multiple regression models to estimate the relationship between current health expenditure and COVID-19 cases. A positive correlation between health expenditure and COVID-19 cases is hypothesized and supported by the linear regression models constructed in this study.

I. Introduction

The COVID-19 pandemic has impacted the entire world, and as time moves on, reflections on the preparedness of countries, and preparations for the next pandemic will need to take place. It is key to understanding the current response to the COVID-19 pandemic in order to develop a relationship between specific factors that have led to different outcomes in countries with similar levels of preparedness. Understanding the relationship between health expenditure as a percent of GDP and the number of COVID-19 cases in a country will be a good indicator of the relationship between a country's healthcare capabilities and it's COVID-19 response.

The definition of current health expenditure is spending on healthcare goods and services, expressed as a percentage of GDP. It excludes capital health expenditures such as buildings, machinery, information technology and stocks of vaccines for emergency or outbreaks. This metric was invented as a way to measure the progress of the United Nation's (UN) Sustainable Development Goal (SDG) three (good health and well-being). This specifically deals with subgoal SDG 3.C which the mission statement is to substantially increase health financing and the recruitment, development, training and retention of the health workforce in developing countries, especially in least developed countries and small island developing States.

This analysis of COVID-19 cases and current health expenditure will be beneficial to not only trying to complete the UN's Sustainable development goals, but to be better prepared for the next pandemic. Understanding the impact of current health expenditure will be beneficial for all countries, as health is also one of the main factors of the Human Development Index (HDI), which is the ultimate criteria for assessing the development of a country. People who live in high HDI ranked countries tend to be happier, healthier, more educated and have a higher standard of living. Improving one aspect will lead to increases in other dimensions of HDI.

This study will utilize cross-sectional data from a variety of sources to create several different simple and multiple regression models. The hypothesis is that as health expenditure in a country increases, the COVID-19 cases in that country will also increase. The rationale behind this is as countries spend more money on healthcare goods and services and therefore improve their healthcare infrastructure the availability and ease of testing will lead to higher COVID-19 cases. There are speculation that actual COVID-19 cases are much higher than confirmed COVID-19 cases in just about every country in the world. So, as countries can test more of their population, they will also reveal more positive COVID-19 cases.

II. Literature Review

To study the effect of public health expenditure on COVID-19 Khan, Awan, Islam and Muurlink (2020) used data from the Johns Hopkins University database, World Bank records and the National Civic Space Ratings 2020 database. They created a regression model that was used to assess the association between case fatality (a ratio of total number of confirmed deaths to total number of confirmed cases) and healthcare capacity index adjusting for other omitted variables. The regression analysis shows that greater healthcare capacity was related to lesser case-fatality with every additional unit increase in the healthcare capacity index associated with a 42% decrease in the case fatality. The paper did address health expenditure, but the variable did not reach statistical significance, even though it was negatively correlated with case fatalities. The research suggests that building effective multidimensional healthcare capacity is the most promising means to mitigate future case fatalities. The data also suggests that government's ability to implement public health measures determines mortality outcomes. Therefore, this paper reached the conclusion that as public health spending increases, the mortality rate of public health crises will drop.

Eissa (2020) completed research and analyzed global health spending patterns pre- and post-COVID-19. They also analyzed country specific case studies and proposed policy recommendations. The study focused on different scenarios for restructuring public health spending to build preparedness and resilience in healthcare systems for pandemic and public health crises. This paper raises awareness for the importance of formulating pandemic preparedness and executing it in all three stages of the pandemic before, during and after. Public health expenditure, a part of pandemic preparedness, along with efficient healthcare systems, plus traditional factors which includes the time element of quick response to the pandemic, are measures of sustainable health. Investment in national healthcare ensures efficient resource allocation. The policy recommendations call for restructuring public expenditure to expand the absorptive capacity of healthcare institutions, eventually leading to sustainability and universal health insurance. This paper did not have regression models, but instead drew conclusions from simple correlation models that dealt with public health expenditure in specific countries. Eissa (2020) formed conclusions through researching SDGs and understanding public health spending patterns. They came to the conclusion that a universal health insurance would be able to restructure and expand healthcare capacity. In relation to COVID-19 this would mean that there would be a lower mortality rate as spending on public health would increase.

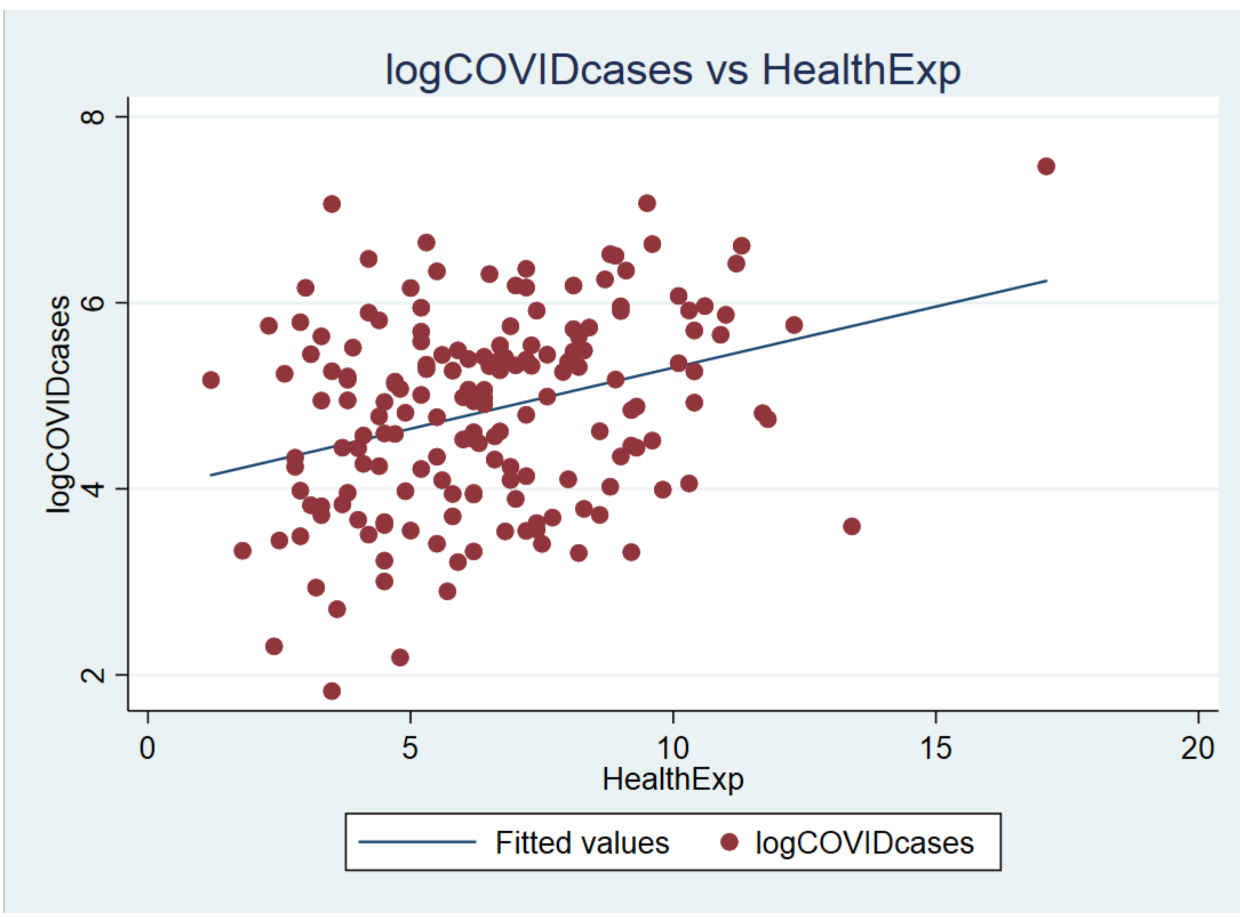
Stribling, Clifton, McGill and Vries (2020) completed a cross-sectional data analysis with the top 36 countries, as ranked by the Global Health Security Index scores. This security index ranks countries based on their health security capabilities which includes their preparedness for an infectious disease outbreak or any other public health crises. This index provided data about each nation's expenditure on health and the nursing workforce, and Stribling, Clifton, McGill and Vries (2020) compared these with mortality data for COVID-19. They reached the conclusion that the extent of a country's pandemic preparedness, magnitude of the nursing workforce and public healthcare expenditure did not significantly impact a nations COVID-19 mortality rate. This can be seen as an unexpected result, because the other two papers found that increasing public health expenditure alone would lead to a sharp decrease in mortality rate. Stribling, Clifton, McGill and Vries (2020) also reached the conclusion that this was something of a paradox. Their conclusion was that arrangements for dealing with future public health crises must involve a range of experts in the field of public health and agencies that can coordinate a response. This paper found that coordination was key in developing an appropriate response to the COVID-19 pandemic, as and preparedness may not matter if not coordinated effectively.

There is a lot of research analyzing the impact of public health expenditure on pandemics, however this study will provide a different perspective on the topic. This paper is focusing on how COVID cases can be impacted by spending on the healthcare expenditure. There will be a main focus on one dependent variable of the natural logarithm of COVID-19 cases from 171 different countries. This dependent variable will go through regression models with up to four different explanatory variables that will help to understand the impact of public health expenditure on COVID cases. Much of the previous research was focused on mortality rate or pandemic preparedness, this paper will focus on how COVID-19 cases in general have a relationship with the Human Development Index, health expenditure, GDP per capita, COVID deaths and life expectancy. For these reasons this study should provide a clearer understanding on the relationship between COVID-19 and public healthcare expenditure.

III. Data

To understand the relationship between current health expenditure and COVID-19 cases, cross-sectional data for around 171 countries was found. The data for the dependent variable, total COVID-19 cases by country was then changed to a natural logarithm of those values to prevent outliers from skewing the data. The data for COVID-19 cases was sourced from the World Health Organization and is continuing to be updated every day. Currently the data used in this study is from March 2021. The primary explanatory variable is current health expenditure as a percentage of Gross Domestic Product (GDP). This data was sourced from the World Bank in 2017. Current health expenditure as a percent of GDP was picked instead of total amount spent on health expenditure in US dollars, as the percent of GDP creates an equal opportunity for all countries to have high percentages spent on healthcare goods and services. This is because richer countries would just be able to spend more money on healthcare even if it is a lower percentage of the GDP, current health expenditure can also give an insight as to how important a country views its own healthcare system. An initial scatter plot of the natural logarithm of COVID cases and current health expenditure as a percent of GDP shows a positive, but mild correlation between the two variables. This scatter plot can be seen below in Figure 1.

Figure 1 – Scatterplot of logCOVIDcases vs HealthExp



In addition to the primary explanatory variable of current health expenditure as a percent of GDP, there are three other secondary explanatory variables for the dependent variable of COVID cases, as well as another dependent variable of the natural logarithm of COVID deaths. These other variables include Human Development Index (HDI) values, life expectancy and the natural logarithm of GDP per capita. The data for the HDI values was sourced from the World Bank and was created in 2019. The HDI is a summary index that measures, achievement in three dimensions of human development: a long and healthy life, being knowledgeable, and having a decent standard of living. The HDI is a score from zero to one and the countries with higher scores have higher levels of human development. Because of the dimensions the HDI value covers it should be predicted in the regression models to have a positive coefficient when regressed with COVID cases. This is because as a countries HDI increases the healthcare system of a country should be well funded and equipped to test a large percentage of the population, which can lead to a higher number of COVID-19 cases. The data on life expectancy was sourced from the UNDESA and the data was obtained in 2017. Life expectancy measures the average period a person in a country will live in years. The data for GDP per capita was obtained from the World Bank and is from 2019. GDP per capita is in the units of 2017's Purchase Power Parity in US Dollars. This data was also changed to a logarithm form to prevent skewing of the data by potential outliers. The data for the second dependent variable, total COVID-19 deaths per country was sourced from the World Health Organization. The data was also obtained in March 2021, and this data was also converted to a logarithm form. A summary of each variable can be found below in Table 1.

Table 1 – Variable Descriptions

Variable Name	Description	Year	Units	Source
HealthExp	Current Health Expenditure	2017	Percentage of GDP	World Bank
HDIindex	Human Development Index	2019	Indices	World Bank
LifeExp	Life Expectancy at Birth	2017	Years	UNDESA
logCOVIDcases	Natural Logarithm of COVID-19 Cases	2021	Number of People	WHO
logCOVIDdeaths	Natural Logarithm of COVID-19 Deaths	2021	Number of People	WHO
logGDPcap	Natural Logarithm of GDP per capita	2019	2017 PPP US$	World Bank

Descriptive statistics for each variable can be found below in Table 2.

Table 2 – Variable Descriptive Statistics

Variable	Observations	Mean	Std. Dev.	Min.	Max.
HealthExp	171	6.48	2.56	1.20	17.10
HDIindex	169	0.72	0.15	0.40	0.96
LifeExp	171	72.21	7.80	52.20	84.30
logCOVIDcases	171	4.84	1.05	1.83	7.47
logCOVIDdeaths	171	3.03	1.18	0	5.73
logGDPcap	166	4.09	0.51	2.88	5.06

Data on HDI values, health expenditure, life expectancy, GDP per capita, COVID cases and COVID deaths were sourced from 171 countries. Country names can be found in Appendix A. The sample sizes in each model remained relatively consistent as each variable apart from HDI values and GDP per capita had 171 observations.

Before creating the regression models the Gauss Markov assumptions have to be checked in order to have the best linear unbiased estimate possible. The five Gauss Markov assumptions and the normality assumption for the Classic Linear Model can be seen below.

Gauss Markov and Classic Linear Model Assumptions:

1. **Linearity:** Model is linear in parameters so that $y = B_0 + B_1X_1 + B_2X_2 + ... + B_nX_n + u$, Where B coefficients are the unknown parameters and u is the error term. All models in Section IV. of the paper satisfy this assumption as the coefficients are linear in parameters, with a u error term.
2. **Random:** Data sourced from random sampling; the data was sourced from countries around the world. There was no deliberation for the countries in the sample. So, this proves that the sampling was random.
3. **Non-Collinearity:** The explanatory variables are not perfectly correlated to each other. STATA has commands to check the collinearity between the variables. The results of the STATA output showed that the variables are not perfectly correlated. The specific values can be found in Appendix B.
4. **Exogeneity:** Zero conditional mean, this means the expected value of the error term, u, is zero given any value for any of the explanatory variables. It is difficult to make sure that the error term is not correlated to COVID cases as the amount of COVID cases could be influenced by other factors not taken into consideration in this study. So, close consideration of the results must be maintained.

5. **Homoskedasticity:** The variance of the error term, u, has a constant variance given any value of an explanatory variable. It is also difficult to make sure the error term variance remains constant as omitted variables may shift or influence the variance of the error term. Again, close consideration of the results must be maintained.
6. **Normality:** The population error u is independent of the explanatory variables and is normally distributed with zero mean and variance. While this assumption may not be completely fulfilled because for some of the variable's normality may not be a good assumption. Regardless this assumption cannot be fully verified in this study therefore close consideration of the results will be maintained.

IV. Results

With the Gauss Markov and Classic Linear Model assumptions being fulfilled, STATA has now been used to compute regression models that can test the hypothesis. All STATA regression outputs can be found in Appendix C.

<u>Model 1:</u>

Model 1 is a simple regression model that is designed to test the relationship between the dependent variable natural logarithm of COVID cases and the primary explanatory variable current health expenditure.

Model 1: $logCOVIDcases = B_0 + B_1(HealthExp) + u$

This model has 171 countries. From the STATA output the estimated equation can be written as:

Estimated Equation 1: $logCOVIDcases = 3.99 + 0.13(HealthExp)$

This model has an R-squared value of 0.10, denoting a weak correlation between health expenditure and COVID-19 cases. The coefficient of health expenditure is positive denoting a positive linear relationship which was predicted. Since a log-level model was created this coefficient can be interpreted to describe that a 1% increase in health expenditure will lead to a 13% increase in COVID-19 cases. This simple regression model shows a positive relationship between health expenditure and COVID cases, which currently proves the hypothesis to be correct. Health expenditure is significant at the 1% level.

This simple linear regression model provides a good starting point for testing and understanding the relationship between COVID cases and health expenditure. But there is still a possible omitted variable bias, and to reduce that bias multiple explanatory variables need to be used in a multiple linear regression. Holding added explanatory variables fixed will lead to a better understanding of the ceteris paribus relationship between COVID cases and health expenditure.

Model 2:

Model 2 is a multiple regression model that was created using all the explanatory variables which now includes GDP per capita, health expenditure, HDI values and life expectancy.

Model 2: *logCOVIDcases* = $B_0 + B_1(HealthExp) + B_2(HDIindex) + B_3(LifeExp) + B_4(logGDPcap) + u$

This model has 165 countries. From the STATA output the estimated equation can be written as:

Estimated Equation 2: *logCOVIDcases = 2.09 + 0.06(HealthExp) + 3.06(HDIindex) + 0.03(LifeExp) - 0.45(logGDPcap)*

This model has an R-squared value of 0.27, indicating a low/mild correlation between the explanatory variables and the dependent variable. This model has the greatest number of explanatory variables leading to the highest R-squared value amongst all the models. Because of the omitted variable bias that is present in Model 1 the value for HealthExp in this model has decreased as compared to the value in Model 1. The omitted variable bias will overestimate the impact of health expenditure on COVID cases. The coefficient of health expenditure is still positive but indicates a 1% increase in health expenditure leads to only an 6% increase in COVID cases. HDI index also has a positive coefficient, indicating a positive relationship between HDI values and COVID cases. The result can be interpreted as every 0.01 increase in the HDI value will lead to a 3.06% increase in COVID cases. The most surprising variable from the regression is GDP per capita because it has a negative coefficient indicating a negative relationship with COVID cases. The result can be interpreted as every 1% increase in GDP per capita leads to a 0.45% decrease in COVID cases. This is the first variable to have a negative relationship with COVID cases, and it indicates that richer countries with higher GDP per capita will also have fewer COVID cases. All the variables are at least significant at the 10% level except for the logarithm of GDP per capita.

Model 3:

Model 3 is a multiple regression model that includes the HDI index variable, health expenditure and GDP per capita. Life Expectancy was dropped from this model because of its low statistical significance at the 10% level and the low coefficient of the variable in Model 2.

Model 3: *logCOVIDcases* = $B_0 + B_1(HealthExp) + B_2(HDIindex) + B_3(logGDPcap) + u$

This model has 165 countries. From the STATA output the estimated equation can be written as:

Estimated Equation 3: *logCOVIDcases = 3.25 + 0.07(HealthExp) + 4.17(HDIindex) – 0.46(logGDPcap)*

This model has an R-squared value of 0.26 still indicating a relatively weak correlation between COVID cases and health expenditure. All the coefficients are relatively the same except for HDI index, this could be caused by the relatively high correlation between HDI index and life expectancy, therefore leading to the regression estimating HDI index is the most significant variable at understanding total COVID cases by country. The new coefficient of HDI index can be interpreted as every 0.01 increase in the HDI value can lead to a 4.17% increase in COVID cases. The regression output is still showing that GDP per capita and COVID cases have a negative relationship. This output is indicating that a higher HDI score leads to more COVID cases, even though GDP per capita is a part of one of the dimensions that are included in HDI values. This just shows that countries may have a high GDP per capita, and will not allocate resources to healthcare spending or other dimensions of the HDI value. These countries could have higher levels of corruption, income inequality or may have different priorities when allocating government spending. HDI index is significant at the 1% level, health expenditure is significant at the 5% level, but GDP per capita is still not significant at the 10% level.

Model 4:

Model 4 is a multiple regression model that includes health expenditure and GDP per capita. It was already established that HDI index is statistically significant and has a large positive coefficient for COVID cases, but the relationship between COVID cases and GDP per capita could still be better understood. Therefore, HDI index values were dropped from this model.

Model 4: $logCOVIDcases = B_0 + B_1(HealthExp) + B_2(logGDPcap) + u$

This model has 166 countries. From the STATA output the estimated equation can be written as:

Estimated Equation 4: $logCOVIDcases = 1.50 + 0.10(HealthExp) + 0.66(logGDPcap)$

This model has an R-squared value of 0.20, which is still a weak correlation. The coefficient of health expenditure increased this could be due to this regression having fewer explanatory variables as compared to Model 2 and Model 3. The regression again overestimates the importance of health expenditure as a variable for COVID cases. The new value can be interpreted as a 1% increase in health expenditure leads to a 10% increase in COVID cases. The most surprising outcome of this regression is that GDP per capita now has a positive coefficient, and the coefficient can be interpreted as a 1% increase in GDP per capita leads to a 0.66% increase in COVID cases. Without the variable of HDI index the regression output could have overestimated the relationship between COVID cases and GDP per capita. Not only did the coefficient become positive but the statistical significance of the variable increased to be significant at the 1% level. The variable of health expenditure was also significant at the 1% level.

Model 5:

Model 5 is another simple regression model that is designed to test the relationship between the dependent variable natural logarithm of COVID deaths and the primary explanatory variable current health expenditure. This regression is meant to see if as health expenditure increases this will lead to an increase or decrease in COVID related deaths.

Model 5: $logCOVIDdeaths = B_0 + B_1(HealthExp) + u$

This model has 171 countries. From the STATA output the estimated equation can be written as:

Estimated Equation 5: $logCOVIDdeaths = 1.97 + 0.16(HealthExp)$

Model 5 has an R-squared value of 0.12 this is the second lowest R-squared value of the five regression models. The coefficient of current health expenditure is positive, meaning that as expenditure increases COVID deaths will also increase. This was not a predicted outcome as the literature review concluded that as health expenditure increased the mortality of COVID-19 should decrease. But this could also be because we are not looking at COVID deaths in relation to the population or number of COVID cases, if these other variables were to be considered health expenditure could have a different relationship with COVID deaths. The high correlation between COVID deaths and COVID cases can be the cause of the very similar relationship the two dependent variables have with the explanatory variable of health expenditure. Currently a 1% increase in health expenditure leads to an 16% increase in COVID deaths. Health expenditure is significant at the 1% level in this model.

A summary of all five regression models can be seen in Table 3.

Table 3 – Regression Models Summary

Dependent Variable: logCOVIDcases					Dependent Variable: logCOVIDdeaths
Independent Variables	Model 1	Model 2	Model 3	Model 4	Model 5
HealthExp	0.13*** (0.03)	0.06** (0.03)	0.07** (0.03)	0.10*** (0.03)	0.16*** (0.03)
HDIindex		3.06** (1.51)	4.17*** (1.38)		
LifeExp		0.03* (0.02)			
logGDPcap		-0.45 (0.40)	-0.46 (0.41)	0.66*** (0.15)	
Intercept	3.99*** (0.21)	2.09* (1.07)	3.25*** (0.85)	1.50** (0.60)	1.97*** (0.23)
No. of Observations	171	165	165	166	171
R-squared	0.10	0.27	0.26	0.20	0.12

*Significant at 10%, **5%, ***1%

V. Extensions

After Model 1 and Model 2 were analyzed and individual significance levels were found, it became important to use an F-test to understand the significance of all the secondary explanatory variables. This includes HDI index, life expectancy and GDP per capita. All the variables are included in Model 2 to create an unrestricted regression, and Model 1 would be used as a restricted regression. Therefore, the hypothesis for Model 2 can be seen below:

$$H_0: B_2 = 0, B_3 = 0, B_4 = 0$$

$$H_1: H_0 \text{ is false}$$

An F-value of 11.67 was calculated using STATA's F-test commands. At the 1% significance level for $F_{(3,160)}$ the critical value is 3.78. Since the calculated F-value is greater than the critical F-distribution value. The null hypothesis is rejected; therefore, all the variables are jointly significant at the 1% level. This reaffirms the significance level found in the Results Section. This also shows that GDP per capita is statistically significant, even though individually it was not significant at the 10% level. This shows that the data selected can lead to reliable conclusions on the topic because of the high statistically significant levels.

The F-test will be completed again with Model 2 as the unrestricted model, but now with Model 4 as the restricted model. Therefore, HDI values and life expectancy will be the variables being tested for joint significance. The hypothesis for Model 2 can be seen below:

$$H_0: B_2 = 0, B_3 = 0$$

$$H_1: H_0 \text{ is false}$$

An F-value of 6.18 was once again found using STATA's F-test commands. The critical value of the F distribution at the 1% level for $F_{(2,160)}$ is 4.61. Just like the previous F-test the critical value is smaller than the calculated F-value, therefore the variables are jointly significant at the 1% level. The null hypothesis is rejected because the variables are jointly significant. With the results of the two F-tests it can be concluded that the three secondary explanatory variables are statistically significant in the models, and therefore have major significance in the results.

After analyzing the correlation and mild positive relationship between COVID cases and the current health expenditure of a country it was decided it would be appropriate to test a different functional form as compared to the models tested in the Results Section. Based off the relationship shown in the initial scatter plot of the logarithm of COVID cases and the health expenditure of a country, the line of best fit may be better estimated as a natural logarithm as compared to a linear function. So, a new explanatory variable of the natural logarithm of health expenditure (*logHealthExp*) was created and was added to the simple regression of Model 1, to construct Model 6.

Model 6: $logCOVIDcases = B_0 + B_1(HealthExp) + B_2(logHealthExp) + u$

This model has 171 countries included in it. The assumptions needed for an unbiased regression output have been met, but the correlation coefficient between the logarithm of health expenditure and health expenditure is 0.96, which is a very high value. This means results need to be analyzed with the consideration of near multicollinearity.

From the STATA output the estimated equation can be written as:

Estimated Equation 6: $logCOVIDcases = 4.26 + 0.16(HealthExp) + -0.82(logHealthExp)$

Model 6 has an R-squared value of 0.10 indicating a very weak relationship between the explanatory and dependent variables. This is the same R-squared value that Model 1 has, showing that there has not been any significant change by adding the logarithm of health expenditure to the model. In this model HealthExp is significant at the 10% level, whereas the new variable logHealthExp is insignificant at the 10% level. In Model 1 HealthExp was significant at the 1% level showing a big drop in significance levels. Because of this the logarithm of health expenditure is not a functional form that will be evaluated further.

VI. Conclusions

The hypothesis of a positive correlation between health expenditure and COVID-19 cases was proven to be true by the regression models. Unfortunately, the regression models also had relatively low R-squared values leading to the conclusion that there was a relatively weak positive correlation between current health expenditure and total COVID-19 cases by country. Increasing healthcare expenditure in a country has led to the discovery of more COVID-19 cases, this is not a negative, as learning more about the virus and testing more of the population is required to find solutions to this public health crisis. Increasing current health expenditure will also lead to better preparedness for the next global pandemic. Increasing current health expenditure can also lead to more countries meeting SDG three, which leads to healthier humans around the world.

All the secondary explanatory variables proved to be statistically significant, and therefore verified that the regressions created reliable outputs. GDP per capita, HDI values and life expectancy were all jointly significant after going through the F-test, and each had very different relationships with the dependent variable of COVID cases. It was very peculiar to see the relationship between GDPs per capita and COVID cases flip from a negative relationship Model 2 and Model 3 to a positive relationship in Model 4. There was also Model 5 where a different dependent variable of COVID deaths was used in a simple regression model with current health expenditure. That model also showed a positive correlation between health expenditure and COVID deaths, but that was to be expected after seeing the high correlation between COVID deaths and COVID cases.

Acquiring new secondary explanatory or secondary dependent variables to extend the models created in this study will be critical to understand more relationships of different variables in the COVID-19 pandemic. In the future, there could be time lags between the explanatory variables and the dependent variables to understand the complete impact of health expenditure on the COVID-19 pandemic. As it takes time for healthcare systems to modernize or for new medical supplies to be shipped around the world. Currently, this data can help countries around the world make informed decisions regarding their current COVID-19 pandemic response, and for future public health crises preparedness.

Appendix A. List of Countries:

Afghanistan	Colombia	Haiti	Monaco	Sierra Leone
Algeria	Comoros	Honduras	Mongolia	Singapore
Andorra	Congo	Hungary	Morocco	Slovakia
Angola	Congo (Democratic Republic of the)	Iceland	Mozambique	Slovenia
Antigua and Barbuda	Costa Rica	India	Myanmar	South Africa
Argentina	Croatia	Indonesia	Namibia	South Sudan
Armenia	Cuba	Iran (Islamic Republic of)	Nepal	Spain
Australia	Cyprus	Iraq	Netherlands	Sri Lanka
Austria	Czechia	Ireland	New Zealand	Sudan
Azerbaijan	CÙte d'Ivoire	Israel	Nicaragua	Suriname
Bahamas	Denmark	Italy	Niger	Sweden
Bahrain	Djibouti	Jamaica	Nigeria	Switzerland
Bangladesh	Dominican Republic	Japan	North Macedonia	Tajikistan
Barbados	Ecuador	Jordan	Norway	Tanzania (United Republic of)
Belarus	Egypt	Kazakhstan	Oman	Thailand
Belgium	El Salvador	Kenya	Pakistan	Togo
Belize	Equatorial Guinea	Korea (Republic of)	Panama	Trinidad and Tobago
Benin	Eritrea	Kuwait	Papua New Guinea	Tunisia
Bhutan	Estonia	Kyrgyzstan	Paraguay	Turkey
Bolivia (Plurinational State of)	Eswatini (Kingdom of)	Latvia	Peru	Uganda
Bosnia and Herzegovina	Ethiopia	Lebanon	Philippines	Ukraine
Botswana	Fiji	Lesotho	Poland	United Arab Emirates
Brazil	Finland	Liberia	Portugal	United Kingdom
Brunei Darussalam	France	Lithuania	Qatar	United States
Bulgaria	Gabon	Luxembourg	Romania	Uruguay
Burkina Faso	Gambia	Madagascar	Russian Federation	Uzbekistan
Burundi	Georgia	Malawi	Rwanda	Venezuela (Bolivarian Republic of)
Cabo Verde	Germany	Malaysia	Saint Lucia	Viet Nam
Cambodia	Ghana	Maldives	Saint Vincent and the Grenadines	Yemen
Cameroon	Greece	Mali	San Marino	Zambia
Canada	Grenada	Malta	Sao Tome and Principe	Zimbabwe
Central African Republic	Guatemala	Mauritania	Saudi Arabia	
Chad	Guinea	Mauritius	Senegal	
Chile	Guinea-Bissau	Mexico	Serbia	
China	Guyana	Moldova (Republic of)	Seychelles	

Appendix B. Correlation Coefficients, Gauss Markov Assumption 3

```
. correlate HealthExp HDIindex LifeExp logCOVIDcases logCOVIDdeaths logGDPcap
(obs=165)

             | Health~p HDIindex  LifeExp logCO~es logCO~hs logGDP~p
-------------+------------------------------------------------------
   HealthExp |   1.0000
    HDIindex |   0.3602   1.0000
     LifeExp |   0.3750   0.8104   1.0000
logCOVIDca~s |   0.3324   0.4681   0.4648   1.0000
logCOVIDde~s |   0.3736   0.3927   0.3990   0.9511   1.0000
   logGDPcap |   0.2592   0.9346   0.7478   0.3965   0.3060   1.0000
```

Appendix C. STATA Regression Outputs

Model 1:

```
. regress logCOVIDcases HealthExp

      Source |       SS           df       MS      Number of obs   =       171
-------------+----------------------------------   F(1, 169)       =     19.44
       Model |  19.2558322         1  19.2558322   Prob > F        =    0.0000
    Residual |  167.415183       169  .990622387   R-squared       =    0.1032
-------------+----------------------------------   Adj R-squared   =    0.0978
       Total |  186.671016       170   1.0980648   Root MSE        =     .9953

------------------------------------------------------------------------------
logCOVIDca~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   HealthExp |   .1313338   .0297886     4.41   0.000     .0725282    .1901394
       _cons |   3.990058    .207497    19.23   0.000     3.580438    4.399678
------------------------------------------------------------------------------
```

Model 2:

```
. regress logCOVIDcases HealthExp HDIindex LifeExp logGDPcap

      Source |       SS           df       MS      Number of obs   =       165
-------------+----------------------------------   F(4, 160)       =     14.80
       Model |  48.7086207         4  12.1771552   Prob > F        =    0.0000
    Residual |  131.610315       160  .822564466   R-squared       =    0.2701
-------------+----------------------------------   Adj R-squared   =    0.2519
       Total |  180.318935       164   1.0995057   Root MSE        =    .90695

------------------------------------------------------------------------------
logCOVIDca~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   HealthExp |   .0627364   .0316635     1.98   0.049     .0002041    .1252687
    HDIindex |   3.057319   1.511868     2.02   0.045     .0715277    6.043109
     LifeExp |   .0275345   .0155361     1.77   0.078    -.0031477    .0582168
   logGDPcap |  -.4482232   .4032883    -1.11   0.268    -1.244678    .3482316
       _cons |   2.092435   1.065456     1.96   0.051    -.0117358    4.196605
------------------------------------------------------------------------------
```

Model 3:

```
. regress logCOVIDcases HealthExp HDIindex logGDPcap

      Source |       SS           df       MS      Number of obs   =       165
-------------+----------------------------------   F(3, 161)       =     18.45
       Model |  46.1249249         3   15.374975   Prob > F        =    0.0000
    Residual |   134.19401       161  .833503171   R-squared       =    0.2558
-------------+----------------------------------   Adj R-squared   =    0.2419
       Total |  180.318935       164   1.0995057   Root MSE        =    .91296

------------------------------------------------------------------------------
logCOVIDca~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   HealthExp |   .0708943   .0315348     2.25   0.026     .0086193    .1331694
    HDIindex |   4.173403   1.383581     3.02   0.003     1.441097    6.905709
   logGDPcap |  -.4562281   .4059355    -1.12   0.263    -1.257873    .3454167
       _cons |   3.252261   .8463628     3.84   0.000     1.580857    4.923665
------------------------------------------------------------------------------
```

Model 4:

```
. regress logCOVIDcases HealthExp logGDPcap

      Source |       SS           df       MS      Number of obs   =       166
-------------+----------------------------------   F(2, 163)       =     20.72
       Model |  36.8624288         2  18.4312144   Prob > F        =    0.0000
    Residual |  144.974752       163  .889415654   R-squared       =    0.2027
-------------+----------------------------------   Adj R-squared   =    0.1929
       Total |   181.83718       165  1.10204352   Root MSE        =    .94309

------------------------------------------------------------------------------
logCOVIDca~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   HealthExp |   .1035095   .0305937     3.38   0.001     .0430984    .1639207
   logGDPcap |   .6581989    .149382     4.41   0.000     .3632255    .9531723
       _cons |   1.497595   .5954358     2.52   0.013     .3218328    2.673357
------------------------------------------------------------------------------
```

Model 5:

```
. regress logCOVIDdeaths HealthExp

      Source |       SS           df       MS      Number of obs   =       171
-------------+----------------------------------   F(1, 169)       =     24.10
       Model |  29.7687936         1  29.7687936   Prob > F        =    0.0000
    Residual |  208.790485       169  1.23544666   R-squared       =    0.1248
-------------+----------------------------------   Adj R-squared   =    0.1196
       Total |  238.559279       170  1.40328988   Root MSE        =    1.1115

------------------------------------------------------------------------------
logCOVIDde~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   HealthExp |   .1632962   .0332665     4.91   0.000     .0976247    .2289677
       _cons |   1.969784   .2317233     8.50   0.000     1.512339    2.427229
------------------------------------------------------------------------------
```

Model 6:

```
. regress logCOVIDcases logHealthExp HealthExp

      Source |       SS           df       MS      Number of obs   =       171
-------------+----------------------------------   F(2, 168)       =      9.83
       Model |  19.5615006         2  9.78075032   Prob > F        =    0.0001
    Residual |  167.109515       168  .994699495   R-squared       =    0.1048
-------------+----------------------------------   Adj R-squared   =    0.0941
       Total |  186.671016       170   1.0980648   Root MSE        =    .99735

------------------------------------------------------------------------------
logCOVIDca~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
logHealthExp |  -.8220074   1.482847    -0.55   0.580    -3.749422    2.105408
   HealthExp |   .1882719   .1069621     1.76   0.080    -.0228911     .399435
       _cons |    4.25845   .5269199     8.08   0.000     3.218213    5.298688
------------------------------------------------------------------------------
```

References

Eissa, Noura. 2020. "Pandemic Preparedness and Public Health Expenditure" *Economies* 8, no. 3: 60. https://doi.org/10.3390/economies8030060

Khan JR, Awan N, Islam MM and Muurlink O (2020) Healthcare Capacity, Health Expenditure, and Civil Society as Predictors of COVID-19 Case Fatalities: A Global Analysis. *Front. Public Health* 8:347. doi: 10.3389/fpubh.2020.00347

Stribling, J., Clifton, A., McGill, G. and de Vries, K. (2020), Examining the UK Covid-19 mortality paradox: Pandemic preparedness, healthcare expenditure, and the nursing workforce. J. Adv. Nurs., 76: 3218-3227. https://doi.org/10.1111/jan.14562

UNDESA (2019a). World Population Prospects: The 2019 Revision. Rev 1. New York. https://population.un.org/wpp/.

"WHO Coronavirus (COVID-19)." *World Health Organization.* https://covid19.who.int/table

World Bank (2020a). World Development Indicators database. Washington, DC. http://data.worldbank.org.