Economic Freedom and COVID-19

AAEC 5307: Term Project

Fall Semester 2020

Don Lim

Contents

1 Introduction and Hypothesis

COVID-19 has affected social norms and the economy. The purpose of this project is to see whether or not economic freedom, GDP per capita, income inequality, and the perception of political corruption has had any effect on the deaths and cases of COVID-19. Initially, from economic theory, I would predict countries with greater economic freedom and GDP per capita would fare better over a pandemic (lower COVID-19 cases and deaths per million) while, on the other hand, greater income inequality and perception of political corruption would lead to higher COVID-19 cases and deaths per million. This is because wealthier and economically freer countries would have more money for research, medical supplies, and more wealth saved up in order to work less and more quickly and efficiently change behavior to mitigate the spread of the disease. Countries with higher corruption perception levels could mean citizens not following the mandates or suggestions of governments in order to mitigate the spread of COVID-19, which could potentially cause higher COVID-19 case and death rates.

2 Data and Economic Setup

For my data, I will be using the 2020 Economic Freedom Index from the Fraser Institute, the GDP per capita (in current USD) and the GINI coefficients from the World Bank, and corruption perception index from Our World Data. Equation (1) shows the model with **cases** as the output and equation (2) shows the model with **deaths** as the output

$$\text{cases}_{c,t} = \beta_1 + \beta_2\text{EF}_{c,t} + \beta_3\text{GDP}_{c,t} + \beta_4\text{GINI}_{c,t} + \beta_5\text{CPI}_{c,t} + e_{c,t} \tag{1}$$

$$\text{deaths}_{c,t} = \alpha_1 + \alpha_2\text{EF}_{c,t} + \alpha_3\text{GDP}_{c,t} + \alpha_4\text{GINI}_{c,t} + \alpha_5\text{CPI}_{c,t} + e_{c,t} \tag{2}$$

where the subscripts c and t refer to country and time, EF is the economic freedom index, GDP is the real GDP per capita, GINI is the GINI coefficient, and CPI is the perceived corruption. However, since the GINI index is calculated sporadically, I will be using the GINI coefficients from 2015, GDP per capita numbers from 2019, the economic freedom index from 2020, the CPI from 2018, and the COVID-19 cases and deaths on November 1st, 2020. Unfortunately, because the GINI coefficient and the CPI are not calculated for every country in 2015, I had to limit the number of countries to those that did have a GINI coefficient or a CPI. This resulted in a reduction from 195 countries to 70.

3 Initial Results

3.a For Case Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-23498.63351	14123.72663	-1.66377	0.100973	Multiple R	0.426382219
EF	5323.346989	2028.247284	2.624605	0.010802	R Square	0.181801797
GDP	0.139316045	0.084895147	1.641037	0.10562	Adjusted R Square	0.131451138
GINI	15.25745872	150.3803198	0.101459	0.919498	Standard Error	8944.649522
CPI	-148.4023636	115.3911906	-1.28608	0.202977	Observations	70

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	1155525843	2.89E+08	3.610713	0.010169261
Residual	65	5200439080	80006755		
Total	69	6355964923			

We see that for every point increase in overall economic freedom, COVID-19 case rates increase by about 5323 per million, ceteris paribus. For every thousand current USD increase in GDP per capita, we see an increase in 0.1393 COVID-19 cases per million, ceteris paribus. For every point increase in the GINI coefficient, we see an increase in 15.257 COVID-19 cases, per million, ceteris paribus. And, findally, for every point increase in the CPI, we see a 148.402 decrease in COVID-19 cases per million. Using $\alpha = 0.05$, we see that none of my variables are significant except for economic freedom. Furthermore, looking at R^2, the variation in EF, GDP, GINI, and CPI only explains 18.18% of the variation in COVID-19 case rates.

3.b For Death Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-540.1858825	419.8732	-1.28655	0.202816	Multiple R	0.277813032
EF	90.99257703	60.29618	1.509094	0.136121	R Square	0.077180081
GDP	0.002590403	0.002524	1.026397	0.308509	Adjusted R Square	0.020391163
GINI	5.417279912	4.470539	1.211773	0.229986	Standard Error	265.9085039
CPI	-2.527609133	3.430375	-0.73683	0.463876	Observations	70

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	384384.7	96096.17	1.359069	0.257872
Residual	65	4595977	70707.33		
Total	69	4980361			

We see that for every point increase in overall economic freedom, COVID-19 deaths increase by about 91 per million, ceteris paribus. For every thousand current USD increase in GDP per capita, we see an increase in 0.00259 COVID-19 deaths per million, ceteris paribus. For every point increase in the GINI coefficient, we see a 5.4172 increase in COVID-19 deaths per million, ceteris paribus. Finally, for every point increase in the CPI, we see see a degrees in 2.527 COVID-19 deaths per million, ceteris paribus. Using $\alpha = 0.05$, we see that none of my variables are significant. Furthermore, looking at R^2, the variation in EF, GDP, GINI, and CPI only explains 7.71% of the variation in COVID-19 death rates.

4 Revision To the Initial Model

Perhaps a good revision to my initial model would be to use the logarithm of GDP per capita:

$$\text{cases}_{c,t} = \beta_1 + \beta_2\text{EF}_{c,t} + \beta_3 log(\text{GDP}_{c,t}) + \beta_4\text{GINI}_{c,t} + \beta_5\text{CPI}_{c,t} + e_{c,t} \quad (1)$$

$$\text{deaths}_{c,t} = \alpha_1 + \alpha_2\text{EF}_{c,t} + \alpha_3 log(\text{GDP}_{c,t}) + \alpha_4\text{GINI}_{c,t} + \alpha_5\text{CPI}_{c,t} + e_{c,t} \quad (2)$$

Doing this, we get the following

4.a For Case Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-47147.38532	15267.63	-3.08806	0.002962	Multiple R	0.48894051
EF	2662.654441	2203.929	1.20814	0.231371	R Square	0.239062822
log(GDP)	11779.73992	4221.358	2.79051	0.0069	Adjusted R Square	0.192235919
GINI	71.71412252	147.0928	0.487543	0.627514	Standard Error	8625.980443
CPI	-221.1712806	108.0683	-2.04659	0.044747	Observations	70

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	1.52E+09	3.8E+08	5.105245	0.001221
Residual	65	4.84E+09	74407539		
Total	69	6.36E+09			

Using $\alpha = 0.05$, we see that none of my variables are significant except for GDP per capita. Furthermore, looking at R^2, the variation in EF, GDP, GINI, and CPI has increased from my initial model of 18.18% to 23.9% for explaining the variation in COVID-19 cases. The coefficients for both case rates and death rates in this section will be discussed later in the results section.

4.b For Death Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-1226.464176	445.545	-2.75273	0.007653	Multiple R	0.415923808
FE	6.556921952	64.31579	0.101949	0.919111	R Square	0.172992614
log(GDP)	363.4825944	123.1891	2.950608	0.004406	Adjusted R Square	0.122099852
GINI	7.749605143	4.29251	1.805378	0.075647	Standard Error	251.726212
CPI	-6.414336372	3.153684	-2.03392	0.046045	Observations	70

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	861565.7	215391.4	3.399159	0.013806
Residual	65	4118796	63366.09		
Total	69	4980361			

Using $\alpha = 0.05$, we see that none of my variables are significant except for GDP per capita and CPI. Furthermore, looking at R^2, the variation in EF, GDP, GINI, and CPI has increased from my initial model of 7.71% to 17.3% for explaining the variation in COVID-19 deaths.

5 Choosing the Model

This seems, then, that using the logarithm of GDP per capita results in a more explanatory model, though the R^2 is still low. To formally check to see which model I should use, I can calculate the Akaike and Bayesian information criterion (AIC and BIC):

	AIC	BIC	R^2	$\bar{R}^2$
Cases	18.23779937	18.36628481	0.181801797	0.131451138
Cases log(GDP)	18.16524556	18.293731	0.239062822	0.192235919
Deaths	11.2064823	11.33496774	0.077180081	0.020391163
Deaths log(GDP)	11.09686181	11.22534726	0.172992614	0.122099852

As we can see, for both cases with log(GDP) and deaths with long(GDP), the AIC and BIC were lower while at the same time, the corresponding R^2 and $\bar{R}^2$ were higher. Therefore, I should choose my revised model with the log(GDP).[1]

6 Dummy Variable: OECD Countries

To see if being an OECD country has any effect on COVID-19 case and death rates, we can include a dummy variable such that:

$$D = \begin{cases} 1 \text{ if OECD country} \\ 0 \text{ if otherwise} \end{cases}$$

[1] I should have tested an additional model with the logarithm of COVID-19 case and death rates. See section 11 Addendum to view what I should have done. Unfortunately, this occurred to me only after doing most of the project.

Doing so, I will examine the effect of an OECD country as an intercept shifter, a slope shifter, and as an intercept and a slope shifter in my model in the following subsections, respectively.

6.a Dummy as an Intercept Shifter

My models with an OECD dummy as an intercept shifter would be:

$$\text{cases}_{c,t} = \beta_1 + \delta_1 D_{OECD} + \beta_2 \text{EF}_{c,t} + \beta_3 log(\text{GDP}_{c,t}) + \beta_4 \text{GINI}_{c,t} + \beta_5 \text{CPI}_{c,t} + e_{c,t} \quad (1)$$

$$\text{deaths}_{c,t} = \alpha_1 + \gamma_1 D_{OECD} + \alpha_2 \text{EF}_{c,t} + \alpha_3 log(\text{GDP}_{c,t}) + \alpha_4 \text{GINI}_{c,t} + \alpha_5 \text{CPIS}_{c,t} + e_{c,t} \quad (2)$$

The results are as follows:

6.a.i For Case Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-44429.9	15572.36	-2.85313	0.005825	Multiple R	0.498894
EF	3017.064	2240.399	1.346664	0.182838	R Square	0.248895
log(GDP)	10641.53	4405.775	2.41536	0.018586	Adjusted R Square	0.190215
GINI	77.80077	147.4267	0.527725	0.599515	Standard Error	8636.766
CPI	-263.425	117.6396	-2.23925	0.028622	Observations	70
OECD	3156.953	3449.114	0.915294	0.363473		

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	1.58E+09	3.16E+08	4.241555	0.002155
Residual	64	4.77E+09	74593721		
Total	69	6.36E+09			

6.a.ii For Death Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-1096.936788	449.4499	-2.44062	0.01744	Multiple R	0.448887439
EF	23.44974342	64.66245	0.362649	0.718062	R Square	0.201499933
log(GDP)	309.230229	127.1596	2.431828	0.017832	Adjusted R Square	0.139117115
GINI	8.039723372	4.255034	1.889461	0.063362	Standard Error	249.2745358
CPI	-8.428342892	3.395318	-2.48234	0.015686	Observations	70
OECD	150.4752384	99.54841	1.511578	0.135563		

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	1003542	200708.5	3.230055	0.011583
Residual	64	3976819	62137.79		
Total	69	4980361			

6.b Dummy as a Slope Shifter

My models with an OECD dummy as a slope shifter would be:

$$\text{cases}_{c,t} = \beta_1 + (\beta_2 + \delta_2 D_{OECD})\text{EF}_{c,t} + (\beta_3 + \delta_3 D_{OECD})log(\text{GDP}_{c,t}) + (\beta_4 + \delta_4 D_{OECD})\text{GINI}_{c,t} + (\beta_5 + \delta_5 D_{OECD})\text{CPI}_{c,t} + e_{c,t}$$

$$\text{deaths}_{c,t} = \alpha_1 + (\alpha_2 + \gamma_2 D_{OECD})\text{EF}_{c,t} + (\alpha_3 + \gamma_3 D_{OECD})log(\text{GDP}_{c,t}) + (\alpha_4 + \gamma_4 D_{OECD})\text{GINI}_{c,t} + (\alpha_5 + \gamma_5 D_{OECD})\text{CPI}_{c,t} + e_{c,t}$$

Here and for the next subsection, I will be using $\beta_1, \ldots, \beta_5$ and $\alpha_1, \ldots, \alpha_5$ for the corresponding intercept, EF, and so on. I will also do so with $\delta_1, \ldots, \delta_5$ and $\gamma_1, \ldots, \gamma_5$ for the corresponding OECD dummy variable interaction since this way, I believe, it will be less confusing. The results are as follows:

6.b.i For Case Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
β_1	-44444.4	15965.51	-2.78378	0.007144	Multiple R	0.505476
β_2	3159.344	2559.415	1.234401	0.221788	R Square	0.255506
β_3	11061.9	5042.699	2.193648	0.032081	Adjusted R Square	0.157867
β_4	91.01537	167.2208	0.544283	0.588231	Standard Error	8807.579
β_5	-340.657	165.7405	-2.05536	0.044132	Observations	70
δ_2	-1167.63	5286.39	-0.22087	0.825928		
δ_3	-288.436	8848.74	-0.0326	0.974103		
δ_4	93.5137	385.1331	0.242809	0.808968		
δ_5	177.7339	273.3376	0.650236	0.517981		

	df	*SS*	*MS*	*F*	*Significance F*
Regression	8	1.62E+09	2.03E+08	2.61685	0.015683
Residual	61	4.73E+09	77573445		
Total	69	6.36E+09			

6.b.ii For Death Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
α_1	-1082.07	445.8509	-2.42698	0.018194	Multiple R	0.508958
α_2	13.70796	71.47393	0.19179	0.848544	R Square	0.259038
α_3	384.8053	140.8218	2.732568	0.00821	Adjusted R Square	0.161863
α_4	8.21703	4.669788	1.759615	0.083487	Standard Error	245.9594
α_5	-14.3239	4.628449	-3.09476	0.002972	Observations	70
γ_2	-20.5767	147.6271	-0.13938	0.889607		
γ_3	-233.773	247.1089	-0.94603	0.347864		
γ_4	11.45103	10.75518	1.064699	0.291208		
γ_5	16.09848	7.633194	2.10901	0.039057		

	df	*SS*	*MS*	*F*	*Significance F*
Regression	8	1290103	161262.9	2.665678	0.01406
Residual	61	3690258	60496.03		
Total	69	4980361			

6.c Dummy as an Intercept and Slope Shifter

My models with an OECD dummy as an intercept and slope shifter would be:

$$\begin{aligned}
\text{cases}_{c,t} &= (\beta_1 + \delta_1 D_{OECD}) + (\beta_2 + \delta_2 D_{OECD})\text{EF}_{c,t} + (\beta_3 + \delta_3 D_{OECD})log(\text{GDP}_{c,t}) \\
&\quad + (\beta_4 + \delta_4 D_{OECD})\text{GINI}_{c,t} + (\beta_5 + \delta_5 D_{OECD})\text{CPI}_{c,t} + e_{c,t} \\
\text{deaths}_{c,t} &= (\alpha_1 + \gamma_1 D_{OECD}) + (\alpha_2 + \gamma_2 D_{OECD})\text{EF}_{c,t} + (\alpha_3 + \gamma_3 D_{OECD})log(\text{GDP}_{c,t}) \\
&\quad + (\alpha_4 + \gamma_4 D_{OECD})\text{GINI}_{c,t} + (\alpha_5 + \gamma_5 D_{OECD})\text{CPI}_{c,t} + e_{c,t}
\end{aligned}$$

6.c.i For Case Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
β_1	-44465	16770.19	-2.65143	0.010237	Multiple R	0.505476
β_2	3160.525	2594.696	1.218071	0.227965	R Square	0.255506
β_3	11064.75	5125.725	2.15867	0.034886	Adjusted R Square	0.143832
β_4	91.10906	169.9594	0.536064	0.593896	Standard Error	8880.671
β_5	-340.71	167.5493	-2.03349	0.046433	Observations	70
δ_1	262.1252	59833.15	0.004381	0.996519		
δ_2	-1180.23	6056.54	-0.19487	0.846154		
δ_3	-327.884	12676.2	-0.02587	0.97945		
δ_4	92.54509	446.8591	0.207101	0.836632		
δ_5	178.377	312.2554	0.571253	0.569962		

	df	*SS*	*MS*	*F*	*Significance F*
Regression	9	1.62E+09	1.8E+08	2.287959	0.027911
Residual	60	4.73E+09	78866311		
Total	69	6.36E+09			

6.c.ii For Death Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
α_1	-1063.23	468.2421	-2.27069	0.026769	Multiple R	0.509208
α_2	12.62758	72.44675	0.174302	0.862215	R Square	0.259292
α_3	382.2066	143.1158	2.67061	0.00973	Adjusted R Square	0.148186
α_4	8.131317	4.745452	1.713497	0.091784	Standard Error	247.958
α_5	-14.2757	4.67816	-3.05156	0.003389	Observations	70
γ_1	-239.801	1670.607	-0.14354	0.886344		
γ_2	-9.05118	169.1052	-0.05352	0.957492		
γ_3	-197.685	353.9334	-0.55854	0.578557		
γ_4	12.33715	12.47679	0.988808	0.326728		
γ_5	15.51021	8.718513	1.778997	0.080306		

	df	*SS*	*MS*	*F*	*Significance F*
Regression	9	1291370	143485.6	2.333737	0.02504
Residual	60	3688991	61483.18		
Total	69	4980361			

7 Instrumental Variables: Daily Caloric Intake and Life Expectancy

For this section, I will explore instrumental variables for economic freedom. In this case, I will use daily caloric consumption from 2017 and life expectancy from 2019 obtained from Our World Data as such instrumental variables. Of course, these are not perfect instrumental variables since they will be correlated with the error term in my model, however, they will be used for demonstrative purposes only.

Thus, the matrix for my model would be:

$$Z_1 = \begin{bmatrix} 1 & EF_1 & log(GDP)_1 & GINI_1 & CPI_1 \end{bmatrix}$$
$$Z_2 = \begin{bmatrix} 1 & EF_2 & log(GDP)_2 & GINI_2 & CPI_2 \end{bmatrix}$$
$$\vdots$$
$$Z_{70} = \begin{bmatrix} 1 & EF_{70} & log(GDP)_{70} & GINI_{70} & CPI_{70} \end{bmatrix}$$

and my matrix with the replacement of economic freedom (EF) with the instrumental variables of daily caloric intake (DCI) and life expectancy (LE) would be:

$$X_1 = \begin{bmatrix} 1 & DCI_1 & LE_1 & log(GDP)_1 & GINI_1 & CPI_1 \end{bmatrix}$$
$$X_2 = \begin{bmatrix} 1 & DCI_2 & LE_2 & log(GDP)_2 & GINI_2 & CPI_2 \end{bmatrix}$$
$$\vdots$$
$$X_{70} = \begin{bmatrix} 1 & DCI_{70} & LE_{70} & log(GDP)_{70} & GINI_{70} & CPI_{70} \end{bmatrix}$$

where 1, 2,. . ., 70 correspond with each of the countries in my data set. Since I have over-identified the number of instrumental variables, I must calculate b_{2SLS} rather than b_{IV}.

7.a For Case Rates

$$b_{2SLS}^{cases} = \begin{bmatrix} \text{-49216.6913924579} \\ \text{3347.51293308166} \\ \text{11171.6093660121} \\ \text{63.0133882562653} \\ \text{-223.707753562140} \end{bmatrix}$$

For the Hausman test, I obtain $m = -0.0042$ with an $OIR = 3.7871$.

H_0 : Daily caloric intake and life expectancy are valid

H_1 : At least one is not valid

Since $OIR < \chi^2_c$, we fail to reject the null hypothesis that the instruments are invalid at $\alpha = 0.05$.

7.b For Death Rates

$$b^{deaths}_{2SLS} = \begin{bmatrix} \text{-2106.11483110898} \\ \text{297.686509127318} \\ \text{104.969633002205} \\ \text{4.05097072458978} \\ \text{-7.49257712557724} \end{bmatrix}$$

For the Hausman test, I obtain $m = -0.8836$ with an $OIR = 4.0146$.

H_0 : Daily caloric intake and life expectancy are valid

H_1 : At least one is not valid

Since $OIR < \chi^2_c$, we fail to reject the null hypothesis that the instruments are invalid at $\alpha = 0.05$.

8 Estimation Issues

There are plenty of estimation issues in this project, some of which stems from the collection and availability of data and other issues from my own limited skills in statistical and econometric techniques. Specifically, because I can only do regressions with complete data sets, I had to remove countries that did not have a GINI coefficient from 2015, who did not report their GDP per capita in 2019, who did not have a CPI, and who did not collect data on COVID-19 case and death rates. This meant I could only run regressions on 70 out of the 195 countries in the world. Moreover, it is wealthier countries who are more able to collect and more willing to share such data consistently. This point is exemplified by the fact that 27 out of the 37 OECD countries are included in my data sample, which is a large over-representation of countries who are, for the most part, democratic and consisting of large and stable economies. Wealthier countries also have more of the population living in dense urban areas, which would allow COVD-19 to spread more easily.

As a result, my models are subject to heavy omission bias, since I could not include most countries from Africa, the Middle East, and Asia. This meant there is an over-representation of Europe in my model. Related to this, some governments are so corrupt that the CPI cannot be estimated, such as North Korea and Russia, and this, again, further biases my data towards democratic and liberal countries.

Additionally, there could be multicollinearity issues with my models. Economic freedom is highly correlated with GDP and even potentially with perceived corruption since the criterion for economic freedom relies on the size of government, legal system and property rights, sound money, free trade, and regulation, all of which are, of course, tied to the role of the government. Yet another problem my model may have is heteroskedasticity. The error terms between each country may vary and therefore, every continent and even every country could have differing variances. Because of this problem, I would need to estimate the errors of each country. In the beginning, I was considering of conducting a time-series analysis from March until November, which would be an attempt to see if whether or not economically freer countries

Lastly, the final problem my models have is the level of aggregation. Although the USA was not included in my data set, but using it as an example, between states, there are differences in economic freedoms, age and other demographics, regulations, and pandemic mandates which would affect COVID-19 cases and deaths, for instance, between Texas and New York. Switzerland, which is in my data set, may have cantons that have dealt with and are dealing with COVID-19 differently and by aggregating data at a country level, the effects of different governance intrastate are lost.

9 Results

My economic models would be

$$\widehat{\text{cases}} = -47147.385 + 2662.654\widehat{\text{EF}} + 11779.74 log(\widehat{\text{GDP}}) + 71.714\widehat{\text{GINI}} - 221.171\widehat{\text{CPI}} \quad (1)$$

$$\widehat{\text{deaths}} = -1226.464 + 6.557\widehat{\text{EF}} + 363.483 log(\widehat{\text{GDP}}) + 7.75\widehat{\text{GINI}} - 6.414\widehat{\text{CPI}} \quad (2)$$

As we can see, the results are almost the opposite of what I hypothesized. For each additional score of economic freedom, cases increases by 2662 per million and deaths increase by 6.55 per million, ceteris paribus. While for every addition percentage increase in GDP per capita increases cases by 11,779 per million and deaths by 363 per million, ceteris paribus. For every additional point gained in the GINI coefficient, this increases cases by 71 per million and deaths by 7.75 per million, ceteris paribus. Lastly, for every additional point in the CPI, cases decrease by 221 per million and deaths decrease 6.4 per million, ceteris paribus. Only the GINI coefficient came out from what I expected: that higher income inequality leads to greater COVID-19 cases and deaths.

The inclusion of an OECD dummy variable, surprisingly, was not statistically significant as an intercept shifter, a slope shifter, or an intercept and a slope shifter. My OECD dummy models as an intercept shifter are:

$$\hat{cases} = -44429.9 + 3156.953\hat{D}_{OECD} + 3017.064\hat{EF} + 10641.53log(\hat{GDP}) + 77.801\hat{GINI} - 263.425\hat{CPI} \quad (1)$$

$$\hat{deaths} = -1096.937 + 150.475\hat{D}_{OECD} + 23.45\hat{EF} + 309.23LOG(\hat{GDP}) + 8.04\hat{GINI} - 8.428\hat{CPI} \quad (2)$$

This means being an OECD country increases case rates by 3157 per million, ceteris paribus, and death rates by 150.475 per million, ceteris paribus. This, again, however, is not significant at $\alpha = 0.05$.

10 Conclusion

Overall I find that economic freedom, GDP per capita, and increased income inequality (the GINI coefficient) contributes to the case and death rates of COVID-19. This may, however, be due to the choice of variables, the multicollinearity of these variables, my limitations in econometrics techniques, and the other correlated factors that wealthier, developed countries have. With wealthier countries more likely to travel and to live in more densely populated metropolitan areas, this creates an environment for contagious diseases to more easily spread. Surprisingly, we see countries with high corruption perception indices actually decrease COVID-19 case and death rates. However, it should be noted that the most corrupt countries do not have a CPI due to the difficulty of even ranking and placing these countries.

The R^2 and $\bar{R}^2$ are both low, however at around 0.2, which means the variation from my model explains little of the variation in COVID-19 cases and deaths per million. I would need to incorporate more relevant variables, try nonlinear regression models, obtain additional and more complete data, and attempt other econometric techniques and tests. The F-tests in both case rates and death rates were significant at a 95% significant level, however, it seems my variables are not very robust, especially economic freedom, which is not significant at $\alpha = 0.05$ in both COVID-19 cases and deaths. In short, I would need to include many more variables in order to see if economic freedom does indeed affect COVID-19 case and death rates. From this term project, it remains inconclusive whether or not greater economic freedom harms or helps in terms of cases and deaths during a pandemic.

11 Addendum

Unfortunately, for whatever reason, I did not think to use the logarithm of COVID-19 cases and deaths for the following models:

$$log(\text{cases}_{c,t}) = \beta_1 + \beta_2\text{EF}_{c,t} + \beta_3 log(\text{GDP}_{c,t}) + \beta_4\text{GINI}_{c,t} + \beta_5\text{CPI}_{c,t} + e_{c,t} \quad (1)$$

$$log(\text{deaths}_{c,t}) = \alpha_1 + \alpha_2\text{EF}_{c,t} + \alpha_3 log(\text{GDP}_{c,t}) + \alpha_4\text{GINI}_{c,t} + \alpha_5\text{CPI}_{c,t} + e_{c,t} \quad (2)$$

Doing so, I obtain

11.a For Case Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-0.500620326	0.919505	-0.54445	0.587998	Multiple R	0.585151246
Overall Score	0.301441858	0.132733	2.271033	0.026465	R Square	0.342401981
loggpd2019	0.706447373	0.254235	2.778722	0.007127	Adjusted R Square	0.30193441
gini2015	-0.00076452	0.008859	-0.0863	0.931493	Standard Error	0.519506533
cpi2018	-0.014294284	0.006508	-2.19625	0.031648	Observations	70

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	9.134213	2.283553	8.461145	1.47E-05
Residual	65	17.54266	0.269887		
Total	69	26.67687			

11.b For Death Rates

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Regression Statistics*	
Intercept	-2.60671841	1.070227	-2.43567	0.017615	Multiple R	0.529061294
Overall Score	0.239642525	0.154491	1.551179	0.125715	R Square	0.279905853
loggpd2019	0.928230914	0.295908	3.136891	0.002565	Adjusted R Square	0.235592367
gini2015	0.004796048	0.010311	0.465144	0.643383	Standard Error	0.604662365
cpi2018	-0.020046977	0.007575	-2.64634	0.010195	Observations	70

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	9.237659	2.309415	6.316494	0.000234
Residual	65	23.76508	0.365617		
Total	69	33.00274			

11.c New AIB and BIC

	AIC	BIC	R^2	$\bar{R}^2$
Cases	18.23779937	18.36628481	0.181801797	0.131451138
Cases log(GDP)	18.16524556	18.293731	0.239062822	0.192235919
Log(Cases) log(GDP)	-1.269574042	-1.1410886	0.342401981	0.30193441
Deaths	11.2064823	11.33496774	0.077180081	0.020391163
Deaths log(GDP)	11.09686181	11.22534726	0.172992614	0.122099852
Log(Deaths) log(GDP)	-0.965992362	-1.080278076	0.279905853	0.235592367

As we can see, when I use the logarithm of both case and death rates along with the logarithm of GDP, the AIC and BIC are lower while R^2 and $\bar{R}^2$ are higher. Therefore, the model I should have used in this paper is:

$$log(\text{cases}_{c,t}) = \beta_1 + \beta_2\text{EF}_{c,t} + \beta_3 log(\text{GDP}_{c,t}) + \beta_4\text{GINI}_{c,t} + \beta_5\text{CPI}_{c,t} + e_{c,t} \tag{1}$$

$$log(\text{deaths}_{c,t}) = \alpha_1 + \alpha_2\text{EF}_{c,t} + \alpha_3 log(\text{GDP}_{c,t}) + \alpha_4\text{GINI}_{c,t} + \alpha_5\text{CPI}_{c,t} + e_{c,t} \tag{2}$$

and the dummy variable and instrument variable techniques I used should have been with these models rather than just the logarithm of GDP alone. Unfortunately, this only occurred to me after I had done most of the work with the less explanatory model.

A Appendix

country	*code*	*gdp2019*	*EF*	*loggdp2019*	*gini2015*	cpi2018	OECD	dummyef	dummygdp	dummygini	dummycpi	calday	life2019	Cases Per Million	Deaths Per Million
Albania	ALB	5352.857	7.8	3.728586	32.9	36	0	0	0	0	0	3400	78.573	7253.805	176.871
Armenia	ARM	4622.733	7.92	3.664899	32.4	35	0	0	0	0	0	3078	75.087	31132.877	459.97
Austria	AUT	50277.28	7.8	4.701372	30.5	76	1	7.8	4.701372	30.5	76	3692	81.544	11834.251	121.802
Belgium	BEL	46116.7	7.56	4.663858	27.7	75	1	7.56	4.663858	27.7	75	3768	81.628	38708.444	1022.467
Benin	BEN	1219.433	6.2	3.086158	47.8	40	0	0	0	0	0	2756	61.771	218.012	3.382
Bulgaria	BGR	9737.601	7.74	3.988452	38.6	42	0	0	0	0	0	2828	75.046	7605.155	184.07
Belarus	BLR	6663.295	6.35	3.823689	25.6	44	0	0	0	0	0	3289	74.791	10422.125	103.711
Bolivia	BOL	3552.069	6.3	3.550481	46.7	29	0	0	0	0	0	2354	71.513	12143.977	747.45
Brazil	BRA	8717.186	6.56	3.940376	51.9	35	0	0	0	0	0	3248	75.881	26042.625	752.185
Botswana	BWA	7961.338	7.6	3.900986	53.3	61	0	0	0	0	0	2340	69.592	2824.43	10.206
Switzerland	CHE	81993.73	8.43	4.913781	32.3	85	1	8.43	4.913781	32.3	85	3413	83.779	17762.54	235.135
Chile	CHL	14896.45	7.96	4.173083	44.4	67	1	7.96	4.173083	44.4	67	3011	80.181	26692.322	743.191
China	CHN	10261.68	6.21	4.011218	38.6	39	0	0	0	0	0	3194	76.912	63.449	3.293
Cote d'Ivoire	CIV	2286.163	6.09	3.359107	41.5	35	0	0	0	0	0	2730	57.783	785.343	4.777
Colombia	COL	6432.388	6.71	3.808372	51.1	36	1	6.71	3.808372	51.1	36	3094	77.287	21110.91	615.413
Costa Rica	CRI	12238.37	7.62	4.087724	48.4	56	0	0	0	0	0	3015	80.279	21587.856	271.882
Cyprus	CYP	27858.37	7.84	4.444956	34	59	0	0	0	0	0	2616	80.982	4984.593	29.684
Czech Republic	CZE	23101.78	7.81	4.363645	25.9	59	1	7.81	4.363645	25.9	59	3276	79.376	31291.676	303.577
Germany	DEU	46258.88	7.85	4.665195	31.7	80	1	7.85	4.665195	31.7	80	3556	81.326	6360.765	125.096
Denmark	DNK	59822.09	8.1	4.776862	28.2	88	1	8.1	4.776862	28.2	88	3384	80.898	8002.309	124.478
Dominican Republic	DOM	8282.116	7.58	3.918141	45.2	30	0	0	0	0	0	2864	74.081	11708.99	206.952
Ecuador	ECU	6183.824	6.46	3.791257	46	34	0	0	0	0	0	2586	77.01	9533.04	718.129
Spain	ESP	29613.67	7.73	4.471492	36.2	58	1	7.73	4.471492	36.2	58	3270	83.565	25359.502	767.365
Estonia	EST	23659.87	7.96	4.374012	32.7	73	1	7.96	4.374012	32.7	73	3245	78.745	3697.592	55.03
Ethiopia	ETH	857.5014	5.61	2.933235	35	34	0	0	0	0	0	2304	66.597	836.517	12.778
Finland	FIN	48685.85	7.76	4.687403	27.1	85	1	7.76	4.687403	27.1	85	3337	81.908	2908.107	64.613
France	FRA	40493.93	7.4	4.60739	32.7	72	1	7.4	4.60739	32.7	72	3558	82.659	20906.26	563.598
United Kingdom	GBR	42300.27	8.08	4.626343	33.2	80	1	8.08	4.626343	33.2	80	3428	81.321	14902.335	685.782
Georgia	GEO	4769.187	8.18	3.678444	36.5	58	0	0	0	0	0	2872	73.767	10209.379	83.977
Greece	GRC	19582.54	6.71	4.291869	36	45	1	6.71	4.291869	36	45	3355	82.239	3765.786	60.059
Honduras	HND	2574.912	7.28	3.410762	49.6	29	0	0	0	0	0	2552	75.27	9834.211	269.773
Croatia	HRV	14853.24	7.36	4.171821	31.1	48	0	0	0	0	0	3083	78.488	12012.858	133
Hungary	HUN	16475.74	7.44	4.216845	30.4	46	0	0	0	0	0	3286	76.875	8198.357	188.295
Indonesia	IDN	4135.569	7.39	3.616535	39.7	38	0	0	0	0	0	2892	71.716	1499.278	50.705
Ireland	IRL	78660.96	8.13	4.895759	31.8	73	1	8.13	4.895759	31.8	73	3717	82.305	12446.039	387.42
Iceland	ISL	66944.83	7.71	4.825717	26.8	76	1	7.71	4.825717	26.8	76	3628	82.993	14256.41	35.165
Italy	ITA	33189.57	7.51	4.521002	35.4	52	1	7.51	4.521002	35.4	52	3522	83.512	11237.338	638.717
Kazakhstan	KAZ	9731.145	7.12	3.988164	26.8	31	0	0	0	0	0	3196	73.597	8008.859	118.604
Kenya	KEN	1816.547	6.84	3.259247	40.8	27	0	0	0	0	0	2124	66.699	1026.421	18.523
Lithuania	LTU	19455.45	8.1	4.289041	37.4	59	1	8.1	4.289041	37.4	59	3419	75.933	5445.413	60.611
Luxembourg	LUX	114704.6	7.75	5.059581	33.8	81	1	7.75	5.059581	33.8	81	3335	82.25	28648.063	250.808
Latvia	LVA	17836.36	7.89	4.251306	34.2	58	1	7.89	4.251306	34.2	58	3169	75.292	3124.798	37.642
Moldova	MDA	4498.521	6.99	3.65307	27	33	0	0	0	0	0	2395	71.901	18849.95	442.493
Malta	MLT	29416.23	7.94	4.468587	29.4	54	0	0	0	0	0	3460	82.53	13683.955	140.418
Myanmar	MMR	1407.813	5.81	3.148545	38.1	29	0	0	0	0	0	2701	67.134	968.686	22.735
Montenegro	MNE	8832.037	6.94	3.946061	39	45	0	0	0	0	0	3481	76.884	29202.531	479.252
Malaysia	MYS	11414.84	7.58	4.05747	41	47	0	0	0	0	0	2909	76.156	974.727	7.693
Namibia	NAM	4957.458	6.52	3.695259	59.1	53	0	0	0	0	0	2431	63.708	5090.684	52.343
Netherlands	NLD	52447.83	7.82	4.719728	28.2	82	1	7.82	4.719728	28.2	82	3249	82.283	20470.767	430.992
Norway	NOR	75419.63	7.6	4.877484	27.5	84	1	7.6	4.877484	27.5	84	3385	82.404	3608.583	52.018
Pakistan	PAK	1284.702	6.07	3.108802	33.5	33	0	0	0	0	0	2326	67.273	1511.913	30.888
Panama	PAN	15731.02	7.8	4.196757	50.8	37	0	0	0	0	0	2839	78.506	30962.963	625.758
Peru	PER	6977.696	7.76	3.843712	43.4	35	0	0	0	0	0	2769	76.736	27371.928	1045.619
Philippines	PHL	3485.084	7.43	3.542213	44.4	36	0	0	0	0	0	2674	71.231	3474.404	65.896
Poland	POL	15595.23	7.04	4.192992	31.8	60	1	7.04	4.192992	31.8	60	3525	78.73	9584.241	148.785
Portugal	PRT	23145.04	7.6	4.364458	35.5	64	1	7.6	4.364458	35.5	64	3492	82.049	13855.355	245.864
Paraguay	PRY	5414.799	7.23	3.733582	47.6	29	0	0	0	0	0	2740	74.254	8858.708	196.845
Romania	ROU	12919.53	7.83	4.111247	35.9	47	0	0	0	0	0	3525	76.054	12545.118	362.206
El Salvador	SLV	4187.25	7.41	3.621929	40.6	35	0	0	0	0	0	2660	73.317	5156.331	150.936
Serbia	SRB	7402.355	7.05	3.86937	40.5	39	0	0	0	0	0	2805	76.001	6900.336	120.507
Slovenia	SVN	25739.25	7.33	4.410596	25.4	60	1	7.33	4.410596	25.4	60	3195	81.324	16502.223	111.115
Sweden	SWE	51610.07	7.58	4.712734	29.2	85	1	7.58	4.712734	29.2	85	3219	82.797	13073.123	594.894
Togo	TGO	675.5422	6.25	2.829652	43.1	30	0	0	0	0	0	2429	61.042	281.565	6.885
Thailand	THA	7808.193	6.75	3.892551	36	36	0	0	0	0	0	2827	77.15	54.212	0.845
Tajikistan	TJK	870.7876	6.05	2.939912	34	25	0	0	0	0	0	2104	71.097	1155.107	8.598
Tunisia	TUN	3317.541	6.07	3.520816	32.8	43	0	0	0	0	0	3467	76.699	5171.078	114.057
Turkey	TUR	9042.493	6.62	3.956288	42.9	41	1	6.62	3.956288	42.9	41	3540	77.691	4450.69	121.557
Ukraine	UKR	3659.031	6.06	3.563366	25.5	32	0	0	0	0	0	3035	72.065	8859.998	164.541
Uruguay	URY	16190.13	7.25	4.20925	40.1	70	0	0	0	0	0	3153	77.911	899.322	16.697
Zambia	ZMB	1291.343	7.19	3.111042	57.1	35	0	0	0	0	0	2013	63.886	893.823	18.984

Table 1: Master Data Set