Indian Institute Of Technology ,Kharagpur

Econometric Analysis Report

The Effect Of GDP On A Country's CO_2 Emission.

Team members:

1. Ujwal Anil Chikane 20HS20053
2. Aniket Patel 20HS20006
3. Harshit Sharma 20HS20024
4. Himanshu Kumar 20HS20026
5. Aayush Jitendra Kumar 20HS20002
6. Yuvraj Tilara 20HS20057

ABSTRACT

The environment has sparked a lot of interest recently. Countries are now collaborating to solve environmental issues, such as the Kyoto Protocol and the Paris Agreement. Carbon dioxide emissions are one of the most significant contributors to worldwide environmental problems. This study examines the relationship between a country's development and the increase in CO2 emissions, hypothesizing a positive relationship.Single and multiple regression models show that there are some substantial positive connections between CO2 emissions and increases in GDP per capita, implying that when countries seek growth, they produce more CO2.

INTRODUCTION

Sustainable development has now become one of the most essential phrases in the global economy. The world has changed dramatically since the industrial revolution as a result of the use of fossil fuels. However, the trend has shifted since many of the difficulties associated with using fossil fuels, which emit CO2, have resulted in major issues that threaten human survival, such as climate change. Many countries began to recognize the dangers of CO2 emissions to the environment and began to limit them.Countries have agreed to work together under the United Nations Framework Convention on Climate Change (UNFCCC) to prevent long-term dangers, as indicated by the Kyoto Protocol and the Paris Agreements. In reality, many of the world's most developed countries have begun to impose legal restrictions on CO2 emissions.

Unlike industrialized countries, which can reduce CO2 emissions through education or technology, most poor countries believe that they must continue to release massive amounts of CO2 because alternative energy sources are too expensive to employ. Many of these countries feel that restricting their use of fossil fuels is unfair after industrialized countries have used their reserves and achieved rapid economic expansion.They further contend that rich countries, not developing countries, are responsible for a large percentage of overall CO_2 emissions, hence prohibiting CO_2 emissions from developing countries is unfair.

Energy is a critical aspect in a country's development, and despite the fact that numerous renewable energies are already in use, fossil fuels still account for a large portion of total energy output. This means that the more the country's ability to create energy, the greater the country's economic growth. Because fossil fuels account for a large share of most energy sources utilized on Earth, it is easy to infer that wealthier countries emit more CO_2 than impoverished countries.The purpose of this article is to determine the true relationship between a country's GDP and CO_2 emissions. If the argument made by most of those developing countries is valid, it is theorized that the higher GDP country will release more CO_2 since it has more capability to do so.

IMPACT AND MEASUREMENT OF VARIABLES

To construct a better regression model with multiple regression, this paper used renewable energy consumption, access to electricity, percentage of manufacturing factor in GDP, and urban population as explanatory variables. The consumption of renewable energy was chosen for the simple reason that if more renewable energy is consumed, there is a greater chance that a country will use less energy that releases CO_2. As previously said, low-income countries may have difficulty getting renewable energy sources, which can lead to a variety of outcomes. In terms of access to power, it is possible that countries with limited access to electricity will experience a lack of energy, resulting in low CO_2 emissions. With regard to the proportion of manufacturing factor, it is reasonable to assume that a country with a high percentage of manufacturing factor will have a greater number of factories, resulting in an increase in CO_2 emissions.The number of urban residents was chosen since cities typically consume more energy than rural areas. When a result, it is believed that as the population of a city grows, more energy would be consumed, resulting in increased CO_2 emissions.A dummy variable is added to the developed and developing countries, depending on whether the country is an OECD member or not. The five important partner countries are also included in the OECD dummy variable.

```
summ logco2
```

Variable	Obs	Mean	Std. Dev.	Min	Max
logco2	148	10.07386	2.064254	5.736572	16.14896

```
summ log_gdppc
```

Variable	Obs	Mean	Std. Dev.	Min	Max
log_gdppc	148	8.769104	1.412213	6.165321	11.66648

```
summ renewable
```

Variable	Obs	Mean	Std. Dev.	Min	Max
renewable	148	32.75393	27.74344	0	96.3837

```
summ accelec
```

Variable	Obs	Mean	Std. Dev.	Min	Max
accelec	148	85.0282	25.32235	0	100

```
summ manuf
```

Variable	Obs	Mean	Std. Dev.	Min	Max
manuf	148	12.81661	6.472073	1.687851	39.91328

```
. summ log_urbanpop
```

Variable	Obs	Mean	Std. Dev.	Min	Max
log_urbanpop	148	15.65014	1.676542	12.07343	20.52947

```
summ oecd
```

Variable	Obs	Mean	Std. Dev.	Min	Max
oecd	148	.2702703	.4456074	0	1

For the simple regression model between **logco$_2$** and **log_gdppc**, and for the multiple regression model with more variables the **Classical Linear Model (CLM)** assumption should be verified.

The assumptions are as below:

Assumption 1: Linear in parameters

The model will follow the assumption that is linear in parameters, as below

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$

Assumption 2: Random Sampling

Since the data were obtained from the World Bank, they were all gathered in a random population and samples from throughout the world, therefore achieving the assumption.

Assumption 3: No perfect collinearity

It is found that there is no perfect collinearity that the value equals to 1. However, there were some high values that were approaching near to 1, therefore more analysis in the robustness testing part would be required.

```
. corr log_gdppc renewable accelec manuf log_urbanpop
(obs=148)
```

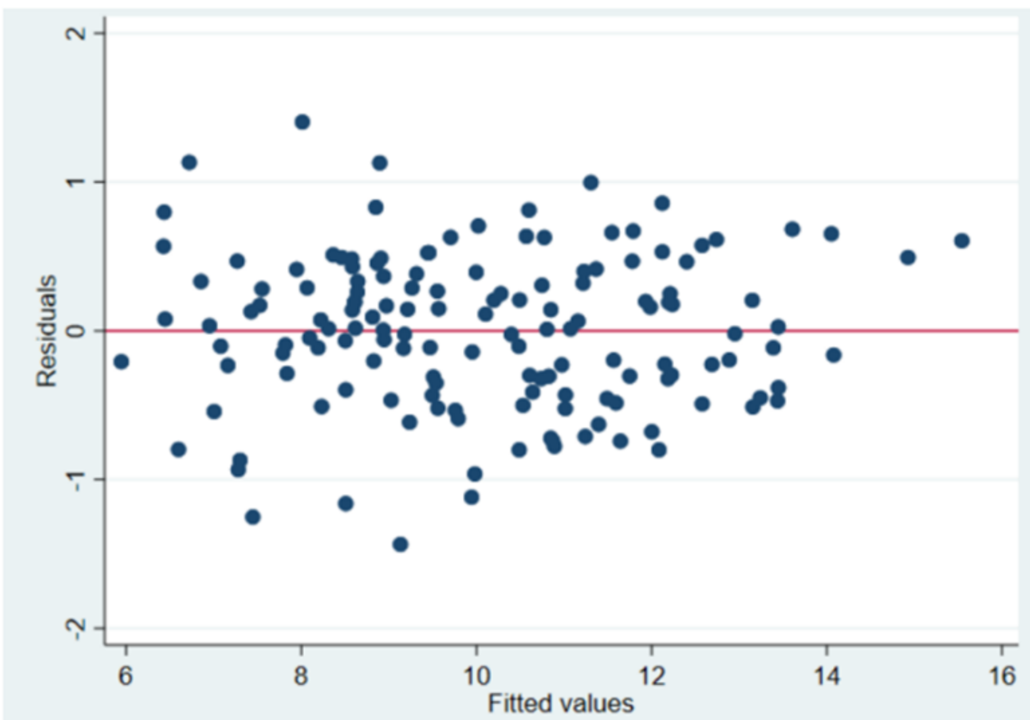
	log_gd~c	renewa~e	accelec	manuf	log_ur~p
log_gdppc	1.0000				
renewable	-0.5976	1.0000			
accelec	0.7011	-0.7360	1.0000		
manuf	0.1499	-0.1423	0.2242	1.0000	
log_urbanpop	0.0087	-0.1245	0.0720	0.2634	1.0000

Assumption 4: Zero conditional mean

True, there will be numerous variable factors that will influence the variables. However, for the multiple linear regressions employed in this study, it is assumed that the residuals have a zero conditional mean, which means that $E[u|x_i] = 0$ for all i = 1, 2,... n. The omitted variable bias i may be used to determine all slopes from the variables. If the bias value is positive, there will be an overestimation, if the bias value is negative, there will be an underestimation.

Assumption 5: Homoskedasticity

Similar to assumption 4, it is assumed for multiple linear regression that the expected variance of residual u is constant for each given dependent variable, implying that $V(u|x_i) = \sigma^2$ for all i = 1,2,... n.



The spread of residual is shown to be equidistantly separated from the zero line within the value of 1. Those points with a value greater than 1 must be carefully studied.

Each regression models will provide the equation and the standard error for each parameter inside the parentheses. Also, the n stands for number of observations and R^2 as the sum of squared residuals.

First, this is a simple regression model that identifies the direct association between CO_2 emissions and GDP per capita in the absence of any other factors. This will show the effect of GDP growth on the rise of CO_2 emissions.

Simple Regression Model 1: $logco_2 = \beta_0 + \beta_1 log\_gdppc + u$

Regressing $logco_2$ on log_gdppc, the equation results as below.

Equation 1: $logco_2 = 4.38 + 0.65\ (log\_gdppc)$

$$\qquad (0.96) \qquad\quad (0.11)$$

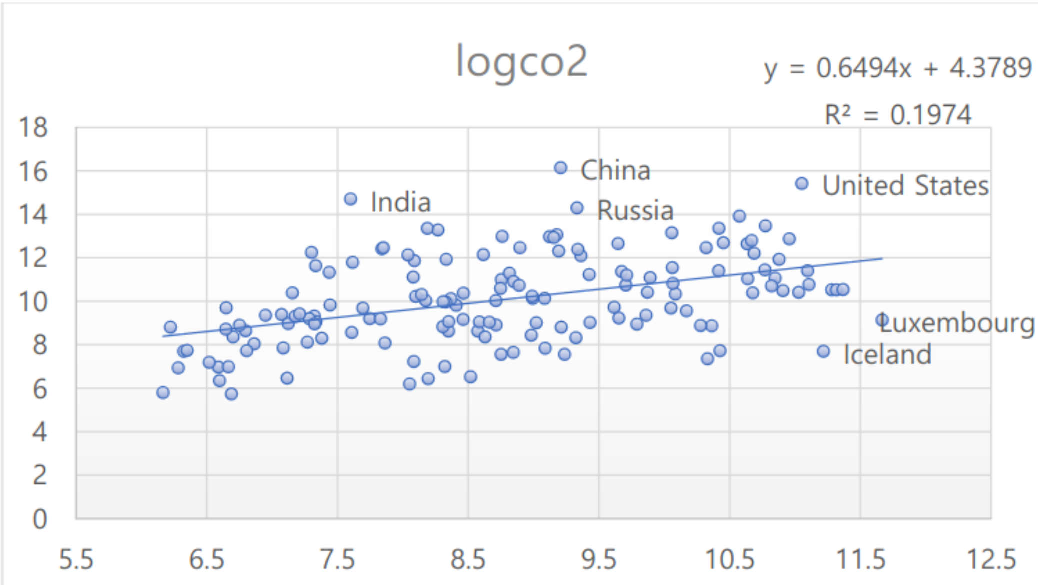
$$n = 148 \qquad R^2 = 0.20$$

```
. regress logco2 log_gdppc

      Source |       SS           df       MS      Number of obs   =       148
-------------+----------------------------------   F(1, 146)       =     35.91
       Model |  123.648035         1  123.648035   Prob > F        =    0.0000
    Residual |  502.740166       146   3.4434258   R-squared       =    0.1974
-------------+----------------------------------   Adj R-squared   =    0.1919
       Total |  626.388201       147  4.26114422   Root MSE        =    1.8556

------------------------------------------------------------------------------
      logco2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .6494337   .1083769     5.99   0.000     .4352436    .8636239
       _cons |   4.378911   .9625309     4.55   0.000     2.476617    6.281205
------------------------------------------------------------------------------
```

This equation illustrates the relationship between CO_2 emission and GDP per capita. There is a positive association between the two, with a 1% rise in GDP per capita resulting in a 0.65% increase in CO_2 emissions. This shows that when a country's economy increases, it emits more CO2 to support its growth. The R-squared value is 0.20, indicating that the regression accounts for 20% of the variance in CO_2 emissions. Furthermore, log_gdppc has a t-value of 5.99 and a p-value of 0.00. This suggests that this regression model is statistically significant at a level of less than 1%, which is highly promising.



Multiple Regression Model 2: $logco_2 = \beta_0 + \beta_1 log\_gdppc + \beta_2 renewable + \beta_3 accelec + \beta_4 manuf + \beta_5 log\_urbanpop + u$

The first multiple regression equation is as below

Equation 2: $logco_2$ = -8.62 + 0.32 (log_gdppc) – 0.014 (renewable) + 0.012 (accelec) +

(0.60) (0.044) (0.0024) (0.0030)

0.0085 (manuf) + 0.97 (log_urbanpop)

(0.0072) (0.027)

n = 148 $R^2 = 0.94$

```
. regress logco2 log_gdppc renewable accelec manuf log_urbanpop

      Source |       SS           df       MS      Number of obs   =       148
-------------+----------------------------------   F(5, 142)       =    418.49
       Model |  586.581398         5  117.31628   Prob > F        =    0.0000
    Residual |  39.8068025       142  .280329595   R-squared       =    0.9365
-------------+----------------------------------   Adj R-squared   =    0.9342
       Total |  626.388201       147  4.26114422   Root MSE        =    .52946

------------------------------------------------------------------------------
      logco2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   log_gdppc |   .3169329   .0442575     7.16   0.000     .2294442    .4044215
   renewable |  -.0138719   .0023777    -5.83   0.000    -.0185722   -.0091715
     accelec |   .0123838   .0030017     4.13   0.000     .0064501    .0183175
       manuf |   .0085263   .0071881     1.19   0.238    -.0056832    .0227359
log_urbanpop |   .9713176    .027276    35.61   0.000      .917398    1.025237
       _cons |  -8.616073   .6027351   -14.29   0.000    -9.807566   -7.424579
------------------------------------------------------------------------------
```

This equation demonstrates the link between a country's CO_2 emissions, GDP per capita, renewable energy consumption, access to electricity, manufacturing factor proportion, and urban population.

The R-squared value improved to 0.94, indicating that all explanatory variables of log_gdppc, renewable, accelec, manuf, and log_urbanpop can explain 94 % of the dependent variable.

This considerable rise in R-squared value can be attributed to the additional explanatory factors. Unlike the link between GDP per capita, which is 0.32 (meaning that a 1% rise leads to a 0.32 % increase in CO_2 emissions), renewable energy consumption has a negative association with CO_2 emissions, with a 1% increase resulting to a 0.014 % drop in CO_2.

This data is intriguing since it appears that the rise in renewable energy has less of an influence on the reduction of CO_2 emissions. This section may be researched further to determine the true impact of renewable energy consumption on CO_2 emissions. The t-values and p-values of each variable may be used to determine the statistical significance of this model. With the exception of manuf, which had a p-value of 0.238, indicating that this is significant at the level of 23.8 % (showing not quite insignificant), all other variables had 0.00 p-values, indicating that everything else is generating promising statistical conclusions.

Independent variables	**Model 1**	**Model 2**
log_gdppc	0.65*** (0.11)	0.32*** (0.044)
renewable		-0.14*** (0.0024)
accelec		0.12*** (0.0030)
manuf		0.0085 (0.0072)
log_urbanpop		0.97*** (0.027)
Intercept	4.38*** (0.96)	-8.62*** (0.60)
Number of observations	148	148
R-squared	0.20	0.94
Adjusted R-squared	0.19	0.93

Checking for Heteroskedasticity

First we selected our dependent and independent variables.

Dependent var.—CO2 emmission

Independent var--GDP per capita, Renewable energy consumption, Access to electricity, Manufacturing factor, urban population

We then performed brush pagon test to find out if there is any heteroskedasticity present or not.

Brush pagon test:

We regressed our model and then squared the residuals.
Then we regressed the squared residual and independent variable and analysed significance f.
We found out that f <<0.1 hence heteroskedasticity was detected.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.967152
R Square	0.935384
Adjusted F	0.933092
Standard E	0.231806
Observatic	147

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	5	109.6777	21.93554	408.2231	5.58E-82
Residual	141	7.576522	0.053734		
Total	146	117.2542			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Ipper 95.0%*
Intercept	-5.256133	0.278876	-18.8476	2.36E-40	-5.807451	-4.704815	-5.807451	-4.704815
X Variable	0.343592	0.042676	8.051236	3.09E-13	0.259225	0.427959	0.259225	0.427959
X Variable	-0.168182	0.02779	-6.051859	1.23E-08	-0.223121	-0.113243	-0.223121	-0.113243
X Variable	0.939432	0.130832	7.180474	3.66E-11	0.680787	1.198077	0.680787	1.198077
X Variable	0.065895	0.0837	0.787279	0.43244	-0.099574	0.231365	-0.099574	0.231365
X Variable	0.983097	0.028372	34.65077	7.43E-71	0.927009	1.039186	0.927009	1.039186

RESIDUAL OUTPUT

Sheet2 Sheet3 Sheet5 econometrics prefinal data (1) Sheet7 Sheet8 Sheet6

Removing heteroskedasticity

We took log of dependent and independent variables. And after that we again followed the same procedure to detect heteroskedasticity in the rectified model and found significance f to be greater than 0.1. Hence heteroskedasticity was rectified from the model.

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.200819
R Square	0.040328
Adjusted F	0.006297
Standard E	0.073883
Observatic	147

ANOVA

	df	SS	MS	F	ignificance F
Regressior	5	0.032344	0.006469	1.18505	0.319536
Residual	141	0.769685	0.005459		
Total	146	0.802029			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	lpper 95.0%
Intercept	0.194307	0.088886	2.186035	0.030463	0.018586	0.370028	0.018586	0.370028
X Variable	-0.007484	0.013602	-0.550232	0.583031	-0.034374	0.019406	-0.034374	0.019406
X Variable	3.49E-05	0.008858	0.003939	0.996862	-0.017476	0.017546	-0.017476	0.017546
X Variable	-0.055007	0.0417	-1.319131	0.189263	-0.137445	0.02743	-0.137445	0.02743
X Variable	-0.006114	0.026678	-0.229192	0.819052	-0.058854	0.046626	-0.058854	0.046626
X Variable	-0.000511	0.009043	-0.056543	0.954989	-0.018388	0.017366	-0.018388	0.017366

RESIDUAL OUTPUT

Sheet2 | Sheet3 | Sheet5 | econometrics prefinal data (1) | Sheet7 | Sheet8 | Sheet6

CONCLUSION

The key log gdp variable remained positive coefficients throughout the results from the various regression models, confirming the basic premise that there will be a positive correlation between the country's CO2 emission and increase in GDP per capita. Even though it appears that several concerns from the robustness test may jeopardize the CLM assumptions, which are crucial, the model did explain a high level of explanation, with R-squared values approaching 0.94.

Three things can be presented as further implications from this work. This research merely explains the existence of a link between CO2 emissions and economic progress in a country. Based on the findings of this work, more research into the causes of nations that are positioned higher than the regression line, as well as the causes of countries with larger CO2 emissions than the average fit line, can be undertaken.

Furthermore, as stated in this paper, there is already a great deal of study being done and in progress on the countries that emit large amounts of CO2, as demonstrated by the four countries discussed in this paper: the United States, China, Russia, and India. However, as illustrated in Appendix H, there are certain countries, such as Luxembourg and Iceland, that have a high GDP per capita while emitting very little CO2. Benchmarking study for those countries can be useful in determining how a country might maintain low CO2 emissions while growing economically.

Finally, because the cross-country research in this paper was limited to data from 2018, there must be a constraint in determining how trends develop over time. As a result, time series analysis among the countries can be added, resulting in a more difficult expansion of the panel data analysis of these countries. Panel data analysis will provide a better knowledge of the actual trend of CO2 emissions and economic development in countries.