Project I

Author: Mohammad Zohaib Saeed

Title: Environmental Impacts of Trade Liberalization in Developing Countries using Logistic Regression

1. Overview

In today's modern world, globalization has enabled countries to adopt and exchange various goods and services beyond their borders with ease. One such example is trade, which is more open nowadays than ever because of removal of trade barriers like tariffs, quotas etc. Due to this, countries are now increasing their trade with a goal to accomplish their economic prosperity. These countries do it by increasing their exports and hence, achieve trade surplus; which is one of the key macroeconomic goals. Trade liberalization or economic openness is thus the total or partial removal of any type of international trade barriers either in the form of tariffs or non-tariffs (Lee, 2005). Trade liberalization enables nations to focus on producing commodities and services in which they have a comparative advantage.

As much as trade liberalization promotes economic growth, increases employment opportunities and makes the countries better-off, it also has some repercussions. One of the major concerns is the adverse impacts on the environment and climate change by releasing greenhouse gas emissions (GHGs), primarily carbon dioxide (CO_2). These gases are harmful to the environment and human health in many ways. It also results in climate change that causes volcano eruption, ozone layer depletion, land degradation, sea level rise, floods and global warming. Most importantly, developing countries are facing severe consequences being vulnerable to climate change. Thus, in order to address whether trade liberalization is adversely impacting the environment in developing countries, this study will empirically analyze the causal relationship between CO_2 emissions and trade liberalization.

1.1 Study Objectives

The objectives of the paper are to find the impacts of trade liberalization on the environment in terms of computing the magnitude of trade liberalization on the primary indicator of environmental quality (CO_2).

2. Empirical Analysis

2.1 Data and Variable Description

To compute the relationship between trade liberalization and environment, the study uses CO_2 as the proxy for environmental quality as a dependent variable. Total trade as a percentage of Gross Domestic Product (GDP) has been used for the trade liberalization. Other control variables included in the model are GDP, FDI, capital, land, labor force and population density. Therefore, a panel data has been collected from 1990 to 2020 for all the available developing countries (147)

from the World Bank database. The list of developing countries has been prepared by the International Monetary Fund (IMF) based on human development index, GNI per capital and total population. The World Bank (2021) designated nations and territories with a GNI of $12,5353 or more as high-income economies. Anything less than that is classified as a developing country.

2.2 Empirical Specification

Before computation, the data was cleaned and organized in MS Excel. All missing values have been countered by adding a zero as there were large number of missing values that would have understated the results. The econometric model is presented in equations 1 as it was used by Onwachukwa, Yan, and Tu (2021). CO_2 is used as an environmental quality indicator to check the variation and direction of the impact, while i and t represents year and time, respectively. It is expected that a positive estimate would mean that higher trade results in increasing the CO_2 and the negative estimate would mean otherwise. Due to a large size of a control variable (capital), the dependent variables have also been converted into the log form.

$$CO_{2it} = \beta_0 + \beta_1 Tr_{it} + \beta_2 GDP_{it} + \beta_3 FDI_{it} + \beta_4 K_{it} + \beta_5 Labr_{it} + \beta_6 Land_{it} + \beta_7 PD_{it} + \boldsymbol{\varepsilon}_{it} \quad (1$$

Table 1: Variable Description - World Bank Dataset (https://data.worldbank.org)

Variable	Abbreviation	Unit
Carbon dioxide emissions	CO_2	Kiloton
Trade openness/ total trade	Tr	As a percentage of GDP
GDP growth	GDP	Annual percentage
Foreign direct investment	FDI	Net inflows as percentage of GDP
Capital	K	Gross fixed capital formation
Land size	Land	Land area as square kilometers
Labor force participation rate	Labr	Percentage of total population age 15+
Population Density	PD	People per square kilometer of land area
Error term	$\mathcal{E}$	
Country and time period	*it*	

2.3 Estimation Technique

For exploratory data analysis, correlation matrix, descriptive statistics and visuals like a scatter plot and matrix scatter plot will use prepared to find the relationship and trend between the variables. Subsequently, a multiple regression technique has been used for estimating the econometric model using XLMiner to capture the impacts of trade liberalization on the environmental quality. Firstly, the magnitude of impacts of trade will be computed against the environmental quality trade and other explanatory variables like capital, GDP, FDI, labor and population density. Additionally, the direction of CO_2 will indicate whether the impact is positive or negative, if there is any. The overall significance of these models will also be taken into consideration. Additionally, we will run an Exhaustive Search method for adding the variables to

select the best model. Based on the results, two best models will be compared and discussed under the results section.

3 Results and Discussion

3.1 Exploratory Data Analysis

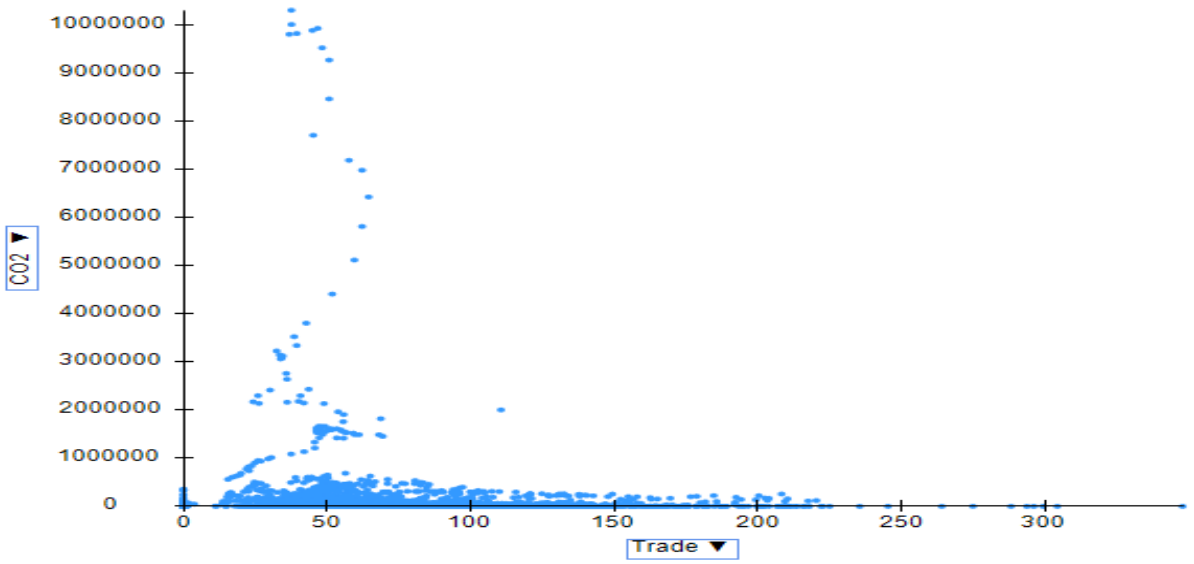

Figure 1: Relationship between CO2 and Trade

Figure 1 represents the relationship between the dependent variable CO_2 and the main explanatory variable, trade using a simple scatter plot. We can see a negative relationship between our variables which shows that as trade increases, carbon dioxide will also decrease based on the data. We will further investigate the causality between these two important variables in regression analysis for drawing a conclusion.

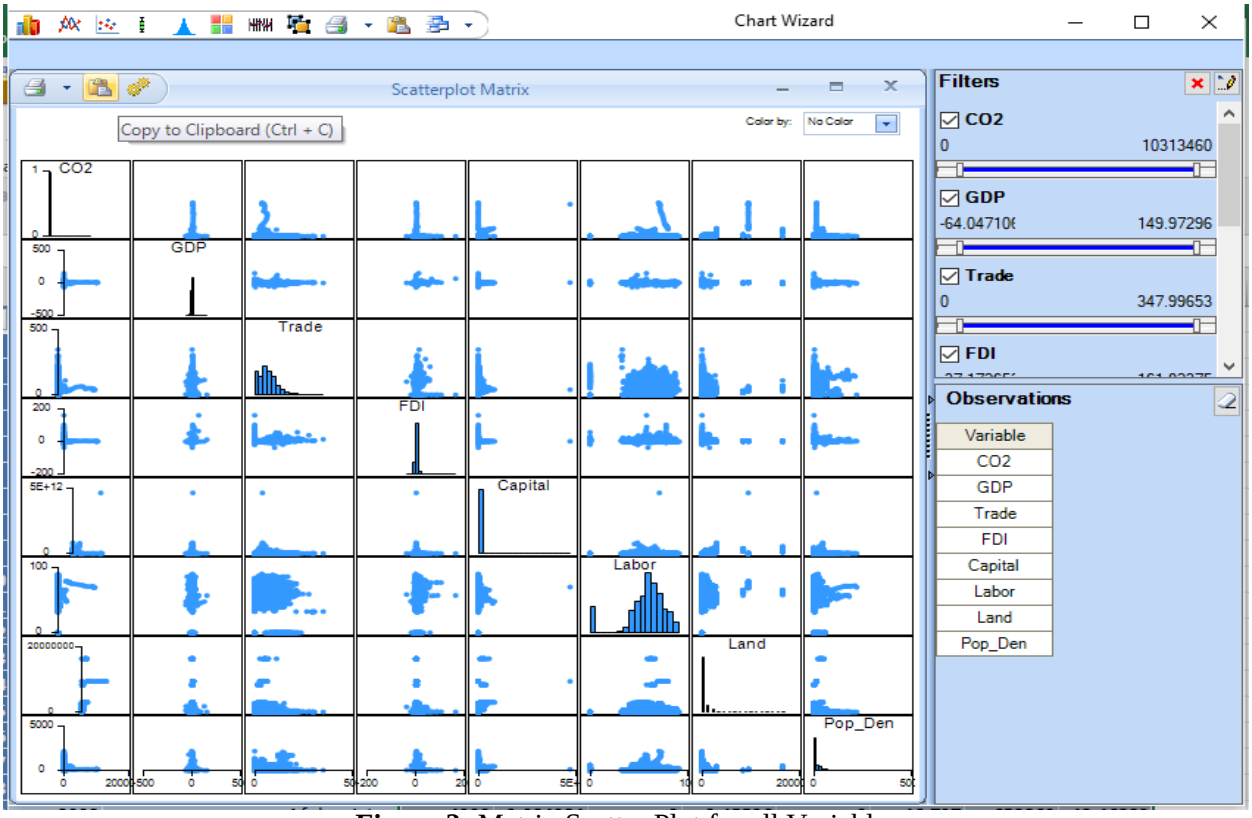

Figure 2: Matrix Scatter Plot for all Variables

Figure 2 shows a matrix scatter plot for all the variables used in this study. We can see that our dependent variable CO_2 has a positive relationship with all the explanatory variables except labor. This means that if labor force has been increased, carbon dioxide should be decreased. We can also witness that trade is positively skewed, while labor is skewed left.

Table 2: Correlation Matrix

	CO_2	GDP	Trade	FDI	Capital	Labor	Land	Pop_Den
CO_2	1							
GDP	0.054636	1						
Trade	-0.04248	0.052977	1					
FDI	-0.02894	0.183255	0.103285	1				
Capital	0.348314	0.009878	-0.05136	-0.04206	1			
Labor	0.058313	0.056821	-0.00773	-0.0466	0.021577	1		
Land	0.558972	0.004938	-0.1081	-0.05833	0.390055	0.105426	1	
Pop_Den	0.026974	-0.01071	0.045858	0.023898	0.028957	-0.08211	0.014715	1

Table 2 shows the correlation matrix for all the variables used in this study. This table also confirms that there is a negative association, but not too high, between carbon dioxide and trade and carbon dioxide and FDI. We can also see that there is no high correlation amongst any explanatory variable, which is a good sign because a high correlation between explanatory

variables leads to multicollinearity problem. However, correlation matrix is not best suited for finding the relationship between the models in a panel data. The reason is that it does not capture the variations for each country during some point in time.

Table 3: Descriptive Statistics

	CO_2	*GDP*	*Trade*	*FDI*	*Capital*	*Labor*	*Land*	*Pop_Den*
Mean	94437.40619	3.341447067	64.77610554	3.616955974	19689533635	56.96337988	660496.3208	127.6237368
Standard Error	8156.246804	0.102613309	0.66926658	0.093986023	1327390661	0.304739648	25909.51056	2.976133599
Median	4130	3.740514173	61.47668809	2.130473335	1116165483	60	155360	64.00744392
Mode	0	0	0	0	0	0	180	0
Standard Dev.	550591.9699	6.926968398	45.17921216	6.344578668	89606243698	20.57161917	1749035.899	200.9055513
Sample Var.	3.03152E+11	47.98289119	2041.161212	40.25367848	8.02928E+21	423.1915155	3.05913E+12	40363.04054
Kurtosis	212.9781456	81.56262284	1.193651253	130.3774655	1581.336083	2.206832451	48.38737056	25.09157496
Skewness	13.54307276	3.164101164	0.650759598	7.84628522	32.00299106	-1.474817491	6.423676144	4.24778074
Range	10313460	214.0200705	347.9965328	198.9964044	4.65629E+12	90.31999969	16389950	2181.516667
Minimum	0	-64.04710697	0	-37.17265386	0	0	0	0
Maximum	10313460	149.9729635	347.9965328	161.8237506	4.65629E+12	90.31999969	16389950	2181.516667
Sum	430351260	15226.97428	295184.7129	16482.46837	8.97252E+13	259582.1221	3009881734	581581.3687
Count	4557	4557	4557	4557	4557	4557	4557	4557

We can see from table 3 that both carbon dioxide and land size show the highest average and standard deviation, which is an indicative of high values in the data. Mode is 180 for land size also shows that many countries have the land size of 180 sq ft. GDP and FDI shows a negative value as it can be expected that for some countries, they experienced a negative growth and FDI. We can also see a large value used in capital as expected. By adding a zero for missing values, our count is same for all the variables.

3.3 Regression Analysis

First, we will use the Exhaustive Search method to find the two best models and then we will run the two models and select the best model between those.

Best Subsets								
Subset ID	**Intercept**	**GDP**	**Trade**	**FDI**	**Capital**	**Labor**	**Land**	**Pop_Den**
Subset 1	1	0	0	0	0	0	0	0
Subset 2	1	0	0	0	0	0	1	0
Subset 3	1	0	0	0	1	0	1	0
Subset 4	1	1	0	0	1	0	1	0
Subset 5	1	1	1	0	1	0	1	0
Subset 6	1	1	1	0	1	0	1	1
Subset 7	1	1	1	1	1	0	1	1
Subset 8	1	1	1	1	1	1	1	1

Best Subsets Details						
Subset ID	**#Coefficients**	**RSS**	**Mallows's Cp**	**R2**	**Adjusted R2**	**Probability**
Subset 1	1	1.38116E+15	2291.379769	-6.66134E-16	-6.66134E-16	0
Subset 2	2	9.49616E+14	154.231261	0.312449584	0.312298639	5.03418E-31
Subset 3	3	9.21966E+14	19.16836086	0.332469347	0.332176184	0.000766762
Subset 4	4	9.18421E+14	3.599825326	0.335035452	0.334597303	0.462957011
Subset 5	5	9.18033E+14	3.675795778	0.335316481	0.334732401	0.642356094
Subset 6	6	9.17725E+14	4.14611409	0.33553991	0.334809895	0.929548893
Subset 7	7	9.17695E+14	6.001905719	0.335560974	0.334684791	0.965181691
Subset 8	8	9.17695E+14	8	0.335561252	0.334538814	N/A

Based on these results, we will use subset 5 and 6 that seems to be the best model because of highest Adjusted R squared value.

Result Subset 5

Coefficients

Predictor	Estimate	Standard Error	T-Statistic	P-Value
Intercept	-54777.3288	12467.74507	-4.39352333	1.141E-05
GDP	3954.71119	961.937135	4.111195057	4.006E-05
Trade	205.813233	148.3557038	1.387295723	0.1654195
Capital	9.4245E-07	8.06449E-08	11.68640766	4.144E-31
Land	0.15762709	0.004150361	37.97913072	1.99E-274

Training: Prediction Summary

Metric	Value
SSE	9.18E+14
MSE	2.01E+11
RMSE	448838.1
MAD	109528.8
R2	0.335316

Result Subset 6

Coefficients

Predictor	Estimate	Standard Error	T-Statistic	P-Value
Intercept	-59409.7942	13017.286	-4.56391557	5.153E-06
GDP	3970.92675	961.9704202	4.127909411	3.726E-05
Trade	196.870767	148.5230862	1.325523	0.1850643
Capital	9.3987E-07	8.06672E-08	11.65119994	6.195E-31
Land	0.15758399	0.004150265	37.9696223	2.66E-274
Pop_Den	41.0331094	33.16999601	1.237054999	0.2161305

Training: Prediction Summary

Metric	Value
SSE	9.18E+14
MSE	2.01E+11
RMSE	448762.6
MAD	109078.9
R2	0.33554

Models Comparison

Subset 5	Subset 6
There are 4 regressors in the model	There are 5 regressors in the model
R square is 0.335316	R square is 0.33554*
Adj R Square is 0.33473	Adj R Square is 0.33480*
SE is 9.18E+14	SE is same due to large value and number of observations
RMSE is 448838.1	RMSE is 448762.6*
MAD is 109528.8	MAD is 109078.9*
All regressors are significant at 20%	4 variables are significant but one variable (pop) is insignificant at 20%
All factors remain same, an increase of 1% of trade results to an increase of 205 kilo tons of carbon dioxide	All factors remain same, an increase of 1% of trade results to an increase of 196 kilo tons of carbon dioxide

4. Conclusion

The objective of the study was to analyze the impacts of trade liberalization on the environmental quality in developing countries. A panel data of 147 developing countries was generated from World Bank Database from a period of 1990 to 2020 for carbon dioxide as a dependent variable, trade, GDP, FDI, Capital, Land, Labor and Population Density as the independent variables. We have used multiple regression technique for the estimation for selecting

the two models that are best due to both economic intuition as well as statistical evidence. Hence, we conclude that model 6 is the best model that predicts the adverse impacts of trade liberalization in deteriorating the environmental quality in developing countries. We can also state that there is a high negative impact of trade liberalization on the environmental quality as a slight increase in free trade would deteriorating the environmental quality at an enormous release of carbon dioxide emissions.

References

Bank, T. W. (2021). *New World Bank country classifications by income level: 2020-2021*. Retrieved from https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2020-2021

Lee, E. (2005). Trade liberalization and employment. *DESA Working Paper No. 5*.

Onwachukwa, C. I., Yan, K.-M. I., & Tu, K. (2021). The Causal Effect of Trade Liberalization on the Environment. *Journal of Cleaner Production*.