CO2 EMISSIONS AND EDUCATION EXPENDITURE ACROSS COUNTRIES

By Jaiveer Raikhy

And

Sviatolsav Shevchenko

EC204 (Empirical Economics II), Fall 2023

Abstract

Increasing greenhouse gas emissions, particularly carbon dioxide (CO_2) emissions in the last few decades emerged as a serious threat to human existence (Eyuboglu and Uzar). This research uses country-year panel data from all countries in the world between the years 1960 and 2022 to question whether higher expenditure on education leads to higher CO_2 emissions. 60.4% of the variation in CO_2 emissions is explained using a logarithmic model in terms of education expenditure, as well as both country and year-fixed effects. A statistically significant positive relationship is found between education expenditure and CO_2 emissions.

I. Introduction

This research paper investigates the possible effect of education expenditure on CO_2 emissions among all countries across 62 years. Researchers previously found mixed findings in regard to the association between education and CO_2 emissions (Sart et al.). On one hand, a higher education level within a society can create greater awareness about environmental issues like global warming caused by increasing CO_2 emissions (Eyuboglu and Uzar). On the other hand, higher education expenditure is required for a higher education level within a society. A by-product of this is also increased productivity levels leading to economic boosts that increase economic growth and the consumption of energy, thereby leading to higher emissions of CO_2 (Zaman et al.).

In this research paper, within the context of past mixed findings on the association between education and CO_2 emissions (Sart et al.), we explore a possible association between education and CO_2 emissions based on our own data and empirical model. Our hypothesis is that higher education expenditures have a positive association with CO_2 emissions. We hypothesize that this is because other economic and industrial factors would have a positive association with education expenditures, thus leading to higher CO_2 emissions as a by-product. The logarithmic model presented in this paper explains 60.4% of the variation in CO_2 emissions explained by education expenditure with country and year-fixed effects. This research concludes that there is a statistically significant positive relationship between education expenditure and CO_2 emissions, however, it suggests an association due to higher education expenditures being likely linked to better economic conditions and greater industrial activity, leading to higher CO_2 emissions.

II. Literature review

The last few decades have seen an exponential rise in greenhouse gas emissions, namely CO_2 emissions, raising concerns about the ongoing environmental issue of global warming among the public and policymakers alike (Liu et al.). These concerns are well justified, as the world average temperature in the year 2017 was 1.5° Fahrenheit higher than between the years 1901 and 2000 (Liu et al.). Albeit there is a consensus on the reduction of CO_2 emissions in recent years, according to BP Statistics, the global CO_2 level of emissions increased rapidly from 11,193 million tonnes (mt) in 1965 to 33,890 mt in 2018 (Eyuboglu and Uzar). In 2021, this figure reached its highest level, equaling about 36,300 mt, thus showing that global warming continues to be a great existential threat to the world (Sart et al.).

The majority of policymakers and researchers in the past aimed to address the issue of global warming through the factors of GDP, trade, improved industrialization, energy efficiency, use of clean energy, financial development, foreign direct investment, urbanization, and population growth (Eyuboglu and Uzar; Sart et al.; Zaman et al.). Although previous studies made a significant contribution through the consideration of these economic and institutional factors as determinants of CO_2, education was often overlooked as a factor of significance in combating environmental issues such as global warming caused by the increase in CO_2 emissions (Eyuboglu and Uzar). However, higher levels of education in society (Eyuboglu and Uzar; Sart et al.; Liu et al.) and a higher expenditure on education (Zaman et al.) proved to be beneficial in countering greater CO_2 emissions.

As results show, development in higher education within a society can be a catalyst for the decrease in CO_2 emissions, however developing education levels requires higher energy consumption, consequently leading to “environmental degradation” partly characterized by higher CO_2 emissions (Eyuboglu and Uzar). Somewhat differently, education coefficients were significant and negative in four models, showing that education has a role of reducing “ environmental degradation” (Liu et al.), likewise increasing education expenditure led to similar results, decreasing CO_2 emissions (Zaman et al.). Overall, education is a significant determinant of "environmental degradation" associated with higher CO_2 emissions, however, country-level findings differ based on country-specific characteristics (Sart et al.).

III. Data Description

This research utilizes panel data at the country level from the years 1960 to 2022 in order to observe a possible relationship between a country's expenditure on education and its CO_2 emissions. Information on each variable has been collected from as many countries as possible, over as many years as possible for a total of 3,443 observations. Country-time period groups for which data was not available were excluded from our analysis.

In order to gather the CO_2 emissions and education expenditure from 1960 to 2022 across all the countries, we use the World Bank Data Bank repository, in particular, the World Development Indicators (WDI) dataset. The World Bank is a global financial institution that extends loans and grants to the governments of countries with low to middle incomes. These funds are intended to support the implementation of capital projects. The World Development Indicators (WDI) represent the principal compilation of development metrics by the World Bank, sourced from internationally recognized official sources. This collection offers the latest and most precise global development information, encompassing estimates at the national, regional, and global levels.

In addition to our two main variables mentioned above, we are also considering several other variables. This is because these variables are either related to activities that would be associated with higher CO_2 emissions or a higher expenditure on education. These variables are our control variables. The data for these variables was also gathered from the World Development Indicators dataset. From the years 1960 to 2022, across all the countries, we had

491 observations. The difference between the number of observations arises from the fact that not all countries report on all of our selected variables. Country-time period groups for which data was not available for all of our variables were excluded from the analysis.

It is essential to highlight the panel data nature of this study, allowing for the examination of both temporal trends and cross-country variations.

Table 1 provides a description of our variables and their units of measurement.

Table 1: Variable Descriptions

CO2_emissions	Emissions of CO_2, measured in Kg per PPP of GDP
education_expenditure	Government expenditure on education, measured as a percentage of GDP
GDP_per_capita	GDP per capita per PPP, measured in constant 2017 US dollars
exports	Value of Exports of goods and services, measured as a percentage of GDP
natural_resources	Value earned from natural resource rents, measured as a percentage of GDP
manufacturing	Value of Manufacturing, measured as a percentage of GDP
industry	Value of Industry including construction, measured as a percentage of GDP
literacy_rate	Literacy rate, ages 18+
population_growth	Population growth, annual percentage
agriculture	Value of agriculture, measured as a percentage of GDP
labor_participation	Labor force participation, percentage of adult population
log_co2	Natural log of CO_2 emissions
log_expenditure	Natural log of expenditure
sq_expenditure	Square of expenditure
expenditure_GDP	Interaction term of expenditure on education and GDP per capita

Table 2: Summary Statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
GDP_per_capita	6,095	18,318	20,568	436.4	157,602
CO2_emissions	5,509	0.319	0.302	0	2.382
education_expenditure	4,925	4.350	1.953	0	44.33
exports	8,701	36.57	29.15	0.00538	433.8
natural_resources	9,235	6.829	11.05	0	88.59
population_growth	13,423	1.746	1.782	-27.72	20.47
industry	8,427	26.70	12.56	2.365	90.51
agriculture	8,485	15.80	14.25	0.0125	89.41
manufacturing	7,707	12.78	7.309	0	49.88
labor_force	5,982	65.84	10.83	32.08	90.96
literacy_rate	987	80.39	21.22	5.405	100

IV. EMPIRICAL MODEL

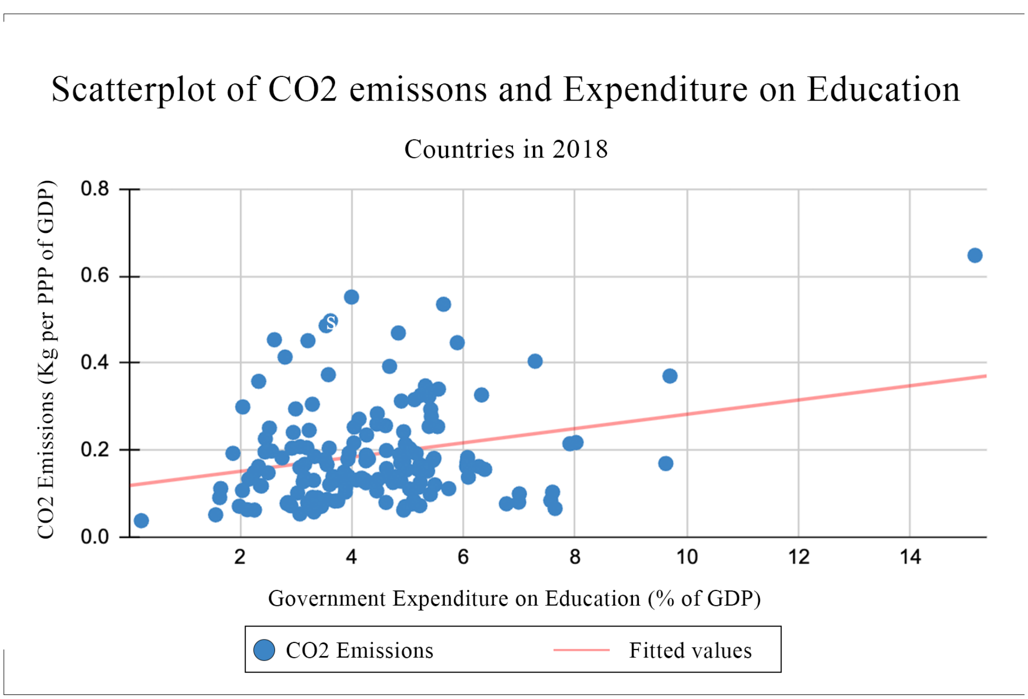
We created a model to estimate the CO_2 emissions of a country as a function of its expenditure on education using country-level and time-fixed effects from our panel data set:

$$CO2emissions = \beta_0 + \beta_1 educationexpenditure_{it} + a_i + u_{it} + dt_t$$

Within our model, while holding the country *i* and year *t* constant, *CO2_emissions* represents the CO_2 emissions per kg per PPP of GDP in country *i* in year *t*. It is also the dependent variable of our model. Whereas, *education_expenditure* is our main independent variable, and it represents the expenditure on education of country *i* in year *t*, as a percentage of their GDP. We estimate that the coefficient β_1 will be positive, as it is predicted that a higher expenditure on education will be associated with higher levels of economic activity, industrialization, and technological development in a given country, leading to higher CO_2 emissions. However, an argument could also be made that the coefficient would be negative. This is because a higher expenditure on education would also be associated with a more educated workforce/population, which would be more environmentally conscious and hence work towards innovations in sustainability, which would lead to lower CO_2 emissions for the country. We also hypothesize that our other control variables would also have a statistically significant relation with the CO_2 emissions when looking at country i in year t. Given that our model relies on panel data spanning from 1960 to 2022, the inclusion of the time dummy variable, denoted as dt_t, accommodates for temporal variations. Additionally, the variable a_i serves to address unobserved time-constant factors.

Figure 1, shown below, is a scatter plot to show the relation between expenditure on education and CO_2 emissions by country in the year 2022. Looking at the plot and the line of best fit, there is evidence that suggests a positive relationship, just as it was hypothesized. The graph only represents the data from 2018, as it gives a better idea about the relationship between CO_2 emissions and education expenditure across countries when time is held constant.

Figure 1



V. RESULTS

Table 3: Regression results

VARIABLES	(1) CO2 Emissions	(2) CO2 Emissions	(3) CO2 Emissions	(4) CO2 Emissions	(5) CO2 Emissions	(6) CO2 Emissions	(7) CO2 Emissions	(8) CO2 Emissions	(9) CO2 Emissions
eudcation_expenditure	0.0199***	0.00174	0.00940**	0.0180***	0.0384***		0.0428***		0.0258*
	(0.00202)	(0.00607)	(0.00368)	(0.00471)	(0.00590)		(0.0123)		(0.0147)
GDP_per_capita				-8.64e-06*	-8.73e-07	-3.11e-06	-1.32e-05*	-1.14e-05***	-2.95e-06
				(5.03e-06)	(3.93e-06)	(3.53e-06)	(4.22e-06)	(4.22e-06)	(3.49e-06)
exports				-0.000345	0.000167				
				(0.000832)	(0.000860)				
natural_resources				0.00223	0.00171				
				(0.00176)	(0.00163)				
population_growth				-0.0145	-0.0123				
				(0.0127)	(0.0124)				
industry				-0.00169	-0.00167				
				(0.00233)	(0.00228)				
agriculture				-0.00141	-0.00130				
				(0.00253)	(0.00234)				
manufacturing				0.00106	-0.000865				
				(0.00267)	(0.00242)				
labor_force				0.00224	0.00404				
				(0.00359)	(0.00342)				
literacy_rate				0.00388***	0.00354***	0.00394***	0.0122**	0.0120**	0.00391***
				(0.00132)	(0.00118)	(0.00120)	(0.00411)	(0.00403)	(0.00121)
expenditure_gdp					-2.69e-06***	-1.24e-06***	-2.59e-06***	-2.74e-06***	-1.48e-06**
					(7.19e-07)	(5.07e-07)	(1.10e-06)	(1.02e-06)	(6.04e-07)
log_expenditure						0.117***		0.239***	
						(0.0320)		(0.0777)	
sq_expenditure									-0.000107
									(0.000556)
Constant	0.192***	0.273***	0.432***	-0.00579	-0.155	-0.114	-2.050***	-2.167***	-0.165
	(0.00954)	(0.0269)	(0.0305)	(0.259)	(0.245)	(0.147)	(0.328)	(0.318)	(0.145)
Observations	3,443	3,443	3,443	491	491	537	537	537	537
Adjusted R-squared	0.030	0.000	0.407	0.586	0.623	0.545	0.599	0.604	0.555
Number of CountryName_encoded		182	182	113	113	124	124	124	124
Country FE	NO	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3 shows the various models we tried/made with our data. Model 1 shows the simple linear regression wherein *CO2_emissions* were regressed on *education_expenditure*. The coefficient in this on our main independent variable education_expenditure was 0.0199, thus hypothesizing a positive relationship between a country's expenditure on education and its CO_2 emissions which was significant with a p-value of 0.000. This model accounts for 3% of the variation in CO_2 emissions explained by education expenditure. For model 2, we decided to add fixed effects for country and run the same regression model. In this model, the regression on our main independent variable changed to 0.00174, but it is statistically insignificant.

In model 3, we added fixed effects for years to our regression, which resulted in the coefficient changing to 0.0094, which was statistically significant at the 5% level. Although the coefficient is smaller than the one in model 1, we decided to keep fixed effects for both country and year in our final model. This is not only because of the statistically significant coefficient, but also because it allows for the data to be compared across all countries in a given year, and to compare the data for a single country across many years. This model still hypothesized a positive relationship. In model 4, we decided to add all our control variables, which are mentioned in table 1. For this model, the coefficient on expenditure changed to 0.018, which is significant at the 1% level. This also highlights that we had a negative omitted variable bias in our model earlier. Additionally, for our control variables, only two had statistically significant coefficients. These being GDP per capita and literacy rates, significant at the 10% and the 1% level respectively. However, the coefficient on both of these variables, especially GDP per capita, was extremely small.

For model 5, we included an interaction term in the regression, this term being expenditure_gdp. This term represents the interaction between a country's GDP per capita and expenditure on education. The interaction between education expenditure and GDP per PPP captures how the economic context influences the relationship between education spending and CO_2 emissions. In this model, the β_1 coefficient is 0.0384, which is significant at the 1% level. Additionally, the coefficient GDP per capita is not significant anymore, however, the interaction term coefficient is significant at the 1% level. The coefficient on literacy rates has changed only minutely, to 0.00354, still significant at the 1% level.

In models 6 - 9, we decided to test whether there is a non-linear relationship between *education_expenditure* and *CO2_emissions*. Additionally, for these models, we decided to take out the majority of our control variables, leaving GDP per capita, literacy rates, and our interaction term expenditure_gdp. We did this as our control variables were statistically insignificant on their own, and after conducting numerous f-tests, we found that they were still insignificant. So we decided to remove them from our further models. However we kept the variable GDP in our models as our interaction term is significant, and it is partly based on GDP.

In model 6, we regressed *CO2_emissions* on the natural log of the *education_expenditure*. In this model, the coefficient on the log of *education_expenditure* was 0.117, significant at the 1% level. This shows that there is indeed a non-linear, logarithmic relation. The coefficient on log_expenditure shows that a 1% increase in education_expenditure is associated with a 0.117 kg increase in CO_2 emissions per PPP of GDP.

For model 7, we regressed the log of *CO2_emissions* on *education_exoenditure*, to form a log-linear model. Our coefficient on *education_expenditure* is 0.0428, significant at the 1% level, and this model accounts for 59.9% of the variation in CO_2 emissions, explained by expenditure on education. Additionally, compared to model 6, our constant is significant at the 1% level. Moving on to model 8, we regressed the log values of *CO2_emissions* on the log values of education_expenditure, resulting in a log-log model. This model accounted for 60.4% of the variation in CO_2 emissions, explained by expenditure on education, making it the most efficacious and efficient model. Our model hypothesizes that a 1% increase in expenditure on education is associated with an increase in CO_2 emissions per PPP of GDP by 0.239%.

In model 9, we tested to see if there was a quadratic relationship between CO_2 emissions and education_expenditure. To do so we generated a squared term of the education_expenditure and ran a regression. The coefficient on the squared term in this case was -0.000107. However, this number was statistically insignificant, meaning that there is no quadratic relationship.

Based on our results and the discussion above, we believe that model 8 is the most optimum one to explain the impact of expenditure on education on CO_2 emissions.

VI CONCLUSION

As the results show, our original hypothesis holds partially true, as there is a positive relation between expenditure on education and CO_2 emissions, but this relation is in fact non-linear. Also unlike the original hypothesis, almost all of our control variables turned out to be insignificant. This was quite surprising as we had anticipated that at least some variables, such as exports and industry would have a sizable impact on CO_2 emissions. Even though we found a statistically significant relationship between education expenditure and CO_2 emissions, we still believe that there might be other variables that would be able to better explain the variance in emissions. This is despite the fact that our model had an R-squared value of 0.604 (or 60.4%). In our models, the R-squared value changes drastically between our different models. In model 1, the R-squared was only 3%, but in model 8, the value was 60.4%. We didn’t take the R-squared into account when considering which model to choose. This is because the R-squared tends to get inflated when using a fixed effects regression. So, it would not be entirely accurate to compare different R-squared values. We too were a bit surprised to see the high values, as even though we had hypothesized that there would be a relation between the expenditure on education and CO_2 emissions, we thought that it would account for a smaller amount, more in line magnitude wise to model 1. As there are a lot of other variables that could better address the CO_2 emissions. Although we can't effectively compare our R squares across all of our models, we can compare it across models 3 to 9, as they all have fixed effects for both year and country.

Our models also confirmed that education expenditure is not associated with a decrease in emissions, as some might have argued. This is because out of our 9 models, not a single one had a negative coefficient on education_expenditure.

One important thing to note is that our model does not suggest that increasing education expenditure will lead to higher emissions, instead, it suggests that higher expenditures on education are associated with a higher level of CO_2 emissions. This is because higher education expenditures are linked to a country's overall economic conditions, and countries with higher levels of economic activity will have higher CO_2 emissions. Instead, however, our model can be used to estimate the emissions of a given country, as expenditure on education is data that all countries record, and can be used to inform future policy decisions regarding CO_2 emissions and climate change.

VII LIMITATIONS AND FUTURE RESEARCH

One aspect of our model and data that can be seen as a particular weakness or limitation is that for not a single year from 1960 to 2022, was there a year which had data for our main variables. This was an even larger issue when we incorporated our other control variables into the model. This can be seen in our results table as before incorporating our control variables we had 3,443 observations. And after adding them, we only had 491 observations. Ideally, we would have had 195 observations per year, that is, one observation per country per year.

There was also no way for us to ensure that the data we had was accurate. Even though our data is from the Work Bank Database, the countries might not have been entirely honest when reporting their data. Or there might be inconsistencies in the way that countries measure/calculate things like their CO_2 emissions, literacy rate, etc.

Future research should take into account the limitations mentioned above and work around them to further explore the relationship between emissions and education expenditure. Another important aspect to consider would be the fact that different countries industrialized at different rates, and some countries are still in the process of doing so. Although our GDP variable does partially account for this, as industrialization is associated with higher GDP, our model doesn't fully consider this due to a lack of data and methods to quantify it. Additionally, future research and models could also take into consideration other variables that we didn't think of that would perhaps have a larger impact on determining a country's CO_2 emissions.

References Cited

Eyuboglu, Kemal, and Umut Uzar. “A New Perspective to Environmental Degradation: The Linkages between Higher Education and CO2 Emissions - Environmental Science and Pollution Research.” *SpringerLink*, Springer Berlin Heidelberg, 19 Aug. 2020, link.springer.com/article/10.1007/s11356-020-09414-8#citeas.

Liu, Nan, et al. “Does Financial Inclusion and Education Limit CO2 Emissions in China? A New Perspective - Environmental Science and Pollution Research.” *SpringerLink*, Springer Berlin Heidelberg, 23 Oct. 2021, link.springer.com/article/10.1007/s11356-021-17032-1. Accessed 07 Dec. 2023.

Sart, Gamze, et al. “Economic Freedom, Education, and CO2 Emissions: A Causality Analysis for EU Member States.” *MDPI*, Multidisciplinary Digital Publishing Institute, 30 June 2022, www.mdpi.com/1660-4601/19/13/8061#B45-ijerph-19-08061. Accessed 06 Dec. 2023.

Zaman, Qamar uz, et al. “Investigating the Nexus between Education Expenditure, Female Employers, Renewable Energy Consumption and CO2 Emission: Evidence from China.” *Journal of Cleaner Production*, Elsevier, 5 June 2021, www.sciencedirect.com/science/article/pii/S0959652621020424#bib77.