1. Introduction

Climate change is one of our most pressing global issues, with far-reaching and potentially devastating consequences for both the natural world and human society. It is now widely accepted that the accelerating climate changes on the planet are mainly a result of human greenhouse gas (GHG) emissions (Världsnaturfonden, n.d). The Paris Agreement was embraced by the United Nations in 2015 to combat global warming. One of the main goals is that GHG emissions should peak before 2025 and then decrease by 43% by 2030 (United Nations Framework Convention on Climate Change (UNFCCC), n.d). Carbon dioxide (CO_2) emissions are the most common GHG and are the main reason for the observed climate changes during the last century. CO_2 emissions result from activities such as the burning of fossil fuels, volcanic eruptions, and breathing. Since the beginning of the 18th century, the atmospheric concentrations of CO_2 have increased by roughly 50% (National grid, n.a). Besides increasing temperature, CO_2 also contributes to the acidification of forests and seas, damaging the ecosystem (Naturvårdsverket, n.a).

In conjunction with the global warming attributed to the release of GHG, numerous nations encounter significant issues regarding air quality, stemming from emissions of substances such as sulfur dioxide (SO_2) and particulate matter (PM). SO_2 emissions are a byproduct of the burning of fossil fuels. Prolonged exposure to high levels of SO_2 can have direct and detrimental effects on human health, such as causing extensive damage to the lungs and respiratory system (American Lung Association, 2022). Aggregated SO_2 emissions worldwide decreased by approximately 31% between 1990 and 2015 (Aas et al, 2019). Particulate matter (PM) is a complex mixture of particles suspended in the air. PM is typically classified according to its diameter, with particles with a diameter less than 10 microns (µm)[1] being labeled as PM_{10}. PM_{10} is inhalable by humans and can pose serious health risks, such as causing damage to the respiratory system. The sources of PM_{10} emissions include the burning of fossil fuels, dust from industrial sites, and incineration of waste materials (California Air Resource Board, n.a.).

[1] $10\mu = 10\times10^{-6}$ meters

In addition to the pressing concern of climate change, poverty poses a significant challenge that requires urgent attention. The World Bank (2022) reported that the Covid-19 pandemic led to a surge in the number of individuals living in extreme poverty, with an estimated 700 million affected. To combat poverty, fostering economic activity and growth in developing nations is essential (World Bank, 2022). Furthermore, economic growth plays a crucial role in developed economies, as it contributes to improvements in various domains, such as education, social welfare, and overall quality of life (Our World in Data, 2021). The United Nations (UN) identified economic growth as a central objective in Goal 8 of its global targets. It aims to achieve a sustainable per-capita growth rate of at least 7% per annum for the least developed countries (UN, 2022).

Economic growth leads to increased production with associated pollution, rendering an activity with significant environmental impact. Nonetheless, the relationship between economic growth and environmental degradation is complex. Grossman and Krueger's (1991) introduction of the environmental Kuznets curve (EKC) depicts an inverted U or conceivably N-shaped relationship between economic growth and environmental degradation. In this regard, economic growth tends to elevate environmental damage initially, but after reaching a certain income threshold, economic growth reduces environmental degradation. This outcome results from factors such as technological advances, increasing environmental concerns, and structural modifications (Grossman and Krueger, 1991). Studies have demonstrated that the relationship between economic growth and environmental degradation is complex, with the long-term relationship potentially diverging from that observed in the short term. Additionally, several country-specific characteristics may influence this relationship (Wang, Yang, & Li, 2023; Torras & Boyce, 1998; Aytun & Akin, 2022; Chen, Huang, & Lin, 2019).

The effects of country characteristics on the shapes of the EKC are important for several reasons. Firstly, understanding how country characteristics interact with the EKC is vital for developing more accurate predictions regarding future environmental degradation. Moreover, if a country's characteristics can be modified, it can pave the way for a more sustainable developmental trajectory for developing nations. A research paper by Chen, Huang, and Lin (2019) established that environmental awareness might be a country-specific characteristic that influences the shape of the EKC. The authors ascertained that countries with greater environmental awareness have a more environmentally friendly course of development (Chen, Huang, and Lin, 2019). Zheng, Khan, Sun, and Danglun (2014) found similar patterns across

Chinese regions and therefore concluded that environmental awareness led to more sustainable development in China.

The present study aims to undertake a more comprehensive empirical investigation that advances the understanding of the relationship between environmental awareness and the shape of the EKC. The findings of this research carry significant implications for resource allocation toward environmental education and awareness-raising initiatives. Moreover, this knowledge has far-reaching implications for forecasting the developmental trajectories of nations, particularly those that are in the process of developing. This arises from the fact that the citizens of these countries may exhibit a higher level of environmental awareness owing to the cumulative knowledge and experiences of earlier-developing countries. The findings of this study can, therefore, provide valuable guidance for policymakers and stakeholders in their efforts to foster more sustainable and environmentally responsible development paths.

The study will contribute to the present literature by evaluating the impact of environmental awareness on the EKC and including more developed countries and thus cubic terms, controlling for other country characteristics and unobserved heterogeneity. Since some earlier studies have shown evidence for an N-shaped EKC, an extension of previous studies on the EKC and environmental awareness with the inclusion of more developed countries should be of interest. This should give a more comprehensive insight to the relationship of interest; this also generates a larger sample yielding higher statistical power of tests. Consequently, this thesis aims to answer the following research question:

- Does environmental awareness mitigate emissions generated from economic development?

The structure of this thesis is organized as follows. Firstly, a literature review will be presented by summarizing the most relevant previous studies related to the present research questions. Following this, a theoretical framework will be outlined that explains the underlying concepts and principles of the EKC and the theoretical foundation for the cause-and-effect relationship that is empirically evaluated in this study. Next, the data utilized in the analysis will be explained, followed by a description of the method employed to address the research question. Following this, the results of the analysis will be presented, followed by an in-depth discussion of the findings. The final section of the study will be dedicated to the presentation of conclusions, summarizing the key findings and their implications for future research and policymaking.

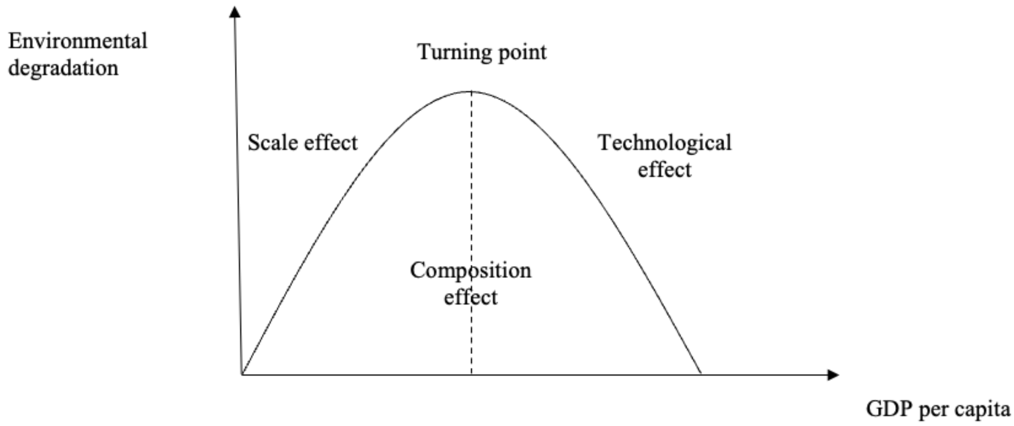
2. Theoretical background

This section will present the theoretical foundations for this thesis. It will start by presenting the theoretical foundations of the EKC, followed by an examination of the drivers behind the N-shaped EKC pattern. Subsequently, it will explain the theoretical rationale supporting the existence of a relationship between environmental awareness and the EKC.

2.1 Environmental Kuznets Curve (EKC)

As previously stated, the EKC explains the empirical relationship between economic growth and environmental degradation. Grossman and Krueger (1991) proposed a framework to explain the theoretical mechanisms behind the EKC, which comprises three distinct parts. The first part is known as the "scale effect," which posits that increased economic output leads to increased pollution levels, all else equal. The second part is the "composition effect" resulting from structural changes in a country, such as industrialization. The composition effect can be either positive or negative. When the shift is from high-polluting to less-polluting industries, it reduces pollution levels and vice versa. At higher income levels, the composition effect typically shifts from industrial sectors towards service sectors, which is generally accompanied by a decrease in environmental degradation. However, at lower income levels, the shift might be from agricultural to industrial sectors, leading to increased environmental degradation. The third part of the development path is the "technological effects", which result from technological advances that promote abatement and increase production efficiency (Grossman & Krueger, 1991). A graphical representation of the EKC is presented in Figure 1 below.

Figure 1. The inverted U-shaped EKC

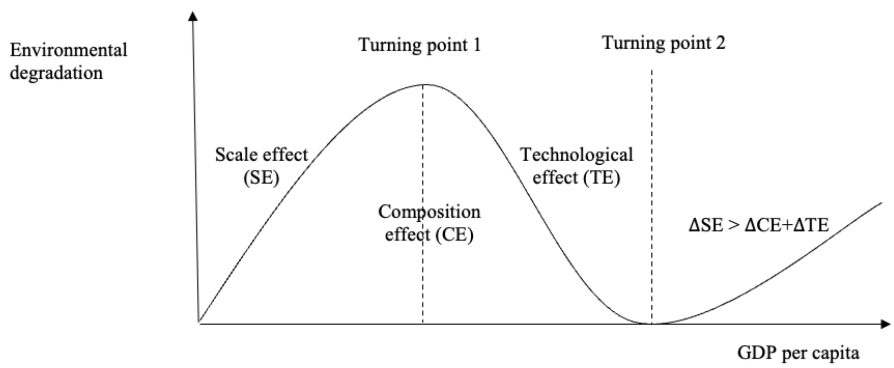


Note: Theoretical shape of the inverted U-shaped EKC
Source: Own.

2.2 N-shaped EKC

Some critique against the inverted-U-shaped EKC has been expressed since it was first proposed. One of the critiques is that the relationship between environmental degradation and growth is unlikely to hold in the long run. As a result, the N-shaped EKC was presented. Torras and Boyce (1998) suggested that this might emerge from the scale effect overcoming the composition and technological effect in the long run. This could be due to diminishing returns from technological advances or that the industry allocation cannot be improved further from an environmental perspective (Torras & Boyce, 1998). A graphical representation of the N-shaped EKC is presented in Figure 2 below.

Figure 2. The N-shaped EKC

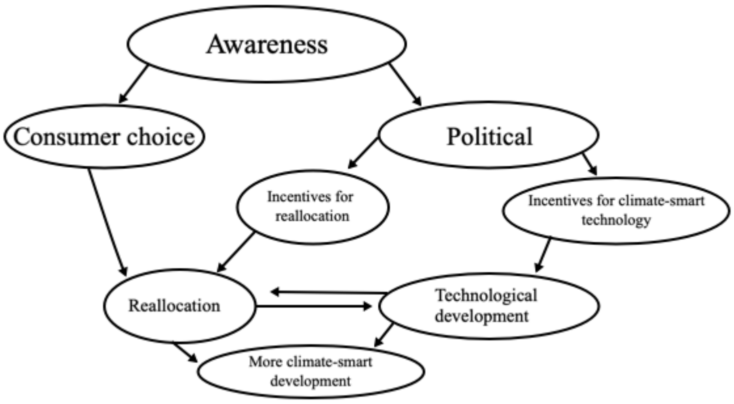


Note: Theoretical shape of the N-shaped EKC
Source: Own.

2.3 Environmental Awareness and the EKC

This section presents a theoretical framework that explains the potential impact of environmental awareness on the EKC. The framework is divided into two distinct perspectives, namely, the political and consumer choice perspectives. Although several other channels also may exist, this section only focuses on these two channels to reduce complexity. Some interactions between the channels are presented with the understanding that several additional interactions are likely to exist within an empirical setting. The theoretical relationship between environmental awareness and climate-smart[2] development is depicted in Figure 3. Nevertheless, it is important to note that, for the empirical part, a reduced form of Figure 3 will be estimated, where awareness is tested on a country's development path in a more direct way.

Figure 3. Theoretical effect of awareness on climate-smart development.



Source: Own

2.3.1 Environmental Awareness and the EKC - the political perspective

According to Aswathanarayana, Harikrishnan, and Kadher-Mohien (2010), public environmental awareness plays a critical role in shaping political actions. This is because the government is unlikely to adopt radical green policies unless they have widespread public support (Aswathanarayana, Harikrishnan, & Kadher-Mohien, 2010). This emphasizes the importance of the political channel of the relationship investigated in this thesis.

Environmentally friendly policies should, assuming they work, reduce environmental degradation. Nonetheless, such policies incur costs that could be reflected in either the reduction

[2] In this thesis, a more climate-smart process means that the process leads to less pollution in relation to a less climate-smart process.

of other public goods like healthcare; or a decrease in disposable income due to tax financing of environmental policies. In this theoretical framework, we assume that the financing of environmental policies is achieved through a decrease in non-green policies[3]. However, the framework can be extended to incorporate a reduction in disposable income as an alternative representation of costs.

The median voter theorem states that under some assumptions, such as single-peaked preferences and a one-dimensional choice, the most preferred policy of the median voter is the Condorcet winner and thus will determine elections (Hindriks & Myles, 2013). So, since the government wants to be re-elected, it should aim toward maximizing the utility for the median voter. Assuming the median voter has a Cobb-Douglas utility function for preferences regarding green and non-green policies, the following representation is possible:

$$U^M = G^{\alpha} N^{1-\alpha} \; s.t. T = GP_G + NP_N, \; \alpha \in [0,1] \tag{1}$$

Where G is the number of green policies, N is the quantity of non-green policies, P is the price per unit of policy, α is a preference parameter for the median voter, and T is the total budget for the government. Furthermore, the preference parameter α is assumed to be an increasing function of environmental awareness. Setting $P_N = 1$ and inserting the remaining condition into the utility function (1) yields a first-order condition for G that is:

$$\frac{\partial L}{\partial G} = aG^{a-1}(t - P_G G)^{1-a} - \frac{(1-a)P_G G^a}{(t-P_G G)^a} = 0 \tag{2}$$

Simplifying and rearranging[4] (2) gives:

$$G = \frac{\alpha t}{P_G} \tag{3}$$

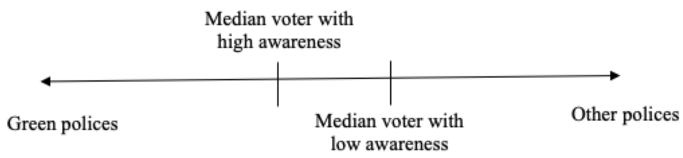
From equation (3), it is apparent that G increase as α increase, all else equal. So, awareness is implicitly connected to the level of green policies through the preference parameter α. Therefore, the theoretical disparity between the allocation of government expenditures between

[3] All policies that do not target environmental issues is counted as green policies. So, policies promoting healthcare, schooling etc. is counted as non-green policies.

[4] Simplifying (2) gives: $a(T - GP_G) - (1 - a)G(T - GP_G)/P_G = 0$, rearranging this gives: $G(T - GP_G)\frac{(a-P_G)}{P_G} = aT$, solving for G gives: $G = \frac{aTP_G}{(a-P_G)(T-GP_G)}$ which then can be simplified into $G = \frac{aT}{P_G}$

green policies and other policies, considering countries with a median voter with high and with low awareness, respectively, is:

Figure 4. Theoretical effect of environmental awareness on the median voter.



Source: Own.

2.3.2 Environmental Awareness and the EKC - the consumer choice perspective

To demonstrate the theoretical effect of environmental awareness on the EKC from a consumer choice perspective, a hypothetical example will be shown where the representative consumer selects between an environmentally harmful product, X, and a more environmentally friendly alternative, Y. Good Y has similar properties but a higher price and lower environmental damage than X. So, the consumer must make a trade-off between the quantity of consumption and the quality of the environment. Geng and He (2021) showed that environmental awareness leads to higher public demand for environmental quality (Geng & He, 2021). So, assuming environmental awareness leads to a greater marginal disutility of environmental degradation for the representative consumer should be reasonable. The larger disutility of environmental degradation will thus lead to the representative consumer being more willing to choose Y over X. The reallocation of consumption towards environmentally friendly goods should create profit opportunities for climate-smart products. The increased demand for climate-smart products might lead to an increased demand for climate-smart technology relative to other technologies since the producers want to maximize production efficiency. This might then yield a technological effect.

There are several ways in which the channels could be able to interact. For example, incentives for improving climate-smart technology can promote production and thus lower prices for low-polluting goods. This should then lead to a reallocation of expenditures for consumers towards the climate-friendlier good, and this effect should increase with awareness. Also, incentives like subsidies for low-polluting goods or taxes for high-polluting goods will create a reallocation between markets. Yielding the same interaction effect as in the climate-smart technology incentive case. So, from a theoretical perspective, the technological effect may be stronger with environmental awareness.

3. Literature review

Several earlier studies have shown evidence for the inverted U-shaped EKC (Ahmed, Uddin, & Sohag, 2016; Leitão, 2010; Chen, Huang, & Lin, 2019; Grossman & Krueger 1991). Evidence for a U-shaped EKC for some country groups has been presented as well (Culas, 2012). Other studies have confirmed the N-shaped EKC (Wang, Yang, & Li, 2023; Torras, & Boyce, 1998). Additionally, inconclusive results about the shape of the EKC have been presented (Allard, Takman, Uddin, & Ali, 2018). Dasgupta, Laplante, Wang & Wheeler, (2002) advocates that the shape of the EKC should differ depending on what environmental indicator is used. For example, PM and SO_2 tend to peak earlier than CO_2. This is mainly a result of the differences in the origins of the emissions and the resulting differences in the connection to different parts of a country's development path (Dasgupta, Laplante, Wang & Wheeler, 2002).

Studies have also been done by evaluating how different country characteristics affect the EKC and ecological footprints of countries. Sharif, Uddin, and Alexiou (2022) found that trade openness affects the shape of the EKC. The results suggest that for most income groups, trade openness initially decreases the impact of GDP on CO_2 but increases the effects in the long run (Sharif, Uddin, & Alexiou, 2022). Mahmoodi and Dahmardeh (2022) evaluated the impact of renewable energy use and institutional quality on the ecological footprints of a country. The study showed that both renewable energy use and institutional quality mitigated ecological footprints (Mahmoodi and Dahmardeh, 2022). Wang, Wang, and Li (2022) showed by a threshold regression analysis that urbanization can increase environmental degradation from increases in economic growth.

Aytun and Akin (2022) showed that education might decrease environmental degradation. Specifically, they identified three channels through which education can impact environmental degradation: income, environmental awareness, and technological advances. They argue that while income may increase environmental degradation, technological advances and environmental awareness can contribute to a reduction in environmental degradation. These findings highlight the potential for education to play a critical role in promoting sustainable development practices and mitigating the negative impacts of economic growth (Aytun & Akin, 2022). Uddin (2014) showed empirical evidence from Bangladesh that education decreases emission production. It was further concluded that the effect of education on emissions emerges from the effect of an increased literacy rate on public environmental awareness (Uddin, 2014).

Geng and He (2021) showed that a rise in environmental awareness among the public in China led to an increase in demand for better environmental quality and, consequently, more eco-friendly policies.

Not many empirical studies examine the effect of environmental awareness on the EKC. However, Zheng et al. (2014) investigated the impact of public concern about environmental issues on the shape of the EKC across cities in China. The results showed that cities with higher environmental concerns had earlier turning points for EKCs with electricity and PM_{10} as outcomes than those with lower environmental concerns (Zheng et al., 2014). Chen, Huang, and Lin (2019) analyzed the effect of emissions on environmental awareness, the effect of environmental awareness on expenditures for environmental protection, and the effect of environmental awareness on the EKC for less developed countries using a panel data approach of 64 countries[5] between 2005-2015. They found that environmental awareness flattens the EKC and makes its turning point occur earlier. Simply meaning that environmental awareness leads to a more climate-smart development. Moreover, the authors provide a theoretical analysis showing why more aware countries should develop more sustainably.

This study will contribute to the present literature by evaluating the impact of environmental awareness on the shape of the EKC, including developed countries, and allowing for a cubic term, which was not done by Chen, Huang, and Lin (2019). Also, CO_2 emissions will be tested as an outcome for the analysis since it is a major cause of global warming. In addition to this, a longer time series will be used.

[5] The study has a total of 64 countries, however only 27 countries are less developed and thus a part of the analysis of the EKC.

4. Data

The present study applies panel data analysis to address the research question stated. Another option would have been to make several time-series analyses to explore whether cross-country differences emerge from differences in environmental awareness. Nevertheless, this would result in limited sample sizes, which could compromise the degrees of freedom and statistical power of the tests conducted. 62 countries will be used, and the sample of countries will be based on the one used by Chen, Huang, and Lin (2019) and is found in Appendix A. The selection of the observation period was based on data availability. The sample size with respect to the time series dimension will also differ depending on the outcome variable due to data availability issues. Specifically, employing PM_{10} and SO_2 as indicators for environmental degradation will consist of yearly observations between 2004 and 2015, while analyses with CO_2 as the outcome will have yearly observations between 2004 and 2020. The data begins in 2004 as prior data on environmental awareness is unavailable. When the panel is balanced[6], each country has 17 observations for all variables except PM_{10} and SO_2, which have 12 observations per country.

4.1 Variables

As Chen, Huang, and Lin (2019) suggested, Google trends data[7] is used to measure environmental awareness (AWA). Google Trends is a database that stores data on the search intensity of various categories and keywords. The index used to measure environmental awareness is based on the interest by country in searches related to "pollution"[8]. According to Google (2023), the interest-by-country index reflects the relative popularity of a topic across different countries. The number of searches for a category/keyword is measured relative to searches on other categories/keywords within each country and compared over a selected time frame. A category comprises searches for keywords related to a specific topic in any language

---

[6] This is done by imputing two GDP values for 2019 and 2020 for Yemen, by assuming the same growth factor as in the last 5 years.

[7] An alternative approach could have involved incorporating data from the World Values Survey, to measure environmental awareness. However, this option was not pursued due to several reasons. Firstly, a significant number of countries included in this study lacked available survey data. Additionally, the only relevant question in the survey focused on the trade-off between economic development and environmental quality, which not necessarily measure environmental awareness. However, an attempt to construct an index based on this question revealed a positive correlation of 0.33 with the awareness index employed in this thesis. Nonetheless, the analysis was limited to the 25 countries that overlapped in both datasets, and therefore, only these countries were included in the correlation calculations.

[8] Table A1 in Appendix A displays a list of countries that exhibit awareness levels below the median and those that exceed the median awareness within the sample.

(Google, 2023). The index ranges from 0 to 100 and is based on search data from 2004 to 2020. The variable is an average over the whole period and is thus not allowed to vary over time. A potential limitation of this study is that the index fails to account for the correlation between income and environmental awareness[9] over time.

Establishing a precise measure of environmental awareness for a given country is challenging. However, Google searches can serve as a reasonable proxy, as individuals who are more aware of environmental issues should be more likely to search for keywords related to the topic. The variable could also be interpreted as environmental concern. Nevertheless, it should be reasonable to conclude that concern emerges from awareness; thus, no further distinction between those will be made in this thesis. Since the index shows interest compared to other topics, differences in internet access between countries should not affect the results. However, a potential problem is that internet users in less developed countries should be concentrated on high-income takers. Therefore, the index might be biased and show the awareness/concern for just the high-income percentiles of the populations in less developed countries.

For the dependent variable, several types of emissions will be tested. That is because emissions differ in both source and consequence. Some types of emission have regional effects while others have global effects, this might affect the investigated relationship. Also, as explained in the introduction, PM_{10} and SO_2 directly affect human health more than CO_2. Analogously to Chen, Huang, and Lin (2019), SO_2 emissions per capita are used. Also, PM_{10} emissions are used, but in contrast to Chen Huang, and Lin (2019), PM_{10} is divided by population instead of GDP. This was mainly done to keep consistency throughout the analysis. This also simplifies the interpretation of results. Moreover, CO_2 emission per capita is also included as a proxy for environmental degradation, as has been done in several earlier EKC studies (Wang, Yang & Li, 2023; Allard et al., 2018; Sharif, Uddin, and Alexiou, 2022). $PM_{2.5}$ will not be included since it shows a high correlation[10] with PM_{10}. SO_2 and PM_{10} were retrieved from The Emission Database for Global Atmospheric Research (EDGAR). According to Crippa et al. (2018), SO_2 and PM_{10} estimations are done by combining activity data[11] and emission factors by country. The CO_2 emissions data was retrieved from Our World in Data (OWID).

[9] However, according to Table 3, cross-sectional correlations between GDP and awareness seem to be moderate and negative for the sample.
[10] 0.996. When analysis was made on this variable, similar results ass PM_{10} were obtained.
[11] Activity data refers to data on activities such as production that produce emissions of the particle/gas (Crippa et al, 2018).

The study incorporates several independent variables, namely GDP per capita (GDP), urban population (URB), institutional quality (INST), and varieties of democracy (VDEM). URB is included as a control variable based on the findings of Wang, Wang, and Li (2022), who suggest that urbanization may have an impact on the EKC. Furthermore, the inclusion of INST as a control variable is supported by the research of Mahmoodi and Dahmardeh (2022), who demonstrate its influence on the EKC. Moreover, considering the theoretical framework of the analysis, INST plays a crucial role in the political channel as it assumes the presence of effective policies. To control for the absence or disturbances in the political channel, VDEM is used as a control variable to address issues with democratic systems.

GDP is purchasing power parity adjusted to current (2023) US dollars and divided by population. URB measures the percentage of the population that lives in urban areas. GDP_{PC} and URB are both retrieved from the World Bank. INST is an index based on principal component analysis (PCA) of the Worldwide Governance Indicators (WGI) from the World Bank. The WGIs consist of measures of the following six dimensions: control of corruption, government effectiveness, political stability and absence of violence/terrorism, regulatory quality, the rule of law, and voice and accountability. The creation of this index is explained in Appendix B. VDEM is an index where countries are assigned a value of 0-1 for each year based on the current level of democracy in the country. According to Herre (2022) the index is based on survey evaluations by experts that can be either academics, civil society, or media. The index evaluates subjects such as rights to assemble and protest, electoral cheating, and citizens' voting rights. The data are summarized in Table 1 below.

Table 1. Summary of data with definitions, units, time periods, and sources.

Variable	Definition	Unit	Time period	Source
GDP	GDP per capita	Constant 2023 US$	2004-2020	OWID
CO_2	Carbon dioxides produce in a country, per capita. Production based.	kg/person	2004-2020	OWID
SO_2	Sulfur dioxide emitted in a country, per capita	mg/person	2004-2015	EDGAR
PM_{10}	Particles emitted with ⌀<10μm, per capita	mg/person	2004-2015	EDGAR
URB	Percentage of population living in urban areas	%	2004-2020	World bank
TRD	Percentage of exports plus imports of GDP	%	2004-2020	World bank
INST	Index created from the WGI.	-10-10	2004-2020	World Bank
VDEM	Index measuring level of democracy	0-1	2004-2020	OWID
AWA	Index created by google search data on the category "pollution"	0-100	2004-2020	Google Trends

4.3 Descriptive Statistics

Summary statistics of data used in the analysis are presented in Table 2.

Table 2. Summary statistics of variables used in the analysis.

Variable	Mean	Median	SD	Min	Max	N
GDP ($/capita)	26,989.65	26,717.23	23,785.14	198.38	120,647.80	1054
CO_2 (kg/capita)	6,687,082.00	6,128,262.00	4,612,739.00	52,529.10	28,000,000.00	1054
SO_2 (mg/capita)	16,900,000.00	10,000,000.00	18,000,000.00	301,727.00	121,000,000.00	744
PM_{10} (mg/capita)	6,758,933.00	4,969,383.00	6,829,759.00	499,359.30	52,600,000.00	744
URB (%)	66.47	68.09	18.30	22.50	100.00	1054
INST	0.00	0.12	2.13	-8.25	4.91	1054
VDEM	0.64	0.77	0.27	0.02	0.92	1054
AWA	13.79	9.00	14.48	3.00	100.00	1054

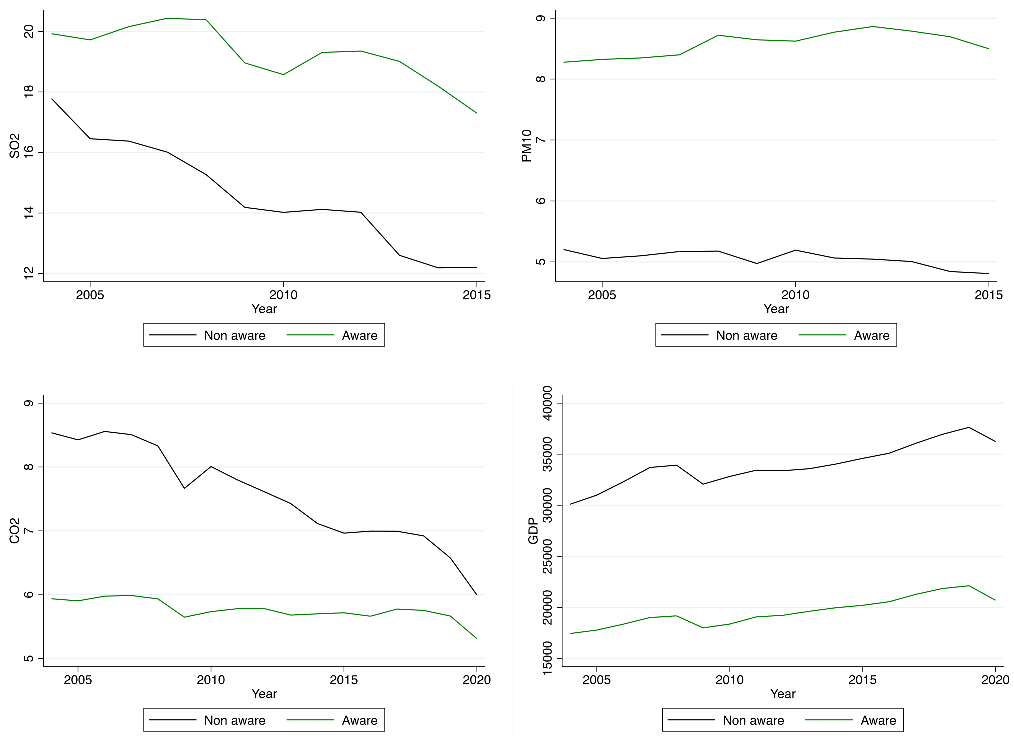
Note: Values are rounded to two decimals.

As seen in Table 2, the development level between countries in the sample differs considerably. The minimum and maximum GDP per capita values are $198 and $120,648, respectively. Thus, the observations to see a large part of the EKC should be present. Similarly, the CO_2 emissions per capita range from 52,529kg to 28,000,000kg. This indicates that there is significant variation in the levels of economic development, urbanization, pollution, and institutional quality across the countries included in the analysis. Another notable feature is that some variables' mean and median values differ substantially. For example, the median CO_2 emissions per capita are 6,128,262 kg, while the mean is 6,687,082 kg. This indicates that there may be some skewness or outliers in the distribution of this variable.

Additionally, the standard deviations for some variables are quite large relative to their means. For example, the standard deviation of PM_{10} per GDP is 4695 mg, which is substantially larger than its mean of 1534 mg/GDP. This suggests that there may be significant variation in this variable across countries. Overall, the descriptive statistics suggest that there is significant heterogeneity in the variables across the sample of countries, which could have important implications for the analysis. For example, accounting for unobserved heterogeneity can be of importance.

The cross-sectional averages for the main variables of interest over time are graphically depicted in Figure 5 below; these are divided by country groups, with the high awareness group as the green line and the low awareness group as the black line. Complete summary statistics for both country groups can be found in Appendix C.

Figure 5. Time-series graphs of cross-sectional averages for variables of interest.



Note: emission variables are converted into different measures, SO_2 is expressed in kg per capita and annum, PM_{10} is expressed in kg per capita and annum, and CO_2 is expressed in kt per capita and year. GDP is expressed in per capita PPP adjusted $.

Figure 5 shows that SO_2 and CO_2 emissions exhibited a negative trend over the observed period. Conversely, PM_{10} emissions did not display a prominent trend during the observed period. GDP demonstrated a positive trend throughout the analyzed timeframe. CO_2 and GDP were higher for the low awareness group for the entire period, while PM_{10} and SO_2 were lower for the low awareness group for the period.

Table 3 presents the Pearson correlation between all variables used in the dataset.

Table 3. Pearson correlation matrix

	GDP	CO_2	SO_2	PM_{10}	URB	INST	VDEM	AWA
GDP	1.00							
CO_2	0.65	1.00						
SO_2	-0.08	0.32	1.00					
PM_{10}	-0.09	-0.02	0.21	1.00				
URB	0.69	0.57	0.16	-0.12	1.00			
INST	0.26	0.13	0.19	0.26	0.21	1.00		
VDEM	0.57	0.21	-0.05	-0.01	0.47	0.38	1.00	
AWA	-0.40	-0.32	-0.05	0.28	-0.25	0.09	-0.29	1

Note: Values are rounded to two decimals.

Table 3 reveals a positive correlation between AWA and PM_{10}, suggesting that awareness increases with PM_{10} emission or vice versa. Conversely, there is no such correlation observed between AWA and SO_2 emissions. And there is a negative correlation between AWA and CO_2 emissions. The table also highlights that GDP positively correlates with CO_2 emissions and shows a weak negative correlation with SO_2 and PM_{10} emissions. As anticipated, URB and VDEM are strongly correlated with GDP.

5. Empirical models

The analysis is divided into two separate parts.

1. Assessing the effect of environmental awareness on total emissions.
2. Assessing the effect of environmental awareness on the form of the EKC.

The first part of the analysis gives important but simplified insights into the relationship. There are mainly two reasons for using the EKC to study the relationship between emissions and environmental awareness. Firstly, this kind of analysis allows controlling for unobserved heterogeneity. This cannot be done when only the awareness dummy is used since a FE model would absorb the dummy variable. Secondly, using the EKC to study the relationship gives a more realistic picture of the complex relationship, making it possible to see that the relationship might differ across GDP levels.

5.1 Awareness and aggregated emissions

To give a first indication of the effect of environmental awareness on emission levels, awareness will be used as an explanatory variable by the following estimation equation:

$$Y_{it} = \gamma_0 + \sum_{k=1}^{3} \gamma_k GDP_{it}^{k} + \delta_t + \gamma_4 D_i + \boldsymbol{\gamma Z_{it}} + u_{it} \tag{4}$$

Where,

$$\begin{cases} D_i = 0, M_{AWA_i} > AWA_i \\ D_i = 1, M_{AWA_i} \leq AWA_i \end{cases} \tag{5}$$

Where, the environmental degradation variable Y_{it} (which varies between CO_2, SO_2, and PM_{10}), is a function of, a third-degree polynomial of GDP_{it}, δ_t that is time-fixed effects, and $\boldsymbol{Z_{it}}$ that is a vector of (URB_{it}, $INST_{it}$ $VDEM_{it}$). And M_{AWA_i} is the median of AWA_i in the sample. The hypothesis will be that countries with high AWA_i produce less emissions:

$$H_0: \gamma_4 > 0, H_a: \gamma_4 \leq 0 \tag{6}$$

5.2 Awareness and the EKC

To evaluate potential differences in the form of EKC between country groups based on environmental awareness, a pooled OLS will be used to estimate the following predetermined threshold polynomial regression:

$$Y_{it} = \begin{cases} \beta_0 + \sum_{k=1}^{3} \beta_k GDP_{it}^k + \delta_t + \boldsymbol{\beta Z_{it}} + u_{it}, \ D_i = 0 \\ \lambda_0 + \sum_{k=1}^{3} \lambda_k GDP_{it}^k + \delta_t + \boldsymbol{\lambda Z_{it}} + u_{it}, \ D_i = 1 \end{cases} \tag{7}$$

To reduce endogeneity issues, (11) will be estimated with a fixed effects (FE) estimator as well. To do this, the within estimator will be applied by estimating country-demeaned versions of the variables.

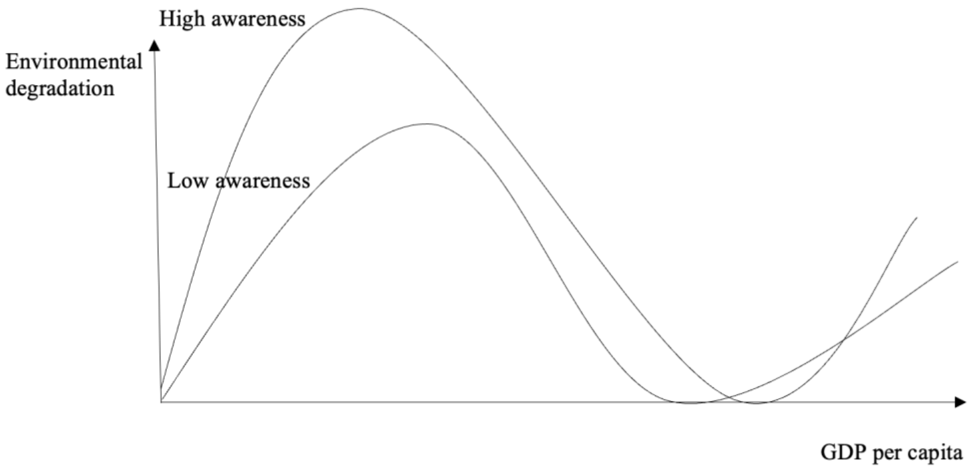
If coefficients for x_{it}^k are found to be significant, the analysis will proceed by evaluating if these differences are statistically significant between the country groups. This will be done by evaluating the following estimation equation:

$$Y_{it} = \alpha_0 + \sum_{k=1}^{3} \alpha_k GDP_{it}^k + \sum_{k=4}^{6} GDP_k x_{it}^{k-3} D_i + \alpha_7 D_i + \delta_t + \boldsymbol{\beta Z_{it}} + u_{it} \tag{8}$$

Where, the joint significance of α_4, α_5, α_6 will be interpreted as that the coefficients are statistically significant between the country groups. When this test is used with FE, $\alpha_7 D_i$ will get dropped since it is a time-invariant dummy variable. So, since the FE estimation does not account for potential shifts in the functions between the group, only the significance of the results should be interpreted and not the magnitude of those. If differences are present, those functions will be illustrated graphically to check whether high awareness tends to be associated with lower emissions.

The analysis in earlier studies regarding EKC and environmental awareness has been based on turning points of the EKC function (Zheng et al., 2014; Chen, Huang, & Lin, 2019). However, this kind of analysis misses important parts of the relationship. Because even if turning points occur in a more climate-smart way, the expected value of aggregated emissions given the GDP level might suggest different results. Consequently, analysis based only on turning points might be misleading. An example of this, in the context of this thesis, would be:

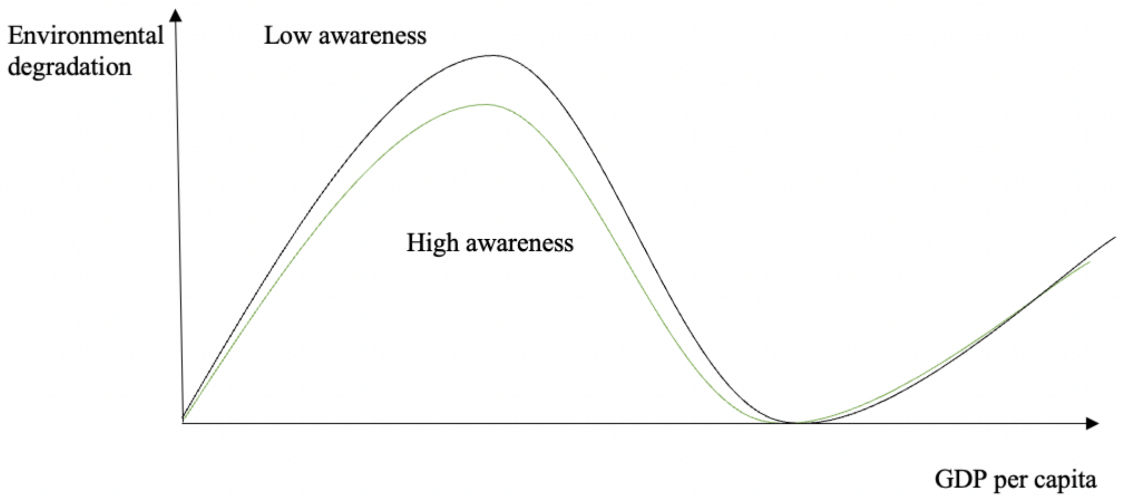
Figure 6. Graphical presentation of issues with turning point analysis of EKC



Source: Own

Figure 6 demonstrates that, despite an earlier occurrence of the first turning point and a later occurrence of the second turning point in the high awareness curve, the trajectory of the low awareness group exhibits lower emissions. Specifically, the function of the low awareness group consistently remains below that of the high awareness group across almost all levels of GDP. Consequently, the hypothesis for this thesis will instead target predicted emissions given the level of GDP. So, with this approach, the hypothesis is that the area below the line, or the integral of the predicted function, for the low-awareness group, should be higher relative to the high-awareness group. If both curves are N-shaped, a graphical representation of the hypothesis could be:

Figure 7. Graphical representation of hypothesis for N-shaped EKC.



Source: Own

Info: The figure only provides an example; the main message should be that the area below the green line should be smaller than the corresponding area below the black line for the hypothesis to be rejected.

The lines presented in the results are thus the expected value for emissions for each specific value of GDP for the country group. These predictions are estimated from (11).

5.3 Data and model testing

To test for stationarity, panel unit root tests will be applied. Panel unit root tests are used to gain additional testing power. Since macroeconomic variables on a country level during times of high globalization are used, cross-sectional dependence is likely to be present. A test that allows for such property is the cross-sectionally augmented Dickey-Fuller (CADF) test by Pesaran (2007). The test assumes that the cross-sectional dependence emerges from a common factor in the error terms, such as:

$$u_{it} = \gamma_i f_t + \varepsilon_{it} \tag{9}$$

Where, f_t is the unobservable common factor, and $\varepsilon_{it}$is an individual specific error. Further, the test assumes that the common factor can be proxied by the cross-sectional arithmetic means ($\bar{y}_t = N^{-1} \sum_{j=1}^{N} y_{jt}$). The test thus evaluates the following auxiliary regression:

$$\Delta y_{it} = \alpha_i + \pi_i y_{i,t-1} + c_{1i}\bar{y}_{t-1} + d_i \Delta \bar{y}_t + u_{it} \tag{10}$$

Where,

$$H_0: \forall i. \pi_i = 0,\ H_1: \exists\ i. \pi_i \neq 0 \tag{11}$$

This is tested by either averaging the individual t-statistics for π_i for each panel or taking each panel separately (Pesaran, 2007). For this analysis, the average t-statistics for π_i will be used to test for stationarity.

The test results for the CADF test can be found in Appendix D. Since rejecting the null hypothesis was not possible, applying a dynamic estimator might have been optimal. However, since T is small, the issues with spurious regression should not be too severe. Also, unit root tests have low power when T is small, which might be a reason for not being able to reject null hypotheses.

Heteroscedasticity and autocorrelation robust standard errors will be used in the analysis, so these properties will not be tested.

6. Results

The results will be presented in two sections, the first section presents the regression with awareness as an independent variable, and the second is presenting the results from the threshold regressions. The second section will also use graphical representations expected value of emissions given the level of GDP to make results easier to interpret. Only statistically significant parameters will be used to illustrate the results graphically.

6.1 Awareness and aggregated emissions

A pooled OLS model was applied according to (4) and (5), and the results are presented in Table 4 below.

Table 4. Effect of awareness on environmental degradation.

Variables	CO_2	SO_2	PM_{10}
GDP	464.5***	1,360***	488.3***
	(33.64)	(277.7)	(83.96)
GDP^2	-0.00861***	-0.0381***	-0.0119***
	(0.000826)	(0.00617)	(0.00177)
GDP^3	5.27e-08***	2.22e-07***	7.01e-08***
	(5.12e-09)	(3.56e-08)	(9.88e-09)
AWA	-264,878	2.517e+06**	2.673e+06***
	(217,609)	(1.193e+06)	(328,057)
INST	283,691***	3.137e+06***	1.251e+06***
	(50,683)	(349,555)	(207,257)
URB	53,735***	299,310***	-86,266***
	(6,251)	(42,664)	(24,798)
VDEM	-7.831e+06***	-2.478e+07***	-4.040e+06**
	(875,683)	(4.999e+06)	(1.787e+06)
Constant	4.034e+06***	1.005e+07***	1.061e+07***
	(615,408)	(3.455e+06)	(2.192e+06)
N	0.572	0.220	0.180
R^2	1,054	744	744
Countries	62	62	62

Note: Heteroskedasticity and autocorrelation-consistent standard errors in parentheses. Year dummies are omitted from the output. Output rounded to two decimals. *** p<0.01, ** p<0.05, * p<0.1.

According to the findings presented in Table 4, AWA has a statistically significant positive effect on PM_{10} and SO_2 emissions, whereas no significant impact was observed for CO_2[12]. So,

[12] When applying a one-sided hypothesis test, the p-value was 0.11. So the null hypothesis cannot be rejected.

the null hypothesis cannot be rejected for any awareness coefficent. Specifically, the results indicate that countries categorized under the high awareness group experience an average of 2,516,413 mg more SO_2 emissions per capita and annum as compared to those in the low awareness group. Moreover, the high awareness group exhibits a higher level of SO_2 emissions per GDP by 1,144,172.25 µg as compared to the low awareness group. This is counterintuitive and can have several explanations. For example, it is possible that the results should be interpreted as countries with higher emissions levels have higher awareness, and not vice versa. The results can also emerge from issues such as unobserved heterogeneity. This will be discussed in greater detail in section 7.

6.2 Awareness and the EKC

Table 5 presents results from the pooled OLS threshold polynomial regressions.

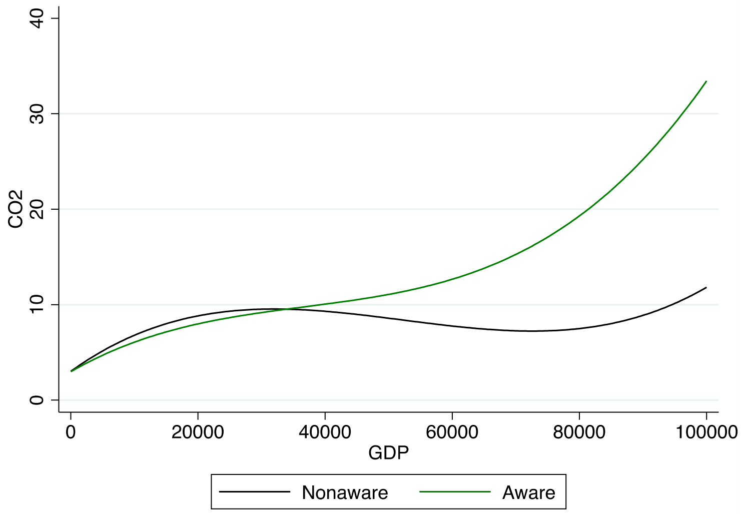
Table 5. Regression output of pooled OLS.

Variables	CO_2		SO_2		PM_{10}	
Awareness	Low	High	Low	High	Low	High
GDP	479.1***	381.6***	-788.8***	2,134***	316.5***	110.0
	(41.52)	(67.64)	(266.4)	(518.4)	(58.71)	(151.6)
GDP^2	-0.0108***	-0.00802***	0.000570	-0.0629***	-0.00911***	0.000173
	(0.000931)	(0.00234)	(0.00487)	(0.0148)	(0.00137)	(0.00438)
GDP^3	6.89e-08***	7.25e-08***	2.44e-08	4.19e-07***	5.69e-08***	-9.97e-09
	(5.43e-09)	(1.98e-08)	(2.79e-08)	(1.01e-07)	(7.90e-09)	(3.08e-08)
URB	43,975***	50,138***	339,732***	403,919***	388,101***	1.895e+06***
	(5,733)	(10,276)	(44,016)	(64,540)	(104,491)	(348,867)
INST	384,286***	347,866***	732,982	4.153e+06***	4.716e+06***	-5.074e+06**
	(66,704)	(67,987)	(603,513)	(373,426)	(970,233)	(2.164e+06)
VDEM	-2.787e+06***	-7.010e+06***	3.793e+07***	-4.065e+07***	-15,938*	-133,362***
	(890,741)	(1.061e+06)	(6.403e+06)	(7.043e+06)	(9,453)	(47,558)
Constant	3.010e+06***	2.979e+06***	-1.196e+07***	1.044e+07**	1.841e+06**	1.724e+07***
	(640,951)	(834,460)	(3.632e+06)	(4.629e+06)	(809,084)	(3.598e+06)
N	544	510	384	360	384	360
R^2	0.68	0.55	0.23	0.26	0.32	0.18
Countries	32	30	32	30	32	30

Note: Heteroskedasticity and autocorrelation-consistent standard errors in parentheses. Year dummies are omitted from the output. *** p<0.01, ** p<0.05, * p<0.1.

For the function with significant differences according to the test introduced in 5.2, the results from Table 5 will be presented graphically and discussed with a focus on the graph. This is done by predicting emission levels throughout all GDP levels that are in the sample. The difference test was done on all variables that showed significant coefficients for some GDP coefficient for both country groups, and full results from those are found in Appendix E. The green lines will represent the country group with high awareness, and the low awareness group will be represented by a black line. Results from the differences test showed that coefficients for GDP and GDP^2 were significant between the country groups, and a graphical representation of CO_2 results will thus be provided.

Figure 8. Visualization of regression results for CO_2.



Info: predicted relationship from resulting function in Table 5.

Figure 8 presents empirical evidence indicating that CO_2 emissions for countries with high awareness can be seen as a monotonically increasing function of GDP, while for countries with low awareness, it should be seen as an N-shaped function of GDP. Moreover, at higher levels of GDP, countries with awareness levels surpassing the median tend to exhibit greater emissions compared to their counterparts with lower awareness levels. Conversely, for lower GDP levels, higher levels of awareness may be associated with reduced emissions. However, as the observed differences in emissions levels are small and thus probably insignificant, no certain conclusions can be drawn regarding low-income levels. This result is counterintuitive and does not support the hypothesis of the investigation. This can be a result of several factors, such as unobserved heterogeneity. However, this will be discussed in greater in the discussion part of the thesis.

For SO_2, the null hypothesis of differences could not be rejected by the joint F-test. However, results from regarding this function did not provide results that can be easily interpreted as predicted emission levels are negative; for this reason, no graphical illustration has been provided for SO_2. However, results from Table 5 suggest that countries with a higher level of awareness have more emissions for all levels of GDP. It also suggests that the country group with a high awareness tends to have an N-shaped EKC, while GDP for the country group with low awareness has a negative but diminishing effect on SO_2. So, the results from Table 5 suggest that the null hypothesis should be rejected for SO_2 as well. These results are counterintuitive and can result from, e.g., a lack of data. A further discussion of this will be presented in section 7. Graphical representation of PM_{10} results will not be presented as the high awareness group does not have any significant results, and the difference between the country groups can, for that reason, not be compared. Overall, as seen in Table 3, PM_{10} and GDP seem to be weakly correlated, and the existence of a relationship between the variables is not obvious.

URB seems to increase all types of emissions for both country groups. INST seems to increase CO_2 emissions for both country groups and SO_2 for the high-awareness group. VDEM decreases CO_2 emissions for both country groups, increases SO_2 emissions for the low awareness group, and increases SO_2 emissions for the high awareness group. Furthermore, VDEM might have a negative effect on PM_{10} as well, especially for the high-awareness group. For all emission types, VDEM has more negative parameters for the high awareness group, this might advocate the importance of the political channel explained in 3.3.1 since democracy is crucial for the political channel to work.

For the next step of the analysis, (11) was estimated with a FE estimator. The results from these models are presented in Table 6 below.

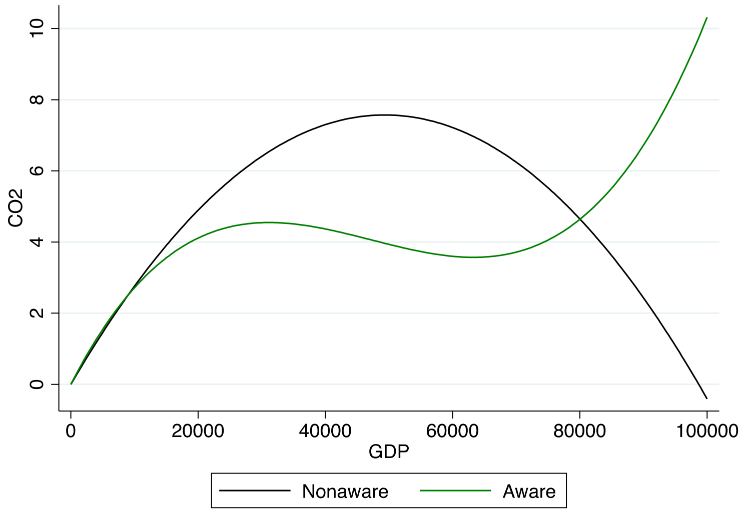
Table 6. Results for fixed effects threshold polynomial regressions.

Variables	CO_2		SO_2		PM_{10}	
Awareness	Low	High	Low	High	Low	High
GDP	306.9***	349.2***	262.2	692.1	152.9*	424.1**
	(78.38)	(91.58)	(543.3)	(427.5)	(81.77)	(168.0)
GDP^2	-0.00311***	-0.00836***	0.00146	-0.0168	-0.00154	-0.00725**
	(0.00113)	(0.00238)	(0.00935)	(0.0142)	(0.00113)	(0.00319)
GDP^3	6,81e-09	5,90e-08***	-1.41e-08	1.10e-07	4.99e-09	3.92e-08**
	(6.95e-09)	(1.71e-08)	(4.35e-08)	(1.04e-07)	(5.02e-09)	(1.88e-08)
URB	103,329	67,590	837,224**	97,763	-136,534	106,623
	(89,321)	(43,220)	(322,961)	(165,910)	(193,793)	(173,699)
INST	333,129**	355,387**	-171,236	-92,441	640,684	1.541e+06
	(154,047)	(171,292)	(748,548)	(739,889)	(3.015e+06)	(1.727e+06)
VDEM	-1.594e+06	531,855	1.239e+07*	6.497e+06*	-2,093	-65,541
	(1.828e+06)	(725,146)	(6.306e+06)	(3.440e+06)	(73,710)	(84,810)
Constant	-2.608e+06	-718,742	-5.765e+07**	6.468e+06	2.077e+06	7.747e+06
	(7.133e+06)	(2.651e+06)	(2.797e+07)	(1.056e+07)	(5.673e+06)	(4.968e+06)
N	544	510	384	360	384	360
R^2	0.499	0.432	0.351	0.146	0.541	0.322
Countries	32	30	32	30	32	30

Note: Heteroskedasticity and autocorrelation-consistent standard errors in parentheses. Year dummies are omitted from the output. Output rounded to two decimals. *** p<0.01, ** p<0.05, * p<0.1.

The results from Table 6 show that when using the within estimator, only EKC with CO_2 as a proxy for environmental degradation has significant coefficients for both country groups. Thus, it is the only one presented in graphical form. This is illustrated in Figure 9, where CO_2 is expressed in kt/person instead of kg/person to make interpretation of results more convenient.

Figure 9. Visualization of FE estimated regression results for CO_2.



Info: predicted relationship from resulting function in Table 6. Note that CO_2 is expressed in kt/person in this table.

The empirical results illustrated in Figure 9 reveal that there appears to be no discernible disparity in emission levels between country groups characterized by different levels of awareness at low-income levels. However, at medium and high-income levels, countries exhibiting higher awareness levels tend to display lower CO_2 emissions. At very high-income levels, high levels of awareness may be associated with greater emissions. Furthermore, it is noteworthy that the form of the relationship between emissions and GDP levels differs between the two groups, with the high awareness group displaying an N-shaped EKC, while the low awareness group exhibits an inverted U-shaped EKC. Since $E[Y_L|x] \geq E[Y_H|x]$ for medium and high-income countries, results suggest that the null hypothesis should not be rejected for those. However, for very high-income levels $E[Y_L|x] \leq E[Y_H|x]$ and the hypothesis should thus be rejected for those levels. The surprising result for the very high-income levels can be a result of, e.g., a low number of observations for these levels of GDP; this will be further discussed in section 7. It is possible to add a shift in the functions by summing the country-specific intercepts for each country group to get a more accurate result. However, this value is 9.73 for countries in the low awareness group and 11.10 for the high awareness group, and this will thus not change results significantly.

When using PM_{10} as the outcome variable, no comparable results were produced as the GDP variables do not have significant parameters for the low awareness group. Additionally, when controlling for unobserved heterogeneity, the parameters for GDP lose their significance for the SO_2 function. This suggests that the parameters from Table 6 when SO_2 was used as an outcome may result from cross-country heterogeneity. This is not surprising given the correlations provided in Table 3.

Almost all coefficients for the control variables lost significance when controlling for unobserved heterogeneity. This can be interpreted as that cross-sectional heterogeneity might have been driving the significant results in Table 5. It is also possible to assume that the within-variation of the dependent variables is too small during the observed period leading to non-significant parameters when controlling for cross-sectional differences.

Results are robust to the exclusion of Afghanistan and Yemen[13]. It is also robust to the exclusion of URB as a control variable.[14]

[13] Since civil war is ongoing in those countries, the relevance of data from those might be questionable, however, as stated in the text, exclusion of those does not change results significantly.
[14] Robustness regarding Urbanization was tested due to its high correlation with GDP and emissions.

7. Discussion

Overall, results are ambiguous depending on the estimation method. However, as countries differ substantially in income levels, culture, geography, etc., accounting for unobserved heterogeneity should be of great importance; thus, the discussion will mainly focus on the results obtained with the FE estimator. The results from section 6.1 suggest that awareness has no significant average effect on CO_2 emissions. Results from the pooled OLS regression in 6.2 produces results that suggest that for high-income levels, countries in the low-awareness group have a lower production of emissions compared to the high-awareness group. However, as was mentioned earlier, these results might be biased due to the likely issue of unobserved heterogeneity. Conclusions from this should thus be drawn with caution. For the FE estimations, the high awareness group exhibits lower levels of emissions for medium and high-income levels. These results suggest that the peak effect should be at income levels per capita of $54,975[15] since this is the point where the difference between the two lines in Figure 9 is maximized.

At income levels per capita exceeding $80 023[16], the data suggests that the CO_2 emissions may be higher for countries in the high awareness group compared to those in the low awareness group. However, caution must be exercised when interpreting these findings due to the limited number of observations with such high income. If these findings are to be accepted, a plausible explanation for the discrepancy could be attributed to the varying stages of technological progress in green technology between the awareness groups. As previously explained in section 2.3, progress in green technology should be the primary driver of the relationship between environmental awareness and the EKC. Hence, countries with high awareness should possess a more advanced (green) technological infrastructure than their low awareness counterparts and hence would experience the diminishing returns of green technology earlier. According to Torras and Boyce (1998), the reason for the second turning point in the EKC is the diminishing returns of technological advancements. It is, therefore, reasonable to assume that the high awareness group experiences the second turning point at a lower income level than the low awareness group because they suffer from diminishing returns on technological advances at lower income levels. So, it is possible that the low awareness group has not yet reached its turning point and that this is the reason for the non-significant cubic term in Table 6.

---

[15] Calculated by solving for x in: $max\{306.9x - 0.00311x^2 - 349.2x + 0.00836x^2 - 0.000000059x^3\}$
[16] Calculated by solving for x in: $306.9x - 0.00311x^2 = 349.2x - 0.00836x^2 + 0.000000059x^3$

Regarding SO_2 and PM_{10}, the results are not robust to the estimation method. Specifically, the first part of the analysis shows that aggregated emissions were higher in the high-awareness group, considering both SO_2 and PM_{10}. A weakness of this analysis is that it implicitly assumes a unidirectional relationship, positing that awareness leads to reduced pollution. In an empirical context, this assumption may not hold true, as it is possible that poor environmental performance triggers increased awareness and concern among citizens This phenomenon could be more pronounced for SO_2 and PM_{10}, which has regional effects that have more direct effects on human health, while CO_2 has global effects that are not as directly connected to human health. Thereby this might be an explanation for the theoretical inconsistency of the results presented in the first part of the analysis. The reversed effect might interfere with the relationship for CO_2 as well. However, as the local effects from CO_2 are smaller and less connected to human health, the effect of CO_2 emissions on awareness might be weaker in relation to the effect of SO_2 and PM_{10} emissions on awareness. This might be an explanation for the differences in results between the outcome variables. Still, it should be noted that the first part of the analysis is not possible to test with the FE estimator since the variable of interest is a dummy variable. So, there is no controlling for unobserved heterogeneity and because issues with unobserved heterogeneity are likely to be present, results should be interpreted with caution.

For the second part of the analysis, no robust results were found when using SO_2 and PM_{10} as dependent variables. The high-awareness group had higher SO_2 emissions than the low-awareness group when estimated by pooled OLS, but these differences were not statistically significant. Also, the significance of the coefficients did disappear when using the FE estimator. So, results for SO_2 in Table 5 might have been driven by unobserved heterogeneity between countries. The absence of significant results for SO_2 and PM_{10} for the EKC analysis can be due to several reasons, such as data quality, omitted variable bias etc. Also, the findings presented in Table 3 indicate a weak correlation between GDP and SO_2 as well as PM_{10}. These results may imply that the variables should not be regarded as a function of GDP. However, this is somewhat contradictory to the findings of Chen, Huang, and Lin (2019), who suggested that environmental awareness might contribute to sustainable development regarding SO_2 and PM_{10} emissions. Since the method and results of Chen, Huang, and Lin (2019) are incomplete, the origin of the dissimilarities in results is not possible to determine. Some potential explanations can be differences in control variables, differences in time periods, and differences in polynomial length[17]. The inclusion of high-income countries might also be an important disparity. Also, this thesis applies a different tool for analysis as it does not use turning points to determine the level of climate-smart development but instead uses the expected value of emissions given GDP for each GDP level in the sample.

The logic behind the unconventional analysis method was explained in subsection 5.2. The main idea is that the expected value of emissions given GDP should be the main interest when analyzing, for example, the effects of country characteristics on the EKC. That is because the turning points only have implications about the marginal effect and do therefore not imply anything regarding total expected emissions for each GDP level which should be the core issue.

The theoretical inconsistency of the results regarding aggregated emissions can, in addition to unobserved heterogeneity, possibly reflect the fact that poor environmental performance raises the awareness and concern of citizens. This phenomenon could be more pronounced in the case of SO_2 and PM_{10}, which has regional effects that have more direct effects on human health, while CO_2 has global effects that are not as directly connected to human health. Since the model implicitly assumes that the
However, it should be noted that the results are not robust to changes in estimation methods, and because issues with unobserved heterogeneity are likely to be present, results should be interpreted with caution.

The results obtained from the unit root test indicate that the null hypothesis, suggesting the absence of unit roots, cannot be rejected. For this reason, the results can be due to spurious regression since the OLS estimator and within estimator are not robust to non-stationary variables. However, as N>T and T are small, unit roots should not be able to create too large problems. Also, since T is small, tests have lower power, and this can be a possible explanation as to why the unit root was not able to be rejected.

Another drawback of the study is that it does not control for inequality due to data availability issues. Since Google searches might be concentrated toward high-income percentiles in very unequal countries due to the lack of availability in low-income percentiles, it is possible to argue that these countries should follow a different pattern than countries with high equality. Also,

[17] Chen, Huang, and Lin (2019), uses a log specification of the model, this was tested in this thesis as well but did not change the results and should thus not be the origin of the differences.

the low-income percentile for these countries is located on a different level on the EKC compared to the high-income percentile. Also, since polynomial regression is used, the models should suffer from high structural multicollinearity. High multicollinearity usually leads to large standard errors in the parameter estimates. Consequently, it is plausible that the lack of statistical significance observed for the coefficient of GDP^3 for the low awareness group in the FE regression analysis of CO_2 emissions may stem from the underlying issue of multicollinearity.

8. Conclusions

This study has investigated the relationship between environmental awareness and the EKC. To achieve this, Google searches categorized as "pollution" were utilized as a proxy for measuring environmental awareness. The data obtained from these searches were analyzed to determine if there are significant differences in the average CO_2, SO_2, and PM_{10} emissions levels among countries with an awareness value above and below the median. The research question was further examined by using threshold polynomial regressions, with the awareness index as a threshold variable and the median as a predetermined threshold. The findings suggest that environmental awareness may reduce CO_2 emissions, particularly for middle and high-income levels, but increase emissions for very high-income levels. However, the results also indicate that countries with a high environmental awareness might exhibit larger SO_2 and PM_{10} emissions compared to countries with a lower environmental awareness; these results are, however, not robust to the estimation method and should be interpreted with caution. A possible explanation for this unexpected outcome is that high levels of awareness may stem from poor environmental performance. So, to answer the research question, the study's empirical findings suggest that environmental awareness might lead to a more climate-smart development path regarding CO_2, at least for some parts of the development path. However, the study finds no support for the hypothesis regarding EKC with emissions of SO_2 and PM_{10}.

As has been argued throughout the text, the discrepancy in the presentation of results between this study and earlier similar studies is motivated by the belief that earlier studies missed important parts of the relationship. Since the distance between the predicted functions should be the main interest and not just the form of each specific function. This critique should be considered one of the main contributions of this thesis since it guides how similar future studies should present their results. Nevertheless, the results of this study suggest that environmental awareness may play a role in successfully tackling the challenges related to global warming. For instance, some of the sample countries with an awareness level below the median are large emitters and will soon reach the GDP level where the awareness effect is maximized. An example of this is Germany, which had a per capita GDP of $51,423 in 2020; promoting awareness in such countries might have effects on global CO_2 emissions in the upcoming years. As such, promoting environmental awareness in such countries may be an effective tool for mitigating the effects of global warming and achieving global targets related to climate change. The results can also have implications regarding predictions of future CO_2 emissions. Assuming all other factors remain constant, nations with higher levels of environmental awareness might exhibit lower CO_2 emissions than similar countries with lower levels of awareness for medium and high-income levels. However, the analysis in this thesis does not focus on the predictive power of the model, and this can be assessed by future research.

A more comprehensive study could be conducted for future research by including more countries. Since GDP seems to have a low correlation with SO_2 and PM_{10}, future research might focus only on CO_2 since it has a stronger relationship with GDP relative to the other emission types. Also, the analysis method used in this paper could be extended by employing bootstrapping or similar methods to get a more formal statistical test. Also, recommendations for future studies are to use analysis techniques that focus on the expected values of emissions rather than just the form of the function.

Further investigations should be conducted to enhance the comprehensiveness of the study, particularly by incorporating additional countries to give more robust results. Since GDP seems to have a low correlation with SO_2 and PM_{10}, future research might concentrate solely on CO_2 emissions. Moreover, it is recommended to address potential concerns related to non-stationarity by employing dynamic estimators to assess the relationship. Additionally, future studies are encouraged to utilize analytical techniques that prioritize the analysis of expected emission values rather than merely focusing on the functional form of the relationship. These suggested avenues for future research should contribute to a better understanding of about the relationship between environmental awareness and the EKC.

References

Aas, W. et al. (2019). "Global and regional trends of atmospheric sulfur". *Scientific Reports*, 9(1), p. 953. Available at: https://doi.org/10.1038/s41598-018-37304-0.

Ahmed, A., Uddin, G.S. & Sohag, K. (2016) "Biomass energy, technological progress, and the environmental Kuznets curve: Evidence from selected European countries". *Biomass and Bioenergy*, 90, pp. 202–208. Available at: https://doi.org/10.1016/j.biombioe.2016.04.004.

Allard, A., Takman, J., G.S, Uddin., & Ali, A. (2018) "The N-shaped environmental Kuznets curve: an empirical evaluation using a panel quantile regression approach". *Environmental Science and Pollution Research*, 25(6), pp. 5848–5861. Available at: https://doi.org/10.1007/s11356-017-0907-0.

American Lung Association (2022). What is Sulfur Dioxide? Available at: https://www.lung.org/clean-air/outdoors/what-makes-air-unhealthy/sulfur-dioxide (Accessed: 15 March 2023).

Aswathanarayana, U., Harikrishnan, T. and Kadher-Mohien, T. (2010) *Green Energy.* 1st ed. CRC Press.

Aytun, C. and Akin, C.S. (2022) "Can education lower the environmental degradation? Bootstrap panel Granger causality analysis for emerging countries". *Environment, Development and Sustainability*, 24(9), pp. 10666–10694. Available at: https://doi.org/10.1007/s10668-021-01877-0.

California Air Resource Board (n.a). *Inhalable Particulate Matter and Health (PM2.5 and PM10).* Available at: https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health (Accessed: 15 March 2023).

Chen, X., Huang, B. & Lin, C.-T. (2019) "Environmental awareness and environmental Kuznets curve". *Economic Modelling*, 77, pp. 2–11. Available at: https://doi.org/10.1016/j.econmod.2019.02.003.

Crippa, M. et al. (2018). Gridded emissions of air pollutants for the period 1970–2012 within EDGAR v4. 3.2. Earth Syst. Sci. Data, 10(4), pp.1987-2013.

Culas, R.J. (2012) "REDD and forest transition: Tunneling through the environmental Kuznets curve". *Ecological Economics*, 79, pp. 44–51. Available at: https://doi.org/10.1016/j.ecolecon.2012.04.015.

Dasgupta, S., Laplante, B., Wang, H., & Wheeler, D. (2002). "Confronting the Environmental Kuznets Curve". *Journal of Economic Perspectives*, 16(1), pp. 147–168. Available at: https://doi.org/10.1257/0895330027157.

Geng, M.-M. & He, L.-Y. (2021) "Environmental Regulation, Environmental Awareness and Environmental Governance Satisfaction". *Sustainability*, 13(7), p. 3960. Available at: https://doi.org/10.3390/su13073960.

Grossman, G. M., & Krueger, A. B. (1991). Environmental Impacts of a North American Free Trade Agreement. *Working Paper No. 3914*.

Herre, B. (2022). *The 'Varieties of Democracy' data: how do researchers measure democracy?* Our World in Data. (n.d.). Available at: https://ourworldindata.org/vdem-electoral-democracy-data.

Hindriks, J. & Myles, G.D. (2013) *Intermediate public economics*. 2nd ed. Cambridge, MA: MIT Press.

Leitão, A. (2010) "Corruption and the environmental Kuznets Curve: Empirical evidence for sulfur". *Ecological Economics*, 69(11), pp. 2191–2201. Available at: https://doi.org/10.1016/j.ecolecon.2010.06.004.

Mahmoodi, M. & Dahmardeh, N. (2022) "Environmental Kuznets Curve Hypothesis With Considering Ecological Footprint and Governance Quality: Evidence From Emerging Countries". *Frontiers in Environmental Science*, 10, p. 849676. Available at: https://doi.org/10.3389/fenvs.2022.849676.

National grid (n.a). *What are greenhouse gases?* Available at: https://www.nationalgrid.com/stories/energy-explained/what-are-greenhouse-gases (Accessed: 15 March 2023).

Naturvårdsverket (n.a). *Koldioxid (CO2).* Available at: https://utslappisiffror.naturvardsverket.se/sv/Amnen/Vaxthusgaser/Koldioxid/ (Accessed: 15 March 2023).

Pesaran, M.H. (2007) "A simple panel unit root test in the presence of cross-section dependence". *Journal of Applied Econometrics*, 22(2), pp. 265–312. Available at: https://doi.org/10.1002/jae.951.

Our World in Data (2021). *What is economic growth? And why is it so important?* Available at: https://ourworldindata.org/what-is-economic-growth (Accessed: 3 March 2023).

Sharif, T., Uddin, M.M.M. and Alexiou, C. (2022) "Testing the moderating role of trade openness on the environmental Kuznets curve hypothesis: a novel approach". *Annals of Operations Research* [Preprint]. Available at: https://doi.org/10.1007/s10479-021-04501-6.

Torras, M. & Boyce, J.K. (1998) "Income, inequality, and pollution: a reassessment of the environmental Kuznets Curve". *Ecological Economics*, 25(2), pp. 147–160. Available at: https://doi.org/10.1016/S0921-8009(97)00177-8.

Uddin, M. M. (2014). Causal relationship between education, carbon dioxide (CO_2) emission and economic growth in Bangladesh. *IOSR Journal of Humanities and Social Science, 19*(4), 60–67. https://doi.org/10.9790/0837-19486067

United Nations. (2022). *Promote inclusive and sustainable economic growth, employment and decent work for all.* Available at: https://www.un.org/sustainabledevelopment/economic-growth/ (Accessed: 15 March 2023).

United Nations Framework Convention on Climate Change (UNFCCC) (n.d). *The Paris Agreement.* Available at: https://unfccc.int/process-and-meetings/the-paris-agreement (Accessed: 6 March 2023).

Världsnaturfonden (n.d). *Klimatförändringar - Vad är det?* Available at: https://www.wwf.se/klimat/klimatforandringar/ (Accessed: 3 March 2023).

Wang, Q., Yang, T. & Li, R. (2023) "Does income inequality reshape the environmental Kuznets curve (EKC) hypothesis? A nonlinear panel data analysis". *Environmental Research*, 216, p. 114575. Available at: https://doi.org/10.1016/j.envres.2022.114575.

Wang, Q., Wang, X. and Li, R. (2022) "Does urbanization redefine the environmental Kuznets curve? An empirical analysis of 134 Countries". *Sustainable Cities and Society*, 76, p. 103382. Available at: https://doi.org/10.1016/j.scs.2021.103382.

World Bank (2022). *Poverty.* https://www.worldbank.org/en/topic/poverty/overview (Accessed: 3 March 2023).

Zheng, S., Khan, M, E., Sun, W., & Danglun, L. (2014). Incentives for China's urban mayors to mitigate pollution externalities: The role of the central government and public environmentalism. *Regional Science and Urban Economics*, 47, pp. 61–71. Available at: https://doi.org/10.1016/j.regsciurbeco.2013.09.003.

Appendix A

Table A1. List of countries divided by awareness.

Low Awareness	High awareness
Austria	Afghanistan
Azerbaijan	Albania
Belarus	Australia
Belgium	Bhutan
Bulgaria	Bolivia
Cyprus	China
Czech Republic	Egypt
Denmark	El Salvador
Estonia	France
Finland	Hong Kong
Georgia	Iceland
Germany	Indonesia
Greece	Iran
Hungary	Italy
Iran, Islamic Republic of	Kazakhstan
Ireland	Latvia
Israel	Lithuania
Japan	Maldives
Kyrgyzstan	Malta
Luxembourg	Mauritius
Netherlands	Moldova, Republic of
Norway	Myanmar
Poland	New Zealand
Russian Federation	Portugal
Slovakia	Romania
Slovenia	Serbia
Sweden	South Africa
Switzerland	Spain
Turkey	Thailand
Ukraine	United Arab Emirates
United Kingdom	
Uzbekistan	
Yemen	

Appendix B

Institutional quality lacks standard measurement, and the WGI are often used as proxies. However, the WGIs comprise six distinct variables, each with relevant information. To address concerns about overparameterization, a principal components analysis was performed to reduce the dimensionality of the WGIs. First, the data values were standardized by:

$$Z_{it} = \frac{X_{it} - \bar{X}}{\sigma_X} \tag{12}$$

The variance-covariance matrix of the standardized values is:

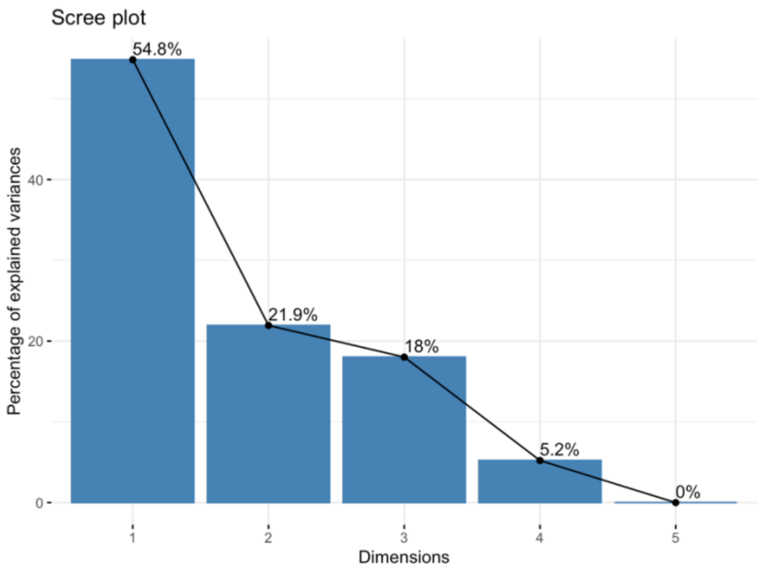
$$COV(\boldsymbol{Z}) = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{16} \\ \vdots & \ddots & \vdots \\ \sigma_{61} & \cdots & \sigma_{66} \end{pmatrix} \tag{13}$$

Eigenvectors and their corresponding eigenvalues were then calculated by:

$$COV(\boldsymbol{Z}) \times \boldsymbol{\theta} = \boldsymbol{\theta} \times \lambda \tag{14}$$

Where, $\boldsymbol{\theta}$ is a 6×1 vector of unknowns calculated by maximizing the sum of squared distances from the projected points to the origin called a principal component. $\boldsymbol{\theta}$ should, when multiplied by the unknown scalar λ be equal to the covariance matrix multiplied by $\boldsymbol{\theta}$. There will be as many principal components as variables; each is calculated by maximizing the sum of squared distances but with the restriction of being perpendicular to the lower-order principal components.

Figure B1. Variance explained by principal components.



As shown in Figure 5, about 55% of the variance is explained by the first principal component (PC1). Since much of the variation is explained by PC1, the index will then rely on this and will thus be created by:

$$\boldsymbol{INST} = \boldsymbol{\theta}_1^T \times \boldsymbol{Z}^T \tag{15}$$

So, INST is created by multiplying the transpose of the eigenvector from the covariance matrix by the transpose of the standardized WGI data, yielding a one-dimensional index representing institutional quality.

Appendix C

Table C1. Summary statistics for the low awareness group

Variable	Mean	Median	SD	Min	Max	N
GDP ($/capita)	33,938.25	33910.24	24954.36	433.235	120647.8	544
CO_2 (kg/capita)	7,554,467	7572381	3786059	340175.4	2.60e+07	544
SO_2 (mg/capita)	14,600,000	9,465,769	13,600,000	867,085.7	83,800,000	384
PM_{10} (mg/capita)	5,052,741	4,523,228	3,188,142	499,359.3	15,300,000	384
URB (%)	70.71	73.46	15.59	28.39	98.08	544
INST	-0.13	-0.19	2.25	-8.25	3.28	544
VDEM	0.71	0.85	0.25	0.11	0.92	544
AWA	5.19	5	1.86	2	8	544

Note: Values are rounded to two decimals.

Table C2. Summary statistics for the high awareness group

Variable	Mean	Median	SD	Min	Max	N
GDP ($/capita)	19,577.81	8076.975	19985.09	198.3769	97757.85	510
CO_2 (kg/capita)	5,761,871	4,604,503	5,202,335	52,529.1	28,000,000	510
SO_2 (mg/capita)	19,300,000	10,700,000	21,500,000	301,727.8	120,100,000	360
PM_{10} (mg/capita)	8,578,871	5,930,019	8,902,839	665,266.1	52,600,000	360
URB (%)	61.94	62.15	19.84	22.5	100	510
INST	0.14	0.22	1.98	-3.92	4.91	510
VDEM	0.58	0.63	0.27	0.02	0.9	510
AWA	18.77	15	11.43	9	49	510

Note: Values are rounded to two decimals.

Appendix D

Table D1 presents the results from the unit root tests performed on the variables in the dataset. The lag length was varied from 1 to 3 for the CADF test to account for potential autocorrelation problems; this did not change the results. The results from the test are thus presented with a lag length of 1.

Table D1. Test results from the CADF test

Variable	$\overline{T}$-statistic
GDP	-2.14
CO_2	-2.05
SO_2	-1.83
PM_{10}	-1.64
URB	-2.05
INST	-2.00
VDEM	-2.53***

Info: time trend is assumed. Lag length = 1. *** p<0.01, ** p<0.05, * p<0.1.

When using a CADF test, table C1 suggests that only VDEM can be concluded to be stationary.

Appendix E

Table E1. Results from tests for differences

Variables	FE CO_2	OLS CO_2	OLS SO_2
GDP	252.4***	562.2***	1,364***
	(64.68)	(36.12)	(236.3)
GDP^2	-0.00293**	-0.0120***	-0.0397***
	(0.00112)	(0.000793)	(0.00472)
GDP^3	6.90e-09	7.41e-08***	2.32e-07***
	(7.39e-09)	(4.69e-09)	(2.71e-08)
GDP* AWA	226.6**	-213.5***	226.5
	(111.3)	(73.25)	(397.4)
GDP^2* AWA	-0.00600**	0.00488*	-0.00905
	(0.00275)	(0.00260)	(0.0134)
GDP^3* AWA	4.87e-08***	-6.27e-09	1.10e-07
	(1.83e-08)	(2.20e-08)	(9.87e-08)
INST	327,779***	405,710***	3.305e+06***
	(112,554)	(49,844)	(358,698)
VDEM	112,035	-5.925e+06***	-2.225e+07***
	(725,701)	(846,626)	(5.916e+06)
URB	108,867***	43,587***	288,554***
	(38,378)	(5,709)	(38,023)
D_i	-	137,812	556,934
	-	(321,453)	(2.420e+06)
Constant	-3.690e+06	3.421e+06***	9.981e+06***
	(2.781e+06)	(561,924)	(3.339e+06)
N	1,054	1,054	744
R^2	0.426	0.617	0.229
Countries	62	62	62
P-value for Joint F-test	0.00	0.01	0.01

Note: Heteroskedasticity and autocorrelation-consistent standard errors in parentheses. Year dummies are omitted from the output. *** p<0.01, ** p<0.05, * p<0.1. The joint F-test regards the hypothesis that all interaction parameters are zero.