Testing the empirical validity of the Kuznets environmental curve on 18 countries from Asia and Oceania using panel regression methods

Gabriel AMMOUR* and Noa LE ROUX*

*2^{nd} year students in Applied Econometrics, Nantes Université

December 30, 2024

Abstract

This study investigates the relationship between economic growth and environmental degradation in 18 Asia-Pacific (APAC) countries, aiming to test the validity of the Environmental Kuznets Curve (EKC) hypothesis and identify potential turning points for environmental improvement. The study employs panel regression methods, utilizing data from the World Bank's open data portal for the period 1970-2015, to analyze the impact of GDP per capita, energy consumption, and trade openness on CO_2 emissions per capita. The analysis reveals a cubic or 'N-shaped' relationship between GDP per capita and CO_2 emissions per capita, suggesting that environmental degradation initially increases with economic growth, then declines, and subsequently rises again at higher income levels. This research contributes to a better understanding of the EKC in the Asia-Pacific region by highlighting the non-linear relationship between economic growth and environmental quality as well as the limits of using per capita CO2 emissions and GDP as the only indicator of environmental and economic health.

Keywords: Environmental Kuznets Curve (EKC) ; Gross Domestic Product (GDP) ; CO2 emissions ; Asia-Pacific (APAC) ; Panel data

1 Introduction

The relationship between economic growth and environmental degradation has been the subject of intense debate for several decades. The Environmental Kuznets Curve (EKC) hypothesis, developed by Grossman and Krueger [1](1991), takes Kuznets'[2](1955) original theoretical model of income inequality and applies it to environmental issues. This hypothesis postulates an inverted 'U-shaped' relationship between environmental pollution and per capita income. It suggests that environmental degradation worsens in the early stages of a country's economic development, but that once a certain level of average income is reached, this trend reverses and environmental degradation begins to slow considerably. According to the EKC, economic growth, rather than being a threat to the environment, is in fact the means to long-term environmental improvement.

The Asia-Pacific (APAC) region represents a highly pertinent context for examining the validity of the Environmental Kuznets Curve (EKC) hypothesis. As the world's largest energy-consuming region and the leading contributor to greenhouse gas emissions, APAC plays a critical role in global environmental dynamics. Over recent decades, the region has witnessed rapid economic growth, accompanied by a significant rise in energy demand and CO_2 emissions. This makes makes an interesting case study for analyzing the relationship between economic development and environmental quality. Furthermore, the APAC region is characterized by a remarkable diversity of countries at various stages of economic development, ranging from low-income nations like Bangladesh to high-income economies such as Australia and Singapore. This heterogeneity provides a good opportunity to assess the applicability and variations of the EKC hypothesis across different economic contexts and levels of development.

To examine the relationship between economic growth and environmental degradation in the APAC region, we'll used the following variables :

- **GDP per capita (*also squared and cubed*):** Measured in constant 2010 US$, GDP per capita serves as the primary indicator of economic growth and living standards. This variable allows for comparisons across countries with varying population sizes. However, GDP per capita, as a measure of economic well-being, has several limitations. It fails to account income distribution, social well-being, and the environmental costs associated with economic activities. The EKC hypothesis predicts a non-linear relationship between CO_2 emissions and GDP per capita, suggesting that emissions initially rise with economic growth but eventually decline after a certain income threshold is reached. Using normal, squared and cubed GDP per capita, we models the non-linear relationship between pollution and per capita income, capturing the initial effect of economic growth on environmental degradation, the turning point and the nuances beyond the inverted 'U-shape' [3].

- **CO_2 emissions per capita** : CO_2 emissions is a widely used indicator of environmental pressure, particularly in relation to climate change [4]. We chose to keep emissions/capita as this measure enables useful comparisons to be made between countries with varying populations, as using total emissions would favor less populous countries. By focusing on per capita emissions, we can isolate the impact of economic activity on the environment independently of demographic factors [3] [5]. Furthermore, it seems more logical since we'll used GDP per capita. CO_2 emissions per capita, while widely used in environmental studies, have notable limitations. As a national average, this metric hide significant regional and socioeconomic disparities within countries, such as differences between urban industrial centers and rural areas. Additionally, it fails to account for historical responsibility, allowing developed nations with significant cumulative emissions to appear less impactful today. To address these gaps, a multidimensional approach is essential. This includes disaggregating emissions data at subnational levels, using complementary indicators like emissions intensity, and incorporating historical contributions to global CO_2 levels.

- **Energy consumption (kg of oil equivalent per capita) :** This variable is measured in terms of energy consumption, expressed in kilograms of oil equivalent per capita. It is included in the Environmental Kuznets Curve (EKC) model to correct for any omitted variable bias. Its inclusion is based on previous work that has highlighted the importance of taking energy consumption into account when analyzing causal relationships between energy consumption, CO_2 emissions and income [6] [7] [5] [8]. It serves as a crucial explanatory variable, given its significant influence on CO_2 emissions, especially in countries heavily reliant on fossil fuels. Higher energy consumption is generally associated with increased CO_2 emissions, emphasizing the importance of transitioning towards cleaner energy sources to mitigate emissions. However, this measure does not distinguish between renewable and non-renewable energy consumption, limiting insights into the potential of energy transitions to decouple economic growth from CO_2 emissions.

- **Trade (% of GDP) :** This variable measures the sum of exports and imports as a percentage of GDP, and is a fundamental indicator for assessing the impact of international trade on emissions. Its inclusion in economic models is justified by several factors. Firstly, trade openness modifies economic structure by encouraging countries to specialize according to their comparative advantages, which can increase pollution in economies centered on polluting industries [3]. Secondly, international trade leads to carbon leakage, as the emissions associated with imported goods are often attributed to exporting countries rather than consumers [9]. Finally, including this variable corrects for an omitted variable bias, as accounting for trade and energy consumption is crucial for analyzing the relationship between emissions and income [10]. Additionnaly, the "pollution haven" hypothesis [11] suggests that countries with weaker environmental regulations may attract polluting industries, potentially leading to higher emissions. However, it is essential to recognize that this indicator does not differentiate between imports and exports. This distinction is crucial, as the impact of trade on emissions is influenced by a country's trade balance. Countries with substantial trade surpluses, like China, may experience increased emissions due to export-oriented production, while countries with trade deficits may observe a more complex interaction between trade and emissions.

Using these variables, the study aims to determine whether economic growth in the APAC region follows an EKC pattern, and if so, to identify the turning point at which environmental improvements begin to occur.

2 Data

Variables mentioned in this study were taken from the World Bank's open data portal. To guarantee a high degree of reproducibility, most variables were extracted directly from *Stata* using the function `wbopendata`. To choose the region's countries, we were inspired by a paper that also discussed the subject of the EKC in this region [12]. However, we added 4 countries from the same region for which data was available from the World Bank. Regarding the years of observation, our analysis begins in 1970 to match the availability of our variable of interest (CO_2 Emissions).

Table I shows the main descriptive statistics for our variables. These statistics include number of observations (Obs), number of missing values (NA), minimum and maximum values, median, mean, as well as measures of distribution shape (skewness and kurtosis).

Table I: **Summary Statistics**

Variable	Obs	NA	Min	Median	Mean	Max	Skewness	Kurtosis
GDP per capita	947	25	69.45	2716.49	10643.96	88428.70	1.82	5.88
CO_2 emissions	954	18	3.05	136.60	715.44	15662.13	5.34	34.83
Energy use	784	188	80.91	971.00	1955.73	9697.02	1.29	3.87
Trade openness	920	52	4.83	55.80	90.50	442.62	2.05	6.59

All the variables studied reveals highly skewed distributions. The positive asymmetry coefficients, particularly pronounced for CO_2 emissions (5.34) and trade openness (2.05), point to the presence of high extreme values, probably characteristic of the most developed or industrialized economies. This observation is reinforced by the substantial gap between means and medians, particularly for GDP per capita (mean 10,643.96 vs. median 2,716.49). High kurtoses, notably for CO_2 emissions (34.83) and trade openness (6.59), confirm leptokurtic distributions with a significant concentration of extreme values, suggesting strong heterogeneity between the countries in the sample.

Table II: **Correlation Matrix**

Variable	GDP/capita†	CO_2 emissions/capita	Energy use	Trade
GDP/capita	1			
CO_2 emissions/capita	0.64	1		
Energy use/capita	0.79	0.91	1	
Trade	0.27	0.01	0.19	1

† It should be noted that correlation between GDP/capita, GDP^2/capita and GDP^3/capita are all above 0.8

The correlation matrix (Table II) highlights the interconnections between the presented variables. A strong positive correlation ($r = 0.79$) is observed between GDP and energy use, suggesting a strong interdependence between economic production and energy consumption. This result corroborates the hypothesis that economic growth is accompanied by an intensification of energy consumption, in line with endogenous growth theories.

The relationship between CO_2 emissions per capita and energy use shows the highest correlation ($r = 0.91$), reflecting the prevalence of carbon-based energy sources in the global energy mix. On the other hand, the correlation between GDP and CO_2 emissions ($r = 0.64$) raises questions about the possibility of a decoupling between economic growth and emissions.

A particularly interesting result concerns trade openness, which shows remarkably low correlations with all other variables. The modest correlation with GDP ($r = 0.27$) might suggest that the benefits of international trade do not necessarily translate into a proportional increase in GDP. More surprisingly, the near-zero correlation with CO_2 emissions ($r = 0.01$) might contradict theories like 'pollution havens'. However, we are unable to affirm this since we do not know the import and export share contained in the variable.

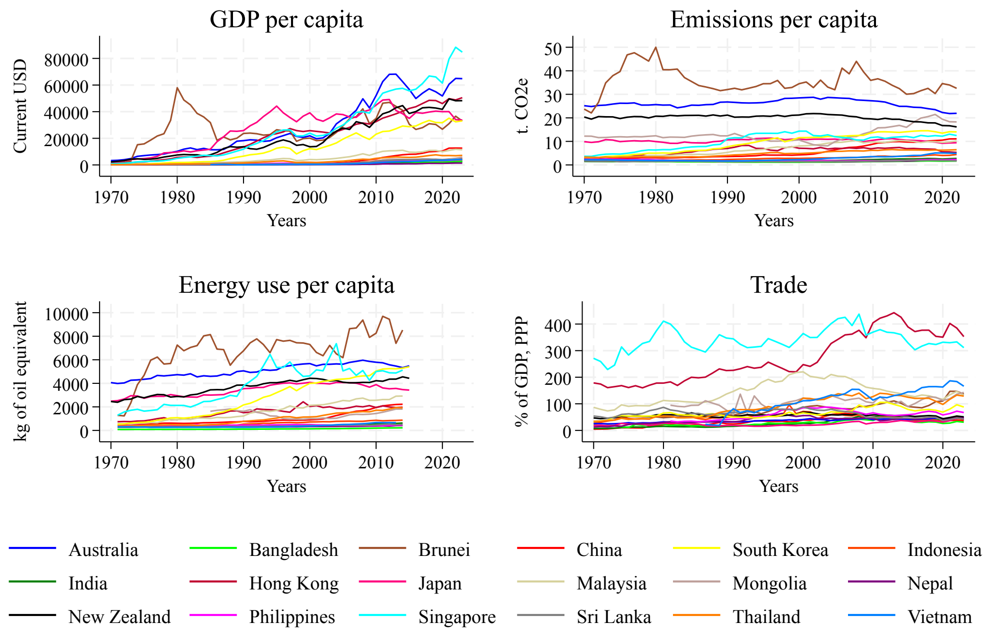

Figure 1: Series over time (1970-2023)
Source : *World Bank Open Data*

Figure 1 shows the evolution of our variables over the period 1970-2023 for the APAC selected countries. In terms of GDP per capita, there is a general upward trend, but with marked disparities. Singapore and Australia stand out with significantly higher levels, exceeding \$60,000 per capita. Hong Kong and New Zealand are above \$40,000. Japan, already industrialized in the post-war period, enjoyed sustained growth until the mid-1990s, after which its growth slowed significantly. At the same time, the *Asian tigers* represented in the chart - Singapore, Hong Kong and South Korea - illustrate their spectacular rise, particularly during the period 1970-1990. Singapore stands out with remarkable growth, reaching the highest levels of GDP per capita in the region. South Korea also demonstrates an impressive economic transformation, moving from a low-income economy in the 1970s to an advanced economy in just a few decades. China, despite its position as the world's second economic power and spectacular growth since 1970, maintains a relatively modest GDP per capita in 2020, mainly due to its huge population and strong regional disparities between developed coastal areas and more rural regions.

When it comes to emissions, Brunei stands out as the main emitter, with levels ranging between 30 and 50 tonnes of CO2 per inhabitant. This position might be explained by its heavy dependence on hydrocarbons, both in its energy mix and its economy centered on oil extraction. Australia and New Zealand form a second group, with relatively stable emissions of between 15 and 20 tons per capita. These high levels probably reflect a Western development model, marked by high levels of motorization and energy consumption. China's emissions have been rising steadily, particularly since 2000, reflecting its rapid industrialization and major economic transformation. Similarly, South Korea and Singapore are following upward trajectories, reflecting their transition to developed economies. Finally, a significant gap remains between the least developed countries, such as Bangladesh and Nepal, and the advanced economies. Bangladesh's emissions remain below 5 tons per capita, a disparity that illustrates the fundamental differences in economic structure, living standards and access to energy, even within the same geographical region.

Figure 1 reveals the importance of international trade in the region, with several economies highly internationalized. While Singapore dominates with a ratio in excess of 300% of GDP, a second group of countries also boasts remarkably high levels : Australia, Malaysia and Vietnam maintain ratios of between 150% and 400% of GDP, reflecting their strong integration into global value chains. Malaysia in particular stands out for an increase since 1990 and then a decline since the mid-2000s, while Vietnam has shown a spectacular increase since the 1990s, as a result of its policy of economic openness. Australia, although fluctuating, consistently maintains high levels, illustrating its major role in regional trade, thanks in particular to its exports of raw materials.

This comparative analysis highlights the different development trajectories in the APAC region, with a general trend towards economic growth accompanied by growing environmental pressures, particularly marked in the major emerging economies such as China and India.

3 Methods [13, 14]

3.1 Simple Linear Model

Panel data combines cross-sectional and temporal dimensions, allowing us to analyse individual behaviour over time and to exploit both temporal and individual variations. This structure allows for more robust statistical analysis especially when capturing unobserved heterogeneity. A simple linear model for panel data can be specified as follows :

$$Y_{i,t} = X_{i,t}\beta + \epsilon_{i,t} \tag{1}$$

where $Y_{i,t}$ is the dependent variable (CO2 emissions per capita), $X_{i,t}$ is the matrix of explanatory variables, β is the vector of coefficients, and $\epsilon_{i,t}$ represents the error term that can be decomposed into two components :

$$\epsilon_{i,t} = u_i + \nu_{i,t} \tag{2}$$

where u_i captures unobserved, time-invariant individual heterogeneity (e.g., cultural or geographic factors specific to each country), and $\nu_{i,t}$ represents idiosyncratic errors that may vary across both time and individuals. While the ordinary least squares (OLS) estimator :

$$\beta = (X'X)^{-1}X'Y \tag{3}$$

is unbiased and consistent under the assumptions of exogeneity ($\mathrm{Cov}(X_{i,t}, \epsilon_{i,t}) = 0$) and no omitted variable bias, these assumptions are often violated in panel data. Especially, if u_i is correlated with $X_{i,t}$, OLS estimates will be biased and inconsistent. Panel data methods address these issues by explicitly modeling or eliminating u_i, providing consistent estimates of β even in the presence of unobserved heterogeneity. The main techniques for dealing with this heterogeneity are the compound error model (or random effects model) and the fixed effects model. These methods differ in their assumptions about the relationship between u_i and $X_{i,t}$, as discussed below.

3.2 The compound error model

The compound error model introduced by Balestra and Nerlove [15] provides a general framework for analysing panel data, while taking account of unobserved heterogeneity. It decomposes the error term $\epsilon_{i,t}$ into two components:

$$\epsilon_{i,t} = u_i + \nu_{i,t} \tag{4}$$

where u_i is the unobserved individual effect, assumed to be uncorrelated with $X_{i,t}$, and $\nu_{i,t}$ is the idiosyncratic error. This decomposition allows for more flexible modelling of individual differences between entities (e.g. countries). The compound error model can be estimated using generalized least squares (GLS), which corrects for the bias and inefficiency of OLS by weighting within-individual and between-individual variations. Two additional estimators can be used : *between* and *within* estimators.

3.2.1 The Between Estimator

The *between* estimator focuses on inter-individual variation by averaging observations over time for each individual and regressing the averaged $Y_{i,t}$ on averaged $X_{i,t}$:

$$\beta_B^* = (X'BX)^{-1}(X'BY) \tag{5}$$

where B is the matrix of averages over individuals. This approach eliminates time-specific variation and isolates long-term differences between individuals, but it ignores within-individual variation and may lead to omitted variable bias if u_i is correlated with $X_{i,t}$.

3.2.2 The Within Estimator

The *within* estimator completes the analysis, it eliminates u_i by centering all variables around their individual means :

$$\beta^*_w = (X^{'}WX)^{-1}(X^{'}WY) \tag{6}$$

where W represents the transformation matrix for demeaning variables. This approach captures within-individual variation over time, ensuring consistent estimation even if u_i is correlated with $X_{i,t}$. GLS combines these two estimators to produce an estimates of β, weighting observations according to the relative variances of u_i and $\nu_{i,t}$.

3.3 Random effects model

The random effects (RE) model is a specific application of the compound error model, with the key assumption that the individual effect u_i is a random variable uncorrelated with the explanatory variables $X_{i,t}$:

$$\text{Cov}(u_i, X_{i,t}) = 0 \tag{7}$$

This assumption allows u_i to be treated as a part of the error term, preserving both within-individual and between-individual variation in the estimation. The random effects model specification is written as :

$$Y_{i,t} = X_{i,t}\beta + u_i + \nu_{i,t} \tag{8}$$

Under these conditions, GLS produces efficient and consistent estimates of β by including both inter-individual and intra-individual information. However, if the orthogonality assumption is violated, the random effects estimator becomes biased and inconsistent. An advantage of the RE model is that it allows the estimation of time-invariant variables, which cannot be directly estimated in the fixed effects framework.

3.4 Fixed effects model

The fixed effects (FE) model addresses the potential correlation issue between u_i and $X_{i,t}$, a common source of bias in panel data. FE model assumes u_i is fixed for each individual and potentially correlated with $X_{i,t}$, making it more robust than RE in such scenarios. The model is written as follows :

$$Y_{i,t} = X_{i,t}\beta + u_i + \nu_{i,t} \tag{9}$$

To estimate β, the fixed effects approach eliminates u_i by applying the within transformation, which centers variables around their individual means :

$$\tilde{Y}_{i,t} = \tilde{X}_{i,t}\beta + \tilde{\nu}_{i,t}, \tag{10}$$

where

$$\tilde{Y}_{i,t} = Y_{i,t} - \bar{Y}_i, \quad \tilde{X}_{i,t} = X_{i,t} - \bar{X}_i, \quad \tilde{\nu}_{i,t} = \nu_{i,t} - \bar{\nu}_i \tag{11}$$

This transformation removes all time-invariant individual effects, ensuring that the resulting estimates of β are consistent even if u_i is correlated with $X_{i,t}$. However, a drawback of the FE model is that it cannot estimate the effects of time-invariant variables.

3.5 Hausman test

The Hausman test [16] determines whether the RE model's orthogonality assumption holds. It compares the FE estimator ($\hat{\beta}_{\mathrm{FE}}$), which is consistent regardless of the orthogonality assumption, with the RE estimator ($\hat{\beta}_{\mathrm{RE}}$), which is efficient but inconsistent if the assumption is violated. The test statistic is computed as :

$$H = (\hat{\beta}_{\mathrm{FE}} - \hat{\beta}_{\mathrm{RE}})' \left[\mathrm{Var}(\hat{\beta}_{\mathrm{FE}}) - \mathrm{Var}(\hat{\beta}_{\mathrm{RE}})\right]^{-1} (\hat{\beta}_{\mathrm{FE}} - \hat{\beta}_{\mathrm{RE}}), \tag{12}$$

where H follows a χ^2 distribution under the null hypothesis. If the test rejects the null hypothesis (typically at a 5% significance level), the FE model is preferred, as it suggests that u_i is correlated with $X_{i,t}$. Otherwise, the RE model is chosen.

3.6 Imputation method for missing values

To deal with the missing values, we used a Predictive Mean Matching[1] (PMM) multiple imputation method. This technique, which is adapted to continuous data, estimates missing values based on predictive regression, then matches them with the most similar existing observations. Imputation was performed for each variable, taking into account the time dimension and individual country specificities. The robustness of the results was enhanced by adding an imputation iteration for each variable, with ($k = 1$) neighbours selected.

4 Results

Initially, we estimate models using the raw data, with imputation-based estimates reserved for later robustness checks. Although multiple imputation can mitigate data gaps, it may introduce bias, and given the relatively large sample size, we focus on raw data initially. That said, the sample used for estimation is composed of 746 observations, from 1970 to 2015.

Table III: Comparison of Fixed Effects and Random Effects Models

Variable	**Fixed Effects (FE)**	**Random Effects (RE)**
GDP/capita	0.00039***	0.00037***
GDP^2/capita	-1.35e-08***	-1.32e-08***
GDP^3/capita	1.31e-13***	1.29e-13***
Energy Use/capita	0.0011***	0.0013***
Trade (% of GDP)	0.0021	0.00066
Constant	5.15***	5.08***
R-squared (Within)	0.48	0.48
R-squared (Between)	0.81	0.84
R-squared (Overall)	0.74	0.77
Corr($u_i, X_{i,t}$)	0.72	0
Hausman Test	$\chi^2(3) = 65.57, \mathrm{Prob} > \chi^2 = 0.0000$	

Note : Hausman test compares FE and RE models : the null hypothesis of no systematic difference is rejected.
Significance levels : *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We begin by comparing the fixed-effects model with the random-effects model (Table III). Both models present very similar results. With the exception of the Trade (% of GDP), all the coefficients are significant at the 1% level. As for signs, we note that GDP/Capita and GDP^3/capita and energy use/per capita are positive when GDP^2/capita is negative. We will discuss these implications below. The R^2 values suggest that both models explain a significant proportion of the variation in CO_2 emissions. Results of the Hausman test presented earlier (*Subsection* 3.5) are displayed in the last row of Table III. The p-value of the test is equal to 0, so we reject the hypothesis H0 which means that the FE model is preferred. This was expected given the correlation model's variables, highlighted by the high correlation (0.72) estimated by FE between individual effects u_i

[1] `mi impute pmm-`, *Stata*

and the explanatory variables $X_{i,t}$.

To account for potential heteroskedasticity, we employ robust variance estimators in the FE models. These estimates were made using robust variance estimators[2]. Additionally, time dummies are included in one specification to capture common temporal shocks or trends across countries. This approach allows us to control for unobserved time-fixed effects, ensuring more reliable coefficient estimates. It would appear that the inclusion of a certain heteroskedasticity has reduced the significance of the coefficients (Table IV). None of the explanatory variables is now significant at the 1% level. The three GDP variables are significant at the 5% level for both models. For the energy consumption, the coefficient is significant at the 5% level for the model with time dummies but only at the 10 % level for the model without. As for the trade, as before, it was not significant for either model. The value and sign of the coefficients and the Corr($u_i, X_{i,t}$) are the same for all three models. Finally, the R^2 of the model without time dummies is slightly higher, around 2 points of percentage.

Table IV: Comparison of Fixed Effects Models with and without Time Dummies

Variable	**FE robust**	**FE robust with Time Dummies**
GDP/capita	0.00039**	0.00055**
GDP^2/capita	-1.35e-08**	-1.81e-08**
GDP^{3}capita	1.31e-13**	1.71e-13**
Energy Use/capita	0.0011*	0.0011**
Trade (% of GDP)	0.0021	0.0045
Constant	5.15***	7.58***
R-squared (Within)	0.48	0.53
R-squared (Between)	0.81	0.76
R-squared (Overall)	0.74	0.72
Corr($u_i, X_{i,t}$)	0.72	0.72

Significance levels : *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

As mentionned, the three FE models presented above are very similar. For the rest of this work, we have decided to retain the FE model robust with time dummies. This choice is motivated by the fact that this model is by nature probably more robust than the classic fixed-effects model, since it takes into account a certain heteroscedasticity. In addition, the model with time dummies is preferred because it captures time fixed effects, such as shocks or trends common to different countries. Moreover, this specification improves the fit of the model, as shown by the increase in the R^2 within from 0.48 to 0.53, and offers better precision of the coefficients since the energy consumption is significant at the 5% threshold.

Theoretical interpretations of the coefficients associated with the explanatory variables could be formulated as follows. The positive coefficient of GDP/capita reflects that economic growth initially increases CO_2 emissions, consistent with the energy-intensive development phase of the EKC. Let's take for instance a country with a standard of living of \$20,000/capita, an increase of \$1,000/capita would have the effect of increasing CO_2 emissions/capita by 0.0312 (Appendix 7.1). The negative coefficient of GDP^2/capita suggests a turning point where economic growth begins to decouple from emissions, while the positive GDP^3/capita coefficient indicates a significant re-intensification of emissions at high-income levels. Energy consumption has a significant effect, showing that higher energy consumption directly increases emissions. All things being equal, an increase of 1,000 units of energy consumption per capita (equal to 1 t of oil equivalent) would lead to an increase in CO_2 emissions of 1.1 t. We note that the sign of the coefficients - associated with energy consumption and GDP/per capita - are consistent with the available literature

Figure 2 suggests an 'N-shaped' (cubic shape) EKC characterised by three distinct phases, identified by inflection points located at \$22,127 and \$48,357 of GDP per capita (Appendix 7.2). The curve is consistent with the signs of the coefficients obtained earlier, i.e. $\beta_1 > 0$, $\beta_1 < 0$, $\beta_3 > 0$. These points mark the transitions between a phase of growth in emissions (Phase 1), a phase of decline (Phase 2) and a new phase of growth (Phase 3). Table V analyses these phases in detail,

[2] `vce(robust)`, Stata

Phase 1 (78.40% of observations) represents the initial growth stage, characterized by a 28.6% increase in CO_2 emissions. Phase 2 (15.12%) reflects a transitional decline (-12.2%), consistent with the downward slope of the EKC, while Phase 3 (6.48%) highlights a recovery in emissions (+108.8%) at high-income levels, suggesting renewed environmental pressures in advanced economies. These phases emphasize the non-linear trajectory of economic development and its environmental trade-offs.

Figure 2: Kuznets Environmental Curve

Note : The vertical lines represent the turning points at $22,127 and $48,357

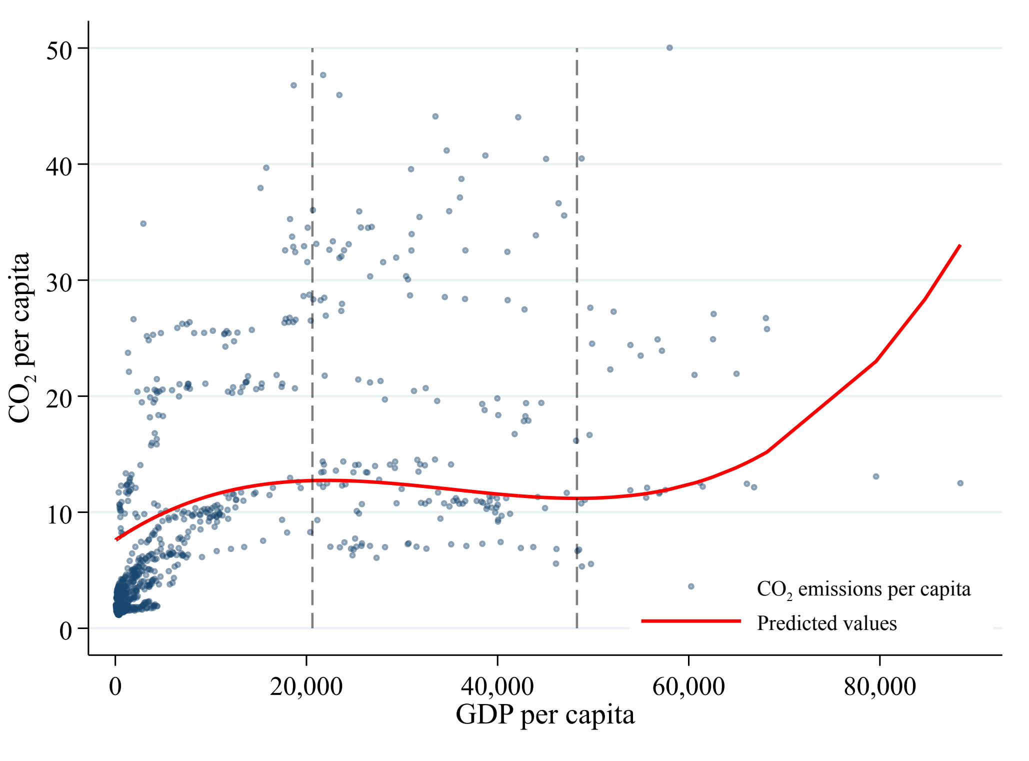

Table V: Analysis of the GDP-CO_2 relationship by development phase

	Phase 1	**Phase 2**	**Phase 3**
GDP/capita ($)	0 - 22,127	22,127 - 48,357	> 48,357
CO_2/capita (t)	9.9 - 12.74	12,74 - 11.19	11.19 - 23.38[†]
Δ CO_2	+2,84	-1,55	+12,18
(%)	(+28,6%)	(-12,2%)	(+108,8%)
GDP-CO_2 relation	Increasing	Decreasing	Increasing
Observations			
N	762	147	63
Share	78,40%	15,12%	6,48%

[†] Estimated value for a GDP/capita of $80,000

5 Conclusion

This study examined the relationship between economic growth and CO_2 emissions in 18 Asia-Pacific countries from 1970 to 2015. The empirical findings confirm the presence of a non-linear, cubic relationship between GDP per capita and CO_2 emissions per capita, supporting an 'N-shaped' EKC trajectory. This dynamic highlights the complexity of the relationship between economic development and environmental degradation. The first phase of the EKC, characterized by a positive correlation between GDP per capita and CO_2 emissions, aligns with the energy-intensive development stage typical of emerging economies. Economic growth during this stage is largely driven by industrialization and increased energy consumption, as reflected in the positive coefficient for energy use per capita. The findings suggest that during this phase, an increase of 1,000 units of energy consumption per capita is associated with a 1.1 tons increase in CO_2 emissions per capita. The second phase indicates a turning point in the relationship, where further economic growth begins to decouple from CO_2 emissions. This transition is marked by a negative coefficient for GDP^2/capita and occurs at a GDP per capita of approximately \$ 22,127. This decline might be attributed to structural economic shifts, technological advancements or maybe an increased environmental awareness leading to cleaner energy practices and policies. However, the third phase, characterized by a sharp increase in emissions at higher income levels, raises critical concerns about the limitations of economic growth as a long-term solution for reducing environmental pressures. This re-intensification of emissions, occurring beyond a GDP per capita of \$ 48,357, may reflect consumption-driven growth, higher energy demands from advanced economies, and the potential for rebound effects. From a policy perspective, these findings suggest that while economic growth can facilitate a transition towards reduced emissions, such outcomes are neither automatic nor guaranteed. Policies targeting energy efficiency, renewable energy adoption, and sustainable consumption patterns are critical to mitigating the re-intensification of emissions in advanced economies. Future research should investigate thr role of additional environmental and development indicators, exploring the role of institutional quality, and examining for instance the long-term impacts of emerging technologies on the EKC trajectory.

6 Discussion

An important limitation of this study lies in the methodological and econometric assumptions underlying. The model's assumption of a uniform functional relationship across all countries does not take account of national particularities, potentially oversimplifying the complex dynamics of the EKC. The turning points identified in the cubic polynomial specification represent averages, which may hide significant heterogeneity among countries with diverse economic and environmental contexts. Furthermore, the imposed cubic functional form, while widely used, is an *a priori* choice that might not accurately reflect the true relationship between economic growth and environmental quality. Exploring alternative specifications, such as nonlinear dynamic models, could reflect better the structural changes and dynamic characteristics of the EKC. Data restrictions also add uncertainty to the results. Although regressions with imputed values produced good results (7.3), these were excluded because a significant portion of missing data occurred after 2015, preventing the model from incorporating critical macroeconomic disruptions such as the COVID-19 pandemic or recent inflation. Econometrically, potential endogeneity between GDP and CO_2 emissions poses a challenge, as the causality could be in both directions. Additionally, the fixed-effects approach captures only time-invariant heterogeneity, leaving temporal structural changes unaddressed. These constraints highligts the need for cautious interpretations.

Moreover, it is essential to have a critical reflection on the underlying conceptual framework of the EKC. While it offers a convenient tool for analyzing the relationship between economic development and environmental degradation, it perpetuates a reductionist view that ties economic wealth and environmental quality primarily to income levels and greenhouse gas emissions. The reliance on GDP/capita and CO2 emissions/capita as primary indicators provides an incomplete view of both economic progress and environmental challenges. While GDP/capita remains a standard metric of economic development, it does not account for critical dimensions such as income inequality, social well-being, or the depletion of natural capital. Similarly, the focus on CO2 emissions/capita, though a crucial marker of global climate impact, omits other environmental pressures, including biodiversity loss, land degradation, deforestation, and air and water pollution. A more holistic approach incorporating a wider range of indicators would allow for a deeper understanding of the multifactorial nature of environmental and economic health.

7 Appendix

7.1 *Marginal effet of GDP/capita on CO_2 emissions*

The marginal effect of GDP/capita on CO_2 emissions is given by the following partial derivative :

$$\frac{\partial \mathrm{CO_2}}{\partial \mathrm{GDP/capita}} = \beta_1 + 2\beta_2 \cdot \mathrm{GDP/capita} + 3\beta_3 \cdot (\mathrm{GDP/capita})^2. \tag{13}$$

Using the estimated coefficients of the model with time dummies :

$$\beta_1 = 0.00055, \quad \beta_2 = -1.81 \times 10^{-8}, \quad \beta_3 = 1.71 \times 10^{-13}, \tag{14}$$

and taking GDP/capita = 20,000 \$, we calculate each term of the marginal effect

$$\begin{aligned} 2 \cdot \beta_2 \cdot \mathrm{GDP/capita} &= 2 \cdot (-1.81 \times 10^{-8}) \cdot 20,000 & (15)\\ &= -0.000724 & (16)\\ 3 \cdot \beta_3 \cdot (\mathrm{GDP/capita})^2 &= 3 \cdot (1.71 \times 10^{-13}) \cdot (20,000)^2 & (17)\\ &= 0.0002052 & (18) \end{aligned}$$

By adding all the terms :

$$\frac{\partial \mathrm{CO_2}}{\partial \mathrm{GDP/capita}} = 0.00055 - 0.000724 + 0.0002052 = 0.0000312 \tag{19}$$

7.2 *Derivation of Turning Points*

The turning points of the Kuznets Environmental Curve are derived from the first derivative of the cubic function relating GDP per capita (GDP) to CO_2 emissions per capita ($CO_2/capita$). The function is expressed as

$$f(GDP) = b_1 \cdot GDP + b_2 \cdot GDP^2 + b_3 \cdot GDP^3 + \epsilon$$

where b_1, b_2, b_3 are the estimated coefficients from the regression.

First Derivative

To identify the turning points, we calculate the first derivative:

$$f'(GDP) = b_1 + 2 \cdot b_2 \cdot GDP + 3 \cdot b_3 \cdot GDP^2$$

Setting $f'(GDP) = 0$ gives the critical points. These are calculated using the quadratic formula:

$$GDP = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where:

- $a = 3 \cdot b_3$
- $b = 2 \cdot b_2$
- $c = b_1$

Second Derivative

To determine the nature of each turning point, we calculate the second derivative

$$f''(GDP) = 2 \cdot b_2 + 6 \cdot b_3 \cdot GDP$$

The second derivative is evaluated at each critical point:

- If $f''(GDP) > 0$, the point is a local minimum.
- If $f''(GDP) < 0$, the point is a local maximum.

7.3 Raw data vs imputed data

Table VI: Comparison of Fixed Effects Models with and without imputation

Variable	**Raw data**	**Imputed data**
GDP/capita	0.00055**	0.00058***
GDP^2/capita	-1.81e-08*	-1.60e-08***
GDP^3capita	1.71e-13**	1.25e-13***
Energy Use/capita	0.0011*	0.00071***
Trade (% of GDP)	0.0045	0.0041*
Constant	7.58***	5.54***
R-squared (Within)	0.48	0.49
R-squared (Between)	0.81	0.67
R-squared (Overall)	0.74	0.62
Corr$(u_i, X_{i,t})$	0.72	0.56
N° Obs	746	972

Significance levels : *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The imputation method allows us to construct a completely cylindrical database (without missing values) of 972 observations. With imputation of the missing values, Trade (% of GDP) becomes significant at the 10% threshold while the variable is not significant in the other models (Table VI). The other coefficients are all significant at the 1% level. We note that, with the exception of energy use/capita, the values of the coefficients are very close. However, the explanatory power of the model seems to be lower with a lower R^2 (0.56<0.72). The correlation between the explanatory variables and the individual effects is also weaker in the imputed model.

References

[1] Gene M Grossman and Alan B Krueger. “2 Environmental Impacts of a North American Free Trade Agreement”. In: *The Mexico-US Free Trade Agreement* 11.2 (1993), p. 13.

[2] Simon Kuznets. “Economic growth and income inequality”. In: *The gap between rich and poor*. Routledge, 2019, pp. 25–37.

[3] Samuel Asumadu Sarkodie and Vladimir Strezov. “A review on environmental Kuznets curve hypothesis using bibliometric and meta-analysis”. In: *Science of the total environment* 649 (2019), pp. 128–145.

[4] David I Stern. “The environmental Kuznets curve after 25 years”. In: *Journal of Bioeconomics* 19 (2017), pp. 7–28.

[5] Hooi Hooi Lean and Russell Smyth. “CO2 emissions, electricity consumption and output in ASEAN”. In: *Applied energy* 87.6 (2010), pp. 1858–1864.

[6] James B Ang. “CO2 emissions, energy consumption, and output in France”. In: *Energy policy* 35.10 (2007), pp. 4772–4778.

[7] Nicholas Apergis and James E Payne. “CO2 emissions, energy usage, and output in Central America”. In: *Energy Policy* 37.8 (2009), pp. 3282–3286.

[8] Mohamed El Hedi Arouri et al. “Energy consumption, economic growth and CO2 emissions in Middle East and North African countries”. In: *Energy policy* 45 (2012), pp. 342–349.

[9] Duc Nha Le. “Environmental degradation and economic growth in ASEAN-10: The perspective of EKC hypothesis”. In: *Malaysian Journal of Economic Studies* 56.1 (2019), pp. 43–62.

[10] Sahbi Farhani et al. “The environmental Kuznets curve and sustainability: A panel data analysis”. In: *Energy Policy* 71 (2014), pp. 189–198.

[11] Boqiang Lin and Mengmeng Xu. “Does China become the “pollution heaven” in South-South trade? Evidence from Sino-Russian trade”. In: *Science of The Total Environment* 666 (2019), pp. 964–974.

[12] Kangyin Dong et al. “Does natural gas consumption mitigate CO2 emissions: testing the environmental Kuznets curve hypothesis for 14 Asia-Pacific countries”. In: *Renewable and Sustainable Energy Reviews* 94 (2018), pp. 419–429.

[13] Emmanuel Duguet. “Econométrie des panels avec applications”. In: *Tiré de http://emmanuel. duguet. free. fr/ouvrage_panel_07_v2_3. pdf* (2009).

[14] Badi Hani Baltagi and Badi H Baltagi. *Econometric analysis of panel data*. Vol. 4. Springer, 2008.

[15] Pietro Balestra and Marc Nerlove. “Pooling cross section and time series data in the estimation of a dynamic model: The demand for natural gas”. In: *Econometrica: Journal of the econometric society* (1966), pp. 585–612.

[16] Jerry A Hausman. “Specification tests in econometrics”. In: *Econometrica: Journal of the econometric society* (1978), pp. 1251–1271.