Spatial Dependence in the Environmental Kuznets Curve

by

Anna Pellicer Carrasquer

NEKN01
Master Essay I
January 2022
Supervisor: Peter Jochumzen

Abstract

This essay aims to detect spatial autocorrelation present in the EKC relationship, as well as to find empirical evidence of said relationship. A panel data of 176 countries for the period 1990-2018 is used for the empirical part. Several spatial models are considered in this essay, namely the SDM, SAR model, SAC model and SEM. The main conclusions are: (i) there exists significant spatial autocorrelation in the CO_2 emissions, (ii) significant evidence for a direct EKC relationship is found, (iii) the renewable energy consumption and the CO_2 emissions are significantly and negatively correlated, (iv) among all models considered in this essay the SAR model with time and spatial fixed effects is the one with less information loss.

Keywords: environmental Kuznets curve, spatial autocorrelation, panel data, CO_2 emissions, renewable energy consumption

Contents

1. Introduction

The pollution-income relationship (PIR) is widely studied by many authors, trying to identify the relationship between environmental deterioration and economic growth. The most supported relationship is the Environmental Kuznets curve (from now on EKC). The EKC is a hypothesis that argues that it exists an inverted-U shaped relationship between pollution and income.

The EKC hypothesis identifies two big phases in an economy. During the early phase there is a domination of the primary production. Since there is a limited economic activity, the waste generated is also limited. The economy presents a positive relationship between pollution and income during this period. Then, the economy moves to an industrial economy, where the secondary sector dominates. We have a stronger positive relationship between the two variables. Both income and pollution keep increasing until it gets to the turning point, where pollution is no longer linked to the increase in income. When the economy reaches that point, economic growth and environmental quality have a negative relationship.

There are five major causes, apart from income, that can explain the EKC relationship. The first one is the income distribution. The basic idea is that there is a higher awareness of environmental issues in a country with a more equitable income distribution, hence environmental regulations are more likely to be imposed. The second one is international trade. This argument deals with the pollution haven hypothesis, which argues that developed countries are able to reduce pollution by importing pollution-intensive goods from (usually low-income) other countries. The third cause is structural change, including the transition from industry to the service sector, and technical progress, including improvements in production techniques. The fourth one is institutional framework. This deals with the idea that environmental improvements depend on public policies. Finally, the last cause is consumers' preferences. This deals with the microeconomic implications of consumers preferences as a partial explanation of an inverted-U pattern (Kaika and Zervas, 2013a).

The remainder of the essay is organized as follows. In Section 2 a brief literature review is presented. Section 3 describes the data used for the empirical part of the essay. Section

4 describes the methodology used. In Section 5 the results are presented. In Section 6 the conclusions are stated.

2. Literature review

The EKC literature can be traced back to Grossman and Krueger (1991, 1995), who studied the relationship between the emissions and the size of an economy. They found that during the first stages of a country's development there is a positive relationship between emissions and size of an economy, while after reaching a certain income level the relationship becomes negative. They split this relationship in three main effects: scale effect, composition effects and technique effect. The scale effect means that a high consumption of fossil-fuel energy is needed to meet the demand for a high level of production output. The composition effect deals with the idea that when an economy moves from an energy-intensive economy towards a service-based economy, it increases its environmental sustainability. The technique effect deals with the idea that countries with higher income per capita levels can invest more money on research, inducing a shift from dirty to clean technologies.

Among the wide EKC literature, the studies present different results regarding the pollution-income relationship. Some studies find a linear relationship between environmental depletion and income per capita (e.g., Azomahou et al., 2006), other studies find the typical inverted-U relationship (e.g., Iwata et al. 2011), while other studies find evidence for an N-shaped relationship (e.g., Friedl and Getzner, 2003). In addition, some studies some studies failed to find evidence of the EKC hypothesis for CO_2 emissions (e.g., Halicioglu, 2009).

From a chronological point of view, the empirical studies of EKC for CO_2 can be split in three main categories. The first category of studies, which starts around 1995, is mainly focused on the effect of income on CO_2 emissions, without accounting for other explanatory variables. The majority of these empirical studies reject the EKC hypothesis. For instance, Moomaw and Unruh (1997) find an increasing and monotonous relationship between income and pollution. Other studies find evidence of an EKC for CO_2 but with a turning point of excessive size in per capita dollars (e.g., Holtz-Eakin and Selden, 1995). This first category of studies has an important lack of explanatory capacity because they are only focused on per capita income levels (Balado-Naves et al., 2018). Some authors, for instance Vincent (1997), argue that the EKC studies should be conducted in single country contexts.

The second category of studies included new explanatory variables, using time series models instead of cross sectional or panel data models. From the late 1990's to the mid 2000's the error correction models and autoregressive models became popular among the EKC studies. One of the main focuses of this category of studies was the effect of structural changes on the EKC relationship (Panayotou et al., 2000; Friedl and Getzner, 2003; Cole, 2004). However, there was no consensus about the results. According to Balado-Naves et al. (2018) it may be due to the quality of the explanatory variables that they used. During this period, Friedl and Getznet (2003) tested the significance of the share of the services sector over GDP and found it non-significant. Lastly, the EKC and the structural change hypotheses were rejected by Azomahou et al. (2006).

The third category of empirical studies can be divided in two subgroups. The first subgroup is characterized for using traditional panel data or time series. These studies also found a positive and monotonous relationship between a country's income per capita and carbon dioxide emissions and large sized turning points (e.g., Kearsley and Riddel, 2010; Franklin and Ruth, 2012 or Zhang and Zhao, 2014). The other subgroup of studies used Autoregressive distributed lag (ARDL) models (e.g., Coondoo and Dinda, 2008; Halicioglu, 2009; Narayan and Narayan, 2010; Iwata et al., 2011; Baek and Kim, 2013; Bölük and Mert, 2014, 2015; Baek, 2015; Al-Mulali et al., 2016). The ARDL studies fail to achieve a concordant result of the EKC hypothesis. However, they found a positive and significant impact of both nuclear and renewable energies on environmental protection.

The majority of studies that have included spatial interactions are focused on China (Auffhammer and Carson, 2008; Chuai et al., 2012; Kang et al., 2016; Wang and Ye, 2017). These studies have not found support for the EKC hypothesis, but they detected a significant spatial autocorrelation among Chinese provinces.

Maddison (2006) used spatial models to analyse the emissions of several pollutants, concluding that there exists spatial autocorrelation between countries, specifically he found that national emissions are influenced by the neighbouring income. Balado-Naves et al., (2018) also used spatial models to analyse the emissions of CO_2, and found support for both direct and indirect EKC for all continents except for Oceania, where they found the opposite effect, an U-relationship between income and CO_2 emissions.

3. Data

The qualitative data used for this essay comes from the World Development Indicators. The sample consists of 5104 observations from 176 countries and covers the period between 1990 and 2018. The dependent variable used is CO_2 emissions per capita. The independent variables are GDP per capita and the renewable energy consumption as a percentage of the total final energy consumption. A natural logarithmic transformation is used for the CO_2 and the GDP data, allowing to interpret the results in percentage terms. Descriptive statistics of the qualitative data used are presented in Table 1.

The spatial data from which the spatial matrix have been developed is from Country borders database.

In figure 1 we can see that there is a positive relation between CO_2 emissions per capita and income per capita. However, for the countries with higher income per capita we can start to see a downward relationship between the two variables. If this negative relationship for high income-per-capita countries is statistically significant, it would imply the existence of an EKC relationship.

CO_2 emissions and renewable energy are negatively related, as can be seen in figure 2. A country with highest renewable energy consumption should be expected to have a lower CO_2 emissions per capita.

TABLE 1: Descriptive statistics

Variable		Mean	Std. Dev.	Min	Max
lnCO2	overall	0.574	1.588	-3.829	3.865
	between		1.564	-3.363	3.563
	within		0.297	-2.720	2.723
lnGDP	overall	8.106	1.595	4.506	12.103
	between		1.502	5.225	11.524
	within		0.546	5.372	10.063
lnGDP2	overall	68.257	26.251	20.301	146.476
	between		24.855	17.391	132.985
	within		8.644	28.690	100.893
Renewable Energy Consumption	overall	33.923	30.444	0.000	98.304
	between		29.815	0.000	94.426
	within		6.541	2.439	83.626

Notes: overall is statistic over time and countries, between is statistic between countries and within is statistic over time within a country.

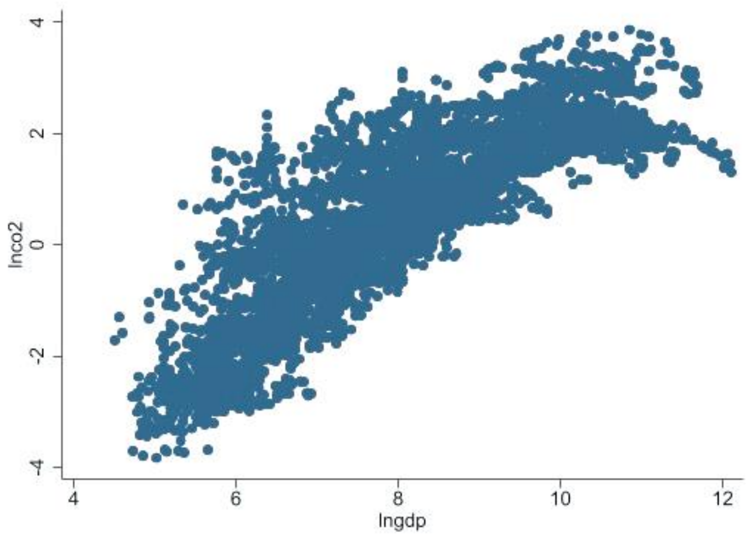

FIGURE 1: Relation between income per capita (in natural logarithm) and CO_2 (in natural logarithm)

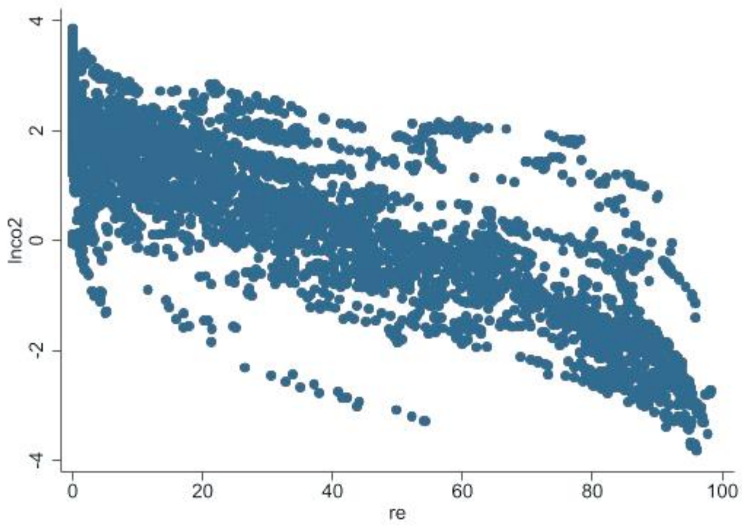

FIGURE 2: Relation between renewable energy consumption and CO_2 (in natural logarithm)

4. Methodology

4.1. Standard EKC

Following the specification from Balado-Naves et al. (2018), the following adjusted version of the general model is used:

$$ln\, e_{it} = \alpha_i + \gamma_t + \beta_1 ln\, x_{it} + \beta_2 (ln\, x_{it})^2 + \beta_3 RE_{it} + \varepsilon_{it}$$

where e_{it} are the CO_2 emissions per capita, x_{it} is the income per capita and RE_{it} represents the renewable energy consumption (% of total final energy consumption) of country i at time t. The individual fixed effects are captured by α_i, the time fixed effects are captured by γ_t and ε_{it} is the error term. Due to data availability, the regressor energy intensity is not included in this essay.

The signs for the parameters are expected to be as follows

i. $\beta_1 > 0$ and $\beta_2 < 0$. If the EKC hypothesis is true, the pollution per capita should increase with income per capita but have an inverse-U shape, implying that the coefficient for the quadratic term should be negative.

ii. $\beta_3 < 0$. It is expected that the higher the renewable energy consumption the lower are the carbon emissions per capita.

4.2. Spatial EKC

By using a spatial regression model, I am able to capture spatial interaction between observations, meaning that data points from one location affect data points from nearby locations.

There are two big groups of spatial effects. The first one considers spatial dependence or autocorrelation, and the second one deals with spatial heterogeneity. This essay will account for spatial autocorrelation. Spatial autocorrelation is expressed by the moment condition

$$cov[y_i, y_j] = E[y_i y_j] - E[y_i]E[y_j] \neq 0 \; for \; i \neq j$$

where y_i (y_j) is the value of y at location i (j). This covariance structure "becomes meaningful from a spatial perspective when the particular configuration of nonzero i, j pairs has an interpretation in terms of spatial structure, spatial interaction or the spatial arrangement of the observations" (Anselin, 2001).

The Moran's I test statistic is computed in order to test for spatial autocorrelation. Following Kondo (2021), Moran's I is given by

$$I = \frac{z^T W z}{z^T z}$$

where z is the (n x 1) demeaned and standardized dependent variable and W is the row-standardized spatial weights matrix. Asymptotically, this statistic tends to zero. Therefore, a positive value suggests positive spatial autocorrelation and a negative value indicates negative spatial autocorrelation. The null hypothesis of Moran's I test is that the data presents spatial randomization. The test statistic is given by

$$z(I) = \frac{I - E[I]}{\sqrt{Var[I]}}$$

where E[I] and Var [I] are computed under the null hypothesis of spatial randomization, which are given by

$$E[I] = -\frac{1}{n-1}$$

and

$$Var[I] = E[I^2] - (E[I])^2$$

The first term in the variance is calculated as follows

$$E[I^2] = \frac{n[(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2] - \frac{m_4}{m_2^2}[(n^2 - n)S_1 - 2nS_2 + 6S_0^2]}{(n-1)(n-2)(n-3)S_0^2}$$

where m_2 is the second moment about the sample mean and m_4 is the fourth moment about the sample mean. The terms S_0, S_1 and S_2 are given by

$$S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \ , \quad S_1 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(w_{ij} + w_{ji})^2 \ , \quad and \ \ S_2 = \sum_{i=1}^{n}(\sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{n} w_{ij})^2$$

The Moran's I test statistic follows a standard normal distribution under the null. According to Anselin (2001), Moran's I test is the locally best invariant and "consistently outperforms other tests in terms of power in simulation experiments".

If Moran's I test confirms that there exist spatial relationships for CO_2 emissions per capita, the next step would be to choose between all spatial models in order to correctly capture the presence of spatial dependence.

The spatial models considered in this essay are the following: spatial Durbin model (SDM), spatial autoregressive model (SAR), spatial autocorrelation model (SAC) and spatial error model (SEM).

The spatial Durbin model includes an autoregressive term for the dependent variable as well as spatially lagged independent variables. The basic equation for the SDM is

$$ln\, e_{it} = \alpha_i + \gamma_t + \rho W ln\, e_{it} + \beta_1 ln\, x_{it} + \beta_2 (ln\, x_{it})^2 + \beta_3 RE_{it} + \theta_1 \sum_{j=1}^{N} w_{ij} ln\, x_{it} + \theta_2 \sum_{j=1}^{N} w_{ij} (ln\, x_{it})^2 + \theta_3 \sum_{j=1}^{N} w_{ij} Re_{it} + \varepsilon_{it}$$

In this model, $\sum_{j=1}^{N} w_{ij} ln\, x_{it}$, $\sum_{j=1}^{N} w_{ij} (ln\, x_{it})^2$ and $\sum_{j=1}^{N} w_{ij} Re_{it}$ represent the spatially lagged regressors. They capture the interaction effects of all j neighbors to country i via income per capita and renewable energy consumption.

The SAC model adds a spatial autoregressive error to the SAR model

$$ln\, e_{it} = \alpha_i + \gamma_t + \rho W ln\, e_{it} + \beta_1 ln\, x_{it} + \beta_2 (ln\, x_{it})^2 + \beta_3 RE_{it} + \nu_{it}$$
$$\nu_{it} = \lambda W \nu_{it} + \varepsilon_{it}$$

The basic equation for the SAR model is

$$ln\, e_{it} = \alpha_i + \gamma_t + \rho W ln\, e_{it} + \beta_1 ln\, x_{it} + \beta_2 (ln\, x_{it})^2 + \beta_3 RE_{it} + \varepsilon_{it}$$

In this model it is assumed that $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$ and $E(\varepsilon_{it}\varepsilon_{is}) = 0$ for $i \neq j$ and/or $t \neq s$.

It can easily be shown that the SDM and the SAC models nest the SAR model. For that reason, the SAR estimated parameters should be unbiased even when the SDM or the SAC are the true model (LeSage, 2014).

Finally, the spatial error model (SEM), which is a special case of the SAC model, it is focused on the spatial autocorrelation in the errors

$$ln\, e_{it} = \alpha_i + \gamma_t + \beta_1 ln\, x_{it} + \beta_2 (ln\, x_{it})^2 + \beta_3 RE_{it} + \nu_{it}$$
$$\nu_t = \lambda W \nu_t + \varepsilon_t$$

In all the models above e_{it} are the CO_2 emissions per capita, x_{it} is the income per capita and RE_{it} represents the renewable energy consumption (% of total final energy consumption) of country i at time t. The individual fixed effects are captured by α_i, the time fixed effects are captured by γ_t and $\varepsilon_{it}$is the error term. W is the spatial weights matrix. β, ρ, λ and θ are the coefficients to be estimated.

The signs for the parameters from all spatial models above are expected to be as follows

i. $\beta_1 > 0$ and $\beta_2 < 0$. If here exists a direct EKC relationship the pollution per capita should increase with income per capita but have an inverse-U shape, implying that the coefficient for the quadratic term should be negative.

ii. $\beta_3 < 0$. It is expected that the renewable energy consumption and the CO_2 emissions per capita are negatively correlated.

iii. If the SDM is the true model generating the data, the following significant parameters are expected: $\rho > 0$, $\theta_1 > 0$, $\theta_2 < 0$, $\theta_3 < 0$. However, λ is expected to be non-significant. This model would suggest the existence of an indirect EKC relationship.

iv. If the SAR model is the true model generating the data, the following signs for the parameters are expected: $\rho > 0$ and significant. On the other hand, θ_1, θ_2, θ_3 and λ are expected to be non-significant.

v. If the SAC model is the true model generating the data, the following signs for the parameters are expected: $\rho > 0$ and $\lambda > 0$ and significant. However, θ_1, θ_2 and θ_3 are expected to be non-significant.

vi. If the SEM is the true model generating the data, it is expected $\lambda > 0$ and significant, while the rest of the parameters, ρ, θ_1, θ_2 and θ_3 to be non-significant.

4.3. Model Selection

The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are computed for model selection. AIC and BIC estimate the quality of every model, and therefore they can be compared choose the most fitting model. The model with the smallest AIC has the least information lost and for that reason it will be the model with the higher quality. The model with minimum BIC is the most likely to have generated the dependent variable, meaning that is the model with the largest posterior probability.

The AIC and BIC are specified as follows

$$AIC = -2\{\log[L(\hat{\beta})] - df(\hat{y})\}, \qquad and \qquad BIC = -2\{\log[L(\hat{\beta})] - \log(N)df(\hat{y})\}$$

where $\hat{\beta}$ is the vector of estimated parameters in the chosen model.

When it comes to choose between both criterions there is no clear choice. On one hand, BIC is asymptotically consistent, meaning that as the sample size goes to infinity the probability that BIC will select the true model approaches to one. On the other hand, AIC minimizes the risk of selecting a bad model from all candidate models (Burnham and Anderson 2004). For these reasons, both criterions will be considered in this essay.

5. Results

To begin with, a Hausman test for the general model is computed in order to test whether a random effects or fixed effects specification should be used. The null hypothesis of the test is that a random effects model is preferred. The alternative hypothesis is that a fixed-effects model is preferred. Table 2 shows the results of said test. The null hypothesis can be rejected at 1% confidence level and therefore a fixed effects specification should be used.

TABLE 2: Hausman test

chi2(3)	129,64
Prob > chi2	0,0000

Table 3 presents the results for the standard EKC model using fixed-effects. This model finds all three independent variables significant at 1% significance level. This model confirms the expectations regarding the coefficient's signs, since $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 < 0$. There is a positive and significant relationship between emissions per capita and income per capita, as well as a negative relationship between emissions per capita and income squared per capita. In addition, there is a negative and significant relationship between the emissions per capita and the renewable energy consumption. This model finds significant evidence of a EKC relationship.

TABLE 3: Standard EKC

lngdp	0.629*** (0.0355)
lndgp2	-0.029*** (0.002)
re	-0.023*** (0.0004)
R2	0.7952
AIC	-1802.418
BIC	-1776.267

Notes: Standard deviation in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

The next step is to test for spatial randomization in our sample. As mentioned in Section 4, the Moran's test is computed for that reason. The results of the test are stated in Table 4. We find the Moran's I test statistically significant at 1% level for each year and therefore the null hypothesis of spatial randomization can be rejected. Additionally, since I > 0 for every year, a positive spatial autocorrelation should be captured by all spatial models considered.

Once the existence of spatial dependence for CO_2 emissions per capita has been confirmed, it is still necessary to choose between all spatial models in order to correctly capture the presence of spatial dependence. As mentioned in Section 4.2, four different spatial models will be used. Specifically, the spatial Durbin model (SDM), the spatial autoregressive model (SAR), the spatial autocorrelation model (SAC) and the spatial error model (SEM).

TABLE 4: Moran's I test

Year	I	z(I)	p-value
2018	0.633	9.155	2.5253E-19***
2017	0.639	9.243	1.1176E-19***
2016	0.635	9.181	1.9874E-19***
2015	0.635	9.183	1.9447E-19***
2014	0.644	9.310	6.0333E-20***
2013	0.644	9.314	5.8102E-20***
2012	0.652	9.425	2.0409E-20***
2011	0.650	9.402	2.5493E-20***
2010	0.646	9.343	4.4223E-20***
2009	0.661	9.547	6.4701E-21***
2008	0.640	9.251	1.0416E-19***
2007	0.635	9.180	1.9933E-19***
2006	0.625	9.040	7.1801E-19***
2005	0.633	9.147	2.7054E-19***
2004	0.635	9.174	2.1156E-19***
2003	0.656	9.472	1.3131E-20***
2002	0.646	9.328	5.101E-20***
2001	0.649	9.382	3.0708E-20***
2000	0.651	9.409	2.383E-20***
1999	0.656	9.475	1.2805E-20***
1998	0.662	9.558	5.7718E-21***
1997	0.659	9.515	8.7316E-21***
1996	0.662	9.564	5.4711E-21***
1995	0.666	9.619	3.224E-21***
1994	0.654	9.441	1.7673E-20***
1993	0.677	9.765	7.8855E-22***
1992	0.677	9.769	7.5788E-22***
1991	0.693	9.997	7.9143E-23***
1990	0.694	10.008	7.1314E-23***

Note: I represents the value of the Moran's I for each year and z(I) is the test statistic value.
Significance level: ***p<0.01, **p<0.05, *p<0.1.

The results for the Spatial Durbin model (SDM) are stated in Table 5. The results for this model present differences depending on the chosen specification. When using a time fixed-effects model there is no statistical evidence of a direct EKC relationship, since β_2, the coefficient for the squared income per capita, is not statistically significant. On the other hand, θ_1 is positive and significant, and θ_2 is negative and significant. The income per capita and the squared income per capita from the j neighboring countries are statistically significant to explain the emissions of CO_2 per capita of country i. This would imply the existence of an indirect EKC relationship. In addition, θ_3 is positive and significant, implying that there might be a positive relationship between the renewable energy consumption from country j to the emissions of CO_2 of country i, which can come as a surprise.

The spatial fixed-effects SDM finds significant support for a direct EKC, since β_1 is positive and significant and β_2 is negative and significant. Additionally, there is a significant and negative relationship between renewable energy consumption and CO_2 emissions. In this model, θ_2 is negative and significant at 5% level, which means that an increase in the income of country j would affect negatively the emissions of country i if they are neighbors.

When using time and spatial fixed-effects, a significant relationship of at least 5% confidence level is found for all three independent variables, while there is no significance in the effects of the regressors of the j neighbors to the emissions per capita of country i.

Nevertheless, all three specifications of the SDM find a positive and significant ρ, confirming that the emissions of one country have a positive effect on their neighbor's emissions.

Regarding the AIC and the BIC for this model, we can rank these models from worst to best, the model with time fixed-effects being the worst and the model with time and spatial fixed-effects being the best. When moving from the time FE specification to the spatial FE specification we have a difference of 9,766,765 for the AIC and a difference of 9.766,764 for the BIC. Choosing the time and spatial FE instead of the spatial FE ends with a difference of 18,238 for both the AIC and the BIC. This difference is definitely smaller than the previous one, but still significant in order to choose a correct specification (Burnham and Anderson, 2004).

TABLE 5: Spatial Durbin model (SDM)

	Time FE	Spatial FE	Time and Spatial FE
lngdp	0.971*** (0.2909533)	0.485*** (0.0426517)	0.476*** (0.144)
lndgp2	-0.189 (0.017)	-0.020*** (0.003)	-0.018** (0.009)
re	-0.019*** (0.002)	-0.021*** (0.0004)	-0.021*** (0.003)
Wlngdp	0.290** (0.123)	-0.126** (0.058)	--0.099 (0.173)
Wlngdp2	-0.035*** (0.010)	0.004 (0.004)	0.006 (0.011)
Wre	0.057** (0.003)	-0.0009 (0.0009)	-0.001 (0.002)
rho	0.409*** (0.526)	0.278*** (0.016)	0.273*** (0.039)
R^2	0.7953	0.8013	0.8230
AIC	7556.455	-2210.31	-2228.548
BIC	7608.757	-2158.008	-2176.246

Notes: Standard deviation in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

Table 6 shows the results for the SAR model. The results for the SAR model indicate a positive and significant at 1% level relationship between emissions per capita and income per capita, as well as a negative and significant at 5% level relationship between emissions per capita and income squared per capita. In addition, the renewable energy consumption affects negatively the emissions per capita. This model also finds significant spatial spillovers, meaning that emissions from one country have a positive relationship with the neighboring countries' emissions.

Again, following the AIC and the BIC, the model specification with time and spatial fixed effects has much lower values than the other two models. This implies that the last model minimizes the information loss among all three specifications, and therefore it is the one closer to the true model.

TABLE 6: Spatial autoregressive model (SAR)

	Time FE	Spatial FE	Time and Spatial FE
lngdp	1.125*** (0.277)	0.438*** (0.134)	0.442*** (0.136)
lndgp2	-0.042*** (0.015)	-0.020*** (0.008)	-0.016** (0.008)
re	-0.019*** (0.002)	-0.022*** (0.003)	-0.021*** (0.003)
rho	0.169*** (0.033)	0.255*** (0.036)	0.275*** (0.037)
R^2	0.8492	0.7968	0.8249
AIC	8092.69	-2185.158	-2230.709
BIC	8125.378	-2152.469	-2198.021

Notes: Standard deviation in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

The results for the SAC model are stated in Table 7. When using the SAC model, we find evidence of a direct EKC since β_1 is positive and significant at 1% confidence level, and β_2 is negative and significant. In the time and spatial fixed-effects specification, β_2 is only significant at 10% confidence level, while for the other specifications is significant at 1% confidence level.

TABLE 7: Spatial autocorrelation model (SAC)

	Time FE	Spatial FE	Time and Spatial FE
lngdp	1.454*** (0.280)	0.464*** (0.146)	0.423*** (0.142)
lndgp2	-0.049*** (0.016)	-0.021*** (0.008)	-0.154* (0.008)
re	-0.018*** (0.002)	-0.022*** (0.003)	-0.021*** (0.003)
rho	0.0579 (0.037)	0.225*** (0.085)	0.296*** (0.081)
lambda	0.438*** (0.061)	0.048 (0.011)	0.040 (0.103)
R^2	0.8523	0.7989	0.8234
AIC	7494.993	-2185.044	-2227.739
BIC	7534.22	-2145.818	-2188.512

Notes: Standard deviation in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

In this model we found disparities regarding the significance in the spatial terms included in the regression. On one hand, when using time fixed-effects λ, which captures the effect of a spatial autoregressive error, is significant at 1%. On the other hand, when using time fixed-effects or both time and spatial fixed-effects ρ is positive and significant while λ is non-significant, meaning that there exists a significant spatial autoregressive term for the dependent variable, but not for the error term.

According to the AIC and BIC, the model that performed best is the one that included time and spatial fixed-effects, and the model taking into account only time fixed-effects is the one that performed worst. Therefore, the evidence for ρ being significant is larger than the evidence for λ being significant.

The results for the spatial error model (SEM) are stated in Table 8. In this model we find significant evidence of a direct EKC since $\beta_1 > 0$ and $\beta_2 < 0$. The renewable energy consumption affects negatively the CO_2 emissions. Additionally, there is a significant and positive λ, meaning that there is a spatial autoregressive error.

By comparing the AIC and the BIC for the three specifications we can rank them from the time fixed-effects being the worst to the time and spatial fixed-effects being the best specification since it is minimizing the information loss.

TABLE 8: Spatial error model (SEM)

	Time FE	Spatial FE	Time and Spatial FE
lngdp	1.421*** (0.282)	0.613*** (0.145)	0.621*** (0.144)
lndgp2	-0.046*** (0.016)	-0.028*** (0.008)	-0.027*** (0.009)
re	-0.019*** (0.002)	-0.022*** (0.003)	-0.021*** (0.003)
lambda	0.485*** (0.051)	0.298*** (0.042)	0.290*** (0.043)
R^2	0.8513	0.7995	0.811
AIC	7533.212	-2117.774	-2142.394
BIC	7565.9	-2085.085	-2109.705

Notes: Standard deviation in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

TABLE 9: Time and Spatial Fixed-Effects Models

	SDM withTime and Spatial FE	SAR with Time and Spatial FE	SAC with Time and Spatial FE	SEM with Time and Spatial FE
lngdp	0.476*** (0.144)	0.442*** (0.136)	0.423*** (0.142)	0.621*** (0.144)
lndgp2	-0.018** (0.009)	-0.016** (0.008)	-0.154* (0.008)	-0.027*** (0.009)
re	-0.021*** (0.003)	-0.021*** (0.003)	-0.021*** (0.003)	-0.021*** (0.003)
Wlngdp	--0.099 (0.173)			
Wlngdp2	0.006 (0.011)			
Wre	-0.001 (0.002)			
rho	0.273*** (0.039)	0.275*** (0.037)	0.296*** (0.081)	
lambda			0.040 (0.103)	0.290*** (0.043)
R^2	0.8230	0.8249	0.8234	0.811
AIC	-2228.548	-2230.709	-2227.739	-2142.394
BIC	-2176.246	-2198.021	-2188.512	-2109.705

Notes: Standard deviation in parentheses. Significance level: ***p<0.01, **p<0.05, *p<0.1.

On the whole, all models that include a spatial term in the specification find strong evidence of spatial autocorrelation in the model, in the form of an autoregressive term in the dependent variable or in the error term, but not both simultaneously. In Table 9 the results for the four considered models with time and spatial fixed-effects are presented for an easier comparison.

Regarding model selection, the models with time and spatial fixed-effects performed best regarding the AIC and BIC, while models with only time-fixed effects performed the worst among all other models. The SAR model with both time and spatial fixed-effects performed the best, with an AIC of -2.230,709; a BIC of -2.198,021 and a R2 of 0,8249. Since it is the model with lower values of AIC and BIC, it is the model with the lowest information loss and therefore it is the model that is closest to the true model generating the data. This model finds the variables income per capita and renewable energy significant at 1% level, while the squared income per capita significant at 5% level. The spatial autoregressive term is also strongly statistically significant. According to this model $\beta_1 > 0$ and $\beta_2 < 0$, and therefore it suggests the existence of a direct EKC relationship. In addition, $\beta_3 < 0$, meaning an increase in renewable energy consumption would decrease CO_2 emissions. Lastly, ρ is also positive and significant, as expected from the Moran's I test. This imply that the CO_2 emissions from one country are positively correlated with the CO_2 emissions from the neighboring countries.

The SDM and SAC models with time and spatial fixed-effects have an AIC and BIC that is pretty close to the SAR model ones. That does not come as a surprise since both models nest the SAR model, and the extra parameters compared to the SAR model are not statistically significant.

6. Conclusions

An analysis for the presence of spatial autocorrelation in the environmental Kuznets curve model is performed using panel data from 176 countries during the period 1990-2018. I started by estimating a standard EKC model without accounting for spatial spillovers, and we found evidence of a EKC relationship between CO_2 emissions and income per capita. I then added spatial autocorrelation terms to the regression, and estimated again the EKC relationship using four different spatial models. Even with spatial interactions, there is statistically significant evidence of an EKC relationship, understood as a positive effect of income and a negative effect of income squared.

Overall, the result of the analysis indicates that there is statistically significant spatial autocorrelation in the EKC. The model selection is based on the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The model that performed best according to both criterions is the spatial autoregressive model. That would imply that the SAR model with time and spatial fixed-effects is the one closer to the true model generating the data among all considered models.

The lack of explanatory variables is the main limitation of this essay. The main objective was to have a panel data set with the maximum countries and years. For that reason, multiple explanatory variables that other studies found significant in order to explain the EKC are not included in this essay. Another limitation is the heterogeneity of the countries considered. Two suggestions are made for future research. The first suggestion is to perform the regressions in smaller groups of countries to better capture the spatial heterogeneity (e.g., by continents). The second suggestion is to perform emissions forecasts given the estimations.

References

Al-Mulali, U., Solarin, S.A., Ozturk, I., 2016. Investigating the presence of the environmental Kuznets curve (EKC) hypothesis in Kenya. An autoregressive distributed lag (ARDL) approach. Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards 80 (3), 1729‑1747.

Anselin, L., 1988. Spatial econometrics: Methods and models. Dordrecht: Kluwer Academic Publishers.

Anselin, L., 2001. Spatial econometrics. A companion to theoretical econometrics. Chapter 14, 310-330.

Anselin, L. 1995. Local indicators of spatial association—LISA. Geographical Analysis 27(2): 93–115.

Auffhammer, M., Carson, R.T., 2008. Forecasting the path of China's CO2 emissions using province-level information. Journal of Environmental Economics and Management 55, 229‑247

Azomahou, T., Laisney, F., Nguyen Van, P., 2006. Economic development and CO2 emissions. A nonparametric panel approach. Journal of Public Economic, 90 (6‑7), 1347‑1363.

Balado-Naves, R., Baños-Pino, J., Mayor, M., 2018. Do countries influence neighbouring pollution? A spatial analysis of the EKC for CO_2 emissions. Energy Policy 123, 266-279.

Baek, J., Kim, H.S., 2013. Is economic growth good or bad for the environment? Empirical evidence from Korea. Energy Economics 36, 744‑749.

Baek, J., 2015. A panel cointegration analysis of CO2 emissions, nuclear energy and income in major nuclear generating countries. Applied Energy 145, 133‑138.

Belotti, F., Hughes, G., 2016. Spatial panel data models using Stata. The Stata Journal ii, 1-37.

Bhandu Murthy, K., Gambhir, S., 2018. Analizing Environmental Kuznets Curve and Pollution Haven Hypothesis in India in the Context of Domestic and Global Policy Change. Bussiness and Finance Journal, 12 (2), 134-156.

Burnham, K., Anderson, D., 2004. Model Selection and Multimodel Inference. A Practical Information-theoretic Approach.

Bölük, G., Mert, M., 2014. Fossil & renewable energy consumption, GHGs (greenhouse gases) and economic growth. Evidence from a panel of EU (European Union) countries. Energy 74 (C), 439‐446.

Bölük, G., Mert, M., 2015. The renewable energy, growth and environmental Kuznets curve in Turkey. An ARDL approach. Renewable and Sustainable Energy Reviews 52 (C), 587‐595.

Chuai, X., Huang, X., Wang, W., Wen, J., Chen, Q., Peng, J., 2012. Spatial econometric analysis of carbon emissions from energy consumption in China. J. Geogr. Sci. 22 (4), 630‐642.

Cliff, A. D., and J. K. Ord., 1970. Spatial autocorrelation: a review of existing and new measures with applications. Economic Geography 46: 269–292.

Cole, M.A., 2004. Trade, the pollution haven hypothesis and the environmental Kuznets curve. Examining the linkages. Ecological Economics, 48 (1), 71‐81.

Coondoo, D., Dinda, S., 2008. Carbon dioxide emission and income. A temporal analysis of cross-country distributional patterns. Ecological Economics 65 (2), 375‐385.

Franklin, R.S., Ruth, M., 2012. Growing up and cleaning up. The environmental Kuznets Curve Redux. J. Applied Geography, 32, 29‐39.

Friedl, B., Getzner, M., 2003. Determinants of CO2 emissions in a small open economy. Ecological Economics, 45 (1), 133‐148.

GeoDataSource. Country borders database. GeoDataSource. Retrieved September 19, 2021, from https://www.geodatasource.com/addon/country-borders

Grossman, G.M., Krueger, A.B., 1991. Environmental Impacts of a North American Free Trade Agreement. National Bureau of Economic Research Working Paper Series, 3914, 1‐57.

Grossman, G.M., Krueger, A.B., 1995. Economic growth and the environment. The Quarterly Journal of Economics, 110 (2), 353–377.

Halicioglu, F., 2009. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. Energy Policy 37 (3), 1156‑1164.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. New York, NY, USA: Springer New York Inc..

Holtz-Eakin, D., Selden, T., 1995. Stoking the fires? CO_2 emissions and economic growth. Journal of Public Economics 57, 85‑101.

Iftikhar, Y., Ahmad, N., Chaudhary, A., 2020. The impact of financial development, political institutions, and urbanization on environmental degradation: evidence from 59 less-developed economies. Environment, Development and Sustainability. 23 (5), 6698–6721.

Iwata, H., Okada, K., Samreth, S., 2011. A note on the environmental Kuznets curve for CO_2. A pooled mean group approach. Applied Energy 88 (5), 1986‑1996.

Kaika, D., Zervas, E., 2013a. The Environmental Kuznets Curve (EKC) theory – Part A: Concept, Causes and the CO_2 emissions case. Energy Policy 62, 1392-1402.

Kaika, D., Zervas, E., 2013b. The Environmental Kuznets Curve (EKC) theory – Part B: Critical issues. Energy Policy 62, 1403-1411.

Kang, Y.-Q., Zhao, T., Yang, Y.-Y., 2016. Environmental Kuznets curve for CO_2 emissions in China: a spatial panel data approach. Ecological Indicators 63, 231‑239.

Kearsley, A., Riddel, M., 2010. A further inquiry into the pollution haven hypothesis and the environmental Kuznets curve. Ecological Economics, 69 (4), 905‑919.

Kondo, K., 2021. Testing for global spatial autocorrelation in Stata.

LeSage, J., 2008. An introduction to spatial econometrics. Revue d'économie industrielle 123, 19-44.

LeSage, J., 2014. What regionals scientists need to know about spatial econometrics. The Review of Regional Studies 44, 13-32.

Narayan, P.K., Narayan, S., 2010. Carbon dioxide emissions and economic growth. Panel data evidence from developing countries. Energy Policy 38 (1), 661‑666.

Maddison, D., 2006. Environmental Kuznets curves. A spatial econometric approach. Journal of Environmental Economics and Management 51 (2), 218‑230.

Moomaw, W., Unruh, G., 1997. Are environmental Kuznets curves misleading us? The case of CO_2 emissions. Environmental and Development Economics, 2 (4), 451‑463.

Panayotou, T., Peterson, A., Sachs, J., 2000. Is the Environmental Kuznets Curve driven by structural change? What extended time series may imply for developing countries. CAER II Discussion Paper 80.

Vincent, J., 1997. Testing for environmental Kuznets curves within a developing country. Environmental and Development Economics, 2 (4), 417-431.

Wang, Z., Ye, X., 2017. Re-examining environmental Kuznets curve for China's city-level carbon dioxide (CO2) emissions. Spatial Statistics 21, 377-389.

World Development Indicators. Data Bank. Retrieved September 20, 2021, from https://databank.worldbank.org/source/world-development-indicators

Zhang, C., Zhao, W., 2014. Panel estimation for income inequality and CO2 emissions. A regional analysis in China. Applied Energy 136 (C), 382‑392.

A. List of countries

TABLE 10: List of countries

Andorra	Gambia, The	Nicaragua
United Arab Emirates	Guinea	Netherlands
Afghanistan	Equatorial Guinea	Norway
Antigua and Barbuda	Greece	Nepal
Albania	Guatemala	Nauru
Armenia	Guinea-Bissau	New Zealand
Argentina	Guyana	Oman
Austria	Honduras	Panama
Australia	Croatia	Peru
Bosnia and Herzegovina	Haiti	Papua New Guinea
Barbados	Hungary	Philippines
Bangladesh	Indonesia	Pakistan
Belgium	Ireland	Poland
Burkina Faso	Israel	Portugal
Bulgaria	India	Paraguay
Bahrain	Iran, Islamic Rep.	Qatar
Burundi	Iceland	Romania
Benin	Italy	Serbia
Brunei Darussalam	Jamaica	Russian Federation
Bolivia	Jordan	Rwanda
Brazil	Japan	Saudi Arabia
Bahamas, The	Kenya	Solomon Islands
Bhutan	Kyrgyz Republic	Seychelles
Botswana	Cambodia	Sudan
Belarus	Kiribati	Sweden
Belize	Comoros	Singapore
Canada	St. Kitts and Nevis	Slovenia
Central African Republic	Korea, Rep.	Slovak Republic
Congo, Rep.	Kuwait	Sierra Leone
Switzerland	Kazakhstan	Senegal
Cote d'Ivoire	Lao PDR	Suriname
Chile	Lebanon	Sao Tome and Principe
Cameroon	St. Lucia	El Salvador
China	Liechtenstein	Syrian Arab Republic
Colombia	Sri Lanka	Eswatini
Costa Rica	Liberia	Chad
Cuba	Lesotho	Togo
Cabo Verde	Lithuania	Thailand
Cyprus	Luxembourg	Tajikistan
Czech Republic	Latvia	Tunisia
Germany	Libya	Tonga
Djibouti	Morocco	Turkey

Denmark	Moldova	Trinidad and Tobago
Dominica	Madagascar	Tuvalu
Dominican Republic	North Macedonia	Tanzania
Algeria	Mali	Uganda
Ecuador	Myanmar	United States
Estonia	Mongolia	Uruguay
Egypt, Arab Rep.	Mauritania	Uzbekistan
Spain	Malta	St. Vincent and the Grenadines
Ethiopia	Mauritius	Venezuela, RB
Finland	Maldives	Vietnam
Fiji	Malawi	Vanuatu
France	Mexico	Samoa
Gabon	Malaysia	Yemen, Rep.
United Kingdom	Mozambique	South Africa
Grenada	Namibia	Zambia
Georgia	Niger	Zimbabwe
Ghana	Nigeria	