UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA MAGISTRALE IN BUSINESS ADMINISTRATION - ECONOMIA E DIREZIONE AZIENDALE

TESI DI LAUREA

"WHAT DRIVES THE INCLINATION TO ACCEPT OR REJECT THE ENVIRONMENTAL KUZNETS CURVE HYPOTHESIS: A META-ANALYSIS OF EMPIRICAL STUDIES"

RELATORE:

CH.MO PROF. THOMAS BASSETTI

LAUREANDO: ENRICO FORATO

MATRICOLA N. 1105776

ANNO ACCADEMICO 2017 – 2018

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Firma dello studente

Contents

1. Introduction

The aim of this thesis is to investigate the factors that influence the relationship between economic growth and environmental degradation through empirical studies. In particular, this work aims at highlighting what drives the inclination to accept or reject the environmental Kuznets curve hypothesis, which displays an inverted U-shaped relationship between income and pollution.

This work combines the results of existing EKC empirical studies with the statistical approach of meta-analysis. This statistical tool allows examining systematic variation across the studies, in order to highlight any common patterns and sources of agreement and disagreement between them.

The issue of global warming is spreading across nowadays society and its impacts and effects are evident and tangible. Therefore, debates about sustainable economic growth have become central in everyday policy decisions. The EKC hypothesis has become a key topic when discussing about economic growth and environmental policies because it describes the change in pollution as a function of income. One of the main aims of international community is to understand the real relationship between economic growth and environmental quality.

The so-called environmental Kuznets curve (EKC) is a hypothesized relationship between various indicators of environmental degradation and income per capita (Stern, 2004). It is based on the findings of Nobel prize Simon Kuznets, who studied the relationship between income inequality and economic growth. He found that as long as economic growth rises, income inequality increases at first, but then, after reaching a turning point, starts its decreasing process, drawing an inverted U-shaped curve (Kuznets, 1955).

The same curve can be found empirically when, instead of income inequality and economic growth, we take into account economic growth and environmental degradation: the first variable is measured on the x-axis, the latter on the y-axis. The two variables are usually considered at a country-level, and data are usually collected on a time interval of several years, depending on the analyst's aim. With regards to economic growth, the most common way to measure it is to capture the yearly levels of income per capita of one nation. On the other side, the environmental degradation may be measured by the levels of a single pollutant over time – carbon dioxide, for instance.

After collecting data for both income per capita and a pollution index, the graph shows the relationship between the two variables and the pattern of the curve.
In general, when there is evidence of an inverted U-shaped relationship between income and pollution, researchers agree that "with further economic growth, services, improved technology and information diffusion limit the material basis of an economy and result in reduced environmental degradation" (Panayotou, 2003). The EKC theory suggests that countries experience a rise of pollution in the early stages of their economic growth; afterwards, for several internal and external reasons, the pollution levels released by these countries start a decreasing process even if their economies continue to grow. Consequently, the relationship between income per capita and pollution levels draw an inverted U-shaped curve.

In an ideal scenario, if the relationship between income and pollution always took the form of a hump-shaped curve, consequences would reveal that economic growth is compatible with environmental improvement. Therefore, there would be no need of regulating dirty emissions. By contrast, if the EKC is not real, the income-pollution relationship can draw either a monotonic increasing or an N-shaped curve. If the relationship is monotonic increasing, economic growth negatively affects the environmental quality. In the case of an N-shaped curve, pollution rises as a country grows, decreases when a medium economic wealth is reached, then experiences a new increase as income per capita continues to rise. In both cases, in the long run, economic growth negatively affects environmental quality. This implies that, if the government of a country aims at providing a sustainable economic growth in terms of pollutant emissions, environmental regulations are needed.
Furthermore, policy implications are completely different, depending on whether the EKC hypothesis is accepted or rejected at a country level. This leads to the necessity to deeply investigate which factors drive its acceptance or rejection. However, Panayotou (1993) suggests that developing countries should strengthen their environmental policies without caring of their future income-pollution relationship. As a matter of fact, developing countries may take several years or decades before experiencing a reduction in emissions.

The empirical studies seeking for the EKC experienced a significant rise in the last decades, due to several reasons. Firstly, the environmental quality has become a daily issue: as a matter of fact, climate change is showing its signs all around the Globe. Secondly, new developing countries are growing at extremely high rates and include the majority of the global population (i.e. BRIC – Brazil, Russia, India, China – countries). Therefore, we need to assess whether their economic development will lead to a somehow environmental improvement or they are just threatening the life quality of the whole world. On the other hand, it is also important to understand the relationship between economic growth and pollution of developed countries, in order to provide some hints for economic policies of the governments. Thus, if the EKC does not hold, it means that post-industrial countries are still polluting a lot. Moreover, even nowadays empirical findings are in some ways conflicting: not all researchers agree on whether EKC holds or not, and under which circumstances. A deeper analysis is therefore necessary in order to shed light on this phenomenon.

Furthermore, empirical EKC studies potentially analyse such a huge range of variables that it becomes extremely difficult to compare different studies one another. In detail, individual studies may differ in the type of countries and pollutants analysed, in the additional variables included in the model and much more. However, taking into consideration just these three differences among studies, it is possible to get a clear view of the problem. Indeed, EKC studies can either analyse one single country – chosen among more than 200 potential nations – or look for the income-pollution relationship across a group of countries. Regarding the pollutants, many pollution indicators are possible and it is clear that it is not easy to compare the income-pollution relationship of two studies with different pollutants analysed. Finally, not all EKC studies include all the possible variables in the models – i.e. trade, waste taxes, technical progress, education and more.

For these reasons, considering the continuous accumulation of EKC empirical studies, there is the need of finding common patterns across the studies and investigate what drives the inclination to accept or reject the EKC hypothesis.

In order to do that, this work gathers information from several EKC empirical studies and exploits a meta-analysis with the aim of highlighting the differences and similarities between them.

Meta-analysis is a statistical procedure used for combining data and results from multiple existing empirical studies. Indeed, it can be used to identify any common effect between the surveys; in alternative, when the effects vary across the studies, meta-analysis shall be used to identify the reasons of these variations. The purpose of applying a meta-analysis about empirical EKC studies is to establish statistical significance with multiple studies that display conflicting results.

With regards to the aim of this thesis, meta-analysis is a useful tool to deeply investigate the main factors that could potentially drive the inclination to accept or reject the environmental Kuznets curve hypothesis on the empirical EKC-related studies. Furthermore, the inconsistency of results can be quantified and analysed.

Practically speaking, each study represents one observation. The dataset is constructed as follows. The dependent variable captures individual studies results. In this work, each study can either accept or reject the environmental Kuznets curve. Thus, the dependent variable takes the form of a dummy variable with either 0 or 1 value, depending on the study result. The explanatory variables, by contrast, capture the differences in the ways of analysing the income-pollution relationship across the studies. For instance, they can be methodological and statistical choices, but also characteristics of the countries, pollutant indexes and other variables included in the models.

The main goal of such a meta-analytic approach is to be able to state sentences such as "collecting data for pollutant *X* rather than pollutant *Y* provides a higher inclination to accept the EKC hypothesis". By contrast, results could even show that "there is no statistical evidence that a change in the pollutant analysed influences the inclination to either accept or reject the EKC hypothesis". Meta-analysis is the tool that can synthesise results of existing studies and provide a clearer picture of the literature.

Following three former meta-analyses of Cavlovic et al. (2000), Li et al. (2007) and Koirala et al. (2011), this investigation collects data from 116 empirical EKC studies from 1998 to 2016. The dependent variable is a dummy variable that displays whether the study accepts or rejects the EKC hypothesis. The regressors consist of 10 independent variables, which control for methodological choices, data and country-specific characteristics. Data are analysed using a linear probability model and a logit model, both with fixed and random effects.

Results of this thesis are more or less in line with the prior meta-analyses: methodological choices and characteristics of the countries affect the acceptance or rejection of the EKC hypothesis. In detail, including more explanatory variables and countries with different degrees of development positively influences the probability of finding the EKC. Interestingly, this work finds that the pollution haven hypothesis is a driver of the environmental Kuznets curve. This means that developed countries that export their dirty activities to developing countries are more likely to exploit an inverted U-shaped relationship between income and pollution. Regarding anthropogenic gases, results show a small evidence that SO_2 positively affects the acceptance of the EKC hypothesis. However, the study confirms the earlier results about the fact that CO_2 does not influence the EKC.

The thesis is structured as follows. Chapter 1 begins with an introduction to the environmental Kuznets curve and meta-analysis, presenting also some of the most important contributes on both topics. Chapter 2 provides an overview of the existing theoretical and empirical literature on the EKC. Chapter 3 presents the data and the way variables are constructed. Chapter 4 deals the methodological approach, and specifically two models for data modelling are presented. Chapter 5 reports and discusses the results of the meta-analysis. Finally, chapter 6 provides some hints for future analyses and policy implications.

2. Literature review

2.1 Theoretical literature

The environmental Kuznets curve was first empirically found (Grossman and Krueger, 1991) and then theorized. By the way, after more than two decades, there is still debate on it and on what are the factors that affect the income-pollution relationship.

Several authors highlighted the logics behind the EKC relationship, in order to explain why we should expect a hump-shaped relationship between income and pollution.

Arrow et al. (1995) suggest that the shape of the curve may reflect the natural progression of a country's development: clean agrarian economies rapidly grow with industrial and polluting businesses until they reach cleaner and more effective ways to produce and grow.

Additionally, Kijima et al. (2010) state that in the early stages of industrialization, when a country is experiencing a fast industrial growth, the demand for a clean environment is low: pollution grows rapidly because people give higher priority to increasing material output rather than improving the environment. By contrast, when citizens achieve a sufficiently high standard of economic wealth – income per capita –, they attribute an increasing value to environmental negative outcomes; therefore, the willingness to pay for a cleaner country increases by a greater proportion than income, thus environmental quality starts improving (Roca, 2003).

Another possible explanation of the EKC phenomenon derives from the fact that advanced economies export their pollution-intensive production processes to developing countries, where there is high demand of economic growth and environmental quality is perceived as a minor concern. Andreoni and Levinson (2001) warn that this process of environmental improvement at the expenses of poorer countries is not indefinitely replicable: the world's poorest countries will not have other poorer countries to which they can export their dirty activities in order to get clean from pollution.

Technical progress has another possible role in causing the environmental Kuznets curve (Stokey, 1998). The reason is that a change in technology that leads to more effective production processes would generally reduce waste and pollution. Furthermore, new technologies – the so-called green technologies – need less carbon-related resources and are able to exploit clean and renewable energy sources.

Grossman and Krueger (1993) also suggest that environmental quality tends to worsen especially when the structure of economy changes from rural to industrial. On the other hand, when energy-intensive industry is replaced by knowledge-based technology-intensive industry, the overall pollution starts decreasing.
Moreover, services contribute to increasing the income per capita but do not significantly affect the pollution level (Grossman and Krueger, 1993).

Furthermore, other scholars developed economic models in order to demonstrate that the inverted U-shaped relationship between income and pollution can be derived also from a theoretical point of view – and not only with empirical data analysis.
Theoretical models that try to explain the EKC may be classified into static and dynamic models. One of the main differences between them is the time dimension: static models are time-independent, while dynamic models take into account changes among time. Moreover, while static models can be for instance production-based or utility-based models, the dynamic ones include also policy-based models (Kijima et al., 2010).
Following, two static and two dynamic models will be presented.

Regarding the static models, Lopez (1994) is one of the first scholars to consider the EKC relationship from a theoretical point of view. In the early 1990s there was very little agreement on the nature of the linkages between trade policy, economic growth and environmental degradation. One of the most important contributes of this model is that it considers the environment as a factor of production.
Lopez's paper is built on three main pillars: first, instead of focusing on interactions between economic growth and environmental degradation, the study considers only a one-way connection that goes from "growth" to environmental degradation, defining growth as "increases in the factors of production and technological change"; second, in order to measure trade policies, the model considers only trade liberalization in a small open developing country, defining liberalization as reduction in tariffs in the manufacturing industry; third, it aims at providing a systematic analysis of polar cases using highly simplified models.
Lopez basically develops a static model with a macroeconomic production function, starting from a function of capital (K), labor (L) and technology (t), $f(K, L, t)$. In order to consider the environment as a factor of production, he extends $f(\cdot)$ to show that industry output is also a function of the environmental factor of production.

In his analysis, Lopez further distinguishes whether the natural resource analysed has stock feedback productive effects and whether (and how) producers and consumers internalize externalities. The stock feedback productive effect of a natural resource means "whether the changes in the stock of the environmental factor play a role or not in the output" (Lopez, 1994): for example, forest resource has a stock feedback productive effect in the long run, while air quality does not.

This theoretical study shows that effects of economic growth on the environmental quality depend on both the nature of the resource stock effects on production and the internalization of these stock effects by producers. In detail, for resources that have a stock feedback productive effect, economic growth and trade liberalization decrease environmental degradation if producers internalize the stock effect (for instance, through government policies). On the other side, for resources that do not have stock effects on production, the relationship between economic growth and environmental degradation is positive and linear, no matter whether producers internalize or not. By the way, it is shown that a non-linear relationship between growth and pollution can be found if individual preferences are non-homothetic – meaning that economic growth increases the value that consumers attribute to environmental quality. Therefore, Lopez claims that, under certain conditions, an inverted U-shaped relationship between economic growth and pollution is possible and realistic.

Andreoni and Levinson (2001) contribute to the EKC theoretical framework with another static model: they present a theoretical model that tries to simplify some prior assumptions that lead to existing theoretical explanations for the EKC pattern.

Their model is a "simple and straightforward static model" (Andreoni and Levinson, 2001) that, in contrast with prior explanations, shows how the inverse U-shaped pattern does not require dynamics, predetermined patterns of economic growth, political institutions or externalities: an environmental Kuznets curve pattern can be derived just from the technological link between the consumption of a good and the abatement of its by-product.

An important innovation of this study is the concept of scale economies applied to the pollution abatement: "the more pollution there is before abatement, the less costly it is to abate one unit of that pollution" (Andreoni and Levinson, 2001).

They present the model starting from the simplifying assumption of an economy with only one person – eventually showing that the results do not change if more persons are added to the model. They draw the utility function of one person derived by the consumption of one private good and a bad called pollution. Also, they derive the pollution function as a positive function of consumption and a negative function of environmental effort. From these two functions, Andreoni and Levinson indicate under which general sufficient conditions an inverted U-shaped pollution-income relationship is real, showing that the EKC seems to hold under reasonable theoretical assumptions (such as considering the private good and the pollution as normal goods).

This model is a big contribution to the theoretical literature of the EKC because of its implications. In contrast with Lopez, Andreoni and Levinson argue that the EKC does not depend on externalities. In general, they suggest that the environmental Kuznets curve is reasonable and may result from simple features of the abatement technology, without depending on sophisticate models.

Regarding the dynamic models, one of the first contributions comes from John and Pecchenino (1994). It is an overlapping-generations model, where each individual takes decisions on the allocation of his income between goods consumption and pollution abatement effort.

John and Pecchenino consider an infinite-horizon economy with perfectly competitive firms and individuals living for two periods. Agents born at date t have preferences defined over consumption in old age ($t+1$) and an index of the quality of the environment when they consume, in $t+1$ as well. They also put in their model a production function described by capital stock and labour, and a function for environmental quality: according to John and Pecchenino, the evolution of environmental quality depends on both the consumption of goods and on the actions made by agents in order to improve it (which are costly).

Considering a developing economy, it is reasonable to imagine little capital available. Therefore, each agent in early generation spends no money for the environment, leading to a worse environmental quality. After a certain period of rime, agents in latter generation get richer (accumulation of capital stock) and start to invest in efforts to improve the environmental quality in their country. In this scenario, an inverted U-shaped relationship between income and environmental degradation is found. John and Pecchenino (1994) show also that an environmental Kuznets curve holds even when considering technological externalities in the model.

Another important dynamic model that analyses the income-pollution relationship comes from Brock and Taylor (2004). They present the empirical findings of the EKC with a theoretical model, the Solow one. Their model is known as “Green Solow Model”.

Brock and Taylor start from the fact that the EKC does exist, and create a model to demonstrate it and link it with the technological progress. They explain the so called rise-and-fall of emissions (inverted U-shape) with the role of technology introduced by Solow in 1956. The Green Solow model predicts that fast initial growth of production overwhelms progress in abatement, causing a period of initially rising emission levels; then, aggregate emissions continue to rise even though emissions per unit of output are falling; finally, technological progress in abatement overwhelms the slowing growth of output (Brock-Taylor, 2004).

The Green Solow model provides also an empirical analysis: the regression model borrows some Solow type regressors like population growth and savings rate, but it includes also a proxy for pollution abatement costs and for technological progress. In their work, Brock and Taylor demonstrate that their model predicts an inverted U-shaped relationship between income and environmental degradation.

One feature of the Green Solow model is that it is country specific; this is a limit for our analysis, since we need to assess what happens in an open economy, including cross-country variables such as trade and foreign direct investments (FDI), for instance. By the way, the Green Solow model is still an important contribution to the EKC literature, since it may explain part of the EKC empirical findings, claiming that technology has a big role in the self-adjustment path of emissions.

2.2 Empirical literature

The first empirical evidence of an inverted U-shaped relationship between economic growth and pollution emerged with Grossman and Krueger's (1991) study of the potential impacts of North American Free Trade Agreement (NAFTA). This was the beginning of a new idea about economic growth: if the EKC hypothesis was true, then economic growth may lead to some environmental improvement (Bhagwati, 1993).

In a couple of years, two additional working papers confirmed Grossman and Krueger's (1991) results of their pioneering study: using cross-country analyses, Shafik and Bandyopadhyay (1992) and Panayotou (1993) found that the relationship between income per capita and some pollution indicators follows a hump-shaped curve. Moreover, Panayotou (1993) first used the term "Kuznets" referring to the shape of the curve that revoked Kuznets' earlier studies (1955).

These findings shed new light on the debate about environmental economics. The first implication was that "growth could be a powerful way for improving environmental quality in developing countries" (Panayotou, 1993). Unfortunately, the existence of the EKC hypothesis is not unconditional, and still nowadays the validity of EKC is controversial: too many factors influence the shape of the curve and it is difficult to control over their effects.

As above mentioned, the environmental Kuznets curve is a phenomenon that can be observed from an empirical and statistical analysis.

In order to study the income-pollution relationship, most analysts apply reduced-form models, where the pollution term is a quadratic function of income per capita. However, it must be underlined that the relationship between income and pollution can draw different shapes. In detail, it may be either monotonic or non-monotonic. In the first case, it is usually a monotonic increasing relationship: as soon as income rises, pollution increases as well. When it is non-monotonic, most analysts claim to find an inverted U-shaped curve, even if some others suggest the existence of an N-shaped relationship between income and pollution.

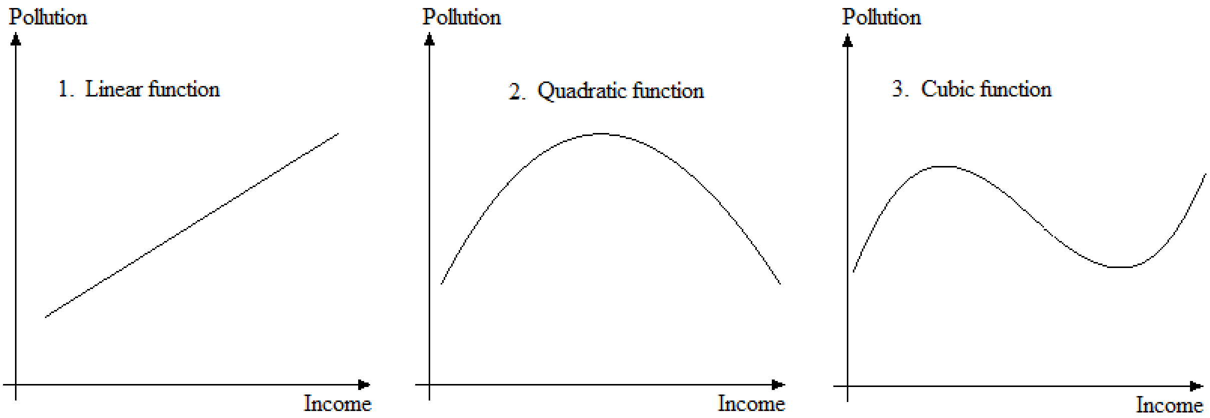
In order to graphically present these possible relationships between income per capita and environmental degradation, Abid (2016) proposes the following model:

$$y_{it} = \alpha_i + \beta_1 x_{it} + \beta_2 x^2_{it} + \beta_3 x^3_{it} + \varepsilon_{it} \tag{2.2.1}$$

Where y is the natural logarithm of environmental indicator, x is the natural logarithm of income and ε is the error term. Additionally, t represents time and i a geographic area responsible for economic and social policies – for instance, a country or region. Finally, β are the parameters to be estimated.

Model 2.2.1 provides the basis for understanding the various shapes that the relationship between economic growth and environmental degradation can take. Figure 2.2.1 shows the three main income-pollution relationships that are mostly found in literature.

Figure 2.2.1 The possible patterns of the income-pollution relationship.



Case 1 represents a linear function, where the parameter β_1 is positive and significant, and $\beta_2 = \beta_3 = 0$. This is the case of a monotonic increasing relationship between income and pollution. Case 2 is a quadratic function: namely, this is the graphic representation of an inverted U-shaped relationship – that is, the so-called environmental Kuznets curve. In this case, the parameter β_2 is negative and significant, while $\beta_1 > 0$ and $\beta_3 = 0$. Case 3 represents a cubic function, where $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$. The parameter β_3 of the cubic term captures another possible pattern of the income-pollution relationship, and in particular it draws an N-shaped curve (Abid, 2016).

Regarding the possible inverted U-shaped ad N-shaped patterns, it is important to highlight that the implications of these two different relationships are completely different, especially from the environmental policies point of view. Indeed, while the environmental Kuznets curve suggests that at some levels of income economic growth can reduce pollution, the N-shaped curve states that the environmental quality will experience a new decrease in the long run (De

Bruyn et al., 1998). If an N-shape between income and pollution is found, the main implication is that economic growth alone cannot provide a clean environment.
The reduced-form is helpful to evaluate the potential quadratic and cubic relationships between income and pollution but does not give any help in understanding which other variables could influence this relationship. For this reason, Kijima et al. (2010) propose the inclusion of other variables – and parameters – that can control for factors which are different from income and environmental indicators. These variables can be added to the model as $\beta_4 Z$ and so on. Each β will be estimated through the use of statistical methods.

One milestone empirical study about the environmental Kuznets curve is Schmalensee and Strocker's (1998). They contribute to the EKC theory with an economic model that investigates the relationship between CO_2 emissions per capita and per capita income. Claiming that carbon dioxide – together with other so-called greenhouse gasses – are likely responsible of Earth's climate becoming warmer, they project CO_2 emissions through 2050 using reduced-form models estimated with national-level panel data for the period 1950-2050. In the model, while using a flexible representation of the per capita CO_2-GDP relationship, they also include time and country fixed effects, handling forecast uncertainty explicitly in their projections. The strength of this model is that it is simple, compared with many other structural simulation models. For instance, the distributional underpinnings of their projections of CO_2 emissions are easy to trace. Schmalensee and Strocker also claim that their estimates provide a benchmark for the construction of simulation models.
The approach of the two scholars is, from one side to model similarities across countries in CO_2 emissions growth with economic development; from the other side, to use fixed effects for the levels of emissions across time and countries. For income and population growth, Schmalensee and Strocker use the same assumptions of the Intergovernmental Panel on Climate Change (IPCC), an intergovernmental body that, among its activities, generates scenarios of future greenhouse gas emissions until the year 2100.
Results from their econometric model are interesting for two reasons: at first, worldwide CO_2 emissions are higher than the IPCC forecasts in almost all the population and income growth scenario. Secondly, Schmalensee and Strocker find evidence of an inverted U-shaped relationship between CO_2 emissions per capita and per capita income.
This empirical study is a milestone of the EKC literature mainly for two reasons: its reference to fixed country and time effects, and its results. Indeed, on the basis of the IPCC

assumptions, Schmalensee and Strocker claim that the environmental Kuznets curve hypothesis holds for CO_2 emissions.
This study has thus become a point of reference for future empirical EKC-related studies.

In the late 1990s, the number of EKC studies continued to grow. As a consequence of these new studies about this complex phenomenon, analysts began to include more variables to the first reduced-form models – presented in equation 2.2.1 – in order to capture the effects of several factors on the income-pollution relationship. For instance, trade, foreign direct investments, technological improvements, environmental regulations are all variables that may influence the EKC shape and must be taken into account (Stern, 2004). In addition, the pollutant analysed was not only CO_2: several researchers started collecting data for water pollutants as well (NO_3 in Grossman and Krueger, 1995 and in Cole et al., 1997), and also for other air pollutants (SO_2 in Jie He, 2006 and Fodha and Zaghdoud, 2010).

The new millennium is still experiencing a continuous growth in the number of EKC studies. Many researchers have started seeking for the EKC in all types of countries, including developing countries, transition countries, developed countries, then collecting data for different time intervals, accounting for a small or big number of variables. The main reason of this increasing number of EKC studies is that literature is still controversial on its existence: results do not always show that the EKC is real, and some analysts warn about the difficulty of doing good econometrics, accounting for all possible variables (Stern, 2004).

Although it has been recognised that an EKC pattern appears as an empirical regularity, there is still debate on it, due to the fact that the income-pollution relationship can be affected by a great number of variables, and not all the researchers include the same ones in their regression models. In fact, sometimes it might even happen that similar studies lead to different results when different variables are considered in the models. For instance, both Halicioglu (2009) and Boluk (2015) look for an EKC relationship between income and CO_2 levels for Turkey. The only differences are that Halicioglu includes in his model a variable for trade openness, and, while Halicioglu collects data from 1960 to 2005, Boluk includes also data for the years 2005-2010. Results state that Boluk finds an EKC for Turkey while Halicioglu does not.
The main issue is that it is relatively simple to collect data and create a short model to find a relationship between income level and environmental degradation, but what is most difficult is to attribute the real impact of the possible causes of the EKC (Stern, 2004).

Following, some of the most important empirical studies results are described, together with the most common choices on what to consider when making an empirical analysis of the environmental Kuznets curve.

The earliest environmental Kuznets curves were simple quadratic functions of the levels of income; but this kind of regression allows levels of indicators to be lower than zero, therefore it is more appropriate to use a logarithmic dependent variable to avoid any negative results. The standard EKC regression model (Stern, 2004) is then:

$$\ln(E/P)_{it} = \alpha_i + \gamma_t + \beta_1 \ln(\mathrm{GDP}/P)_{it} + \beta_2(\ln(\mathrm{GDP}/P))^2_{it} + \varepsilon_{it} \qquad (2.2.2)$$

Where E represents the emissions, GDP the gross domestic product, P the population, and *ln* indicates the natural logarithm. The first two terms, instead, are intercept parameters that capture differences across countries and years. Hence, the basic assumption is that income elasticity is the same in all countries.

The model can be estimated with either time series, cross sectional or panel data. Time series data are usually used when the study is analysing one single country over a period of time; cross sectional data are used when observing several countries at a certain point of time; panel data, instead, is a combination of both time series and cross sectional data, and therefore it is the most common way to estimate the model. When deciding on which data to collect, really few researchers take cross sectional data into account. Jiang et al. (2014), for instance, use a cross sectional dataset from the Chinese manufacturing sector in order to investigate which factors are related to emission intensity in China. Their findings highlight an inverted U-shaped relationship between GDP per capita and three pollutants – SO_2, waste water, soot. By the way, almost all other authors use either time series data – when analysing one single country – or panel data – when computing data from more than one country together.

Many researchers try to make regressions including both the fixed and random effects models. For the fixed effects model, α_i and γ_t are regression parameters, for the random effects model, α_i and γ_t are components of the random disturbance. The issue with the random effects model is that α_i and γ_t and the explanatory variables cannot be correlated, otherwise the model will not be consistent. It is possible to compute the Hausman test in order to test the

inconsistency of the random-effects models – testing for correlation between the error and the regressor. Stern (2004) warns that many studies, after running the Hausman test and finding that the random-effects model is inconsistent, simply estimate the fixed-effects model without being aware of the differences between the two models. In detail, the parameters of the fixed effects model depends on the country and time effects of the sample of data (Hsiao, 1986), therefore they cannot be used to tell the future behaviour of other samples of data.
One example of how some regressions analyses are developed comes from Bernauer and Koubi (2009): they use both the fixed and random effects estimations, but since the Hausman test shows that the differences in the coefficients between the fixed and random effects estimations do not differ significantly, they report only the random effects estimations.

After having estimated the model, if an inverted U-shaped relationship between income and pollution is found, the turning point can be detected as:

$$\tau = \exp(-\beta_1/(2\beta_2)) \tag{2.2.3}$$

The turning point may vary across countries, depending on many factors, like the type of pollutant analysed, the type of country, the degree of trade openness, the number of years analysed.

Formula 2.2.2 is a reduced-form. This is helpful for directly measuring the impact of income on environmental degradation, but on the other side it is not possible to catch the underlying structural functions that lead to this kind of relationship (Grossman and Krueger, 1995). For this reason, some authors contribute in adding more variables to the basic formula. Following, some of the main variables included in the EKC analysis will be presented.

Still referring to formula 2.2.2, the term *E* refers to the pollution in general. Several types of pollutants exist in nature, and not all the researchers look for the same ones in their studies. Pollutants in general can be divided into three main categories: air, water and solid. Among air pollutants, we can further distinguish local from global ones. In the first group we find for example SO_2, CO, suspended particles (SPM) and N_2O. Lopez (1994) argues that local air pollutants are more likely to fit the environmental Kuznets curve hypothesis. The reason is that local impacts may be internalized by a stand-alone economy, leading to the institution of environmental policies to correct those negative externalities of economic growth. By the way, it is possible to look for EKC also studying other indicators of local pollution: Hilton and Levinson (1998) estimate a model for automotive lead emissions in 20 Countries. On the other side, the most common global pollutant analysed when looking for EKC is CO_2, since it is considered to be one of the major causes of the greenhouse effect. It is still controversial whether an EKC exists for carbon dioxide, and this may be another reason why the majority of studies are still looking for a relationship between income and CO_2. Ansuategi and Escapa (2002) argue that CO_2 emissions do not have local and recognizable effects on the environment.

The second category is composed of water pollutants. Usually, water quality indicators data are taken from rivers. One of the most used indicators of water pollution is NO_3 (Grossman and Krueger, 1995). Also Cole et al. (1997) estimate a model for NO_3, while Hettige et al. (2000), Managi et al. (2009) and Orubu et al. (2011) control for organic water pollutants, using Biochemical Oxygen Demand (BOD) as measure for it: BOD is defined as the amount of oxygen required by aquatic bacteria to break waste down (Orubu et al., 2011). Shafik (1994) finds an interesting N-shaped relationship between pollution and income. This result seems to be confirmed by Orubu et al. (2011) as well: they show empirical evidence of an N-shaped relationship between water and income in Africa.

The third category of pollutants should consider all the solid pollution, but very few literature exists in the branch of the environmental Kuznets curve. Instead, many studies include several different indicators of pollution, in order to assess whether the validity of EKC hypothesis is influenced by the different types of pollutants (Faehn and Bruvoll, 2009; Kearsley and Riddel, 2010; Jiang et al., 2014; Cheng, 2016). Asici and Acar's study (2016) is also noteworthy: "Ecological Footprint" is used as indicator for pollution (EF is the biologically productive area needed to provide for consumers use).

By the way, it must be cleared out that the equation "same pollutant analysed = same result" does not hold: for instance, Akbostancı et al. (2008) do not find EKC for SO_2 in Turkey, while Park and Lee (2011) find an EKC for SO_2 in Korea Republic.

Noteworthy is also the analysis of Fodha and Zaghdoud (2010): it studies the relationship between income and both SO_2 and CO_2 emissions in Tunisia. The results of the study show that an EKC exists for SO_2 but not for CO_2. Fodha and Zaghdoud explain this fact arguing that SO_2 is a local pollutant and CO_2 has to be considered as a global pollutant; therefore, only local pollutants display an inverted U-shaped curve. However, this explanation sounds a little bit tricky, because we still need to consider that Tunisia is a transition country, and also many

other analysts have already found an environmental Kuznets curve for CO_2 in a lot of other countries: for example, Bin Hitam and Binti Borhan (2012) show an EKC for CO_2 in Malaysia, and Shahbaz et al. (2015) prove the existence of an EKC for CO_2 in India.
In general, it seems that carbon dioxide is the most difficult pollutant to analyse under the EKC pattern, since the majority of studies and results are conflicting. Cole et al. (1997) for instance find EKC for CO_2 only in high levels of income. This result is also confirmed by Bouznit and Pablo Romero (2016), who show an EKC for Algeria with a turning point at 220% of the current income.

Since the early 1990s, some researchers (for instance Shafik, 1994) started adding a cubic term in order to control for possible N-shaped relationships between income and environmental degradation.
The cubic term is interesting, since several researchers disagree about the shape of the curve when wider time-ranges are studied (Levinson, 2000). The main idea of assessing the cubic term is that, if an N-shaped relationship is found, it means that a country's pollution increases as income grows, decreases after reaching the turning GDP point, and then starts a new increase as income continues to grow. It is still debated whether environmental degradation actually begins to decrease for good (inverted U-shaped relationship) or the decline is only temporary (N-shape). Ren et al. (2014) explain the N-shaped relationship between income and CO_2 in China with the effect of the rise of foreign direct investments in the long run. Clearly, the policy implications of an N-shaped rather than a U-shaped relationship between income and emissions are extremely different. When an N-shaped curve is found, one cannot just believe that the economy is able to self-protect from environmental degradation.

Many authors claim that international trade may influence the EKC pattern (Suri and Chapman, 1998; Cole, 2004). Some researchers define trade openness as the ratio of the value of total trade to real GDP (Baek et al., 2009), some others describe it as a function of imports and exports of goods (Bento and Moutinho, 2016).
International trade has an impact on pollution through four main effects: scale, composition, technique, technological (Ben Jebli and Ben Youssef, 2015). Scale effect links production to pollution, and measures the increase in environmental degradation that would occur if the economy were simply scaled up, keeping all the rest unchanged (Copeland and Taylor, 2004). Composition effect states that the production of some dirty goods may be moved to countries with less severe environmental regulations. Technique effect captures the pollution decrease

due to the available technologies. Technological effect posits that there is the possibility that international trade encourages R&D investments on green technologies as a reaction to competition. For this reason, trade influences environmental degradation, in a positive (i.e. scale effect) or negative (i.e. technological effect) way, and many researchers add to their models the variable of trade openness in order to capture its effects when seeking for an EKC (Tamazian and Rao, 2010; Kohler, 2013). Cole (2004) shows that trade openness in OECD countries is negatively related with environmental degradation, while Ben Jebli and Ben Youssef (2015) find that in Tunisia trade causes CO2 emissions.

These findings are consistent with the so called pollution haven hypothesis, which states that when developed countries look for externalizing their production sites, they pay attention to profits; since pollution control is costly, they externalize their production plants where pollution control costs are low; knowing this, developing countries, in order to attract foreign investments, set their environmental standards below socially-efficient levels (Levinson and Taylor, 2004). The hypothesis that weak environmental protection acts as a "pollution haven" for migrating industries is strictly related with the environmental Kuznets curve, and may affect it with a double effect: on one side, PHH is able to describe the increase of dirty emissions in developing countries as an inflow of waste from developed economies; on the other side, that transfer of waste could cause a reduction of pollution in the post-industrial countries, showing an inverted U-shaped relationship between income level and pollution (Copeland and Taylor, 2004).

Weak environmental regulation in developing countries generates either an export advantage or attracts FDI. Export advantage means that developing countries, thanks to lower abatement costs, are able to produce and export goods at a lower price; trade can then be a good proxy to control for this fact. Foreign direct investments are a measure of how many companies decide to relocate their production plants in developing regions, where environmental regulations are weaker. For this reason, not only trade but also the level of FDI is a good proxy to check whether pollution haven hypothesis can explain the environmental Kuznets curve or not. Empirical evidence of FDI affecting emission levels are found for example by Jie He (2006) for SO_2 in China, Pao and Tsai (2011) for CO_2 in BRIC – Brazil, Russia, India, China – countries and Atici (2012) for CO_2 in ASEAN – Association of Southeast Asian Nations – countries.

Again, this view is not universal. Tang and Tan (2015) control for FDI and do not find any evidence of the pollution haven hypothesis for Vietnam, and also Lee and Oh (2015) find insignificant results for PHH in China. However, Cheng (2016) deeply investigates and states that green technologies prevent pollution havens for China. These last findings may influence Chinese government's future environmental policies.

Empirical literature shows that researchers decide to compute the EKC analysis both on single and multiple countries, without any particular preference. Burnett et al. (2013), for instance, analyse one country (United States); however it is possible to compute a panel data of many countries, like in Aller et al. (2015), where 177 countries are taken into consideration. In between, we can find studies with different numbers of countries: the most common choice is to analyse a peer group of states with common characteristics. Cole (2004) finds EKC for 21 OECD countries, Narayan (2010) shows conflicting results when analysing 43 developing countries. Another possible common characteristic may be geographic: Orubu (2011) collects data for 47 African countries, finding that EKC holds for SPM but not for water. An interesting research has been conducted by Baek (2015): he seeks for EKC for 12 nuclear generating countries displaced all over the world, and finds out that there is no inverted U-shaped relationship between income and CO_2 for those countries; surprisingly, the relationship found is not even monotonically increasing, but monotonically decreasing. Sinha and Sen (2016) instead look for an EKC for CO_2 in the BRIC countries, and their results shows evidence of a hump-shaped relationship for Brazil and India but not for Russia and China, ceteris paribus.

Another important choice of what to analyse is regarding the type of country. There is a big difference between looking for EKC in a developing country rather than in a high income country. Theoretically speaking, studies for post-industrial countries are more likely to find an EKC than in developing ones, due to production decentralization, regulations and green technologies. Iwata et al. (2010), for instance, find an EKC for CO_2 in France, while Akbostancı et al. (2009) do not find EKC for CO_2 in Turkey (less economically developed country than France). What Ben Jebli and Ben Youssef (2015) find for Tunisia is noteworthy: they explain the results of the study (no EKC) with the fact that Tunisia may have not reached the turning point yet; they control also for renewable and non-renewable energy consumption and predict that "a continuous economic growth will encourage the use of renewable energy, leading to a reduction in per capita CO_2 emissions and an inverted U-shaped EKC" (Ben Jebli and Ben Youssef, 2015). Similar conclusions are found by Nasir and Rehman (2011): EKC does not hold in the short-run for carbon emissions in Pakistan, but it holds in the long-run, suggesting that EKC is a long-run phenomenon.

The empirical contributes to the EKC literature are many; the majority of authors agree on the fact that, if an environmental Kuznets curve for a particular country is found, it depends not only on what happens inside that country (Green Solow model), but also on the interactions with the rest of the world. In particular, trade has a big impact on it. Composition effect may lead to the pollution haven hypothesis, so that high regulation countries will lose part of the dirty industries and poor countries will get them (Dinda, 2004); technological effect states that international trade permits the diffusion of clean technology (Reppelin-Hill, 1999).

By the way, the EKC literature is not free of critiques. As a response to the mixed results of the EKC-related studies, some authors have started moving critiques to the general concepts represented by the environmental Kuznets curve.
Many of them derive from the fact that a great number of factors that influence the EKC are interdependent, therefore it is difficult to determine which ones may dominate and govern the shape of EKC (Ezzati et al., 2001).
Two main types of critiques are conceptual and methodological. Regarding the conceptual critique, it looks quite clear that the EKC cannot be generalized: too many factors and variables may influence the EKC pattern. Regarding the methodological critique, Dinda (2004) warns that there are no studies that consider as indicator for pollution the global environmental degradation level instead of country's pollution or emission levels. The use of Environmental Degradation Index (EDI) and an appropriate measure of economic development (e.g. Human Development Index) could allow the creation of a global EKC model (Jha and Bhanu Murth, 2003). Other methodological critiques point at the fact that "it is easy to do bad econometrics" (Stern, 2004): too little effort has been put to the statistical properties of the data used, therefore some results on the EKC models may not be reliable. He states that when he uses appropriate techniques he finds evidence against the EKC hypothesis (Perman and Stern, 2003). Furthermore, Stern (2004) warns that, even if a relationship between income and pollution is found, econometrics should further test whether it is valid or it is a spurious correlation.

Other authors' critiques start from the conflicting results, trying to highlight the reasons behind their occurrence. Regarding the mixed results when different pollutants are analysed, Lee and Oh (2015) claim that local pollutants are more likely to decrease when income rises; that is because their negative outcomes are local and citizens feel a bigger need to clean such pollutants up. Concerning the disagreements on the shape of the curve, Levinson (2000) suggests that for most countries and pollutants the real income-pollution relationship is N-shaped. In his opinion, the reason why most studies find evidence of a hump-shaped curve is that they do not analyse wider time intervals, otherwise they would likely find a new pollution increase.

Suri and Chapman (1998) add a noteworthy intuition to the EKC empirical literature, warning that a global pollution reduction is unlikely to occur. Indeed, they claim that if the pollution haven hypothesis is mainly responsible of the pollution reduction in wealthy nations, during their development, poorer countries will not be able to export their pollution abroad. For this reason, looking at the EKC at the country level does not solve the global issue of environmental degradation (Suri and Chapman, 1998).

Considering the significant increasing number of EKC studies, and taking into account the differences in variables considered in the models and the conflicting results, there is the possibility to exploit a meta-analysis collecting data and results from existing EKC studies, in order to try to deeply investigate into what is hidden behind the differences across the studies.

Meta-analysis is a statistical technique used for combining findings from multiple independent studies.

The basic idea of meta-analysis is that each individual study cannot reach the unknown truth because of limited data and statistical errors. For this reason, combining the results of different empirical studies may lead to the unknown truth thanks to the use of this statistical approach. The aim of meta-analysis is then to highlight all the hidden relationships among the studies that are not visible when they are considered individually.

Glass (1976) defines meta-analysis as the “analysis of the analyses”. It is a useful tool when dealing with a high number of studies that show mixed variables and results. Moreover, meta-analysis can provide more generalised and reliable results rather than a single study (Cavlovic et al., 2000).

Three notable meta-analyses about the EKC relationship have been conducted in the past: Cavlovic et al. (2000), Li et al. (2007) and Koirala et al. (2011).

Cavlovic et al. (2000) combine the results found in the 1990s by the first EKC analysts, testing the validity of the EKC hypothesis. Their contribute is significant for two reasons: firstly, for the first time 25 EKC-related studies are statistically summarized; secondly, the meta-analysis results are helpful to predict new income turning points (ITP) for eleven different pollutants. Results of this meta-analysis indicate that both methodological choices and pollutant types affect income turning points. In detail, those studies that investigate the empirical EKC relationship for developed countries tend to find lower ITPs. Moreover, the Cavlovic et al. suggest that studies including trade effects as an explanatory variable rather than income alone tend to find higher ITPs. This paper provides a deeper focus on the ITPs rather than the methodological variables that may drive the EKC. In general, Cavlovic et al. demonstrate that methodological choices of researchers can significantly influence results.

Li et al. (2007) extend Cavlovic's survey. In this meta-analysis, they collect data from 77 empirical EKC studies from 1992 to 2005. Li's contribute is from one side in the data expansion, from the other side in two new modelling approaches. This meta-analysis uses a multinomial logit model in order to investigate the general pattern in the EKC relationships, and then applies a tobit model in order to estimate the ITPs, especially for the greenhouse gases. Greenhouse gases are divided into two categories: "anthropogenic activity-related" and "chemically active". CO_2 is part of the first group, and no statistically significant evidence of the existence of an EKC is found for the anthropogenic activity-related gases, therefore including CO_2. Regarding the chemically active gases, instead, the EKC relationship is statistically significant, but the income turning point is seven times the world average GDP per capita. With regards to the methodological choices, Li et al. find that controlling for more variables, using longer time periods and gathering data for multiple countries all increase the probability of accepting the environmental Kuznets curve.

The third notable contribute to the EKC meta-analysis literature comes from Koirala et al. (2011). This study manages to further enlarge the dataset, collecting data from 103 empirical EKC studies from 1992 to 2009. Koirala uses cluster estimation techniques in order to correct for heterogeneity, and takes into consideration twelve different environmental quality indicators, including six air pollutants. Results show that the type of environmental quality indicator affects the presence of an EKC. The inverted U-shaped relationship with income seems to be confirmed for landscape degradation, water pollution, agricultural wastes and some indicators for air pollution. However, when considering CO_2, this study is consistent with Li et al. (2007): the income turning point is too high (ten times the world average GDP per capita) and holds only for extremely high growth scenarios. On the other side, Koirala et al. confirm that some methodological choices affect the probability of finding the EKC. In detail, the number of observations, the use of multiple countries, adding a control for the country development status and a variable for trade all positively influence the inclination to accept the EKC hypothesis.

3. Data

3.1 Dataset description

The aim of this work is to find the reasons why there is such an important mismatch of empirical results when scholars look for an EKC relationship between income and environmental degradation. To pursue this objective, I created a dataset collecting information from 116 EKC-related studies which belong to 29 journals, controlling for methodological and country-specific variables in order to highlight any possible features that may affect the acceptance or rejection of the environmental Kuznets curve hypothesis.

This work consists of a panel dataset composed of 134 observations from 116 EKC studies. Several studies estimate more than one model, taking into account different features – for instance, different pollutants or various time intervals. Therefore, each model with different features and results is counted as a single observation. This is the reason why the dataset is composed of a higher number of observations rather than studies.

Table 3.1 shows the publication years of the 116 studies. The oldest study that belongs to this dataset was published in 1998, while the earliest one was published in 2016.
It is important to notice that 67.24% of the studies I considered were published after 2011, when the most recent notable EKC meta-analysis was published by Koirala et al., collecting observations from 103 EKC-related studies.
Furthermore, the selection of data is different from the previous meta-analyses: I considered published studies only, while previous meta-analyses gathered data also from unpublished papers (Koirala et al., 2011), working manuscripts (Li et al., 2007) and book chapters (Li et al., 2007; Koirala et al., 2011).

Table 3.1: Publication years of the EKC studies.

Publication Year	Number of studies	Percent	Cum.
1998	2	1.72	1.72
2000	1	0.86	2.59
2001	1	0.86	3.45
2003	2	1.72	5.17
2004	1	0.86	6.03
2005	3	2.59	8.62
2006	2	1.72	10.34
2007	1	0.86	11.21
2008	5	4.31	15.52
2009	10	8.62	24.14
2010	3	2.59	26.72
2011	7	6.03	32.76
2012	8	6.90	39.66
2013	10	8.62	48.28
2014	13	11.21	59.48
2015	22	18.97	78.45
2016	25	21.55	100.00
Total	116	100.00	

3.2 Variables

The dependent variable I will use in this work shows whether the environmental Kuznets curve is found in each study. Therefore, the dependent variable named “EKC” is defined in the dataset as a dummy variable with either 1 (if the study finds evidence of an environmental Kuznets curve relationship between income and pollution) or 0 (if the study does not find evidence of any EKC) values.

The following independent variables collect information from how scholars study the EKC relationship. Variables descriptions are summed up in table 3.2.
The first independent variable in the model controls for the types of dataset used by scholars. In particular, this variable – named PANEL – captures the choice of running a panel data analysis rather than using time series or cross-sectional data.
Usually, in empirical EKC-related studies, panel data are suitable to evaluate information over time from several countries – or regions within the same country, while time series are mainly used when the study collects data over time for a single country or region.

Another variable measures the length of time coverage in data and it is named N_YEARS. If an environmental Kuznets curve is found in a given country, it will clearly show up over a certain time span: as long as income per capita grows, pollution at first increases, and then, after several years, will start its decreasing process. Therefore, it is possible to expect that the likelihood of finding an EKC increases when a higher number of years is taken into account.

It is normal to expect that looking for an EKC for China collecting data from 1950 to 1980 may lead to different results rather than collecting data from 1980 to 2010, even if the data coverage period is 30 years in both cases. This is one of the reasons why it is important to control also for the historic period considered in the study. In my dataset, the PERIOD variable has three categories: for data collected after 1970, for data collected after 1985 and for broader time intervals.

With the variable named N_COUNTRIES, I control for the number of countries analyzed in the EKC studies. Some researchers collect data for a single country, some others group more countries in their models. Usually, when there are several countries analysed in one study, they have something in common: they may be grouped together because of geographic or economic similarities, for example.

The acceptance or rejection of the EKC hypothesis may also be affected by the type of countries analysed (D_COUNTRIES). It is possible to expect that the likelihood of finding an EKC increases when both developing and developed countries are considered in one study. In fact, developing countries may show a positive relationship between income and pollution (which is the rising branch of an environmental Kuznets curve), while developed countries may draw the falling branch of the curve, because the relationship between income and pollution for developed countries could be negative, due to the fact that they have access to green technologies (Brock and Taylor, 2004) and care more about pollution (Lopez, 1994).
For my analysis, each country is attributed to a category: developing, transition, high income, based on the World Bank List of Economies (July 2016).

The POLLUTANTS variable captures the differences in pollutants analysed by the studies, to check whether the choice of pollutant may affect the relationship between income and environmental degradation. There are plenty of possible pollutant indexes, but I decided to control for the two most studied and controversial air pollutants, SO_2 and CO_2. It is expected that papers controlling for SO_2 are more likely to find an environmental Kuznets curve since SO_2 is a local pollutant and its local impacts may be internalized by a stand-alone economy (Lopez, 1994). On the other side, CO_2 is a global pollutant and it is responsible for the greenhouse effect: hence, it is extremely important to assess its influence on the environmental Kuznets curve, since global economic policies could be affected.

It is possible to expect that the more variables are added to a model (CONTROLS), the more accurate the results on the EKC will be. The reason lays on the fact that adding more controls could lead to a better understanding of interdependencies between the factors that influence the income-pollution relationship, avoiding on the other side the issue of omitted variables bias (Stern, 2004). This variable shows then the number of independent variables included in each EKC model.

Several regression models can be used to analyse panel data in an EKC-related study. In particular, Stern (2004) warns on the fact that the early EKC researchers preferred to run fixed effects models in order to find an inverted U-shaped relationship between income and pollution because random effects models used to give inconsistent results. Therefore, this variable – named METHOD – aims at testing whether there is still some influence of the fixed effects models on the probability of accepting the EKC hypothesis also in the recent studies.

The so-called FDI variable controls for both trade and foreign direct investments. Some theories suggest that trade affects environmental degradation, both in a positive (scale effect) and negative (technological effect) way (Ben Jebli and Ben Youssef, 2015). FDI, instead, measures a developing country's degree of attractiveness to Foreign direct investments. This variable shows which papers include in their models a controlling variable for FDI or for trade.

The so-called pollution haven hypothesis tries to explain part of the EKC theoretical framework: it suggests that developed countries get "cleaner" at the expense of developing countries, where dirty activities and waste are sent to (Levinson and Taylor, 2004). Therefore, the PHH variable provides information on which studies confirm the pollution haven hypothesis.

Table 3.2: Definitions and descriptions of independent variables.

Variable Definition	Variable Description
PANEL	Panel data indicator variable equals to 1 if the study uses panel data, otherwise 0.
N_YEARS	Data coverage period (years).
PERIOD	Time period variable equals to 1 if the time interval analysed begins after 1970, 2 if it begins after 1985, 3 if it covers a broader time period.
N_COUNTRIES	Number of countries analysed in the EKC study.
D_COUNTRIES	Country indicator variable equals to 1 if the EKC study includes developing countries only; 2 if it includes transition countries only; 3 if it includes high-income countries only; 4 if it includes countries that belong to more than one category.
POLLUTANTS	Pollutant indicator variable equals to 1 if the study analyses CO_2 only; 2 if it analyses SO_2 only; 3 if it analyses many pollutants at once, otherwise 4.
CONTROLS	Number of independent variables in the EKC model.
METHOD	Method indicator variable equals to 0 if the EKC study uses no panel data; 1 if it uses runs both fixed effects and random effects models; 2 if it runs RE only; 3 if it runs FE only, otherwise 4.
FDI	Trade indicator variable equals to 1 if the EKC study controls for FDI; 2 if it does not control, 3 if it controls for trade.
PHH	Policy indicator variable equals to 0 if the EKC study does not consider the pollution haven hypothesis; 1 if its findings do not support the PHH; 2 if there is empirical evidence of PHH.

3.3 Descriptive statistics

This study estimates the effects of methodological factors on the acceptance or rejection of the environmental Kuznets curve in 116 empirical studies. The dataset contains 134 observations from 29 EKC-related journals, one dependent variable and a set of 10 independent variables. Summary statistics are reported in table 3.3.

For what concerns the dependent variable named "EKC", it is interesting to note that 73% of EKC-related studies considered in my dataset finds empirical evidence of the environmental Kuznets curve. Even if most of the studies find empirical evidence of an EKC, there is still a lot of debate on its nature.

Regarding the independent variables, 69.4% of the studies collected PANEL data for their analyses.

The average data coverage period of the studies (N_YEARS) is 31 years, with a standard deviation of 26.7. The smallest data coverage period is 1 year and the highest one is 173. The third quartile indicates that 75% of the studies collected data from periods that are smaller than 39 years.

Regarding the PERIOD variable, almost half of the EKC studies collect data from recent years: 46% of the studies, in fact, have a period of analysis starting after 1985 (PERIOD_1), while 36% of them have a time interval starting after 1970 (PERIOD_2). A small portion of the studies (17% - PERIOD_3) consider broader time periods.

The average number of countries analysed is 20, with a standard deviation of 34.6. The N_COUNTRIES variable has a wide range of values, from one single country of analysis up to 177. It is important to note that the median is 4: compared with the average number of countries (20), this highlights the fact that a big portion of EKC studies tends to analyse a few – or even one – countries. In fact, the first quartile is still 1, which means that more than 25% of the studies in my dataset run an environmental Kuznets curve analysis at a single-country level.

EKC studies consider different types of countries in their models: 18.6% of them seek for the environmental Kuznets curve for developing countries (D_COUNTRIES_1), and 18.6% as well for high income countries (D_COUNTRIES_3), while the biggest portion of the studies in my dataset make their analyses for transition countries (44% - D_COUNTRIES_2), maybe because transition countries are the most controversial and interesting to analyse in terms of economic growth and pollution. Only 17.9% of the studies take into consideration countries belonging to different categories.

More than two-thirds of the EKC studies consider pollution levels of CO_2 (POLLUTANT_1), while only 11% analyse SO_2 (POLLUTANT_2). Such a big difference could be explained by the fact that researchers prefer to investigate the effects of income on CO_2 levels because it is a global pollutant and it is considered as one of the main causes of global warming.

The number of independent variables (CONTROLS) considered in each EKC-related study has an average of 5.37 and a standard deviation of 3.10, with a minimum value of 1 and a maximum number of independent variables of 19. The value of the third quartile – 7 – shows that the majority of researchers prefer to include few extra variables in their models, which is not really wise if one wants to investigate the deep causes of EKC.

For what concerns the method used to analyse panel data, 30.5% of the studies do not use panel data (METHOD_0), 6.7% run both fixed effects and random effects models finding no significant differences between them (METHOD_1), 11.9% run a RE model only (METHOD_2), while 35.1% run a FE model (METHOD_3). This suggests that the fixed effects regression is the most used in my dataset of EKC studies. 15.7% of them used other types of regression models (METHOD_4).

From the FDI variable, we learn that researchers tend to add few extra independent variables in their models. In fact, only 26.1% of the studies control for foreign direct investments (FDI_1), and 29.9% added a trade variable (FDI_3). Almost half of the studies (44.0%) in my dataset do not consider the idea of adding a trade variable (FDI_2).

Regarding the pollution haven hypothesis, 36.6% of the studies do not investigate it (PHH_0), 52.2% of them do not find empirical support for pollution haven hypothesis (PHH_1), while 19.4% find significant evidence of the presence of pollution havens (PHH_2).

Table 3.3: Summary statistics

Variable	N	Mean	sd	Min	p25	p50	p75	Max
EKC	134	.7313433	.4449242	0	0	1	1	1
PANEL	134	.6940299	.4625463	0	0	1	1	1
N_YEARS	134	31.26866	26.70394	1	16	29	39	173
PERIOD_1	134	.3656716	.4834252	0	0	0	1	1
PERIOD_2	134	.4626866	.5004767	0	0	0	1	1
PERIOD_3	134	.1716418	.3784837	0	0	0	0	1
N_COUNTRIES	134	20.52985	34.60164	1	1	4	25	177
D_COUNTRIES_1	134	.1865672	.3910255	0	0	0	0	1
D_COUNTRIES_2	134	.4477612	.4991295	0	0	0	1	1
D_COUNTRIES_3	134	.1865672	.3910255	0	0	0	0	1
D_COUNTRIES_4	134	.1791045	.3848786	0	0	0	0	1
POLLUTANTS_1	134	.6865672	.4656293	0	0	1	1	1
POLLUTANTS_2	134	.1119403	.3164761	0	0	0	0	1
POLLUTANTS_3	134	.0820896	.2755311	0	0	0	0	1
POLLUTANTS_4	134	.119403	.3254789	0	0	0	0	1
CONTROLS	134	5.365672	3.096302	1	3	4	7	19
METHOD_0	134	.3059701	.4625463	0	0	0	1	1
METHOD_1	134	.0671642	.2512454	0	0	0	0	1
METHOD_2	134	.119403	.3254789	0	0	0	0	1
METHOD_3	134	.3507463	.4789943	0	0	0	1	1
METHOD_4	134	.1567164	.3648973	0	0	0	0	1
FDI_1	134	.261194	.4409338	0	0	0	1	1
FDI_2	134	.4402985	.4982857	0	0	0	1	1
FDI_3	134	.2985075	.4593204	0	0	0	1	1
PHH_0	134	.3656716	.4834252	0	0	0	1	1
PHH_1	134	.5223881	.5013728	0	0	1	1	1
PHH_2	134	.1940299	.3969359	0	0	0	0	1

4. Methodology

4.1 Meta-analysis

Meta-analysis is the statistical approach I use to investigate which factors drive the inclination to accept or reject the environmental Kuznets curve hypothesis in the empirical literature.

Meta-analysis was defined by Glass in 1976 as the “analysis of the analyses”. It refers to the statistical way of analysing a large sample of results collected from empirical studies (Glass, 1976).

The main purpose of meta-analysis is to integrate the results of a group of studies on a defined subject and identify common patterns that may be hidden if the studies are only considered individually. Sources of agreement and disagreement among the results or any other interesting relationships may become known in the context of multiple studies.

Meta-analysis is a useful tool when the number of empirical studies grows at a very high rate and the field of analysis is wide and complex (Glass, 1976). For instance, empirical studies on the EKC hypothesis may suffer from differences in variables analysed and countless other factors.

The main difference between a primary analysis and a meta-analysis approach is that primary analysis is the original analysis of data in a research study, where the researcher collects primary data and applies a statistical method to study the phenomenon. A meta-analysis, instead, takes into consideration several research studies together: observations are individual study results, where one specific outcome is the dependent variable, and characteristics of individual studies compose the independent variables (Glass, 1976). Running a meta-analysis can therefore lead to multiple advantages: rather than relying on individual research studies, it can create the possibility of improving statistical inferences, aiming at providing tangible conclusions that may be able to influence policy decisions on multiple levels (Hunt, 1997).

On the other side, meta-analysis is not that easy to handle with, since many methodological concerns could lead to misleading results, like improper comparison of variables, publication bias due to wrong selection of papers, or heterogeneity in data (DeCoster, 2004; Koirala et al., 2011).

In this work, in order to avoid the earliest studies, which may contain econometric and methodological mistakes that can possibly lead to biases and wrong results, my dataset is composed of studies published during a time interval that goes from 1998 to 2016.

The first meta-analytic approach we have trace of is a paper published in 1904 by Karl Pearson, while the term "meta-analysis" was coined by Glass in 1976, who applies it to educational research. Lately, the use of meta-analysis spread among several different fields, from medical research to psychology, economics, biology and more.

The first meta-analysis on EKC-related studies was conducted by Cavlovic et al. (2000), with the aim of examining the relationship between economic growth and environmental degradation. Cavlovic et al. mainly control for several environmental pollution indexes and find significant evidence of EKC for CO_2, SO_2 and hazardous waste. Li et al. (2007) implement Cavlovic's study, adding new observations, new variables and new modelling approaches for the data analysis. With respect to anthropogenic activity-related gases, Li et al. find no statistical evidence of any EKC between income and environmental degradation. The research of Koirala et al. (2011) represents the latest meta-analysis conducted so far on the EKC-related studies. Koirala et al. highlight the presence of an environmental Kuznets curve for several liquid and solid local pollutants and for some air pollutants as well, but results for CO_2 show an inverted U-shape relationship only for unrealistic scenarios.

Since then, the number of empirical EKC-related studies has continued to grow at extremely high rates.

4.2 Linear regression

In order to investigate what drives the inclination to accept or reject the environmental Kuznets curve hypothesis, the first model I present is based on the following linear regression:

$$Y_i = \beta_0 + \beta_j X_{ij} + \varepsilon \tag{4.2.1}$$

For journal i (i=1…n) where X_j (j = 1…k, with k = 10) is a set of ten control variables presented in chapter 3. The dependent variable Y takes the form of a dummy variable, with either 0 or 1 value.

The above approach is also called Linear Probability Model (Soderbom, 2009). If we take expectations on both sides of the equation 4.2.1, we get:

$$E\ (Y \mid X_j; \beta_j) = \beta_0 + \beta_j X_j \tag{4.2.2}$$

Then, we can conclude that, just like under unconditional probabilities $E\ (Y) = \Pr\ (Y = 1)$, the conditional probability that Y equals one is equal to the conditional expected value of Y:

$$\Pr\ (Y = 1 \mid X) = E\ (Y \mid X_j; \beta_j)$$

$$\Pr\ (Y = 1 \mid X) = \beta_0 + \beta_j X_j \tag{4.2.3}$$

Equation 4.2.3 is a binary response model. Here, the probability of success (Y = 1) is a linear function of the explanatory variables in the vector X (Soderbom, 2009).

The main reasons why it is common to run a linear probability model lay on the fact that it is based on the well-known and widely used linear and multiple regression models, which are straightforward to estimate.

The LPM is also easy to interpret. As shown above, the predicted value of Y is the probability that the dependent variable equals one, given X. Furthermore, the parameter β_j indicates the change in the "success probability" Y = 1, resulting from a one-unit change of the independent variable, holding everything else fixed.

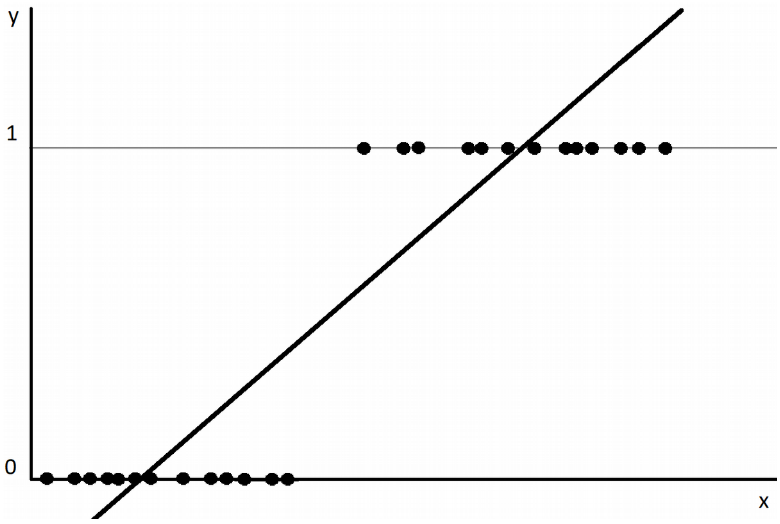
$$\Delta\Pr\ (Y = 1 \mid X) = \beta_j \Delta X_j \tag{4.2.4}$$

On the other side, the use of a linear probability model for a dummy dependent variable has to deal with some important drawbacks.

First of all, the LPM does not bound the predicted probability in a 0-1 unit interval. As shown in figure 4.2.1, the linear probability model may produce estimates which are either less than 0 or greater than 1. Since estimates are probabilities, it is hard to give them a meaning if they do not fall within a 0-1 range.

A related problem is that it does not make sense to state that a probability is "linearly related" to a continuous explanatory variable: if it were, the independent variable would inevitably lead P (Y = 1 | X) outside the unit interval. For this reason, since the relationship between the predictors and the dependent variable is likely to be non-linear, a regression line does not fit the data so accurately, as shown in figure 4.2.1.

Table 4.2.1: A linear regression model with a dummy dependent variable.



Another shortcoming of the likear probability model is that error terms are heteroskedastic by definition (Soderbom, 2009). Indeed, if Y takes the value of 1 or 0, then the variance of the errors will be:

$$var\ (\varepsilon) = P\ (1 - P)$$
$$var\ (\varepsilon) = (\beta_0 + \beta_j X_j)\ (1 - \beta_0 - \beta_j X_j) \qquad (4.2.5)$$

Which varies with the independent variables X_j. One way of overcoming this issue is to obtain estimates of the standard errors that are robust to heteroskedasticity.

An additional issue with the LPM has to deal with the distribution of the error term. Since the error term can only take two values, it has a binomial distribution instead of a normal one. Because of that, hypothesis testing in this model may be inaccurate (Maddala, 1983).

In the linear probability model of this work, in order to solve the problem of heteroskedasticity, I use heteroskedasticity-consistent robust standard error estimates.

Two different statistical models have been developed for inference from a collection of studies, the fixed effects and the random effects models (Konstantopoulos, 2006). In order to assess whether the most appropriate model is the fixed or the random effects one, it is not possible to simply run the common Hausman test (1978) when dealing with heteroskedasticity-consistent robust standard errors.

A test of fixed vs random effects can also be seen as a test of overidentifying restrictions (Sargan, 1975; Hansen, 1982). This test is known as Sargan-Hansen test. The fixed effects estimator uses the orthogonality conditions that the regressors are uncorrelated with the idiosyncratic error; the random effects estimator uses other additional orthogonality conditions that the regressors are uncorrelated with the group-specific error. These additional orthogonality conditions are the so-called overidentifying restrictions (Baum et al., 2006).

This overidentification test uses the artificial regression approach described by Arellano (1993), where a random effects equation is re-estimated, increased with additional variables consisting of the original regressors transformed into deviations-from-mean form. The test statistic is a Wald test of the significance of these additional regressors.

Under homoskedasticity, this test statistic is asymptotically equivalent to the above-mentioned Hausman test; but in addition to that, the Sargan-Hansen test can extend to heteroskedastic robust versions (Baum et al., 2006).

The null hypothesis of the Sargan-Hansen test is that the random effects model is consistent. A high p-value of the statistic would mean a failure in rejecting the null hypothesis. By contrast, a small p-value would lead to a rejection of the null hypothesis that the random effects is consistent, switching the preference to the fixed effects model.

4.3 Logistic regression

In this work I also estimate a logit model in order to analyze which factors influence the acceptance or rejection of the environmental Kuznets curve hypothesis. The logit model of the probability is given by the following equation:

$$\Pr(Y_i=1|X_i)=\frac{\exp\left(\beta_0+\beta_j X_{ij}\right)}{1+\exp\left(\beta_0+\beta_j X_{ij}\right)} \tag{4.3.1}$$

Where the dependent variable *Y* takes the form of a dummy variable, with either 0 or 1 value. $\Pr(Y_i=1|X_i)$ is the probability that the EKC category *Y* falls in alternative 1 (empirical evidence of the environmental Kuznets curve) for journal *i*.

X_j ($j = 1 \ldots k$, with $k = 10$) is a set of ten control variables presented in chapter 3, and β_0 and β_j are vectors of the parameters.

As can be observed from equation 4.3.1, the logit model does not show a linear relationship between the regressors and the dependent variable. Indeed, the above formula derives from a cumulative distribution function (Soderbom, 2009).

In order to highlight the differences between a linear probability model and the logit model, we first consider the following equation:

$$\Pr(Y = 1 \mid X) = G(\beta_0 + \beta_j X_j) \tag{4.3.2}$$

Where G is a function that takes values between zero and one: $0 < G(\beta_0 + \beta_j X_j) < 1$, for all real values of *X*. G is the so-called cumulative density function, monotonically increasing, with these characteristics:

$$\Pr(Y = 1 \mid X) \rightarrow 1 \text{ as } (\beta_0 + \beta_j X_j) \rightarrow +\infty$$

$$\Pr(Y = 1 \mid X) \rightarrow 0 \text{ as } (\beta_0 + \beta_j X_j) \rightarrow -\infty$$

Therefore, G cannot be a linear function. In literature, the most common non-linear function used for the logit model is the logistic distribution:

$$G=\frac{\exp\left(\beta_0+\beta_j X_j\right)}{1+\exp\left(\beta_0+\beta_j X_j\right)} \tag{4.3.3}$$

Which ranges between zero and one for all values of *X*.

Thus, the first formal difference between the linear probability model and the logit one is that the LPM assumes that the probability *P* is a linear function of the independent variables, while the logistic model assumes that the natural logarithm of the odds $P/(1-P)$ - probability of an event occurring divided by the probability that it will not occur – is a linear function of the independent variables:

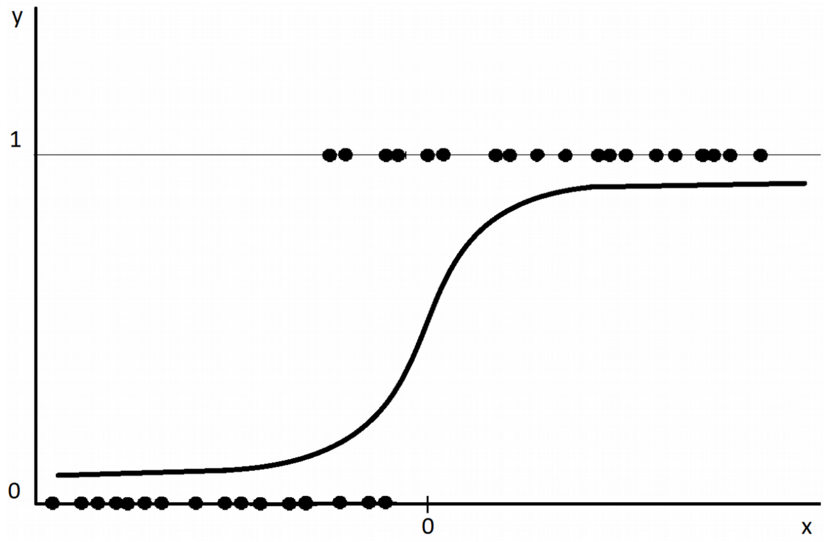
$$\ln\left[\frac{P(Y|X)}{1-P(Y|X)}\right]=\beta_0+\beta_j X_j \tag{4.3.4}$$

The logit model comes in response to the main issues of the linear probability model presented in chapter 4.2.

One of the most important concerns related to the linear probability model is boundedness: probabilities have to fit within a 0-1 range, while the LPM can predict values outside it. As previously shown, the logit model is able to bound the predicted probabilities within a 0-1 interval thanks to a transformation of a linear model, $\beta_0 + \beta_j X_{ij}$, which can draw values of the dependent variable from $-\infty$ to $+\infty$, into a model that can only range between 0 and 1 (figure 4.3.1).

When looking at figure 4.3.1, it is evident that the regression line is non-linear, giving a more realistic description of the data, with little change in the response variable at the extreme values that the independent variable can take. This is due to the fact that the relationship between the regressors and the dummy dependent variable is likely to be non-linear, and the logistic regression does not assume a linear relationship between the variables. The LPM assumes that the probability *P* is a linear function of the explanatory variables, while the logistic model assumes that the natural log of the odds $P/(1 - P)$ is a linear function of the regressors.

Table 4.3.1: The logit model with a dummy dependent variable.



Other advantages of the logit model lay on the fact that binary logistic regressions, by design, overcome many of the restrictive assumptions of linear regressions. For instance, the error terms do not need to be normally distributed and homogeneity of variance does not need to be satisfied. The condition of homoskedasticity is then not assumed in the model (Maddala, 1983).

On the other side, the logit model has its own drawbacks as well. Even if it better fits the data rather than the linear probability model, it is clear that it is not straightforward to interpret. This job would be much easier if the results of the logit model look like those from the LPM: the marginal effect of changing X, the probability of getting $Y = 1$. Instead, since the regressors are a linear function of the natural logarithm of the odds $P/(1 - P)$ - equation 4.3.4 -, there is not an intuitive link between the marginal effect and probability P, because it depends on the values of X.

The estimation of the logistic regression is made with the so-called maximum likelihood function.

Unlike with the linear probability model, with the logistic regression the use of heteroskedasticity-consistent robust standard error estimates is not needed, since the logistic model does not need to assume homoskedasticity.

Hausman (1978) shows a specification test in order to determine the best estimator between fixed and random effects.

Hausman's specification error test is a common tool used for testing whether the predictor variables are uncorrelated with the error term. H_0 is the hull hypothesis under which there is no misspecification; H_1 is the alternative hypothesis under which there is misspecification.

Let $\hat{\beta}_{RE}$ be the estimator for the random-effects model and $\hat{\beta}_{FE}$ the estimator for the fixed-effects model. Under the null hypothesis H_0 they are both consistent estimators, but under the alternative H_1 $\hat{\beta}_{RE}$ is inconsistent (O'Brien and Patacchini, 2003). If there is no correlation between the explanatory variables, there is no systematic difference between the two estimators $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$.

The Hausman test statistic H is a measure of the difference between the two estimates:

$$H = (\hat{\beta}_{RE} - \hat{\beta}_{FE})' [\mathrm{Var}(\hat{\beta}_{FE}) - \mathrm{Var}(\hat{\beta}_{RE}]^{-1} (\hat{\beta}_{RE} - \hat{\beta}_{FE}) \qquad (4.3.5)$$

Under the null hypothesis, H is distributed as a χ^2 with k degrees of freedom, that is the number of independent variables in the model.

If the p-value of the statistic is not significant ($p > 0.05$), the null hypothesis holds. On the other side, if the p-value of the statistic is significant ($p < 0.05$), it means that the random and fixed-effects models are different enough to reject the null hypothesis and prefer a fixed-effects model.

5. Results

This chapter presents the results of the data analysis estimated using a linear probability and a logit model.

Table 5.2 shows the outcomes of the two models that are used to compute the meta-analysis of the studies: the first two columns are dedicated to the linear regression, displaying results of the fixed effects and the random effects models, respectively. The last two columns show results of the logistic regression, with both fixed and random effects outcomes as well.

The linear probability model is run with estimates of the standard errors that are robust to heteroskedasticity, in order to overcome the structural issue of heteroskedasticity in the linear probability model.

The logit model with fixed effects shows results of 115 observations and 12 journals only, due to the fact that 17 groups (and 19 observations) have been dropped because of all positive or all negative outcomes.

R-squared and pseudo r-squared are displayed for linear probability and logit models, respectively.

In order to analyse and discuss the outcomes, two tests – one for the linear probability model and one for the logit model – are run.

Referring to the linear probability model with robust standard errors, to understand which of the two estimation techniques – fixed or random effects – is the most appropriate, a test for overidentifying restrictions is run (Sargan, 1975 and Hansen, 1982). The null hypothesis of the Sargan-Hansen test is that the difference in coefficients is not systematic and RE is consistent. Results indicate a χ^2 (10) = 32.158 and a p-value = 0.0004. This large value of the test statistic – and small p-value – means that the null hypothesis is rejected and therefore the random effects model is inconsistent, leading to a preference for the fixed effects model.

With regards to the logit model, the choice between the fixed and the random effects is helped by the Hausman test. The null hypothesis is that the random effects is the consistent and efficient estimator. The Hausman test on the logistic regression indicates a significant difference between the fixed and random effects estimators (χ^2 (19) = 35.57 and p-value = 0.0119). Therefore, I reject the null hypothesis and state that the results of the fixed effects model are more consistent on the logistic regression.

Furthermore, it is possible that results are misleading because of the correlation between the explanatory variables. For this reason, I propose a correlation analysis of the variables (table 5.1). Three notable correlations come to light: N_YEARS and PERIOD have a pairwise correlation coefficient of 0.67, PANEL and METHOD have a pairwise correlation coefficient of 0.88, and N_COUNTRIES and D_COUNTRIES' pairwise correlation coefficient is 0.62.

Table 5.1: Pairwise correlation coefficients of the variables.

	EKC	N_ YEARS	N_COU NTRIES	CON- TROLS	PANEL	PERIOD	ME- THOD	FDI	D_COU NTRIES	POLLU- TANTS	PHH
EKC	1.0000										
N_YEARS	-0.0268	1.0000									
N_COUN- TRIES	-0.0024	-0.1427	1.0000								
CONTRO LS	0.1701*	-0.2715*	0.1181	1.0000							
PANEL	-0.0005	-0.4042*	0.2987*	0.2677*	1.0000						
PERIOD	-0.0949	**0.6714***	-0.0999	-0.2618*	-0.4801*	1.0000					
METHOD	-0.0284	-0.3268*	0.2176*	0.2068*	**0.8751***	-0.3252*	1.0000				
FDI	-0.0148	0.0543	0.0198	0.0783	-0.1836*	0.1974*	-0.2183*	1.0000			
D_COUN- TRIES	0.1012	0.0538	**0.6151***	0.2108*	0.2261*	0.1326	0.1653	0.2772*	1.0000		
POLLU- TANTS	0.0133	-0.2321*	0.1658	0.2223*	0.2458*	-0.2553*	0.0482	0.2351*	0.1844*	1.0000	
PHH	0.1665	-0.0376	0.1337	0.0261	0.1259	-0.1283	0.0638	-0.3141*	0.0477	0.1405	1.0000

The correlation between N_YEARS and PERIOD is expectable. The way the PERIOD categorical variable is constructed – one category for time interval starting after 1970, one for time interval beginning after 1985, one for broader time periods – correlates it with the number of years almost "by definition". Regarding PANEL and METHOD variables, the high correlation means that the way dataset is constructed (PANEL variable) is linked to the way panel data are analysed (METHOD variable). When the study does not use panel data, both variables are 100% correlated. Therefore, a high correlation when considering all the studies is still expectable. Finally, correlation between the two variables N_COUNTRIES and D_COUNTRIES suggests that the choices of the number and types of countries to be analysed in one study are in some ways correlated.

One way to check whether a high correlation between the variables is responsible for misleading results is to drop the two variables individually and compute the regressions again. After the estimation of the models with all the variables and the results discussion, the last part of this chapter deals with the analysis of the potential effects of these variables correlation.

Table 5.2: Regression results of the linear probability model and logit model.

Variables	Linear Probability Model		Logit model	
	FE	RE	FE	RE
N_YEARS	0.001 (0.001)	0.001 (0.002)	-0.002 (0.015)	0.005 (0.014)
N_COUNTRIES	-0.005*** (0.002)	-0.002 (0.002)	-0.061** (0.025)	-0.011 (0.013)
CONTROLS	0.038** (0.015)	0.013 (0.014)	0.551** (0.229)	0.103 (0.100)
PANEL	-0.125 (0.101)	-0.184* (0.102)	-0.775 (0.897)	-0.995 (0.728)
PERIOD_2	-0.175* (0.096)	-0.149 (0.095)	-0.883 (0.895)	-0.883 (0.697)
PERIOD_3	-0.252 (0.167)	-0.251 (0.186)	-1.350 (1.238)	-1.323 (1.014)
METHOD_1	0.140 (0.243)	0.116 (0.183)	1.550 (1.991)	0.450 (1.203)
METHOD_2	-0.021 (0.185)	0.013 (0.176)	0.900 (1.488)	-0.057 (0.957)
METHOD_3	0.018 (0.133)	0.107 (0.117)	-0.091 (0.980)	0.531 (0.704)
FDI_2	-0.183 (0.166)	-0.159 (0.143)	-0.987 (1.421)	-0.827 (0.947)
FDI_3	0.074 (0.182)	-0.013 (0.147)	0.431 (1.127)	-0.118 (0.791)
D_COUNTRIES_2	-0.061 (0.113)	-0.031 (0.088)	-0.187 (0.865)	-0.242 (0.621)
D_COUNTRIES_3	0.089 (0.132)	0.171 (0.114)	0.889 (1.158)	0.929 (0.787)
D_COUNTRIES_4	0.353** (0.172)	0.186 (0.161)	5.548** (2.776)	1.073 (1.244)
POLLUTANTS_2	0.138* (0.074)	0.164** (0.078)	1.589 (1.342)	1.522 (1.143)
POLLUTANTS_3	-0.149 (0.141)	-0.170 (0.117)	-1.941 (1.472)	-0.952 (1.011)
POLLUTANTS_4	-0.023 (0.106)	-0.082 (0.105)	-0.370 (0.911)	-0.403 (0.737)
PHH_1	0.308 (0.207)	0.138 (0.203)	1.991 (1.546)	0.756 (1.026)
PHH_2	0.326** (0.130)	0.244*** (0.094)	2.159* (1.145)	1.812** (0.907)
CONS	0.582** (0.225)	0.776*** (0.188)	- -	1.262 (1.177)
Id (no. of journals)	29	29	12	29
Number of obs.	134	134	115	134
R2 / pseudo-R2	0.208	0.172	0.299	0.151

Note: Standard errors in parenthesis (robust s.e. for LPM), *p < 0.1; **p < 0.05; ***p < 0.01

The following discussion refers to the results of the linear and logistic regressions presented in table 5.2. Coefficients of the linear regression can be read as a change in the probability of finding empirical evidence of an EKC, due to the linear relationship between the regressors and the dummy dependent variable. On the other hand, a quantitative interpretation of the logistic regression results is not so intuitive. Therefore, it is only possible to make a qualitative analysis of these outcomes.

The estimated coefficient of the N_COUNTRIES variable is negative and significant at the 1% level in the linear probability model with fixed effects and negative and significant at the 5% level in the logit model with fixed effects. This suggests that the number of countries analysed in the EKC-related studies negatively affects the probability of accepting the environmental Kuznets curve hypothesis. In other words, according to the coefficient of the LPM, adding one additional country of analysis into an EKC study would lead to an increase in the probability of rejecting the EKC hypothesis by 0.5%. This result is in contrast with the earlier meta-analyses of Li et al. (2007) and Koirala et al. (2011): they show a positive relationship between the coverage of multiple countries and the acceptance of the EKC hypothesis.

The estimated coefficient of the CONTROLS variable is positive and significant at the 5% level both in the linear probability and logit models with fixed effects estimates. It means that those studies that add more control variables in their datasets – and not only one environmental indicator and the country's GDP level, for instance – are the ones where an EKC pattern shows up more likely. More technically speaking, for each extra control variable added in the model – that means, controlling for an additional factor – the probability of accepting the EKC hypothesis would increase by 3.8%, according to the LPM estimates. This finding could be seen from the perspective of the hypothesis that the EKC subject is way more complex than just a few variables; more precisely, a high number of factors influence the shape of the curve (Ezzati et al., 2001; Dinda, 2004). Also, this is in line with the earlier meta-analysis results (Li et al., 2007; Koirala et al., 2011): the number of observations in each study positively influences the probability of accepting the EKC hypothesis.

There is small evidence of significance on the estimated coefficient of the PANEL variable, and only in the linear probability model with RE. For this reason, I will not take it into account as a reliable result.

The estimated coefficient of PERIOD_2 is negative and significant at the 10% level in the linear probability model with FE. It suggests that collecting data starting after the year 1985 instead of collecting data starting after 1970 (base category) would increase the probability of rejecting the EKC hypothesis of 17.5%. However, there is no statistical evidence in the logit model.

The estimated coefficient of D_COUNTRIES_4 is positive and significant at the 5% level both in the linear probability and logit models with FE. The base category of this factor variable is "developing countries only" (chapter 3, table 3.2). Results indicate that collecting data from different types of countries (D_COUNTRIES_4) instead of developing countries only would lead to an increase in the probability of finding empirical evidence of EKC of 35.3%, according to the LPM coefficient. This finding is particularly interesting because it shows that, testing against the base category of developing countries, the categories of transition countries (D_COUNTRIES_2) and high-income countries (D_COUNTRIES_3) do not show any statistical significance, and only including countries from different categories in the model would increase of the probability of accepting the EKC hypothesis by 35.3%. A first hypothesis is that the EKC is a phenomenon that comes to light more likely when different types of countries are considered.

The estimated coefficient of POLLUTANTS_2 (SO_2) is positive and significant both in the two linear probability models, with FE and RE, at the 10% and 5% level respectively. Therefore, there is evidence – but only in the linear regression – that studies that collect data for SO_2 instead of CO_2 (base category) would more likely tend to accept the EKC hypothesis. With respect to the earlier meta-analyses (Cavlovic et al., 2000; Li et al., 2007; Koirala et al., 2011), this could be something new. The earlier findings, indeed, suggest that all anthropogenic gases do not show a U-shaped relationship with economic growth and, if they do, their income turning point is far beyond the realistic values. By contrast, this work shows some evidence in support of a positive relationship between SO_2 and the inclination of accepting the EKC hypothesis. SO_2 is a local pollutant, therefore its local impacts may be internalized by a stand-alone economy (Lopez, 1994). By the way, the statistical significance is found only in the linear probability model and not in the logit model. Therefore, future analyses will be needed to get a better understanding of this phenomenon.
On the other side, the fact that CO_2 is not a driver of the environmental Kuznets curve is highlighted by all meta-analyses conducted so far, including this work.

Regarding the pollution haven hypothesis, the estimated coefficient of PHH_2 is positive and significant in all the four models. In particular, it is significant at the 5% level in the LPM with fixed effects, at the 1% level in the LPM with random effects, at the 10% level in the logit model with FE and at the 5% level in the logit model with RE. These results indicate that the pollution haven hypothesis is a driver of the environmental Kuznets curve. In details, when there is evidence of pollution haven hypothesis (PHH_2) in one study – tested against the base category of "PHH not analysed in the study" - the probability of accepting the EKC hypothesis increases by 32.6%, according to the LPM coefficient. Again, these findings draw a link between two different categories of countries when speaking about EKC: the pollution haven hypothesis, indeed, states that developed countries get "cleaner" at the expense of the developing ones.

It is noteworthy that variables N_YEARS, METHOD and FDI do not show any statistical significance in all the models. Therefore, for the sample of studies analysed in this work, the number of years that each study considers for its EKC research does not influence the inclination to accept or reject the environmental Kuznets curve hypothesis, and that is in contrast with the findings of Li et al. (2007). The statistical methodology of analysing data is also not significant, and, finally, the inclusion of a trade or FDI variable in the studies does not influence the acceptance or rejection of the EKC hypothesis. This last result is in line with the meta-analysis of Koirala et al. (2011) but not with Li et al. (2007).

In the following part, the potential consequences of the variables correlation showed in table 5.1 will be analysed. Regarding the three pairs of variables with high correlation, for each pair, the two variables are dropped individually from the model. In this way, the new regressions without each variable highlight any noteworthy difference in significance of the other variables coefficients.
First of all, N_YEARS is dropped. The two models show no significant differences in the other variables coefficients and, in particular, coefficients and significance of the PERIOD variable remain overall unchanged. Then, PERIOD variable is dropped. Again, no big changes in the overall significance are detected. The N_YEARS variable was not statistically significant in the full model, and it does not gain any significance even without the PERIOD variable.

The same procedure is applied to the PANEL and METHOD variables. When each of the two variables is individually dropped, nothing notable happens to the other variables.

Finally, N_COUNTRIES and D_COUNTRIES variables are individually dropped, keeping all the other explanatory variables in the model. They both have significant coefficients, thus it is interesting to check what happens when one of the two variables is dropped.

Table 5.3 shows the regression results of the linear probability and logit models when N_COUNTRIES is dropped. What is notable is that D_COUNTRIES_4 loses significance in the coefficient and, more generally, there is no statistical significance in all the D_COUNTRIES categories.

Table 5.4 shows the regression results of the linear probability and logit models when D_COUNTRIES is dropped. It is possible to notice that the estimated coefficient of N_COUNTRIES is not significant for all the four estimated models.

The conclusion of this variables correlation analysis is that the number of countries and the type of countries are both important for the models, even if there is some correlation between them. Indeed, N_COUNTRIES variable is statistically significant only if the information about the type of countries is also considered in the models; on the other side, D_COUNTRIES_4 is significant only if the number of countries is included in the models as well.

Table 5.3: Regression results of LPM and logit model after dropping N_COUNTRIES.

Variables	Linear Probability Model		Logit Model	
	FE	RE	FE	RE
N_YEARS	0.001	0.001	0.005	0.006
	(0.001)	(0.002)	(0.015)	(0.014)
CONTROLS	0.037**	0.015	0.348**	0.115
	(0.016)	(0.014)	(0.162)	(0.098)
PANEL	-0.152	-0.193*	-0.858	-1.039
	(0.102)	(0.103)	(0.833)	(0.723)
PERIOD_2	-0.156	-0.119	-0.923	-0.768
	(0.109)	(0.098)	(0.825)	(0.677)
PERIOD_3	-0.276	-0.236	-1.493	-1.306
	(0.193)	(0.192)	(1.170)	(1.001)
METHOD_1	0.102	0.110	0.380	0.355
	(0.259)	(0.183)	(1.602)	(1.171)
METHOD_2	0.009	0.023	-0.024	-0.080
	(0.172)	(0.174)	(1.233)	(0.950)
METHOD_3	0.004	0.095	0.090	0.493
	(0.128)	(0.116)	(0.882)	(0.700)
FDI_2	-0.223	-0.169	-1.237	-0.879
	(0.165)	(0.141)	(1.324)	(0.944)
FDI_3	0.070	-0.007	0.419	-0.077
	(0.192)	(0.149)	(1.015)	(0.787)
D_COUNTRIES_2	-0.005	-0.008	-0.003	-0.097
	(0.112)	(0.082)	(0.787)	(0.595)
D_COUNTRIES_3	0.097	0.174	0.471	0.935
	(0.140)	(0.113)	(1.004)	(0.783)
D_COUNTRIES_4	0.021	0.060	0.494	0.356
	(0.225)	(0.128)	(1.277)	(0.853)
POLLUTANTS_2	0.172**	0.191***	1.492	1.647
	(0.079)	(0.072)	(1.220)	(1.131)
POLLUTANTS_3	-0.233**	-0.183	-1.756	-1.015
	(0.110)	(0.112)	(1.331)	(0.990)
POLLUTANTS_4	-0.075	-0.085	-0.373	-0.395
	(0.104)	(0.101)	(0.908)	(0.734)
PHH_1	0.321	0.146	2.045	0.817
	(0.222)	(0.204)	(1.463)	(1.024)
PHH_2	0.318**	0.239**	2.086**	1.805**
	(0.137)	(0.096)	(1.061)	(0.908)
CONS	0.548**	0.727***	-	0.978
	(0.229)	(0.184)	-	(1.120)
Id (no. of journals)	29	29	12	29
Number of obs.	134	134	115	134
R2 / pseudo-R2	0.174	0.151	0.218	0.147

Note: Standard errors in parenthesis (robust s.e. for LPM), *p < 0.1; **p < 0.05; ***p < 0.01

Table 5.4: Regression results of LPM and logit model after dropping D_COUNTRIES.

Variables	Linear Probability Model		Logit Model	
	FE	RE	FE	RE
N_YEARS	0.001	0.002	0.003	0.009
	(0.001)	(0.002)	(0.014)	(0.013)
N_COUNTRIES	-0.002	-0.000	-0.015	-0.002
	(0.002)	(0.001)	(0.011)	(0.008)
CONTROLS	0.038**	0.016	0.365**	0.110
	(0.016)	(0.015)	(0.165)	(0.097)
PANEL	-0.107	-0.143	-0.491	-0.734
	(0.093)	(0.097)	(0.839)	(0.699)
PERIOD_2	-0.156	-0.104	-0.947	-0.646
	(0.114)	(0.099)	(0.805)	(0.647)
PERIOD_3	-0.250	-0.210	-1.341	-1.113
	(0.155)	(0.168)	(1.172)	(0.964)
METHOD_1	0.120	0.108	1.043	0.393
	(0.239)	(0.185)	(1.812)	(1.152)
METHOD_2	-0.010	0.027	-0.055	-0.038
	(0.166)	(0.168)	(1.217)	(0.942)
METHOD_3	0.008	0.111	-0.070	0.562
	(0.130)	(0.124)	(0.878)	(0.687)
FDI_2	-0.201	-0.164	-1.203	-0.880
	(0.167)	(0.146)	(1.349)	(0.952)
FDI_3	0.128	0.063	0.881	0.314
	(0.163)	(0.134)	(1.084)	(0.750)
POLLUTANTS_2	0.141	0.166**	1.378	1.482
	(0.088)	(0.072)	(1.199)	(1.118)
POLLUTANTS_3	-0.209*	-0.177	-1.688	-0.994
	(0.118)	(0.117)	(1.402)	(0.977)
POLLUTANTS_4	-0.068	-0.065	-0.534	-0.290
	(0.087)	(0.098)	(0.916)	(0.722)
PHH_1	0.366*	0.201	2.549*	1.150
	(0.203)	(0.189)	(1.502)	(1.001)
PHH_2	0.370***	0.277***	2.607**	1.989**
	(0.117)	(0.083)	(1.145)	(0.907)
CONS	0.508***	0.633***	-	0.437
	(0.183)	(0.152)	-	(0.964)
Id (no. of journals)	29	29	12	29
Number of obs.	134	134	115	134
R2 / pseudo-R2	0.182	0.153	0.233	0.131

Note: Standard errors in parenthesis (robust s.e. for LPM), *p < 0.1; **p < 0.05; ***p < 0.01

6. Conclusion

The goal of this work is to investigate which factors influence the inclination to accept or reject the environmental Kuznets curve hypothesis in the empirical EKC-related studies.
With regards to the environmental Kuznets curve – a hypothesized inverted U-shaped relationship between income and environmental degradation –, the empirical literature shows mixed results among the studies. A part of it tends to find statistically significant evidence of a hump-shaped relationship between income and pollution: at a country level, as soon as income grows, at some point environmental degradation starts a decreasing process. Another part of literature finds a linear and positive relationship between income and pollution. Therefore, it is necessary to investigate which other factors drive the environmental Kuznets curve.

This work applies a meta-analysis of 116 empirical studies in order to highlight any possible systematic patterns across them. One of the main advantages of the meta-analysis is that its results are more reliable than those of a single study (Cavlovic et al., 2000).
In the regression models, the dependent variable takes the form of a dummy variable with either 0 or 1 values, depending on whether the study accepts or rejects the environmental Kuznets curve hypothesis. The 10 explanatory variables included in the model control for methodological choices, environmental quality and country-specific characteristics. Most of them derive from existing literature, while some others are brand new – for instance, the inclusion of factor variables for the type of countries and the pollution haven hypothesis.
Following the three prior meta-analyses (Cavlovic et al, 2000; Li et al., 2007; Koirala et al., 2011), this work uses a larger sample of studies and a particular focus on methodological variables, allowing a smaller focus on all the different types of pollutants analysed. Indeed, this study mainly controls for CO_2 and SO_2, which are the two most important anthropogenic gases.
For the data processing, this thesis uses a linear probability and a logit model, both with fixed and random effects.

Results of the meta-analysis indicate that several methodological choices and characteristics of the countries analysed have significant effects on finding evidence of the environmental Kuznets curve.
In detail, this study confirms the findings of the earlier meta-analyses, particularly on the fact that there is a positive relationship between the number of explanatory variables in each study and the probability of accepting the EKC hypothesis. This could suggest that some research analysts do not find the EKC because they do not control for enough variables. Therefore, EKC literature may be affected by methodological biases, as claimed by Stern (2004).
Data results also show that the number and type of countries analysed in the studies influence the EKC. In contrast with Li et al. (2007) and Koirala et al. (2011), this study finds a small and negative relationship between the number of countries and the probability of finding the EKC. On the other side, this work contributes in highlighting that also the type of countries is relevant: including both developing and developed countries in the same dataset provides statistically significant evidence of increasing the probability of finding the EKC.

However, one of the most interesting results regards the pollution haven hypothesis. For the first time, a meta-analysis on the EKC-related studies includes a variable for it, showing that the pollution haven hypothesis is one of the drivers of the environmental Kuznets curve. More technically speaking, there is statistical significance that accepting the pollution haven hypothesis increases the probability of accepting the EKC hypothesis. This relationship suggests that part of the EKC could be explained by the fact that developed countries decrease their emissions exporting dirty activities to developing countries. This is clearly not the best situation, since the issue of pollution simply moves from developed countries to developing ones, where regulations are less strict and demand for environmental quality is low.

In general, what appears from this meta-analysis is that the environmental Kuznets curve shows up more likely when more variables are taken into account and different countries interact with each other. These interdependencies highlight the fact that, even if measured at a country-level, the EKC phenomenon cannot be seen as just part of a closed economy. For this reason, future EKC studies should control for a high number of variables, in order to catch other potential factors that are able to influence the EKC, like technical progress and environmental policies changes, for instance.

Regarding the category of anthropogenic gases, there is some evidence that SO_2 positively affects the acceptance of the EKC hypothesis on the studies – but only in the linear probability model. A better understanding of the real role of SO_2 in the EKC pattern is crucial. If SO_2 really drives the EKC, it means that this local pollutant can positively influence the presence of an inverted U-shaped relationship between income and environmental degradation. Therefore, a stand-alone economy could be able to overcome the negative environmental outcomes of economic growth. By contrast, if SO_2 is not significant – as shown in the logit model –, local environmental regulations would be needed to provide sustainable economic development. Unfortunately, due to the mixed results between the linear probability and logit models on the SO_2, it is not possible to properly assess the role of SO_2 as a significant driver of the EKC hypothesis. However, this could be an interesting focus for future and more specific analyses.

With regards to the main global anthropogenic gas (CO_2), this study is consistent with the results of the prior meta-analyses (Cavlovic et al., 2000; Li et al., 2007; Koirala et al., 2011). Indeed, results confirm that CO_2 does not significantly influence the EKC; for this reason, considering the importance of CO_2 on the greenhouse effect and global warming, we can state that countries' CO_2 emissions will not "adjust" with economic growth alone. National and international policy actions, together with technological improvements, will likely be two cardinal tools for providing a sustainable economic growth in terms of carbon dioxide emissions for the future.

References

ABID, M., 2016. Impact of economic, financial, and institutional factors on CO_2 emissions: Evidence from Sub-Saharan Africa economies. *Utilities Policy (2016) 1-10.*

AKBOSTANCI, E., TURUT-ASIK, S., TUNC, G. I., 2009. The relationship between income and environment in Turkey: Is there an environmental Kuznets curve? *Energy Policy 37, 861–867.*

ALLER, C., DUCTOR, L., HERRERIAS, M. J., 2015. The world trade network and the environment. *Energy Economics 52, 55–68.*

ANDREONI, J., LEVINSON, A., 2001. The simple analytics of the Environmental Kuznets Curve. *Journal of Public Economics 80, 269-286.*

ANSUATEGI, A., ESCAPA, M., 2002. Economic growth and greenhouse gas emissions. *Ecological Economics 40, 23–37.*

ARELLANO, M., 1993. On the testing of correlated effects with panel data. *Journal of Econometrics, 59, 87-97*.

ARROW, K., BOLIN, B., COSTANZA, R., DASGUPTA, P., FOLKE, C., HOLLING, C. S., JANSSON, B. O., LEVIN, S., MALER, K. G., PERRINGS, C., PIMENTEL, D., 1995. Economic growth, carrying capacity, and the environment. *Science, 268, 520-521.*

ASICI, A. A., ACAR, S., 2016. Does income growth relocate ecological footprint? *Ecological Indicators 61, 707–714*.

ATICI, C,. 2012. Carbon emissions, trade liberalization, and the Japan–ASEAN interaction: A group-wise examination. *J. Japanese Int. Economies 26, 167–178.*

BAEK, J., CHO, Y., KOO, W. W., 2009. The environmental consequences of globalization: A country-specific time-series analysis. *Ecological Economics 68 (2009) 2255–2264.*

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

BAUM, C. F., SCHAFFER, M. E., STILLMAN, S., 2006. Enhanced routines for instrumental variables/GMM estimation and testing. *Boston College Economics, Working Paper No. 667*.

BEN JEBLI, M., BEN YOUSSEF, S., 2015. The environmental Kuznets curve, economic growth, renewable and non-renewable energy, and trade in Tunisia. *Renewable and Sustainable Energy Reviews 47, 173–185.*

BENTO, J. P. C., MOUTINHO, V., 2016. CO2 emissions, non-renewable and renewable electricity production, economic growth, and international trade in Italy. *Renewable and Sustainable Energy Reviews 55, 142–155*.

BERNAUER, T., KOUBI, V., 2009. Effects of political institutions on air quality. *Ecological Economics 68, 1355-1365.*

BHAGWATI, J., 1993. The case for free trade. *Scientific American, 42-49.*

BIN HITAM, M., BINTI BORHAN, H., 2012. FDI, Growth and the Environment: Impact on Quality of Life in Malaysia. *Procedia – Social and Behavioral Sciences 50, 333-342.*

BOLUK, G., MERT, M., 2015. The renewable energy, growth and environmental Kuznets curve in Turkey: An ARDL approach. *Renewable and Sustainable Energy Reviews 52, 587–595.*

BOUZNIT, M., PABLO ROMERO, M. 2016. CO2 emission and economic growth in Algeria. *Energy Policy 96, 93–104.*

BROCK, W., A., TAYLOR, M. S., 2004. The Green Solow Model. *University of Calgary, Discussion Paper 2004-09.*

BURNETT, J. W., BERGSTROM, J. C., DORFMAN, J. H., 2013. A spatial panel data approach to estimating U.S. State-level energy emissions. *Energy Economics 40, 396–404*.

CAVLOVIC, T., A., BAKER, K., H., BERRENS, R., P., GAWANDE, K., 2000. A Meta-Analysis of Environmental Kuznets Curve Studies. *Agricultural and Resource Economics Review 29/1 (April 2000) 32-42*.

CHENG, Z., 2016. The spatial correlation and interaction between manufacturing agglomeration and environmental pollution. *Ecological Indicators 61, 1024–1032.*

COLE, M. A., RAYNER, A. J., BATES, J. M., 1997. The environmental Kuznets curve: an empirical analysis. *Environment and Development Economics, 2, 401-416.*

COLE, M., A., 2004. Trade, the pollution haven hypothesis and the environmental Kuznets curve: examining the linkages. *Ecological Economics 48, 71– 81.*

COPELAND, B., R., TAYLOR, M. S., 2004. Trade, Growth and the Environment. *National Bureau of Economic Research, Working Paper 9823.*

DE BRUYN, S. M., VAN DEN BERGH, J. C. J. M., OPSCHOOR, J. B, 1998. Economic Growth and Emissions: reconsidering the empirical basis of environmental Kuznets curves. *Ecological Economics, vol. 25, 161-175.*

DECOSTER, J., 2004. Meta-analysis notes. *Department of Psychology, University of Alabama, 1-50. Retrieved on November, 16th, 2017 from http://www.stat-help.com/notes.html.*

DINDA, S., 2004. Environmental Kuznets Curve Hypothesis: A Survey. *Ecological Economics 49, 431– 455.*

EZZATI, M., SINGER, B. H., KAMMEN, D. M., 2001. Towards an Integrated Framework for Development and Environment Policy: The Dynamics of Environmental Kuznets Curves. *World Development Vol. 29, No. 8, 1421-1434*.

FAEHN, T., BRUVOLL, A., 2009. Richer and cleaner—At others' expense? *Resource and Energy Economics 31, 103–122.*

FODHA, M., ZAGHDOUD, O., 2010. Economic growth and pollutant emissions in Tunisia: An empirical analysis of the environmental Kuznets curve. *Energy Policy 38, 1150–1156.*

GLASS, G., V., 1976. Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher, Vol. 5, No. 10, 3-8.*

GROSSMAN, G. M., KRUEGER, A. B., 1991. Environmental Impacts of a North American Free Trade Agreement. *National Bureau of Economic Research, Working Paper 3914.*

GROSSMAN, G. M., KRUEGER, A. B., 1995. Economic Growth and the Environment. *Quarterly Journal of Economics, Vol. 110, No. 2 (May, 1995), 353-377.*

HALICIOGLU, F., 2009. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy 37, 1156–1164.*

HANSEN, L. P., 1982. Large sample properties of generalized method of moments estimators. *Econometrica, Vol. 50, No. 4 (July, 1982), 1029-1054.*

HAUSMAN, J., A., 1978. Specification Tests in Econometrics. *Econometrica, Volume 46, Issue 6 (Nov. 1978), 1251-1271.*

HETTIGE, H., MANI, M., WHEELER, D., 2000. Industrial pollution in economic development: the environmental Kuznets curve revisited. *Journal of Development Economics Vol. 62, 445–476.*

HILTON, F. G. H., Levinson, A. M., 1998. Factoring the environmental Kuznets curve: evidence from automotive lead emissions. J*ournal of Environmental Economics and Management, 35, 126-41.*

HSIAO, C., 1986. Analysis of Panel Data. *Cambridge University Press.*

HUNT, M., 1997. How science takes stock: the story of meta-analysis, *The Russell Sage Foundation, New York*.

IWATA, H., OKADA, K., SAMRETH, S., 2010. Empirical study on the environmental Kuznets curve for CO2 in France: The role of nuclear energy. *Energy Policy 38, 4057–4063*.

JHA, R., BHANU MURTH, K. V., 2003. An inverse global environmental Kuznets curve. *Journal of Comparative Economics 31, 352–368*.

JIANG, L., LIN, C., LIN, P., 2014. The determinants of pollution levels: Firm-level evidence from Chinese manufacturing. *Journal of Comparative Economics 42, 118–142.*

JIE HE, 2006. Pollution haven hypothesis and environmental impacts of foreign direct investment: The case of industrial emission of sulfur dioxide (SO2) in Chinese provinces. *Ecological Economics 60, 228-245.*

JOHN, A., PECCHENINO, R., 1994. An Overlapping Generations Model of Growth and the Environment. *The Economic Journal, Vol. 104, No. 427. (Nov., 1994), pp. 1393-1410.*

KEARSLEY, A., RIDDEL, M., 2010. A further inquiry into the Pollution Haven Hypothesis and the Environmental Kuznets Curve. *Ecological Economics 69, 905–919.*

KIJIMA, M., KATSUMASA, N., OHYAMA, A., 2010. Economic Models for the Environmental Kuznets Curve: a Survey. *Journal of Economic Dynamics & Control 34 (2010), 1187-1201.*

KOHLER, M., 2013. CO2 emissions, energy consumption, income and foreign trade: A South African perspective. *Energy Policy 63, 1042–1050.*

KOIRALA, B., S., LI, H., BERRENS, R., P., 2011. Further Investigation of Environmental Kuznets Curve and Meta-analysis. *International Journal of Ecological Economics & Statistics (IJESS) Year 2011, Vol. 22, No. S11.*

KONSTANTOPOULOS, S., 2006. Fixed and Mixed Effects Models in Meta-Analysis. *Northwestern University, Discussion Paper No. 2198*.

KUZNETS, S.. 1955. Economic Growth and Income Inequality. *American Economic Review 45 (March), 1–28.*

LEVINSON, A., 2000. The Ups and Downs of the Environmental Kuznets Curve. *The UCF/CentER conference on Environment, November 30-December 2, 2000, Orlando, FL.*

LEVINSON, A., TAYLOR, M. S., 2004. Unmasking the pollution haven effect. *National Bureau of Economic Research, Working Paper 10629*.

LEE, S., OH, D. W., 2015. Economic growth and the environment in China: Empirical evidence using prefecture level data. *China Economic Review 36, 73–85.*

LI, H., GRIJALVA, T., BERRENS, R., P., 2007. Economic growth and environmental quality: a meta-analysis of environmental Kuznets curve studies. *Economics Bulletin, Vol. 17, No. 5 pp. 1-11.*

LOPEZ, R., 1994. The Environment as a Factor of Production: the Effects of Economic Growth and Trade Liberalization. *Journal of Environmental Economics and Management 27, 163-184.*

MADDALA, G. S., 1983. Limited dependent and qualitative variables in econometrics. *Cambridge University Press, 257-291.*

MANAGI, S., HIBIKI, A., TSURUMI, T., 2009. Does trade openness improve environmental quality? *Journal of Environmental Economics and Management 58, 346–363.*

NARAYAN, P., K., NARAYAN, S., 2010. Carbon dioxide emissions and economic growth: Panel data evidence from developing countries. *Energy Policy 38, 661–666.*

NASIR, M., REHMAN, F., U., 2011. Environmental Kuznets Curve for carbon emissions in Pakistan: An empirical investigation. *Energy Policy 39, 1857–1864.*

O'BRIEN, R., PATACCHINI, E., 2003. The Hausman Test for Correlated Effects in Panel Data Models under Misspecification. *University of Southampton, Discussion Papers in Economics and Econometrics, 302*.

ORUBU, C., O., OMOTOR, D., G., 2011. Environmental quality and economic growth: Searching for environmental Kuznets curves for air and water pollutants in Africa. *Energy Policy 39, 4178–4188.*

PANAYOTOU, T., 1993. Empirical tests and policy analysis of environmental degradation at different stages of economic development. *Working Paper WP238, Technology and Employment Programme, International Labor Office, Geneva.*

PANAYOTOU, T., 2003. Economic Growth and the Environment. *United Nations Economic Commission for Europe, 1-49.*

PAO, H., T., TSAI, C., M., 2011. Multivariate Granger causality between CO2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): Evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. *Energy 36, 685- 693.*

PARK, S., LEE, Y., 2011. Regional model of EKC for air pollution: Evidence from the Republic of Korea. *Energy Policy 39, 5840–5849.*

PERMAN, R., STERN, D. I., 2003. Evidence from panel unit root and cointegration tests that the environmental Kuznets curve does not exist. *Australian Journal of Agricultural and Resource Economics, Vol. 47.*

REN, S., YUAN, B., MA, X., CHEN, X., 2014. The impact of international trade on China's industrial carbon emissions since its entry into WTO. *Energy Policy 69, 624-634.*

REPPELIN-HILL, V., 1999. Trade and Environment: an Empirical Analysis of the Technology Effect in the Steel Industry. *Journal of Environmental Economics and Management 38, 283–301*.

ROCA, J.,2003. Do individual preferences explain the environmental Kuznets curve? *Ecological Economics 45 (1), 3–10.*

SARGAN, J. D., 1975. Testing for misspecification after estimating using instrumental variables. *Mimeo. London School of Economics.*

SCHMALENSEE, R., STROCKER T. M., JUDSON, A., 1998. World carbon dioxide emissions: 1950–2050. *Review of Economics and Statistics 80,15–27.*

SHAFIK, N., 1994. Economic development and environmental quality: an econometric analysis. *Oxford Economic Papers, 46, 757-73.*

SHAFIK, N., BANDYOPADHYAY, S., 1992. Economic Growth and Environmental Quality: Time Series and Cross-Country Evidence. *Background Paper for the World Development Report 1992*.

SHAHBAZ, M., MALLICK, H., MAHALIK, M., LOGANATHAN N., 2015. Does globalization impede environmental quality in India? *Ecological Indicators 52, 379–393.*

SINHA, A., SEN, S., 2016. Atmospheric consequences of trade and human development: A case of BRIC countries. *Atmospheric Pollution Research 7, 980-989.*

SODERBOM, M., 2009. Applied Econometrics, Lecture 10: Binary Choice Models. *University of Gothenburg.*

STERN, D. I., 2004. The Rise and Fall of the Environmental Kuznets Curve. *World Development Vol. 32, No. 8, pp. 1419–1439.*

STOKEY, N. L., 1998. Are there limits to growth? *International Economic Review 39 (1), 1–31.*

SURI, V., CHAPMAN, D., 1998. Economic growth, trade and energy: implications for the environmental Kuznets curve. *Ecological Economics 25, 195-208*.

TAMAZIAN, A., RAO, B., B., 2010. Do economic, financial and institutional developments matter for environmental degradation? Evidence from transitional economies. *Energy Economics 32, 137–145.*

TANG, C., F., TAN, W., T., 2015. The impact of energy consumption, income and foreign direct investment on carbon dioxide emissions in Vietnam. *Energy 79, 447- 454.*

THE WORLD BANK GROUP, 2016. World Bank List of Economies. *July, 2016.*