International Journal of Energy Economics and Policy
Vol. 3, Special Issue, 2013, pp.52-61
ISSN: 2146-4553
www.econjournals.com

The Relationship between Environment and Income: Regression Spline Approach

Rabia Ece Omay
Anadolu University, Science Faculty, Department of Statistics,
Yunus Emre Campus, 26200, Eskisehir, Turkey.
Telephone: +902223350580/4683. Email: reayar@anadolu.edu.tr

ABSTRACT: This paper examines the relationship between economic growth and carbon dioxide (CO_2) emissions, the dominant contributor to the greenhouse effect, in Turkey. For time series data for the period 1980-2009 taking into account the Environmental Kuznets Curve (EKC) was modeling. The results show that there are N-shaped relationship between CO_2 emissions and economic growth, so these findings do not support EKC hypothesis with inverted-U shaped curve. At this point, because of the presence of a cubic polynomial, *regression spline* method is used as a different approach and a new model has been proposed. The model gives better results according to different model selection criteria.

Keywords: Environmental Kuznets curve; CO_2 emissions; regression spline; thin plate regression spline
JEL Classifications: C14; C22; Q53; Q56

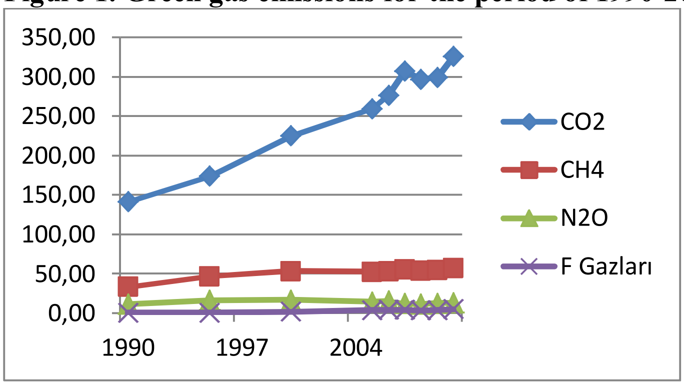
1. Introduction

One of the most remarkable environmental impacts is the emission of greenhouse gases, which combustions of fossil fuels. The most important of these gases is carbon dioxide, CO_2. All three coal, petroleum and natural gas, which is main fossil fuels, contains carbon. CO_2 accumulates in the atmosphere and causes significant detrimental impacts on the world's climate such as global warming, rises in the ocean levels, increased intensity of tropical storm, loses in biodiversity. (Sweeney, 2000: 21)

According to results of greenhouse gas emissions inventory in 2010 total greenhouse gas emissions in Turkey were estimated to be 401.9 million tons on a carbon equivalent basis. 81.23% of these emissions is CO_2. Greenhouse gas emissions for the year 2010 are shown in Figure 1.

In 2010, the largest proportion of greenhouse gas emissions on a carbon equivalent basis is energy-based emissions wit 71% and secondly is industry-based emissions with 13%. Waste and agricultural activities, respectively, have a share of 9% and 7%.

Figure 1. Green gas emissions for the period of 1990-2010



In 2010, considering energy-based CO_2 emissions, it is seen that 41% of total energy-based CO_2 emissions originate energy sector. Additionally, 20 % industry, 16% transportation and 23% other sectors of these originate. (TURKSTAT, 2011)

Total greenhouse gas emissions, equivalently CO_2, in 2010 compared to 1990 increased by 115%. While CO_2 equivalent emission per capita was calculated as the value of 5.51 tons/capita in 2010, this value was 3.39 tons/capita in 1990.

Some of the pollutants, which is also included CO2, are natural by-products of economic activities as electric consumption, operation of motor vehicles etc.. Increasing of economic activities will lead to increase emissions of these pollutants. However companies and owners can control these pollutants to some extent by choosing to be used technology. Cleaner technologies produce less pollution per product. A society becomes more wealthy, the members of the society can concentrate their demand of a more healthy and sustainable environment. And in this case governments may want to put into effect more hard environmental controls (Grossman ve Krueger, 1991). This relationship between income and environmental degradation is called *Environmental Kuznets Curve* (EKC) and the shape of the curve is inverted-U. According to EKC, environmental quality deteriorates in early stage of economic development/growth and improves in later stage as an economy develops. In other words, environmental pressure increases faster than income in the early stage of development and slows down relative to gross domestic products (GDP) growth in higher income levels (Dinda, 2004).

Economic growth affects the environmental quality in three different channels: scale effects, technological effects and composition effects (Grossman ve Krueger, 1991). While increasing part of EKC is related to scale effects, expression of decreasing part of EKC is used technological and composition effects (Saatçi and Dumrul, 2011).

Increasing output requires more input and thus more use of natural resource at production process. More input also implies more waste and emissions as by-product and hence these lead to degrade environmental quality. Therefore economic growth exhibits scale effects with a negative impact on environment. On the other hand, economic growth has also positive impact through composition effects. As income grows, structure of the economy tends to change and increases cleaner activities that produce less pollution. While structure of the economy changes from rural to urban or agricultural to industrial, environmental degradation starts to increase. However the environmental degradation starts to decrease with structural change from energy based technology to service and information based technology. Technological progress occurs with economic growth and the dirty and obsolete technologies are replaced by upgraded new and cleaner technology, which improves environmental quality. This is the technique effect of economic growth (Dinda, 2004).

This study aims to investigate existence of EKC hypothesis by using parametric and nonparametric regression models for the relationship economic growth and environmental quality in Turkey. For this aim, the second section of the study was made a brief literature review, especially including the studies dealing with air quality-income relation in Turkey. At the third section of the study, the methodology to be used to construct and test the models was explained. Data and models, belongs to empirical study about air quality-income relation in Turkey, were explain in the fourth section. The statistical findings and the results about these models were in the fifth and the sixth sections, respectively.

2. Literature Review

In the last decades, empirical studies about EKC have visibly grown up. In these studies, generally GDP per capita has been used as an income variable. All the same time, as an air pollution variables, Greenhouse gases, carbon dioxide $C0_2$, methane CH_4, nitrous oxide N_2O, hydro fluorocarbons HFCs, per fluorocarbons PFCs and sulfur hexafluoride SF_6, and sulfur dioxide SO_2, particulate matter, etc. have been used. In addition to these variables, some of the models have also included such as population density, energy consumption etc.. Using the different periods and variables for the specific countries leads to different results for the existence of EKC hypothesis and thresholds.

At this part of the study, the studies, which take into consideration CO_2 emissions, were mentioned for Turkey. In the studies time series and/or panel data were used and the parametric models were constructed for EKC hypotheses.

Akbostancı et al. (2009) examined relationship between air pollution and income in Turkey at two levels. At the first level, CO_2 emissions and GDP per capita relationship were examined by using time series for 1968-2003 periods. At the second level, the relationship between SO2, PM10 and GDP per capita and the population density were investigated for Turkish province. In the part of the study for

1992-2001 periods were used panel data. In both cases, the findings support the existence of the EKC hypothesis could not be reached.

Atıcı and Kurt (2007) examined interaction with Turkey's foreign trade and environment for the time series includes 1968-2000 period. They constructed a model includes CO2 emissions, water pollution, national income, total and agricultural export and import data as income and environmental quality variables. They concluded that the relationship between Turkey's national income and emissions confirmed EKC hypothesis.

Başar and Temurlenk (2007) investigated the relationship between CO_2 and income in Turkey by using the time series for the 1950-2000 period. They constructed the models separately for CO_2 from fossil fuel, solid fuel, fuel oil, CO_2 per capita and income. They found inverted N shape relationship between CO_2 per capita and income, and CO_2 from solid fuel and income. In the period for these variables, there was no evidence for the existence of the EKC in Turkey.

Gürlük and Karaer (2004) took into consideration the 1975-2000 period for time series data in Turkey. They constructed separately the regression models for relationship between GDP per capita and CO_2, SO_2 and NO_2 for the period. They found inverted U shaped relationship between CO_2 and GDP per capita, and NO_2 and GDP per capita. On the other hand, according to Gürlük and Karaer, N shaped relationship exists between SO_2 and GDP per capita. They concluded that emissions of CO_2 and NO_2 have started to decrease. However SO2 emissions have tended to increase over the next 20-year period.

Halıcıoğlu (2009) investigated the relationship between CO_2 emissions, energy consumption, foreign trade and income in Turkey by using EKC hypothesis. For this aim, she considered the time series data for the period of 1960-2005. She found that most significant variable affecting to CO_2 emissions in Turkey was income. Energy consumption and foreign trade followed that one.

Öztürk and Acaravcı (2010) examined the long run and causal relationship between CO_2 emissions, economic growth, energy consumption and employment ratio in Turkey over the period of 1698-2005. They concluded that CO_2 emissions and energy consumption did not cause economic growth but employment ratio caused economic growth in short run. In addition, EKC hypothesis was not valid in Turkish case.

Saatçi and Dumrul (2011) investigated the relationship between Gross National Product and CO_2 emissions through 1950-2007 period in Turkey. They found an inverted U-shaped relationship between these variables. In other words, they concluded that for the period the hypothesis of EKC was valid.

Yıldırım (2013) used panel data which included 20 countries for the period of 1990-2009. He investigated the impact of trade openness and income on pollution. In the study, CO_2 as an indicator of pollution, and the rate of total trade to GDP as an indicator of trade openness were used. He observed that there is a significant increase in pollution emissions due to trade of developing countries. In addition, for all country, also including Turkey, EKC hypothesis was confirmed and valid.

Of course, for the CO2-income relationship, the studies on the EKC hypothesis is not only limited to Turkey. Some of the most imported studies for the other countries are briefly mentioned bellow.

Egli (2004) examined the existence of the EKC hypothesis for Germany by using the time series data. Over the period of 1966-1999, Egli considered other pollutant, like as nitrogen oxide (NOx), carbon monoxide (CO), ammonia (NH3), methane (CH4), particulate matter (PM) and non-methane volatile organic compounds (NMVOC), as well as CO_2. He concluded that for nitrogen oxide and mostly for ammonia, an EKC or N-shaped pattern was found. Thus, for these two pollutants, the results of most cross-country studies could be confirmed. However, and more importantly, the other six pollutants did not show clear results.

Dinda (2004) reviewed the study which was about the EKC hypothesis. The study included an overview of the EKC literature, background history, conceptual insights, policy and the conceptual and methodological critique. He has indicated that the common point of all the studies is the assertion that the environmental quality deteriorates at the early stages of economic development or growth and subsequently improves at the later stages.

Nasir and Rehman (2011) investigated the relationship between carbon emissions, income, energy consumption, and foreign trade in Pakistan for the period 1972–2008. They found that there was a quadratic long-run relationship between carbon emissions and income, confirming the existence of Environmental Kuznets Curve for Pakistan. On the other hand the short-run results have denied the existence of the Environmental Kuznets Curve.

Shahbaz et al. (2013) probed the dynamic relationship between economic growth, energy consumption and CO_2 emissions for period of 1980–2010 in case of Romania. Their results confirmed long-run relationship between economic growth, energy consumption and energy pollutants. The empirical evidence revealed that Environmental Kuznets curve (EKC) was found both in long-and-short runs in Romania. In addition, energy consumption is major contributor to energy pollutants.

Ahmed and Long (2012), by considering EKC hypothesis, investigated the relationship between CO_2 emission, economic growth, energy consumption, trade liberalization and population density in Pakistan with yearly data from 1971 to 2008. The results supported the hypothesis both in short-run and long-run. They found the inverted U-shaped relationship between CO_2 emission and growth.

All studies have reviewed so far in this study has been only used parametric regression models. However non/semiparametric models have been used especially in the last decade. These studies are very few in the literature and the some of the most important ones are given bellow.

Wang (2011) proposed a flexible nonparametric method to analyze the EKC using spline smoothing techniques which was partially linear model by using kernel method. The proposed approach could be applied to testing the structural change for time series data from many other areas. In the cross-country EKC analysis, the cross-country panel data based on OECD-countries offered some empirical support for the EKC story. However, there was no statistically evidence of the inverted-U from the non-OECD countries. The EKC hypothesis was tested by performing a spline trend analysis. The trend analysis showed that the EKC varies across individual nations. There is no single EKC that fits all.

Dijkgraaf ve Vollebergh (2005) investigated the empirical results based on panel estimations of an inverted U-shaped relationship between per capita GDP and pollution. They constructed nonparametric regression model by using splines for the EKC. Using a new dataset for OECD countries on carbon dioxide emissions for the period 1960–1997, they found that the crucial assumption of homogeneity across countries was problematic. Their results challenged the existence of an overall EKC for carbon dioxide emissions.

Aldy (2005) examined the existence of EKC hypothesis for the relationship between CO2 and income in U.S.A. by using time series and panel data. In the study, he used both of the parametric and the nonparametric regression models with cubic splines. He emphasized that the production-based CO_2 EKC followed a standard inverted-U shape in the more flexible specifications; the consumption-based CO_2 EKC followed a peak and plateau shape.

3. Methods

***Environmental Kuznets Curve (EKC)*:** In general, the following model form, Eq.1, is used to test the various possible relationships between pollution level/environmental pressure and income:

$$y_{it} = \alpha_i + \beta_1 x_{it} + \beta_2 x_{it}^2 + \beta_3 x_{it}^3 + \beta_4 z_{it} + \varepsilon_{it} \quad (1)$$

where y is the environmental indicator, x is income and z are the other variables relation to environmental degradation. The subscripts i and t are the country and time, respectively. Eq.1 provides us to test different relationship forms between environment and income:

i) $\beta_1 = \beta_2 = \beta_3 = 0$ (no relationship between x and y)
ii) $\beta_1 > 0$ ve $\beta_2 = \beta_3 = 0$ (a linear relationship)
iii) $\beta_1 < 0$ ve $\beta_2 = \beta_3 = 0$ (a monotonic decreasing relationship)
iv) $\beta_1 > 0, \beta_2 < 0$ ve $\beta_3 = 0$ (an inverted U-shaped relationship, i.e. EKC)
v) $\beta_1 < 0, \beta_2 > 0$ ve $\beta_3 = 0$ (a U-shaped relationship)
vi) $\beta_1 > 0, \beta_2 < 0$ ve $\beta_3 > 0$ (a cubic polynomial or N-shaped curve)
vii) $\beta_1 < 0, \beta_2 > 0$ ve $\beta_3 < 0$ (an inverted-N shaped relationship)

Hence, Eq.1 identifies an EKC only if it refers to pattern (iv) as described below. From (iv), the turning point of EKC is obtained at $\frac{-\beta_1}{2\beta_2}$ (Dinda, 2004).

***Regression Splines*:** Basically a model with smooth function which has only one regressor is defined as follows:

$$y_i = f(x_i) + \varepsilon_i \quad (2)$$

where y_i is a response variable, x_i is a covariate, f is a smooth function and ε_i is the random variables which has independent identical Normal distribution with 0 mean and σ^2 variance. One way to estimate the function f, Eq.1 is defined as a parametric model. This operation can be performed by selecting *a base* in the functions space where *f* is also element of. Selection of such a base is equivalent to selecting some of the *base function*. If $b_j(x)$ is jth base function, then the function *f* is defined as follows:

$$f(x) = \sum_{j=1}^{q} b_j(x)\,\beta_j \qquad (3)$$

Substituting Eq.3 into Eq.2 clearly yields a linear model (Wood, 2006).

f can be explained by using a *cubic spline*. A cubic spline is a curve made up of sections of cubic polynomial joined together. They are continuous in value as well as first and second derivatives. The points at which the sections join are defined as the *knots* of the spline. For a conventional spline, every datum is handled as a knot, but for the regression splines the location of the knots must be chosen. Choosing the degree of model smoothness is kept under control by determining the basis dimension. A way of keeping under control the smoothness is to add penalty term to least squares. Therefore, the penalized regression spline fitting problem is to minimize

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\boldsymbol{\beta}^T\mathbf{S}\boldsymbol{\beta} \qquad (4)$$

with respect to $\boldsymbol{\beta}$. Where, $\mathbf{S}$ is a matrix of known coefficients and λ is a smoothing parameter. The problem of estimating the degree of smoothness for the model is now the problem of estimating the λ. The minimizer of Eq.4, the penalized least squares estimator of $\boldsymbol{\beta}$, is

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\,\mathbf{X} + \lambda\mathbf{S})^{-1}\mathbf{X}^T\mathbf{y}. \qquad (5)$$

For the model, influence matrix **A** can be written

$$\mathbf{A} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X} + \lambda\mathbf{S})^{-1}\mathbf{X}^T. \qquad (6)$$

Where, $\widehat{\boldsymbol{\mu}} = \mathbf{Ay}$.

For practical compitation, note that

$$\left\|\begin{bmatrix}\mathbf{y}\\ \mathbf{0}\end{bmatrix} - \begin{bmatrix}\mathbf{X}\\ \sqrt{\lambda}\mathbf{B}\end{bmatrix}\boldsymbol{\beta}\right\|^2 = \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\boldsymbol{\beta}^T\mathbf{S}\boldsymbol{\beta}, \qquad (7)$$

where $\mathbf{B}$ is the square root of $\mathbf{S}$ such that $\mathbf{B}^T\mathbf{B} = \mathbf{S}$. $\mathbf{B}$ can be obtained easily by Choleski decomposition, and once obtained the augmented least squares problem can be solved using orthogonal methods, in order to solve the penalized least squares problem and fit the model (Wood, 2006).

Thin Plate Regression Splines: Consider the problem of estimating the smooth function $g(\mathbf{x})$, from *n* observations $(y_i,\ \mathbf{x}_i)$ such that

$$y_i = g(\mathbf{x}_i) + \varepsilon_i \qquad (8)$$

where ε_i is a random error term and **x** is a *d*-vector. Thin plate spline smoothing estimates g by finding the function $\hat{f}$ minimazing

$$\|\mathbf{y} - \mathbf{f}\|^2 + \lambda J_{md}(f) \qquad (9)$$

where $\mathbf{y}$ is the vector of y_i and $\mathbf{f} = [f(\mathbf{x}_1), f(\mathbf{x}_2), \dots, f(\mathbf{x}_n)]^T$. $J_{md}(f)$ is a penalty functional measuring the wiggliness of f. It can be shown that the function minimizing Eq.9 has the form

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{n} \delta_i\eta_{md}\,(\|\mathbf{x} - \mathbf{x}_i\|) + \sum_{j=1}^{M} \alpha_j\emptyset_j(\mathbf{x}), \qquad (10)$$

Defining matrix **E** by $E_{ij} \equiv \eta_{md}\left(\|x_i - x_j\|\right)$, the thin plate spline fitting problems becomes,

$$\min\|\mathbf{y} - \mathbf{E}\boldsymbol{\delta} - \mathbf{T}\boldsymbol{\alpha}\|^2 + \lambda\boldsymbol{\delta}^T\mathbf{E}\boldsymbol{\delta}\,,\ \mathbf{T}^T\boldsymbol{\delta} = \mathbf{0} \qquad (11)$$

with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$. Where $\boldsymbol{\alpha}$ ve $\boldsymbol{\delta}$ are vectors of the coefficients to be estimated, $T_{ij} = \emptyset_j(x_i)$, $\emptyset_j$ are linearly independent ploynomials and η_{md} are the base functions.

Let $\mathbf{E} = \mathbf{UDU}^T$ be the eigen-decomposition of $\mathbf{E}$, where $\mathbf{D}$ is a diagonal matrix of eigenvalues of $\mathbf{E}$ arranged so that $|D_{i,i}| \geq |D_{i-1,i-1}|$ and the columns of $\mathbf{U}$ are the corresponding eigenvectors. Let $\mathbf{U}_k$ denote the matrix consisting of the first k columns of $\mathbf{U}$ and let $\mathbf{D}_k$ denote the top right $k \times k$ submatrix of $\mathbf{D}$. By writing $\boldsymbol{\delta} = \mathbf{U}_k\boldsymbol{\delta}_k$, Eq.11 becomes

$$\min\|\mathbf{y} - \boldsymbol{U}_k\,\mathbf{D}_k\,\boldsymbol{\alpha}_k - \boldsymbol{T\alpha}\|^2 + \lambda\boldsymbol{\delta}_k^T\mathbf{D}_k\boldsymbol{\delta}_k\,,\ \mathbf{T}^T\mathbf{U}_k\boldsymbol{\delta}_k = \mathbf{0} \qquad (12)$$

with respect to $\boldsymbol{\delta}_k$ and $\boldsymbol{\alpha}$. Any orthogonal column basis, $\mathbf{Z}_k$, is found such that $\mathbf{T}^T\mathbf{U}_k\mathbf{Z}_k = \mathbf{0}$. One way to do this is to form the QR decomposition of $\mathbf{U}_k^T\mathbf{T}$: the final M columns of the orthogonal factor give a $\mathbf{Z}_k$. Restricting $\boldsymbol{\delta}_k$, by writing $\boldsymbol{\delta}_k = \mathbf{Z}_k\widetilde{\boldsymbol{\delta}}$, yields the unconstrained problem that must be solved to fit the rank k approximation to the smoothing spline

$$\min\|\mathbf{y} - \boldsymbol{U}_k\,\mathbf{D}_k\,\boldsymbol{Z}_k\widetilde{\boldsymbol{\delta}} - \boldsymbol{T\alpha}\|^2 + \lambda\widetilde{\boldsymbol{\delta}}^T\mathbf{Z}_k^T\mathbf{D}_k\mathbf{Z}_k\widetilde{\boldsymbol{\delta}} \quad (13)$$

with respect to $\widetilde{\boldsymbol{\delta}}$ and $\boldsymbol{\alpha}$. Having fitted the model, evaluation of the spline at any point is easy by evaluating $\boldsymbol{\delta} = \mathbf{U}_k\mathbf{Z}_k\widetilde{\boldsymbol{\delta}}$ and by using Eq.10 (Wood, 2006).

4. Data and Models

Legal arrangements, agreements and protocols about improving to air quality in Turkey have started to come into force since 1980s: In 1986, regulation on the protection of air quality; in 2004, regulation on air pollution control from industry; in 2005, regulations on air pollution control from heating, exhaust gas emission control, gasoline and diesel fuel quality; in 2008, regulation of air quality assessment and management has been published.

Turkey in the international arena, in 1983, long-range cross-border air pollution agreement; in 1990, Montreal Protocol, the United Nations Framework Convention on Climate Change (UNFCCC); in 2009, Kyoto Protocol has signed.

In the study, a *time series* data for the period of 1980-2007 in Turkey are used by taking into account these legal arrangements. CO_2 emissions per capita, as an indicator of air pollution, and GDP per capita, as an indicator of economic growth, are considered. CO_2 per capita and GDP per capita in constant 2000 US dollars are from World Bank World Development Indicators (WDI) (CO_2: kg per 2000 US$ of GDP and GDP: constant 2000 US$).

The aim of the study is to investigate the existence of EKC hypothesis in Turkey for the 1980-2007 period is by using *regression spline* method. In this study, in order to investigate the impact of the income on the air pollution, impact of the GDP on the CO_2 has modeled with the cubic splines, on the basis of which a cubic polynomial, instead of the linear regression approach uses a cubic polynomial for the EKC. Existence of the EKC for Turkey examines with both of the parametric and nonparametric regression models. These parametric and nonparametric regression models are as follows, respectively.

$$C_t = \beta_0 + \beta_1 G_t + \beta_2 G_t^2 + \beta_3 G_t^3 + \varepsilon_t \quad (14)$$

$$C_t = \beta_0 + f(G_t) + \varepsilon_t \quad (15)$$

Where C is the CO_2 emissions (co2), G is the GDP per capita (gdp). Subscript t is the time and $t = 1, 2, \ldots, 28$. ε_t is the random variables which has independent identical Normal distribution with 0 mean and σ^2 variance. Both of the variables in the models are used by taking the logarithm.

5. Empirical Results

First, polynomial parametric linear model, Model1, in order to investigation impact of GDP on CO_2 was constructed for the period of 1980-2007. Some of the results about Model1 are listed in Table 1.

$$\text{Model 1: } co2_t = \beta_0 + \beta_1 gdp_t + \beta_2 gdp_t^2 + \beta_3 gdp_t^3 + \varepsilon_t \quad (16)$$

Table 1. Some of the statistics and estimation results for Model 1

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-728.8645	349.0423	-2.088	0.0476	*
gdp	259.9108	127.6549	2.036	0.0529	.
gdp^2	-30.8813	15.5562	-1.958	0.0587	.
gdp^3	1.2221	0.6313	1.935	0.0649	.
Adjusted R-squared: 0.6906 F-statistic: 21.09 on 3 and 24 DF, p-value: 6.513e-07 D-W Statistic: 1.952181 p-value: 0.428					

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Then, in order to test unit root Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for the null hypothesis that it is level or trend stationary was used. KPSS test has equally efficiency to diagnose existence of unit root for both of the linear and nonlinear time series (Kwiatkowski et al., 1992). The value of the KPSS test statistics and the corresponding probability value are found to be 0.5036 and 0.0406, respectively. Thus null hypothesis is rejected and hence suspicion of spurious regression arises. According to Granger and Newbold (1974), a good indicator for the spurious regression is R-squared and Durbin Watson d statistic: If the value of the R-squared grated than the Durbin Watson d statistic, estimated regression equation might be spurious regression. Table 1 shows that the suspicion of the spurious regression is quite weak, because of $R^2 < d$.

A way to avoid spurious regression, nonstationary series are made stationary by taking the difference. However, when the relationship between long-term rather than changes between two periods is examined, the way is not to be suitable. Hence in such a case, working with cointegrated series is more convenient. Thus existence of cointegration is tested by using Phillips-Ouliaris Cointegration test for the null hypothesis that it is not cointegrated (Phillips and Ouliaris, 1990). The value of the test statistic and the corresponding probability value are -27.0499 and 0.15, respectively. According to the result, null hypothesis cannot be rejected.

Adding a *time trend* in the model has been decided due to these results. Spurious regression problem (if present) can also be overcome by adding the time trend to the model. If the model does not include a time variable, the neglected variable problem can arise. The model which includes also the time variable, Eq.17, is constructed and some statistics and estimates are listed in Table 2.

Model 2: $co2_t = \beta_0 + \beta_1 gdp_t + \beta_2 gdp_t^2 + \beta_3 gdp_t^3 + time_t + \varepsilon_t$ (17)

In Table 2, it is seen that the parameters of the model except the time variable are statistically significant. For the Eq.17, Model 2, it is found that $\beta_1 > 0$, $\beta_2 < 0$ and $\beta_3 > 0$. This result states that the relationship between CO_2 emissions and GDP in Turkey follows an N-shaped pattern.

Table 2. Some of the statistics and estimation results for Model 2

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-7.528e+02	3.950e+02	-1.906	0.0692	.
gdp	2.687e+02	1.446e+02	1.858	0.0760	.
gdp^2	-3.196e+01	1.764e+01	-1.812	0.0831	.
gdp^3	1.266e+00	7.165e-01	1.767	0.0905	.
time	5.053e-04	3.599e-03	0.140	0.8896	
Adjusted R-squared: 0.6774 F-statistic: 15.18 on 4 and 23 DF, p-value: 3.3e-06 Deviance: 0.01896 D-W Statistic: 1.9396 p-value: 0.354					

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

At this stage of the study, on the ground that Eq.17, Model 2, is a cubic polynomial, modeling the relationship between CO_2 and GDP in Turkey by using cubic splines, which is based on the cubic polynomials, is suggested. For this aim, a semi parametric regression model, Eq.18, with regression spline is constructed. Some of the estimation results and statistics for the Eq.18, Model 3, are listed Table 3.

Model 3: $co2_t = \beta_0 + \beta_1 time_t + f(gdp_t) + \varepsilon_t$ (18)

In Table 3, statistically significant nonparametric effect of gdp points out, but time variable is not significant yet.

After the time variable is removed, the nonparametric regression model can be re-constructed. But this process can cause neglect of the time trend. Thus removing the time variable is not true way most of the time. By taking into consideration the variation of depending on the time of GDP, constructing a model that includes an interaction effect of these two variables could be solution for us. *Thin plate regression spline* model is in such a model, i.e. Eq.19. Some statistics and estimation results for Eq.19, Model 4, take part in Table 4.

Model 4: $co2_t = \beta_0 + f(gdp_t, time_t) + \varepsilon_t$ (19)

Table 3. Some of the statistics and estimation results for Model 3

Parametric coefficients					
Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept) time	-0.2825 0.0003	0.0473 0.0032	-5.971 0.093	7.75e-06 0.924	***
Approximate significance of smooth terms					
	edf	Ref. df	F value	p value	
s(gdp)	6.014	7.171	9.461	5.93e-06	***
Adjusted R-squared: 0.754 Deviance: 0.0125 D-W Statistic: 2.3266 p-value: 0.434					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

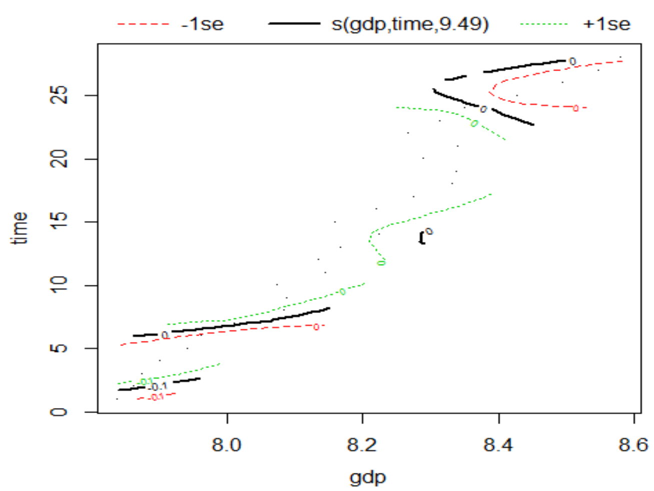
Table 4 shows that the nonparametric interaction effect of GDP and time variables on CO_2 is statistically significant, so this relationship shows a curvature. The curvature increases and decreases over time according to GDP and does not follow U-shaped pattern, which express EKC hypothesis. The curvature is observed in Figure 2. In Figure 2, the dashed lines/curves are the limit of 95% confidence and the solid line/curve is the thin plate regression spline curve.

Table 4. Some of the statistics and estimation results for Model 4

Parametric coefficients					
Coefficients	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-0.278095	0.004269	-65.14	<2e-16	***
Approximate significance of smooth terms					
	edf	Ref. df	F value	p value	
s(gdp, time)	9.492	12.05	9.125	6.71e-07	***
Adjusted R-squared: 0.80 Deviance: 0.0089 D-W Statistic: 2.6458 p-value: 0					

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 2. Thin plate regression spline model, Model 4



It is seen that, when the Model 2, 3 and 4 are compared, Model 4 has the highest adjusted R-squared value, 0.80, and the lowest deviance, 0.0089. Furthermore Myers et al. (2002) have mentioned the ratio of deviance to degrees of freedom as a rule of thumb. If the ratio closes to 1, the model is considered to be adequate. These ratios for Model 2, 3 and 4 are calculated as 0.00082, 0.00174 and 0.00074, respectively. According to these results, the most adequate model is thin plate regression spline model, Model 4.

6. Conclusion

In this study, the relationship between income and environmental quality in Turkey was examined in order to investigate existence of EKC hypothesis. For this aim, GDP per capita, as an income variable, and CO2 per capita, as an environmental quality variable, were considered by using the time series data over the 1980-2007 periods.

In order to investigate the impact of the income on the air pollution, impact of the GDP on the CO_2 has modeled with the cubic splines, on the basis of which a cubic polynomial, instead of the linear regression approach uses a cubic polynomial for the EKC. Existence of the EKC for Turkey examined with both of the parametric and the nonparametric regression models. Constructed models included a time variable in order to avoid spurious regression problem.

Looking at the problem in terms of the parametric regression model, Model 2, although the time variable was not statistically significant, N-shaped relationship between income and air quality was observed. Therefore, the finding of the parametric regression model did not support EKC hypothesis for this data set. The other studies like Akbostancı et al. (2009), Başar and Temurlenk (2007), Öztürk and Acaravcı (2010) and Yıldırım (2013) also found same results about relationship between income and air quality in Turkey.

Afterwards the relationship between air quality and income is modeled by using cubic splines, on the basis of which a cubic polynomial, because of the cubic polynomial EKC equation. For this aim, a semi parametric regression model, Model 3, with regression spline method was constructed. In the model statistically significant nonparametric effect of GDP pointed out, but time variable was not significant. Thus, by taking into consideration the variation of depending on the time of GDP, thin plate regression spline model, Model 4, that includes an interaction effect of these two variables was constructed. According to the model results, the nonparametric interaction effect of GDP and time variables on CO_2 is statistically significant, so this relationship shows a curvature. The curvature increases and decreases over time according to GDP and does not follow U-shaped pattern, which express EKC hypothesis.

As a result, both the parametric and the nonparametric models have shown that there was no evidence about existence of EKC hypothesis for the relationship between CO2 and GDP in Turkey over the 1980-2007 periods. On the other hand, for the modeling of this relationship, the thin plate regression spline model is better convenient than the others.

References

Akbostancı, E., Aşık, S.T., Tunç, G.İ. (2009). The relationship between income and environment in Turkey: Is there an environmental Kuznets curve? Energy Policy, 37, 861-867.

Aldy, J.E. (2005). An environmental Kuznets curve analysis of U.S. State-level carbon dioxide emissions, The Journal of Environment & Development, 14, 48-72.

Atıcı, C., Kurt, F. (2007). Turkey's foreign trade and environmental pollution: An environmental Kuznets curve approach, Turkish Journal of Agricultural Economics, 12(2), 61-69. (in Turkish)

Başar, S., Temurlenk, M.S. (2007). Environmental Kuznets Curve: An Empirical Analysis for Turkey, Journal of Economics and Administrative Sciences, 21(1), 1-12. (in Turkish)

Dijkgraaf, E., Vollebergh, H.R.J. (2005). A test for parameter homogeneity in CO_2 Panel EKC Estimations. Environmental & Resource Economics, 32, 229–239.

Dinda, S. (2004). Environmental Kuznets curve hypothesis: A survey", Ecological Economics, 49(4), 431–455.

Egli, H. (2004). The environmental Kuznets curve - evidence from time series data for Germany, Institute of Economic Research, ETH Zurich, Working Paper 03/28.

Granger, C.W.J., Newbold, P. (1974). Spurious regressions in econometrics, Journal of Econometrics 2, 111-120.

Grossman, G.M., Krueger, A.B. (1991). Environmental Impacts of the North American Free Trade Agreement, NBER Working Paper Series: 3914, 1-57.

Gürlük, S., Karaer, F. (2004). On the examination of the relation between economic growth and environmental pollution, Turkish Journal of Agricultural Economics, 10(2), 43-54. (in Turkish)

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

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics, 54, 159–178.

Myers, R.H., Montgomery, D.C., Vining, G.G. (2002). Generalized Linear Models with Applications in Engineering and the Science, Wiley Series in Probability and Statistics, USA.

Nasir, M., Rehman, F.U. (2011). Environmental Kuznets curve for carbon emissions in Pakistan: An empirical investigatin, Energy Policy, 39, 1857-1864.

Öztürk, İ., Acaravcı, A. (2010). CO2 emissions, energy consumption and economic growth in Turkey, Renewable and Sustainable Energy Reviews, 14, 3220–3225.

Phillips, P.C.B., Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration, Econometrica, 58, 165–193.

Saatçi, M., Dumrul, Y. (2011). The relationship between economic growth and environmental pollution: the estimation of environmental Kuznets curve with a cointegration analysis of structural breaks for Turkish economy, Journal of Erciyes University Faculty of Economics and Administrative Sciences, 66(37), 65-86. (in Turkish)

Shahbaz, M., Mutascu, M., Azim, P. (2013). Environmental Kuznets curve in Romania and the role of energy, Renewable and Sustainable Energy Reviews, 18, 165–173.

Sweeney, J.L. (2000). Economics of energy, Article 48, 4(9), 1-29. Available at: http://www.localenergy.org/pdfs/Document%20Library/Energy%20Economics.pdf

TURKSTAT, (2011). National Greenhouse Gas Emissions Inventory Report, Period: 1990-2009. (in Turkish)

Wang, L. (2011). A nonparametric analysis on the environmental Kuznets curve, Environmetrics, 22(3), 420-430.

Wood, S.N. (2006). Generalized additive models: An introduction with R, Chapman&Hall/CRC, USA.

World Bank World Development Indicators (WDI), http://data.worldbank.org/data-catalog/world-development-indicators

Yıldırım, B. (2013). Trade openness and climate change: Cross-country analysis, International Journal of Social Science, 6(1), 1611-1621.