

Kasioumi, Myrto; Stengos, Thanasēs

Article

The Environmental Kuznets Curve with recycling: A partially linear semiparametric approach

Journal of Risk and Financial Management

Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

Suggested Citation: Kasioumi, Myrto; Stengos, Thanasēs (2020) : The Environmental Kuznets Curve with recycling: A partially linear semiparametric approach, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 13, Iss. 11, pp. 1-26, <https://doi.org/10.3390/jrfm13110274>

This Version is available at:

<https://hdl.handle.net/10419/239353>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Article

The Environmental Kuznets Curve with Recycling: A Partially Linear Semiparametric Approach

Myrto Kasioumi and Thanasis Stengos *

Department of Economics and Finance, Gordon S. Lang School of Business and Economics, University of Guelph, Guelph, ON N1G 2W1, Canada; mkasioum@uoguelph.ca

* Correspondence: tstengos@uoguelph.ca

Received: 2 September 2020; Accepted: 3 November 2020; Published: 10 November 2020



Abstract: This paper is the first to study a comparatively new Environmental Kuznets Curve which traces empirically the relationship between environmental abatement and real GDP. Our model is a partial linear semi parametric model that allows for two way fixed effects to eliminate the bias arising from two sources. We use data for recycling and real GDP, for fifty states of the United States for the years between 1988 and 2017. We find evidence that this relationship is characterized by an increasing curve which confirms the existence of a J curve, a finding that agrees with the predictions from recent theoretical models.

Keywords: growth; recycling; Environmental Kuznets Curve (EKC); semi-parametric analysis

JEL Classification: C14; O47; Q53

1. Introduction

The Environmental Kuznets Curve (EKC) hypothesis has been one of the most debated research questions for both environmental economists as well as growth researchers. The initial Kuznets curve was first developed to capture the relationship between economic growth and income inequality, suggesting the presence of an inverse-U shape to characterize it. The Kuznets' hypothesis implies that in the early development of an economy, those who already have capital to invest can increase their wealth since investment opportunities are greater. Hence, as economic growth increases, income inequality also increases in poor countries while it is being reduced in wealthier countries and as a result, this rise of inequality will be reversed after the economy reaches a certain level of average income. Following the steps of the original curve, the EKC was introduced in the literature to depict the pattern that exists between economic growth and environmental degradation. As before, the relationship between these two variables is characterized by an inverse-U curve, suggesting that that as economic growth increases, environmental degradation follows the same trend until the economy reaches a specific level of average income after which environmental quality starts to improve. The concept of the EKC was first introduced in the literature by [Grossman and Krueger \(1991\)](#) who explored the empirical relationship between pollution and economic growth. They focused on the effects that trade liberalization and particularly the NAFTA agreement may have had on income and pollution on the participating countries. They concluded that this agreement could lead to larger scale of production and higher income which is accompanied by an increase in pollution as well.

In the present paper, we study the EKC from a new perspective, by examining the relationship between recycling and economic growth. We differentiate our analysis from the existing literature by substituting the environmental degradation with the transfers to recycling from total waste. It is important to examine this relationship not only because recycling—as a measure of pollution abatement—is the opposite of pollution, so we can test the validity of the EKC from a different

angle, but also because recycling measures both the improvement of the environmental quality today and tomorrow by reducing the waste pile of a country. Our analysis is based on a partial linear semi parametric panel data model using data for the 50 states of the US for the period between 1988 and 2017 and this is the first study that employs semiparametric techniques to explore this research problem.

The structure of our study is as follows: in Section 2 we discuss the existing literature of the EKC with recycling in addition to some fundamental papers for the basic EKC which will help us understand the theory that we will work on as well. In Section 3 we introduce the data and the model that we will use. In Section 4 we present our results, some tables from the regressions, and some graphs followed up by Section 5 in which we present our conclusions.

2. Literature Review

Very little research has been done examining the relationship between GDP per capita and recycling, hence the literature of the EKC with recycling is a fairly new topic and apart from a few theoretical models there are no studies that examine this relationship empirically. Thus, we will first focus our attention to the literature of the standard EKC with pollution and later we will describe some of the theoretical papers of the EKC with recycling. We can think of recycling as a measure of environmental abatement, and the exact opposite of pollution or environmental degradation. Hence, it is important to understand the literature of the basic EKC in order to analyze the extension with recycling that is the focus of our study.

2.1. The Classical EKC

Grossman and Krueger (1991) were the first to explore the empirical relationship between pollution—specifically air quality—and economic growth using daily data for Sulphur dioxide for 42 countries in 1977, 1982 and 1988, for suspended particles for 29 countries for the same years as well as data for dark matter (dust) for 19 countries. The countries they chose to use in their analysis were both developed and developing countries from different regions of the world. What they found was an inverse-U curve that describes the relationship between air quality and GDP (EKC), with the turning point being between 4.000 to 5.000 US dollars. Subsequently, Grossman and Krueger (1995) using again data for air and water quality for many developed and developing countries from all over the world, examined the relation between growth and pollution. As in their previous study, they mostly focussed on data for sulfur dioxide since it affects lung quality and to measure the water quality, they used data for lakes, groundwater aquifers and river basins. They also included the state of the oxygen in water, the pathogenic contamination of water and the heavy metals content as indicators for water quality and urban air pollution as an indicator for air quality. They confirmed the presence of the (EKC) of an inverse-U relationship between GDP per capita and pollution with a turning point at an income level less than 8.000\$.

Esteve and Tamarit (2012) also showed that the EKC holds for Spain even though previous studies had found the relationship not to follow an inverse-U curve as EKC would suggest, but linear. The data they used was for the period 1857 to 2007 for per capita CO₂ and per capita income for Spain and the main difference of their research from previous studies was the empirical approach they adopted based on a threshold time series model and they found that the shape of EKC is indeed inverse-U.

Cole et al. (1997) introduced more environmental indicators than most studies used until then to check if the results of earlier studies still held true. They found that carbon dioxide and methane are not statistically significant while the relationship between GDP per capita and municipal waste, energy consumption and traffic volumes was monotonically increasing. Moreover, the inverse-U shape of EKC was only supported by local air pollutants, while pollutants of a more global nature have really high turning points with large standard errors a result also confirmed by Stern (1998). Roca et al. (2001) also found that the inverse-U curve should not be generalized neither for all type of emissions nor for global pollutants, but rather for some local ones.

The above mentioned papers found evidence that either fully or at least partially supported the

EKC following an inverted U pattern. However, there are also papers that refute entirely the evidence of the inverse-U shape and support instead an N curve pattern. An N-shaped curve was suggested by [De Bruyn and Opschoor \(1997\)](#) to be a more appropriate shape that describes the relationship between pollution and economic growth using data for 19 Western European countries for 1966 to 1990. [Harbaugh et al. \(2002\)](#) using a more updated version of the dataset on air pollution than that of [Grossman and Krueger \(1995\)](#) found that the inverse-U relationship was not supported by these updated data. Similarly, [Dinda \(2004\)](#) also concluded that the EKC is better characterized by an N-shape curve. Specifically, the inverse-U relationship between pollution and income per capita was verified mainly by air quality indicators and it should not be generalized for other types of pollutants. Water pollution, population pressure and other factors should also be considered for environmental quality. The literature so far assumed that the initial increase in the pollution is temporary while the decrease of it is permanent. This result is debatable as the decrease of the pollution after a certain level of income per capita may not be permanent, which leads to an N-shape curve for EKC. Most papers have used emission levels as a pollution index even though global environmental degradation is assumed to be a better factor for global EKC. Moreover, although country specific methods are more appropriate for the examination of the EKC hypothesis, the most popular analysis used in most of empirical studies was based on cross-section data.

2.2. The EKC with Recycling

The existing literature of EKC with recycling is quite thin and all contributions are theoretical in nature. [Selden and Song \(1995\)](#) using a simple growth model and assuming that the utility of the representative agent is affected positively by consumption and negatively by pollution levels, derived the shape of the EKC to be an inverted U curve and that the corresponding curve for abatement to be a J shape curve. However, pollution in that model may not be optimal. An additional study that dealt with recycling and growth is the paper by [Di Vita \(2001\)](#), which showed that economic growth is higher when one introduces materials in production which come from waste recycling compared to the case that an economy uses only capital, labour and technology to produce its final good. Using values for the exogenous variables consistent with macro empirical evidence they showed that developed countries increase their welfare by selling recyclable waste to developing countries which can produce secondary materials cheaper. [Pittel \(2006\)](#) introduced a new EKC model which varies in two main aspects from the standard EKC model. The model is very simple, assuming that there is a closed economy which functions under a material balance condition, i.e., material cannot be destroyed and it can only be converted through recycling. What he found is that EKC, under certain conditions, might arise during the transition to the long run balanced growth path. More recently, [George et al. \(2015\)](#) found that the EKC does not hold in a closed economy with two factors of production, namely a recyclable input as well as a polluting resource and assumed that employing the recyclable input had no cost for the firm while the polluting input had, which if increased a lot led to the substitution of the more expensive polluting input by the cheapest recyclable input. As a result, economic growth did not have a positive effect on environmental quality as the EKC would suggest and as the economy would become richer environmental degradation would increase. [Kasioumi \(2020\)](#) extended [George et al. \(2015\)](#) by relaxing its main assumptions, namely that the two inputs are independent and that the recycling input is costless and reversed its results by confirming the presence of the EKC to be an inverted U curve and the corresponding curve for abatement to be a J shape curve as in [Selden and Song \(1995\)](#). They also included two extensions of their basic model adding technological progress in the production of the polluting input and a dynamic cost for that input respectively, both of which confirm the previously mentioned results.

2.3. The Use of Non- and Semi-Parametric Methods in the EKC Literature

Non parametric and semi parametric methods have become widely used in applied work as they provide a lot of functional flexibility and offer robustness shield against deviations from parametric

functional form assumptions. Most of the earlier studies made use of single time series or single cross sectional data and it is only more recently that panel data have become to be used more frequently. [Millimet et al. \(2003\)](#), examined the validity of the parametric models—usually used to explain the inverse-U relationship of pollution and income (EKC)—by using more robust semi parametric partially linear regression models as means of comparison. Their research focused on specific types of emissions, i.e., SO₂ and NO_x for US states for the years between 1929 to 1994. They concluded that both techniques confirm the existing pattern of the inverted-U shape of EKC but a comparison of the results of the different econometric methods led them to conclude that semiparametric models are more appropriate to use. [Yavapolkul \(2005\)](#) used a nonparametric approach to explore the shape of the EKC and make comparisons with previous studies using parametric methodologies. Using data for SO₂ for 44 countries for the period 1972 to 2001 and data for CO₂ for 103 countries for the period between 1975 and 1996 he found that the inverse U shape of the EKC holds for both pollutants, even though some previous studies found that the EKC under CO₂ does not hold using parametric models. More recently, [Shahbaz et al. \(2017\)](#) found that the EKC holds for most of the G7 economies¹. Using non parametric unit root tests, non parametric cointegration, regression analysis and causality tests, they examined the relationship between CO₂ emissions and economic growth for the previously mentioned countries for the years between 1820 and 2015- the biggest dataset ever employed in a non parametric as well as parametric analysis of this type. They found that there is a long run relationship between pollution and GDP (nearly two centuries) and most importantly, the inverse U shape of the EKC holds for all the G7 countries except Japan. In addition, they found that even though there is no causality between the two variables for Japan, there is a one-way causality for Canada, Germany, U.K. and the U.S. while there is a dual causality for France and Italy.

Even though the previously mentioned papers conclude that the inverse U shape of the EKC holds irrespective of the use of nonparametric or semi parametric models, there are also some other studies which do not share this conclusion. Within a nonparametric framework [Criado \(2008\)](#) tested the homogeneity assumption on spatial as well as temporal dimensions. Using a panel data set for 48 Spanish provinces and four air pollutants CH₄, CO, CO₂, NMVOC he found that the inverse U curve for the EKC holds for all four air pollutants. However, he also found that the semi parametric and parametric fixed effects models failed to account for spatial heterogeneity which resulted in the shape of the curves for the EKC to be monotonic (increasing or decreasing depending on the air pollutant used), missing the curvature that is implied by the fully nonparametric model. [Azomahou et al. \(2006\)](#) were the first to base their analysis on a non parametric panel data model with individual effects and the questioned the existence of the EKC as an inverse U curve. Using data for 100 countries for the period between 1960 and 1996, for CO₂ emissions per capita and GDP per capita they applied a local linear kernel regression model to test the constancy of the relationship of the two variables over time using a poolability test, while to test a monotonic versus the polynomial functional form they used a monotonicity test. They found that the EKC holds under a parametric model only, while under the nonparametric and the first difference specification, the EKC does not hold as these models suggest that the relationship between pollution and GDP is monotonic and upward slopping. They also showed that there is stability in the relationship between the two variables for the time period they are studied. More recently, [Fakih and Marrouch \(2019\)](#) examined the EKC using data for the Middle East and North Africa (MENA) under a nonparametric panel data framework. Specifically, using data for per capita carbon dioxide emissions and GDP per capita for the period between 1980 and 2010, their results are in agreement with the pessimistic view on the EKC, i.e., that the relationship between those two variables is monotonically increasing and it does not follow an inverse U shape.

¹ The G7 economies are the following: Canada, France, Germany, Italy, U.K., U.S. and Japan.

2.4. Non- and Semi-Parametric Models with Fixed Effects

Li and Stengos (1996) derived a \sqrt{N} consistent estimator for a general partially linear semiparametric panel data model. They used a kernel instrumental variable method and they assumed that some of the regressors can be correlated with the errors. They concluded that this estimator works best for sample size between 50 and 100 observations as well as for a time span of three years. Baltagi and Li (2002) extended Li and Stengos (1996) using a series method to obtain a consistent estimator which does not suffer from the curse of dimensionality and estimates the original unknown function of the model and not the function which arises after first differencing. This estimator is based on a dynamic partially linear semiparametric panel data model with fixed effects. Henderson et al. (2008) introduced an iterative nonparametric kernel estimator to estimate nonparametric panel data models with fixed effects. They found that the estimation results are consistent whether they ignore the within subject correlation or not. They also extended their analysis to the case of a semi parametric partially linear fixed effects model, for which they found that the estimation results are consistent only when they incorporate the within subject correlation. Sun et al. (2009), using a local linear regression approach, derived an asymptotically normally distributed estimator for a varying coefficient panel data model with fixed effects. The benefit of this estimator is that it removes the fixed effects by using kernel-based weights and not by applying first differences in the initial model. Lin et al. (2014) obtained a test statistic for a linear fixed effects panel data model (null hypothesis) against a nonparametric fixed effects panel data model (alternative hypothesis). Finally, Boneva et al. (2015) estimated a semiparametric two ways fixed effects panel data model which allows the nonparametric covariate effect to be heterogeneous across individuals, by applying the local linear kernel estimation method. The Nadaraya–Watson smoother can also be used instead of the local linear methodology leading to same results. They carried out an empirical investigation with weekly data for the FTSE 100 and FTSE 250 for the time period between May 2008 and June 2011.

3. Model

We will estimate a partially linear semi-parametric model using data for 30 years and 50 US states. We will focus first on the overall data set and estimate the relationship between Recycling and Real GDP per capita (the Recycling EKC) and then we will proceed to divide our data to three categories, separating the states into rich, middle rich and middle poor and re-estimate the model for each category. We will also present our results for two different time periods for all the above data categories, i.e., for the years 1988 to 1997 and for the years between 1998 to 2007.

3.1. Data

The data we will use in our analysis is for the period between 1988 to 2017, for the 50 American states. Data on transfers to recycling from total waste have been collected by the Toxics Release Inventory (TRI) of the United States Environmental Protection Agency (EPA) and they are measured in pounds. Real GDP per capita was originally published in the Bureau of Economic Analysis—U.S. Department of commerce (BEA) and it is measured in dollars, using 2012 as base year. The data for the total Real GDP are also from the Bureau of Economic Analysis—U.S. Department of commerce (BEA) and it is measured in millions of dollars, using 2012 as base year. We are interested to study the relationship between recycling and real GDP for both total and per capita values. However, our data for recycling are calculated in total values only and hence we need to convert it to per capita units and to do so, we use population data from the United States Census Bureau. Finally, both recycling and real GDP are annual data collected by State².

² A list with all the states used can be found in the Appendix A as well as the categorization of the states depending on their total real GDP.

We will use this data set to create some smaller data sets and we will base our analysis to the per capita values as well as the total values of these data sets. Our final 8 data sets that we will base our analysis are the following: a data set which includes all the remaining states (without the outliers) for all the years, a data set which includes the rich states only for all the years, a data set which includes middle rich states for all the years and finally a data set which includes the middle poor states for all the years, for the per capita and total values.

In the following table, we present the descriptive statistics of our main dataset that includes the data from all the states used in the analysis covering the period 1988 to 2017³. For Table 1, we used data on Recycling and Real GDP for the final data set we used, which does not include the outliers (the 3 richest and the 3 poorest states). Descriptive statistics for the rest of the categories (Rich, Middle Rich and Middle Poor States) can be found in Appendix B.

Table 1. Descriptive Statistics of the data for All States of U.S.

	Mean	var	Median	Range
Recycling total	3.501382×10^7	2.096566×10^{15}	2.100000×10^7	$0.0 \times 10^0 - 4.3 \times 10^8$
Recycling pc	7.941477×10^0	7.280822×10^1	5.871753×10^0	$0.0 \times 10^0 - 94.15841 \times 10^0$
Real GDP total	1.813785×10^5	1.601082×10^{10}	1.528180×10^5	$18,571.02 \times 10^0 - 499,694.90 \times 10^0$
Real GDP pc	4.383338×10^4	1.028631×10^8	4.269534×10^4	$23,829.92 \times 10^0 - 79,894.00 \times 10^0$

3.2. Empirical Method

The main purpose of our paper is to explore the best possible shape of the curve that best characterizes the relationship between recycling and real GDP without making assumptions about specific functional forms. To do so, we will use a semi parametric regression model that avoids any pre-conceived assumption such as an inverted U shape or an N shape for this relationship and allows the data to speak for themselves. Following the empirical methodology of Millimet et al. (2003), we will estimate the following semi parametric partial linear two-way fixed effects panel data model. It is important to use both time and individual effects in our analysis because in that way we can eliminate biases that stem from two potential sources. We want to control for some unobserved factors which differ over time but are constant over space (states) such as different environmental legislation (time effects) as well as for factors that vary across states but are constant over time but they also affect recycling (individual effects). There are few papers that study the basic EKC under non parametric or semi parametric models, but there is no other paper, to our best knowledge, i.e., following the same empirical methodology on recycling and we hope to be able to uncover the underlying relationship and provide evidence to assess the theoretical models on the subject. We proceed to analyse the EKC with recycling with a semi parametric partial linear two way fixed effects panel data model, which can be seen in the following formula:

$$R_{it} = \lambda_i + \mu_t + g(y_{it}) + u_{it}, \quad i = 1, 2, \dots, 50 \text{ and } t = 1988, \dots, 2017$$

where R_{it} is the transfers to recycling from total waste in State i at time t and y_{it} is the real GDP in state i at time t . In addition, λ_i is the state effects while μ_t is the time effects of our model and these two variables are the ones that make our model partial linear. R_{it} is the dependent variable while y_{it} , λ_i and μ_t are the independent variables of our model. Furthermore, u_{it} is the error term which we assume is uncorrelated with $g(\cdot)$ and has zero mean. Finally, $g(\cdot)$ is an unknown function.

³ Please note that for the calculation of the descriptive statistics the unit of measurement for our data is as follows: total recycling is measured in pounds, recycling per capita is measured in millions of pounds, Real GDP total is measured in millions of dollars while Real GDP per capita is measured in dollars. We have excluded the 3 richest and the 3 poorest states, so the table includes the data from the remaining 44 states

We will start our analysis by exploring the possibility that a simple linear specification and/or a quadratic specification may provide adequate representations of the data. In fact, some of the papers dealing with this topic suggest that the EKC with recycling should be characterized by a J curve, with recycling on the vertical axis and GDP on the horizontal axis, supporting the notion that the EKC with recycling should take a quadratic form. Yet, others reject this idea as the claim is that EKC with recycling may not exist or if it does it may be more complicated. For this reason, we believe that our semi parametric model is more suitable than the linear and/or a quadratic alternative, since it allows the data to reveal which pattern can best characterize the relationship between our two main variables. However, in order to proceed we will test these two benchmark parametric specification against the semiparametric alternative using the consistent specification test by [Racine et al. \(2006\)](#) which tests the null of a parametric model against a (semi) nonparametric alternative. The two benchmark models and the semi parametric partial linear panel data model that we estimate, expressed in logarithmic form to interpret the marginal responses in percentages as it is commonly the practice in the literature, are given below as

$$\log(R_{itv}) = \lambda_i + \mu_t + \log(y_{itv}) + u_{it}, \quad i = 1, 2, \dots, 50 \text{ and } t = 1988, \dots, 2017$$

$$\log(R_{itv}) = \lambda_i + \mu_t + \log(y_{itv}) + \log(y_{itv})^2 + u_{it}, \quad i = 1, 2, \dots, 50 \text{ and } t = 1988, \dots, 2017$$

$$\log(R_{itv}) = \lambda_i + \mu_t + g[\log(y_{itv})] + u_{it}, \quad i = 1, 2, \dots, 50 \text{ and } t = 1988, \dots, 2017$$

where i represents the states and it can take the values between 1 and 50, t represents the years and is between 1988 to 2017 while v stands for the type of our data and it can either be total values or per capita.

There are alternative ways to remove the fixed effects from the above model, see [Henderson et al. \(2008\)](#). The standard approach is based on taking first differences to remove the fixed effects. Alternatively, if the number of units is not large as in our case, one can use a two stage procedure, where in the first stage one removes the effect of the fixed effects by regressing the recycling and income variables on the fixed units and then using the residuals in the second stage as the new variables. The findings that we present are based on the second approach, even though both methods yield very similar results.

Since our model is a semi parametric model and we make no assumptions about the distribution of our data, we will use a local constant kernel estimation approach based on a Gaussian choice of kernel and cross validation for bandwidth selection. We are interested in the relationship between our two variables for both total but also per capita values. We first want to examine the pattern that exists between recycling and real GDP in total values in order to establish the general link between these variables. This is important because we will be able to see how recycling in each State reacts to changes of their total income level respectively. We are also interested in the per capita relationship that exists between these two variables because we want to see how recycling by the average person in each State reacts to changes in their income level.

Our data will be separated in the different categories according to the income level of the states that were mentioned earlier and in the next section we present our results⁴. Thus, except for the initial data set which includes all the states and the years, we also have the smaller data sets for the richest and middle rich states only as well as for the states which belong to the lower middle income level states, which we will call middle poor states. The poor states will be included only in the overall data set that includes all years and states and they will not make it into a separate category of their own, due to very low levels of recycling in these states that do not allow for a meaningful separate estimation.

⁴ A list with all the states used can be found in the Appendix A as well as the categorization of the states depending on their total real GDP. In our analysis we will also exclude from the original data set some outliers that belong to the three poorest and the three richest states

4. Results

In this section, we present the semi-parametric regression results for the relationship between recycling and real GDP per capita for the eight different data sets that we have created. We will first discuss the results when we use the entire data set for 44 states and for 30 years. We will then discuss the results for the same regression but for the smaller data sets, i.e., for the eleven richest, the eleven middle rich and the eleven middle poor states. All the above regressions are carried out for both per capita and total values which led us to create the eight different data sets. The estimated 95 per cent confidence intervals is added to each estimated curve in order to ascertain their statistical significance. The construction of these confidence bands is based on the bootstrap⁵.

We start the analysis by first testing the null hypothesis of the linear and quadratic specifications against the semiparametric model alternative. The test results from the test by [Racine et al. \(2006\)](#) suggest strong rejections of both the linear and quadratic models for all different data sets that we use in our estimations. As it can be seen from [Table 2](#) all the *p*-values taken from the linear and the quadratic model are zero, for both the per capita and the total values. Thus, we conclude that the semi parametric model that we use offers the best specification for our empirical investigation and we proceed to present the results from this model.

Table 2. *p*-values of the linear and quadratic models.

	Linear (pc)	Quadratic (pc)	Linear (Total)	Quadratic (Total)
All States	0.000	0.000	0.000	0.000
Rich States	0.000	0.000	0.000	0.000
Middle Rich States	0.000	0.000	0.000	0.000
Middle Poor States	0.000	0.000	0.000	0.000

We first present the results for the data sets with per capita units and later on we will present the plots for the total values. We notice that initially, when Real GDP per capital is comparatively small, recycling is being increased dramatically up to a point, after which while per capita income increases, recycling remains comparatively constant at its highest values. Overall, the shape that characterizes the relationship between recycling and real GDP per capita is increasing as we can see in [Figure 1](#)⁶.

When we focus on the data set for the rich states only, the curve we obtain is similar with the previous one but with some small differences. It is clear that this curve suffers from more fluctuations at the beginning of the plot but it still follows the same pattern as the curve for the overall data from all states. Mainly, the recycling level increases as real GDP increases until it reaches its highest point and then it remains constant. Hence, as in the previous plot we see that the relationship between real GDP per capita and recycling per capita is characterized by an increasing curve ([Figure 2](#)).

The figure we obtain for the middle rich states is subject to more fluctuations than the previous two cases as we can see in [Figure 3](#). In particular, there are many fluctuations in the first half of the plot, while after that point the plot is smoother. However, in that case the pattern between recycling and income level is not increasing anymore and it seems that in the final part of the plot the curve is starting to fall.

⁵ The bootstrap has only been recently used to construct confidence intervals in panel data models, see for example [Karlsson \(2009\)](#) who based his analysis on a weighted nonlinear quantile regression. The main finding there is that bias corrected confidence intervals failed to achieve a true coverage which is at least as large as the nominal coverage, since they tend to over correct the bias. However, confidence intervals which were calculated with the percentile method, are the only ones that perform well for both coverage and length. We used R to construct our confidence intervals and specifically the package np.

⁶ Please note that the vertical axis of this plot includes negative values, since we are using the logarithmic form of our data.

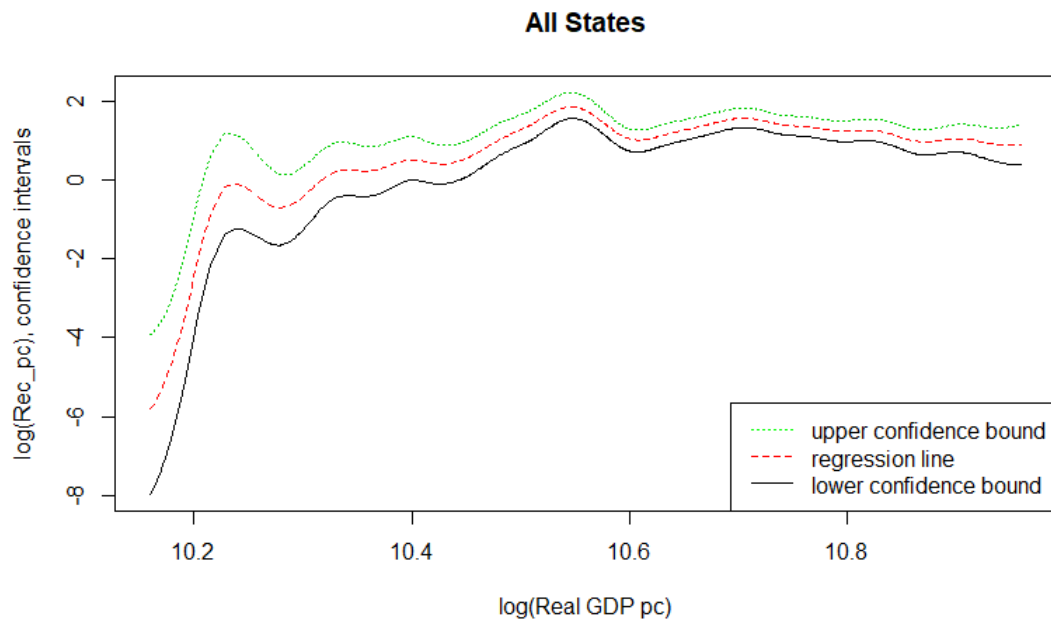


Figure 1. Relationship between Recycling per capita and Real GDP per capita.

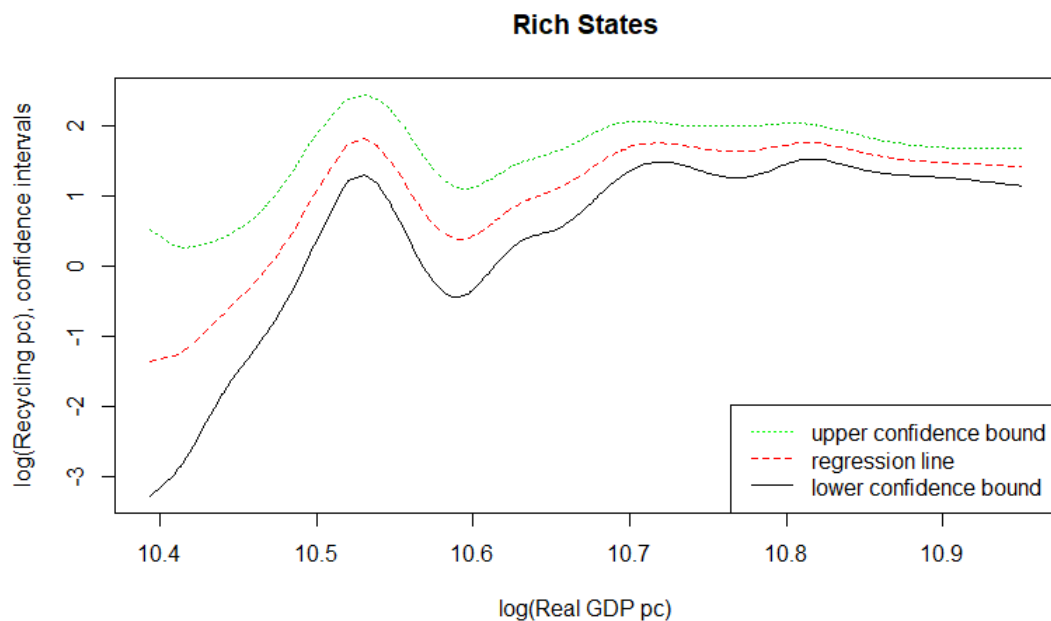


Figure 2. Relationship between log of Recycling per capita and log of Real GDP per capita for rich States.

Finally, for the middle poor states, the shape that describes the relationship between recycling and real GDP per capita is shown in Figure 4. The recycling level increases dramatically with income at the beginning as in the case of all states. This is followed by some small fluctuations, but in general the recycling level remains at a high level. In other words, as the level of income increases, recycling increases too for the middle poor states observing again an increasing pattern.

Moving on to the case where we analyze the relationship between our two variables for the total values, we will again present plots for the four categories that we study. To start with, the pattern between our two variables for all the states is still increasing and is very similar to the one for the per capita values (Figure 5). The only difference between the two plots for all the states is that the

plot for total values displays some more fluctuations than the one with the per capita values. On the other hand, the curves for the rich and middle rich states have less fluctuations than before. However, the main pattern is still the same, i.e., an increasing curve for the rich states and an inverse U curve for the Middle Rich states as we can see in Figures 6 and 7 respectively. Last but not least, the plot for the middle poor states, see Figure 8, is similar to the corresponding plot for the per capita values but smoother and it still follows an increasing pattern.

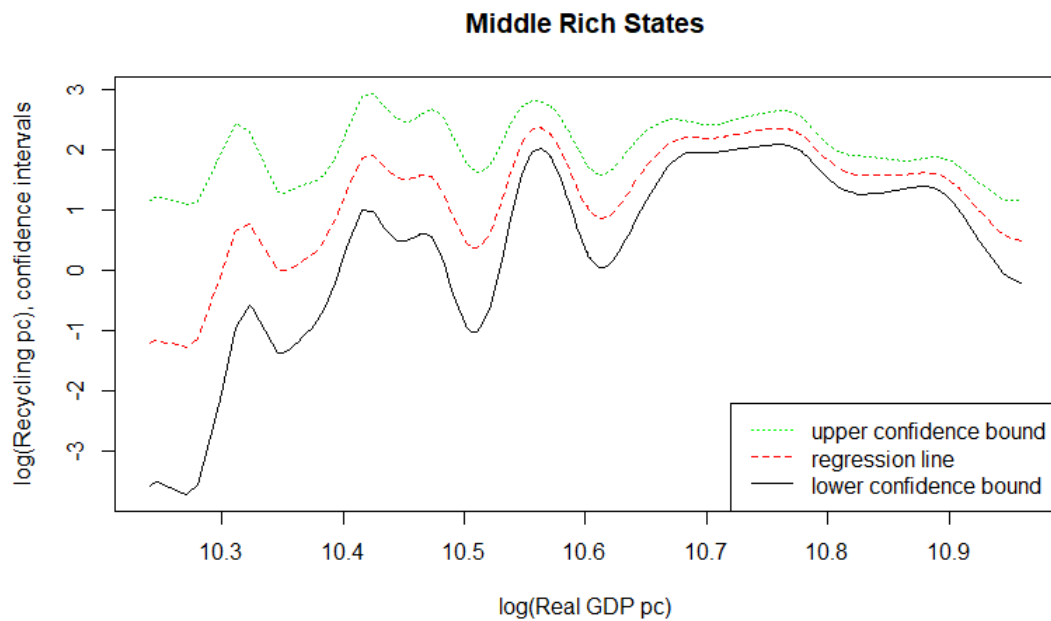


Figure 3. Relationship between log of Recycling and log of Real GDP per capita for middle rich States.

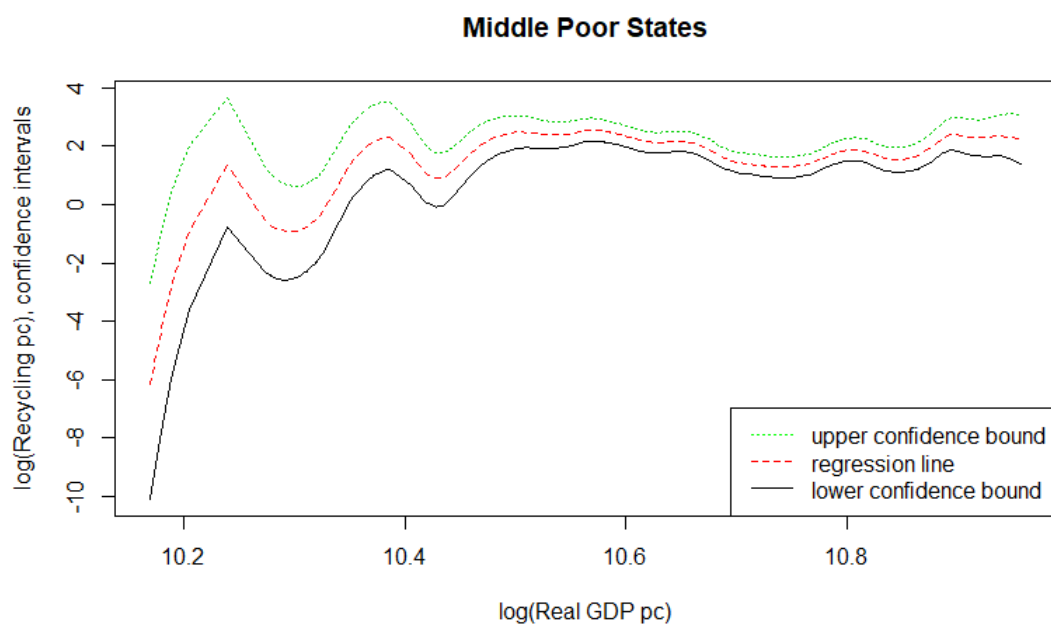


Figure 4. Relationship between Recycling per capita and Real GDP per capita for middle poor States.

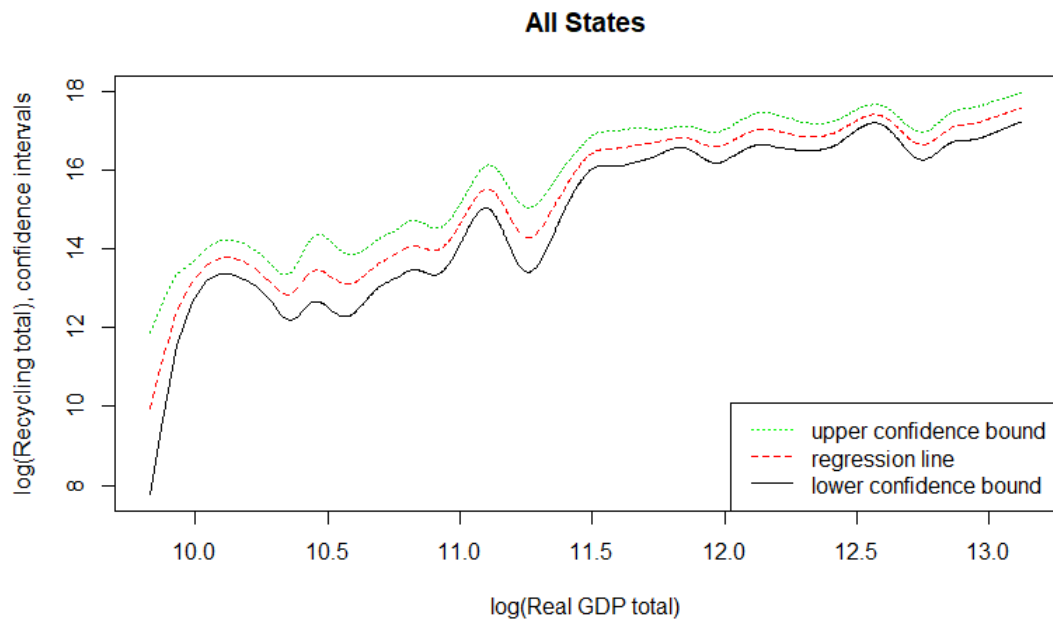


Figure 5. Relationship between log of total Recycling and log of total Real GDP.

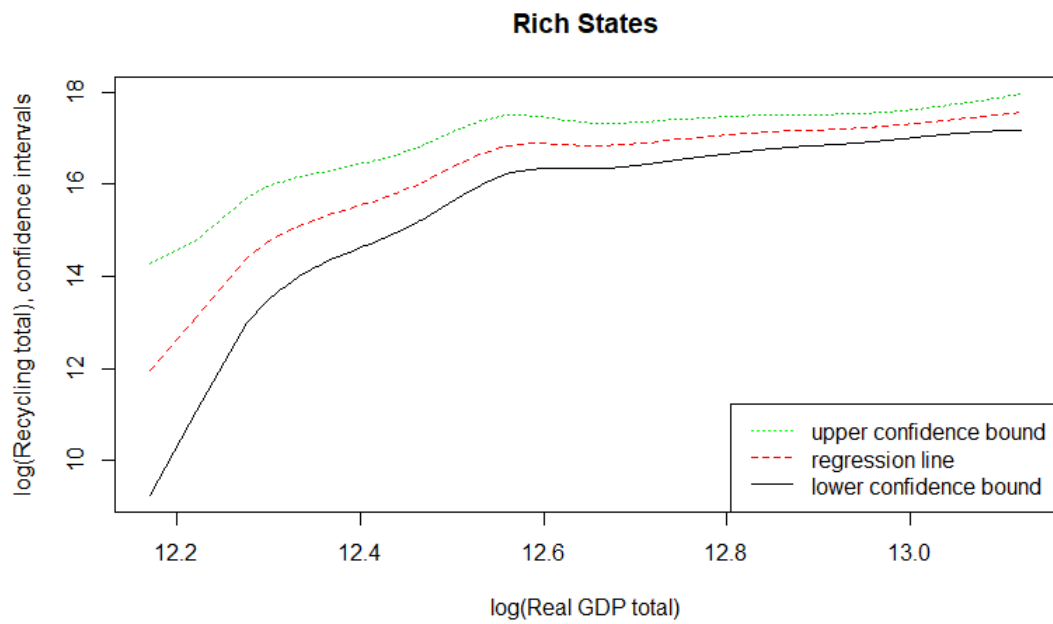


Figure 6. Relationship between log of total Recycling and log of total Real GDP for rich States.

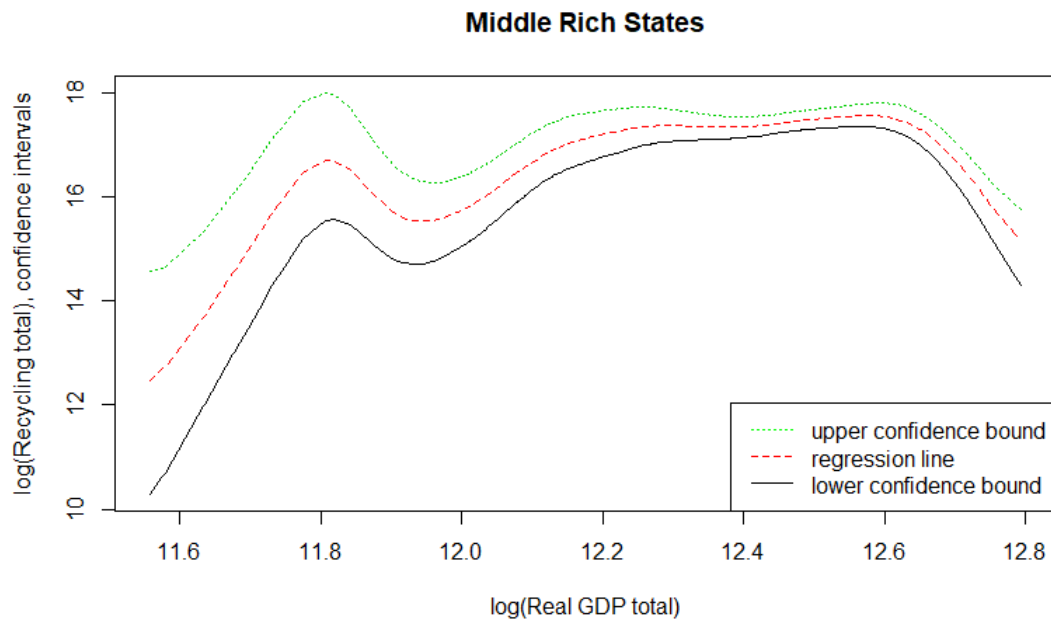


Figure 7. Relationship between log of total Recycling and log of total Real GDP for middle rich states.

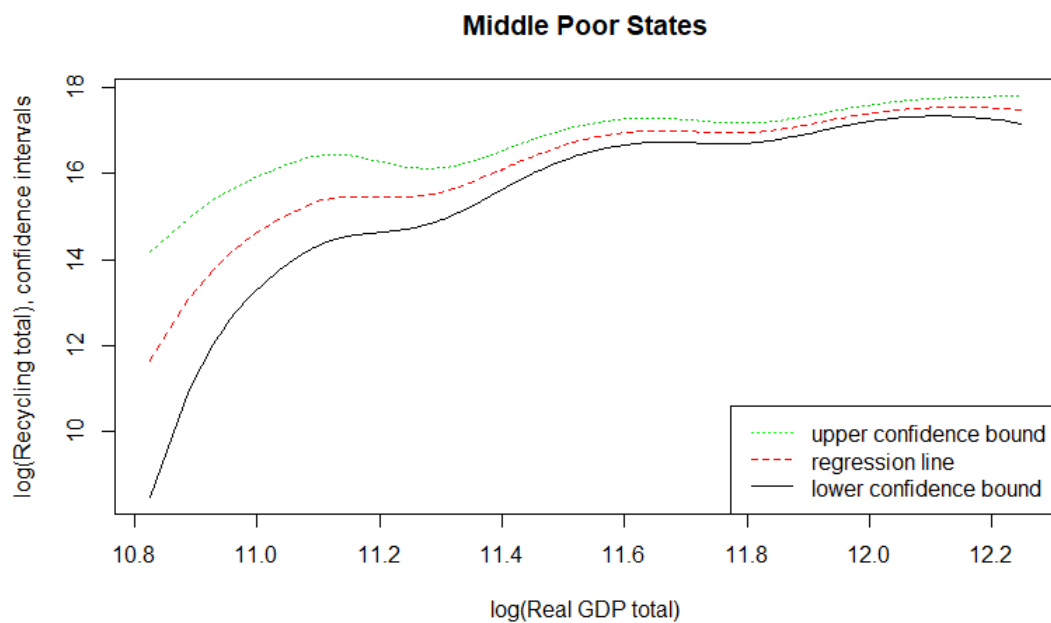


Figure 8. Relationship between total Recycling and total Real GDP for the middle poor states.

We see that for all the cases discussed above, apart from some differences in fluctuations, the dominant pattern between recycling and real GDP is increasing. We can also conclude that all the plots which are based on the total values are smoother than the corresponding ones for the per capita values and finally, the plots for the middle rich states have the most fluctuations in both per capita and total values, than the plots for all the other cases.

4.1. Separation of the Data According to Years

We are also interested to see how the recycling level of the states changes with real GDP during certain specific time periods. For that reason, we include one more case in our analysis. We separate our data into two categories each one including 10 years in order to test whether the relationship between recycling and real GDP changes over time or not. Again, we check how the pattern between our two variables evolves for both per capita and total values as before. In the first category we include years 1988 to 1997 while in the second one years 1998 to 2007. The plots of these regressions are presented in the Appendix C and they show similar patterns.

As we can see from all the previous cases, the main pattern between abatement and real GDP is characterized by an increasing curve. Under both main categories—total and per capita values—the results we found have not changed, for all the different categories according to the level of the real GDP or the years under consideration. Hence, we can conclude that our results confirm the existence of an increasing curve for the EKC with recycling, as other theoretical papers found as well. Our paper is in line with the results of [Selden and Song \(1995\)](#) and [Kasioumi \(2020\)](#), both of which confirm the existence of a J curve for the EKC with abatement using two different theoretical models. The present paper contributes to the existing literature in two ways. Firstly by introducing the first empirical paper on that topic and secondly by confirming the theoretical results of some recent papers which support the view that the EKC with abatement is a J curve.

5. Conclusions

The purpose of our paper is to study a comparatively new Environmental Kuznets Curve, which examines the relationship between recycling and income level. The existing literature on this topic is quite limited and it includes only theoretical papers. Our contribution is to approach this topic from an empirical perspective and specifically to use a very flexible empirical methodology. Hence, we use a partial linear semi parametric two way fixed effects panel data model and we use data for recycling and real GDP, for the fifty states of the United States for the years 1988 to 2017. In that way we can control for unobserved factors which are different in every State and affect recycling, as well as some macroeconomic factors that are same across the states but are different across years. We make use of a standard local constant kernel estimator and we also use the cross validation method to select our bandwidth.

This model allows the data to show us which pattern can best characterize the relationship between the two variables we are studying. In other words, we do not force the data to follow a specific curve structure, rather we let them reveal this curve to us. The rejection of the null hypothesis of both the linear and quadratic models against the semi parametric model that we use confirms that the latter is the best and most adequate model specification to explain the relationship between recycling and real GDP.

Hence, we base our analysis in eight different data sets based on the income level of the states. First of all, we use all the data we have for all American states with available data over the period, that is for 44 states and for 30 years for the period between 1988 and 2017. Then, we use data for the 11 richest, 11 middle rich and 11 middle poor states for the same years. The separation of the data is based on the comparison between the average real GDP that each State had during the years we are studying. Finally, our regressions are based on the per capita and total values of the above data sets and we also separate all the above data sets into two smaller periods, for the years between 1988 to 1997 and 1998 to 2007.

We find the relationship between recycling and real GDP to be characterized by an increasing curve for most of the data sets we used, as the results do not change significantly among different data sets and the general pattern remains that of an increasing curve. That means that as the states become richer, they decide to increase the level of recycling they do. This may be happening for many reasons. First of all, the states may realize that higher production coming from non green companies, leads to a great pollution problem which they want to avoid for environmental reasons, and as such

recycling can be one way to reduce environmentally harmful waste. Another reason might be that as real GDP increases and states become richer, they are able to invest to more recycling firms as well as new more effective ways of recycling. Over time more people become environmentally conscious and governments are forced to pay greater attention to environmental issues— one of which is recycling. Our results confirm the predictions of recent theoretical models, see [Selden and Song \(1995\)](#) and [Kasioumi \(2020\)](#) that find an increasing pattern in the form of a J curve for the EKC with abatement.

Our analysis involves data for the United States only, thus for future research some possible extensions could include data for different years and different countries. In addition, we should also include more variables in our regression such as education level, emission level and data about temperature. In that way, we will be able to check the robustness of the EKC with recycling in different countries in order to have more confidence in the results we found in this paper.

Author Contributions: Conceptualization, T.S. and M.K.; methodology, T.S. and M.K.; software, M.K.; validation, T.S. and M.K.; formal analysis, T.S. and M.K.; investigation, M.K.; data curation, M.K.; writing—original draft preparation, M.K.; writing—review and editing, T.S.; supervision, T.S.; project administration, T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Separation of the States According to Their Average Income Level

We used eight different data sets in our previous analysis which are differentiated according to the average income level that the states had between the years of 1988 and 2017. The first two data sets we used include the per capita values for all 50 states of the US and the total values of these states respectively. The next two data sets include data for the richest states for the per capita and total values again. The following two data sets include data for the middle rich states for the total and per capita values of our data while the last two sets include data for the middle poor states for the per capita and total values once again. Here, we present the separation of the states into four categories according to their average real GDP. We have the richest, the middle rich, the middle poor and the poorest states. Even though we do not use separately the last category in our analysis, we believe that is important to state which states are included in it as these states are still being used in the first two data sets we mentioned above (sets for all the states). Last but not least, since we did not include some outlier states in our analysis, the following categories do not include the three richest and the three poorest states. Hence, the states included in the four categories for the total values are the following:

Richest States: Washington, Virginia, Massachusetts, North Carolina, Georgia, Michigan, New Jersey, Ohio, Pennsylvania, Illinois, Florida

Poorest states: South Dakota, Montana, Rhode Islands, Alaska, Idaho, Maine, Delaware, New Hampshire, West Virginia, Hawaii, New Mexico

Middle rich states: Alabama, Louisiana, Arizona, Connecticut, Colorado, Wisconsin, Tennessee, Missouri, Minnesota, Indiana, Maryland

Middle poor states: Nebraska, Mississippi, Arkansas, Utah, Nevada, Kansas, Iowa, Oklahoma, Oregon, South Carolina, Kentucky

Appendix B. Descriptive Statistics for the Rich, Middle Rich and Middle Poor States

In that section we present the descriptive statistics for the smaller datasets used in our analysis. The following tables (Tables [A1–A3](#)) are for data for the 3 different groups of states, the Rich States, the Middle Rich and finally the Middle Poor States for the years 1988 to 2017.

Table A1. Descriptive Statistics of the data for Rich States.

	Mean	var	Median	Range
Recycling total	5.318674×10^7	3.029723×10^{15}	3.100000×10^7	$4455 \times 10^0 - 0.320 \times 10^9$
Recycling pc	5.958935×10^0	3.203852×10^1	3.939069×10^0	$4.728654 \times 10^{-4} - 4.642474 \times 10^1$
Real GDP total	3.785252×10^5	6.268160×10^9	3.975409×10^5	$193,029.6 \times 10^0 - 499,694.9 \times 10^0$
Real GDP pc	4.602757×10^4	7.314679×10^7	4.471650×10^4	$32,657.4 \times 10^0 - 71,153.0 \times 10^0$

Table A2. Descriptive Statistics of the data for Middle Rich States.

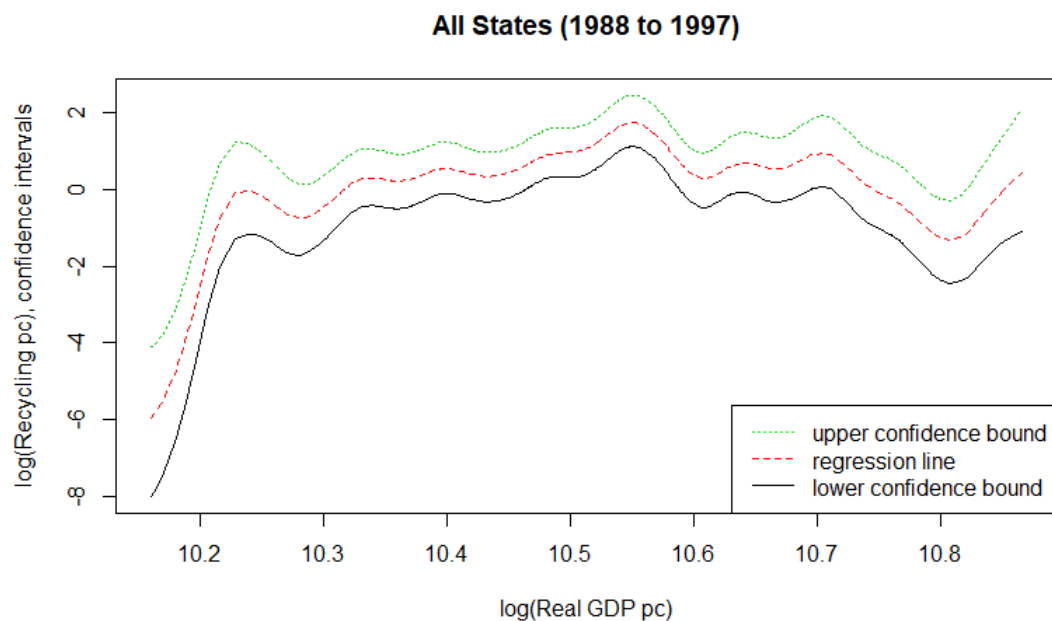
	Mean	var	Median	Range
Recycling total	5.046606×10^7	3.256133×10^{15}	3.500000×10^7	$2.5 \times 10^3 - 4.3 \times 10^8$
Recycling pc	9.461599×10^0	9.282100×10^1	7.598915×10^0	$5.208583 \times 10^{-4} - 7.492570 \times 10^1$
Real GDP total	2.280674×10^5	3.044786×10^9	2.333420×10^5	$104,638.5 \times 10^0 - 360,030.1 \times 10^0$
Real GDP pc	4.514498×10^4	8.389377×10^7	4.438300×10^4	$28,025.36 \times 10^0 - 74,031.00 \times 10^0$

Table A3. Descriptive Statistics of the data for Middle Poor States.

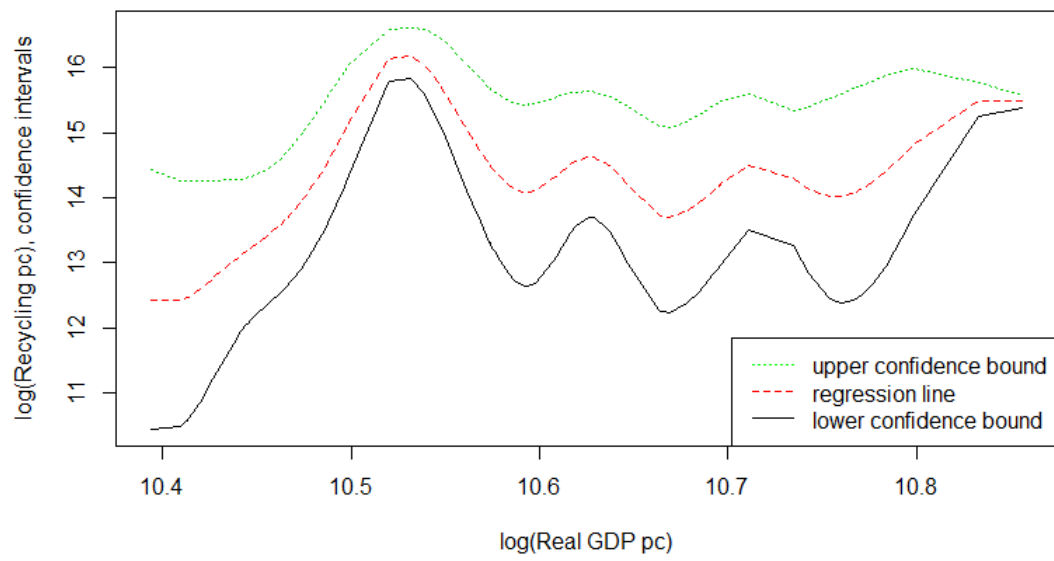
	Mean	var	Median	Range
Recycling total	3.524089×10^7	9.129975×10^{14}	2.900000×10^7	$0.0 \times 10^0 - 2.5 \times 10^8$
Recycling pc	1.150730×10^1	9.564348×10^1	1.117272×10^1	$0.0 \times 10^0 - 94.15841 \times 10^0$
Real GDP total	1.188691×10^5	1.300694×10^9	1.133653×10^5	$50,288.41 \times 10^0 - 208,625.80 \times 10^0$
Real GDP pc	3.975868×10^4	6.212606×10^7	3.900900×10^4	$23,829.92 \times 10^0 - 57,982.00 \times 10^0$

Appendix C. Plots for the Different Categories Based on Year Separation

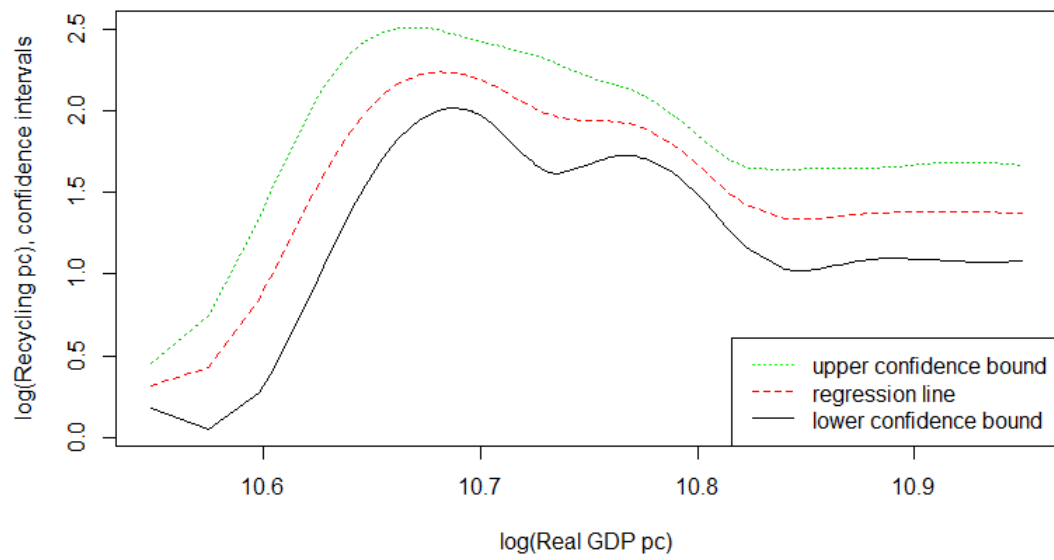
In this section, we will present some of the most representative plots that we are getting from the regression between recycling and real GDP for the total and per capita values, when we use the data sets which are separated according to the different time periods, as we mentioned previously in Section 4.1. Again, we will begin with the plots for the per capita units for all the states as well as for some smaller categories—rich, middle rich, etc states—and later we will present for the same categories the plots for the total values.



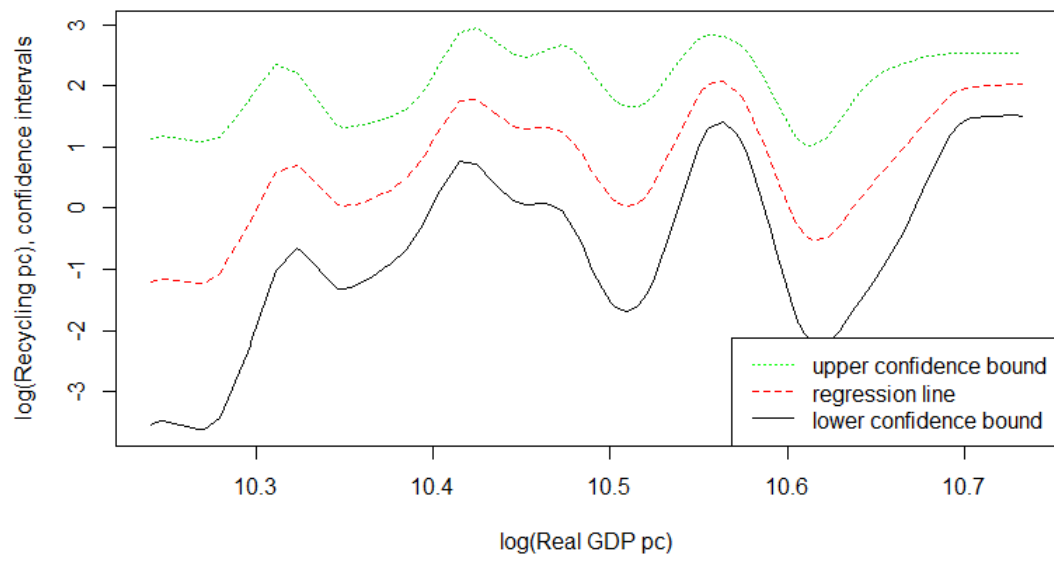
Rich States (1988 to 1997)



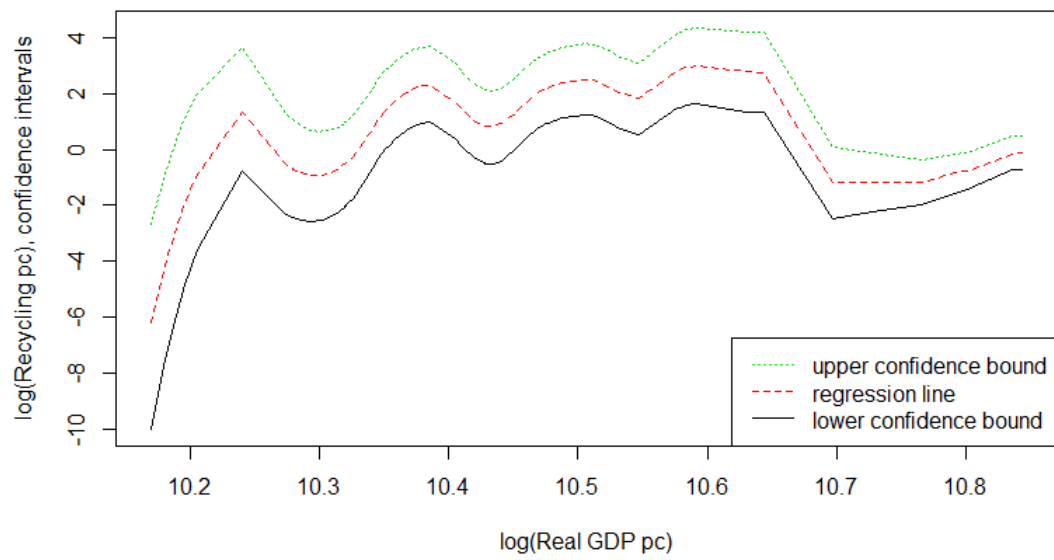
Rich States (1998 to 2007)



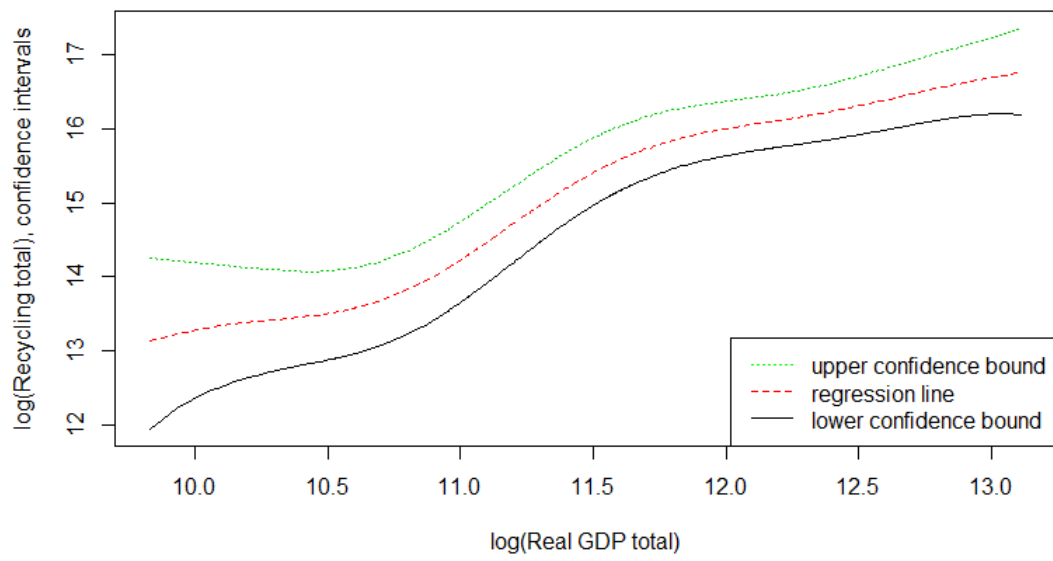
Middle Rich States (1988 to 1997)



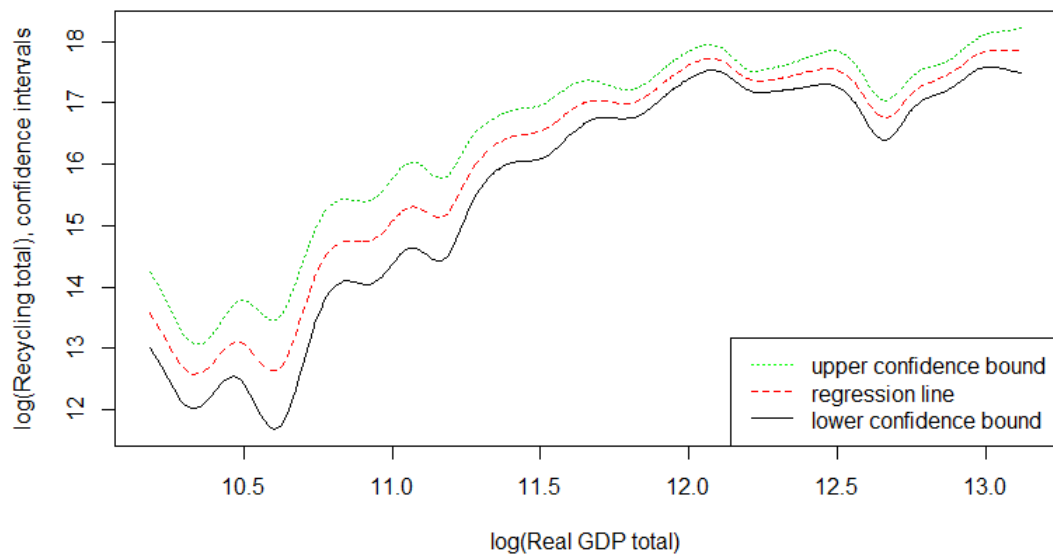
Middle Poor States (1988 to 1997)



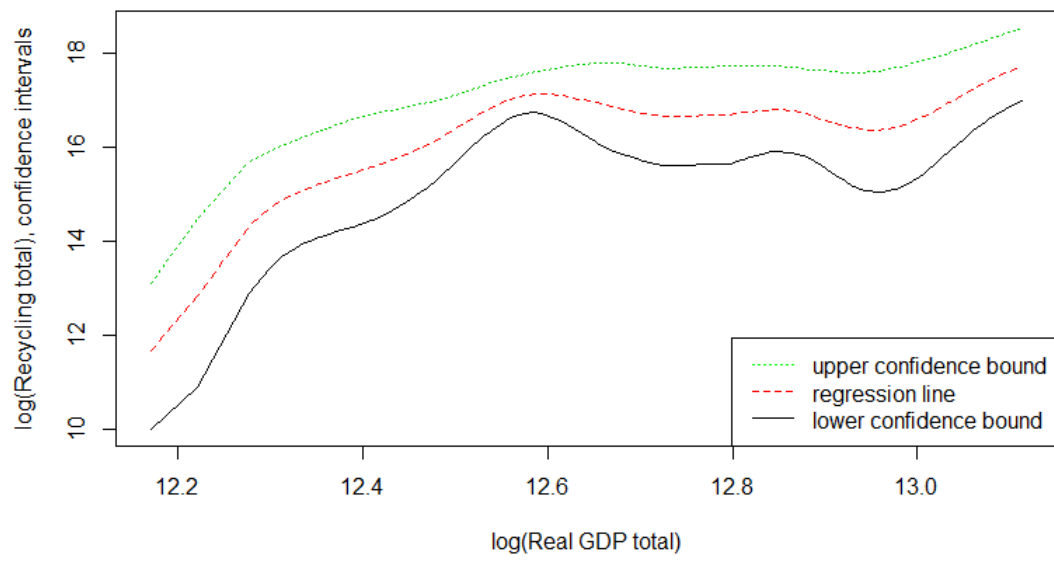
All States (1988 to 1997)



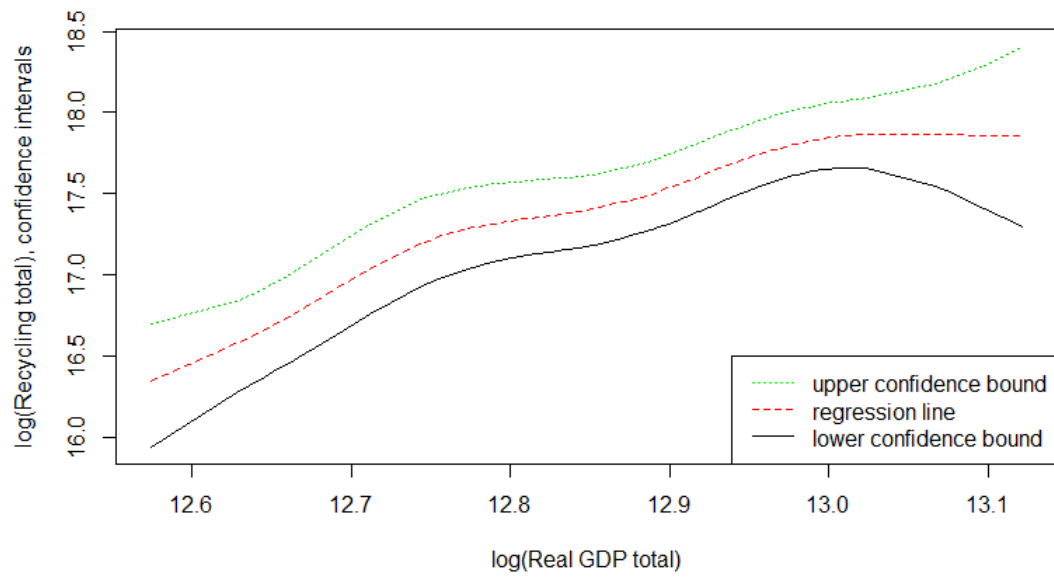
All States (1998 to 2007)



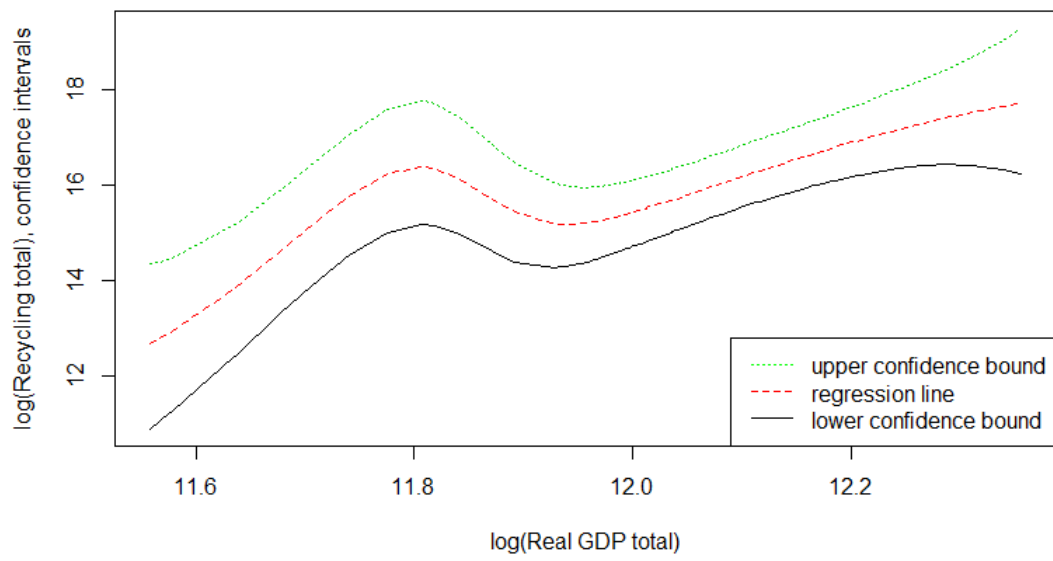
Rich States (1988 to 1997)



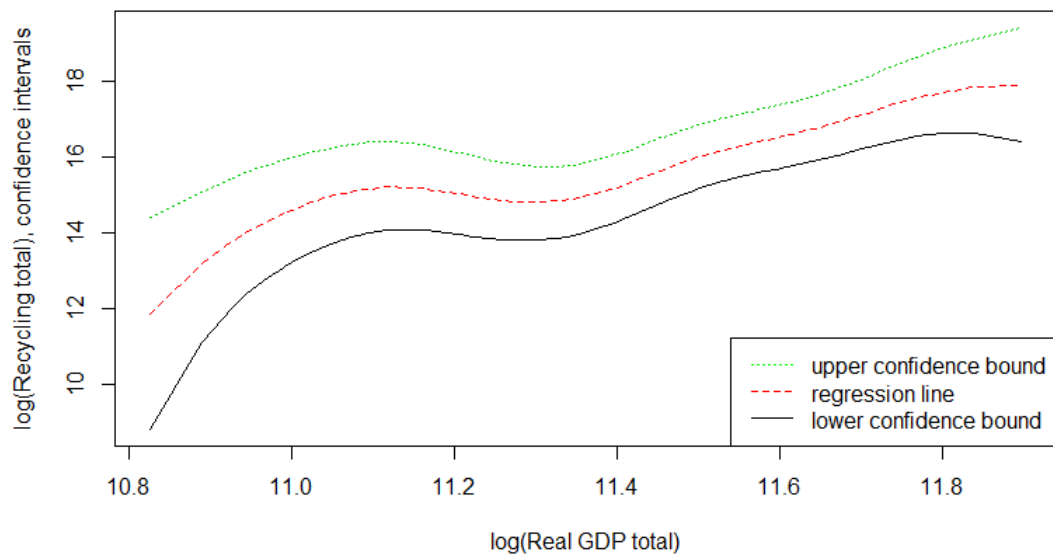
Rich States (1998 to 2007)

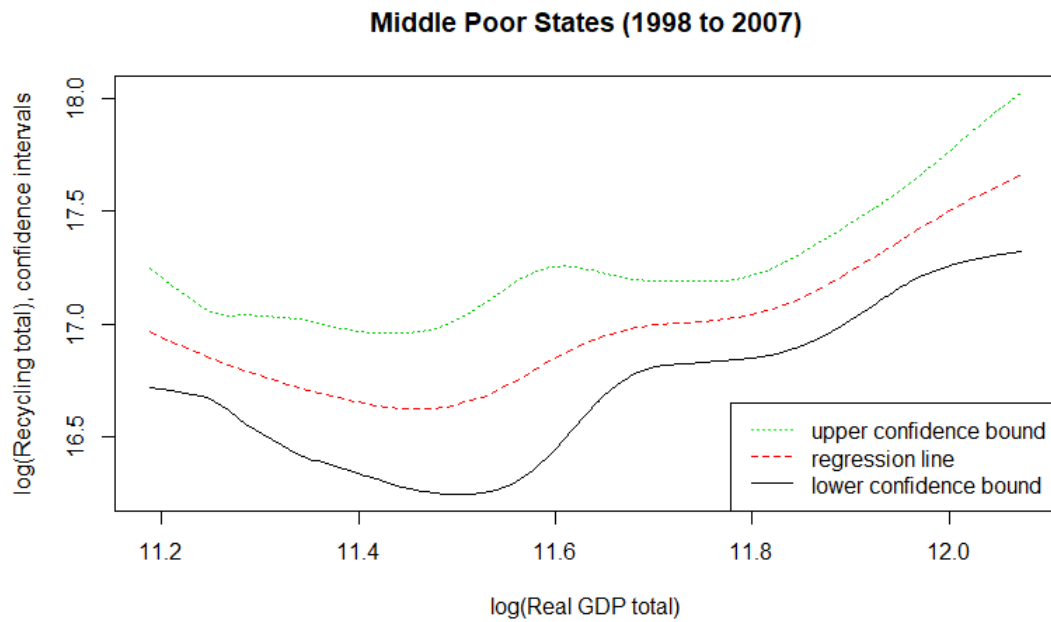


Middle Rich States (1988 to 1997)



Middle Poor States (1988 to 1997)

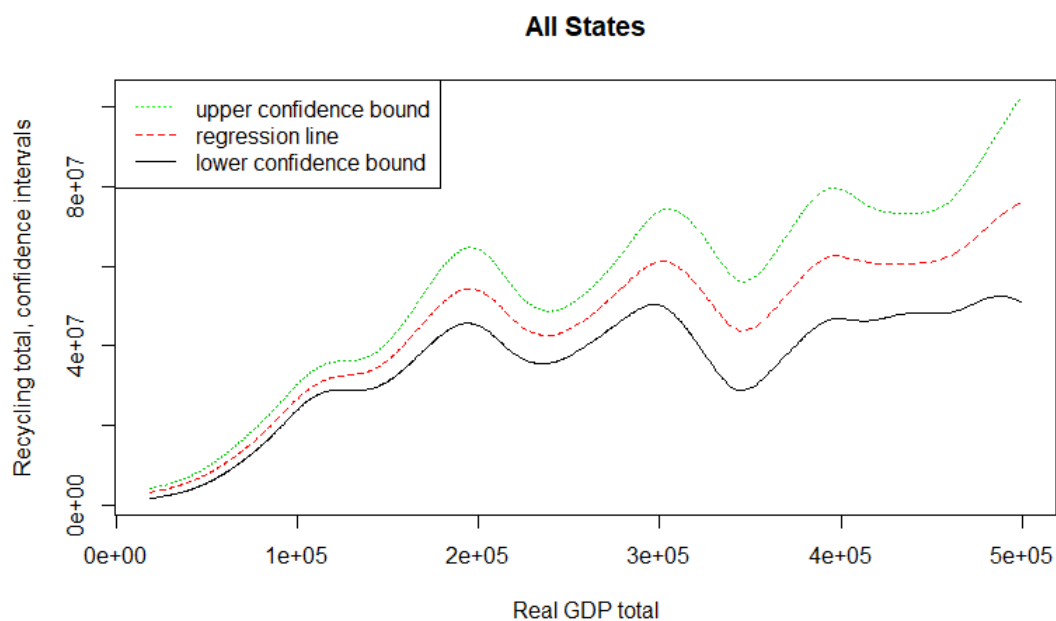




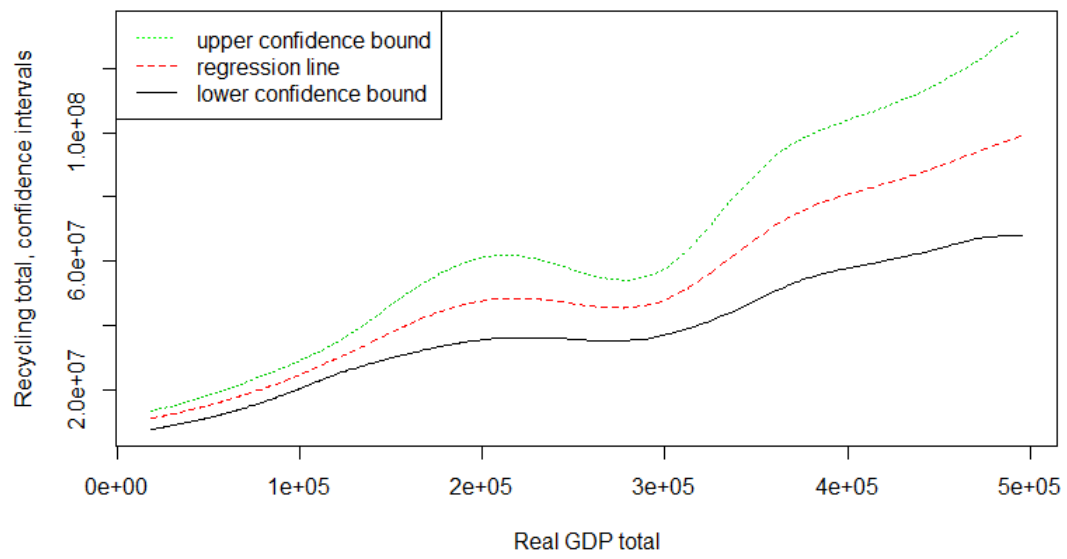
As we can see from all the previous plots, the dominant pattern between recycling and real GDP is an increasing curve. In addition, we see that some of these plots suffer from more fluctuations while other are smoother, but in general they keep the same pattern as the plots we presented in the main part of this paper.

Appendix D. Plots for All the Categories for the Non Log Values of Our Data

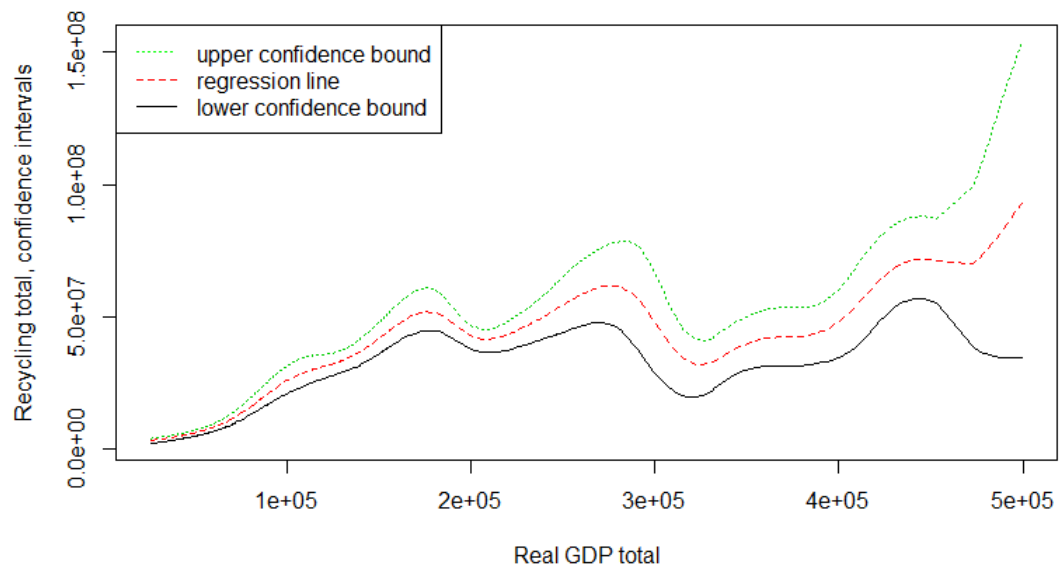
Last but not least, we will present the most representative plots we get for all the previously mentioned categories (separation based on the average income of the states as well as the different time periods for both total and per capita values). Once again, the pattern between recycling and real GDP does not change significantly and it remains an increasing curve, something that boost our previous results.



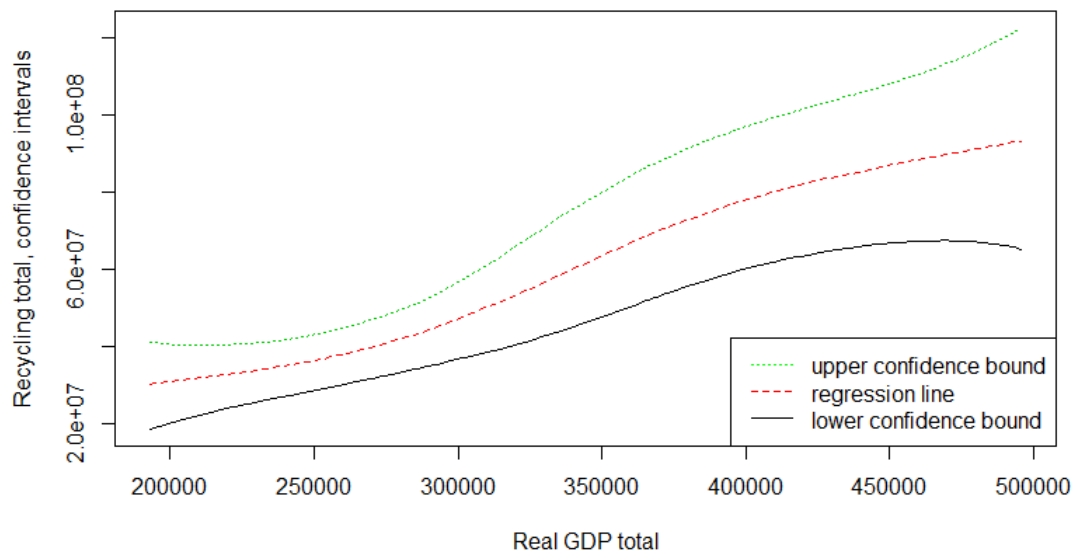
All States (1988 to 1997)



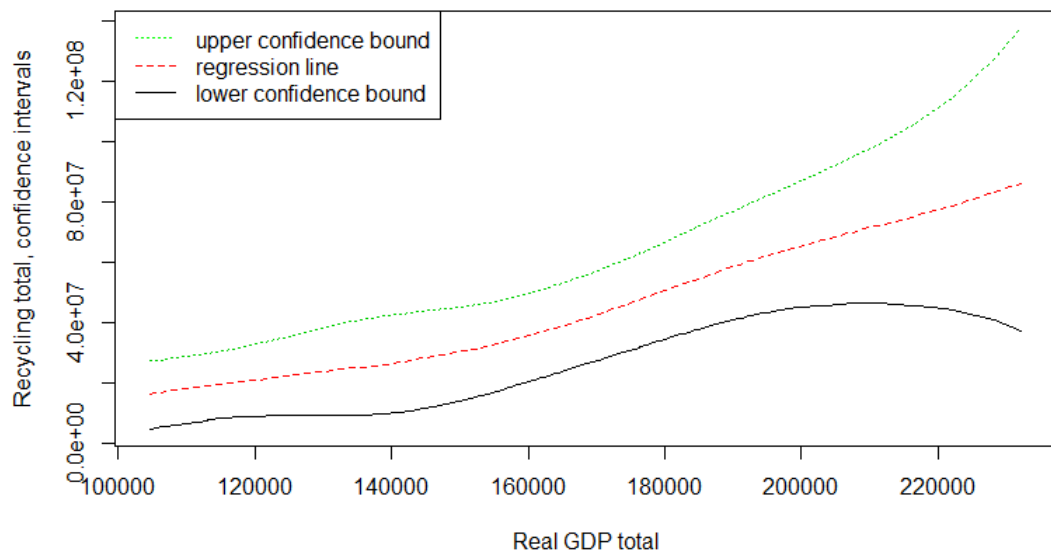
All States (1998 to 2007)



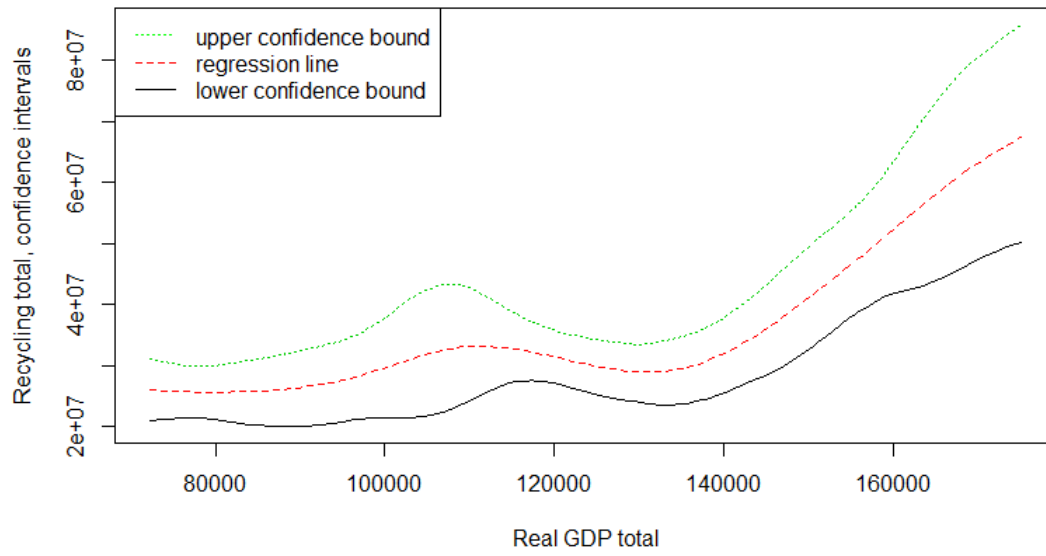
Rich States (1988 to 1997)



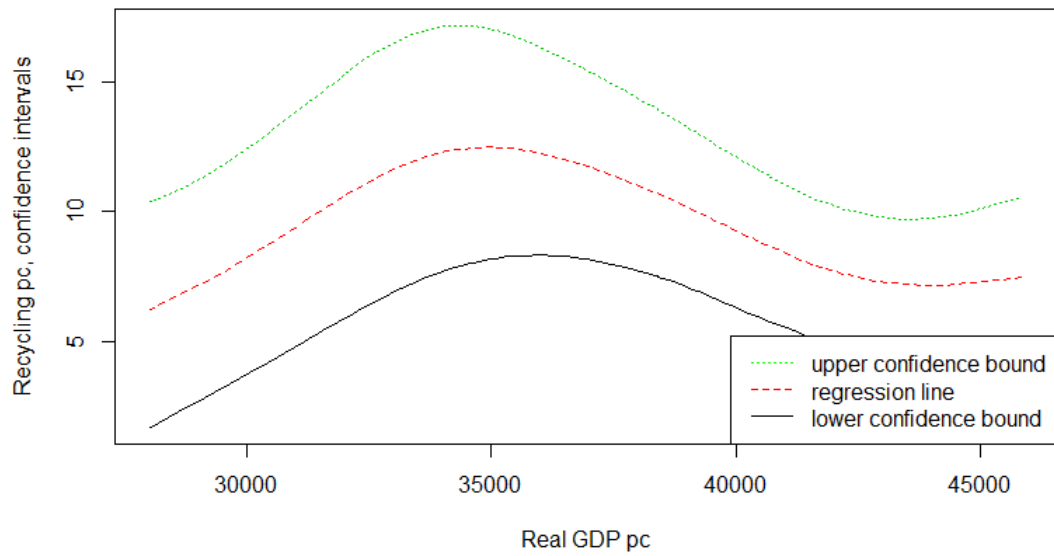
Middle Rich States (1988 to 1997)

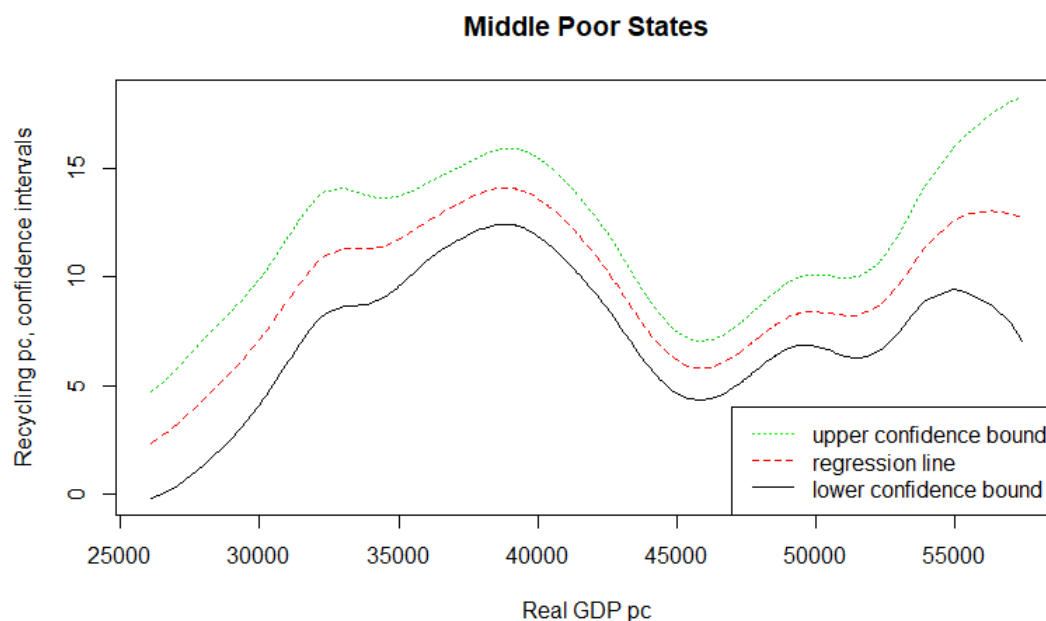


Middle Poor States (1998 to 2007)



Middle Rich States (1988 to 1997)





References

- Azomahou, Theophile, François Laisney, and Phu Nguyen Van. 2006. Economic development and CO₂ emissions: A nonparametric panel approach. *Journal of Public Economics* 90: 1347–63. [\[CrossRef\]](#)
- Baltagi, Badi H., and Dong Li. 2002. Series estimation of partially linear panel data models with fixed effects. *Annals of Economics and Finance* 3: 103–16.
- Boneva, Lena, Oliver Linton, and Michael Vogt. 2015. A semiparametric model for heterogeneous panel data with fixed effects. *Journal of Econometrics* 188: 327–45. [\[CrossRef\]](#)
- Cole, Matthew A., Anthony J. Rayner, and John M. Bates. 1997. The environmental kuznets curve: An empirical analysis. *Environment and Development Economics* 2: 401–16. [\[CrossRef\]](#)
- Criado, Carlos Ordás. 2008. Temporal and spatial homogeneity in air pollutants panel ekc estimations. *Environmental and Resource Economics* 40: 265–83. [\[CrossRef\]](#)
- De Bruyn, Sander M., and John B. Opschoor. 1997. Developments in the throughput-income relationship: theoretical and empirical observations. *Ecological Economics* 20: 255–68. [\[CrossRef\]](#)
- Dinda, Soumyananda. 2004. Environmental kuznets curve hypothesis: A survey. *Ecological Economics* 49: 431–55. [\[CrossRef\]](#)
- Di Vita, Giuseppe. 2001. Technological change, growth and waste recycling. *Energy Economics* 23: 549–67. [\[CrossRef\]](#)
- Esteve, Vicente, and Cecilio Tamarit. 2012. Threshold cointegration and nonlinear adjustment between CO₂ and income: The environmental kuznets curve in Spain, 1857–2007. *Energy Economics* 34: 2148–56. [\[CrossRef\]](#)
- Fakih, Ali, and Walid Marrouch. 2019. Environmental kuznets curve, a mirage? A non-parametric analysis for mena countries. *International Advances in Economic Research* 25: 113–19. [\[CrossRef\]](#)
- George, Donald AR, Brian Chi-ang Lin, and Yunmin Chen. 2015. A circular economy model of economic growth. *Environmental Modelling & Software* 73: 60–63.
- Grossman, Gene M., and Alan B. Krueger. 1991. *Environmental Impacts of a North American Free Trade Agreement*. Technical Report. Cambridge: National Bureau of Economic Research.
- Grossman, Gene M., and Alan B. Krueger. 1995. Economic growth and the environment. *The Quarterly Journal of Economics* 110: 353–77. [\[CrossRef\]](#)
- Harbaugh, William T., Arik Levinson, and David Molloy Wilson. 2002. Reexamining the empirical evidence for an environmental kuznets curve. *Review of Economics and Statistics* 84: 541–51. [\[CrossRef\]](#)
- Henderson, Daniel J., Raymond J. Carroll, and Qi Li. 2008. Nonparametric estimation and testing of fixed effects panel data models. *Journal of Econometrics* 144: 257–75. [\[CrossRef\]](#)

- Karlsson, Andreas. 2009. Bootstrap methods for bias correction and confidence interval estimation for nonlinear quantile regression of longitudinal data. *Journal of Statistical Computation and Simulation* 79: 1205–18. [CrossRef]
- Kasioumi, Myrto. 2020. *Economics of Recycling*. Guelph: Department of Economics, University of Guelph.
- Li, Qi, and Thanasis Stengos. 1996. Semiparametric estimation of partially linear panel data models. *Journal of Econometrics* 71: 389–97. [CrossRef]
- Lin, Zi, Qi Li, and Yiguo Sun. 2014. A consistent nonparametric test of parametric regression functional form in fixed effects panel data models. *Journal of Econometrics* 178: 167–79. [CrossRef]
- Millimet, Daniel L., John A. List, and Thanasis Stengos. 2003. The environmental kuznets curve: Real progress or misspecified models? *Review of Economics and Statistics* 85: 1038–47. [CrossRef]
- Pittel, Karen 2006. A Kuznets Curve for Recycling. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=928331 (accessed on 30 October 2020).
- Racine, Jeffrey Scott, Jeffrey Hart, and Qi Li. 2006. Testing the significance of categorical predictor variables in nonparametric regression models. *Econometric Reviews* 25: 523–44. [CrossRef]
- Roca, Jordi, Emilio Padilla, Mariona Farré, and Vittorio Galletto. 2001. Economic growth and atmospheric pollution in Spain: Discussing the environmental kuznets curve hypothesis. *Ecological Economics* 39: 85–99. [CrossRef]
- Selden, Thomas M., and Daqing Song. 1995. Neoclassical growth, the j curve for abatement, and the inverted u curve for pollution. *Journal of Environmental Economics and Management* 29: 162–68. [CrossRef]
- Shahbaz, Muhammad, Muhammad Shafiullah, Vassilios G. Papavassiliou, and Shawkat Hammoudeh. 2017. The CO₂-growth nexus revisited: A nonparametric analysis for the G7 economies over nearly two centuries. *Energy Economics* 65: 183–93. [CrossRef]
- Stern, David I. 1998. Progress on the environmental kuznets curve? *Environment and Development Economics* 3: 173–96. [CrossRef]
- Sun, Yiguo, Raymond J. Carroll, and Dingding Li. 2009. Semiparametric estimation of fixed-effects panel data varying coefficient models. In *Nonparametric Econometric Methods*. Bingley: Emerald Group Publishing Limited.
- Yavapolkul, Navin 2005. *Environmental Kuznet Curve: Empirical Investigation Using Non-Parametric Approach*. Berkeley: Department of Agricultural and Resource Economics, University of California.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).