CORE DISCUSSION PAPER
2004/51

The Environmental Kuznets Curve Semi-Parametrically Revisited

Luisito BERTINELLI[1,2], Eric Strobl[2,3]

July 2004

Abstract

This paper re-examines the existence of an Environmental Kuznets Curve (EKC) across countries using a semi-parametric regression estimator, which places no restrictions on the functional form. Our results using cross-country panel data on Sulfur and Carbon Dioxide strongly suggest that the relationship between wealth and environmental degradation is not bell-shaped, as suggested by an EKC. Rather that there is a positive link for the very poorest countries and no clear relationship for richer countries.

JEL Code: O13, Q32, Q56
Keywords: environmental Kuznets curve, semi-parametric kernel regression

[1]*CREA, Université du Luxembourg.*
[2]CORE, *Université Catholique de Louvain, Belgium.*
[3] Corresponding author: Eric Strobl, CORE, Université Catholique de Louvain,Voie du Roman Pays 34, B-1348 Louvain-la-Neuve, Belgium. Email: strobl@core.ucl.ac.be

Comments by Thierry Brechet are gratefully acknowledged.

1. Introduction

Since the seminal paper by Grossman and Kruger (1991) there has been considerable interest in the relationship between economic growth and environmental pollution. Importantly the authors showed that the link between these follows an inverted U-shaped pattern, now commonly referred to as the Environmental Kuznets Curve (EKC). This finding suggests that lower income regions are 'too poor to be green', and only when these become rich enough will the benefits from a clean environment outweigh its costs. Much of the subsequent literature has focused on estimating the actual turning point and/or investigated whether the shape may depend on the econometric techniques and assumptions employed. One of the main concerns for the latter has been over the appropriate underlying functional form (see Dasgupta et al. (2002)), where normally researchers allow for the possible non-linearity by introducing higher order terms. However, recently Millimet et al (2003) have shown with data for US states that such parametric modeling can be rejected in favour of a semi-parametric estimator, which does not impose any a priori restriction on the functional form of the relationship.

In this paper we employ such a semi-parametric estimator to investigate the existence of the EKC in a cross-country context. So far, results on cross-country studies measuring the relationship between economic growth and pollution have led to rather mixed results concerning the existence of an EKC (see for instance Shafik (1994), Selden and Song (1994), Grossman and Kruger (1995), Holtz-Eakin and Selden (1995), Stern et al (2001), Hettige et al. (2000), Harbaugh and Turton (2002) among others for recent evidence). However, it must be noted that all of these have imposed relatively restrictive functional forms. Our findings using the semi-parametric estimator suggest that in a cross-country sense, at least in terms of measuring pollution by sulfur and carbon

dioxide, the link between environmental pollution and economic growth is actually monotonically increasing for low levels of GDP/capita, and flat thereafter.[4]

2. Methodology and Data

Most studies examining the EKC have been concerned with estimating the following equation:

$$P_{it} = \alpha + g(Y_{it}) + Z_{it}\delta + v_i + u_{it} \qquad (1)$$

where P_{it} is some proxy of environmental degradation (per capita) in country i at time t, Y_{it} is a measure of wealth, usually real per capita GDP at the start of the period, Z_{it} is a vector of variables that controls for other factors, v_i is a unit-specific residual and u_{it} is a disturbance term.

In order to allow for the possible non-linearity of $g(Y)$ most analyses have simply included a second and third order polynomial of Y. As in Millimet *et al.* (2003) we instead implement Robinson's (1988) semi-parametric Kernel regression estimator (see Blundell and Duncan (1998) for details and a helpful discussion of the implementation of this method). Accordingly, if we allow $g()$ in (1) to be a smooth and continuous, possibly non-linear, function of Y, and assume that the other control variables captured by the vector Z have a linear effect on P, then the estimation of $g(Y)$ can be made by:

$$\hat{g}(Y) = \hat{m}_P(Y) - \hat{\delta}\,\hat{m}_Z(Y) \qquad (2)$$

where $\hat{m}_P(Y)$ and $\hat{m}_Z(Y)$ are the (non-parametric) Nadaraya-Watson estimates (Nadaraya (1964) and Watson (1964)) of $E(P/Y)$ and $E(Z/Y)$, such that, for a given continuous, bounded, and real shape function, $K_h()$ integrating to one with a smoothing parameter h, $\hat{m}_P(Y)$ (and similarly $\hat{m}_Z(Y)$) is defined as:

[4] This result echoes the recent skepticism raised by Stern (2004) over the existence of an EKC internationally in his review of the literature.

$$\hat{m}_P(Y) = n^{-1} \frac{\sum_{i=1}^{n} K_h(y - Y_i) P_i}{\sum_{i=1}^{n} K_h(y - Y_i)} \tag{3}$$

and $\hat{\delta}$ is the OLS estimator of:

$$P - \hat{m}_P(Y) = \delta(Z - \hat{m}_Z(Y)) + \varepsilon \tag{4}$$

The appeal of the estimator (2) lies in its very flexible approach to non-linearity by allowing the relationship between *P* and *Y* to vary over all values of *Y* after purging the effects of other explanatory variables. Specifically, this technique entails first purging the effect of the other factors *Z* from the relationship between *P* and *Y* and then estimating the regression function of *P* on *Y* at a particular point by locally fitting constants to the data via weighted least squares, where those observations closer to the chosen point have more influence on the regression estimate than those further away, as determined by the choice of *h*.[5]

One should note that, given its semi-parametric nature, the estimate of $\hat{g}(Y)$ cannot be subjected to the kind of standard statistical tests (such as an *F-test* or a *t-test*) of parametric regressions. However, it is possible to calculate upper and lower point-wise confidence intervals, as suggested by Haerdle (1990). Specifically, we calculated bands at the *1st* and *99th* percentiles along the range of initial income and at every fifth percentile in between. Choosing points according to the distribution of observations also allows one to gauge how the density of the sample affects the approximation bias, since these are inversely related. It should also be noted that the use of the semi-parametric estimator just described leads to a problem of non-identification of an unrestricted intercept term, which leads to a scaling issue when comparing our semiparametric results with any

[5] For all estimations we use a Gaussian kernel for *Kh* and the optimal smoothing parameter suggested by Fox (1990).

parametric alternative. As in Millimet et al (2003) we deal with this issue by standardising our data (relatively to the full sample).

The data used for our analysis consists of annual information on sulfur emissions taken from the Historical Global Sulfur Emissions database (Lefohn et al, 1999), carbon dioxide emissions taken from World Resource Institute (People and Ecosystems CD-rom), and GDP per capita figures (real GDP per capita in constant dollars, base year 1985) taken from the Penn World Tables 6.1. As other explanatory variables, *Z,* we included time and country specific dummies.[6] Together this provides us with a total sample size of 3976 observations consisting of 122 countries (among which 95 LDCs), and 3336 observations for 108 countries (among which 81 LDCs) for the carbon dioxide sample and the sulfur sample respectively, over period 1950-1990.[7]

3. Results

The graph of $\hat{g}(Y)$ for annual sulfur emissions along with the confidence bands is shown in Figure 1. Accordingly, there is little evidence of a bell shaped link between carbon dioxide emissions and GDP/capita.[8] Rather, in contrast to an 'EKC', our estimate appears to be decidedly linear, environmental pollution increasing with country wealth for low levels of GDP/capita, and is flat thereafter. The only exception of a change in this trend is for very high GDP per capita values, but, as the distance between the confidence intervals suggest, this portion of the curve is likely to be very poorly estimated because of the lower number of observations around it and the fact that it is near the endpoint. The accuracy of the estimate of *g(Y)* at *Y* is positively related to the density of other observations around that point. Furthermore, the approximation bias is larger at the boundaries (see Wand and Jones, 1995).

[6] We also experimented with other controls, such as openness and population growth rate. However, these made little qualitatively or quantitative difference and only reduced sample size.
[7] A list of all the countries is provided in the appendix.
[8] One should note that our panel is unbalanced. However 80 per cent of the countries had more than three quarters of the observations in the carbon dioxide sample (70 per cent for the sulfur sample). Missing data are essentially concentrated in former colonial countries for years before 1960.

The estimates using carbon dioxide emissions as a proxy of pollution show a similar picture in Figure 2. The emission of pollutants slightly increases as countries with very low GDP/capita grow richer, but flattens thereafter. The only part of the estimated curve that shows again some signs of positive link is in the case around observations of pollution for the very richest countries, but, under similar argument as before, these are likely to be poor estimates of the relationship.

Our estimates for both sulfur and carbon dioxide emissions suggest that the relationship between pollution and wealth is in fact linear. We also estimated (1) but simply including the pollutant proxy of order one for both cases and plotted our estimated regression line in the graphs. The adjusted line is almost horizontal and crosses the y-axes at about 0 for the two pollutants cases.[9] Thus, these reiterate the linearity of the relationships.

Finally, we estimated specification (1) imposing $g(Y) = \beta Y + \beta' Y^2$, using OLS and fixed effect estimators. Results are displayed in Table 1. In both pollutant cases, we found a bell-shaped curve, supporting hence the idea of the EKC. Maximum values of the bell-shaped curve are always in the upper range of GDP/capita (last row of Table 1). However, given our semi-parametric estimation, we have seen that this result is essentially driven by the shape of the curve for the very poorest countries, but does not appear over the whole range of feasible values of GDP/capita.

4. Concluding Remarks

Using a semi-parametric estimator we estimate the Environmental Kuznets Curve for a panel of countries. Our results strongly suggest the absence of a relationship over almost the whole range of GDP/capita values, in contrast to the often-argued bell-shaped link. Thus, historical evidence seems to indicate that the result according to which richer countries spur environmental awareness is not robust.

[9] The slope coefficients are 0.0000002 and 0.0004 in the sulfur respectively in the carbon dioxide case.

References

Blundell, R. and A. Duncan (1998), Kernel Regression in Empirical Microeconomics, *Journal of Human Resources*, 33(1), pp.62-87.

S. Dasgupta, B. Laplante, H. Wang and D. Wheeler (2002), Confronting the environmental Kuznets curve, *Journal of Economic Perspectives*, 16, 147-68.

J. Fox (1990), Describing univariate distributions, in J. Fox and J.S. Long (Eds.), Modern methods of data analysis (Sage).

G. M. Grossman and A. B. Krueger (1991), Environmental Impacts of a North American Free Trade Agreement, National Bureau of Economic Research Working Paper 3914, NBER, Cambridge MA.

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

W. Haerdle (1990), Applied Nonparametric Regression, Cambridge University Press.

W. Harbaugh, A. Levinson, and D. M. Wilson (2002), Reexamining the empirical evidence for an environmental Kuznets curve, *Review of Economics and Statistics*, 84, 541-551.

H. Hettige, M. Mani, and D. Wheeler (2000), Industrial pollution in economic development: the environmental Kuznets curve revisited, *Journal of Development Economics*, 62, 445-76.

D. Holtz-Eakin and T. M. Selden (1995), Stoking the fires? CO2 emissions and economic growth, *Journal of Public Economics*, 57, 85-101.

Kuznets, S. (1955), Economic growth and income inequality, *American Economic Review,* 49, 1-28.

AS. Lefohn, JD. Husar and RB. Husar (1999), Estimating historical anthropogenic global sulfur emission patterns for the period 1850-1990, *Atmospheric Environment* 33: 3435-3444.

DL. Millimet, JA. List and T. Stengos (2003), The Environmental Kuznets Curve: Real Progress or Misspecified Models?, *Review of Economics and Statistics* 85 (4): 1038-1047.

E. Nadarya (1964), On estimating regression, Theory of Probability and its Applications 10: 186-190.

P. Robinson (1988), Root-N-Consistent Semiparametric Regression, *Econometrica* 56: 931-54.

DI. Stern (2004), The Environmental Kuznets Curve, in Encyclopedia of Energy, C. J. Cleveland (Ed.), Academic Press, San Diego CA.

Stern, D. I. and M. S. Common (2001), Is there an environmental Kuznets curve for sulfur? *Journal of Environmental Economics and Management*, 41, 162-78.

T. M. Selden and D. Song (1994), 'Environmental quality and development: Is there a Kuznets curve for air pollution?' *Journal of Environmental Economics and Management*, 27, 147-62.

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

MP. Wand and M.C. Jones (1995), Kernel Smoothing, Chapman & Hall.
GS. Watson (1964), Smooth regression analysis, Sankhya, Series A 26: 359-37

Figure 1: Sulfur (1950-90)

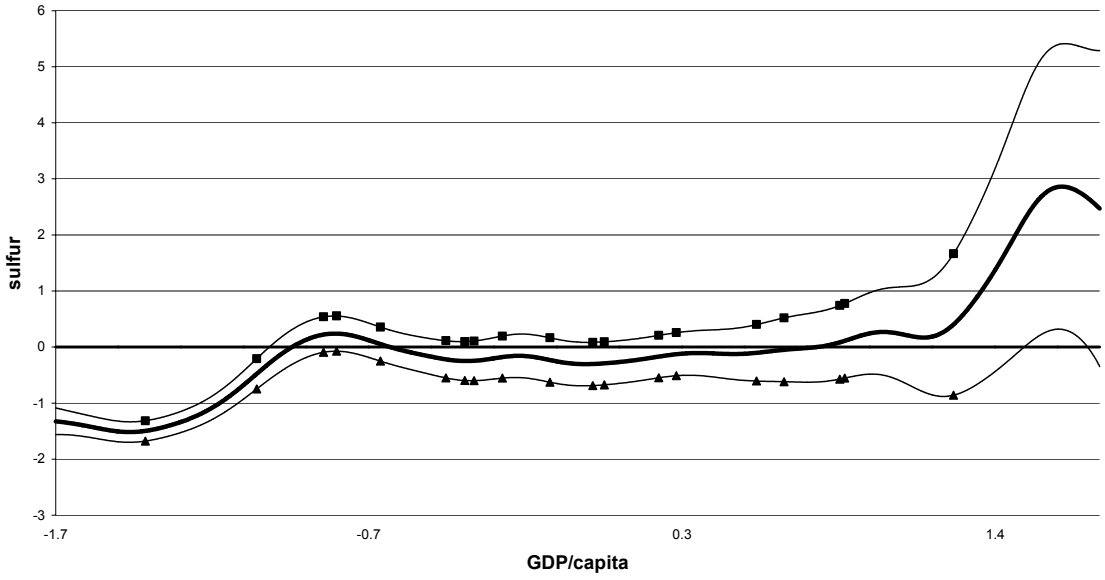

Figure 2: Carbon dioxide (1950-90)

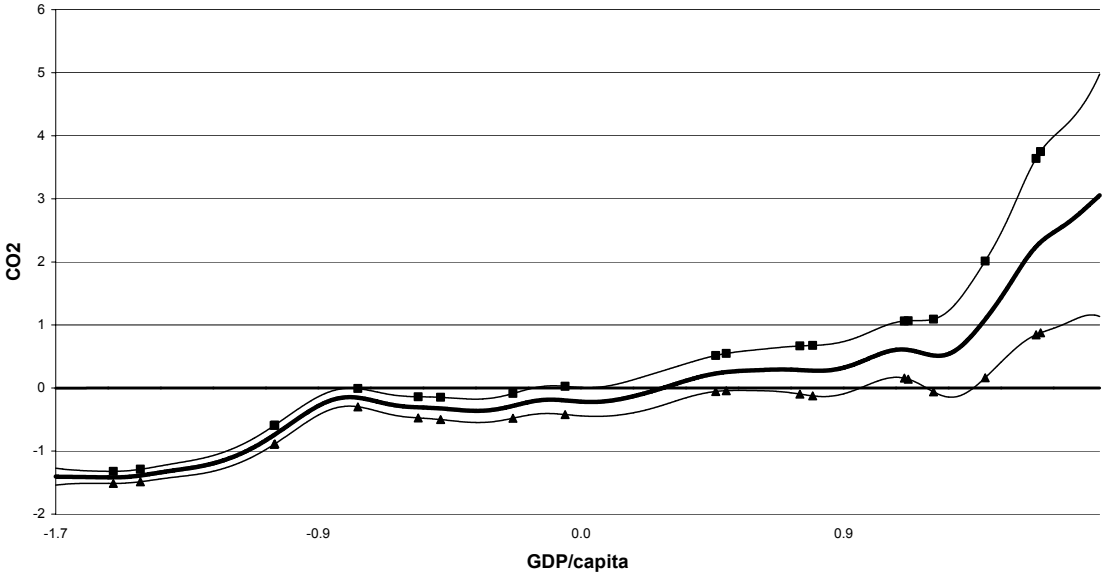

Table 1: OLS and fixed effect regressions

	Sulfur		Carbon dioxide	
	OLS	**Fixed effects**	**OLS**	**Fixed effects**
	(1)	**(2)**	**(3)**	**(4)**
GDP/capita	*0.000****	*0.000****	*0.001****	*0.001****
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
(GDP/capita)^2	*-0.000****	*-0.000****	*-0.000****	*-0.000****
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*
Constant	*0.001*	*0.007****	*-0.940****	*-0.509***
	(0.001)	*(0.002)*	*(0.068)*	*(0.218)*
N° of obs.	*3336*	*3336*	*3976*	*3976*
R-squared	*0.13*	*0.028*	*0.69*	*0.50*
N° of groups		*106*		*122*
Maximum	*$15315*	*$13765*	*$37927*	*$26467*

Notes: (1) Standard errors in parentheses. (2) ***, **, and * indicate 1, 5, and 10 per cent significance levels.

Appendix: Country list

Algeria
Angola
Antigua Barbuda
Argentina
Australia
Austria
Bangladesh
Barbados
Belgium
Belize
Benin
Bolivia
Botswana
Brazil
Burkina Faso
Burundi
Cameroon
Canada
Cape Verde
Central African Republic
Chad
Chile
China
Colombia
Comoros
Congo
Costa Rica
Ivory Coast
Cuba
Cyprus
Czechoslovakia
Denmark
Dominica
Dominican Republic
Ecuador
Egypt
El Salvador
Equatorial Guinea
Ethiopia
Fiji
Finland
France
Gabon
Gambia
Ghana
Greece
Grenada
Guatemala
Guinea
Guinea Bissau
Guyana
Haiti
Honduras
Hong Kong
Hungary
Iceland
India
Indonesia
Iran
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kenya
South Korea
Lebanon
Macau
Madagascar
Malawi
Malaysia
Mali
Mauritania
Mauritius
Morocco
Mozambique
Nepal
Netherlands
New Zealand
Nicaragua
Niger
Nigeria
Norway
Pakistan
Panama
Papua New Guinea
Paraguay
Peru
Philippines
Poland
Portugal
Romania
Rwanda
Sao Tome Principe
Senegal
Seychelles
Sierra Leone
Singapore
South Africa
Spain
Sri Lanka
St Lucia
St Vincent Grenadines
Sweden
Switzerland
Syria
Tanzania
Thailand
Togo
Trinidad Tobago
Tunisia
Turkey
Uganda
United Kingdom
United States
Uruguay
Venezuela
Vietnam
Yemen
Zambia
Zimbabwe