# Documents
# de travail

« **A Structural nonparametric reappraisal of the CO$_2$
emissions-income relationships** »

Auteurs

**Theophile T. Azomahou, Micheline Goedhuys, Phu Nguyen-Van**

CNRS

Nancy-Université
*Université Nancy 2*

UNIVERSITÉ DE STRASBOURG

# A structural nonparametric reappraisal of the $CO_2$ emissions-income relationship

Théophile T. Azomahou[a]*, Micheline Goedhuys[a], Phu Nguyen-Van[b]


[a] UNU-MERIT, Maastricht University
Keizer Karelplein 19, 6211 TC Maastricht, the Netherlands

[b] BETA-CNRS, Université de Strasbourg
61 avenue de la Forêt Noire, F-67085 Strasbourg Cedex, France

December, 2009

**Abstract**

Relying on a structural nonparametric estimation, we show that $CO_2$ emissions clearly increase with income at low income levels. For higher income levels, we observe a decreasing relationship, though not significant. We also find that $CO_2$ emissions monotonically increases with energy use at a decreasing rate.

*Keywords*: Nonparametric triangular systems, EKC; Energy use; $CO_2$ emissions

*JEL Classification*: C14; O13

*Correspondence: Azomahou T., Tel. +31 433884440, Fax +31 433884499, e-mail: azomahou@merit.unu.edu; M. Goedhuys (micheline.goedhuys@skynet.be), P. Nguyen-Van (nguyen-van@unistra.fr)

# 1    Introduction

The concept of the Environmental Kuznets Curve (EKC), introduced by Grossman and Krueger (1995), is a hypothesized relationship among various indicators of environmental degradation and income per capita. In its basic specification, it assumes that during the early stages of economic development, environmental damage and pollution increase. Beyond some level of income per capita, also termed turning point, the trend reverses and economic development leads to environmental quality improvement.

While the inverted U-shape of the EKC has been confirmed for several environmental quality indicators (see Azomahou et al., 2006, for a literature review), for $CO_2$ emissions a lot of controversy remains. Indeed, the majority of studies mainly based on reduced-form single-equation models find emissions to monotonically increase with income. These models however do not account for possible feedback effects of the environment to economic growth, or for the fact that the economy and the environment are jointly determined, as explained by Perrings (1987). Omission to account for feedback effects may lead to simultaneity bias and inconsistent estimates (Stern et al. 1996). As far as we know, Liu (2005) provided the first study of the relationship between $CO_2$ emissions and income based on a parametric two-equations system. The author underlined the crucial role of energy use in the system and concludes on a negative relationship between $CO_2$ emissions and income. However, given the sample used by Liu (2005) – 24 OECD countries over the period 1975-1990 – this result may not be representative to conclude on the existence of a $CO_2$ emissions-EKC.

In this study, we propose a structural nonparametric estimation of the emissions-EKC, relying on the nonparametric triangular system of Newey et al. (1999). By using a structural model, we improve the specification to account for simultaneity between income and emissions. In addition, by relying on a nonparametric framework we allow for non-linearities of unknown form in the income-environment relationship. It is worth noticing that the nonparametric methodology has been applied to reduced form EKC models (Millimet et al., 2003; Bertinelli and Strobl, 2005; Azomahou et al., 2006; Nguyen Van, 2009) showing indeed important non-linearities. However, to the best of our knowledge, our study is the first to analyse the $CO_2$ emissions-EKC in a structural nonparametric specification.

We apply this methodology to panel data of 107 countries, both low and high income countries, over a 44 year period (1961-2004), thus having an excellent coverage in time and income dispersion. Although our results are not supportive for the existence of an EKC for $CO_2$ emissions, we find that $CO_2$ emissions firstly increase with income at low income levels and then become delinked with income at high income levels. We also find that $CO_2$ emissions monotonically increases with energy use at a decreasing rate.

# 2    Structural nonparametric specification

We consider the triangular nonparametric simultaneous specification of Newey et al. (1999):

$$y = m(x, z_0) + \varepsilon \tag{1}$$
$$x = \pi(\mathbf{z}) + u, \qquad \mathbb{E}(\varepsilon|u, \mathbf{z}) = \mathbb{E}(\varepsilon|u) \neq 0, \quad \mathbb{E}(u|\mathbf{z}) = 0 \tag{2}$$

where $y$, $x$ and $z_0$ denote respectively $CO_2$ emissions per capita, GDP per capita and energy use per capita; $\mathbf{z}$ is a set of instruments that includes $z_0$. The system (1)-(2) is a generalization of the limited information simultaneous equations model to allow for structural nonparametric relation $m(x, z_0)$ between variables $y$, $x$ and $z_0$, and a nonparametric reduced form $\pi(\mathbf{z})$. The conditional expectation of equation (1) yields the integral equation:

$$\mathbb{E}(y|\mathbf{z}) \equiv \pi(\mathbf{z}) = \mathbb{E}[m(x, z_0)|\mathbf{z}] = \int m(x, z_0) F(dx|\mathbf{z}) \tag{3}$$

where $F$ denotes the conditional cumulative distribution function of $x$ given $\mathbf{z}$. Thus, functions $\pi$ and $F$ are the nonparametric generalization of the reduced forms for $y$ and $x$. Newey et al. (1999) discussed the identification of the system (1)-(2).[1] Starting from a preliminary estimation of the reduced forms $\hat{\pi}$ and $\hat{F}$:

$$\hat{\pi}(\mathbf{z}) = \int m(x, z_0) \hat{F}(dx|\mathbf{z}), \tag{4}$$

the authors developed an estimator for $\hat{m}$ that overcomes the well known ill-posed problem.[2] In order to apply this methodology to analyze the EKC, we specify a generalized additive model (hereafter GAM) for fixed effects panel data.[3] For equation (1), the GAM is

$$y_{it} = \sum_{j=1}^{p} m_j(w_{it}^j) + \mu_i + \varepsilon_{it}, \qquad i = 1, \cdots, N, \quad t = 1, \cdots, T \tag{5}$$

where $w_{it}^j$ is the $j$th component ($j = 1, \cdots, p$) of $\mathbf{w}_{it} \equiv (x_{it}, z_{0it})$. For equation (2) we use a semiparametric GAM specification the structure of which is given by

$$x_{it} = \sum_{k=1}^{q} \pi_j\left(z_{1it}^k\right) + \mathbf{z}_{2it}'\gamma + \lambda_i + u_{it}, \qquad i = 1, \cdots, N, \quad t = 1, \cdots, T \tag{6}$$

where $z_{1it}^k$s is the $k$th component ($k = 1, \cdots, q$) of the set of continuous instruments $\mathbf{z}_1$ and $\mathbf{z}_{2it}$ corresponds to other instruments which do enter linearly in the specification. The unobserved fixed effects $\mu_i$ and $\lambda_i$ can be eliminated by first differences:

$$y_{it} - y_{i,t-1} = \sum_{j=1}^{p} \left[m_j(w_{it}^j) - m_j(w_{i,t-1}^j)\right] + \varepsilon_{it} - \varepsilon_{i,t-1} \tag{7}$$

$$x_{it} - x_{i,t-1} = \sum_{k=1}^{q} \left[\pi_j\left(z_{1it}^k\right) - \pi_j\left(z_{1i,t-1}^k\right)\right] + (\mathbf{z}_{2it} - \mathbf{z}_{2i,t-1})'\gamma + u_{it} - u_{i,t-1} \tag{8}$$

Observe that the method of Newey et al. (1999) consists of estimating equation (7) by including an additional control variable which is the first difference residuals $\hat{u}_{it} - \hat{u}_{i,t-1}$ computed from equation (8). Therefore, estimation of equation (7) involves in total five univariate unknown functions associated to $x_{it}$, $x_{i,t-1}$, $z_{0it}$, $z_{0i,t-1}$, and $\hat{u}_{it} - \hat{u}_{i,t-1}$. We perform estimation in two steps: (i) we construct semiparametric first differences residuals $\hat{u}_{it} - \hat{u}_{i,t-1}$ of equation (8) where the parametric estimates $\hat{\gamma}$ are obtained using the Robinson's (1988) procedure. (ii) We estimate

---

[1] Identification is needed as $\pi$ and $F$ are functionals of the distribution of observables $(y, x, \mathbf{z})$.

[2] The ill-posed inverse problem follows from non-continuity of $\hat{m}$. Indeed, lack of continuity of $\hat{\pi}$ and $\hat{F}$ can translate into large inaccuracies in $\hat{m}$.

[3] See, e.g., Hastie and Tibshirani (1990) for further details on GAM.

the nonparametric model in equation (7) using the residuals $\hat{u}_{it} - \hat{u}_{i,t-1}$ from (i) as additional regressor. In practice, we base our nonparametric estimation on the 'backfitting algorithm' (Hastie and Tibshirani, 1990). Furthermore, as $m_j$ is estimated twice, denoted as $\hat{m}_j^{(1)}$ and $\hat{m}_j^{(2)}$ for $w_{it}^j$ and $w_{i,t-1}^j$ respectively, a simple and more precise estimator of $m_j$ can be obtained by a weighted average: $\hat{m}_j = (\hat{m}_j^{(1)} + \hat{m}_j^{(2)})/2$.

# 3 Data and estimation results

## 3.1 Data

The data used for the analysis consist of an unbalanced panel of 107 countries, both developed and developing countries and spanning the period 1961-2004. The data are obtained from the *World Development Indicators* database 2007. For equation (6), GDP per capita, measured in constant 2000 US dollars, is used as dependent variable. The variables included in $\mathbf{z}_1$ are primary energy use per capita (in kilotons of oil equivalent), foreign direct investment (net inflows), population density, trade openness (imports plus exports divided by GDP). Variables included in $\mathbf{z}_2$ are regional dummies (East Asia & Pacific, Europe & Central Asia used as reference, Latin America & Caribbean, Middle East & North Africa, North America, South Asia, Sub-Saharan Africa) and year dummies, to control for autonomous technological change and macroeconomic effects. For equation (5), we use $CO_2$ emissions per capita (in metric tons) as dependent variable. Similar to Liu (2005), the regressors included in equation (5) are GDP per capita and primary energy use per capita. Statistics in Table 1 (for 1961, 1982, 2004 and for the whole sample) show that on average, GDP and $CO_2$ emissions increase over time while energy use remains stable.

**Insert Table 1 here**

## 3.2 Estimation results

Estimation results are presented in Figure 1 and Table 2. The curves displayed in Figure 1 correspond to the structural nonparametric functions discussed in the previous section.[4] We use the 'gain' statistic (see Hastie and Tibshirani, 1990, for further details) to test the significance of non-linearity in the econometric specification.[5] Table 2 summarizes the 'gain' statistics. As can be checked from the $p$-values, all the parametric functions are strongly rejected in favor of the nonparametric counterparts, meaning that our nonparametric specification provides a better approximation of the data.

**Insert Table 2 here**

From Figure 1(a) we observe a positive and significant effect of income on $CO_2$ emissions for low income levels. The turning point is located near 16500 USD per capita GDP, beyond

---

[4]We do not report the results of the reduced-form estimation, since the control variables in equation (8) are only used for the sake of instruments. The results are available from the authors upon request. Moreover, in estimations all the nonparametric functions are normalized to have zero means.

[5]The 'gain' is the difference in normalized deviance between the GAM and the parametric linear models. Its distribution is approximated by a $\chi^2$ ($df = df_g - df_l$), where $df_g$ denotes the degree of freedom of the GAM and $df_l$ is the degree of freedom of the analogue parametric linear model.

this point the relationship turns negative.[6] Nevertheless, this decreasing part is not significant. The proportion of observations located beyond the turning point is about 10%.[7] Compared to Liu (2005) who found a downward slope in a panel of 24 OECD countries over the period 1975-1990, we equally observe such a trend for high income levels implying that the negative effect of income on $CO_2$ emissions is at least neutralized.

<div align="center">**Insert Figure 1 here**</div>

In Figure 1(b), we plot the estimated curve for $CO_2$ emissions and energy use. As outlined by Liu (2005), the latter can be viewed as a proxy to account for differences in industrial structure across nations which may explain their ability to reduce emissions. This relationship is monotonically increasing with a concave pattern, meaning that $CO_2$ emissions increase with energy use but at a decreasing rate. We can interpret this finding as the presence of a learning effect, a technological improvement, and/or changes in energy composition (shifts from fossil energies, which are sources of $CO_2$ emissions, to non-fossil energies) that allow for limiting $CO_2$ emissions when using energy. It seems that more energy intensive economies are more likely to implement cleaner technologies and stringent environmental policies which in turn might neutralize the positive effect of income on $CO_2$ emissions.

## 4    Conclusion

We show that $CO_2$ emissions clearly increase with income at low income levels. For higher income levels, we observe a non significant decreasing slope. This finding reconciles previous results based on different specifications and partial data. Moreover, $CO_2$ emissions rise with energy use but at a decreasing rate. These results show that for a given industrial structure of the economy, higher income countries are likely to achieve the delinking of $CO_2$ emissions from income.

## References

Azomahou, T., Laisney, F. and P. Nguyen Van, P., 2006. Economic development and $CO_2$ emissions: A nonparametric panel approach, Journal of Public Economics 90, 1347-1363.

Bertinelli, L. and E. Strobl, 2005. The environmental Kuznets curve semi-parametrically revisited, Economics Letters 88, 350-357.

Grossman, G.M. and A.B. Krueger, 1995. Economic growth and the environment, The Quarterly Journal of Economics 60, 353-377.

Hastie, T.J. and R.J. Tibshirani, 1990. Generalized Additive Models, Chapman and Hall, London, New York.

Liu, X. 2005. Explaining the relationship between $CO_2$ emissions and national income-The role of energy consumption, Economics Letters 87, 325-328.

---

[6]The exact value of the turning point is computed at 16457.47 USD per capita GDP.

[7]The corresponding countries are Australia, Austria, Belgium, Canada, Denmark, Finland, Iceland, Israel, Kuwait, the Netherlands, Norway, Sweden, Switzerland.

Millimet, D.L., J.A. List and T. Stengos, 2003. The environmental Kuznets curve: Real progress or misspecified models? Review of Economics and Statistics 85, 1038-1047.

Newey, W.K., J.P., Powell and F. Vella, 1999. Nonparametric estimation of triangular simultaneous equations models, Econometrica 67, 565-603.

Nguyen-Van P., 2009. Energy consumption and income: A semiparametric panel data analysis, Energy Economics, forthcoming.

Perrings, C., 1987. Economy and environment: A theoretical Essay on the interdependence of economic and environmental systems. Cambridge: Cambridge University Press.

Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56, 931-954.

Stern, D.I., M.S., Common and E.B. Barbier, 1996. Economic growth and environmental degradation: The environmental Kuznets curve and sustainable Development, World Development 24, 1151-1160.

| | GDP per capita | | | | $CO_2$ emissions per capita | | | | Energy use per capita | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mean | std. | min. | max. | mean | std. | min. | max. | mean | std. | min. | max. |
| 1961 | 2996.034 | 4020.431 | 77.662 | 19959.91 | 2.404 | 4.726 | 0.008 | 36.319 | 0.00243 | 0.0021 | 0.0003 | 0.01 |
| 1982 | 5163.583 | 7156.795 | 124.515 | 39368.63 | 4.701 | 9.238 | 0.027 | 90.425 | 0.00239 | 0.0034 | 0.0001 | 0.023 |
| 2004 | 6678.001 | 9977.794 | 86.45 | 49996.1 | 5.424 | 8.05 | 0.013 | 69.159 | 0.00237 | 0.003 | 0.0002 | 0.02 |
| pooled | 5258.663 | 7722.851 | 56.468 | 59182.83 | 4.927 | 11.02 | 0.0005 | 183.836 | 0.00238 | 0.003 | 0.0001 | 0.036 |

Table 1: Descriptive statistics

Table 2: Gain statistics test

| Form[a] | Gain value | Degree of freedom | $p$-value |
|---|---|---|---|
| $\hat{m}_1(x_{it})$ | 2.518 | 4.486 | 0.007 |
| $\hat{m}_1(x_{i,t-1})$ | 4.311 | 6.972 | 0.000 |
| $\hat{m}_2(z_{0it})$ | 50.863 | 7.979 | 0.000 |
| $\hat{m}_2(z_{0i,t-1})$ | 49.628 | 4.665 | 0.000 |
| $\hat{m}_3(\hat{u}_{it} - \hat{u}_{i,t-1})$ | 2.617 | 4.996 | 0.005 |

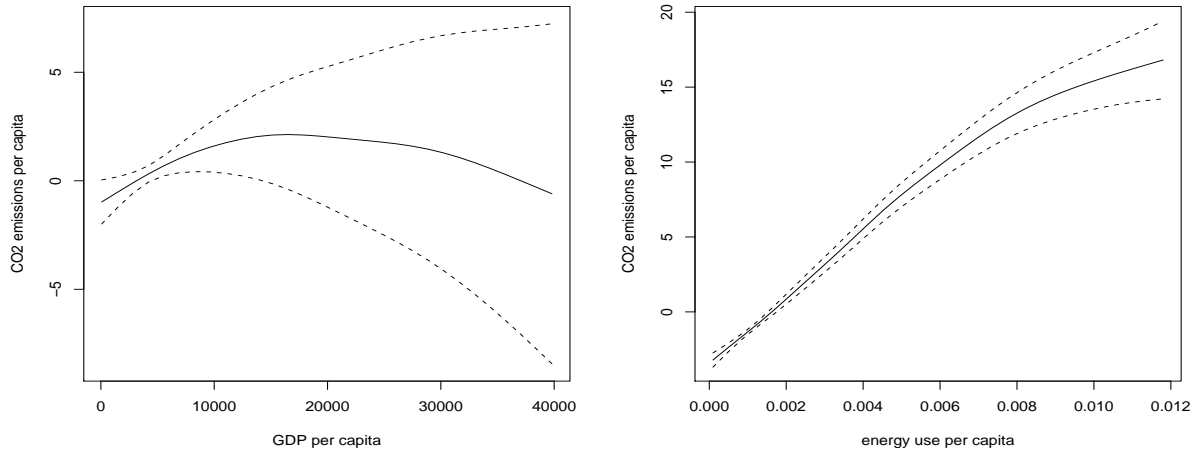[a]: $x$ is GDP, $z_0$ is energy use, and $u$ is the residuals of equation (4).



Figure 1: Structural nonparametric estimations. [left (a)]: $CO_2$ emissions – income relationship. [right (b)]: $CO_2$ emissions – energy use relationship. The solid line represents the nonparametric fit. The dashed lines correspond to the 95% pointwise confidence interval.

# Documents de travail du BETA

_____

2010–**01**     *The Aggregation of Individual Distributive Preferences through the Distributive Liberal Social Contract : Normative Analysis*
Jean MERCIER-YTHIER, janvier 2010.

2010–**02**     *Monnaie et Crise Bancaire dans une Petite Economie Ouverte*
Jin CHENG, janvier 2010.


2010–**03**     *A Structural nonparametric reappraisal of the $CO_2$ emissions-income relationships*
Theophile AZOMAHOU, Micheline GOEDHUYS, Phu NGUYEN-VAN, janvier 2010.




La présente liste ne comprend que les Documents de Travail publiés à partir du 1er janvier 2010. La liste complète peut être donnée sur demande.
*This list contains the Working Paper writen after January 2010, 1rst. The complet list is available upon request.*

_____