



**Bureau
d'économie
théorique
et appliquée
(BETA)**
UMR 7522

Documents de travail

« Energy consumption and income : a semiparametric panel data analysis »

Auteur

Phu Nguyen-Van

Document de Travail n° 2009 - 26

Juin 2009

Faculté des sciences économiques et de gestion

Pôle européen de gestion et
d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secrétariat du BETA

Géraldine Manderscheidt
Tél. : (33) 03 90 24 20 69
Fax : (33) 03 90 24 20 70
g.manderscheidt@unistra.fr
<http://cournot2.u-strasbg.fr/beta>



Nancy-Université
Université Nancy 2



Energy consumption and income: a semiparametric panel data analysis[§]

Phu Nguyen-Van*

BETA-CNRS, Université de Strasbourg

June 12th, 2009

Abstract

This paper proposes a semiparametric analysis for the study of the relationship between energy consumption per capita and income per capita for an international panel dataset. It shows little evidence for the existence of an environmental Kuznets curve for energy consumption. Energy consumption increases with income for a majority of countries and then stabilizes for very high income countries. Neither changes in energy structure nor macroeconomic cycle/technological change have significant effect on energy consumption.

Key words: Energy consumption, environmental Kuznets curve, semiparametric panel model, nonstationarity

JEL classification: C14, C23, Q40

[§] I am also indebted to Monique Flasaquier, Lise Patureau, Francesco Ricci, and Jamel Trabelsi for useful comments. Financial supports from ANR grant n° ANR-05-JCJC-0134-01 (CEDEPTE) are also acknowledged.

* Correspondence: P. Nguyen-Van, BETA, CNRS and Université de Strasbourg, 61 avenue de la Forêt Noire, F-67085 Strasbourg cedex, France; phone: +33 (0)3 90 24 20 39; fax: +33 (0)3 90 24 20 71; e-mail: nguyen-van@unistra.fr

1 Introduction

The Environmental Kuznets Curve (EKC) hypothesis, which suggests an inverted U-shaped relationship between environmental degradation and income, has been extensively investigated in the literature. Various environmental degradation indicators have been examined: emissions or concentrations of pollutants (CO, CO₂, SO₂, NO_x,...), deforestation rate, water quality, etc. Results on the existence of an EKC are mixed and much of them depend largely on the econometric methodology.

Energy constitutes of course an important subject as it is considered as a source of many serious environmental problems. The literature on the relationship between economic growth and energy consumption is dominated by parametric cross-country modeling and time series analysis. For example, by using standard parametric panel specifications (random and fixed effects models), Cole et al. (1997) found that total energy consumption per capita and transport energy use per capita increased with income per capita for OECD countries during the period 1970-1992. Agras and Chapman (1999) did not find evidence of an EKC for energy use for an international panel data during the period 1970-1990. Suri and Chapman (1998) used parametric panel models and showed that the relationship between energy consumption and income displays an increasing pattern (and the turning point is outside the data sample). Richmond and Kaufmann (2006a,b), by using parametric specifications for panel data, found little evidence for an EKC for energy consumption. They showed that energy consumption increases with income at a decreasing rate. In a recent study, Esmaeili and Abdollahzadeh (2009) found an EKC for oil exploitation (not energy use) for a panel data of oil producing countries during 1990-2000.

Existing time series studies on energy include, among others, Stern (2000), Altinay and Karagol (2005), Lee (2005), Lee and Chang (2005, 2007a), Richmond and Kaufmann (2006b), Lee et al. (2008), Huang et al. (2008), Chiou-Wei et al. (2008), Hu and Lin (2008), Yuan et al. (2008), Soytas and Sari (2009), Wolde-Rufael (2009), Bowden and Payne (2009), and papers from a

recent issue of *Energy Economics* (volume 29(6), 2007).¹ They investigated nonstationarity, cointegration and causality between energy and economic series. Causality has been found to be uni- or bi-directional between income and energy consumption, depending on the country considered.

This paper aims to provide a robust estimation of the profile of the relationship between energy consumption and income, which would help us to intervene convincingly in the discussion for the existence of an EKC for energy.² To this end, we use a semiparametric partially linear panel model, which has the advantage to avoid the misspecification problem that may arise in parametric EKC studies as pointed out by Taskin and Zaim (2000), Roy and van Kooten (2004), Bertinelli and Strobl (2005), Millimet et al. (2003), and Azomahou et al. (2006). To our knowledge, the only work on energy that deals with nonparametric techniques is Luzzati and Orsini (2009). The authors used a single regression model and found an increasing relationship between energy consumption and GDP per capita (both in natural logarithms). However, when energy consumption per capita was used as dependent variable instead of energy consumption, there was clearly an inverted U-shaped relationship.

Moreover, our modeling is flexible enough as it enables us to control for other variables such as energy structure and time trend by parametrically including them into the regression. The purpose is to provide results that can

¹Almost all papers of the issue of *Energy Economics*, i.e. Zamani (2007), Narayan and Singh (2007), Soytas and Sari (2007), Lise and Montfort (2007), Yuan et al. (2007), Squalli (2007), Francis et al. (2007), Zachariadis (2007), and Yamaguchi (2007), are time series analyses on selected countries (Iran, the Fiji Islands, Turkey, China, OPEC countries, Caribbean countries, Japan). One paper of this issue, i.e. Lee and Chang (2007b), proposed a panel VAR study.

²In this respect, the paper is more related to cross-country parametric studies than time series ones. Indeed, we are more concerned by the correlation between energy consumption and income than by the causality relationship between them. Furthermore, incorporating nonstationarity in a nonparametric panel data framework is very complex (and the literature on this question is still developing) but may constitute an interesting question to be investigated in the future.

be compared with those of existing works (Richmond and Kaufmann 2006b, Luzzati and Orsini, 2009). We follow Richmond and Kaufmann (2006b) by accounting for changes in the structure of final energy consumption (or changes in energy mix as called by these authors). These authors argued that structural changes (e.g. from coal to oil/natural gas and from oil/natural gas to hydro and nuclear electricity) allow for higher energy efficiency (i.e. lower energy consumption for a given level of economic activity). They also showed that the presence of these structural changes in regressions reduces the size of the turning point. Concerning time trend, which represents the effect of macroeconomic cycle (common international cycle, changes in energy prices, etc.) and technological change, Luzzati and Orsini (2009) found a positive impact on energy consumption, which is rather counter-intuitive.

The next section presents the data and the econometric model. Section 3 discusses estimation results and Section 4 concludes.

2 Data and method

2.1 Data

The data, collected from the Energy Information Administration (EIA), cover a balanced panel of 158 countries and territories for the period 1980–2004 (3950 observations). Variables are total primary energy consumption per capita (measured in millions British thermal units, Btu) and GDP per capita (in thousands real 2000 U.S. dollars). Total primary energy consumption includes consumptions of petroleum, natural gas, coal, hydroelectric power, nuclear power and renewable electric power (geothermal, solar, wind, wood and waste). It also includes net electricity imports (i.e. imports minus exports). GDP distribution shows that most of the observations correspond to low income countries (about 2800 observations corresponding to incomes per capita lower than 10,000 dollars).

Insert Table 1 here

Insert Figure 1 here

We calculate the shares of coal, petroleum and gas, and hydroelectric, nuclear and renewable electric power in total energy consumption. It should be noted that the sum of these three shares, measured in percentage, might not be equal to 100 due to independent rounding.

2.2 Econometric model

We propose the following semiparametric partially linear panel model

$$y_{it} = m(x_{it}) + z'_{it}\gamma + \delta t + u_{it}, \quad i = 1, \dots, N, t = 1, \dots, T, \quad (1)$$

$$= m(x_{it}) + w'_{it}\eta + u_{it}, \quad w_{it} \equiv (z'_{it}, t)', \quad (2)$$

where y_{it} is the energy consumption per capita (in log) of country i at year t , x_{it} is GDP per capita (also in log), m is an unknown function, identifiable up to an additive constant, z_{it} contains other observed time-varying regressors, t is the time trend, u_{it} is the error term that includes unobserved factors. Values of energy consumption and income per capita are expressed in logarithm for providing results which can be compared with existing studies, in particular the work of Luzzati and Orsini (2009). The unknown form of m , which is smooth, continuous, and possibly nonlinear, avoids the use of a pre-specified parametric functional form (polynomial or other parametric forms) on the relationship between energy consumption and income, which is a source of possible misspecification. This model is flexible as it combines a non-parametric regression function (m), a linear part ($z'_{it}\gamma + \delta t$), and the panel structure of our data sample. The nonparametric function m allows for a robust test of existence of the EKC for energy use. We consider m as a univariate function because of the well-known curse of dimensionality when m includes several regressors. The linear part of the model enables us to analyze the effects of other variables.

We assume for instance that u_{it} is i.i.d. in the i index and there is no restriction in the t index. This assumption includes the case of the one-way error component model with $u_{it} = \mu_i + \varepsilon_{it}$ where μ_i is the individual effect and ε_{it} is the standard error term, both of them are uncorrelated with x_{it} and w_{it} ,

i.e. $E(\varepsilon_{it}|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) = E(\mu_i|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) = 0$. In fact, the model discussed here is more general than this well-known random effects model (see e.g. the model of Luzzati and Orsini, 2009) as it allows for ε_{it} being serially correlated and conditionally heteroskedastic (Li and Stengos, 1996). It also includes the usual fixed effect specification, $u_{it} = \mu_i + \varepsilon_{it}$, where $E(u_{it}|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) = E(\mu_i|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) \neq 0$.

Regressors included in z correspond to the share of coal consumption and the share of petroleum and natural gas consumption. The share of hydroelectric, nuclear, and renewable electric power is considered as the reference category. These variables capture structural changes in energy consumption. Time trend variable t accounts for the macroeconomic effect common to all countries. It may also represent the effect of energy prices in the international market. Nevertheless, this variable does not distinguish between these effects (price effect, international economic cycle, etc.) and the effect of technological change on energy consumption.³

Consider the case $E(\mu_i|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) = 0$ (random effects models included). Li and Stengos (1996) proposed an instrumental semiparametric estimator for this model. First, taking the expectation of (1) conditional on x_{it} and then calculating the difference of it with (1), we obtain

$$y_{it} - E(y_{it}|x_{it}) = (w_{it} - E(w_{it}|x_{it}))' \eta + u_{it}. \quad (3)$$

Assuming there exists an instrumental variable q_{it} (such that $E(u_{it}|q_{it}) = 0$) of the same dimension than w_{it} , Li and Stengos (1996) proposed an instrumental variable estimator for η , $\hat{\eta} = (Q'W)^{-1}Q'Y$, where $Q_{it} = q_{it} - E(q_{it}|x_{it})$, $Y_{it} = y_{it} - E(y_{it}|x_{it})$, and $W_{it} = w_{it} - E(w_{it}|x_{it})$. For simplicity's sake, we choose $q_{it} = w_{it}$ as recommended by Li and Stengos (1996). Once $\hat{\eta}$ is available, m may be estimated by

$$\hat{m}(x_{it}) = E((y_{it} - w'_{it}\hat{\eta})|x_{it}) = E(y_{it}|x_{it}) - E(w_{it}|x_{it})'\hat{\eta}. \quad (4)$$

³Another variable that would be interesting to be controlled for is energy prices observed at the country level. However, such a variable is not available for all countries, and using it would considerably reduce the sample size. It would therefore make the nonparametric method unattractive because it requires a large sample.

In estimations we use the local constant kernel estimator (i.e. the Nadayara-Watson estimator) with the Epanechnikov kernel and the rule-of-thumb bandwidth (see Silverman, 1986) to calculate $E(q_{it}|x_{it})$, $E(w_{it}|x_{it})$, and $E(y_{it}|x_{it})$.⁴

We turn now to the case of the fixed effects model where $E(\mu_i|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) \neq 0$. We can consider first differences to eliminate the fixed effects μ_i :

$$y_{it} - y_{i,t-1} = \Psi(x_{it}, x_{i,t-1}) + (z_{it} - z_{i,t-1})'\gamma + \delta + u_{it} - u_{i,t-1}, \quad (5)$$

where $\Psi(x_{it}, x_{i,t-1}) := m(x_{it}) - m(x_{i,t-1})$. As Ψ is a very general function, which may include a constant, we will not consider separately δ and Ψ in estimations (or in other words, δ is not separately identified with Ψ).

This model is the same as (1) and may be estimated by the method of Li and Stengos (1996) detailed above, except that variables in level are replaced by their first differences, the univariate function m being now replaced by a bivariate function Ψ , and instrumental variables $q_{it} = w_{it}$ being replaced by $q_{it} = z_{i,t-1}$. When an estimation of Ψ for this model is obtained, i.e. $\hat{\Psi}(x_{it}, x_{i,t-1}) = E((y_{it} - y_{i,t-1}) - (z_{it} - z_{i,t-1})'\hat{\gamma}|x_{it}, x_{i,t-1})$, we can use the marginal integration method to compute the univariate function m , which is identifiable up to an additive constant. This method, developed by Linton and Nielsen (1995), was applied in the case of CO₂ emissions by Azomahou et al. (2006). The main idea of marginal integration can be described as follows. For simplicity's sake, let us rename the arguments of $\hat{\Psi}$ as u and v . We can write

$$E_v [\hat{\Psi}(u, V)] = \int \hat{\Psi}(u, v) f(v) dv \quad (6)$$

$$= m(u) - E_v [m(V)] \quad (7)$$

$$= m(u) - k, \quad (8)$$

⁴Oversmoothing (corresponding to a higher value of the bandwidth) and undersmoothing (smaller bandwidth) give the same patterns as \hat{m} obtained with the rule-of-thumb bandwidth. We also perform estimations with the local linear kernel estimator which has a smaller bias at the data boundary than the Nadayara-Watson estimator (see, e.g., Pagan and Ullah, 1999). The results still remain very similar.

and similarly,

$$E_u [\hat{\Psi}(U, v)] = \int \hat{\Psi}(u, v) f(u) du \quad (9)$$

$$= k - m(v). \quad (10)$$

We obtain estimators of $m(x_{it})$ and $m(x_{it-1})$ up to the same constant by taking the sample averages

$$\hat{m}^{(1)}(x_{it}) = \frac{1}{N(T-1)} \sum_{j=1}^{N(T-1)} \hat{\Psi}(x_{it}, x_j). \quad (11)$$

By the same way, we can obtain an estimator for $m(x_{it-1})$, i.e.

$$\hat{m}^{(2)}(x_{it-1}) = -\frac{1}{N(T-1)} \sum_{j=1}^{N(T-1)} \hat{\Psi}(x_j, x_{it-1}). \quad (12)$$

A more precise estimator of m can be obtained by a weighted average between $\hat{m}^{(1)}$ and $\hat{m}^{(2)}$, and a simple estimator is given by $\hat{m}(x) = [\hat{m}^{(1)}(x) + \hat{m}^{(2)}(x)] / 2$.

3 Estimation results

As the data period is long enough, we first examine the stationarity of variables. We use the panel data unit root test proposed by Pesaran (2007). The test (denoted as PCADF), based on the augmented Dickey-Fuller statistics, allows for cross-section dependence. We implement the test for all series with only constant on the one hand, and with both constant and trend on the other hand. The statistic is distributed standard normal under the null hypothesis that all the individual series are non-stationary. As reported in Table 2, the PCADF test rejects the nonstationary hypothesis for energy consumption per capita in both cases, with constant only and with both constant and trend. Other variables are nonstationary with constant and trend. When performing the same test for first-differenced data, we find that all differenced variables are stationary.

Insert Table 2 here

According to the Pasaran's test, we can now study the long-run relationship between energy consumption and income in order to validate the environmental Kuznets curve for energy. We use the tests proposed by Westerlund (2007) to investigate the existence of a linear cointegration relation between variables. These tests are based on a parametric error correction model for panel data. The long-run cointegration relation corresponding to this model is the parametric version of model (1):

$$m(x_{it}) = b_0 + b_1x_{it} + b_2x_{it}^2 + b_3x_{it}^3 \quad \text{and} \quad u_{it} = \mu_i + \varepsilon_{it}. \quad (13)$$

Westerlund (2007) proposed four statistics, two group-mean statistics (G_τ and G_α) and two panel statistics (P_τ and P_α), which have limiting normal distributions under the null hypothesis of no cointegration.⁵ Group-mean statistics correspond to the null hypothesis of no cointegration against the alternative of cointegration for some countries. By contrast, panel statistics relate to the alternative hypothesis of cointegration for all countries. The computed values of these four statistics, G_τ , G_α , P_τ , and P_α , are -3.433, -8.005, -32.222, and -6.615, respectively. The associated p -values computed by bootstrap (with 500 replications), which are robust to cross-sectional dependence, are 0.01, 0.05, 0.03, and 0.28, respectively. All statistics, except the last one, reject the null hypothesis of no cointegration at the 5% level. We can therefore conclude that our variables are cointegrated at least for some countries.

Model (1) with *parametric* specification (13) represents the long-run relationship between energy consumption per capita, GDP per capita, and energy shares. It can be estimated by GLS (random effects model), within, and first-difference estimators (fixed effects model). It should be noted that the first two estimators define a long-run average relation such as underlined by Phillips and Moon (2000), and constitute special cases of the long-run average variance estimator proposed by Sun (2004) for nonstationary panel

⁵Westerlund (2007) remarked that his tests shows higher power than the usual cointegration test developed by Pedroni (2004) for panel data. See also Persyn and Westerlund (2008) for an implementation of Westerlund's test in Stata.

time series. The first-difference estimator itself does not require any particular attention because it is implemented on first-differenced variables, which are stationary. Estimation results are reported in Table 3.

As noted previously, the underlying assumption behind the GLS and within estimators is $E(\varepsilon_{it}|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) = 0$, being well-known as the strict exogeneity assumption. Compared to the within estimator, the GLS estimator has the additional assumption $E(\mu_i|x_{i1}, \dots, x_{iT}, w'_{i1}, \dots, w'_{iT}) = 0$ that can be tested by a Hausman test. The computed statistic, equal to $28.56 > 12.59$ (value of $\chi^2(6)$ at the 5% level), allows us to reject the GLS estimator (i.e. rejecting the random effects model) in favor of the within estimator.

Insert Table 3 here

A Hausman test is also computed to compare the within and the first-difference estimators of the fixed effects parametric model. First-difference of the parametric model in (13) is

$$\begin{aligned} y_{it} - y_{i,t-1} &= b_1(x_{it} - x_{i,t-1}) + b_2(x_{it}^2 - x_{i,t-1}^2) + b_3(x_{it}^3 - x_{i,t-1}^3) \\ &\quad + (z_{it} - z_{i,t-1})'\gamma + \delta + (u_{it} - u_{i,t-1}) \end{aligned} \quad (14)$$

We remark that the new constant of this model is δ while b_0 is eliminated from the regression. As for the within estimator, b_0 can be identified by the usual additional assumption that $\sum_{i=1}^N \mu_i = 0$.⁶ Furthermore, we always have the strict exogeneity assumption with the within estimator (the null hypothesis) whereas we have a much weaker assumption with the first-difference estimator, called first-difference assumption, i.e. $E(\varepsilon_{it} - \varepsilon_{it-1}|x_{it}, x_{i,t-1}, w'_{it}, w'_{i,t-1}) = 0$, $i = 1, \dots, N$, $t = 2, \dots, T$.⁷ The

⁶In this case, an estimate of b_0 , \hat{b}_0 , is given by $\bar{y} - \bar{X}'\hat{b} - \bar{z}'\hat{\gamma}$ where $X \equiv (x, x^2, x^3)'$, $\bar{\varphi}$ means the sample average of quantity φ , and $\hat{\zeta}$ an estimate (by within or first-difference) of ζ .

⁷As pointed out by Azomahou et al. (2006), an extension of the predeterminedness assumption $E(\varepsilon_{it}|x_{i1}, \dots, x_{it}, w'_{i1}, \dots, w'_{it}) = 0$ that yields this first-difference assumption is $E(\varepsilon_{it}|x_{i1}, \dots, x_{i,t+1}, w'_{i1}, \dots, w'_{i,t+1}) = 0$, $i = 1, \dots, N$, $t = 1, \dots, T - 1$.

Hausman test statistic, which compares estimators of b_1 , b_2 , b_3 , and γ , is equal to $138.18 > 11.07$ (value of $\chi^2(5)$ at the 5% level). We can conclude that the within estimator is rejected. Therefore the first-difference estimator in (14) is the best estimator in the parametric case.

The next step of the investigation of the long-run average relationship between our variables is to check for its robustness. This naturally leads to consider the *semiparametric* setup in the previous section, that does not impose any known form about the relation between energy consumption and income. Some recent parametric studies highlighted the presence of non-linearity in the relationship between energy consumption and income. In particular, Lee and Chang (2007a) found that a threshold regression model provides a better empirical fit to the data on energy consumption in Taiwan than the standard linear model. Hu and Lin (2008) confirmed the nonlinear long-run equilibrium relationship between disaggregated energy consumption and GDP in Taiwan. Concerning the nonparametric framework, as we underlined previously, to the best of our knowledge, no work except the study of Luzzati and Orsini (2009) has proposed an investigation on the nonparametric relationship between energy consumption and income.

We first use the Hausman-type test proposed by Li and Stengos (1992) to compare the two semiparametric estimators of γ , the one obtained under the null (obtained from the equation (1)) and the other under the alternative (equation (5)). The coefficient related to the time trend (δ) is excluded for the reason that, as underlined previously, δ is not separately identified with the nonparametric component Ψ . The test statistic is a $\chi^2(k)$, with $k = \dim(\gamma)$. The computed value of the statistic is equal to 0.001 much lower than 5.99, the value of $\chi^2(2)$ at the 5% level, implying that the semiparametric model given in (1) is preferred.

Finally, we implement the nonparametric test of Li and Wang (1998). The null hypothesis is the parametric model given in (13) and the associated first-difference estimator and the alternative is the semiparametric model in level given in (1) that was selected by the Li and Stengos test as shown

previously. The Li and Wang test is based on the residuals of the ‘mixed’ regressions under the null and the alternative hypotheses. The statistic is given by

$$I = \frac{1}{n^2 h^\kappa} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \hat{u}_i \hat{u}_j K_{ij} \quad (15)$$

with $n = NT$ and \hat{u} corresponding to the parametric residuals of the ‘mixed’ regressions, i.e. $\hat{u} = y - \hat{m}(x) - w'\hat{\eta}$ where $\hat{m}(x) = \hat{b}_0 + \hat{b}_1 x_{it} + \hat{b}_2 x_{it}^2 + \hat{b}_3 x_{it}^3$ obtained under the null (given by the first-difference estimator) and $\hat{\eta}$ obtained under the alternative. It should also be noted that κ is the dimension of x and in our case $\kappa = 1$ as x is univariate. $K_{ij} = K\left(\frac{x_i - x_j}{h}\right)$ where K is the kernel function (we use the Epanechnikov kernel) and h is the smoothing parameter (obtained by the rule of thumb). Under the null, $nh^{\kappa/2}I \rightarrow N(0, \Omega)$, as $n \rightarrow \infty$, where $\Omega = 2 \left[\int K^2(v) dv \right] E \left[f(x) (E(\sigma^2(x, z)|x))^2 \right]$ with $\sigma^2(x, z)|x = E(u^2|x, z)$, $u = y - m(x) - w'\eta$. Ω is consistently estimated by $\hat{\Omega} = (2/n^2 h^\kappa) \sum_i \sum_{j \neq i} \hat{u}_i^2 \hat{u}_j^2 K_{ij}^2$. It follows that $J := nh^{\kappa/2}I/\sqrt{\hat{\Omega}} \rightarrow N(0, 1)$. The computed value of the Li and Wang test statistic is very high ($= 472.1$), much higher than 1.96, and clearly implies a rejection of the parametric model at the 5% level. We can conclude that the more suitable model for our data is the semiparametric model in (1).

Differences between the parametric model (first-difference estimation) and the semiparametric model given in (1) in terms of estimations of m might be viewed graphically in Figure 2. The parametric curve based on the first-difference estimator is increasing and monotonous. The nonparametric curve also presents an increasing pattern but it is more shaped than the first-difference parametric curve. It fits quite well the data, and clearly better than the first-difference parametric curve, explaining the high value of the Li and Wang statistic. This comparison urges us to be careful when interpreting any parametric relationship between energy and income, as the estimator of this relationship is subject to a downward bias. We should also be cautious when interpreting the result at the boundaries of the sample where few observations are available (i.e. typically when log of GDP per capita is lower than about -0.5 or higher than 3.5).

Insert Figure 2 here

We observe that the nonparametric curve is increasing for almost all values of income, with however a stabilizing shape for very high values of income (log GDP per capita around 3-3.5 or income per capita about 20000-33000 dollars). For a majority of countries and territories (particularly developing ones), our estimation results suggest that their energy consumption would rapidly increase with economic development. For very high income countries (of which the values of log income are between 3 and 3.5, corresponding to income per capita between about 20,000 and 33,100 dollars) the curve is less sloping (or even horizontal), which may be explained by the effectiveness of environmental policies and energy-saving technologies during the period of the study.⁸

The increasing pattern obtained here (both parametric and nonparametric curves) contrasts with the inverted-U shaped curve (or EKC) as obtained by Luzzati and Orsini (2009) mainly because our econometric specification is more general and our estimator accounts for weakly exogenous regressors, which were excluded from the usual within estimation. In particular, our

⁸Indeed, the period of the study (1980-2004) is marked by the introduction of several environmental policies in a number of countries (most of them being part of the developed world). For example, the amount of environmental measures (directives, regulatory acts, or recommendations) adopted by the European Union rose from about 15 at the beginning to 25 at the end of the 80s; it even amounted to more than 40 in 1991 (Lévêque, 1996, Figure 2.1, p. 10). Regarding the energy issue, various tax and fiscal policies promoting energy efficiency (carbon taxes, direct and indirect subsidies for energy users, tax credits, economic incentives such as loans and subsidies for purchasing energy-efficient equipment, regulatory measures, etc.) have been used during the same period (World Energy Council, 2008). There are also integrated policies like energy or greenhouse gas emissions mitigation programs that took place at the international level as voluntary agreements, emission permits market, etc. (the most famous example is the Kyoto protocol). Moreover, according to Mitchell (2007) there are about 33 agreements, 11 amendments and 18 protocols concerning the energy question at the multilateral level for the period 1923-2005. When searching for this period, and for air pollution including climate change (which is mainly due to energy consumption), there were 7 agreements, 21 amendments and 13 protocols at the multilateral level.

nonparametric results differ from the nonparametric EKC obtained by Luzzati and Orsini (2009, Figures 2 and A2). In our view, these authors specified in fact a semiparametric model with the usual random effects. However, our results appear consistent with the increasing curve obtained by the same authors when the dependent variable is energy consumption (in log) instead of energy consumption per capita (also in log) and the time trend is excluded. This result is also consistent with the findings of Cole et al. (1997), Suri and Chapman (1998), Agras and Chapman (1999), and Richmond and Kaufmann (2006a,b).

Insert Table 4 here

The estimates of the effects of the share of coal consumption and the share of petroleum and natural gas consumption in the semiparametric model (1) are respectively 0.011 (standard error = 0.212) and 0.002 (.207), which are not significant compared to the share of hydroelectric, nuclear and renewable electric power (see Table 4). Changes in energy structure (or energy mix) have then no effect on energy consumption, contrary to the results of Richmond and Kaufmann (2006a). We may conclude that policies promoting green energies, that can abate greenhouse gas emissions produced from notably the use of fossil energies, do not likely contribute to the reduction of total energy use. Finally, the effect of time trend is not significant.⁹ It seems that international macroeconomic cycles and technological change do not have an impact on final energy consumption for the period of the study.

4 Concluding remarks

The EKC hypothesis is not confirmed by our analysis. Energy consumption rises with income and then stabilizes for very high income levels. The shape of the increasing part of the nonparametric curve is even higher than that of

⁹Luzzati and Orsini (2009) found a counter-intuitive result that time trend has a positive effect on energy consumption.

the parametric counterpart, suggesting that energy consumption in developing countries would rise more rapidly than expected by parametric studies. It would result in the near future in serious economic and environmental problems in these countries such as rapid increase of greenhouse gas emissions due to energy use, excessive pressure on the provision of energy resources, etc. Furthermore, the stabilizing part of the curve corresponding to very high income values suggests that stringent environmental policies may play an important role in reducing or stabilizing energy use. Therefore, a rough policy recommendation that we can express in this study is the application of these policies, in particular policies for promoting energy efficiency, in all countries in order to reduce future energy use.

The structure of models used in this paper, which relies on weaker assumptions (unknown functional form, weakly exogenous regressors) than those of standard parametric panel data models (polynomial functional forms, strict exogeneity), may be applied in the study of other environmental indicators. In particular, the assumption of weakly exogenous regressors allows us to obtain results robust to omitted variables that usually induce a correlation between the regression error and regressors, which is not taken into account by the assumption of strictly exogenous regressors often adopted by existing studies. Our model is not however able to explicitly estimate the possible feedback effects, e.g. that of energy use on income. Another drawback is that we cannot perform a forecasting analysis as in existing parametric studies. These issues would be interestingly addressed in our further works. It would be also promising to investigate the determinants of the energy EKC by considering other variables as in Gassebner et al. (2006) for estimating the ceteris paribus impact of GDP on energy consumption and then making more precise policy prescriptions.

Appendix: List of countries and territories

Antigua and Barbuda, Afghanistan, Algeria, American Samoa, Argentina, Australia, Austria, Bahrain, Barbados, Botswana, Bermuda, Belgium, The

Bahamas, Bangladesh, Belize, Bolivia, Burkina Faso, Burma, Benin, Solomon Islands, Brazil, Bhutan, Brunei, Burundi, Canada, Cambodia, Chad, Congo (Brazzaville), Congo (Kinshasa), China, Chile, Cayman Islands, Cameroon, Comoros, Colombia, Costa Rica, Central African Republic, Cuba, Cape Verde, Cyprus, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, Equatorial Guinea, El Salvador, Ethiopia, Fiji, Finland, France, French Guiana, Gabon, The Gambia, Ghana, Greece, Grenada, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hong Kong, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Japan, Jordan, Kenya, Kiribati, North Korea, South Korea, Kuwait, Laos, Lebanon, Lesotho, Liberia, Libya, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Martinique, Mauritania, Mauritius, Mexico, Mongolia, Morocco, Nepal, Netherlands, Netherlands Antilles, New Zealand, Nicaragua, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Portugal, Puerto Rico, Qatar, Reunion, Rwanda, Saint Kitts and Nevis, Saint Lucia, Saint Vincent/Grenadines, Samoa, Sao Tome and Principe, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Singapore, Solomon Islands, Somalia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Swaziland, Sweden, Switzerland, Syria, Taiwan, Tanzania, Thailand, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Uganda, United Arab Emirates, United Kingdom, United States, Uruguay, Vanuatu, Venezuela, Vietnam, US. Virgin Islands, Yemen, Zambia, Zimbabwe.

References

- [1] Agras J., Chapman D. (1999), “A dynamic approach to the Environmental Kuznets Curve hypothesis”, *Energy Economics* 28, 267–277.
- [2] Altinay G., Karagol E. (2005), “Electricity consumption and economic growth: Evidence from Turkey”, *Energy Economics* 27, 849–856.
- [3] Azomahou T., Laisney F., Nguyen Van P. (2006), Economic development and CO₂ emissions: a nonparametric panel approach”, *Journal of Public*

Economics 90, 1347–1363.

- [4] Bertinelli L., Strobl E. (2005), “The environmental Kuznets curve semi-parametrically revisited”, *Economics Letters* 88, 350–357.
- [5] Bowden N., Payne J.E. (2009), “The causal relationship between U.S. energy consumption and real output: A disaggregated analysis”, *Journal of Policy Modeling* 31, 180–188.
- [6] Chiou-Wei S.Z., Chen C.-F., Zhu Z. (2008), “Economic growth and energy consumption revisited - Evidence from linear and nonlinear Granger causality”, *Energy Economics* 30, 3063-3076.
- [7] Cole M.A., Rayner A.J., Bates J.M. (1997), “The environmental Kuznets curve: An empirical analysis”, *Environment and Development Economics* 2, 401–416.
- [8] Esmaeili A., Abdollahzadeh N. (2009), “Oil exploitation and the environmental Kuznets curve”, *Energy Policy* 37, 371–374.
- [9] Francis B.M., Moseley L., Iyare S.O. (2007), “Energy consumption and projected growth in selected Caribbean countries”, *Energy Economics* 29, 1224–1232.
- [10] Gassebner M., Lamla M., Sturm J.-E. (2006), “Economic, demographic and political determinants of pollution reassessed: A sensitivity analysis”, working papers 06-129, ETH Zurich.
- [11] Hu J.-L., Lin C.-H. (2008), “Disaggregated energy consumption and GDP in Taiwan: A threshold cointegration analysis”, *Energy Economics* 30, 2342–2358.
- [12] Huang B.-N., Hwang M.J., Yang C.W. (2008), “Causal relationship between energy consumption and GDP growth revisited: A dynamic panel data approach”, *Ecological Economics* 67, 41-54.
- [13] Lee C.-C. (2005), “Energy consumption and GDP in developing countries: A cointegrated panel analysis”, *Energy Economics* 27, 415–427.

- [14] Lee C.-C., Chang C.-P. (2005), “Structural breaks, energy consumption, and economic growth revisited: Evidence from Taiwan”, *Energy Economics* 27, 857–872.
- [15] Lee C.-C., Chang C.-P. (2007a), “The impact of energy consumption on economic growth: Evidence from linear and nonlinear models in Taiwan”, *Energy* 32, 2282–2294.
- [16] Lee C.-C., Chang C.-P. (2007b), “Energy consumption and GDP revisited: A panel analysis of developed and developing countries”, *Energy Economics* 29, 1206–1223.
- [17] Lee C.-C., Chang C.-P., Chen P.-F. (2008), “Energy-income causality in OECD countries revisited: The key role of capital stock”, *Energy Economics* 30, 2359–2373.
- [18] Lévêque F. (1996), *Environmental Policy in Europe: Industry, Competition and the Policy Process*, Edward Elgar, Cheltenham, UK.
- [19] Li Q., Stengos T. (1992), “A Hausman specification test based on root- N -consistent semiparametric estimators”, *Economics Letters* 40, 141–146.
- [20] Li Q., Stengos T. (1996), “Semiparametric estimation of partially linear panel data models”, *Journal of Econometrics* 71, 389–397.
- [21] Li Q., Wang S. (1998), “A simple consistent bootstrap test for a parametric regression function”, *Journal of Econometrics* 87, 145–165.
- [22] Linton O., Nielsen J.P. (1995), “A kernel method of estimating structured nonparametric regression based on marginal integration”, *Biometrika* 82, 93–100.
- [23] Lise W., Van Montfort K. (2007), “Energy consumption and GDP in Turkey: Is there a co-integration relationship?”, *Energy Economics* 29, 1166–1178.
- [24] Luzzati T., Orsini M. (2009), “Investigating the energy-environmental Kuznets curve”, *Energy* 34, 291–300.

- [25] Millimet D.L., List J.A., Stengos T. (2003), “The environmental Kuznets curve: real progress or misspecified models?”, *Review of Economics and Statistics* 85, 1038–1047.
- [26] Mitchell R.B. (2007), *International Environmental Agreements Database Project (Version 2007.1)*, available at <http://iea.uoregon.edu>.
- [27] Narayan P.K., Singh B. (2007), “The electricity consumption and GDP nexus for the Fiji Islands”, *Energy Economics* 29, 1141–1150.
- [28] Pagan A., Ullah A. (1999), *Nonparametric Econometrics*, Cambridge University Press, New York.
- [29] Pedroni P. (2004), “Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis”, *Econometric Theory* 3, 579-625.
- [30] Pesaran M.H. (2007), “A simple panel unit root test in the presence of cross-section dependence”, *Journal of Applied Econometrics* 22, 265–312.
- [31] Persyn D., Westerlund J. (2008), “Error correction based cointegration tests for panel data”, *Stata Journal*, forthcoming.
- [32] Phillips P.C.B., Moon H.R. (2000), “Nonstationary panel data analysis: An overview of some recent developments”, *Econometric Reviews* 19, 263–286.
- [33] Richmond A.K., Kaufmann R.K. (2006a), “Energy prices and the turning points: the relationship between income and energy use/carbon emissions”, *Energy Journal* 27, 157–180.
- [34] Richmond A.K., Kaufmann R.K. (2006b), “Is there a turning point in the relationship between income and energy use and/or carbon emissions?”, *Ecological Economics* 56, 176–189.
- [35] Roy N., van Kooten G.C. (2004), “Another look at the income elasticity of non-point source air pollutants: A semiparametric approach”, *Economics Letters* 85, 17–22.
- [36] Silverman B.W. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York.

- [37] Soytas U., Sari R. (2007), “The relationship between energy and production: Evidence from Turkish manufacturing industry”, *Energy Economics* 29, 1151–1165.
- [38] Soytas U., Sari R. (2009), “Energy consumption, economic growth, and carbon emissions: Challenges faced by an EU candidate member”, *Ecological Economics* 68, 1667-1675.
- [39] Squalli J. (2007), “Electricity consumption and economic growth: Bounds and causality analyses of OPEC members”, *Energy Economics* 29, 1192–1205.
- [40] Stern D. (2000), “Multivariate cointegration analysis of the role of energy in the US macroeconomy”, *Energy Economics* 22, 267–283.
- [41] Sun Y. (2004), “Estimation of the long-run average relationship in non-stationary panel time series”, *Econometric Theory* 20, 1227–1260.
- [42] Suri V., Chapman D. (1998), “Economic growth, trade and energy: implications for the environmental Kuznets curve”, *Ecological Economics* 25, 195–208.
- [43] Taskin F., Zaim O. (2000), “Searching for a Kuznets curve in environmental efficiency using kernel estimation”, *Economics Letters* 68, 217–223.
- [44] Westerlund J. (2007), “Testing for error correction in panel data”, *Oxford Bulletin of Economics and Statistics* 69, 709–748.
- [45] Wolde-Rufael Y. (2009), “Energy consumption and economic growth: The experience of African countries revisited”, *Energy Economics* 31, 217–224.
- [46] World Energy Council (2008), *Energy Efficiency Policies around the World: Review and Evaluation*, World Energy Council Report, London.
- [47] Yamaguchi K. (2007), “Estimating energy elasticity with structural changes in Japan”, *Energy Economics* 29, 1254–1259.
- [48] Yuan J.-H., Kang J.-G., Zhao C.-H., Hu Z.-G. (2008), “Energy consumption and economic growth: Evidence from China at both aggregated and disaggregated levels”, *Energy Economics* 30, 3077–3094.

- [49] Yuan J.-H., Zhao C.-H., Yu S.-K., Hu Z.-G. (2007), “Electricity consumption and economic growth in China: Cointegration and co-feature analysis”, *Energy Economics* 29, 179–1191.
- [50] Zachariadis T. (2007), “Exploring the relationship between energy use and economic growth with bivariate models: New evidence from G-7 countries”, *Energy Economics* 29, 1233–1253.
- [51] Zamani M. (2007), “Energy consumption and economic activities in Iran”, *Energy Economics* 29, 1135–1140.

Table 1: Descriptive statistics

Variable	Units	Mean	Std. Dev.	Min.	Max.
Energy consumption per capita	millions British thermal units (Btu)	88.904	174.09	0.12	2507.34
GDP per capita	thousands real 2000 U.S. dollars	7.89	8.04	0.07	44.07
Coal share	percent	7.67	15.86	0	84.65
Petroleum and natural gas share	percent	78.11	23.32	4.28	100.54
Hydroelectric, nuclear & renewable power	percent	14.19	18	-3.05	91.51

Notes: Balanced panel data on 158 countries and territories observed for the period 1980–2004 (3950 observations).

Data source: Energy Information Administration (EIA).

Table 2: PCADF unit root test

	constant	constant & trend
Energy cons. per capita	-9.020*	-4.265*
GDP per capita	2.906	1.752
Squared GDP per capita	0.083	4.987
Cubic GDP per capita	1.675	4.237
Coal share	18.972	15.205
Petroleum and gas share	7.661	2.933

Notes: Energy consumption per capita and GDP per capita are expressed in logarithm. The Z_{tbar} statistic is provided. Z_{tbar} is distributed standard normal under the null hypothesis that all the individual series are non-stationary. Significant coefficients at the 5% level are starred.

Table 3: Parametric regressions

	GLS ^a		Within ^b		First-difference ^c	
	Coef.	Std.Err	Coef.	Std.Err	Coef.	Std.Err
GDP, linear term	0.856*	0.026	0.801*	0.026	0.385*	0.040
GDP, quadratic term	0.009	0.011	-0.009	0.011	0.022	0.011
GDP, cubic term	-0.015*	0.003	-0.011*	0.003	-0.012*	0.004
Coal share	0.001	0.001	0.001	0.001	0.005*	0.001
Petroleum and gas share	-0.005*	0.001	-0.005*	0.001	0.004*	0.001
Time trend	0.009*	0.001	0.010*	0.001	0.013*	0.001
Intercept	2.255*	0.103	2.383*	0.079	2.162*	0.082

Notes: ^aGLS estimation of the random effects model. ^bwithin estimation of the fixed effects model. ^cfirst-difference estimation of the fixed effects model. Energy consumption and GDP per capita are measured in log. Significant coefficients at the 5% level are starred.

Table 4: Nonparametric regressions

	Level ^a		First-difference ^b	
	Coef.	Std.Err	Coef.	Std.Err
Coal share	0.011	0.212	0.032	1.076
Petroleum and gas share	0.002	0.207	0.037	1.520
Time trend	0.003	0.160	–	–

Notes: ^a Li and Wang' (1996) estimator for equation in level, i.e. equation (1).

^bLi and Wang' (1996) estimator for equation in first-difference, i.e. equation (5).

In the first-differenced model, the coefficient of the time trend δ is not separately identified from the nonparametric component Ψ . Significant coefficients at the 5% level are starred.

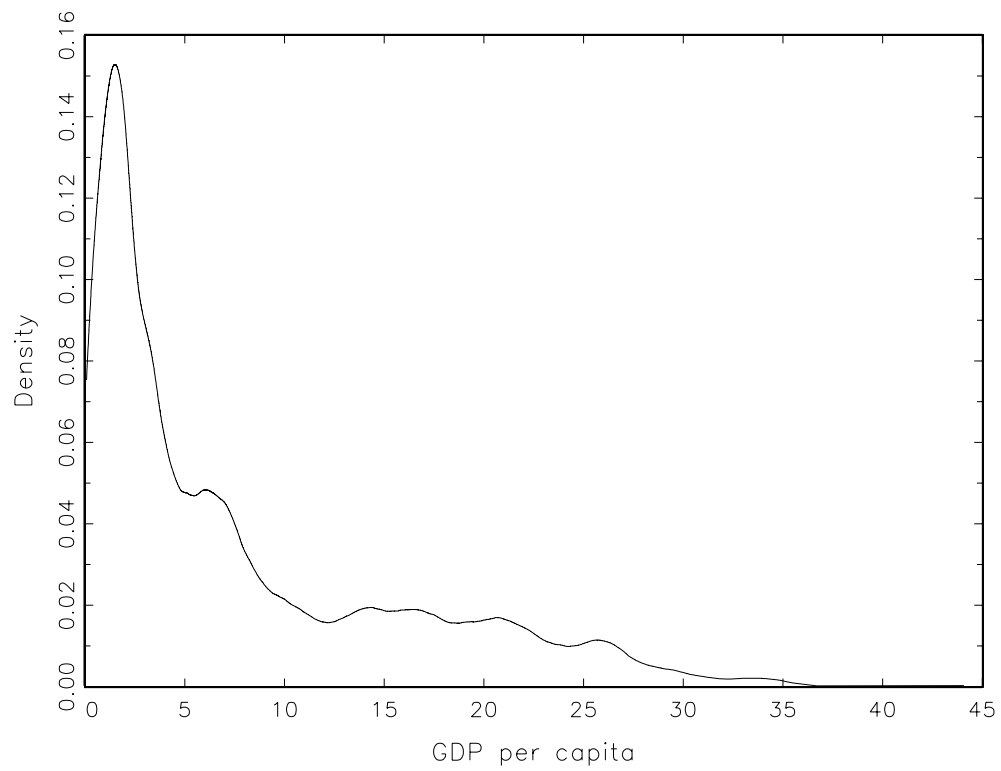


Figure 1: Kernel density estimation for GDP per capita (in thousands real 2000 US dollars).

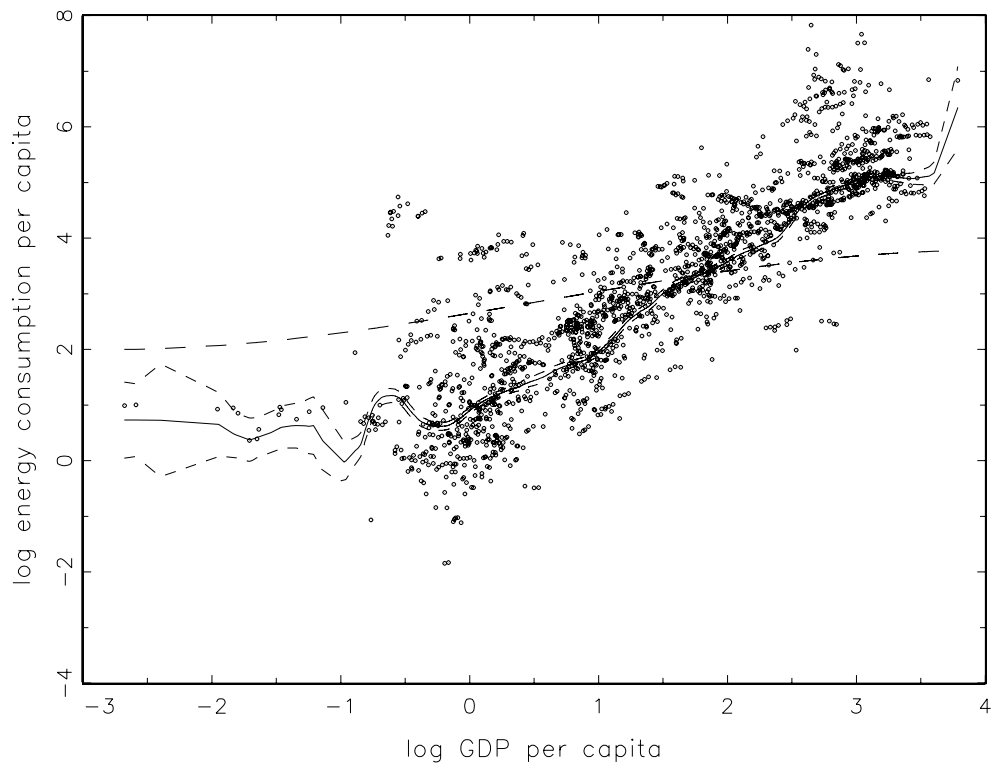


Figure 2: Relation between log energy consumption per capita (in millions Btu) and log GDP per capita (in thousands real 2000 US dollars). The solid curve is the nonparametric estimation of $m(x)$. The short dashes curves correspond to its 95% confidence interval. The long dashes curve corresponds to the first-difference estimation of the parametric model with $m(x_{it}) = b_0 + b_1x_{it} + b_2x_{it}^2 + b_3x_{it}^3$.