Environmental Kuznets curves: Bayesian evidence from switching regime models

George E. Halkos*, Efthymios G. Tsionas

*Department of Economics, University of Thessaly, Volos, Greece
bCouncil of Economic Advisors, Economic Research and Policy Analysis Unit, Ministry of National Economy, Athens, Greece

Abstract

The purpose of the paper is to test empirically the existence of an environmental Kuznets curve (EKC), using switching regime models along with cross-sectional data and Bayesian Markov chain Monte Carlo methods to perform the computations. The models are based on the normal and Student’s t distributions. These methods allow us to present exact, finite-sample posterior distributions of switching regime model parameters as well as exact probabilities of separation of countries into regimes of high and low environmental degradation. Our evidence indicates a monotonic relationship between environmental degradation and income and thus rejects the existence of an EKC. Additionally, we find that several economic and demographic variables cannot explain the distinction between low and high damage countries. © 2001 Elsevier Science B.V. All rights reserved.

JEL classifications: C11; C15; Q2

Keywords: Damage evaluation; Bayesian analysis; Markov chain Monte Carlo

1. Introduction

Kuznets (1965, 1966) showed that during the various economic development stages, income disparities first rise and then begin to fall. In this paper we examine the concept of an environmental Kuznets curve (hereafter EKC) in a critical way with an eye towards proposing policies compatible with sustainable development. The EKC indicates an inverted U-shaped relation between environmental degradation (pollution or deforestation) and per-capita income. Degradation seems to be
lower in the most developed countries compared to many middle-income countries. Similarly, it tends to be higher in many middle-income countries in comparison to less developed countries. A number of studies have proposed that nations will go through a period of high environmental degradation followed by a period of lower degradation as they develop.

Stern et al. (1996), Arrow et al. (1995), Ekins (1997) and Ansuategi et al. (1998) provide a number of reviews and critiques of the EKC studies. Grossman and Krueger (1992, 1995) and Shafik and Bandyopadhyay (1992) suggest that at high income levels, material use increases in a way that the EKC is N-shaped. IBRD (1992) and Shafik and Bandyopadhyay (1992) claim that CO₂ emissions increase with income. However, Stern et al. (1996) demonstrate the sensitivity of the results to the datasets used. That is, using the World Bank data cannot represent an EKC curve while the UN data can. Pezzey (1989) presents arguments for an N-shaped EKC and proposes that the optimal path of environmental degradation may be monotonically increasing with the level of development. However, in our dataset this is not the case.

The greenhouse effect is the most serious problem to sustainable development but no-one suggests that an inverted U-shaped curve applies for greenhouse gases. The levels of several pollutants per unit of output in specific processes have declined in the developed countries over time with the use of strict environmental regulations. Stern et al. (1996) claim that the mix of effluent has shifted from sulfur and NOₓ to CO₂ and solid waste, in a way that aggregate waste is still high and even if per unit output waste has declined, per capita waste may not have declined. A number of authors have estimated econometrically the EKC using OLS analysis. Stern et al. (1996) identified seven major problems with some of the main EKC estimators and their interpretation, namely: (a) econometric problems; (b) the assumption that changes in trade relationship associated with development have no effect on environmental quality; (c) the assumption of unidirectional causality from growth to environmental quality and the reversibility of environmental change; (d) the mean–median income problems; (e) the interpretation of particular EKCs in isolation from EKCs for other environmental problems; (f) asymptotic behavior; and (g) ambient concentrations against emissions. Stern (1998) reviews these problems in detail, integrating the main concepts and insights from other critiques and showing where progress has been made in empirical studies.

The use of OLS is not an appropriate technique in modeling the EKC. First of all, none of the empirical studies presents diagnostic statistics of the regression residuals. Second, an alternative form of the EKC hypothesis suggests that environmental degradation as a function of income is not a stable relationship but may depend on the level of income. This is because in this alternative form, there may exist one relationship for poor and another for rich countries. On the aggregate this would give an inverted U-like curve, however, approximating it by a quadratic will not yield satisfactory results because of the non-linearity involved. See also Zang (1998) for an analysis of the intertemporal stability of the EKC.

We account for this fact by using switching regime models, which should be more
appropriate in the presence of such types of non-linearity. In addition we use Bayesian methods to perform inference because we suspect that relying on asymptotic distributions is not enough and exact, finite-sample results are needed in order to test for the existence of an EKC with reasonable precision. Obviously this presupposes that a possible turning point is constrained within the sample. Otherwise, switching regime, spline regression and quadratic regression models will all fail to detect it. In Section 2 we fit a quadratic in our data, following standard practice. The regression equation turns out to be heavily misspecified and, therefore, totally unreliable as a model for testing the EKC hypothesis. This is exactly the kind of setup we would expect if a switching regime model were the true data-generating process but instead we fit a quadratic regression.

The EKC estimates for any dependent variable (e.g. SO$_2$, NO$_x$, deforestation, etc.) peak at income levels which are around the world mean income per capita. As expected, income is not normally distributed but skewed (with a lot of countries below mean income per capita). Cropper and Griffiths (1994) and Selten and Song (1994) conclude that the majority of countries in their analyses are below their estimated peak levels for air pollutants and thus economic growth may not reduce air pollution or deforestation. Because of this problem estimating the left part of the EKC is easier than estimating the right-hand part. Thus, use of OLS is not likely to yield accurate estimates of the peak levels. We propose to address the problem in a formal way, using switching regime models. In this paper:

1. We use switching regime models to test the existence of an EKC based on a cross-section of 61 countries. Our first model is a structural-break formulation. The second model is a separating hyperplane formulation that performs discriminant analysis based on a set of explanatory variables. Both normal and Student’s $t$ distributions are used to model the disturbances. The Student’s $t$ distribution is used to allow for the possible existence of a few outliers. The presence of undetected outliers would seriously affect the ability of the model to detect a structural break in the sample.
2. We use Markov chain Monte Carlo techniques to perform the computations associated with the Bayesian analysis of the model.
3. The method is capable of providing a clustering of countries according to their degree of environmental degradation based on multivariate discrimination using our separating hyperplane formulation.
4. We provide exact posterior distributions for the structural break (i.e. the peak), thus avoiding reliance on asymptotic theory which would be a bad approximation in our small sample.
5. We explore in a systematic manner the role of industrialization, urbanization, mortality, and growth rates for the existence of an EKC and their discriminatory power, i.e. their ability to separate the sample in high and low environmental degradation countries.

The remainder of the paper is organized as follows: Section 2 discusses the existing work. Section 3 presents the econometric models used in the study. The
cross-sectional empirical evidence and technical details are presented in Section 4. The data are presented in Section 5, and the empirical results are reported in Section 6. The final section concludes the paper.

2. Previous work

The existing empirical evidence suggests that EKCs exist for pollutants with semi-local and medium-term impacts (Arrow et al., 1995; Cole et al., 1997; Ansuategi et al., 1998). The empirical analysis of the EKC has focused on whether a given index of environmental degradation shows an inverted-U relationship when it is related with income per capita. As a result the ‘turning point’ can be calculated by the level of per capita income at which the EKC peaks.

Shafik and Bandyopadhyay (1992) estimated EKC for 10 different indicators of environmental degradation (lack of clean water, ambient sulfur oxides, annual rate of deforestation, etc.). The study uses three different functional forms (log-linear, log-quadratic in income, logarithmic cubic polynomial in GDP/c and a time trend). GDP was measured in PPP and other variables included were population density, trade, electricity prices, dummies for locations, etc. Deforestation was found to be insignificant in relation to income ($R^2$ adjusted $\approx 0$).

Panayotou (1993, 1995) employed cross-sectional data and GDP in nominal US dollars. The equations for the pollutants considered were logarithmic quadratics in income per capita. Deforestation was estimated against a translog function in income/c and population density. All the estimated curves were inverted Us with turning point for deforestation at $823$ per capita. Panayotou used current exchange rates instead of using a PPP approach which lowers the income levels of developing countries in comparison to some developed ones.

Cropper and Griffiths (1994) estimated three regional EKCs for deforestation only. The regressions were for Africa, Latin America and Asia. They used pooled time series cross-section data on a regional basis. The results for Latin America and Africa show an adjusted $R^2$ of 0.47 and 0.63, respectively. Both the population growth and time trend were insignificant in all cases. None of the coefficients in the Asian regression were significant and the adjusted $R^2$ was only 0.13. One of their main conclusions was that economic growth does not solve the problem of deforestation.

Stern et al. (1996) used data from the Human Development Report for 1992 (UNDP, 1992), the greenhouse index for 1988–1989 and the income per capita in PPP adjusted US dollars for 1989, and fitted a quadratic to this data with the addition of the national average annual temperature as regressor. They found that greenhouse index increases linearly with income with an adjusted $R^2$ equal to 0.3255. Regressing per capita energy consumption on income and temperature gave them an inverted U-shaped relationship between energy and income. Fitting a quadratic in income gave them a significant negative coefficient for the squared income term with an adjusted $R^2$ equal to 0.8081. Energy consumption peaked at $14,600$. The authors stated that the results depend on the income measure used. If
PPP-adjusted income is used, the coefficient on squared income was positive but small and insignificant. If income per capita was measured using official exchange rates, the fitted energy income relationship was an inverted U-shape with squared income coefficient negative and significant (with an adjusted $R^2 = 0.6564$). Energy use per capita peaked at income $23,900$.

Dijkgraaf and Vollebergh (1998) estimated EKCs for CO$_2$ emissions relying on a panel data of OECD countries and time series regressions for each of the countries in the panel. They estimated fixed time and country effects for OECD countries and found a turning point at 54% of maximal GDP in the sample. They claim that although, analyzing the whole dataset, there is not a meaningful EKC for carbon emissions, for some individual countries this relationship may be significant. Their main result was that the coefficients in the individual regressions seem to differ widely. They find that for the panel estimate the residuals are serially correlated while for the case of individual countries this is not the case.

Most EKC studies have used panel data and either fixed or random effects estimators. Only a minority use cross-sectional or time series data. Stern et al. (1998) and Dijkgraaf and Vollebergh (1998) found a turning point for carbon within the sample mainly due to the use of data on OECD countries only, while other studies are more global in scope.

Holtz-Eakin and Selten (1995) confirmed Shafik’s (Shafik, 1994) results by estimating quadratic EKCs for CO$_2$ emissions using panel data and finding high turning points of $35,000$ in terms of levels regression and $8$ million in a logarithmic regression. Similarly a study by Schmalensee et al. (1995) found an in-sample turning point for carbon using a more extensive version of the Holtz-Eakin and Selten (1995) dataset. They used a piecewise or spline regression to estimate a carbon EKC. Various econometric and data related issues were treated in the context of EKC estimation by Zang (1998), Matyas et al. (1998) and Wang et al. (1998).

Finally, Kahn (1998) uses 1993 California vehicle emissions data to show that a non-monotonic emissions–income relation exists at the household level. He concludes that richer households may create more vehicle emissions as they own more cars and drive more. Poorer households maintain their cars less and they may pollute more.

These studies do not provide diagnostics so we cannot be certain that the peak levels provided — and the policy implications suggested — are accurate. Based on our data set the following results were obtained:

\[
\text{Deforestation} = -2.566 + 1.363\log\text{GNP} - 0.232\left[\frac{1}{2}(\log\text{GNP})^2\right]
\]

Harvey test for heteroskedasticity, $\chi^2(2) = 9.306$; RESET test for misspecification, $F_{3,55} = 2.99$; BP test for heteroskedasticity, $\chi^2(2) = 1.399$; Jarque–Bera test for normality, $\chi^2(1) = 4.41$.

Standard errors are presented in parentheses. These results indicate the existence of an EKC. However, the diagnostics imply the specification is totally unreliable as we seem to have heteroskedasticity, misspecification and non-normal-
ity problems.\(^1\) If indeed the EKC hypothesis is not accurately described by a quadratic but by a switching regime model, it is reasonable to expect that simple quadratics are heavily misspecified as it turns out to be the case in our sample.

Our econometric models, proposed next, are non-linear. If they are more faithful to the data compared to the linear models (that previous work has employed) then OLS applied to linear models would yield biased results. We also condition on several economic, social and demographic factors and we try to verify the existence of an EKC conditionally on these factors.

3. Econometric models

Our modeling strategy is based on the concept of switching regimes. Switching regimes are prominent in modeling time series with a change in regime (e.g. Hamilton, 1989, 1994, ch. 22). Bayesian work in the field includes Albert and Chib (1993), Geweke and Terui (1993) and Muller et al. (1997).

Two formulations will be used that are versions of a switching regime model with different assumptions about the switching process. In the first approach, we have

\[
y_i = x_i'\beta_1 + u_{i1}, \quad \text{if } z_i \leq z^*, \\
y_i = x_i'\beta_2 + u_{i2}, \quad \text{if } z_i > z^*, \quad i = 1, \ldots, n
\] 

where \(x_i\) is a \(k \times 1\) vector containing data on exogenous variables (see Section 5); \(z^*\) is a break point and \(z_i\) is a certain variable that defines the structural change. We call this model an exogenous break switching regime model (EBSR). If \(z_i\) is income, then we assume in advance that income induces a break (i.e. we impose Kuznets’ hypothesis) and that \(z\) is the critical level of income. A posteriori, it is possible to reject the validity of this model if we find that \(z^*\) is too close to \(\min(z_i, i = 1, \ldots, n)\).

Second, assume that there is a separating variable \(S\) such that when \(S\) exceeds a certain limit (which can be normalized to zero) the model undergoes a structural change. In other words,

\[
y_i = x_i'\beta_1 + u_{i1}, \quad \text{if } S_i > 0 \\
y_i = x_i'\beta_2 + u_{i2}, \quad \text{if } S_i \leq 0
\] 

where

\[
S_i = z_i'\gamma + v_i, \quad i = 1, \ldots, n
\] 

\(u_{i1}, u_{i2}\) and \(v_i\) are independent \(N(0, \sigma_1^2)\), \(N(0, \sigma_2^2)\) and \(N(0, \sigma^2)\), respectively, \(x_i\) and \(z_i\) are \(k \times 1\) and \(m \times 1\) vectors of explanatory variables (having some, none or all variables in common), \(y_i\) is the dependent variable and \(\beta_1\), \(\beta_2\) and \(\gamma\) are parameter vectors conformable with \(x_i\) and \(z_i\). Without loss of generality we may

\(^1\)Results using CO\(_2\) emissions as dependent variable were equally poor.
set \( \sigma = 1 \) in order to identify the \( \gamma \)'s. We call this model a normal separating hyperplane switching regime model (N-SHSR). Later on we will consider a Student’s \( t \) distribution as an alternative.

The EKC hypothesis holds for this model provided income can be included in the separating function. If this is not the case, the transition from one model to the other does not depend on the level of income, violating Kuznets’ law. One can think of the SHSR model as a generalization of the EBSR model, in the sense that the switching point is made a stochastic function of certain explanatory variables.

The likelihood function for the SHSR, is given by

\[
L(\beta_1, \beta_2, \gamma, \sigma_1^2, \sigma_2^2; y, X) = \prod_{i=1}^{n} \left\{ (2\pi \sigma_1^2)^{-1/2} \exp\left[-(y_i' - x_i' \beta_1)^2 / 2\sigma_1^2\right] \right\} 
\times [1 - \Phi(-z_i' \gamma)] 
+ (2\pi \sigma_2^2)^{-1/2} \exp\left[-(y_i' - x_i' \beta_2)^2 / 2\sigma_2^2\right] \Phi(-z_i' \gamma) \right) \tag{4}
\]

where

\[
\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp(-t/2)dt
\]

is the cumulative distribution function (cdf) of a standard normal random variable, \( y = [y_1, \ldots, y_n] \) and \( X = [x_1', \ldots, x_n'] \). Alternatively, it can be assumed that \( S_i \) is distributed according to a leptokurtic distribution to take account of the cross-section heteroskedastic fluctuations of a possible separating hyperplane.\(^2\) To that end we adopt a Student’s \( t \) distribution with fixed degrees of freedom \( (v) \), in which case the cdf is given by:

\[
\Phi_v(x) = \frac{\Gamma[(v + 1)/2]}{\Gamma(1/2)\Gamma(v/2)v^{1/2}} \int_{-\infty}^{x} (1 + u^2/v)^{-(v+1)/2} du \tag{6}
\]

As \( v \to \infty \), \( \Phi_v(x) \to \Phi(x) \) and we get the normal distribution. The corresponding model is called Student’s \( t \) separating hyperplane switching regime (ST-SHSR).

The likelihood function for the EBSR is given by:

\[
L(\beta_1, \beta_2, \gamma, \sigma_1^2, \sigma_2^2; y, X) = \prod_{i \in A(z^*)}^{n} (2\pi \sigma_1^2)^{-1/2} \exp\left[-(y_i' - x_i' \beta_1)^2 / 2\sigma_1^2\right] \prod_{i \in B(z^*)}^{n} (2\pi \sigma_2^2)^{-1/2} \exp\left[-(y_i' - x_i' \beta_2)^2 / 2\sigma_2^2\right] \tag{7}
\]

where \( A(z^*) = \{i \in I: z_i \leq z^*\} \) and \( B(z^*) = \{i \in I: z_i > z^*\} \), \( I = \{1, 2, \ldots, n\} \).

\(^2\)When a large number of different countries are pooled together, residuals might show evidence of long tails (longer than normal tails).
Non-informative priors are used throughout. These priors are of the form
\[
\pi(\beta_1, \beta_2, \gamma, \sigma_1^2, \sigma_2^2) \propto (\sigma_1 \sigma_2)^{-1}
\] (8)
for the SHSR model and
\[
\pi(\beta_1, \beta_2, \gamma, \sigma_1^2, \sigma_2^2, z^*) \propto (\sigma_1 \sigma_2)^{-1} 1(k \leq z^* \leq n - k + 1)
\] (9)
for the EBSR model, where \(1(\cdot)\) denotes the indicator function. Informative (e.g., normal) priors may be used for regression coefficients but at this stage one could argue that use of such priors biases the results for or against the EKC hypothesis. Therefore, their use is not advisable. Model posteriors are analyzed using Markov chain Monte Carlo (MCMC) methods, as detailed next.

4. Bayesian computations

The purpose of Markov Chain Monte Carlo methods is to produce a sample of draws \(\{\theta^{(i)}\}\) for the parameters of a posterior kernel \(\pi(\theta|\text{Data})\) such that \(\{\theta^{(i)}\}\) converges in distribution to \(\pi\). The Metropolis algorithm and the Gibbs sampler are leading numerical posterior simulators that can be used to accomplish this task. The problem in its most general form can be described as generating random draws from a general density \(\pi(x)\), \(x \in \mathbf{X}\).

For the SRSF, the posterior distribution was analyzed using a random walk Markov Chain Monte Carlo method. First, in the Metropolis–Hastings algorithm (Metropolis et al., 1953; Hastings, 1970; Tierney, 1994; Tsionas, 1999) consider a candidate transition kernel with density \(q(x,y)\), \(x,y \in \mathbf{X}\) which generates potential transitions for a discrete time Markov chain evolving on \(\mathbf{X}\). A candidate transition to \(y\) generated according to the density \(q(x,.\) is then accepted with probability \(\alpha(x,y)\) given by
\[
\alpha(x,y) = \min \left\{ 1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)} \right\} \quad \text{if} \quad \pi(x)q(x,y) > 0
\]
\[
\alpha(x,y) = 1 \quad \text{if} \quad \pi(x)q(x,y) = 0
\]

Thus, actual transitions of the Hastings chain, take place according to a law with transition probability \(q(x,y)\alpha(x,y)\) (\(y \neq x\)) and a probability of remaining at the same point given by
\[
r(x) = \int q(x,y)[1 - \alpha(x,y)]dy
\]

A particularly simple approach that can implement the above method is to symmetric transition density \(q(x,y)\) in which case
\[
\alpha(x,y) = \min \{1, \pi(y)/\pi(x)\} \quad \text{if} \quad \pi(x)q(x,y) > 0.
\]
A convenient transition density is a uniform distribution centered at the current state \( x \). In our implementation the range of the transition density is adjusted every 50 passes to ensure that acceptance rates are not too high or too low. This is the approach originally suggested by Metropolis et al. (1953).

For the EBSR model, a Gibbs sampler (Gelfand and Smith, 1989; Tanner and Wong, 1987) has been used. The Gibbs sampler is an iterative Monte Carlo technique for numerical posterior integration in high-dimensional Bayesian models. For a posterior distribution \( \pi(\theta | y, X) \), the Gibbs sampler starts from an arbitrary initial parameter vector \( \theta^{(0)} \) and produces parameter draws \( \{ \theta^{(i)}, i = 1, \ldots, M \} \) that converge in distribution to the posterior \( \pi(\theta | y, X) \). These random drawings are produced as follows. For \( i = 1, 2, \ldots, M \):

1. Draw \( \theta_1^{(i)} \) from \( \pi(\theta_1 | \theta_{-1}^{(i-1)}, y, X) \)
2. Draw \( \theta_2^{(i)} \) from \( \pi(\theta_2 | \theta_{-2}^{(i-1)}, y, X) \)

\[ \vdots \]

3. Draw \( \theta_k^{(i)} \) from \( \pi(\theta_k | \theta_{-k}^{(i-1)}, y, X) \)

where \( \theta_{-i} = [\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k] \).

This requires that univariate conditional distributions are in a form suitable for practical random variate generation. Generally, the degree to which this can be accomplished varies greatly from application to application.

The required conditional distributions of the parameters for the EBSR model, are as follows:

\[
\begin{align*}
\beta_1 | \beta_2, \sigma_1^2, \sigma_2^2, z^*, y, X & \sim N\left[ \hat{\beta}_1, \sigma_1^2 (X_1'X_1)^{-1} \right] \\
\beta_2 | \beta_1, \sigma_1^2, \sigma_2^2, z^*, y, X & \sim N\left[ \hat{\beta}_2, \sigma_2^2 (X_2'X_2)^{-1} \right]
\end{align*}
\]

\[
\begin{align*}
(y_1 - X_1 \beta_1)'(y_1 - X_1 \beta_1)/\sigma_1^2 & \text{ is } X^2(n^*) \quad \text{given } \beta_1, \beta_2, \sigma_2^2, z^*, y, X \\
(y_2 - X_2 \beta_2)'(y_2 - X_2 \beta_2)/\sigma_2^2 & \text{ is } X^2(n - n^*) \quad \text{given } \beta_1, \beta_2, \sigma_2^2, z^*, y, X
\end{align*}
\]

where \( y_1, X_1 \) are the observations corresponding to \( A(z^*) = \{ z_i \leq z^* \} \), \( y_2, X_2 \) are the observations corresponding to \( B(z^*) = \{ z_i > z^* \} \) and \( n = \#A(z^*) \), and \( \bar{\beta}_j \) \((j = 1, 2)\) are obvious least squares estimators within the corresponding subsamples. Finally, the conditional distribution of \( z \) is

\[
\begin{align*}
\pi(z^* | \beta_1, \beta_2, \gamma, \sigma_1^2, \sigma_2^2, z^*, y, X) & \propto \\
\prod_{i \in A(z^*)} (2\pi \sigma_1^2)^{-1/2} \exp\left[-(y_i - x_i'\beta_1)^2/2\sigma_1^2\right] \\
\prod_{i \in B(z^*)} (2\pi \sigma_2^2)^{-1/2} \exp\left[-(y_i - x_i'\beta_2)^2/2\sigma_2^2\right]
\end{align*}
\]
With the exception of the conditional distribution of $z^*$, all other conditional distributions are in standard form and random sampling is particularly easy. To get a random draw from $\pi(z)$ we have used a griddy Gibbs sampler in the interval $[z_{\text{min}}, z_{\text{max}}]$, see Ritter and Tanner (1992).

4.1. Convergence

For the SREB model we have used 100 points for the griddy random number generator of the conditional distribution of $z^*$. We have used 5000 Gibbs sampler passes with an initial burn-in phase consisting of 3000 passes. For the SRSH model, we have used 10,000 Metropolis–Hastings passes with a burn-in period consisting of 7000. Convergence has been assessed using standard criteria (Gelman and Rubin, 1992; Geyer, 1992). For the application of this paper, convergence was found to occur in the first few hundred passes of the Markov chain Monte Carlo samplers. For the SREB we have used an informative but locally uniform prior for $\gamma$. This prior is of the form

$$\gamma \sim N(0, 10^{-5}I)$$

Results were not sensitive to the choice of prior.

4.2. Computation of average posterior regime probabilities

For each parameter draw, the probability that a given country belongs to regime 2 can be recorded. At the end of the MCMC sampling scheme there is a number of draws from the posterior distribution of this probability. Posterior means were computed based on this distribution. One can compute other measures of the distribution (for example standard deviations) and, of course, the exact distribution of the probability. Such information, however, cannot be presented in a satisfying manner when the number of countries is large, as in our case. Based on estimates of the posterior mean of the regime 2 probability, regime 1 otherwise, these separation results are available from the authors upon request. What is the important aspect of these results, is how sensitive the results are to different distributional specifications and alternative dependent variables in EKC switching regression models. These results are presented in Section 6.1.

5. Data

Many EKC studies have chosen ambient concentrations as the dependent variable in their regressions as the negative impact of emissions in terms of air quality is positively correlated with the quantity of the pollutant per unit of area (Grossman and Krueger, 1992, 1995; Shafik, 1994; Kaufmann et al., 1997). It is worth mentioning that if we want to find a causal relationship between environmental damage and economic activity, ambient concentrations do not provide the most proper indicator of environmental impact. The alternative is the use of
emission estimates as well as an approximation of environmental degradation like deforestation.

The EKC concept is dependent on the state of the economic activity. The main explanatory variable is the per capita income. As population grows and economic growth takes off, forests are being cut to provide materials for construction, land for cultivation, etc. Usually, the larger the size of economic activity (approximated by GDP per capita) the larger is the depletion of natural resources. By the laws of thermodynamics the use of natural resources implies the production of waste. The EKC literature assumes that there are no limits to growth.

Income per capita represents consumer’s purchasing power and approximates the consumption patterns. We accompany income per capita with population density as explanatory variables to capture both scale and composition of consumption activities. We also consider the importance of some variables like the distribution of GDP in manufacturing in order to examine the importance of specific sources of pollution in the analysis. We also take into consideration some other socio-economic variables to see their influence in this kind of analysis. Finally, dealing with the long-run character of CO$_2$ emissions we include in the regressors the growth rate of the economies of the countries under examination.

Thus, we have used the following variables: GDP per capita in purchasing power parity (million $, 1991); the average annual growth rate (%) of GDP in 1980–1991; population density (1993, per 1000 hectares); infant mortality rate (per 1000 live births in 1990–1995); urban population (% of total, 1995); distribution of GDP in manufacturing (% of total, 1995); deforestation (average % change, 1980–1990); and CO$_2$ per capita (emissions in tons). The source of the last variable is Rodenburg et al. (1995) while the rest of the variables were obtained from World Resources WRI, 1991, 1994. Frequency distributions for the main variables of the study are reported in Fig. 1. 3

6. Discussion of results

6.1. Separating hyperplane model

6.1.1. Magnitudes of $\beta_1$ and $\beta_2 - \beta_1$

In the deforestation model estimates of $\beta_1$ are $-0.429$ and $-0.469$ (see Table 1) for the normal and Student’s $t$, respectively, i.e. higher GNP/c leads to lower deforestation. The estimate of $\beta_2 - \beta_1$ is not statistically significant thus the two

---

3The countries considered in our study are: Ethiopia, Mozambique, Bangladesh, Madagascar, Zambia, Uganda, Niger, Benin, India, Kenya, Pakistan, Nigeria, Ghana, Sri Lanka, Indonesia, Senegal, Zimbabwe, Philippines, Morocco, Dominican R., Honduras, Thailand, El Salvador, Nicaragua, Guatemala, Cameroon, Jamaica, Paraguay, Botswana, Ecuador, Tunisia, Colombia, Chile, Peru, Mexico, Brazil, Panama, Hungary, Algeria, Portugal, Venezuela, Greece, Spain, Ireland, Singapore, former Czechoslovakia, UK, Italy, Australia, Belgium, the Netherlands, Austria, France, Germany, Finland, Denmark, Canada, Japan, Norway, USA and Switzerland.
regimes are not different in terms of coefficients. In the CO$_2$ model the estimates of $\beta_1$ are 1.078 and 0.992, respectively. Again $\beta_2 - \beta_1$ is not significant. This can be seen also from 95% Bayes probability intervals for $\gamma_3$. They are, generally, very dispersed and do include the origin. We conclude that the two regimes do not differ in terms of $\beta_1$ and $\beta_2$. 

Fig. 1. Frequency distributions for main variables of the study.
Table 1

Posterior results for the SHSR\textsuperscript{a}

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Dependent variables</th>
<th>CO\textsubscript{2}</th>
<th>\textit{CO\textsubscript{2}}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Normal</td>
<td>Student’s \textit{t}</td>
</tr>
<tr>
<td>\textit{Deforestation}</td>
<td></td>
<td>1.07</td>
<td>0.99</td>
</tr>
<tr>
<td>\textit{Normal}</td>
<td>-0.43</td>
<td>0.23</td>
<td>0.25</td>
</tr>
<tr>
<td>\textit{Student’s \textit{t}}</td>
<td>-0.20</td>
<td>0.46</td>
<td>0.47</td>
</tr>
<tr>
<td>\textit{Normal}</td>
<td>-0.12</td>
<td>3.44</td>
<td>3.56</td>
</tr>
<tr>
<td>\textit{Student’s \textit{t}}</td>
<td>0.4</td>
<td>4.36</td>
<td>4.21</td>
</tr>
<tr>
<td>\textit{Normal}</td>
<td>1.98</td>
<td>5.42</td>
<td>5.27</td>
</tr>
<tr>
<td>\textit{Student’s \textit{t}}</td>
<td>0.16</td>
<td>3.42</td>
<td>3.49</td>
</tr>
<tr>
<td>\textit{Normal}</td>
<td>1.67</td>
<td>4.66</td>
<td>4.52</td>
</tr>
<tr>
<td>\textit{Student’s \textit{t}}</td>
<td>-0.88</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>\textit{Normal}</td>
<td>1.83</td>
<td>0.18</td>
<td>1.02</td>
</tr>
<tr>
<td>\textit{Student’s \textit{t}}</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}Notes. Student’s \textit{t} has 1 d.f. (i.e. \(\nu = 1\)). Posterior standard deviations in parentheses.

6.1.2. Magnitudes of \(\sigma_1\) and \(\sigma_2 - \sigma_1\)

In the deforestation case the difference \(\sigma_2 - \sigma_1\) is 1.833 and 0.592 for the normal and Student’s \textit{t} models with S.E.s 0.175 and 1.016. Thus according to the normal specification \(\sigma_1 \neq \sigma_2\) but not according to the Student’s \textit{t} specification. In the CO\textsubscript{2} model the variance differences are \(-0.807 and -0.815 with S.E.s 0.083 and 0.086, respectively, i.e. highly significant. Thus we may conclude that the two regimes differ in terms of error variances of the environmental degradation equation.

This means that the variances of deforestation or CO\textsubscript{2} equations are different and therefore policy changes, which affect GNP will not have the same result in high-income and low-income countries because of the different variances. When we use CO\textsubscript{2} as the dependent variable we find that \(\sigma_1 > \sigma_2\) so low-income countries have higher dispersion. Therefore policy measures which increase GNP will decrease CO\textsubscript{2} emissions with higher certainty in high-income countries. The opposite is true if we use deforestation as a dependent variable.

6.1.3. Magnitudes of \(\gamma_1 \ldots \gamma_5\) (separating hyperplane coefficients)

As can be seen from their posterior means and posterior standard errors none of the regressors (urbanization, industrialization, mortality, growth or GNP) appears to be a separator. The sign of \(\gamma_5\) (the coefficient of GNP in the separating hyperplane) is positive when we consider the deforestation equation and negative when we consider the CO\textsubscript{2} equation although statistically insignificant in both
cases. In this sense there is no evidence to support separation according to GNP or other economic and demographic variables, which implies no support for an EKC.

6.1.4. Discussion of country membership to regimes
We found that the two regimes are different in terms of $\sigma^2$ although this cannot be explained by economic and demographic variables. Due to the fact that estimates of $\gamma_i$ appear to be insignificant, allocating countries to the two regimes according to the separating hyperplane should be viewed with conservatism. However, it is useful to compare separation probabilities across models and across dependent variables.

6.1.5. Discussion of differences between normal and Student’s $t$ specifications
Fig. 2a,b presents graphs of the probability of regime 2 for the Student’s $t$ vs. the normal specification, using deforestation and $CO_2$ as dependent variables, respectively. Fig. 2c,d plots the same probability for $CO_2$ vs. deforestation for the Student’s $t$ and normal distributions, respectively. From Fig. 2a,b, we see that Student’s $t$ and normal distributions give comparable results only for $CO_2$. For deforestation, there is significant disagreement in the separation probabilities — although they tend to be positively correlated, and strongly so under the Student’s $t$ specification.

Fig. 2c,d reveals significant divergence in the separation properties of different dependent variables. Under a normal specification (Fig. 2c) probabilities of regime 2 for $CO_2$ and deforestation, show significant scattering around their regression line. Under a Student’s $t$ specification not only that happens but they also tend to be negatively correlated. Therefore it seems that not only the choice of dependent variable, but also the stochastic specification of switching regression models, affects separation significantly according to environmental degradation and economic-demographic factors. Thus applied researchers who venture into testing for the existence of an EKC, should not only examine alternative dependent variables and alternative sets of regressors, but they should also examine different stochastic specifications for their switching regime or linear regression models. This adds another dimension of model uncertainty, yet it is necessary practice because EKCs are sensitive to all these factors. To put things in a different manner, given the choice of dependent and explanatory variables, EKCs will be sensitive to alternative distributional assumptions about the error term.\footnote{Distributional assumptions are important. This is the case in that logs vs. levels or fixed effects vs. random effects can have an impact on the results in panel regression studies.} In cross-sectional studies, this is something that applied researchers should expect, and therefore they should properly account for it in order to avoid excess sensitivity of results to the nature of the disturbance term.

6.2. Exogenous break model

The most striking result that emerges from the exogenous break model is our
Fig. 2. Probabilities of regime 2: (a) deforestation; (b) CO₂; (c) normal distribution; and (d) Student’s t.
Table 2
Posterior results for the EBSR (dependent variable: deforestation)*

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Regime 1 coefficients</th>
<th>Regime 2 coefficients</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>8.16 (2.40)</td>
<td>0.220 (9.81)</td>
<td>7.94 (9.98)</td>
</tr>
<tr>
<td>Urbanization</td>
<td>−0.00038 (0.014)</td>
<td>−0.0084 (0.016)</td>
<td>0.008 (0.02)</td>
</tr>
<tr>
<td>Industrialization</td>
<td>0.00086 (0.016)</td>
<td>0.015 (0.047)</td>
<td>−0.14 (0.051)</td>
</tr>
<tr>
<td>Mortality</td>
<td>−0.018 (0.0086)</td>
<td>0.162 (0.246)</td>
<td>−0.180 (0.246)</td>
</tr>
<tr>
<td>Growth</td>
<td>−0.080 (0.088)</td>
<td>0.180 (0.389)</td>
<td>−0.260 (0.403)</td>
</tr>
<tr>
<td>log GNP</td>
<td>−0.843 (0.345)</td>
<td>−0.181 (1.007)</td>
<td>−0.662 (1.049)</td>
</tr>
<tr>
<td>σ²</td>
<td>1.24 (0.27)</td>
<td>0.202 (0.48)</td>
<td>1.037 (0.572)</td>
</tr>
</tbody>
</table>

*Posterior standard deviations in parentheses.

failure to find a posterior mean for the break \( z^* \) significantly lower than the maximum value of \( z \) (i.e. log GNP/c in our specification). Estimates of \( z^* \) are 9.418 and 9.453 with posterior S.E.s 0.572 and 0.430 for deforestation and CO\(_2\) equations, respectively. These are approximately 94% of the maximum value of log GNP/c. Moreover, the differences of \( \beta_{11} - \beta_{21} \) appear to be insignificant. The same is true for the difference \( \sigma_{2}^2 - \sigma_{1}^2 \) (estimates are 1.037 and 0.217 with S.E.s 0.572 and 0.43, respectively — see Tables 2 and 3). Thus, we find no evidence for a structural break from log GNP/c.

For the deforestation model, mortality rates and log GNP/c are significant in the first regime but none of the variables is statistically significant in the second regime. For the CO\(_2\) model only industrialization and log GNP/c are significant in the first regime. Notice that higher GNP implies lower deforestation in the first regime but higher CO\(_2\) emissions. We conclude that measuring environmental degradation using deforestation or CO\(_2\) makes a difference.

Only for the deforestation model we find that \( \sigma_{2}^2 - \sigma_{1}^2 \) is close to significance (1.037 and with S.E. 0.572). So it turns out that we find very little evidence to support the existence of two regimes with differences in coefficients and/or error variances.

6.2.1. Marginal posterior distribution of \( z^* \)
Marginal posterior distributions of \( z^* \) for the deforestation and CO\(_2\) model are shown in Fig. 3. From inspection it follows that the mass is concentrated between 90% and 95% of the maximum value of \( z = \log \text{GNP} \). It is remarkable that for
Table 3  
Posterior results for the EBSR (dependent variable: CO$_2$)$^a$

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Regime 1 coefficients</th>
<th>Regime 2 coefficients</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>$-6.23$</td>
<td>$6.02$</td>
<td>$-12.25$</td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td>(14.33)</td>
<td>(14.31)</td>
</tr>
<tr>
<td>Urbanization</td>
<td>0.0063</td>
<td>0.0076</td>
<td>$-0.0013$</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.0203)</td>
<td>(0.0217)</td>
</tr>
<tr>
<td>Industrialization</td>
<td>0.025</td>
<td>$-0.022$</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.045)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Mortality</td>
<td>$-0.0057$</td>
<td>0.362</td>
<td>$-0.368$</td>
</tr>
<tr>
<td></td>
<td>(0.0049)</td>
<td>(0.278)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Growth</td>
<td>0.0448</td>
<td>0.434</td>
<td>$-0.389$</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.365)</td>
<td>(0.369)</td>
</tr>
<tr>
<td>log GNP</td>
<td>0.771</td>
<td>$-0.731$</td>
<td>1.502</td>
</tr>
<tr>
<td></td>
<td>(0.185)</td>
<td>(1.468)</td>
<td>(1.467)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.406</td>
<td>0.188</td>
<td>0.217</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.418)</td>
<td>(0.430)</td>
</tr>
</tbody>
</table>

$^a$Posterior standard deviations in parentheses.

both the log CO$_2$ and deforestation models these distributions are approximately the same.

Fig. 3. Posterior of break point (relative to maximum log GNP).
7. Concluding remarks

In this paper we tested empirically the existence of an EKC using switching regime models and Bayesian Markov chain Monte Carlo methods. The evidence indicates that higher GNP per capita leads to lower deforestation. A monotonic relationship exists between environmental degradation and income because economic growth does reduce deforestation and therefore in that sense there is an EKC. The measurement of environmental deterioration using deforestation or CO₂ does make a difference. Additionally, we find that the demographic and economic variables considered here could not explain the difference between low and high damage countries.

Our main conclusions were that the two regimes do not differ in terms of their parameters. They do differ in terms of error variances at the environmental degradation equations and there is no evidence to support separation according to GNP or other economic and demographic variables which does not provide support for an inverted-U-shaped EKC.

The rejection of the EKC implies that we may not have a certain environmental degradation along a nation’s development route. Economic growth appears to improve environmental quality in developing countries and so policies to stimulate growth are recommended. There is empirical evidence that trade liberalization is positively correlated with growth. To that extent our evidence supports the view that higher rate of liberalization is associated with low degradation.

Acknowledgements

The paper benefited from the comments of an anonymous referee.

References

Stern, D.I., Auld, M.S., Common, M.S., Sanyal, K.K., 1998. Is there an environmental Kuznets curve for...


