Environmental Kuznets curve for some countries - regression and agent-based approach

Pawel Bartoszczuk¹,², Tieju Ma¹ & Yoshiteru Nakamori¹
¹Japan Advanced Institute of Science and Technology
School of Knowledge Science, Hokuriku, Japan,
²Systems Research Institute of the Polish Academy of Science, Warsaw, Poland

Abstract

In the past few years there has been a lot of research on the topic of environmental quality and economic growth. This paper describes the theory of Environmental Kuznets Curve and illustrates the concept with special reference to some developed countries. One of the objectives of this study is to test the shape of the relationship between income and environmental degradation. We use regression and agent based approach to identify the models of income and environmental abuse. This study uses data to try and capture some of the effects that have happened in recent years in some countries. Our goal is to attempt if model given by the agents can fit the data better than regression models. The paper emphasises ecological consequences of economical development.

1 Introduction

The relationship between economic development and pollution is a very complex issue. Natural resources serve as inputs into production of many goods and services. If the methods of production were not changed then damage to the environment would be linked to the scale of global activity. For this reason, in the last few years several studies have tried to characterize this problem as an empirical reduced form relationship.
Some evidence suggests that at very low level of economic activity environmental impacts are generally low but as development proceeds, pollution increases rapidly. At higher level of development, economic growth eventually redress the environmental impact of the early stages of development and that growth lead to further environmental improvements in developed countries or tend to fix environmental problems. The forces leading to change in the composition and techniques of production may be sufficiently strong to offset the adverse effects of increased economic activity on the environment.

This has come to the relationship called the Environmental Kuznets Curve. This correlation proposes there is an inverted U shape relation between various indicators of environmental degradation and income per capita (Fig. 1). It draws its inspiration originally from the income distribution theory developed by Kuznets [1], who found an inverted-U shape between an indicator of income inequality and the level of income.

The objective of this work is to critically review existing knowledge from the literature and attempt to find the relationship between economic growth and the pollution for some developed countries from available data. Such attempt will be conducted both by agent and regression method.

2 Theoretical background of environmental Kuznet’s relationship

The origins of the Environmental Kuznets Curve (EKC) hypothesis are somewhat ambiguous, and appear to be product of various studies conducted in early 1990s. The first of studies were made by Grossman and Krueger in 1991, and by Shafik and Bandyopadhyay in 1992 of air pollution measures in cross section of countries for different years, cited by [2].

Shafik and Bandyopadhyay [3] discovered a turning point for deforestation around 2000 dollars. Two air pollutants conform to EKC hypothesis and the turning point was found between 3000-4000 dollars. In the
contrary, lack of clean water and lack of urban sanitation did not correspond to EKC, but declined monotonically. Grossman and Kreuger [4] found turning point for SO$_2$ and dark matter for about 4000-5000 dollars. For water quality this indicator equals at least 7500 dollars and for dissolved oxygen - 10000 dollars.

Shafik [5] used data from a wide range of countries at different levels of development, and he found an inverted U relationship for the suspended particulate matter (SPM) and sulphur dioxide. He observed the turning point for SO$_2$ around 3000 dollars. In another study, Grossman and Kreuger [6] established the turning point for an income of country equalled to 8000 dollars. The EKC curve was found also for ambient levels of air pollutants, SO$_2$ and SPM, with turning point in the range 3000 to 5000 dollars of per capita GDP.

A number of explanations exist for the observed inverse-U relationship. First, it could be explained by transformation from agrarian economies through industrial economies to clean service ones [7]. At low level of economic development, the share of industry first is very low and rises with income increase. Corresponding environmental pollution first rise and then, when there is a transformation from industry to service, falls with income growth, controlling for all other influences transmitted through income.

The other most common explanation is that demand for environmental quality rises faster with income than demand for other goods and services. This would be the case, for instance, if there exists a threshold level of income below which resources are devoted to environmental protection. Countries living on low level of economic development may find it extremely difficult to protect environment. When income grows, people presumably become both more able and more willing to sacrifice some consumption to protect the environment. Income-elastic demand for environmental quality is therefore one element that could generate a pollution path that eventually turns downward.

Some researchers note that decisive evidence of an EKC relationship applies only to a few pollutants, thus making it difficult to use this evidence to speculate more generally. Others have pointed out that, even for those pollutants displaying EKC characteristics, aggregate global emissions are projected to rise over time, demonstrating that existence of an EKC relationship for such pollutants does not necessarily imply that, at the global level, any associated environmental damage is likely to disappear with economic growth. There is some evidence that turning points of global pollutants, such as CO$_2$ emissions and other greenhouse gases are estimated at considerably higher incomes than more local ones. One interpretation of this is that people do not care much about global warming and climate change. Global warming and depletion of the ozone layer are rather recent public concerns. Based on existing examinations, some policy analysts have concludes that developing countries will automatically become cleaner as their economics grow. Others have argued that it is natural for the poorest countries to be more polluted as they develop.

From the other side, many authors articulate that economic growth is not a panacea for environmental quality. Protecting the capacity of ecological systems to sustain welfare is of as much of importance to poor countries as it is to those that are rich. A lot of researchers claim that the turning point achieved by
developed countries, if not yet in all environmental indicators could be a result of the migration of polluting industries to developing countries. That is, some of the dirtier industries, especially mining and raw materials processing, are now disappearing from the richer countries but becoming more important to the poorest ones.

Since, eventually, there will be no countries, to which pollution making will be exported, there may be a limit to the ability of nations to reduce pollution, or at least the costs will be more obvious for later entrants into the game. Therefore, the process of environmental improvement will not be indefinitely replicable, as the world's poorest countries will never have even poorer countries to which they can export their pollution.

From the other side, it can be opposite case. Developing countries may find it easier to pass the peak of the EKC, because of technologies that were not available at the time the developed countries were at the same stage of development. However, the are lots of other questions arising, like: at what level of per capita income is the turning point and if irreversible damages take place before environmental degradation turns down, and how they can be avoided [8].

3 Model specifications of relationship between emissions and income

Basically all researchers assume that empirical reduced-form relationship between per capita CO₂ emissions and GDP can be adequately described by parametric model, and specifically by a polynomial function of income [8]. The equation is either quadratic or cubic. It is estimated econometrically using cross-section data, i.e. from many countries at a single point of time, or panel data. Up most, an average income per capita is taken as the income variable. The environmental variable may be explicitly averaged, as by taking all reported hourly polluter concentrations reading for a country in the year and calculating their mean.

The most simple model specification shows a relationship between an environmental indicator (E) and the income per capita (y). The following linear, quadratic, log-linear and log quadratic forms are normally present in the studies on Environmental Kuznets Curve, eqns (1)-(4).

\[
E_{it} = B_0 + B_1 y_{it} + \epsilon_{it}, \tag{1}
\]

\[
E_{it} = B_0 + B_1 y_{it} + B_2 y_{it}^2 + \epsilon_{it}, \tag{2}
\]

\[
E_{it} = B_0 + B_1 \ln(y_{it}) + \epsilon_{it}. \tag{3}
\]

\[
E_{it} = B_0 + B_1 \ln(y_{it}) + B_2 (\ln(y_{it}))^2 + \epsilon_{it}. \tag{4}
\]

where:

- B -parameter to be estimated
- E-environmental indicator
error term
\( y \)-income per capita
\( t \)-time trend.

Now the linear-in variables model will be considered: eqn (5)

\[
E_t = B_0 + B_1 y_t + B_2 y_t^2 + B_3 y_t^3. \tag{5}
\]

Two first derivatives are of special interest. By setting them to zero, turning point is received. The first and second derivatives are:

\[
\frac{dE_t}{dx} = B_1 + 2B_2 y + 3B_3 y^2, \\
\frac{d^2E_t}{dx^2} = 2B_2 + 6B_3 y. \tag{6}
\]

Let assume the coefficient of \( B_3 \) in equation is equal to zero. In this case eqn (6) exhibits an inverted U-shape, if additionally \( B_2 < 0 \). If also \( B_1 + 2B_2 x = 0 \), one can obtain the abscissa of the turning point.

\[
y_{tp} = -\frac{B_1}{2B_2}. \tag{7}
\]

This \( x_{tp} \) is positive as long as \( B_2 < 0 \). Hence the appearance of the EKC hinge upon the sign of \( B_2 \). It is seen that linear in \(-\)variables form emerges a closed form for the turning point of EKC equation.

3 Agent–based model of relationship between emissions and income

Every analysed country has its own relationship between income and environmental abuse. Just using quadratic or cubic form to map the relationship is too roughly, and it is a hard task to build the model for different country by regression analysis. The first advantage of this approach is that instead of proposing an individual formula in advance by researcher, what can be difficult in most cases, the agents will find the right one. The second advantage is that the model given by the agents can fit the data very well, especially when data pattern is very complex and any formula that we can find can’t fit the data well.

A number of agents are sent to the data space of two dimensions which are economic dimension and pollution dimension, every agent tries to build local linear model by competition with each other in micro-level, and then in macro-level all agents build the global structure by cooperation with each other. In
micro-level the regression method and the related coefficient are used as agents’ built in knowledge and local rules along with the agents’ behavior of dying and expanding their territories, while in macro-level membership function is used to simulate the cooperation between agents along with agents’ location and strength.

The agents will evolve during the progress of building model and their level is 1 at the beginning. Every agent will build local linear model according to the specified algorithm.

Step 1: First, it counts the number of objects in its vision. If such number is smaller than 3, it dies, otherwise it conducts second step:
Step 2: linear regression analysis of the all the data objects in its vision. If the coefficient of correlation $r \geq p_2$, it goes to step four, otherwise goes to step 3.
Step 3: It decreases its vision by 1, that means $p_t^{t+1} = p_t^t - 1$, where $t$ is time. If $p_t^{t+1} = 0$ it dies, otherwise this procedure is repeated from the fist step.
Step 4: It builds local linear model and evolve into an agent with level 2.

Those agents who evolve into level 2 agents try to expand their territory. Here an agent’s territory means the subspace which includes all the data objects based on which the agent builds its local linear model. Further, the agent of level 2 expands its territory according to the following steps:

Step 1: The agent finds all frontier data objects, and memorizes them in list $f_0$.
Step 2: The agent moves to the location of a frontier data object of its territory.
Step 3: The agent looks for new data objects in its vision. If it finds new data objects, it goes to step 4, otherwise it goes to step 5.
Step 4: The agent does linear regression analysis with the data objects and those new data objects. If the related coefficient $r \geq p_2$, the agent succeeds in expanding its territory, that is to say those new data objects are included in this agent’s territory, otherwise it fails to expand its territory.
Step 5: If there are another frontier data objects in list $f_0$ that the agent hasn’t visit, it moves to one of them and goes to step 3. If the agent has visited all the data objects in list $f_0$, it memorizes its current territory as $T_1$ then goes to step 6.
Step 6: If $T_1 = T_2$ the agent stops expanding its territory and involve into level 3, otherwise it goes to step 1.

If an agent’s territory is included in another agent’s territory, it will die. When there are only level-3 agents left in the data space, they begin to cooperate with each other to build the global model. Further, integration is done by the way of introducing membership function [9].
Suppose there are $H$ level-3 agents left in the data space, the set of these agents is $A = \{A_1, A_2, \cdots, A_H\}$. For every agent $A_i \in A(i = 1, 2, \cdots, H)$, following local linear model is hold $y_i = f_i(x) = a_i x + b_i$. For a given $x^*$, its membership value for $A_i$ is $w_i$, then the output is $y^*$:

$$y^* = \sum_{i=1}^{H} w_i f_i(x^*) / \sum_{i=1}^{H} w_i$$

(8)

$w_i$ is defined in the following way, eqn (9):

$$w_i = M_i / N e^{\|x^* - C_i\|}$$

(9)

Where

$M_i$ - is the number of data objects in agent $A_i$’s territory,

$N$ - is the number of all data objects.

$C_{ix}$ - is agent $A_i$’s center in $x$-dimension, defined by eqn (10).

$$C_{ix} = \sum_{j=1}^{M_i} x_j / M_i$$

(10)

In the above, $p_1$, $p_2$, $p_3$ are design parameters. They can be set in three ways. The first is some experts set them. Here the experts mean those who have much knowledge and experience about the problem and this agent-based approach. The second is users set the parameters by trail and error. And the third is let the agent system walk through the space $(p_1, p_2, p_3)$ and find the best values for these three parameters. In this case we need set some measurements for the agent system. We are still working to improve the performance of this agent-based approach.

4 Example of application of presented method

The employed data set consists of observations on countries over 36 years period beginning in 1960 and ending in 1996. The figures concerning income are described by Gross Domestic Product Per Capita Purchasing Power Parities (GDP PPP) are taken from Demographia (available in the web). Purchasing power parity eliminate the differences in price levels between countries. The PPP’s are given in national currency unit per US dollar. The data relating to CO$_2$ emissions from fossil fuels comes from Oak Ridge National Laboratory.

It is very risky to take cross section data, while this can be similar to the result of regressing market data on quantity purchased per period on price data
from the same period. In the absence of other data, the prognosis can be very hazardous. Therefore, in conducted research a time-series data of carbon dioxide emission are applied. Carbon dioxide is global pollutant and measured in thousand metric tons.

To illustrate the concept of agent based modelling, we use data from very developed countries. The first step in building the agent-based model is to standardize raw data and putting them into a data space, eqns (11) - (12).

\[ x'_i = (x_i - x_{min}) \times \frac{\text{SpaceXSize}}{x_{max} - x_{min}}, \]  \tag{11} \\
\[ y'_i = (y_i - y_{min}) \times \frac{\text{SpaceYSize}}{y_{max} - y_{min}}. \]  \tag{12} \\

At this point, we found following characteristics of the data: \( x_{max}=19,366, \ y_{min}=12924, \ y_{max}=25156 \). We set the size of the data space \( \text{SpaceXSize} = \text{SpaceYSize} = 200 \) (Fig. 3). Every point in that space is called a data object. Then we sent 37 agents to the data space and each agent mount on a data object. Every agent has a vision \( p_1 \), which means its ability to observe the environment and is a design parameter (Fig 3). We simply set \( p_1 = 40 \).

Fig 3: Data space for Sweden and five local linear models.

After expanding territories, there are 5 agents left, which build the local linear model (Fig. 3). After integrating these local models and changing \( x' \) and \( y' \) back to \( x \) and \( y \), the global model is obtained (Fig. 4). For comparison, we present regression models of pollution-income model of some countries. Results of regression show there is inverted U-shape relationship between pollution and economic development.

The data were approximated by polynomial equation. This study has identified a bell shaped curve for the pollution intensity of GDP. The behavior of presented data implies that, starting from low income levels, emissions or concentrations start to increase. After reaching so called turning point they start to decline as income further increase (Fig. 4). Every analyzed country has its own relationship between income and environmental abuse.
The turning point is around 15 000 USD- PPP in three from fourth analyzed countries. Nevertheless, emissions of CO₂ in that point differ from country to country. They equal 16000 for Spain, but are much higher for France – 120000.

5 Conclusions

Results of conducted research show that agent based simulation technique applied in this study can fit the data very well, especially when data pattern is very complex and any mathematical formula of relationship between pollution and income is difficult to obtain. The outcome of this paper is in agreement with the theory and previous research, for presented developed countries. Nevertheless, both the shape of the inverted U-curve and value of turning point in pollution level corresponding to GDP is varying among countries. Agent based models shows, more precisely relationship between pollution and income. Local models are connected into one global model. This specific approach requires more investigation in future.

Empirical relationship between country’s GDP and the emissions of a
major greenhouse gas CO₂, ought to undergo a more careful examination. More work is required to improve the agent-based approach. Simplistic arguments that income growth will by itself take care of the pollution problems of the world are inaccurate. If the economic incentives facing producers and consumers do not change with higher incomes, pollution will continue to rise unchecked alongside the increasing scale of economic activity. More regular shape of the EKC curve is observed for countries where environmental issues are concerned, like for example Sweden. This suggests that other things besides GDP are important in predicting emissions levels (like policy). An important policy implication that comes from the results is that the low-income countries have an opportunity to learn from this history and avoid mistakes of developed countries.

Acknowledgments

We are thankful to Japan Society for the Promotion of Science for supporting part of this research.

References