

Is there an Environmental Kuznets Curve for South Africa? A Co-Summability Approach Using a Century of Data

Adnen Ben Nasr^a, Rangan Gupta^{b,*}, João Ricardo Sato^c

^a*Laboratoire BESTMOD, ISG de Tunis, Université de Tunis, Tunisia.*

^b*Department of Economics, University of Pretoria, South Africa.*

^c*Center of Mathematics, Computation and Cognition, Universidade Federal do ABC, Brazil.*

Abstract

There exists a huge international literature on the, so-called, Environmental Kuznets Curve (EKC) hypothesis, which in turn, postulates an inverted u-shaped relationship between environmental pollutants and output. The empirical literature on EKC has mainly used test for cointegration, based on polynomial relationships between pollution and income. Motivated by the fact that, measured in per capita CO_2 equivalent emissions, South Africa is the world's most carbon-intensive non-oil-producing developing country, this paper aims to test the validity of the EKC for South Africa. For this purpose, we use a century of data (1911-2010), to capture the process of development better compared to short sample-based research; and the concept of co-summability, which is designed to analyze non-linear long-run relations among persistent processes. Our results, however, provide no support of the EKC for South Africa, implying that to reduce emissions without sacrificing growth, policies should be aimed at promoting energy efficiency.

JEL Codes: C01; C22; Q53

Keywords: Environmental Kuznets Curve, CO_2 emissions; Output; Co-summability; South Africa

*Corresponding author:

Email addresses: `adnen.bennasr@isg.rnu.tn` (Adnen Ben Nasr),
`rangan.gupta@up.ac.za` (Rangan Gupta), `joao.sato@ufabc.edu.br` (João Ricardo Sato)

1. Introduction

There exists a huge international literature, both for developed and emerging economies, that focuses on the environmental pollutants (such as CO_2 , NO_x , and SO_2) and output nexus, which, in turn, is essentially involved in testing the validity of the so-called Environmental Kuznets Curve (EKC) hypothesis.¹ The hypothesis argues that the relationship between these two variables is inverted U-shaped, implying that environmental degradation increases with output during the early stages of economic growth, but declines with output after reaching a certain threshold. This shape is understandable since agents living in poor economies are more concerned with employment and income than with the environment; as a result, environmental regulation is limited at early stages of development. However, as economies gain in wealth, agents start to value the environment more, production technology tends to become cleaner, and more efficient regulatory institutions are formed (Dasgupta et al., 2002). Understandably, the implication of this hypothesis is that environmental degradation can be slowed at some point by policies that not only protect the environment, but also promote economic development.

The literature on EKC uses three different channels to explain the u-

¹The reader is referred to Aroui et al., (2012a, b), Berenguer-Rico and Gonzalo (2013), Duarte et al., (2013), and Ajmi et al., (forthcoming) for detailed literature reviews dealing with the EKC. It must be said that evidence is, at best, mixed, with the same depending upon the estimation techniques, the time periods and the country characteristics (Ajmi et al., forthcoming).

shaped relationship between pollutants and output: scale, composition, and technique effects (Grossmann and Krueger, 1995; and Brock and Taylor, 2005). Ceteris paribus (i) as scale of economic activity increases, emissions tend to rise; (ii) when the goods produced in an economy become cleaner, emissions fall through the composition effect, and; (iii) finally, emissions fall as the technology involved in production becomes less contaminating. The EKC hypothesis depends on the relative strength of the three effects. Ideally, to identify these three channels, one should resort to structural modelling, however, the empirical literature on the EKC has mainly used a reduced form approach, where by, one attempts to test for cointegration using polynomial relationships between pollution and income, with the former being treated as the dependent variable. Though, there does not seem to be a clear agreement about the order of the polynomial to be used (Berenguer-Rico and Gonzalo, 2014), the literature has primarily focussed on a quadratic structure (Arouri et al., 2012a, b).

In this paper, we use South Africa as our case study to test the existence or non-existence of the EKC, which, as discussed above, would tend to indicate whether the threshold level of development has been reached so as to lead the authorities in the country to adopt more efficient methods of production. An obvious question is: Why South Africa? To answer this, we need to look at some figures relating to emissions in South Africa. Measured in per capita CO_2 equivalent emissions in 2010, and excluding island states, South Africa is the world's most carbon-intensive non-oil-producing

developing country (EIA, 2010). Also, with 42 percent of the continents emissions coming from South Africa alone, it is the largest emitter of GHGs in Africa. Furthermore, South Africa is also a bigger emitter of CO_2 than all other Sub-Saharan African (SSA) countries combined (EIA, 2010). In this regard, it is important to note that South Africa is a signatory to the 1992 UNFCCC and its Kyoto Protocol. Under the Kyoto Protocol, the biggest emitters of GHGs are encouraged to implement measures that leads to energy efficiency, and also motivate energy sustainability policies. South Africa is classified as a non-annex developing country, and therefore does not have mandatory emission reduction targets. However, the country is committed to the fight against climate change (while simultaneously aiming for higher growth and employment, and reduced poverty), and has thus, put into place several national-level policies to reduce GHG emissions. Clearly then, South Africa can be considered as a compelling candidate for a separate study that investigates the possibly (nonlinear) relationship between CO_2 emissions and output?

However, to the best of our knowledge, there exists only one study by Kohler (2013), which has explicitly looked at the EKC in South Africa.²

²However, there are couple of studies for South Africa that does analyze causal relationships between output and CO_2 emissions. While Menyah and Wolde-Rufael (2010) uses the bounds testing approach (controlling for energy consumption, labor and capital besides the two variables of concern) to show that unidirectional causality runs from emissions to output, Cowan et al., (2014) analyzed the causality between output and CO_2 emissions for the BRICS in a panel setting (controlling for electricity consumption), and detected one-way causality running from output to emissions for South Africa. Clearly then evidence are contradictory.

This study, relying on the linear bounds testing approach (controlling for commercial energy-use and degree of openness), could not provide evidence of the existence of the EKC over a 50 year period of 1960 to 2009, given the statistical insignificance in the cointegrating relationship between output and its squared value with CO_2 emissions as the dependent variable.³

Against this backdrop, the objective of this study is to revisit the existence of the EKC in South Africa. We aim to extend the work of Kohler (2013) in two ways: First, we consider an unique data set on output and CO_2 emissions spanning a century of data, i.e., 1911-2010. Long span data sets are ideal for testing the EKC as it allows us to consider the transition process of an economy covering early stages of development to its current status; and second, we improve the linear methodological framework of Kohler (2013) and the literature in general, by testing for EKC using the concept of co-summability. The basic idea behind co-summability can be explained as follows: There is no doubt that co-integration theory is an ideal framework to study linear relationships among persistent (non-stationary) economic time series. However, the inherent linearity in the concepts of integration and co-integration makes it unsuitable to study non-linear long-run relations among persistent processes, which is clearly the case when testing the EKC. Thus, in this paper, we use the idea of co-summability, developed by Berenguer-Rico and Gonzalo (2014), which is built upon the concept order of summabil-

³Note that, Menyah and Wolde-Rufael (2010) could not detect a cointegrating relationship for output with CO_2 emissions, given energy consumption, labor and capital.

ity (Berenguer-Rico and Gonzalo, 2014), which, in turn, was developed to address non-linear transformations of persistent processes. Theoretically, a co-summable relationship is balanced, in the sense that the variables involved have the same order of summability, and describes a long run equilibrium that can be non-linear, given that the errors have a lower order of summability. To the best of our knowledge, this is the first attempt to use the concept of co-summability to test for the EKC in South Africa using a century of data. In fact, the only other paper that applies this approach to test for the EKC in the US is by Berenguer-Rico and Gonzalo (2013). The remainder of the paper is organized as follows: Section 2 lays out the basics behind the concept of co-summability. Section 3 presents the empirical model to be tested with a discussion of the data and the results. Finally, Section 4 concludes.

2. Summability, Balancedness and Co-Summability

2.1. Summability

The concept of Summability was proposed in Gonzalo and Pitarakis (2006) and more recently formalised in Berenguer-Rico and Gonzalo (2013) and Berenguer-Rico and Gonzalo (2014). A stochastic process $\{y_t\}$ is said to be summable of order δ , denoted as $S(\delta)$, if there exist a nonrandom sequence $\{m_t\}$ such that

$$S_T = \frac{1}{T^{\frac{1}{2}+\delta}} L(T) \sum_{t=1}^T (y_t - m_t) = O_p(1) \text{ as } T \rightarrow \infty,$$

where δ is the minimum real number such that S_T is bounded in probability and $L(T)$ is a slowly varying function.

This concept generalizes the concept of integration in the linear case and allows for establishing the order of summability for a number of nonlinear models. Indeed, if a linear time series y_t is $I(d)$, then it is also summable of order d , i.e. $S(d)$. In case where y_t is a nonlinear transformation, this necessitates the adoption of the concept of summability. In our empirical application we will estimate the order of summability of all variables to be included in the polynomial specifications.

2.2. Balancedness

On the basis of the analysis of summability, the ‘balance’ condition of the empirical relationship should be tested, that is testing whether both sides of the empirical equation of the model have the same order of summability; the equation $y_t = g(x_t, \theta)$ is said to be balanced if $y_t \sim S(\delta_y)$; $f(x_t, \theta) \sim S(\delta_f)$, and $\delta_y = \delta_f$: Thus, the null hypothesis of balancedness can be stated as $H_0 : \delta_y - \delta_f = 0$. Notice that under the null of balance, the associated confidence interval contains zero. Testing for balancedness is important for the validity of the empirical specification.

2.3. Co-summability

In addition to the balancedness test, co-summability is another pre-estimation testing as to the validity of the empirical model specification to be used. Two summable stochastic processes, $y_t \sim S(\delta_y)$ and $x_t \sim S(\delta_x)$, are said to be

co-summable if there exists $f(x_t, \theta_f) \sim S(\delta_y)$ such that $u_t = y_t - f(x_t, \theta_f)$ is $S(\delta_u)$, with $\delta_u = \delta_y - \delta$ and $\delta > 0$. In short, $(y_t, x_t) \sim CS(\delta_y, \delta)$.

Of course, the parametric function $f(\cdot, \theta_f)$ can be replaced with a general nonlinear function. As δ_y , δ_x , and δ are unknown in practice, Berenguer-Rico and Gonzalo (2014) proposed a consistent estimator with a slow convergence rate of $1/\ln(T)$. Notice that the "strong co-summability" will imply that the order of summability of u_t , δ_u , is statistically close to zero. Under the null of co-summability the confidence interval includes zero.

3. The Empirical Model, Data and Results

As discussed above, the Environmental Kuznets curve (EKC) hypothesis postulates an inverted-U-shaped relationship between per capita emissions and per capita income. Formally, following Berenguer-Rico and Gonzalo (2014), the relationship between pollution and income can be expressed in a polynomial form as follows:

$$p_t = \theta_0 + \theta_1 y_t + \theta_2 y_t^2 + \dots + \theta_k y_t^k \quad (1)$$

where p_t is a measure of pollution and y_t is a measure of income. It is important to highlight the following issues, related to the above equation: (i) In terms of the measures chosen for p_t , the most often used measures of air pollution is carbon dioxide (CO₂) emissions, which is what we use due to data availability for the entire period of 1911-2010; (ii) For the curvature of the

EKC, the order of the polynomial considered in the literature has either been quadratic (Holtz-Eakin and Selden, 1995) or cubic (Grossman and Krueger, 1995). Following Berenguer-Rico and Gonzalo (2014), we look till $k = 4$, and; (iii) p_t and y_t are sometimes used in levels (Grossman and Krueger, 1995), while at other times log transformations is used (Hong and Wagner, 2008), and, in some cases both are compared (Holtz-Eakin and Selden, 1995). Hence, we test the EKC both for the raw data and the same in natural logarithms of per capita real Gross Domestic Product (GDP) at constant 2005 rand-values, measuring y_t , and per capita CO_2 emissions. While data on real GDP per capita is obtained from the Global Financial Database (GFD), data on CO_2 comes from the Carbon Dioxide Information Analysis Centre, and is measured in thousand metric tons of carbon. The series is converted into its per capita form by dividing with the population figures, also obtained from the GFD. While, total CO_2 emissions data for South Africa dates back to 1884, reliable and continuous measures of real GDP per capita and population are only available from 1911, thus governing the starting point of our sample. While, the end point (2010), is determined by the latest data on CO_2 emissions. We plot these data, both in levels and in logarithms, in figure 1.

Next, we turn our attention to the discussion of the results, starting with the order of summability. Table 1 provides estimates and corresponding subsampling confidence intervals of the order of summability for all model variables in (1) for $k = 4$. The order of summability of GDP per capita

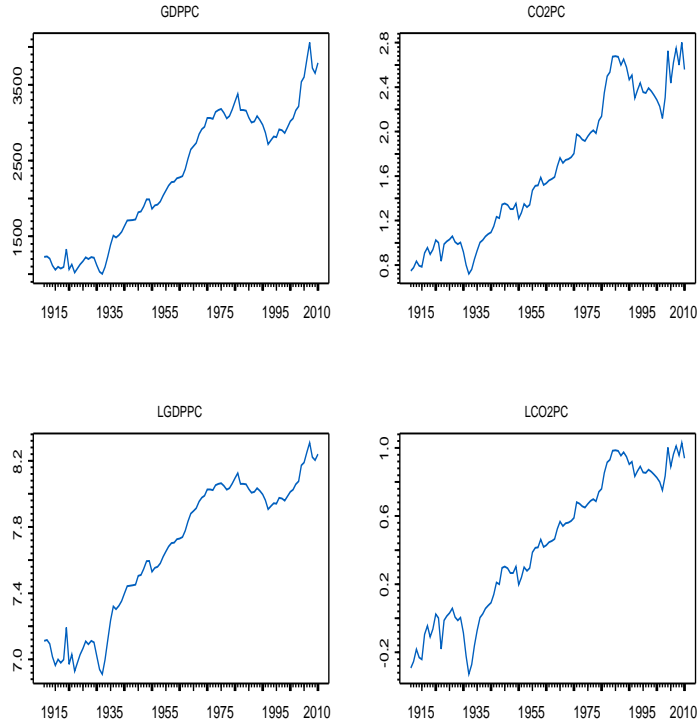


Figure 1: Evolution of per capita CO_2 emissions and per capita GDP

increases when higher orders are added to the regression. For all variables, except $\ln(CO_2pc)$, confidence intervals do not contain zero, suggesting that persistence is a feature of the data that needs to be taken into consideration.

Results of balancedness tests, taking the variables both in levels and in logarithms, are reported in Table 2. The results show that balancedness is only achieved under the linear specification when variables are in logarithms since zero is included in the corresponding confidence intervals. Understand-

Table 1: Order of Summability estimation

Variables	$\hat{\delta}$	I_{low}	I_{up}
CO_2pc	0.916	0.187	1.645
GDPpc	1.195	0.356	2.034
GDPpc ²	1.308	0.501	2.115
GDPpc ³	1.425	0.671	2.178
GDPpc ⁴	1.546	0.637	2.456
$\ln(CO_2pc)$	0.575	-0.061	1.211
$\ln(GDPpc)$	1.082	0.443	1.721
$\ln(GDPpc)^2$	1.097	0.523	1.672
$\ln(GDPpc)^3$	1.113	0.556	1.669
$\ln(GDPpc)^4$	1.128	0.542	1.713

Note: $\hat{\delta}$ denotes the estimated order of summability. I_{low} and I_{up} indicate the lower and upper bounds of the corresponding 95% confidence intervals. All the variables have been partially detrended.

ably, we do not need to consider the case of the raw data any further.

Table 3 reports the results of testing for co-summability when the variables are taken in logarithms. Notice that co-summability is not rejected for all considered specifications except the first one, which is a linear form without deterministic trend. However, as noted above, balancedness is only achieved under the linear specification, so we need to use the linear specification where a deterministic trend is included and the variables are in logarithms; which is the only specification that satisfies balancedness and co-summability. Based on these results, we can however, conclude, as in Kohler (2013), that the EKC does based on CO_2 emissions does not hold for South Africa. In other words, relative to Kohler (2013), using a longer

Table 2: Testing for Balancedness

Balancedness	$\hat{\beta}_n = \hat{\delta}_y - \hat{\delta}_g$	I_{low}	I_{up}
CO_2pc			
GDPpc	-2.126	-3.616	-0.637
GDPpc ²	-4.264	-6.894	-1.635
GDPpc ³	-6.331	-10.076	-2.586
GDPpc ⁴	-8.372	-13.309	-3.435
$\ln(CO_2pc)$			
$\ln(GDPpc)$	-0.469	-1.180	0.241
$\ln(GDPpc)^2$	-1.191	-1.944	-0.437
$\ln(GDPpc)^3$	-1.827	-2.920	-0.733
$\ln(GDPpc)^4$	-2.429	-3.844	-1.014

Note: $\hat{\delta}_y$ and $\hat{\delta}_g$ denote the estimated order of summability of the endogenous variable and the sum of the explanatory variables, respectively. The variables have been partially detrended. I_{low} and I_{up} denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals.

span of data and a more sophisticated methodology that allows us to capture nonlinearity in the long-run relationship between output and emission, does not help us to detect evidence in favor of the EKC.⁴

⁴Following the concern raised by Balcilar et al., (2010) and Ajmi et al., (forthcoming) that ignoring time-varying patterns in the relationships between income and CO_2 emissions may lead to erroneous conclusions on the EKC hypothesis and, hence, environmental policy, we also used a time-varying Granger causality analysis, as in Sato et al., (2007) to investigate whether there are dynamic relationships between income and CO_2 emissions for South Africa. The time-varying approach can easily be motivated, given that we are using a long-span of data, over which the economy has undergone changes in economic conditions, local and intergovernmental legislations, and technologies. The approach of Sato et al., (2007) involves using the so-called “curve causality” graphs to detect the nonlinear relationship between emissions and output over time; but the first step of this method requires one to detect the existence of time-varying Granger causality. However, we could not detect time-varying causality in any direction for either the first-differenced raw data or the data in its growth rate form (i.e., first-differences of natural logarithms), thus, once again corroborating the lack of the existence of EKC for South Africa. In addition,

Table 3: Testing for Co-summability

EKC	ln(CO2pc)	ln(CO2pc)	ln(CO2pc)	ln(CO2pc)	ln(CO2pc)	ln(CO2pc)
1	-6.668	-3.275	6.774	-1.417	250.849	181.993
t		0.008		0.007		0.007
ln(GDPpc)	0.927	0.434	-2.636	-0.068	-99.432	-72.785
ln(GDPpc) ²			0.235	0.034	13.010	9.626
ln(GDPpc) ³					-0.561	-0.421
$\hat{\delta}_{\hat{\epsilon}}$	0.483	0.407	0.511	0.394	0.503	0.271
I_{low}	0.025	-0.475	-0.016	-0.255	-0.071	-0.592
I_{up}	0.942	1.290	1.037	1.044	1.076	1.134

Note: $\hat{\delta}_{\hat{\epsilon}}$ denotes the estimated order of summability of the residuals calculated from the regression $y_t = f(x_t, \hat{\theta}_f) + \hat{\epsilon}$ (Berenguer-Rico and Gonzalo (2013)). I_{low} and I_{up} denote the lower and upper bounds of the corresponding 95% subsampling confidence intervals.

4. Conclusions

There exists a huge international literature on the so-called Environmental Kuznets Curve (EKC) hypothesis, which in turn, postulates an inverted u-shaped relationship between environmental pollutants and output. The empirical literature on EKC has mainly used a reduced form approach, where by, one attempts to test for cointegration using polynomial relationships between pollution and income, with the former being treated as the dependent variable. Motivated by the fact that, measured in per capita CO_2 equivalent emissions in 2010, and excluding island states, South Africa is the world's

as in Kohler (2013), using the Johansen and Juselius (1992), cointegration tests, we too failed to detect any cointegrating relationship between output and its squared value with CO_2 emissions for our data, based on both the raw data and its natural logarithms. Of course, before applying the cointegration tests, we first tested for unit root tests, and could not reject the null of non-stationarity for the variables in levels and log-levels. Further details on these results are available upon request from the authors.

most carbon-intensive non-oil-producing developing country, this paper aims to test the validity of the EKC for South Africa. For this purpose, we use a century of data, covering the period of 1911-2010, to capture better the process of development; and the concept of co-summability, which is designed to analyze non-linear long-run relations among persistent processes. Our results, however, provide no support of the EKC for South Africa. Stated alternatively, we find that a linear specification where a deterministic trend is included tends to capture best the relationship between the natural logarithms of per capita CO_2 emissions and per capita real GDP. Our results imply that for South Africa to reduce emissions, it will need to sacrifice growth. This is not a feasible solution for a country plagued with high unemployment, poverty and inequality, hence, policies aimed at promoting energy efficiency should be implemented in order to decrease CO_2 emissions without adversely affecting economic growth.

References

- [1] Ajmi A. N., Hammoudeh, S., Nguyen, K. N. and J. R. Sato, (forthcoming): “A new look at the relationships between CO2 emissions, energy consumption and income in G7 countries: the importance of time variations ” *Energy Economics*.
- [2] Arouri, M. E-H., Ben Youssef, A., M'henni, H., and C. Rault, (2012a): “Energy consumption, economic growth and CO2 emissions in Middle East and North African countries” *Energy Policy*, 45, 342-349.

- [3] Arouri, M. E-H., Ben Youssef, A., M'henni, H., and C. Rault, (2012b): “Empirical Analysis of The EKC Hypothesis for Sulfur Dioxide Emissions in Selected Middle East and North African Countries” *The Journal of Energy and Development*, 37, 207-226.
- [4] Balcilar, M., Ozdemir, Z. A. and Arslanturk, Y., (2010). “Economic growth and energy consumption causal nexus viewed through a bootstrap rolling window” *Energy Economics*, 32(6), 1398-1410.
- [5] Berenguer-Rico, V. and J. Gonzalo, (2013): “Co-summability: from linear to non-linear co-integration.” *Universidad Carlos III de Madrid Working Paper 13-12*.
- [6] Berenguer-Rico, V., and J. Gonzalo, (2014): “Summability of stochastic processes: A generalization of integration for non-linear processes,” *Journal of Econometrics*, 178, 331-341.
- [7] Brock, W. A., and M. S. Taylor (2005): “Economic growth and the environment: A review of theory and empirics, ” *Handbook of Economic Growth*, 1, 1749-1821.
- [8] Cowan, W. N., Chang, T., Inglesi-Lotz, R., and R. Gupta, (2014): “The nexus of electricity consumption, economic growth and CO2 emissions in the BRICS countries, ” *Energy Policy*, 66, 359-368.
- [9] Dasgupta, S., Laplante, B., Wang, H., and D. Wheeler (2002): “Con-

- fronting the Environmental Kuznets Curve,” *The Journal of Economic Perspectives*, 16, 147-168.
- [10] Duarte, R., Pinilla, V., and A. Serrano (2013): “Is there an environmental Kuznets curve for water use? A panel smooth transition regression approach,” *Economic Modelling*, 31, 518-527.
- [11] Energy Information Administration (EIA). (2010): “Energy Statistics for Countries.”
- [12] Gonzalo, J. and J. Pitarakis (2006): “Threshold Effects in Cointegrating Relationships,” *Oxford Bulletin of Economics and Statistics*, 68, 813-833.
- [13] Grossman, G. M., and A. B. Krueger (1995): “Economic growth and environment,” *Quarterly Journal of Economics*, 110, 353-377.
- [14] Holtz-Eakin, D., and T. M. Selden (1995): “Stoking the fires? CO2 emissions and economic growth,” *Journal of Public Economics*, 57, 85-101.
- [15] Hong, S. H., and M. Wagner (2008): “Non-linear cointegration analysis and the environmental Kuznets curve,” *Mimeo*.
- [16] Johansen, S., and K. Juselius (1992): “Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK,” *Journal of Econometrics*, 53, 211-244.

- [17] Kohler, M (2013): “CO2 Emissions, Energy Consumption, Income and Foreign Trade: A South African Perspective ” Economic Research Southern Africa (ERSA), Working Paper No. 356.
- [18] Menyah, K., and Y. Wolde-Rufael (2010): “Energy consumption, pollutant emissions and economic growth in South Africa ” Energy Economics 32, 1374-1382.
- [19] Sato, J.R., Morettin, P.A., Arantes, P.R., and J.E. Amaro (2007): “Wavelet based time-varying vector autoregressive modelling ” Computational Statistics and Data Analysis, 51, 5847-5866.

Appendix:

Table A1: Estimation results of linear and quadratic model specifications for subsample B.

Variable	$\ln(CO_2pc)$	$\ln(CO_2pc)$	$\ln(CO_2pc)$	$\ln(CO_2pc)$
1	-1.815** (0.845)	-3.958*** (0.849)	-6.937 (69.050)	81.357 (55.974)
t		-0.006*** (0.001)		-0.006*** (0.001)
$\ln(GDPpc)$	0.338*** (0.105)	0.615*** (0.107)	1.603 (17.051)	-20.489 (13.844)
$\ln(GDPpc)^2$			-0.078 (1.052)	1.305 (0.856)

Note: Standard errors are given in parentheses.

***, ** and * denote statistical significance at 1%, 5%, and 10% respectively.