Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)

Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

Nr. 180

A General Framework for Estimating Global CO_2 Emissions

Inmaculada Martínez-Zarzoso

February 2009

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173

e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

A GENERAL FRAMEWORK FOR ESTIMATING GLOBAL CO_2 EMISSIONS

Inmaculada Martínez-Zarzoso*

Department of Economics and Ibero America Institute for Economic Research, Universität Göttingen (Germany) and Instituto de Economía Internacional, Universitat Jaume I (Spain)

* Address for correspondence: Ibero America Institute for Economic Research, Universität Göttingen, 37077 Göttingen, Germany. E-mail address: martinei@eco.uji.es. Tel: 0049551399770. Financial support from Fundación Caja Castellón-Bancaja and the Ministry of Education is gratefully acknowledged (P1-1B2005-33 and SEJ 2007-67548).

Acknowledgments: I would like to thank Simon Dietz, Stephan Klasen, Ingo Geishecker and David Maddison, as well as the participants in the CESifo Summer School Program held in Venice in July 2008, for their helpful comments and discussion.

A GENERAL FRAMEWORK FOR ESTIMATING GLOBAL CO_2 EMISSIONS

Abstract

This paper proposes a new analytical framework with which to analyze the determinants of global CO_2 emissions. It contributes to the existing literature by examining the determinants of CO_2 emissions using a flexible functional form (transcendental logarithmic model), taking into account the presence of dynamic effects and allowing for heterogeneity in the sample of countries. The sample covers 121 countries and the period analyzed extends from 1975 through 2003. Two main results emerge. First, a static specification is rejected against a dynamic model. Second, the data also reject a general specification for all countries; hence slope-heterogeneity in the estimated coefficients has to be modeled. Conversely, the STIRPAT model is generally accepted for high-income countries, whereas for developing countries several interaction terms also play a role in explaining CO_2 emissions.

JEL classification: Q25, Q4, Q54

Keywords: CO_2 emissions, developing countries, panel data, population growth, urbanization

Abbreviations: IPAT, STIRPAT, EKC, WDI, GDP, IND, EI, FGLS, PPP

A GENERAL FRAMEWORK FOR ESTIMATING GLOBAL CO_2 EMISSIONS

1. Introduction

Climate change, with the attendant need to stabilize contributing global emissions, is one of the most challenging problems of our times and a matter of great concern among policy makers. Some aspects of the projected impact, such as global warming, increasing desertification, rising sea levels, and rising average temperatures, might have a disproportionate impact on developing countries, which least contributed to the cause of climate change.

While many factors have been adduced for climate change, energy consumption, as affluence grows, is singled out as having the most adverse impact on the environment. However, this impact becomes more severe when accompanied by demographic growth, given that population increases lead to increases in energy consumption and, consequently, to greater atmospheric pollution. A number of factors, namely, the increase in life expectancy, reduced child mortality, and improved farming methods, have resulted in rapid and exponential growth of the world population over the last 150 years. World population is currently growing by approximately 1.5 percent or by 78 to 80 million per year. According to the latest UN world population projections (2006 Revision), the world population will increase from the current 6.7 billion to 9.2 billion by 2050. Population growth is expected to be concentrated in the developing regions of the world, mainly Africa and Asia, while in the developed regions, growth will be very slow. In fact, the population of developed countries as a whole is expected to remain virtually unchanged and at about 1.2 billion between 2007 and 2050.

The main greenhouse gas in terms of quantity is CO_2, which, according to the Intergovernmental Panel on Climate Change (IPCC) (2007), accounted for about 76.7 percent of total anthropogenic greenhouse gas emissions in 2004. Although the reduction commitments of CO_2 emissions were seen as a task predominantly for developed countries

(United Nations Framework Convention on Climate Change [UNFCCC], 1997), based on the consensus that they are the largest contributors to global CO_2 emissions, there have been recent calls for developing countries to play an active role in global emissions reduction (Winkler, Spalding-Fecher, Mwakasonda & Davidson, 2002). The level of CO_2 emissions from developing countries has been rapidly exceeding that of the developed countries, and in 2003 accounted for almost 50 percent of the world's CO_2 emissions (Figure 1). This trend is expected to grow if the current path, in terms of energy consumption, is maintained. Since CO_2 is one of the main contributors to global emissions, it is of great interest to determine which policy measures will be most effective in curbing CO_2 emissions. However, given the abovementioned differences between developed and developing countries, those policy measures cannot be homogeneous and must be designed for specific country-groups.

In the last two decades, a number of researchers have investigated the determinants of CO_2 emissions within the framework of the Environmental Kuznets Curve (EKC) hypothesis without reaching conclusive evidence supporting the hypothesis (See Stern, 2004, for a survey). The focal point of this strand of the literature has been to determine whether or not the pollution-income relationship behaved as an inverted-U. Advances in the environment-development literature usually examined additional explanatory factors, such as structural change, trade or geography.

Other recent developments stem from studies using decomposition analysis and efficient-frontier methods, taking into account as explanatory variables not only affluence, but also energy use intensity, technical change, and structural change. Some of these variables are based on the IPAT[1] framework suggested by Erlich and Holdren (1971) and include population as an explanatory variable. However, in most cases changes in per-capita CO_2 emissions are explained with changes in income per capita, energy intensity, and structural change in the economy, assuming implicitly that population has a unitary elasticity with

[1] Impact=Population, Affluence, Technology (IPAT)

respect to emissions. Relatively little effort has been devoted to investigating the impact of demographic factors on the evolution of CO_2 emissions and most of the existing studies assume that this impact is comparable for all countries (e.g., MacKellar Lutz & Prinz, 1995; Dietz & Rosa, 1997; York, Rosa & Dietz, 2003; Cole & Newmayer, 2004). Two exceptions to this general assumption are the studies of Shi (2003), which grouped countries according to income levels, and Martínez-Zarzoso, Bengochea-Morancho & Morales-Lage (2007), which studied the impact of population growth for old and new European Union (EU) members.

This paper makes two primary contributions to the current state of the art. First, the paper contributes by refering to the functional form; this is the first paper exploring the role of non-linearities and interactions in a systematic way by estimating a transcendental logarithmic (translog) model in a dynamic framework. Second, it is the first paper to explore the role of urbanization in explaining CO_2 emissions for subgroup of countries.

We specify a model in which CO_2 emissions are related to the level of income per capita, the population size, the percent of urban versus rural population, the industrial structure and the energy efficiency of each country. The study involves four groups of countries classified by the World Bank as high, upper, middle, or low-income countries and analyzes the behavior of each group separately.

The results show important disparities among groups. First, whereas the effect of non-linearities and interaction terms is negligible for high-income countries, this is not the case for upper, middle, and low-income countries. An inverted U-shaped relationship is found with population for low-middle-income countries; several interaction terms play a role in explaining CO_2 emissions, as well. Second, urbanization shows a very heterogeneous impact on emissions. For low, lower-middle, and upper-middle-income countries, urbanization growth has a positive impact on CO_2 emissions, whereas in high-income countries, the impact is negative.

The paper is organized as follows: Section 2 briefly reviews the relevant literature, Section 3 presents the theoretical framework and specifies the model, Section 4 describes the empirical analysis and discusses the main results, and Section 5 concludes.

2. Literature Review

Erlich and Holdren (1971) suggested a suitable framework for analyzing the determinants of environmental impact known as the equation IPAT: *I=PAT,* where *I* represents environmental impact, *P* is the population size, *A* is the level of population affluence, and *T* denotes the level of environmentally damaging technology. The impact of human activity in the environment is viewed as the product of these three factors.

The IPAT model can be expressed as an identity where *A* could be defined as consumption per capita and *T* as pollution per unit of consumption. As stated by MacKellar et al. (1995), the IPAT identity is an approach which suggests that environmental impact is due to multiple, rather than to a single factor. However, these authors outline the limitations of testing this identity related to the choice of variables and the interactions between them. They compare households (*H*) with total population levels, as the demographic unit used to forecast future world CO_2 emissions, showing how each choice leads to different predictions in all the regions of the world, always increasing the impact on emissions for the *I=HAT* model, where the term *households*, replaces the term *population*.

The first studies which considered the IPAT framework to explain the sources of air pollution were based upon cross-sectional data for a sole time period. In this line of research, Cramer (1998, 2002) and Cramer and Cheney (2000) evaluated the effects of population growth on air pollution in California and found a positive relationship for some sources of emissions but not for others. Dietz and Rosa (1997) and York, Rosa, and Dietz (2003) studied the impact of population, affluence and other factors on cross-national carbon dioxide emissions and energy use within the framework of the IPAT model. These authors designated their model with the term, STIRPAT (Stochastic Impacts by Regression on Population, Affluence, and Technology). The results from these studies indicate that the elasticity of CO_2 emissions and energy use with respect to population size are close to unity. Affluence monotonically increases both CO_2 emissions and energy use. Finally, indicators of modernization (urbanization and industrialization) are associated with high impacts.

In a panel data context, Shi (2003) quantified the impacts of changes in population, income level, and energy efficiency of economic production on emissions. He found a direct relationship between population changes and carbon dioxide emissions in 93 countries between 1975 and 1996. He further determined that the impact of population on emissions varies with the levels of affluence and has been more pronounced in lower-income countries than in higher-income countries. Also using panel data, Cole and Neumayer (2004) considered 86 countries during the period from 1975 to 1998 and found a positive link between CO_2 emissions and a set of explanatory variables including population, urbanization rate, energy intensity, and smaller household size; however, the authors assumed that the effect of population and urbanization is equal for all income levels. Previous research also outlined the negative environmental impact caused by demographic pressure (Daily & Ehrlich, 1992; Zaba & Clarke, 1994), but they failed to analyze this impact within an appropriate quantitative framework.

In addition, several studies have discussed and tested the existence of an EKC where the relationship between pollution and income is considered to have an inverted U-shape. These models frequently take emissions per capita for different pollutants as an endogenous variable, assuming implicitly that the elasticity *emission-population* is unitary. A few of them considered population density as an additional explanatory variable (e.g., Cole, Rayner & Bates, 1997; Panayotou, Peterson & Sachs, 2000). However, their tests are not based upon an underlying theory, and the practice of testing variables individually is subject to the problem of omitted-variables bias. The results obtained within this framework are far from homogeneous and their validity has been questioned in recent surveys of the EKC literature (e.g., Stern, 1998 and 2004). Most of the criticisms are related to the use of nonappropriated techniques and to the presence of omitted-variables bias. In fact, Perman and Stern (2003) state that when diagnostic statistics and specification tests are taken into account and the proper techniques are used, the results indicate that the EKC does not exist. Borghesi and Vercelli (2003) consider that the studies based on local emissions present acceptable results, whereas those concerning global emissions do not offer the expected outcomes, and therefore the EKC hypothesis cannot be generally accepted. In addition, the existence of an inverted U-shape relationship between emissions and income contradicts the monotonicity in the income assumption underlying the IPAT model.

There are two approaches that go beyond the EKC literature. They are based on decomposition analysis and are known as *index number decompositions* and *efficient frontier methods*. The first approach requires detailed sectoral data and does not allow for stochasticity, whereas the second (frontier models) is based upon the estimation of econometric models, allows for random errors, and estimates factors common to all countries. Decomposition methods have been applied to an increasing number of pollutants in developed and developing countries (e.g., Hamilton & Turton, 2002; Bruvoll & Medin, 2003). Emissions are typically decomposed into scale, composition, and technique effects. Scale effects are measured with income and population variables, composition effects refer to changes in the input or output mix, and technique effects are proxied by energy intensity (the effect of productivity on emissions) and global technical progress. Hamilton and Turton (2002) concluded that income per capita and population growth are the two main factors increasing carbon emissions in OECD countries, whereas the decrease in energy intensity is the main factor reducing them. Bruvoll and Medin (2003) covered 10 pollutants and determined that in all cases, technique effects were dominant in offsetting the increase in scale. The authors concluded that, whereas structural change explains the increase in energy intensity from 1913 through 1970, technical change has been the main factor in reducing energy intensity after

1970. Shifts in the fuel mix is the main factor explaining carbon emissions per unit of energy used.

Efficient frontier methods are closely related to the abovementioned STIRPAT models and have also been extensively applied to different pollutants and countries. In this line, Lantz and Feng (2006) modeled carbon dioxide (CO_2) emissions in Canada over the period from 1970 through 2000 as a function of income per capita, population, and technological change. This study tests a more flexible model introducing squared terms for all the explanatory variables. Findings indicate that income per capita is unrelated to CO_2 emissions in Canada, that an inverted U-shaped relationship exists to population, and that a U-shaped relationship exists to technology.

In this paper a step forward is made in the same direction as in Lantz and Feng (2006). In addition to squared terms for all explanatory variables, interaction terms are added to the model under the framework of the translog model. In any empirical model, interaction terms and squared terms may enter as proxies for each other unless all terms are initially included in the specification. It is therefore very important to test whether either of these two possibilities has empirical support.

Finally, it would also be desirable to model and estimate the emissions process as a dynamic production process. Although Agras and Chapman (1999) previously illustrated the importance of modeling dynamics in EKC analyses, this issue has been ignored in most of the subsequent literature[2]. In a simple dynamic framework, current emissions depend upon lagged emissions. This assumption will be tested and incorporated into the model specification.

3. Basic Framework of Analysis

Building upon Ehrlich and Holdren's (1971) basic foundation, Dietz and Rosa (1997) formulated a stochastic version of the IPAT (STIRPAT) equation with quantitative variables

[2] An exception is Auffhammer and Carson (2008) which used a dynamic model to forecast China's CO_2 emissions using province-level information.

containing population size (*P*), affluence per capita (*A*), and the weight of industry in economic activity as a proxy for the level of environmentally damaging technology (*T*). The specification of the STIRPAT model is given by the following equation,

$$I_i = \lambda_0 P_i^{\lambda_1} A_i^{\lambda_2} T_i^{\lambda_3} \varepsilon_i \qquad (1)$$

where I_i, P_i, A_i, and T_i are the variables defined above, λ_k are parameters to be estimated, and ε_i is the random error.

This paper proposes a generalization of the STIRPAT model as the reference theoretical and analytical framework with which to analyze the income-emissions relationship. In most cases researchers follow a model selection strategy which begins with a simple specification and seeks to refine it by adding variables. Here, the opposite approach is suggested, starting with a general specification and seeking to refine it by imposing the appropriate restrictions.

One of the most popular flexible functional forms, widely used in modern studies of demand and production, is the translog model which allows one to model interactions and second-order effects. This model can be interpreted as a second-order approximation to an unknown functional form and was introduced formally in a series of papers by Christensen, Jorgenson, and Lau in the early 1970s[3].

To derive the translog model from Equation 1, we first write $I_i=f(P_i, A_i, T_i)$. Then, $\ln I_i=\ln f(P_i, A_i, T_i)=g(P_i, A_i, T_i)$. Since by a trivial transformation, $P_i=\exp(\ln P_i)$, $A_i= \exp(\ln A_i)$, and $T_i=\exp(\ln T)$, the function can be interpreted as a function of the logarithms of the variables. Hence, $\ln I_i= g(P_i, A_i, T_i)$. In order to simplify the equation, we assume that X_k are the

---

[3] L. R. Christensen, D. W. Jorgenson, and L. J. Lau, "Conjugate Duality and the Transcendental Logarithmic Production Function," *Econometrica* 39*,* July 1971, 255-56, "Transcendental Logarithmic Production Frontiers" *Rev. Econ. Statist*. 55, Feb. 1973, 28-45. “Transcendental Logarithmic Utility Functions,” *The American Economic Review* 65(3), Jun. 1975, 367-383.

explanatory variables P_i, A_i, and T_i. Now ln I_i is expanded in a second-order Taylor series around unity; thus the logarithm of each variable is zero:

$$\ln I_i = g(0) + \sum_{k=1}^{K} \frac{\partial \ln I_i}{\partial \ln X_{ik}} \ln X_{ik} + \frac{1}{2} \sum_{k=1}^{K} \sum_{l=1}^{K} \frac{\partial^2 \ln I_i}{\partial \ln X_{ik} \partial \ln X_{il}} \ln X_{ik} \ln X_{il} + \mu_i \quad (2)$$

where $X_{ik} = P_i, A_i, T_i.$

Since the function and its derivatives evaluated at the fixed value of zero are constants, these derivatives are interpreted as coefficients. By imposing symmetry on the cross-term derivatives, the model becomes

$$\ln I_i = \alpha + \sum_{k=1}^{K} \beta_k \ln X_{ik} + \frac{1}{2} \sum_{k=1}^{K} \sum_{l=1}^{K} \gamma_{kl} \ln X_{ik} \ln X_{il} + \mu_i \quad (3)$$

The log-linear model is a special case of this formulation in which γ_{kl}=0. Hence, Equation 3 is a more general formulation of the income-emissions relationship, for which the STIRPAT model, as well as the EKC relationship[4], can be derived as special cases.

Finally, the traditional model is extended with two additional explanatory variables, X_{ik}: urbanization and industrial activity. The first could be considered as an indicator of the spatial distribution of population within countries. On the one hand, poor rural countries have fewer chances to pollute than advanced, industrialized countries. On the other hand, highly urban, developed countries may require less personal transport since public transport is available. The second variable, industrial activity, could also be considered as a proxy for structural change in the economy.

4. Econometric estimation

[4] The EKC model is a special case of this formulation in which γ_{kk}=0.

4.1 Data sources, model specification, and main results

Equation 3 is estimated for a sample of 121 countries[5] during the period from 1975 to 2003. The countries under analysis are classified into four income groups according to data from the World Development Indicators (WDI), 2007. Low-income economies are those in which 2005 GNI per capita was $875 US or less (54 countries). Lower-middle-income economies are those in which 2005 GNI per capita was between $876 and $3,466 (58 countries). Upper-middle-income economies are those in which 2005 GNI per capita was between $3,466 and $10,725 (40 countries). Finally, high income countries are those in which 2005 GNI per capita was $10,726 or more (36 countries). The sample of countries is considerably reduced when energy efficiency is included as an explanatory variable since data for this variable are not available for many developing countries[6]. There are also some countries for which income data are missing and transition economies only report data since the early 1990s, when their economies began the opening-up process. Countries considered in each group are listed in Table A.2 in the Appendix (WDI, World Bank, 2007). A summary of the data, as well as the simple correlation coefficients between the variables in the model, is shown in Table A.1 in the Appendix.

In order to test whether the evolution of the factors considered in the TRALIPAT (Translog Impacts by Regression on Population, Affluence, and Technology) model influences the level of CO_2 emissions through time and across countries, a dynamic version of the translog model described in the previous section (Equation 3) is specified as follows:

[5] The countries are listed in the Appendix. Data are available for 24 low-income, 39 lower-middle-income, 25 upper-middle-income, and 33 high-income economies.
[6] Energy Efficiency data are available for 31 low-income countries, 38 lower-middle-income countries, 26 upper-middle-income countries, and 35 high-income countries.

$$\ln CO2_{it} = \alpha_i + \phi_t + \pi_i \ln CO2_{i,t-1} + \beta_1 \ln P_{it} + \beta_2 \ln YH_{it} + \beta_3 \ln EI_{it} + \beta_4 \ln PUPC_{it} + \beta_5 \ln IA_{it} + \gamma_{11}\left(\frac{1}{2}\ln^2 P_{it}\right) + \gamma_{12}(\ln P_{it} * \ln YH_{it}) +$$
$$+ \gamma_{13}(\ln P_{it} * \ln EI_{it}) + \gamma_{14}(\ln P_{it} * \ln PUPC_{it}) + \gamma_{15}\ln(P_{it} * \ln IA_{it}) + \gamma_{22}\left(\frac{1}{2}\ln^2 YH_{it}\right) + \gamma_{23}(\ln YH_{it} * \ln EI_{it}) + \gamma_{24}(\ln YH_{it} * \ln PUPC_{it}) +$$
$$+ \gamma_{25}(\ln YH_{it} * \ln IA_{it}) + \gamma_{33}\left(\frac{1}{2}\ln^2 EI_{it}\right) + \gamma_{34}(\ln EI_{it} * \ln PUPC_{it}) + \gamma_{35}(\ln EI_{it} * \ln IA_{it}) + \gamma_{44}\left(\frac{1}{2}\ln^2 PUPC_{it}\right) + \gamma_{45}(\ln PUPC_{it} * \ln IA_{it}) +$$
$$+ \gamma_{55}\left(\frac{1}{2}\ln^2 IA_{it}\right) + \mu_{it} \quad (4)$$

where the sub-index i refers to countries and t refers to the different years. $CO2_{it}$ is the amount of CO_2 emissions in tons, P_{it} denotes population, YH_{it} is the Gross Domestic Product (GDP) per capita expressed in constant PPP (purchasing parity prices) ($2000 US), EI_{it} is proxied with energy efficiency (EI) measured as GDP at constant PPP prices divided by energy use, where *energy use* refers to apparent consumption (production+imports-exports), $PUPC_{it}$ denotes urbanization rates, IA_{it} is the percentage of the industrial activity with respect to the total production measured by the GDP. Finally, α_i and ϕ_t capture the country and time effects, respectively, and μ_{it} is the error term. Since the model is specified in natural logarithms, the coefficients of the explanatory variables can be directly interpreted as elasticities. The time effects, ϕ_t, can be considered as a proxy for all the variables that are common across countries but which vary over time. Within the context of decomposition analysis, these effects are sometimes interpreted as the effects of emissions-specific technical progress over time (Stern, 2002) and can also be interpreted as a proxy for energy prices (Agras and Chapman, 1999).

Equation 4 was first estimated for the whole set of countries under analysis. Table 1 shows estimation results from a set of baseline models, as well as the best models according to goodness of fit and model selection criteria (rmse).

Table 1. *Selected Estimation Results for All Countries in the Sample (1975-2003)*

All the models are estimated including country and time fixed-effects, since these effects are statistically significant (as indicated by the respective LM tests). The result of the Hausman test indicates that the country effects are correlated with the residuals and therefore only the fixed-effects estimates are consistent. Since the time dimension of the panel is relatively large (29 years), serial correlation is almost certainly present in our data. This hypothesis is confirmed by performing the Wooldridge autocorrelation test for panel data. In order to get consistent estimates, feasible generalized least squares (FGLS) techniques can be used to estimate the first-order autocorrelation coefficient, which is represented by rho. Models 1 to 6 are adjusted for first order autocorrelation.

Models 1 and 2 are static baseline models. Model 1 estimates a typical EKC, with emissions per capita as the dependent variable. The model presents a very low R^2; autocorrelation in the residuals and the implicit turning point for income per capita is out of sample. In Models 2 to 7 the dependent variable is total emissions instead of emissions per capita, given that population is now added as an explanatory variable in the models in order to test for the hypothesis of unitary emissions elasticity with respect to population, which was implicitly assumed in Model 1. The estimated coefficient for population in Model 2 is 1.68 and a Wald test rejects the null of unitary emissions elasticity with respect to population. Model 3 augments the second model by adding a pooled lag of emissions. The fit of the equation improves considerably (the R^2 almost quadruple) and the estimated AR(1) denoted as rho_ar at the end of Table 1, is considerably reduced. The Baltagi-Wu LBI statistic and the Bhargaba et al. Durbin-Watson statistic both accept the null hypothesis of no first-order autocorrelation.

Models 4 and 5 list estimation results from the unrestricted and restricted translog models, respectively. Model 4 shows the estimates obtained from the specification given in Equation 4. The estimates indicate that there are non-linearities for income per capita and energy efficiency, since the coefficients of the corresponding squared terms are statistically significant at five-percent level. Although an inverse U-shape relationship is found between

CO_2 emissions and per capita income, the turning point is, as before, out of sample ($268,337 at constant PPP 2000). With respect to energy intensity, the squared term is also negative, reinforcing the effect of the variable in levels. With respect to the interaction terms, four out of ten are statistically significant at the one or five-percent level. To evaluate the presence of non-linearities and interaction terms as explanatory variables of CO_2 emissions, two hypotheses are tested. First, we test for the joint significance of the squared terms and, second, we test for the joint significance of the interaction terms. Whereas the result of the first test indicates that the null of zero coefficients on the squared terms can be marginally accepted (p-value=0.03), the null of zero coefficients on the interaction terms is strongly rejected (p-value=0.00). In both cases we use a Wald test.

Since adding a lagged dependent variable to the list of explanatory variables generates some estimation complications, we use alternative techniques to control for endogeneity of the lagged dependent variable and to correct the bias which can affect the estimated coefficients.

Models 6 and 7 report the results obtained by using instrumental variables techniques and by using the bias-corrected least-squares dummy variable estimator (BCLSDV), respectively. Model 6 is estimated by using the generalized method of moments (GMM) with a restricted set of instrument and with fixed effects, the estimates are robust to autocorrelation and heteroskedasticity (Schaffer, 2007). Model 7 is estimated using the BCLSDV estimators for the standard autoregressive panel-data model using the bias approximations in Bruno (2005), who extends the results by Bun and Kiviet (2003), Kiviet (1999), and Kiviet (1995) to unbalanced panels. Kiviet and Bun (2001) suggest a parametric bootstrap procedure to estimating the asymptotic variance-covariance matrix of the BCLSDVC, which is superior to the analytical expression and can be applied to any version of least-squares dummy variable estimator (LSDVC). The estimated coefficients in Column 7 are similar to those obtained in Columns 5 and 6. The main difference is that the squared terms for income and energy intensity are not any more statistically significant in Column 7. However, these results have to be taken with caution since the standard errors reported for the BCLSDV estimators are based on a bootstrap variance-covariance matrix for LSDVC using 50 repetitions.

With respect to the estimated short-run elasticities evaluated at the mean, according to Model 7 the population elasticity is 0.29 (0.693-0.042*3.40-0.029*8.92) and decreases with the level of income and industrial activity, since the corresponding interaction coefficients are negative. The percentage of urban population also has a positive effect on CO_2 emissions for average levels of income and energy efficiency; the estimated elasticity for urbanization is 0.84 in the short run, but this effect decreases with the level of income and increases with energy efficiency. The estimated elasticity for income per capita (income per capita squared is not significant in Model 7) is 1.5 in the short-run, and an increase in energy efficiency leaves emissions almost unchanged. Finally, the effect of the percentage of industrial activity is positive (0.35), with an almost unitary elasticity in the long-run. The time effects are only significant and show a decreasing trend in the early 1980s and early 1990s, reflecting perhaps the effects of the business cycle. In some cases the null hypothesis of zero time effects were marginally accepted (Models 3, 4, and 5).

The negative coefficients for the interaction terms *income-urbanization* and *income-population* suggest that the marginal effect of urbanization (population) on emissions diminishes as income per capita goes up. For example, for a country with GDP per capita equal to $1,000, a one-percent increase in population raises CO_2 emissions by 0.35 (0.693-0.029*6.908-0.042*3.40), while for a country with GDP per capita equal to $20,000, a one-percent rise in population increases emissions by 0.26 (0.693-0.029*9.903-0.042*3.40). Similarly, for a country with GDP per capita equal to $1,000, a one-percent increase in urbanization increases CO_2 emissions by 0.040 (0.948-0.171*6.908+0.17*1.61), whereas for a country with GDP per capita equal to $20,000, a one-percent increase in urbanization reduces CO_2 emissions by 0.47 (0.948-0.171*9.903+0.17*1.61).

The other two interaction terms that are significant are *population*industrial activity* and *urbanization*energy intensity*. The negative coefficient obtained for the first term suggests that the marginal effect of population on emissions also diminishes with higher levels of industrialization, whereas the positive coefficient obtained for the second term indicates that the marginal effect of urbanization on emissions increases with higher levels of energy intensity.

The significance of the interaction terms indicates that the effect of the factors explaining CO_2 emissions is heterogeneous across groups of countries. Therefore, the sample is divided into different groups of countries according to the income level according to the World Bank classification. Countries are grouped as low, lower-middle, upper-middle, and high income. High income countries are further divided into two groups (OECD and non-OECD countries); transition countries are considered separately. Table 2 shows the results for the preferred specifications. In most cases the best estimation method, in terms of explanatory power and forecasting accuracy, was the GMM-dynamic-fixed-effects specification corrected for autocorrelation and heteroskedasticity. The LSDVC estimates and the model estimated with fixed effects and corrected for autocorrelation showed similar results that are available upon request.

Important differences between the six sets of results are observed. The first one concerns the model specifications, which differ markedly across groups. Applying the general-to-specific methodology, the existence of non-linearities (squared terms) is always rejected by the tests with only one exception: an inverted U-shaped relationship between CO_2 emissions and population for middle-low-income countries. The Kuznets Curve hypothesis (emissions-income inverted-U relationship) is rejected for all country-groups.

Table 2. *Estimation Results for Sub-Groups of Countries* (1975-2003)

A second source of differences concerns the interaction terms. Only some of them are statistically significant for low-middle, middle-low, and upper-middle income countries but none of them is significant for high income countries (OECD and non-OECD sub-groups). As a result, the restricted model is greatly simplified for high income countries; it reduces to a dynamic-extended version of the STIRPAT model, adding urbanization as a regressor. In addition, the variable *industrial activity* is only significant for low-middle, middle-low, and upper-middle income countries and it changes with the level of income and urbanization.

A third source of differences between the six sets of results concerns urbanization. The elasticity emissions-urbanization is negative and significant for the OECD and non-OECD high-income groups, whereas for upper-middle, lower-middle, and low-income countries, it is positive and significant. The emissions-urbanization coefficient is higher than unity (1.51) for low-income countries but the effect decreases for higher levels of industrial activity. It is even higher for middle-low-income countries (1.88) and increases with energy efficiency. It is almost one for upper-middle-income countries (0.98) but decreases with higher levels of energy efficiency and industrial activity. The inclusion of urbanization in the model does not change the estimated coefficients of the other explanatory variables. The model was also estimated without this variable and the only difference was that the Log-Likelihood was lower in magnitude.

Concerning population, the population coefficient is not significant for low income countries or for upper-middle-income countries, whereas for lower-middle-income countries, an inverted U-shaped relationship is found (turning point=89 million inhabitants). Figure A.3 in the appendix shows the inverted-U curve. The elasticity for OECD and high-income non-OECD countries is lower than one and only statistically significant for the former.

Some differences have also been observed in the other explanatory variables. An increase of 1 percent in the GDP per head causes a 1.83 (3.044-0.234*3.24-0.144*3.11) percent increase in CO_2 emissions of low-middle-income countries and a 0.90 percent increase in CO_2 emissions of low-income countries. The negative contribution of energy efficiency to emissions is also different: for low-middle-income countries, the impact is also higher than for low-income (the elasticities are -1.24 and -0.31, respectively). To summarize, the environmental impact caused by population, urbanization, and affluence variables (income effect) seems to be higher in low-income and middle-low-income countries than in others, whereas the contribution of energy efficiency is more similar for all countries.

4.2 Robustness checks

As a first robustness check, we tested for cross-sectional dependence. This dependence may arise due to the presence of common shocks and unobserved components that become part of the error term when they are not modeled. We use the tests proposed by Pesaran (2004) and Frees (1995, 2004). Pesaran's cross-sectional-dependence test (CD test) is valid under a wide range of panel data models, including dynamic panels and unbalanced panels. However, since it involves the sum of the pair-wise correlation coefficients of the residual matrix, rather than the squared correlations used in the classical LM test, it can fail to reflect cases of cross-sectional dependence where the sign of the correlations alternates (positive and negative correlations cancel out when averaging). Frees's proposed test is based on the sum of the squared correlation and is therefore not subject to this drawback. But, unlike Pesaran's CD test, it was devised for static panels and the asymptotic properties have been derived only for static panels.

The tests were applied to the sub-groups of countries considered in Table 2 and, although the Pesaran test indicated that the null of cross-sectional independence cannot be rejected, the Frees test indicates that the null hypothesis is always rejected at conventional significance levels. Table 3 shows the results of the tests. Given that the results of the tests are contradictory, to account for the possible existence of cross-sectional dependence, the model was re-estimated with the Common Correlated Effects (CCE) model proposed by Pesaran (2006).

Table 3. *Tests for Cross-Sectional Dependence*

This method consist of approximating the linear combinations of the unobserved factors by cross-section averages of the dependent and explanatory variables and then running standard panel regressions augmented by the cross-section averages. This approach also yields consistent estimates when the regressors are correlated with the factors. The results, reported in Table 4, indicate that although some of the added regressors are statistically significant in terms of goodness of fit (R^2) and forecasting performance (root mean squared error), estimates presented in Table 2 show a better performance. In fact, the average value of the off-diagonal elements of the variance-covariance matrix of the estimated residuals was around 0.25 for all subgroups, indicating that cross-dependencies are not severe.

Table 4. *Estimation Results for Group of Countries with Cross-Sectional Dependency*

As a second robustness check, we evaluated the forecasting accuracy of the different specifications using two criteria: the rmse, and the Stavins and Jaffe goodness of fit (SJ). Stavins and Jaffe (1990) proposed a goodness-of-fit statistic equal to one minus Theil's U-statistic based on comparing predicted and actual values for the dependent variable (S&J goodness of fit). The Theil inequality coefficient lies between 0 and 1 and a value of zero indicates a perfect fit. This measure is also scale invariant. We can compare models in Table 1 using both statistics. The rmse and S&J (1990) goodness-of-fit values are shown at the bottom of Table 1. Model 6 shows the lowest rmse, followed by Model 5, whereas Model 7 presents the best S&J goodness of fit with an SJ equal to 0.9864, followed by Model 6 with an SJ equal to 0.9774.

Finally, we analyzed graphically the regression results in order to identify outliers. Namibia and Cameroon were identified as outliers and therefore, the model was re-estimated excluding both countries. The primary results were practically unchanged.

5. Conclusions

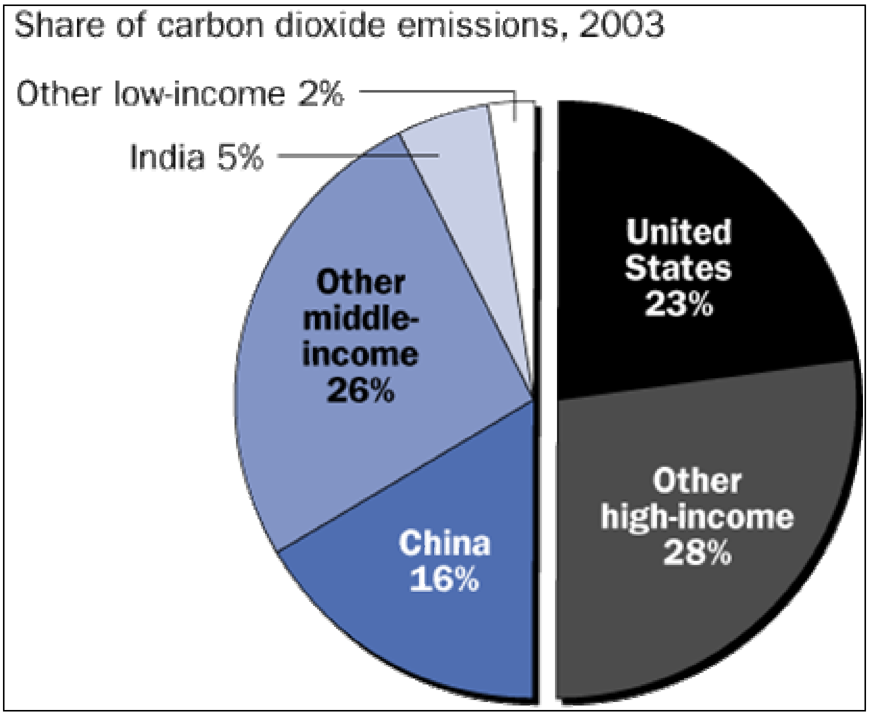
In this paper a multivariate analysis of the determinants of global carbon dioxide emissions during the period of 1975 to 2003 has been conducted. A generalization of the STIRPAT model as the reference theoretical and analytical framework is proposed. In the proposed model, population is introduced as a predictor, together with affluence per capita, and the level of environmentally damaging technology, proxied with the weight of the industrial sector in the GDP and with energy intensity. We have added urbanization and industrial activity as predictors and used several model specifications and estimation methods in a panel data framework. This paper is the first to explore the role of non-linearities and interactions in a systematic way by estimating a dynamic translog model to investigate the relationship between CO_2 emissions and income.

The empirical discussion suggests several general conclusions. It appears that the commonly hypothesized Kuznets Curve is in general rejected by the model, once country-heterogeneity is considered, except for the inverted U-shaped relationship found for population when the model was estimated for low-middle-income countries.

On the other hand, the STIRPAT model is generally accepted for high-income countries, whereas for developing countries, several interaction terms also play a role in explaining CO_2 emissions. This is largely due to the fact that the effects of technological and structural change on emissions vary with the level of income and urbanization, even for subgroups of countries (low, middle-low, and upper-middle-income countries). In fact, concerning urbanization, the results show very different patterns for low, lower-middle and upper-middle-income countries and the rest. For the first three sets of countries, the elasticity emission-urbanization is positive, whereas for OECD countries, high-income non-OECD countries, and transition economies, the elasticity is negative, although non-significant for the latter. This result has a very important policy implication: Once urbanization reaches a certain level, its effect on emissions becomes negative, contributing to reduced environmental damage. In 2008 more than half of the world's human population (3.3 billion people) is living in urban areas. By 2030, this is expected to increase to almost 5 billion. Although many of these cities will be poor, no country in the industrial age has ever achieved significant economic growth without urbanization. Cities may concentrate poverty, but they also represent the best hope of escaping it. Despite the fact that cities cause considerable environmental damage, namely by increasing emissions due to transportation, energy consumption, and other factors, policymakers and experts increasingly recognize the potential value of cities to long-term sustainability. It could be that these potential benefits of urbanization outweigh the disadvantages.

We leave for further research the application of the estimation framework proposed herein with respect to other pollutants. It would also be desirable to take dynamics into account to forecast future emissions.

Figure 1. Carbon dioxide emissions in 2003



Source: World Development Indicators 2007

Table 1

Selected Estimation Results for All Countries in the Sample (1975-2003)

MODELS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LYH	1.783***	1.563***	0.681***	2.310***	2.308***	2.357***	2.237***
	4.437	3.907	5.554	6.179	8.100	4.341	3.391
LYH^2	-0.073**	-0.058*	-0.026***	-0.103**	-0.092***	-0.087**	-0.081
	-3.022	-2.393	-3.57	-2.814	-3.630	-2.652	-1.398
LP		1.678***	0.408***	1.137***	0.783***	0.794***	0.693***
		9.510	10.362	3.587	7.504	3.589	3.300
LCO2(-1)			0.759***	0.670***	0.691***	0.623***	0.720***
			70.822	53.037	57.422	6.565	42.499
(1/2)LP^2				-0.018			
				-0.887			
LPUPC				1.434**	0.791**	0.998**	0.948*
				2.669	3.133	2.448	1.672
(1/2)LPUPC^2				0.001			
				0.012			
LEI				-0.566	-0.639***	-0.842**	-0.679**
				-1.813	-4.223	-2.634	-2.321
(1/2)LEI^2				-0.099*	-0.109**	-0.130**	-0.073
				-2.180	-2.966	-3.110	-1.313
LIA				0.942**	0.803***	0.776**	0.827**
				2.623	3.508	2.636	2.479
(1/2)LIA^2				0.134*	0.050		
				2.25	0.089		
LYH*LPUPC				-0.132*	-0.139***	-0.171**	-0.158*
				-2.284	-3.898	-2.832	-1.940
LYH*LP				-0.023	-0.037***	-0.029***	-0.035
				-1.740	-4.472	-3.558	-1.539
LYH*LIA				-0.027			
				-0.932			
LYH*LEI				-0.017			
				-0.607			
LP*LPUPC				-0.029			
				-1.016			
LP*LIA				-0.056***	-0.045**	-0.042*	-0.046**
				-3.369	-3.137	-2.372	-2.164
LP*LEI				-0.012			
				-0.858			
LPUPC*LIA				-0.063			
				-1.044			
LPUPC*LEI				0.143**	0.135***	0.173**	0.136*
				2.763	3.594	2.611	1.848
LIA*LEI				0.053			
				1.407			
R-squared	0.022	0.149	0.835	0.822	0.832	0.82	
N	3067	3067	2975	2639	2639	2532	2760
Ll	1578.106	1587.610	1566.202	1467.415	1462.876	1499.638	
Rmse	0.145	0.145	0.144	0.140	0.140	0.138	0.141
Rho_ar	0.839	0.830	-0.024	-0.040	-0.007		
d1	0.331	0.350	2.035	2.004	2.011	J=2.383	
LBI	0.492	0.509	2.097	2.078	2.087	J(p)=0.304	
(1-U-Theil)	-1.6210	0.696	0.9701	0.9738	0.9785	0.9774	0.9864

*Note: LYH denotes per-capita income, LP denotes population, LPUPC is the percentage of urban population over total population, LEI is energy efficiency, and LIA is the percentage of industrial activity over total GDP. All the variables are in natural logarithms. t-statistics reported. *, **, and *** denote significance at the 10, 5, and 1% level, respectively.*

Table 2

Estimation Results for Group of Countries (GMM-Dynamic-FE Model)

	Restricted Translog Function					
	Low Income	Mid-low	Up-mid	OECD	High-noOECD	Transition
LCO2(-1)	0.634***	0.499***	0.720***	0.672***	0.658***	0.047
	12.901	3.636	14.739	20.446	7.659	0.596
LP	0.533	2.618***	-0.093	0.593***	0.116	1.365***
	1.676	3.666	-0.658	6.090	1.028	3.632
LPUPC	1.506***	1.885**	0.980**	-0.186*	-0.693*	-0.190
	4.600	2.458	2.298	-1.707	-2.233	-0.067
LYH	-0.300	3.044***	0.158*	0.318***	0.057	3.375**
	-0.959	4.237	1.700	6.470	0.773	2.669
LEI	-0.305***	-0.645	-0.990*	-0.307***	-0.047	-2.060**
	-3.547	-1.216	-1.729	-7.421	-0.854	-2.884
LIA	-0.395	1.150*	0.738*	0.069		-0.310
	-0.680	1.690	1.907	1.536		-0.205
LY*LPUPC		-0.414**				-0.303
		-2.733				-1.056
LPUPC*LEI		0.497***	-0.155*			
		3.055	-1.983			
LYH*LIA	0.290**	-0.144*				-0.244
	2.940	-1.812				-1.263
LPUPC*LIA	-0.480***	0.316	-0.154			0.467
	-4.398	1.546	-1.542			0.989
(1/2)LP^2		-0.143***				
		-3.946				
LY*LEI		-0.234***	0.143**			-0.189
		-4.150	2.258			-1.523
LIA*LEI						0.722*
						2.435
R-Squared	0.816	0.796	0.904	0.898	0.883	0.944
N	501	776	491	586	231	197
Ll	232.59	297.66	469.99	910.25	193.89	274.09
Rmse	0.16	0.17	0.10	0.05	0.11	0.07
J	6.68	0.41	6.16	3.76	3.90	0.58
Jp	0.08	0.94	0.10	0.29	0.27	0.90

*Note: t-statistics reported. *, **, and *** denote significance at the 10, 5, and 1% level, respectively. LYH denotes per-capita income, LP denotes population, LPUPC is the percentage of urban population over total population, LEI is energy efficiency, and LIA is the percentage of industrial activity over total GDP. All the variables are in natural logarithms.*

Table 3

Tests for Cross-Sectional Dependence

Cross sectional independence	**Low income**	**Mid-low**	**Up-mid**	**OECD**	**Transition**
Pesaran's test	-1.799	5.459	-1.223	-1.991	-2.44
Probability	*1.928*	*0.000*	*1.779*	*1.954*	*1.985*
Frees's test	0.996	4.851	0.521	0.950	0.678
Critical values (alpha=0.05) (from Frees's Q distribution)	*0.433*	*0.568*	*0.343*	*0.284*	*0.284*

Note: Both tests are described in De Hoyos and Sarafidis (2006).

Table 4

Estimation Results for Group of Countries with Cross-Sectional Dependency

Common Correlated Effects model						
	Low income	Mid-low	Up-mid	OECD	High-noOECD	Transition
LCO2(-1)	0.707***	0.451**	0.751***	0.710***	0.422**	0.057
	12.215	2.534	12.655	17.478	2.865	0.482
AVRLCO2	1.238**	0.501***	0.453*	0.707**	0.849**	0.178
	3.176	3.945	1.807	3.057	2.451	0.747
AVRLYH	-1.031	0.100	-0.501	-0.975**	0.489	-0.256
	-1.692	0.261	-1.515	-2.946	0.991	-1.376
AVRLP	5.837*	8.904**	-1.838	-7.326***	-0.617	-4.141
	1.951	2.046	-0.74	-3.867	-0.403	-1.481
AVRLPUPC	-13.463**	-17.590*	0.154	-1.591	3.815	6.946
	-2.761	-1.974	0.054	-0.741	1.096	1.206
AVRLEI	0.283	-0.344	-0.036	0.007	0.303	-0.210
	0.83	-1.574	-0.229	0.023	1.428	-0.949
AVRLIA	0.36	0.283	0.157	0.609**	-1.030**	-0.193
	1.043	1.180	0.603	2.484	-2.478	-0.624
LP	0.524	2.937**	-0.092	0.517***	0.009	1.605***
	1.689	3.238	-0.661	4.899	0.059	4.488
LPUPC	1.253***	1.736*	0.12	-0.221	-1.045**	-0.803**
	3.667	2.200	1.506	-1.918	-3.091	-3.291
LYH	-0.341	2.815***	0.190*	0.306***	-0.011	1.055***
	-1.081	4.000	2.179	6.058	-0.133	7.163
LEI	-0.208*	-0.453	-1.150*	-0.280***	-0.008	-2.176***
	-2.217	-0.814	-1.986	-6.437	-0.120	-5.043
LIA	-0.443	1.298**	0.105	0.041	0.126**	0.008
	-0.751	2.012	1.399	0.842	2.316	0.114
LY*LIA	0.258**	-0.142*				
	2.545	-1.824				
LPUPC*LIA	-0.399***					
	-3.551					
(1/2)LP^2	-0.154***	-0.154***				
	-3.646	-3.646				
LY*LPUPC	-0.340**	-0.340**				
	-2.513	-2.513				
LY*LEI	-0.245***	0.095	0.095			
	-4.180	1.590	1.590			
LPUPC*LEI	0.459**					
	2.88					
LIA*LEI	0.396***					0.396***
	3.71					3.710
R-Squared	0.815	0.798	0.905	0.896	0.856	0.942
N	501	776	491	586	231	197
Ll	222.447	291.020	463.678	893.767	141.141	257.568
Rmse	0.162	0.173	0.098	0.054	0.117	0.070
J	0.169	0.326	6.056	0.236	1.396	0.511
Jp	0.919	0.849	0.048	0.889	0.498	0.775

*Note: t-statistics reported. *, **, and *** denote significance at the 10, 5, and 1% level, respectively. LYH denotes per-capita income, LP denotes population, LPUPC is the percentage of urban population over total population, LEI is energy efficiency, and LIA is the percentage of industrial activity over total GDP. All the variables are in natural logarithms. AVR denotes averages of the variables across countries.*

References

Agras J and Chapman D (1999) A dynamic approach to the environmental Kuznets curve hypothesis. Ecol Economics 28: 266-277.

Auffhammer, M. and Carson R.T. (2008) Forecasting The Path of China's CO2 Emissions Using Province Level Information. Journal of Environmental Economics and Management. 47(1): 47-62.

Borghesi S, Vercelli A (2003) Sustainable globalisation. Ecol Economics 44:77-89.

Bruvoll A, Medin H (2003) Factors behind the Environmental Kuznets Curve: a decomposition of the changes in air pollution. Env and Res Economics 24(1): 27-48.

Bruno, G. S. F. (2005) Approximating the Bias of the LSDV Estimator for Dynamic Unbalanced Panel Data Models. *Economics Letters,* 87: 361-366.

Bun, M. J. G. and J. F. Kiviet (2003) On the diminishing returns of higher order terms in asymptotic expansions of bias. *Economics Letters* 79: 145–152.

Cole MA, Rayner AJ, Bates J M (1997) The Environmental Kuznets Curve: an empirical analysis. Env and Dev Economics 2(4):401-416.

Cole MA, Neumayer E (2004) Examining the impact of demographic factors on air pollution. Pop and Dev Review 2(1):5-21.

Cramer C J (2002) Population growth and local air pollution: methods, models and results. In Lutz W, Prkawetz A, Sanderson WC (eds.) Population and Environment. Supplement to Pop and Dev Review, vol. 28: Population Council, New York, 53-88.

Cramer CJ (1998) Population growth and air quality in California. Demography 35(1): 45-56

Cramer CJ, Cheney RP (2000) Lost in the ozone: population growth and ozone in California. Pop and Env 21(3):315-337.

Daily GC, Ehrlich PR (1992) Population, sustainability and earth's carrying capacity. Biosciences 42:761-771.

De Hoyos, R.E. and Sarafidis, V (2006) Testing for cross–sectional dependence in panel–data models Stata Journal Volume 6 Number 4.

Dietz T, Rosa EA (1997) Effects of population and affluence on CO_2 emissions. In: Proceedings of the Natl Acad of Sciences USA 94(1): 175-179.

Ehrlich PR, Holdren JP (1971) Impact of population growth. Science 171:1212-1217

Friedman.

Frees, Edward W., 1995. "Assessing cross-sectional correlation in panel data," Journal of Econometrics, Elsevier, vol. 69(2), pages 393-414.

Frees, E. (2004). Longitudinal and panel data. Cambridge: Cambridge University Press.

Galeotti M, Lanza A (1999) Richer and cleaner? A study on carbon dioxide emissions in developing countries. FEEM Working Paper 87.

Hamilton C, Turton H (2002) Determinants of emissions growth in OECD countries. En Pol 30:63-71.

Intergovernmental Panel on Climate Change, Working Group III Report: Mitigation of Climate Change (2007), Http://www.ipcc.ch/ipccreports/Ar4-Wg3.

Kiviet, J. F. (1995) On Bias, Inconsistency and Efficiency of Various Estimators in Dynamic Panel Data Models. *Journal of Econometrics* 68: 53–78.

— (1999) Expectation of Expansions for Estimators in a Dynamic Panel Data Model; Some Results for Weakly Exogenous Regressors. In *Analysis of Panels and Limited Dependent Variable Models*, eds. L.-F. L. C. Hsiao, K. Lahiri and M. H. Pesaran, 199–225. Cambridge: Cambridge University Press.

Lantz, V., Feng, Q. (2006) Income, population, and technology impacts on CO_2 emissions in Canada: where's the EKC? Ecological Economics 57(2): 229-238.

MacKellar L, Lutz W, Prinz C, et al (1995) Population, households and CO_2 emissions. Pop and Dev Review 21(4):849-865.

Martínez-Zarzoso I, Bengochea-Morancho A, Morales-Lage, R (2007). The Impact of Population on CO2 Emissions: Evidence from European Countries. Env and Res Economics 38: 497-512.

Panayotou T, Peterson A, Sachs J (2000) Is the Environmental Kuznets Curve driven by structural change? What extended time series may imply for developing countries?. CAER II Discussion Paper 80.

Perman R, Stern DI (2003) Evidence from panel unit root and cointegration tests that the Environmental Kuznets Curve does not exist. Aus J Agr and Res Economics 47:325-347.

Roberts JT, Grimes PE (1997) Carbon intensity and economic development 1962-91: a brief exploration of the Environmental Kuznets Curve. World Dev 25(2):191-198.

Pesaran, M.H. (2004) "General diagnostic tests for cross-section dependence in panels" CESifo Working Paper Series No. 1229; IZA Discussion Paper No. 1240.

Pesaran, M. H. (2006) Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure, Econometrica 74(4): 967-1012.

Sarafidis, V. and Robertson, D. (2008) On the impact of cross-sectional dependence in short dynamic panel estimation, mimeo, University of Sydney, Australia.

Schaffer, M.E., (2007) xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. http://ideas.repec.org/c /boc/bocode/s456501.html.

Shafik N, Bandyopadhyay S (1992) Economic growth and environmental quality: time series and cross-country evidence. Background paper for the World Dev Rep 1992, Working Paper No 904, The World Bank, Washington DC.

Shi A (2003) The impact of population pressure on global carbon dioxide emissions, 1975-1996: evidence from pooled cross-country data. Ecol Economics 44:29-42.

Stern DI (1998) Progress on the environmental Kuznets Curve?. Env and Dev Economics 3:173-196.

Stavins, R. N. and Jaffe, A. B. (1990), "Unintended Impacts of Public Investments on Private Decisions: The Depletion of Forested Wetlands", *American Economic Review*, 80, 337-352.

Stern DI (2004) The rise and fall of the Environmental Kuznets Curve. World Dev 32(8):1419-1439.

United Nations Population Division (2008) World Population Prospects: The 2006 Revision, UN, New York.

Winkler H, Spalding-Fecher R, Mwakasonda S, Davidson O (2002) Sustainable Development Policies and Measures: Starting From Development to Tackle Climate Change, in Baumert K, Blanchard O, Llosa S, Perkaus J eds. Building a Climate of Trust: The Kyoto Protocol and Beyond (61-87). Washington D.C.: World Resources Institute.

World Bank (2007), World Development Indicators 2001 CD-ROM, Washington, DC.

York R, Rosa EA, Dietz T (2003) STIRPAT, IPAT and ImPACT: analytic tools for unpacking the driving forces of environmental impacts. Ecol Economics 46(3):351-365.

Zaba B, Clarke, JI (1994) Introduction: current directions in population-environment research. In: Zaba B, Clarke JI (eds) Environment and population change. Derouaux Ordina Editions, Liège.

Appendix

Table A1. Summary statistics and simple correlations

Variable		Mean	Std. Dev.	Min	Max	Observations
Co2 emissions kt	overall	165280.20	542623.00	7.33	5794672.00	N = 3523
	between		534362.60	858.15	4933594.00	n = 122
	within		100321.40	-1037847.00	1964952.00	T-bar = 28.87
Population (1000)	overall	40003.24	128828.10	218.00	1288400.00	N = 3535
	between		128110.70	252.43	1112220.00	n = 122
	within		17368.97	-181063.70	269875.90	T-bar = 28.97
% Urban population	overall	8970.26	8501.44	479.74	50147.74	N = 3199
	between		7923.01	514.39	32013.76	n = 122
	within		2554.22	-3566.52	27104.23	T-bar = 26.22
GDP real per head	overall	55.34	22.18	4.80	100.00	N = 3538
	between		21.77	9.11	100.00	n = 122
	within		4.65	30.33	75.54	T = 29
Energy eficiency	overall	5.00	2.48	0.69	24.96	N = 3050
	between		2.34	0.74	11.16	n = 122
	within		1.00	0.89	23.08	T-bar = 25
% Industrial Activity	overall	33.24	11.10	6.25	79.09	N = 2950
	between		9.39	11.86	60.93	n = 121
	within		5.57	-3.39	59.56	T-bar = 24.38

Correlations	**LCO2**	**LYH**	**LP**	**PUPC**	**LEI1**	**LIA**
LCO2	1					
LYH	0.5044	1				
LP	0.6501	-0.2093	1			
LPUPC	0.3858	0.7461	-0.2246	1		
LEI	0.0027	0.3084	0.0354	0.2025	1	
LIA	0.3583	0.351	-0.0648	0.3735	-0.0804	1

Note: LYH denotes per-capita income, LP denotes population, LPUPC is the percentage of urban population over total population, LEI is energy efficiency, and LIA is the percentage of industrial activity over total GDP. All the variables are in natural logarithms.

Table A.2. Lists of countries in each group

Low income	**Lower-middle income**	**Upper-middle income**	**High income**	**Transition Economies**
Bangladesh	Albania	Argentina	**OECD**	Albania
Benin	Algeria	**Belize**	Australia	Bulgaria
Chad	Angola	Botswana	Austria	Czech Republic
Comoros	Armenia	Chile	Belgium	Estonia
Congo, Dem. Rep.	Azerbaijan	Costa Rica	Canada	Georgia
Cote d'Ivoire	Belarus	Croatia	Denmark	Hungary
Eritrea	Bolivia	Czech Republic	Finland	Latvia
Ethiopia	Brazil	**Equatorial Guinea**	France	Lithuania
Ghana	Bulgaria	Estonia	Germany	Moldova
Guinea	Cameroon	Gabon	Greece	Poland
Guinea-Bissau	**Cape Verde**	**Grenada**	Iceland	Rumania
Haiti	China	Hungary	Ireland	Russia
India	Colombia	Latvia	Italy	Slovak Republic
Kenya	Congo, Rep.	Lebanon	Japan	
Kyrgyz Republic	**Djibouti**	Lithuania	Korea, Rep.	
Lao PDR	**Dominican Republic**	Malaysia	Luxembourg	
Madagascar	Ecuador	**Mauritius**	Netherlands	
Malawi	Egypt, Arab Rep.	Mexico	New Zealand	
Mali	El Salvador	Oman	Norway	
Mauritania	Fiji	Panama	Portugal	
Mongolia	Georgia	Poland	Spain	
Mozambique	Guatemala	Romania	Sweden	
Nepal	**Guyana**	Russian Federation	Switzerland	
Niger	Honduras	**Seychelles**	United Kingdom	
Nigeria	Indonesia	Slovak Republic	United States	
Pakistan	Iran, Islamic Rep.	South Africa	**Non OECD**	
Rwanda	Jamaica	**St. Kitts and Nevis**	Bahrain	
Sao Tome and Principe	Jordan	**St. Lucia**	Cyprus	
Senegal	Kazakhstan	**St. Vincent and the Grenadines**	Hong Kong, China	
Sierra Leone	**Kiribati**	Trinidad and Tobago	Israel	
Sudan	**Lesotho**	Turkey	Kuwait	
Tajikistan	Macedonia, FYR	Uruguay	Malta	
Tanzania	**Micronesia, Fed. Sts.**	Venezuela, RB	Saudi Arabia	
Togo	Moldova		Singapore	
Uganda	Morocco		Slovenia	
Uzbekistan	Namibia		United Arab Emirates	
Vietnam	Nicaragua			
Yemen, Rep.	Paraguay			
Zambia	Peru			
Zimbabwe	Philippines			
	Samoa			
	Sri Lanka			
	Suriname			
	Swaziland			
	Syrian Arab Republic			
	Thailand			
	Tonga			
	Tunisia			
	Turkmenistan			
	Ukraine			

Source: World Development Indicators 2007. For countries in **bold** energy intensity was not available. Industrial activity was not available for Israel.

A.3 Estimated effect of population on emissions for low-middle income countries