THE AUSTRALIAN NATIONAL UNIVERSITY

Centre for Resource and Environmental Studies

Ecological Economics Program

Working Papers in Ecological Economics

Number 9901 **June 1999**

THE ENVIRONMENTAL KUZNETS CURVE:
IMPLICATIONS OF NON-STATIONARITY

by

Roger Perman and David I. Stern

CRES, ANU, Canberra ACT 0200, AUSTRALIA
Tel: +61 2 6249 4277
Fax: +61 2 6249 0757
E-Mail: dstern@cres.anu.edu.au
WWW: http://cres.anu.edu.au/~dstern/anzsee/EEP.html

The Ecological Economics Program (EEP) was established in CRES as the result of a successful application to the Institute of Advanced Studies 1994 Strategic Initiatives round. EEP's objective is to promote the development of ecological economics by conducting research and acting as a focal point for ecological economics activities in the Canberra region. To this end, EEP:

- hosts the office of ANZSEE, the Australia New Zealand Society for Ecological Economics
- runs a seminar series
- runs a working paper series on the World Wide Web
- can serve as a host for individuals wishing to do ecological economics research as visitors to the ANU (Contact Mick Common for more information).
- supervises research students. People interested in doing a research degree in ecological economics should contact Mick Common or David Stern for more information.

The **working paper series** reports on research conducted in EEP by academic staff, visitors, and students, and invited seminar presenters.

Working Papers in Ecological Economics

9601 David Stern. Progress on the environmental Kuznets curve?
9701 Michael Young. Water rights: an ecological economics perspective.
9702 David Stern and Robert Kaufmann. Time series properties of global climate variables: detection and attribution of climate change.
9703 Ken Menz and Katie Ellis. Fire as an economic disincentive to smallholder rubber planting in Imperata areas.
9704 Roger Attwater. Property entitlements and land reform in upland Thai catchments.
9705 Mick Common, Tim Bull, and Natalie Stoeckl. The travel cost method: an empirical investigation of Randall's difficulty.
9706 Colin Hunt. Economic instruments for environmental and natural resource conservation and management in the South Pacific.
9707 David Alden. Recreational user management of parks: an ecological economic framework.
9708 David Stern and Robert Kaufmann. Is there a global warming signal in hemispheric temperature series?
9801 Daniel McKenney. Resource economists should do more cost analysis and less benefit analysis.
9802 Roy Darwin. FARM: A global framework for integrated land use/cover modeling.
9803 David Stern. A multivariate cointegration analysis of the role of energy in the U.S. macroeconomy
9804 David Stern, Tony Auld, Mick Common, Kali Sanyal. Is there an environmental Kuznets curve for sulfur?
9805 Simon Niemeyer. Consumer-based carbon reduction incentives.
9806 Denise Mulholland and Elizabeth Wilman. Bioprospecting and biodiversity contracts.
9901 Roger Perman and David Stern. The environmental Kuznets curve: implications of non-stationarity.

The Environmental Kuznets Curve: Implications of Non-Stationarity

Roger Perman*

Department of Economics
University of Strathclyde
100, Cathedral Street
Glasgow
SCOTLAND
E-mail: r.perman@strath.ac.uk

David I. Stern

Centre for Resource and Environmental Studies
Australian National University
Canberra ACT 0200
AUSTRALIA
E-mail: dstern@cres.anu.edu.au

* Corresponding author

16 June 1999

Abstract

In this paper, we apply time series techniques for panel data to the environmental Kuznets curve (EKC) model. Within the literature that estimates emissions-income relations in the EKC context, little attention has been paid to the time series properties of the data and in particular to whether the variables could be integrated time series. We estimate the EKC for sulphur emissions using a panel data set for 74 countries over 30 years. Using individual unit root tests, we find that both sulphur emissions and GDP per capita are integrated variables in the majority of countries. This result is confirmed by panel unit root tests that find that the panel series are integrated. Individual cointegration tests show that EKC relations in most countries do not cointegrate. Results of a number of panel cointegration statistics are mixed. Even if there is cointegration in the panel many of the individual EKC functions are U shaped or monotonic in income. There is no single cointegrating vector common to all countries. The results show that the EKC may be a problematic concept, as simple global EKC models are misspecified.

Acknowledgements

The authors thank Peter Pedroni for providing econometrics code, and feedback on some of our results and an earlier draft of the paper. We also thank Mick Common and Yongcheol Shin for useful comments. Kali Sanyal assembled the data set.

1. INTRODUCTION

The environmental Kuznets curve (EKC) hypothesis proposes that there is an inverted U-shape relation between various indicators of environmental degradation and income per capita. Several studies have attempted to test the Environmental Kuznets Curve (EKC) hypothesis empirically. The majority of these studies use panel data in conjunction with a static fixed and/or random effects panel estimator. Little or no attention has been given in the literature to the time-series nature of the panel data, nor to its dynamics. A key maintained (but untested) assumption in much of this body of research is that the variables used in the regression analysis are covariance stationary in the time-series dimension of the panel.

The statistical properties of estimators are very different in the case where variables are stationary than where they are not stationary (see, for example, Hamilton, 1994). Many of these differences carry over to panel estimation. It is, therefore, important to test whether the variables used in EKC studies are stationary; and if they are not, it is important to take this non-stationarity into account in subsequent estimation and statistical inference. Given this, what is the statistical validity of the existing body of work, and what does good econometric practice require that we do in further studies of this type? It is these questions that we are addressing in this paper.

The paper is structured as follows. Section 2 surveys the concept of the EKC hypothesis and establishes the sense in which that term is used in this paper, with particular reference to the EKC for sulphur emissions. Section 3 describes the data used in the analysis. Section 4 demonstrates that we are unable to reject the null hypothesis that the variables used in our study (logarithms of emissions and income per capita) are integrated time series. This finding implies that much of the existing knowledge about EKCs is based on a statistically unsound foundation.

We then test for cointegrating relationships between the variables of interest in Section 5. We allow for greater heterogeneity among individual members of the panel than is commonly permitted, allowing, not only for intercepts to differ among countries and over time, but also variation in the long-run parameters of the emissions/income relationship among countries. In Section 6 we test whether the individual estimates converge to a common cointegrating vector over the whole sample (or some sub-samples of interest), in the context of dynamic panels specified using an error correction representation of an autoregressive distributed lag panel model. Section 7 concludes.

2. THE NATURE OF THE EKC

The EKC is a reduced form relationship that may arise from one or more different structural relationships. Actual changes in emissions must be due to changes in: scale of the economy, input mix, output mix, and technological change. But these "proximate factors" may be driven by a variety of underlying factors such as environmental regulation and preferences for environmental quality that in turn may evolve with the level of income. Ideally, modelling of the income-pollution relationship would make direct use of these structural relationships. However, difficulties in obtaining the necessary data and in specifying those relationships have led researchers to work with the income-emissions reduced form. Several different “structures” could generate a quadratic inverted U shaped emissions-income relationship and several of these are discussed in the literature (Lopez, 1994; Selden and Song, 1995; John and Pecchenino, 1994; and John *et al*., 1995; McConnell, 1997).

In Lopez's (1994) model, if producers pay the social marginal cost of pollution then the relation between emissions and income depends on the properties of technology and preferences. If preferences are non-homothetic, the response of pollution to growth depends on the elasticity of substitution in production between pollution and the conventional inputs, and the degree of relative risk aversion i.e. the rate at which marginal utility declines with rising consumption of produced goods. The faster marginal utility declines with rising income and the more substitution is possible in production the less pollution will tend to increase with production. For empirically reasonable values of these two parameters, pollution may increase at low levels of income and fall at high levels - the inverted U. Selden and Song (1995) derive an inverted U curve from a model that is somewhat similar to Lopez's. While Lopez (1994) and Selden and Song (1995) both develop models based on infinitely lived agents, John and Pecchenino (1994) and John *et al*. (1995) develop models based on overlapping generations (McConnell, 1997). Therefore, the pollution externality is only partially internalised in these models. In addition, pollution is generated in these latter models by consumption rather than production activities. All these models can generate inverted U shape curves under appropriate conditions. McConnell (1997) and Ansuategi *et al*. (1996) also develop models of consumption pollution. McConnell (1997) uses his model to argue that there is no defining role for the income elasticity for environmental quality in the EKC model. While a higher elasticity will lead, *ceteris paribus*, to a faster reduction in pollution, pollution can decline even if the elasticity is non-positive.

Institutional factors, such as power and income inequalities (Torras and Boyce, 1998), or structural changes such as urban and industrial decentralisation (Stern *et al*., 1996), might mediate between the underlying structures described above and proximate factors such as output structure or the state of technology. Other mechanisms may generate EKC-like relationships that imply that the EKC is a statistical artefact rather than a functionally defined path that all countries proceed along. For example, per capita emissions of carbon rose strongly in many developed countries up till the oil price shocks of the 1970s but subsequently declined in many countries (Moomaw and Unruh, 1997). The relation between trade and development provides an alternative explanation. As development proceeds the trade specialisation of countries changes. The first countries to develop increasingly specialise towards human and manufactured capital intensive products and "outsource" production of labour and resource intensive production to developing countries. But because the total number of countries is limited, countries that develop later will find it harder to reduce environmental impacts as there will be less scope to specialise away from resource intensive production (Stern *et al*., 1996).

These explanations for the empirically observed EKC relation propose that the observed EKC is spurious, as it is not due to behaviour consistent with the EKC theory. Throughout the rest of this paper, we shall be dealing with the EKC hypothesis that all countries follow a similar, if not identical, development path over time. The basic EKC model is given by:

$$\ln\left(\frac{M}{P}\right)_{it} = \alpha_i + \chi_t + \delta_i t + \beta_{1,i} \ln\left(\frac{Y}{P}\right)_{it} + \beta_{2,i}\left[\ln\left(\frac{Y}{P}\right)\right]^2_{it} + \varepsilon_{it} \qquad (1)$$

where M is some measure of pollutant emissions or concentration (or some other index of environmental pressure), Y is constant price national income, P denotes a country's population size, and t is a deterministic time trend. The variables are observed over a panel of countries (i = 1,…,N) and time periods (t=1,…,T). The random disturbances are assumed to be independent across countries. Non zero α_I terms allow for country-specific "effects"; the χ_t terms are time-specific dummy variables, usually interpreted as disturbances affecting all countries in the panel at some point in time in a common way; and $\delta_i t$ allows heterogeneous linear time trends over the sample of countries. Some or all of these country-specific or time-specific effects, or time trends, may be restricted either on the basis of prior information or after some specification search process.

A weaker version of the EKC hypothesis is that the EKC has a "common form", equivalent to the restriction that $\beta_{1i} > 0$ and $\beta_{2i} < 0$ for all i, but that these parameters have different values in different countries. A stronger version of the hypothesis is that the weak condition is satisfied and $\beta_{1i} = \beta_1$ and $\beta_{2i} = \beta_2$ for all i. There are no strong theoretical grounds for believing that the latter set of restrictions is satisfied. Some previous studies on other datasets (e.g. Dijkgraaf and Vollebergh, 1998), together with preliminary examination of the data used in this study (Stern and Common, 1998) suggest rejection of long-run parameter homogeneity. Therefore, in contrast to the usual practice when using fixed effects or similar panel estimators, we do not impose these restrictions *a priori*. But we do examine whether individual country cointegrating vectors converge on a global cointegrating vector. This allows a test of the usual assumptions of the strong hypothesis.

Our model consists, then, of a heterogeneous panel, with variation between countries in the parameters of the long-term relationships, some of the deterministic components, and the short-run dynamics. Neither does it impose equality of error variances over countries. Because of this, the distributions of the panel statistics we use are non-standard (see Pedroni, 1998).

The EKC model described by (1) is a static model. All adjustment to any shock takes place within the time period in which this occurs. This is justified, either if adjustment processes are really very fast, or if equation (1) (without its disturbance term) represents an equilibrium relationship. It is inconceivable that the adjustment process in the relationship we are studying is actually ‘instantaneous’. On the contrary, the stories we tell to explain the EKC suggest slow adjustments. Given this, where time series or panel data is used, a statistically sound approach requires estimating a dynamic model of some form. As this has not been done before for panel data, we conclude tentatively that previous EKC models are misspecified. Even if (1) is an equilibrium condition, that relationship could be estimated consistently by simple static regressions only in very special circumstances. Specifically, if all variables in the regression were covariance stationary, then the static regression would require that all omitted variables (in this case omitted lagged values of variables) were uncorrelated with their current dated levels, a most unlikely condition. However, where the data are integrated of order one in the time series dimension, we can obtain consistent (although possibly highly biased) estimates of the long-run parameters from static regressions (see Banerjee *et al*, 1993). We show below that the assumption that the data is stationary is probably incorrect, and that once the non-stationarity of the data is recognised, a different approach is needed for estimating EKCs and testing hypotheses about them.

The usual fixed effects and random effects estimators transform the data to eliminate specific country effects and time effects common to all countries. Hence, if there are trends in the data that are common to all countries - whether variable specific or common to a group of variables - these will be eliminated from the data used in the regression procedure. However, these methods cannot eliminate trends that are country specific. It seems very unlikely that countries at very different levels of development will share the same trend contemporaneously, even if the parameters of any cointegrating vectors may be common to all countries.

Grossman and Krueger (1991) seem to have been the first of several studies that use pooled time-series/cross-section datasets. Occasionally, use is made of simple cross-section analysis (e.g. Roberts and Grimes, 1997), whilst other studies have focussed on individual country time series analysis (e.g. Dijkgraaf and Vollebergh (1998)). The literature on this topic is now fairly large because it has investigated various indices of environmental pressure. The evidence for an EKC of the postulated form is rather mixed, and overall is relatively weak (Ekins, 1997). However, for a number of environmental pressure variables, it is widely argued that data supports the EKC hypothesis. For good recent surveys of the empirical evidence, see Ansuategi *et al*., 1998; Ekins, 1997; and Stern, 1998.

3. THE DATA USED IN THIS STUDY

The dataset that we use is relatively large in both the N and T dimensions. Estimated sulphur emissions for a broad set of 74 countries (OECD and non-OECD) covering 81% of world population are taken from a database constructed by A.S.L and Associates. Income is measured in constant price, PPP adjusted, income, and is taken from the Penn World Table for 1960-1990.

Further details of the data are provided in Stern and Common (1998). Both emissions and income are transformed to logarithms. We choose to examine sulphur emissions because previous studies (Cole *et al*., 1997; de Bruyn, 1997; de Bruyn *et al*., 1998; Grossman and Krueger, 1991; Kaufmann *et al*., 1998; Panayotou, 1993, 1995, 1997; Selden and Song, 1994; Shafik, 1994; Shukla and Parikh, 1992; Torras and Boyce, 1998; Vincent, 1997) suggest that sulphur emissions are a likely candidate for finding an inverted U shape EKC.

4. TIME-SERIES PROPERTIES OF THE DATA: UNIT ROOT TESTING

CONVENTIONAL INDIVIDUAL (UNIVARIATE) ADF TESTS

The results of univariate unit root tests on our dataset are reported in Table 1. The columns labelled Y1, Y2 and Y3 refer to ln(Y/P), and those labelled M1, M2 and M3 to ln(M/P). We do not report test statistics for $[\ln(Y/P)]^2$, as these are virtually identical to those for ln (Y/P). Columns M1 and M2, and Y1 and Y2 report augmented Dickey-Fuller (ADF) statistics on the time series for individual countries after common time means have been subtracted from the data. This transformation is equivalent to estimating the ADF regressions with individual dummies χ_t included for each period of time. The lag lengths in the ADF regressions were chosen separately for each country using the Hall (1991) procedure. [1]

Columns Y1 and M1 give ADF t test statistics from ADF regressions that do not include trends; those in Y2 and M2 report ADF t test statistics from regressions that do include time trends. The distribution of the statistics under the null hypothesis of a unit root depends not only on the form of the estimated regression but also on the true but unknown nature of the data generating process (Hamilton 1994). The relevant distributions are reported in the notes to Table 1.

It is not clear in advance whether deterministic time trends should be included in the ADF regressions. It is, therefore, not clear which of the columns is relevant, nor which of the distributions should be used in obtaining critical values. Consequently, we have supplemented these statistics with

---

[1] This is a conventional step-down procedure that begins with a preselected maximum lag in the ADF regression. This is followed by a sequential search procedure: one-step reductions of the lag length are made until they can no longer be rejected in testing for the significance of the final included lag using a t test.

a full sequential unit root testing procedure for each time series for each country. This procedure is based on that suggested in Campbell and Perron (1991), and elaborated by Holden and Perman (1994).[2] Columns Y3 and M3 report only one item of inference drawn from these searches, namely whether the series in question is a random walk (i.e. contains a single unit root) or is an I(0) stationary process. Some of these stationary series also contain a deterministic time trend and some of the random walk processes contain a drift term. We used this information to decide whether to report the statistics from ADF regressions with or without a time trend. Where the sequential search procedure suggests that a deterministic time trend is (is not) in the DGP for a series, only the ADF t statistic with (without) a time trend is reported in Table 1.

The univariate test statistics strongly support the view that both variables are I(1) processes. Using columns Y1 and Y2, the unit root null for income per capita is rejected in only 18 out of the 74 countries. The full search procedure (column Y3) results in only 6 definite rejections and four borderline cases.

Matters are less clear for emissions per capita. The two simple ADF tests reject the unit root null for 37 countries out of 74. However, the full search procedure yields only six clear rejections and four borderline cases. There are two reasons for this apparent discrepancy. First, the full search tests are conducted on data that did not have time means removed. Second, the full search procedure allows for a greater variety of paths in arriving at a final inference.

PANEL UNIT ROOT TESTS

In recent years, an alternative framework has been developed for implementing unit root tests in panel data (see Quah, 1994, Levin and Lin, 1993, Im, Pesaran and Shin, 1996, and Pedroni, 1995). These tests were developed, at least partially, in response to some important weaknesses of existing single time series tests. Conventional unit root tests on single time series, such as the Dickey-Fuller or augmented Dickey-Fuller (ADF) procedures, often suffer from unacceptably low power when applied to series of moderate length, and are susceptible to large size distortions (especially in the presence of moving average errors). Panel data exploits more information and so

[2] The procedure can be found in a RATS routine URADF.SRC that can be downloaded from the Estima home page.

improves test power. In effect, pooling allows the researcher to exploit the information that, under the null hypothesis of stationarity of a variable over the whole panel, the autoregressive root is unity each country. This restriction is imposed in the pooled unit root tests, and so increases test power. However, the pooling we use does not constrain the properties of the disturbance terms or equation dynamics in the ADF regressions to be homogeneous over countries. So, for example, the number of lagged values of the differenced dependent variable in the ADF regressions, the equation variances, and the long-run covariance matrices, may be country-specific.

We compute and report two forms of panel unit root test statistic, one similar in spirit to the Levin and Lin (1993) testing framework (hereafter called 'panel' statistic), and the other based on the group mean t statistic developed by Im, Pesaran and Shin (IPS) (1996) (hereafter called the 'group' statistic). We used RATS code made available by Peter Pedroni to implement these tests. The panel statistics reported below are derived from regressions including time dummies to eliminate common time effects. We report the group statistics with and without time dummies. Each of these statistics is constructed to have an asymptotic normal distribution, and is mean and variance adjusted so that the distribution is standard normal. Given this, the unit root null for the panel as a whole is rejected if the statistic is smaller than the one-tailed 5% significance critical value of -1.645.

With just one exception, the statistics point to each of the three series for all countries in the panel containing a single unit root (Table 2). This conclusion is most robust in the case of the two income per capita variables, with all statistics being far from their critical values at conventional levels of significance. The statistics reinforce the findings of the individual country ADF test statistics, suggesting that the widespread failures to reject the null of non-stationarity are not attributable to low power. Inference is also strong but not unanimous for the emissions per capita series. The only statistic suggesting a (trend) stationary process is the group statistic from a model including both heterogeneous trends and time dummies for ln(M/P).

5. COINTEGRATION ANALYSIS

Any attempt to estimate the parameters of supposed EKC relationships and to test hypotheses about these must take into account the fact that the regression is among integrated data. If the variables used in EKC analysis are not I(0) variables, then the regressions may be spurious in the statistical sense, and inference using classical methods and distributions would be invalid. Thus, for example, in the case of a spurious regression for a single country (N = 1), the conventionally calculated t statistic does not have a t distribution, indeed it does not have any limiting distribution, and diverges as the sample increases in the T dimension. Moreover, the F test also does not have its standard distribution and diverges with T. The R^2 statistic cannot be interpreted in the conventional way, and the DW statistic does not have its usual distribution. Whilst matters are different in the case of regressions among I(1) variables that do cointegrate (see, for example, Kao (1997)), it remains the case that conventional inference is not appropriate. So again, for example, t and F statistics are divergent, making the probability that a null will be rejected go to one as N increases. In general, the only way in which valid inference can be drawn about (and restrictions tested on) parameters of long-run relationships is when these are embedded in a more completely specified dynamic model. Banerjee *et al* (1993) provides a thorough account of these results for the single country spurious case. Given all this, a substantial part of the extant empirical literature on the EKC may have used invalid techniques of statistical inference.

In general, regressions among non-stationary variables will be spurious regressions. However, if some linear combination of the variables is stationary, then this will correspond to a long-run equilibrium, cointegrating relationship among the variables. Despite the fact that testing hypotheses about the parameters of these regressions is not possible using conventional techniques, ordinary least squares (OLS) estimates will be superconsistent. If we find evidence of cointegration in the EKC relation, then OLS estimates will be consistent. Incidentally, in the case of cointegrating regressions, evidence that the dependent variable has feedback effects on regressors, does **not** invalidate the use of simple OLS regression techniques. Tests for simultaneity bias (as in Holtz-Eakin and Selden (1995) and Cole *et al*. (1997) for example) are not necessary.

Second, conventional criteria of statistical misspecification, such as the absence of serial correlation in the residuals, are often taken as evidence of fundamental equation misspecification. Whilst this is certainly true in the case of a regression among stationary variables, matters are more subtle for regression among non-stationary variables. In the latter case, the static model constitutes a cointegrating (or spurious) regression. Consistent parameter estimates are obtained if the variables cointegrate. Cointegration requires that the residuals are stationary, but not the stronger condition that they are uncorrelated.

One would expect that these regressions **do** show conventional evidence of statistical misspecification because the static cointegration regressions are dynamically misspecified: the dynamic adjustment processes are not modelled in the static regression. For this reason, the parameter estimates can be badly biased in finite samples even though the estimator is consistent (Banerjee *et al*, 1993). The non-stationary time series literature has long recognised that an appropriate modelling strategy is to estimate a dynamic model of the relationships among the variables such as the Engle-Granger or Johansen procedures.

RESIDUALS-BASED SINGLE EQUATION COINTEGRATION TESTS

For any single country, estimating a "cointegrating regression" by OLS and testing for a unit root in the regression residuals can test the null hypothesis of non-cointegration. We distinguish between three special cases (labelled cases 1, 2 and 3 below), depending on which combination of deterministic trend and time dummies is included.[3]

CASE 1 Model without heterogeneous deterministic trends or time dummies:

$$\ln\left(\frac{M}{P}\right)_{it} = \alpha_i + \beta_{1,i}\ln\left(\frac{Y}{P}\right)_{it} + \beta_{2,i}\left[\ln\left(\frac{Y}{P}\right)\right]^2_{it} + \varepsilon_{it} \qquad \text{for i = 1,..,N (2a)}$$

CASE 2: Model without heterogeneous deterministic trends but including time dummies:

$$\ln\left(\frac{M}{P}\right)_{it} = \alpha_i + \chi_t + \beta_{1,i}\ln\left(\frac{Y}{P}\right)_{it} + \beta_{2,i}\left[\ln\left(\frac{Y}{P}\right)\right]^2_{it} + \varepsilon_{it} \qquad \text{for i = 1,..,N (2b)}$$

CASE 3: Model with heterogeneous deterministic trends and time dummies:

$$\ln\left(\frac{M}{P}\right)_{it} = \alpha_i + \chi_t + \delta_i t + \beta_{1,i}\ln\left(\frac{Y}{P}\right)_{it} + \beta_{2,i}\left[\ln\left(\frac{Y}{P}\right)\right]^2_{it} + \varepsilon_{it} \qquad \text{for i = 1,..,N (2c)}$$

---

[3] If attention is restricted only to individual equation analysis, then the inclusion of time dummies is not necessary. These dummies are designed to proxy for common time effects over the countries in the panel; if we look **only** at individuals in the panel, there is nothing lost by ignoring such cross-country effects. However, since we will later be examining the panel as a whole, these cross-country effects must be controlled for so as to avoid cross-country dependence in the errors. To maximise comparability of results throughout this paper, we report single equation results with time dummies included (by demeaning the data; but see footnote 2 again).

Denoting the residuals from the least squares regression as as $\hat{\varepsilon}_{it}$, the second step consists of estimating an auxiliary regression using these residuals. For ADF statistics, the augmented Dickey-Fuller (ADF) regression on the cointegrating regression residuals is estimated:

$$\Delta \hat{e}_{it} = (r - 1)\hat{e}_{it-1} + \sum_{t=1}^{t=q} d_{it} \Delta \hat{e}_{it-q} + v_{it} \tag{3}$$

In this regression as many lags of the dependent variable should be included as are necessary to ensure that the residuals v_{it} are serially uncorrelated.[4] The ADF test procedure uses a t test of the null hypothesis that ρ = 1, which implies that $\hat{\varepsilon}_{it}$ is a unit root process and so the “cointegrating regression” does not cointegrate, against the alternative that ρ < 1. Non-standard tables of critical values must be used.

Table 3 summarises our cointegration test results. Those which relate to the three specifications described above are listed in the column headed 'Quadratic'. These test statistics do not provide strong evidence for cointegration in individual countries in the panel. In particular, in just under one half of all cases (35 out of 74 countries), none of the models exhibited cointegration between emissions per capita and the first and second powers of income per capita. Cointegration is found most frequently where the regressions include a deterministic trend.

Statistics listed in the column headed 'Linear' are for models which exclude the second power of income per capita, and impose a monotonic form on the EKC. Note that the support for cointegration is substantially weaker in this case than where the relationship allows for non-linearity.

Table 4 lists some qualitative results concerning the estimates from all individual quadratic cointegrating regressions. However, these estimates are derived from regressions that may not cointegrate in all cases. Furthermore, we have no means of knowing which individual estimates are statistically significant, as conventional t ratios and F tests have non-standard distributions in regressions among non-stationary variables. Nevertheless, they are useful in giving us a quick

[4] Where we use the phrase “heterogeneous lags” in this paper, this refers to the fact that the number of lags q used in the ADF regression is selected separately for each country. “Hall’s method” refers to the criterion used for choosing that number.

summary of the extent to which the signs on the variables expected under the EKC are found in the data.

Looking at the first row (the regression without time trends included), we see that only 42 out of 74 of the countries have parameter signs which conform with the EKC hypothesis at the single country level. Moreover, over one third of the countries appear to have U shaped (rather than inverted U shaped) emissions/income relationships (if indeed any relationship exists at all). The relative proportions are even less favourable to the EKC hypothesis when time trends are included.

Figure 1 gives a feel for the distribution of the individual pairs of parameter estimates over the whole sample of countries. The figure cross-plots the estimates of β_1 and β_2. There is a striking negative linear relationship between the parameters over the sample. There is also a very strong correlation between the intercept terms and the two GDP parameters. These correlations are close to 1 or –1. The larger the constant is, the greater in absolute value are both the GDP parameters. Most of the EKC effect is taken up by the individual country means and the GDP parameters in the individual country regressions then adjust to cope with these different intercepts. Basic differences between countries are much more important in explaining the EKC than is growth within countries. These correlations are similar to the inconsistency of the panel data random effects models (Stern and Common, 1998). In this case, though, the correlation is between the intercepts and the coefficients of the GDP variables while in the panel estimate the relevant correlation is between the intercepts and the GDP variables themselves.

Note also that a strict interpretation of the EKC requires parameter combinations to lie only in the lower left quadrant of Figure 1.

PANEL COINTEGRATION TESTS

Like unit root tests, single individual (country) cointegration tests suffer from low power. This low power can provide another interpretation of the results in the previous section. Even if the postulated EKC relationship was generally true, the relatively short spans of data (T=31) suggests that low power might lead a researcher to reject cointegration far more often than should be done. A panel cointegration testing approach might provide a firmer base for inference.

Panel tests for cointegration are a development from the panel unit root testing literature that we have discussed and used earlier. A good survey of this literature can be found in Pedroni (1997, 1998). As in the case of panel unit root tests, proponents of panel cointegration tests typically use improved power as the basis for their advocacy. Panel tests can improve power by exploiting information from pooling (but still allowing dynamics and fixed effects to be heterogeneous). The tests we use in this paper are those formulated by Pedroni (1997, 1998). They are particularly appropriate for panels in which both N and T are of moderately large dimension (which is arguably the case here), so that GLS estimators are not easily applicable and individual country cointegrating regressions are likely to suffer from limited power. Moreover, our estimation and testing framework satisfies some basic characteristics of the panel data by allowing for heterogeneity among panel members in both long-run relationships and short run dynamics (and deterministic trends, where appropriate). Other than being adapted to the context of a panel, the tests we use here are relatively conventional. They are residuals-based tests of the null of no cointegration.

Because panel cointegration tests test a particular form of hypothesis, power comparisons are somewhat questionable (Maddala, 1998). We can envisage these tests in the following way: the null hypothesis asserts that for each individual panel member, the variables are not cointegrated, whilst the alternative asserts that for each individual there exists a single cointegrating vector, although this vector may be unique for each individual.

One set of tests used below - adapting a framework originally developed by Levin and Lin (1993), and hereafter called 'panel' statistics - is based on pooling over the so-called *within* dimension. Numerator and denominator components of the test statistics are summed separately over the N dimension. A second set - in the spirit of Im, Pesaran and Shin (1996), and which are hereafter called 'group' statistics - is based on pooling among the *between* dimension, obtaining the ratio of numerator to denominator for each country prior to aggregating over the N dimension. An advantage which has been claimed for the between statistics over the within statistics is that they impose fewer untested homogeneity restrictions. Specifically, the *within* statistics constrain the autoregressive roots in the cointegrating equation residuals regression to be common across all members of the panel under **both** the unit root null and the stationary alternative hypotheses. In contrast, the *between* statistics constrain the autoregressive roots to be common under the unit root null, but permit the roots to differ over the panel under the stationary alternative hypothesis. In the framework of equation (3), the within statistics test the null $\rho = 1$ for all i against the alternative $\rho_i = \rho < 1$ for all i, whereas the between statistics test the null $\rho = 1$ for all i against the alternative $\rho_i < 1$ for all i. Another - and possibly more intuitive - way of thinking about the difference is as follows. A panel test derives a single test statistic directly from the pooled data; a group statistic is derived by obtaining a statistic separately for each cross-section unit, and then forming a single statistic as a standardised average of the N individual statistics.

Full details of the test statistics, together with an examination of some properties of the tests in Monte Carlo simulations, are given in Pedroni (1997, 1998). Except in special circumstances (which are not applicable to our dataset) none of these seven statistics clearly dominates the others in terms of power and size properties. However, Pedroni's results suggest that the panel variance statistic is dominated by the others in many circumstances (and should be regarded as unreliable). We report all statistics and base inference on our judgement about the implications of the set as a whole. All test statistics were computed using RATS code provided by Pedroni.

In contrast to the case of our tests for cointegration at the individual country level, the results shown in Table 5(a) - together with others not reproduced here - do not yield a strong, robust conclusion about the existence of cointegration over the panel. Our preferred model is that in the middle row, which includes time dummies (to eliminate cross-country common time effects that would otherwise create cross-equation dependence in the error terms) but does not incorporate country-specific deterministic time trends. We have reported statistics for the other two cases to allow comparison with the results of other studies and to show sensitivity of our results to modelling assumptions.

For the preferred model, five out of the seven statistics suggest cointegration over the panel as a whole at the 5% level or better. However, the two ρ-based statistics suggest no cointegration in this specification (or either of the others). The evidence for cointegration is considerably weaker in the specification with no time dummies or trends, but this is of little practical importance given the consensus that time dummies are necessary to validate the conventional estimation assumption of cross-section independence. Inclusion of deterministic trends does little to alter our inference regarding cointegration.

It is important to note that inference is rather sensitive to the choice of the maximum lag length allowed for in the testing procedure (see the notes to Tables 5(a) and (b) for more on this matter). If a smaller lag truncation is used (than the value of 3 in Tables 5(a) and (b)), test statistics typically fall in absolute value, thus weakening evidence in favour of cointegration. The opposite happens when the truncation is increased: in particular, selection of an excessively large truncation length ("overfitting") leads to misleadingly high absolute values of the parametric test statistics. This presents the researcher with something of a quandary: whilst there are well-accepted routines for choosing lag lengths in single regressions, there is no robust equivalent when dealing with panel estimation. We used a general-to-specific pre-testing procedure, beginning with a maximum truncation of 6 lags. The final choice of maximum lag truncation (here found to be three) is based on inspecting the behaviour of the endogenously chosen lag length for each country as the maximum truncation changes. Judgement cannot be avoided in this process, and different researchers would not necessarily arrive at the same conclusion for any given data set.

Subject to all these qualifications, there is some support for the hypothesis that there is a cointegrating quadratic relationship between emissions per capita and first and second powers of income per capita over the panel as a whole. This contention may be surprising given the country-by-country findings that gave little or no support for cointegration. However, this is exactly why doing panel testing may be important: the low frequency of cointegration in individual country regressions could easily reflect the very low power of tests in that context. The growing literature that empirically examines the purchasing power parity (PPP) hypothesis supports this interpretation. While long-run PPP is typically using individual country data, panel cointegration approaches provide much greater support for the hypothesis (see Pedroni, 1997).

This does not necessarily mean that where panel and single country inferences diverge we should believe the former. But given the well-known poor power and size properties of residuals-based cointegration tests derived from relatively small-sample individual country regressions, we would be wise to attach a reasonably high weight to the panel results.

We note two further points at this stage:

- Though each country may have a cointegrating relationship of the type indicated, we have not yet tested whether it is possible to restrict the relations to a single, common cointegrating vector. We conduct that test in the following section. However, it seems most unlikely that we will be able to accept the restrictions given the huge variability of point estimates found in the single country equations.
- For a substantial proportion of the countries in our sample, preliminary investigation of point estimates suggests that the implied cointegrating relationships are U shaped or monotonic. So even if the panel as a whole cointegrates (that is, that every country has a, possibly distinct, cointegrating vector), this is **not** equivalent to a verification of the EKC itself.

Pursuing the latter observation a little further, we investigate the possibility that the "EKC" relationship is linear over the whole panel by calculating panel cointegration statistics from models (1) to (3) excluding the square of the logarithm of income per capita. We report these results in Table 5(b) and the signs obtained on the income parameters in these regressions in Table 9. Again, the results are mixed. Whilst support for the hypothesis of cointegration over the panel is weaker than in the case of non-linear models, it is not non-existent. In our preferred specification (the second model), three out of the seven statistics cannot reject cointegration (compared with five for the non-linear specification). As the existence of an inverted U shaped EKC at the individual country level depends on a negative coefficient on the second order term in income, the failure of our tests to **decisively** differentiate between the two specifications casts considerable doubt on the EKC hypothesis.

6. DYNAMIC EKC MODELS AND TESTING FOR A COMMON LONG-RUN VECTOR

We proceed under the assumption that there is a cointegrating relationship between emissions and the first and second powers of income for each country in our panel. We estimate an unrestricted dynamic EKC model for each country, and test whether the individual countries' emissions/income relationships converge to a common cointegrating vector. The estimated equation is an autoregressive distributed lag (ADL) model, parameterised in error correction form:

$$\Delta Y_{it} = \left\{\alpha_i Y_{it} - \beta_{1,i} X_{1,it} - \beta_{2,i} X_{2,it}\right\} + \sum_{j=1}^{p-1} \gamma_{ij} \Delta Y_{i,t-j} + \sum_{j=0}^{q-1} \delta_{1,ij} \Delta X_{1,i,t-j} + \sum_{j=0}^{r-1} \delta_{2,ij} \Delta X_{2,i,t-j} + \mu_i + \eta_t + \varepsilon_{it} \tag{4}$$

For compactness of notation, we have used the symbols Y, X_1 and X_2 for ln(M/P), ln(Y/P) and $[\ln(Y/P)]^2$ respectively. μ_i and η_t are country and time specific intercepts. The autoregressive lag length p, and the distributed lags q and r, were selected separately for each country from a maximum lag of four using the Akaike information criterion. Recalling our earlier finding that there is a (possibly distinct) cointegrating relation between Y, X_1 and X_2 for each country, the term in braces in equation (4) is a stationary variable, as are all other terms in that specification. Hence (subject to a few minor qualifications), classical estimation and inference procedures can be used in the context of this specification [5] and the statistical adequacy of this regression model can be assessed using conventional diagnostic statistics (such as RESET, normality, heteroscedasticity, and serial correlation tests). This is in contrast to static regressions (including fixed effects) where, as we explained earlier, they are inapplicable.

Another property of this specification follows from its reparameterised form

$$\Delta Y_{it} = \alpha_i \left\{ Y_{it} - \left(\frac{\beta_{1,i}}{\alpha_i}\right) X_{1,it} - \left(\frac{\beta_{2,i}}{\alpha_i}\right) X_{2,it} \right\} + \sum_{j=1}^{p-1} \gamma_{ij} \Delta Y_{i,t-j} + \sum_{j=0}^{q-1} \delta_{1,ij} \Delta X_{1,i,t-j} + \sum_{j=0}^{r-1} \delta_{2,ij} \Delta X_{2,i,t-j} + \mu_i + \eta_t + \varepsilon_{it} \tag{4b}$$

in which the nature of the error-correction mechanism is more readily apparent. The term within braces is the period t "disequilibrium error"; whenever the system is out of equilibrium, this error will be non-zero. α_i is the i^{th} country's error correction coefficient, providing information about the speed of adjustment of the system back to equilibrium. The existence of a stable equilibrium (or cointegration in other words) implies that this coefficient should be negative (so that if Y is above its target value it should then fall), and lie in the interval $\{0 < \alpha_i \leq 1\}$. An alternative test of cointegration is based on the ability to reject the hypothesis that $\alpha_i = 0$ in favour of the one-tailed alternative that it is less than zero. If the null could not be rejected, then there is no mechanism restoring an out-of-equilibrium system back towards equilibrium.

We estimate the model using maximum likelihood.[6] Table 6 presents regression results for the whole sample of countries, including those for various restricted cases of equation (4), and Tables 7 and 8 present results for OECD and non-OECD sub-samples. Some further details of the restricted models are given below.

Ignoring fixed effects, we obtain the long run equilibrium relationship for each country by solving either (4) or (4b):

$$Y_i = \left(\frac{\beta_{1,i}}{\alpha_i}\right) X_{1,i} + \left(\frac{\beta_{2,i}}{\alpha_i}\right) X_{2,i} \qquad (5)$$

Given that the parameters are all indexed by i, the long-run relationship in this specification is not restricted to be common over countries. Setting $\alpha_i = \alpha$ and $\beta_i = \beta$ for all i imposes homogeneity over all long-run parameters but permits dynamics and fixed effects to be heterogeneous over the panel. For the whole sample of countries this involves 146 restrictions. Imposing these leads to the maximised log likelihood falling from 1828 to 1513, with a likelihood ratio statistic of 631 and a p value indistinguishable from zero. So, likelihood ratio tests reject the hypothesis that the long-run EKC parameters converge to a common cointegrating vector at any level of significance. The hypothesis of a common cointegrating vector is decisively rejected when the sample is restricted to either OECD or non-OECD countries alone.

However, small sample bias and inefficiency may lead to very unreliable estimates of the individual equations (Pesaran *et al.*, 1998), and there may be a large dispersion of coefficient estimates even when the true coefficients are much more tightly clustered. A Hausman (1978) test is sometimes used to examine the possible homogeneity of panels as an alternative to classical tests of restrictions.

---

[5] Note, however, that in this specification (unlike in the case of static cointegrating regressions, considerations about (weak) exogeneity of regressors become relevant. We have not tested weak exogeneity assumptions here. Note previous findings in this respect.

The intuition is that the mean group estimator provides a consistent estimate of the average of the individual slope estimates in individual unrestricted regressions. On the other hand, under the null of common slopes, a pooled estimator that imposes those restrictions (such as the pooled mean group estimator) is consistent and efficient. A Hausman test can therefore be applied to the difference between these two sets of estimates. At conventional levels, the Hausman statistic is insignificant for the whole sample and each of the two sub-samples (although it is significant at 11% in the non-OECD country set). These findings could give weak support to the claim that the long-run vector is common, but we prefer to give higher weight to the likelihood ratio tests given the overwhelming strength of hypothesis rejection that they imply.

Inspection of the regression statistics also supports rejection of homogeneity. To save space, we summarise these results. In the unrestricted model, the error correction coefficient is correctly signed and significantly different from zero, and the diagnostic test statistics show little sign of equation misspecification in most of the dynamic models. This is much less true in the equations with common long-run coefficients imposed. The greater differences in performance occurred within the OECD sample of countries. These differences in equation performance with and without homogeneity restrictions imposed are not necessarily surprising, given that there are some very substantial differences between groups of countries like Canada, the US and Australia compared to Japan and the UK. Australia even has rising per capita emissions. In contrast, for non-OECD countries there was substantially less difference in equation performance between restricted and unrestricted cases.

In tables 6, 7 and 8 the left-hand column of statistics refers to the unrestricted model. In this case, we impose no restrictions on any of the parameters across countries, and we allow the error variances to differ across countries. In effect, there are separate regressions, and so separate sets of parameter estimates, for each country. The single point estimates in that column are what Pesaran, Shin and Smith call *mean group estimates*, each of which is the simple average of the individual country coefficient estimates. This is a consistent estimate of the mean of the individual parameters. But this averaging requires caution in interpreting estimates of turning points from the mean group

---

[6] The estimation technique is described in Pesaran, Shin and Smith (1998). These authors have made their GAUSS programme available on http:///www.econ.cam.ac.uk/faculty/pesaran.

estimates. Our rejection of the homogeneity restrictions implies that each country has a unique $\{\beta_1, \beta_2\}$ parameter pair; strictly speaking, each country has, therefore, a unique turning point in its EKC.

The *pooled mean group estimates* in the next column are derived under the null that the long-run parameters are constant over the panel but all other parameters including the speed of adjustment parameter and error variance varying over countries. The final column shows statistics from the familiar static fixed effects estimator with full homogeneity of parameters other than fixed effects, and equality of variances imposed. The last two sets of estimator results (pooled mean group and fixed effects) are statistically invalid given our rejection of long-run parameter homogeneity, and are included only to show what results occur if those misspecified models are estimated.

For the whole panel, the static fixed effects model yields a turning point at $82,746. This is not only way out of our sample range but is also approximately one order of magnitude higher than that implied by the parameter averages from the unrestricted model. Note also that for the whole panel and for the two sub-samples, imposing restrictions has a dramatic effect on the magnitude of the estimated speed of adjustment parameter. The restricted dynamic models (pooled mean group) have a speed of adjustment roughly half that given by the average of the unrestricted estimates. Neither the fixed effects nor pooled mean group estimators are properly specified models. Depending on which cointegration statistics are referred to, the unrestricted estimates may or may not be properly specified. But the variety of functional shapes means that the mean turning point cannot be taken seriously as a single global turning point. Similar caveats apply to the OECD and non-OECD estimates too.

However, the results for the OECD countries for the unrestricted and pooled mean group models differ less from the fixed effects results than is the case for the global panel. The pooled mean group turning point estimate is greater than the fixed effects estimate. In addition, the average of the individual unrestricted estimates is about twice that of the fixed effects estimate. All three of these turning points are within sample.

For the non-OECD sample, the pooled mean group turning point estimate is $28 792 vs. $116 619 for the fixed effects estimator. Both estimates are out of sample and imply a monotonic emissions-income relation. The turning point implied by the average of the unrestricted estimates is a minimum at $403. This, too, implies an essentially monotonic relation. Only Tanzania and Myanmar had income lower than this and even then only for a few years in the 1960s.

Finally, is our evidence consistent with a hypothesis that there is a monotonically increasing emissions-income relationship in the non-OECD countries and a monotonically decreasing emissions-income relationship in the OECD countries? Looking at the signs on the income parameters in Table 9, but bearing in mind that these are probably not cointegrating relationships, we conclude that even if the first part of this hypothesis is true, the latter part is not. Even among the higher-income OECD countries, there is little difference in the proportions with negative and positive emissions-income relationships.

7. CONCLUSIONS

In this paper, we have focussed on five main issues. Our findings with respect to each are briefly listed below:

(a) Does it make an important difference to estimation and inference techniques if the data are non-stationary rather than stationary? The answer to this is unambiguously in the affirmative.

(b) Are the data in this study stochastically non-stationary? They clearly are. We also conjecture that many other indices of environmental pressure are also integrated variables.

(c) Are there cointegrating relations between sulphur emissions per capita and income per capita in individual countries? Whilst our findings are somewhat mixed, there is weak support for the contention that such cointegrating relations do exist in the panel as a whole.

(d) Are the cointegrating relationships that are found consistent with the so-called EKC hypothesis? We find that for many countries they are not. A large minority of countries has basic shapes of emission/income relationships that do not have the EKC form.

(e) Do all individual members converge to a common cointegrating vector, given that some (possibly heterogeneous) cointegrating relationship exists for all individuals? Matters are rather less clear here, and depend on the relative weight one wishes to attach to likelihood ratio versus Hausman tests, but on balance, the evidence seems to reject the existence of a common vector.

(f) Is dynamic as opposed to static modelling important? We conclude that it is. Static regressions such as simple fixed or random effects are badly misspecified, and are inappropriate for statistical inference (because of (a) and (b) above).

REFERENCES

Ansuategi A., E. B. Barbier and C. A. Perrings (1998), 'The environmental Kuznets curve', in J. C. J. M. van den Bergh and M. W. Hofkes (eds) Theory and Implementation of Economic Models for Sustainable Development, Dordrecht: Kluwer.

Arrow, K., Bolin, B., Costanza, R., Dasgupta, P., Folke, C., Holling, C.S., Jansson, B.O., Levin, S., Maler, K.G., Perrings, C. and Pimental , D. (1995): Economic Growth, Carrying Capacity, and the Environment. Science, 269, pages 520-21.

Banerjee, A., Dolado, J., Galbraith, J.W. and Hendry, D.F. (1993): Co-integration, Error Correction, and the Econometric Analysis of Non-Stationary Data. Oxford University Press, Oxford.

Beckerman, W. (1992): Economic Growth and the Environment. Whose growth? Whose Environment? World Development 20, 481-496.

Campbell, J.Y. and Perron, P. (1991): Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots. NBER Macroeconomics Annual. MIT Press, Cambridge, Mass.

Cole M.A., A.J. Rayner, and J.M. Bates, The environmental Kuznets curve: an empirical analysis, *Environment and Development Economics* **2**(4), 401-416 (1997).

de Bruyn S. M., Explaining the environmental Kuznets curve: structural change and international agreements in reducing sulphur emissions, *Environment and Development Economics* **2**(4), 485-503 (1997).

de Bruyn S. M., J. C. J. M. van den Bergh, and J. B. Opschoor, Economic growth and emissions: reconsidering the empirical basis of environmental Kuznets curves, *Ecological Economics* **25**, 161-175 (1998).

Dijkgraaf, E. and Vollebergh, H.R.J. (1998): The Environmental Kuznets Curve Revisited. Erasmus University, Rotterdam, Discussion Paper.

Ekins P. (1997), 'The Kuznets curve for the environment and economic growth: examining the evidence', Environment and Planning A., 29, 805-830.

Grossman G. M. and A. B. Krueger, *Environmental Impacts of a North American Free Trade Agreement*, National Bureau of Economic Research Working Paper 3914, NBER, Cambridge MA (1991).

Hall, A. (1991): Testing for a Unit Root in Time Series with Pretest Data Based Model Selection. Department of Economics, North Carolina State University. Mimeo.

Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press, Princeton, New Jersey.

Holden and Perman (1994) "Units Roots and Cointegration: An Introduction", in Rao, W.W. (ed) Cointegration: Expository Essays for the Applied Economist. Macmillan.

Holtz-Eakin D. and T. M. Selden (1995), 'Stoking the fires? CO_2 emissions and economic growth', Journal of Public Economics, 57, 85-101.

Im, Pesaran and Shin (1996): Testing for Unit Roots in Heterogeneous Panels, Discussion Paper, University of Cambridge. (Revised version available, 1997).

John A. and R. Pecchenino (1994) An overlapping generations model of growth and the environment, *Economic Journal* 104, 1393-1410.

John A., R. Pecchenino, D. Schimmelpfennig, and S. Schreft (1995) Short-lived agents and the long-lived environment, *Journal of Public Economics* 58, 127-141.

Kao, C. (1997): Spurious Regression and Residual-Based Tests for Cointegration in Panel Data. Working Paper, Department of Economics, Syracuse University. [Revised December 1997].

Kaufmann R. K., B. Davidsdottir, S. Garnham, and P. Pauly, The determinants of atmospheric SO_2 concentrations: reconsidering the environmental Kuznets curve, *Ecological Economics* **25**, 209-220 (1998).

Kuznets, S. (1955): Economic Growth and Income Inequality. American Economic Review, 49, pages 1-28.

Levin, A. and Lin, C. (1993): Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. University of California, San Diego, Discussion Paper.

Lopez R. (1994) The environment as a factor of production: the effects of economic growth and trade liberalization, *Journal of Environmental Economics and Management* 27, 163-184.

Maddala, G.S. and Kim, I-M. (1998): Unit Roots, Cointegration and Structural Change. Cambridge University Press, Cambridge.

McConnell K. E. (1997) Income and the demand for environmental quality, *Environment and Development Economics* 2(4).

Moomaw W. R. and G. C. Unruh (1997) Are environmental Kuznets curves misleading us? The case of CO2 emissions, *Environment and Development Economics* **2**(4), 451-463

Panayotou T., "Empirical Tests and Policy Analysis of Environmental Degradation at Different Stages of Economic Development", Working Paper WP238, Technology and Employment Programme, International Labour Office, Geneva (1993).

Panayotou T., Demystifying the environmental Kuznets curve: turning a black box into a policy tool , *Environment and Development Economics* **2**(4), 465-484 (1997).

Panayotou T., Environmental degradation at different stages of economic development, in: "Beyond Rio: The Environmental Crisis and Sustainable Livelihoods in the Third World" (I. Ahmed and J. A. Doeleman, Eds.), MacMillan, London (1995).

Pedroni, P. (1988) Approximate Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors . UCSD Working Paper, forthcoming Oxford Bulletin of Economics and Statistics.

Pedroni, P. (1995): Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests With an Application to the PPP Hypothesis. Indiana University Working Papers in Economics, No 95-013.

Pedroni, P. (1997): Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests With an Application to the PPP Hypothesis: New Results. Indiana University Working Paper in Economics, Revised April 1997.

Pesaran, M.H., Shin, Y. and Smith , Ron P. (1998): Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. Dept of Economics, University of Cambridge, Mimeo, August 1997, Revised November 1998.

Quah, D. (1994): Exploiting Cross-Section Variation for Unit Root Inference in Dynamic Data. Economic Letters, 18, pages 381-386.

Roberts, J.T. and Grimes, P.E. (1997): Carbon Intensity and Economic Development 1962-91. World Development, 25(2), pages 191-198.

Selden T. M. and D. Song (1995) Neoclassical growth, the J curve for abatement and the inverted U curve for pollution, *Journal of Environmental Economics and Environmental Management* 29, 162-168.

Selden T. M. and D. Song, Environmental quality and development: Is there a Kuznets curve for air pollution? *Journal of Environmental Economics and Environmental Management* **27**, 147-162 (1994).

Shafik N., Economic development and environmental quality: an econometric analysis, *Oxford Economic Papers* **46**, 757-773 (1994).

Shafik, N. and Bandyopadhyay, S. (1992): Economic Growth and Environmental Quality. Journal of Environmental Economics and Management, 27, pages 147-62.

Shukla V. and K. Parikh, The environmental consequences of urban growth: cross-national perspectives on economic development, air pollution, and city size, *Urban Geography* **12**, 422-449 (1992).

Stern D. I. (1998) Progress on the environmental Kuznets curve? *Environment and Development Economics* 3, 173-196.

Stern, D.I. and Common, M.S. (1998): Is there an Environmental Kuznets Curve for Sulfur? Manuscript, CRES, Australian National University.

Torras M. and J. K. Boyce, Income, inequality, and pollution: a reassessment of the environmental Kuznets curve, *Ecological Economics* **25**, 147-160 (1998).

Vincent J. R., Testing for environmental Kuznets curves within a developing country, *Environment and Development Economics* **2**(4), 417-431 (1997).

TABLE 1: UNIVARIATE UNIT ROOT TEST STATISTICS: LN (Y/P) [(COLUMNS Y1, Y2 AND Y3] AND LN(M/P) [(COLUMNS M1, M2 AND M3].

Country	ln(Y/P) ADF t statistic from model without time trend	ln(Y/P) ADF t statistics from model with time trend	ln(Y/P) Full search decision	ln(M/P) ADF t statistic from model without time trend	ln(M/P) ADF t statistic from model with time trend	ln (M/P) Full search decision
	Y1	Y2	Y3	M1	M2	M3
Canada	-0.06		RW	-2.91 *		RW
U.S.A.		-0.76	RW	-1.90 *		RW
Japan	-2.77*		Stationary	-1.79 *		RW
Austria	2.49	0.07	RW	-2.05 *		RW
Belgium	1.12		RW	-1.55		RW
Denmark	-1.26		RW	-1.66 *		RW
Finland	0.18		RW	-1.95 *		RW
France		-1.54	RW (?)	-0.14		RW
W. Germany	0.86		RW	-1.26		RW
Ireland	2.07		RW	-2.19 *		RW
Italy	-0.75		RW (?)	-0.59		RW
Luxembourg		-0.98	RW (?)	-2.92 *		RW
Netherlands	-1.85 *		RW	-1.85 *		RW
Norway	-0.49		RW	-1.62		RW
Spain	-1.01		RW	-0.47		RW
Sweden		-1.37	Stationary	-0.63		RW
Switzerland	-2.25 *		RW	-0.47		RW
U.K.	-1.02		RW		-3.22 *	Stationary
Australia	-1.80 *		RW		-1.41	Stationary ?
N.Z.	-1.51		RW	-1.34		RW
Greece		-2.17 *	Stationary	-0.03		RW
Portugal	0.46		RW	2.88		RW
Turkey	-1.04		RW	0.45		RW
Algeria	-2.76 *		RW	-1.54		RW
Egypt	-1.34		RW	-0.97		RW
Ghana	-1.42		RW	-2.91 *		RW
Kenya	-3.58 *		RW		-2.39 *	Stationary
Madagascar	-0.54		RW	-1.69 *		RW
Morocco	-2.94 *		RW	-1.05		RW
Mozambique	0.25		RW	0.51		RW ?

Namibia	-1.07		RW	-0.70		RW
Nigeria	-1.76 *		RW	-3.68 *		Stationary
Safrica	0.30		RW	0.06		RW
Tanzania	-1.89 *		RW	-2.51 *		RW
Tunisia	-0.09		RW	-1.81 *		RW
Zaire	0.20		RW	-0.42		RW
Zambia	0.36		RW	-0.19		RW
Zimbabwe	-0.45		RW	-2.28 *		RW
Barbados	-2.10*		Stationary ?	-2.29 *		RW
Guatemala	0.89		RW	-2.33 *		RW
Honduras	-0.92		RW	-2.98 *		RW
Mexico	-2.84 *		RW	-0.86		RW
Nicaragua	0.68		RW	-0.39		RW ?
Trinidad	-0.91		RW	-0.34		RW
Argentina	1.70		RW	-2.67 *		RW
Bolivia	-0.55		RW	-1.37		RW
Brazil	-0.96		RW	-1.06		RW
Chile	-1.36		RW	-1.84 *		RW
Colombia	-1.00		RW	-1.95 *		RW
Peru	0.92		RW	-2.65 *		RW
Uruguay		-3.05 *	Stationary	-1.09		RW
Venezuela	0.38		RW	-2.43 *		RW
China	-0.67		RW		-5.66 * **	Stationary
Hong Kong	0.21		RW	-1.46		RW
India		-1.47	Stationary	-0.47		RW
Indonesia		-2.39 *	RW ?	-0.16		RW
Korea	2.47		RW	-0.41		RW
Malaysia	0.59		RW	-3.23 *		RW
Myanmar	-1.95 *		RW	-1.48		RW
Philippines	-2.03 *		RW	-1.84 *		RW
Singapore	0.38		RW	-2.68 *		RW
Sri Lanka	-1.87 *		RW		-2.85 *	Stationary
Taiwan	3.40		RW	0.35		RW
Thailand	2.66		RW	-1.47		RW ?
Cyprus	-0.57		RW	-1.79 *		RW
Czechoslov.	-1.22		RW		-1.92 *	Stationary
Romania	-1.28		RW	-0.96		RW
USSR	-1.44		RW	-2.71 *		RW
Yugoslavia	-1.62		RW	0.91		RW
Iran	-1.20		RW	-2.66 *		RW
Israel	-2.07 *		RW	-3.12 *		RW
Kuwait	1.07		RW	-2.61 *		RW
Saudi Arabia	-1.46		RW		-2.45 *	Stationary ?

Syria	-1.43		RW	-1.45		RW
	15 rejects of unit root at 5%	3 rejects of unit root at 5%	6 stationary + 4 borderline	31 rejects of unit root at 5%	6 rejects of unit root at 5%	6 stationary + 2 borderline

Notes to Table 1:

All statistics are based on models with heterogeneous lags, with a maximum lag truncation of 5 lags. Appropriate lag lengths for each series for each country are obtained using the Hall (1991) method. In drawing inference from the unit root tests, ‘large’ negative values imply rejection of the null of a single unit root in the series. [Separate unit root tests, not reported here, indicate that none of the series was integrated of an order higher than one.]

Columns Y1 and M1: (ADF regression with constant but no trend; corrected for time means): * denotes unit root rejected at 5% level, using the standard normal distribution (critical value [cv] = -1.645) . Distribution is standard normal under true null of random walk + drift (cv = -1.645). The distribution is non-standard DF (cv = - 2.95) under true null of pure random walk without intercept which (as it implies zero asymptotic means) is not applicable to these datasets.

Columns Y2 and M2 (regression with constant and trend, corrected for time means): If the series truly did not contain a deterministic trend, relevant distribution would be non-standard DF (cv = -3.55). However, all entries in these columns are found to include a trend term. In that case, * denotes rejection at 5% under assumption that there is a significant deterministic trend, in which case distribution is standard normal.

A ? symbol adjacent to a finding (RW or stationary) indicates that the inference is either borderline (in the sense of being very close to the accept/reject null margin) or not robust (in the sense that the sequential search procedure did not lead to an unambiguous outcome).

TABLE 2: FULL PANEL UNIT ROOT ADF T-TYPE TEST STATISTICS

ln(Y/P)		
Panel: Regression without trends	8.93	Do not reject unit root null
Panel: Regression with trends	4.34	Do not reject unit root null
Group: Without trends and without common time dummies	1.41	Do not reject unit root null
Group: Without trends but with time dummies	8.86	Do not reject unit root null
Group: With trends and common time dummies	1.27	Do not reject unit root null
In(M/P)		
Panel: Regression without trends	1.23	Do not reject unit root null
Panel: Regression with trends	1.41	Do not reject unit root null
Group: Without trends and without common time dummies	-0.05	Do not reject unit root null
Group: Without trends but with time dummies	-0.23	Do not reject unit root null
Group: With trends and common time dummies	-2.51	Reject unit root null
YOP2		
Panel: Regression without trends	9.23	Do not reject unit root null
Panel: Regression with trends	3.97	Do not reject unit root null
Group: Without trends and without common time dummies	2.67	Do not reject unit root null
Group: Without trends but with time dummies	9.27	Do not reject unit root null
Group: With trends and common time dummies	0.91	Do not reject unit root null

TABLE 3: SIGNIFICANT ADF STATISTICS IN INDIVIDUAL REGRESSIONS (N=74, T=31):

Model used:	Proportion of ADF t statistics significant at 10% or better:	
	Quadratic	Linear
Case 1: No trends and no time dummies	17/74	3/74
Case 2: No trends but with time dummies	16/74	10/74
Case 3: With trends and time dummies	31/74	14/74
Countries with cointegration in all 3 models	7/74	1/74
Countries with cointegration in 2 models	11/74	9/74
Countries with cointegration in 1 model only	21/74	6/74
Countries with cointegration in no model	35/74	58/74

Note: All models estimated with heterogeneously chosen lags up to a maximum of 3 lags.

TABLE 4: INDIVIDUAL EKC REGRESSIONS : INCLUDING TIME DUMMIES AND WITH AND WITHOUT HETEROGENOUS TIME TRENDS.

	$\beta_1 > 0$ and $\beta_2 < 0$ (Inverted U shaped emissions/income relationship)	$\beta_1 < 0$ and $\beta_2 > 0$ (U shaped emissions/income relationship)	$\beta_1 > 0$ and $\beta_2 > 0$ or $\beta_1 < 0$ and $\beta_2 < 0$
Model:			
With time dummies but no trend	42/74	26/74	6/74
With time dummies and trend	34/74	36/74	4/74

Notes to Table 4: The models we estimated did not incorporate time dummy variables in the form specified in Cases 2 and 3 above. Given the span of our data, this would have made estimation infeasible. Instead, we have adopted the common practice of controlling for common time effects by demeaning the data by subtracting sample averages taken over the time dimension. This is identical to using time dummies in the case where long-run coefficients are homogenous over countries. Where they are not (as in our sample), this procedure will remove most but not all of the common time effects (see Pesaran, Shin and Smith, 1998). Given computation limitations, this is the best we can do.

TABLE 5(a): PANEL COINTEGRATION TEST STATISTICS: QUADRATIC MODELS

	'Panel' statistics				'Group' statistics		
	Panel V	Panel RHO	Panel PP	Panel ADF	Group RHO	Group PP	Group ADF
MODEL:							
No time dummies or trends	0.64	0.95	-0.95	-0.57	0.93	-2.58**	-3.72**
Time dummies included but not tends	2.37**	-0.02	-2.16*	-1.87*	1.06	-2.83 **	-4.44**
Both trends and time dummies included	1.23	-0.004	-3.90**	-5.33**	1.68	-3.80 **	-8.84**

TABLE 5(b): PANEL COINTEGRATION TEST STATISTICS: LINEAR MODELS

	'Panel' statistics				'Group' statistics		
	Panel V	Panel RHO	Panel PP	Panel ADF	Group RHO	Group PP	Group ADF
MODEL:							
No time dummies or trends	0.89	0.28	-0.97	0.34	0.94	-1.37	-0.01
Time dummies included but not tends	0.05	-0.13	-1.80*	-1.45	1.11	-1.92 *	-2.32**
Both trends and time dummies included	-0.78	1.33	-1.66*	-1.63	2.17*	-1.76 *	-4.34**

Notes to Tables 5(a) and (b):

All regressions were run with individually chosen lag lengths, from a maximum lag of 3[i]

Panel V denotes a non-parametric variance ratio statistic; Panel (or group) RHO is a non-parametric test statistic analogous to the Phillips and Perron (PP) rho statistic. PP denotes a non-parametric statistic analogous to the PP t statistic. ADF denotes a parametric statistic analogous to the augmented Dickey-Fuller statistic.

All 7 statistics are standardised so as to be distributed as standard normal as T and N grow large. Rejection of the null of no cointegration is one-sided and involves:

- variance ratio ; large positive values imply cointegration (at 5% significance, reject null of no cointegration if V > 1.645)

- other six ; large negative values imply cointegration (at 5% significance, reject null of no cointegration if statistic < -1.645)

* Denotes that the null of no cointegration is rejected at the 5% level

** Denotes that the null of no cointegration is rejected at the 1% level

TABLE 6: WHOLE PANEL

	Estimator		
	Unrestricted model (Mean group parameter estimator)	Pooled mean group estimator (homogeneous long-run coefficients)	Static fixed effects
Long-run parameter estimates:			
ln(Y/P)	16.56 (1.30)	8.786 (19.18)	3.85 (3.67)
in$(Y/P)^2$	-0.89 (-1.30)	-0.48 (-17.63)	-0.17 (-2.70)
Error correction:	-0.366 (-12.7)	-0.228 (-7.00)	na
Implied turning point	$ 10,974	$ 9,434	$82,746
Hausman		0.37 (p = 0.83)	
lnL	1512.89	1828.19 (p = 0.00)	-1805.23

Notes:
t ratios in parentheses
t ratios for fixed effects based on robust standard errors

TABLE 7: OECD COUNTRIES ONLY

	Estimator		
	Unrestricted model (Mean group parameter estimator)	Pooled mean group estimator (homogeneous long-run coefficients)	Static fixed effects
Long-run parameter estimates:			
ln(Y/P)	19.78 (2.11)	34.59 (13.31)	12.84 (4.73)
in$(Y/P)^2$	-1.02 (-1.91)	-1.85 (-12.65)	-0.71 (-4.62)
Error correction:	-0.300 (-5.57)	-0.163 (-3.16)	na
Implied turning point	$ 16253	$ 11483	$ 8452
Hausman		2.71 (p = 0.26)	
lnL	923.22	843.00 (p = 0.00)	- 8.21

Notes:
t ratios in parentheses
t ratios for fixed effects based on robust standard errors

TABLE 8: NON-OECD COUNTRIES ONLY

	Estimator		
	Unrestricted model (Mean group parameter estimator)	Pooled mean group estimator (homogeneous long-run coefficients)	Static fixed effects
Long-run parameter estimates:			
ln(Y/P)	-1.56 (-0.23)	5.75 (11.91)	3.50 (2.73)
in$(Y/P)^2$	0.13 (0.31)	-0.28 (-9.95)	-0.15 (-1.89)
Error correction:	-0.331 (-11.09)	-0.221 (-7.40)	na
Implied turning point	$ 403 minimum point	$ 28792 maximum	$ 116619 maximum
Hausman		4.51 (p = 0.11)	
lnL	922.32	751.56 (p = 0.00)	- 1464.01

Notes to Table 8:
t ratios in parentheses
t ratios for fixed effects based on robust standard errors

Panel cointegration test statistics - specifically those which are fully parametric in form (here the panel and group ADF statistics) - can be highly sensitive to the value chosen for the maximum lag lengths allowed prior to the selection of the appropriate lag for each regression. Selection of an excessively large truncation length ("overfitting") leads to misleadingly high absolute values of the parametric test statistics. We used a general to specific search procedure, beginning with a maximum truncation of 6 lags, to obtain the appropriate initial truncation (here found to be three lags). This procedure is not required in the process of generating cointegration test statistics for individual countries, in which the appropriate lag length was optimally chosen (from a maximum of 5) on a country-by-country basis.

TABLE 9: SIGNS ON INCOME PARAMETER IN ESTIMATED LINEAR RELATIONSHIPS

OECD		Non-OECD	
$\beta_1 > 0$	$\beta_1 < 0$	$\beta_1 > 0$	$\beta_1 < 0$
13/23	10/23	38/51	13/51

Figure 1. Relation between Parameters in Individual Country Regressions