ECOLOGICAL ECONOMICS 62 (2007) 648–660

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/ecolecon

ANALYSIS

Exploring the environmental Kuznets hypothesis: Theoretical and econometric problems☆

Georg Müller-Fürstenberger[a], *Martin Wagner*[b,*]

[a]*University of Bern, Department of Economics, Switzerland*
[b]*Institute for Advanced Studies, Vienna, Austria*

A R T I C L E   I N F O

Article history:
Received 18 April 2006
Accepted 10 August 2006
Available online 6 October 2006

Keywords:
Carbon Kuznets curve
Integrated assessment models
Regressions with integrated variables
Nonstationary panels

JEL classification:
C12; C13; Q20

A B S T R A C T

Focussing on the prime example of CO_2 emissions, we discuss several important theoretical and econometric problems that arise when studying environmental Kuznets curves (EKCs). An EKC refers to an inverse U-shaped relationship between a pollutant and economic activity. The dominant theoretical approach is given by integrated assessment modelling, which consists of economic models that are combined with environmental impact models. We critically evaluate the aggregation, model dynamics and calibration aspects and their implications for the validity of the results. We then turn to a discussion of several important econometric problems that go almost unnoticed in the literature. The most fundamental problems relate to nonlinear transformations of nonstationary regressors and, in a nonstationary panel context, to neglected cross-sectional dependence. We discuss the implications of these two major and some minor problems that arise in the econometric analysis of Kuznets curves. Our discussion shows that EKC modelling as performed to date is subject to major drawbacks at both the theoretical and the econometric level.

1. Introduction

More than 80% of the world's current primary energy demand is met by fossil fuels (see Energy Information Administration, 2004). Their use yields carbon dioxide – CO_2 – as a joint product, which once released into the atmosphere contributes to climatic change with potentially irreversible negative impacts on the world economy. A key issue in environmental economics is to project these man-made CO_2 emissions for a given scenario describing inter alia population growth and technological progress.

This issue closely relates to the 'environmental Kuznets curve' (EKC) discussion, which investigates the quantitative relation between per capita emissions of some pollutants and economic activity. In this paper we focus on carbon dioxide emissions, hence we address this relation as the 'carbon Kuznets curve' (CKC). Our discussion of theoretical and econometric problems, however, applies to other pollutants as well.

The CKC hypothesis refers to an inverse U-shaped relationship between economic activity, usually measured in terms of per capita GDP, and per capita CO_2 emissions. Thus, it conjectures emissions to first rise with growing GDP, to pass through a peak at a certain income level and to decline afterwards with income increasing further, for example because the willingness to pay for environmental quality increases with

☆ We gratefully appreciate the helpful comments of two anonymous reviewers and the editor, which have led to substantial improvements of the paper. Furthermore, the comments of Gregor Bäurle, Robert Kunst, Klaus Neusser and Reto Tanner are gratefully acknowledged. We thank Benito Müller for providing the data used in this study and for discussions on the topic. The usual disclaimer applies.

* *Corresponding author*. Institute for Advanced Studies, Stumpergasse 56, A-1060 Wien, Austria. Tel.: +43 1 59991 150; fax: +43 1 59991 163.
E-mail address: Martin.Wagner@ihs.ac.at (M. Wagner).

0921-8009/$ - see front matter
doi:10.1016/j.ecolecon.2006.08.005

income. The reference to Kuznets is reminiscent of Kuznets (1955), who postulated an inverse U-shaped relationship between the level of economic development and the degree of income inequality in his presidential address to the American Economic Association in 1954.

This paper contributes to the CKC discussion by critically reviewing two main strands of analyzing the GDP-CO_2 emissions relationship in the economic literature. These are the computable general equilibrium (CGE) approach on the one hand and the econometric approach on the other. Of course, both approaches overlap to a certain extent, but for the sake of illustration we essentially separate the discussion in two sections.[1]

The CGE approach consists of a fully specified general equilibrium model which is calibrated on economic data of the real-world economy under consideration. Adding a carbon cycle model and a climate change impact model results in an Integrated Assessment Model (IAM) which can account for the feed-back effects of CO_2 emissions on economic activity and welfare. Section 2 discusses key choices when setting up an IAM. These include the choice between a *bottom-up* or *top-down* approach and choices with respect to equilibrium concepts, functional specifications and calibration. We show that equally reasonable assumptions concerning the discount rate or the rate of technological change lead to substantially different conclusions concerning the CKC hypothesis. Thus, the CGE approach is seen to be subject to large uncertainties that are usually not discussed in the corresponding literature.

We continue by using our prototype IAM to generate data on which we 'estimate' a CKC. In one of the scenarios we find an inverse U-shaped relationship between GDP and emissions. However, this effect is solely driven by exogenous technical progress underlying that scenario; it is not due to increasing willingness to pay for environmental quality at higher income levels. This illustrates the pitfalls that arise when attaching structural interpretations to reduced form relationships and also illustrates the danger associated with CGE models in general: *confabulation*. This term refers to the fact that CGE model results are often interpreted in ways that do not correspond to the mechanisms present in the model.

In Section 3 we focus on the single-equation econometric approach to estimate the CKC relationship. There is a huge literature applying time series and panel data techniques to estimate the relationship between GDP and pollutants.[2] Surprisingly, several important econometric problems have largely gone unnoticed in the empirical literature. The most fundamental problems relate to nonlinear transformations of nonstationary regressors, and in a panel context additionally to the effects of cross-sectional dependence. The implications of these major problems and some additional minor problems are discussed in detail. The companion paper Wagner (2005) contains a detailed discussion of as well as a potential solution to some important econometric problems arising in EKC analysis when using time series or panel data. In this paper we merely want to highlight the problems and want to indicate that a more thoughtful application of standard econometric techniques should lead researchers to be more cautious about their findings than what is commonly observed. Section 4 briefly concludes and summarizes the paper.

2. A CGE simulation of the carbon Kuznets curve

Historically, CGE models in climate change economics originate from detailed energy technology assessment models, like ETA-Macro (see Manne, 1977), which were designed to analyze energy-economy interactions. Assigning carbon content coeff-cients to different types of fossil fuels allows these models to simulate carbon emissions along economic growth and structural change paths. 'Integrated Assessment Models' go one step ahead by adding two sub-models. These are a carbon cycle model and a climate change impact model. The carbon cycle model computes the atmospheric concentration path of CO_2 in parts per million by volume (ppmv) along a global CO_2 emissions path. The climate change impact model – which often reduces to a single 'damage function' – translates atmospheric carbon concentrations via changes of mean surface temperatures into economic damages. The two sub-models establish a feed-back loop from carbon emissions to economic damages. Due to the long-term nature of the climate problem, Integrated Assessment Models typically involve time-horizons of more than hundred years.

The CGE based integrated assessment was pioneered by Nordhaus (1992) with the DICE model and by Manne, Mendelsohn and Richels (1995) with MERGE. Variations and extensions of these and related types of models define the current state of the art in CGE based climate economics.

The purpose of such models is threefold. First, they can identify optimal carbon emission paths by weighting the benefits of avoided climate damages against abatement costs. Second, they serve to define a baseline scenario, which projects future atmospheric carbon concentrations under business-as-usual (BAU) assumptions. In the BAU case, climate change is a negative externality which unfolds without policy intervention. Third, these models are used to quantify costs and benefits of climate policy interventions. The BAU case is of prime importance since it serves as yard stick to measure costs and benefits of policy interventions. Moreover, the relation between GDP and CO_2 emissions in the BAU scenario is closely related to the CKC issue and the related question whether the climate change problem relaxes when the world economy gets richer. For these reasons, our focus in this paper is on BAU scenario simulation.

Defining a BAU scenario requires a considerable number of assumptions reflecting the modelers' expectations about the future evolution of the economy. Key assumptions include population growth rates and rates of technological progress. The choices necessary to define the BAU case are generally recognized as a key issue in CGE modelling, see Intergovernmental Panel on Climate Change (2001). Nevertheless, there

[1] For example McKibbin, Ross, Shackleton and Wilcoxen (1999) present a CGE model where some key elasticities are estimated from time series. In this sense, they combine econometric analysis and CGE modelling. For a recent survey on the relationship between economic growth and the environment see Brock and Taylor (2005), who present several theoretical models as well as empirical evidence.

[2] Thus, note in particular that our discussion of the econometric problems illustrated in Section 3 applies to the econometric environmental Kuznets curve (EKC) literature in general.

are large and arbitrary variations among different models and no consensus seems to be reached.

We show in the discussion below that different assumptions concerning the rate of technological progress lead to the presence, respectively absence of an inverse U-shaped relationship between GDP and CO_2 emissions for our model. When it occurs, the inverse U-shape is driven by exogenous technological improvements and technology diffusion. In the model thus the inverse U-shape is not at all due to increased willingness to pay for environmental quality at higher income levels and therefore growth enhancing policies do not lead to decreasing per capita emissions (due to exogenous carbon intensities).

2.1. Some general remarks on the CGE approach

There are two fundamental decisions at the outset of any CGE study. The first decision is about the level of aggregation, the second about the model dynamics. Let us start with the level of aggregation. Highly aggregated models are usually termed *top-down* models. They assume a highly aggregated macroeconomic production function with a single consumption–investment commodity as output at the top of the production structure. Additional structure is added by nesting production sub-sets into the macro production function. The case in point for a top-down approach is the DICE model.

The *bottom-up* approach starts from a detailed description of the economy, in particular with respect to production sectors and energy transformation technologies. Examples are IGSM of Prinn et al. (1999), GEMINI-E3 of Bernard and Vielle (1998) or WIAGEM of Kemfert (2002). The detailed modelling of sectoral production structures and the engineering based descriptions of key technologies qualifies these down-to-earth models to assess the structural change and choice-of-technology impacts of environmental policies. Both approaches, however, overlap as it is possible to endow production sub-sets in a top-down approach with a detailed and technically backed fundament.

Thus, by construction the top-down approach is not suitable for analyzing substitution effects between competing technologies and intersectoral adjustments. These issues can only be addressed by custom-made bottom-up models. Therefore, the choice between the approaches is dictated by the specific question at hand.

The second choice is about model dynamics, in particular with respect to intertemporal decision making. This issue often directly relates to the aggregation issue since there is a trade-off between a low level of aggregation and a sophisticated intertemporal decision and equilibrium concept. To keep highly disaggregated models computationally tractable, bottom-up approaches are usually not formulated as perfect foresight or rational expectations models but as a sequence of temporary equilibria where growth is driven by a fixed savings rate, for example. Hence they cannot be used to solve for optimal capital accumulation paths or for the optimal timing of environmental expenditures and adaptation measures.

Top-down models explicitly model capital accumulation and climate policy as results of maximizing intertemporal welfare. Tractability of the dynamic problem necessitates a high degree of aggregation in usually fewer regions and with a smaller number of sectors than in bottom-up models. A prime example is the version of MERGE published by Manne (1999).

IAMs produce a large amount of output whose correct interpretation requires a profound understanding of the underlying model. Such model results are therefore vulnerable to *confabulation*,[3] which means that CGE modelers tend to provide ‘economic’ intuitions for their model results which are completely unrelated to what actually happens in the model. In the sequel we present a scenario where an inverse U-shape occurs in the simulated data. However, as already mentioned, it is not consistent with the model to attribute this to increasing willingness to pay for environmental quality as income grows.

Next, we present a prototypical top-down model to highlight the most important decisions in such a modelling process by a concrete example. Illustrating the problems with a small scale model corresponds to focusing on single-equation econometric analysis of the CKC in Section 3. It is important to note that similar problems as discussed in this paper for top-down models and single-equation econometric analysis are even more prevalent in bottom-up modelling or when econometrically specifying systems of equations to study the nonlinear relationships between economic activity and pollutants.

2.2. A simple top-down model

For simplicity we consider only two regions, which we call North and South, indexed $i = N, S$, and assume discrete time $t = 0, 1, \ldots, \infty$. For the numerical implementation, North is thought of as comprising all members of the OECD in 1990. South covers the remaining countries of the 107 countries listed in Table 4 in the Appendix.[4] The world economy is set up as a one-good two-region Ramsey-type growth model. This simplification allows to simulate a fully dynamic model.

It is assumed that the production decisions are carried out under perfect competition. Consumption and investment decisions are considered as if a representative agent maximizes intertemporal utility, given a complete set of present value prices for the consumption/investment good and factor prices. Such a complete system of markets is a fiction with respect to real world economies, even for the most industrialized countries in what we call North. This strong assumption with respect to the completeness and competitiveness of the market system is generally not questioned in dynamic CGE modelling. The problem in this respect is that theories of incomplete markets (see e.g. Magill and Quinzii, 1996) are not yet developed sufficiently to allow for implementation within CGE models. Limitations in this direction are a clear lacuna in dynamic CGE modelling.

To transform a general equilibrium model into a computable one, it is necessary to specify functional forms for production, utility etc. With respect to production it is common to adopt a constant elasticity of substitution (CES) function; it is easily

[3] This term was used by R.A. McDougall in an introductory note on a CGE course at Purdue University.

[4] Thus, North and South are not to be taken literally in a geographic sense but North represents the developed countries and South the less developed countries.

calibrated, see below, and it comprises both the Cobb–Douglas and the Leontief production functions as special cases. In our example, we describe regional production by a nested CES aggregator f_i, with physical capital k_{it}, labor l_{it}, and energy g_{it} as production factors:

$$f_i(k_{it}, l_{it}, g_{it}) = (a_i^1(l_{it}^{\vartheta_i} k_{it}^{1-\vartheta_i})^{\tau_i} + a_i^2 g_{it}^{\tau_i})^{\frac{1}{\tau_i}}, \tag{1}$$

where a_i^1 and a_i^2 are factor productivities, ϑ_i is a technical parameter determining the value share of labor in value addition, and τ_i relates to the elasticity of substitution between value added and energy. This formulation implies that there are no endogenous changes in the production technology, which would be the case if the factor productivities a_i were subject to endogenous technological progress. It is thus assumed impossible to foster technological progress by policy intervention.

Production output is spent on consumption, investment, energy production or to fix climate damages:

$$y_{it} = f_i(k_{it}, l_{it}, g_{it}) = c_{it} + i_{it} + g_{it} + \theta_{it} y_{it}. \tag{2}$$

We denote gross output by y_{it}, consumption by c_{it} and investment by i_{it}. The marginal costs of energy supply are constant and normalized to one. Climate damages in terms of % losses of gross output are given by θ_{it}, specified below in Eq. (6).

Note that by rearranging Eq. (2)

$$(1-\theta_{it}) f_i(k_{it}, l_{it}, g_{it}) = c_{it} + i_{it} + g_{it}, \tag{3}$$

climate damages can be interpreted as negatively affecting total factor productivity. This is common practice in the literature, for example in the MERGE and DICE models. This rests, however, on several implicit assumptions; for example that climate damages have no impact upon the marginal rate of substitution between production factors.

Capital accumulates according to

$$k_{it+1} = (1-\delta) k_{it} + \omega i_{it}, \tag{4}$$

with δ as the depreciation rate, and ω as the linear production coefficient in the investment technology. Concerning the accumulation of atmospheric carbon we rely upon the widely used Nordhaus (1991) equation

$$S_{t+1} = \phi_1 \sum_i \eta_{it} g_{it} + \phi_2 S_t, \tag{5}$$

where ϕ_1 and ϕ_2 are climate system parameters. ϕ_2 is the natural decay rate of atmospheric carbon dioxide and ϕ_1 is estimated in Nordhaus (1991) by OLS regression and provides a good fit of historic data. The coefficients η_{it} are emissions coefficients describing the emission intensity of energy production. It must be kept in mind, however, that the Nordhaus equation violates physical and chemical principles, which matters in case of large perturbations of the climate system, see Joos, Müller-Fürstenberger and Stephan (1999). The accuracy of the approximation deteriorates rapidly when the actual carbon emission path deviates substantially from the one generating the data on which the parameter ϕ_1 is estimated. This means that both the BAU as well as the policy intervention paths have to be in the vicinity of the Nordhaus (1991) path. For emission paths that deviate by a large amount from this reference, the results may be highly misleading, since the carbon accumulation and hence climate damage paths are modelled with potentially large biases. A more detailed and robust but still tractable carbon cycle model may be important for integrated assessment modelling, in particular for cases where emissions keep growing rapidly.

Accumulated atmospheric carbon S_t translates into damages according to a so called damage function. A common formulation is given by

$$\theta_{it} = \left(\frac{\Delta S_t}{\Omega_i}\right)^2, \text{ with } \Delta S_t = \max(0, S_t - S_0), \tag{6}$$

where Ω_i is a key parameter whose calibration is a highly debated issue. In our model these numbers are calibrated as follows: we assume 2% GDP loss in the OECD (i.e. in our North) and 5% loss of GDP for the less developed countries at twice the pre-industrial level of atmospheric carbon, i.e. at 560 ppmv.[5] Inserting in Eq. (6) and assuming *no damages* at the current level S_0=360 ppmv gives Ω_N=1979.9 ppmv and Ω_S=1252.2 ppmv.

The damage estimates given above submerge several very heterogeneous types of damages into a single number: agricultural damages and benefits (in areas where due to warming agricultural conditions improve), land-losses and associated damages in coastal areas, drinking water availability, species loss, increased necessity of air conditioning, migration due to environmental catastrophes. Some potentially important damages are hardly accessible, like increasing human mortality and morbidity. To account for a possible underestimation of climate damages, some authors assume higher damage estimates that amount to up to 4% at double pre-industrial carbon concentrations for industrialized countries, see Kopp (2004). Without further investigations into the nature and composition of damages, however, any such number remains to a large extent arbitrary. Obtaining more reliable damage estimates is thus of prime importance.

Finally, the quadratic form of the damage function is chosen solely for computational simplicity; it results in linear marginal damage functions, which simplifies the solution of the model.[6]

The model is completed by specifying the objective functions of the regional benevolent central planners and the solution concept. The objective functions are given by

$$W_i = \sum_{t=0}^{\infty} \beta^t \ln c_{it}, \tag{7}$$

with β denoting the time discount rate. Note that in the model described here β is calibrated to replicate the global average GDP growth rate of the base year. Thus, calibration against the growth rate observed in the base year uniquely determines the utility discount rate of the central planner over the entire future.

[5] For an overview of damage estimates see Tol and Fankhauser (1998).

[6] Note that more elaborate specifications of damage functions exist, see e.g. Dumas and Ha-Duong (2004), who present an abrupt stochastic damage function. Such more sophisticated formulations, however, are not common practice in CGE modelling because of computational constraints.

In the BAU case the decision makers maximize utility neglecting the *endogenous* impact of fossil fuel use upon accumulated atmospheric carbon. This assumption reflects the global common pool character of the climate system. However, in equilibrium they anticipate future total factor productivities correctly, which includes a correct assessment of climate damages. This assumption accounts for correct anticipation of the future and hence guarantees time-consistency. This is quite a difference to an optimal policy case, where regional decision makers do not take the sequences θ_{it} as exogenously given but take into account their endogenous nature, being the result of fossil fuel use g_{it} (via Eqs. (5) and (6)). In the BAU equilibrium the emissions paths have to be such that the associated concentrations according to Eq. (5) yield the equilibrium paths of regional damages. Hence, solving for the BAU path involves two nonlinear optimal control problems as well as one global equilibrium (or consistency) condition. It is this combination of maximization and consistency conditions which is usually not tractable in bottom-up models. Note that the optimal policy case is more easily solved since it can be modelled as a joint maximization problem taking Eqs. (5) and (6) into account.

To sum up, our BAU scenario rests on one particularly important assumption: it is based on a non-cooperative solution concept, hence implies – by construction – that no mitigation policy occurs. Potential global gains from a mitigation policy can be identified in an alternative scenario of global cooperation where global welfare is maximized by a single global central planner and climate damages are internalized. From a computational point of view, the cooperative equilibrium is much easier to simulate. To illustrate the numerical differences between our BAU scenario and a cooperative policy approach, we present results from corresponding computational experiments below (with the corresponding scenarios labelled 'Cooperation'). Which scenario qualifies as reference, however, cannot be answered solely on theoretical grounds. Note that the two scenarios represent fundamentally different views about global cooperation in climate policy. An assessment concerning which is more likely is subjective and therefore it is important to highlight the differences between non-cooperative and cooperative solutions in a sensitivity analysis.

As we have already mentioned, CGE modelling requires parameter calibration. The usual approach (compare Shoven and Whalley, 1992) is to calibrate the model such that it replicates a given set of benchmark data of the base year. The stringent assumptions put on the model structure allow to determine several parameters easily, namely all parameters of the production function except for the substitution elasticities τ_i.

To be more specific, the first order conditions for profit maximization under perfect competition in combination with the income data in Table 1 uniquely determine a_i^1, a_i^2 and ϑ_i. As is common, the elasticities $\tau_i = -0.5$ are taken from the literature (from Manne et al., 1995). However, not only the values for τ_i are taken from the literature, also other parameters are set to specific values, as opposed to being e.g. estimated for the data set at hand. These include the depreciation rate δ, the investment technology parameter ω and also the carbon contents of energy η_{it}.

The calibration procedure as outlined suffers from two problems. First, the models are generally under-determined, i.e. some of the parameters have to be 'taken from the literature'

Table 1 – Key benchmark data

Key data for calibration

Data (base year 1998)	North	South
Labor income (trillion \$(1995))	16.780	3.011
Capital income (trillion \$(1995))	7.193	1.291
Energy expenditures (trillion \$(1995))	1.262	0.731
Carbon emission ($GtCO_2$)	11.406	7.955
Population (billion)	1.005	3.746
Annual population growth rate (2000–2015)	.015	.025
Emission coefficient η	9.04	10.88
Exogenous decarbonization of energy ('NTP') (2000–2015)	0	0
Exogenous decarbonization of energy ('TP') (2000–2015)	.02	–
Parameters		
Depreciation rate δ	0.05	0.05
Investment technology ω	0.2	0.2
Elasticity of substitution τ_I	−0.5	−0.5
Discount rate β	.975	.975
Climate damage at 560 ppmv in % output loss	2	5
Climate system parameter ϕ_1	0.302	
Climate system parameter ϕ_2	0.99	

'NTP' indicates the control scenario with no technological progress and 'TP' indicates the technological progress and diffusion of technology scenario, in which decarbonization of energy in the South is endogenous, see above. 560 ppmv corresponds to twice the pre-industrial level of atmospheric carbon concentration. For data sources see the Appendix.

or are subject to 'educated guesses'. This introduces a certain arbitrariness in the modelling approach, that can have, as will become clear later, important impacts on the results. Second, the calibration of the model economy to a specific base year makes the results potentially vulnerable to particularities of the chosen year. The growth rate of the economy, for example, which determines the utility discount rate, varies over time as discussed above. Static calibration with respect to a single time slice of the economy is a clear short-coming of current practice. In our data set the annual average growth rate varies between 1.4% and 4.6%, thus the choice of any single base year and the associated growth rate appears arbitrary. Furthermore, the usual calibration approach also necessitates to assume identical discount rates for all regions, despite clear interregional growth differences. Notwithstanding its clear drawbacks static calibration is common practice to date.

To assess the effects of choices concerning parameters that are not determined by calibration, we focus on two scenarios concerning the carbon content of energy, i.e. on the parameters η_{it}. These two scenarios are additionally used later to illustrate the potential pitfalls in attaching a structural interpretation to reduced form relationships (at the end of Section 2.3). In the scenario no technical progress (labelled 'NTP') we assume that the carbon content of energy remains constant in both North and South, reflecting base-year data. Thus, any initial technology differentials persist as these parameters are assumed to be constant and hence no decarbonization occurs.

However, assuming no technological progress and no diffusion of technology might be considered unrealistic and contrary to historical experience. Therefore we add these two elements in the scenario technical progress and diffusion of

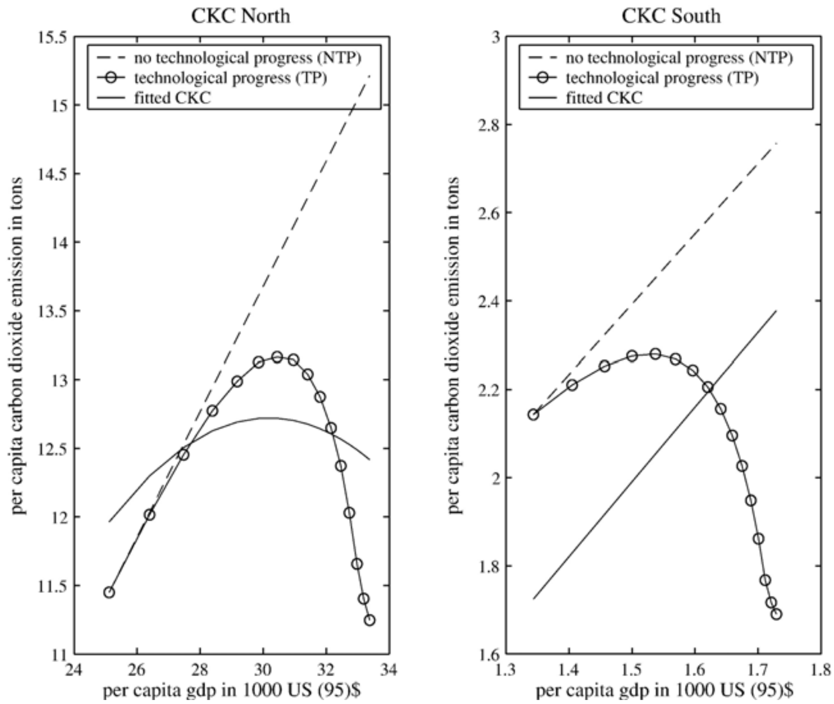

Fig. 1 – Simulation results for the GDP-CO_2 relation. The solid lines give the fitted CKC (Eq. (9)), based on a panel with cross-sectional dimension two on data from the 'technological progress' scenario. The turning point is at 30,000$ at 1995 prices.

technology (labelled 'TP'), where we assume a rate of decarbonization of 2% per year in the North, i.e. $\eta_{Nt+1} = \eta_{Nt}{}^{*}0.98$, and in addition complete catching-up of the South within 14 years. These two assumptions determine the sequence η_{St}.

The results for both scenarios are shown in Fig. 1, where we depict the GDP-CO_2 emissions relationships. The left picture in the figure shows the results for North and the right for South. The results for scenario 'NTP' are displayed in dotted lines and

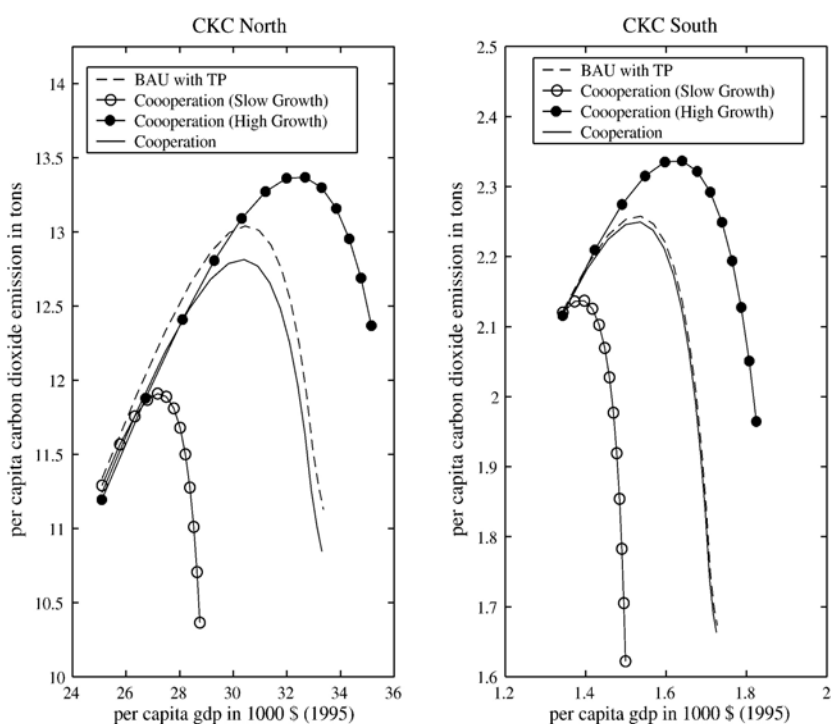

Fig. 2 – Simulation results for the GDP-CO_2 relation with technical progress (TP) and cooperation. See the text for explanations.

for scenario 'TP' in circled lines. In 'NTP' the GDP-emissions relationship is almost linear in both regions. This follows immediately from the constant carbon content of energy and increasing GDP. On the contrary, both regions exhibit an inverse U-pattern due to the exogenous decarbonization of energy and technology convergence in scenario 'TP'. Due to the low per capita incomes in the South the quadratic relationship appears to be almost linear in the right picture. As discussed, in the model this inverse U-shape is solely driven by technology induced decarbonization and not by income growth.

Fig. 2 shows the GDP-CO_2 emissions relationship for three additional scenarios based on 'TP', i.e. technological progress with respect to the carbon content of energy. BAU is identical to the one shown in Fig. 1, scenario 'Cooperation' gives the results for fully internalized climate effects. 'Cooperation (Slow Growth)' differs from 'Cooperation' by a higher utility discount rate, which has been calibrated to replicate the lowest growth rate in the observed time span, i.e. 1.4%. Similarly, the scenario 'Cooperation (High Growth)' is based on the highest observed growth rate of 4.6%. Comparing scenarios 'BAU with TP' and 'Cooperation' we observe, especially in the South, only minor differences. This is due to the short time span. However, the choice of the discount rate has major implications on the shape and location of the inverse U-shaped relationship between GDP and CO_2 emissions. In our model this arises because of the exogenous rate of decarbonization. Thus, a higher growth rate leads to (at any point time) larger CO_2 emissions due to exogenously given carbon intensity. Therefore, higher growth leads to higher total emissions and to a turning point at a higher income level. Current practice of calibrating CGE results with respect to a benchmark year is thus seen to be a critical issue. It is unfortunately not clear how to overcome this clear technical, but also conceptual, limitation of CGE modelling.

2.3. 'Estimating' the CKC for the CGE results

In this sub-section we 'estimate' a usual quadratic carbon Kuznets curve (CKC) to show that an inverse U-shaped pattern can emerge that is driven entirely by exogenous decarbonization; since as discussed before it is not due to a causal link between GDP and CO_2 emissions via e.g. increased willingness to pay at higher income levels. Before doing so, however, we start with a brief discussion concerning the usual single-equation approach to the EKC.

The most prominent single-equation approach to the EKC is to estimate a polynomial relationship (up to degree three) between emissions (as the dependent variable) and GDP on cross-section, time series or panel data. This approach dates back at least to the seminal work of Grossman and Krueger (1991, 1993, 1995) who find evidence for an inverse U-shaped relationship between per capita GDP and measures of fourteen pollutants.[7] Summary discussions of the empirical literature like Stern (2004) or Yandle, Bjattarai and Vijayaraghavan (2004) report more than 100 refereed publications of this type. The standard parametric EKC regression model is given by

$$\ln(e_{it}) = \alpha_i + \theta_t + \beta_1 \ln(y_{it}) + \beta_2 (\ln(y_{it}))^2 + u_{it} \tag{8}$$

where e_{it} and y_{it} denote per capita emissions and GDP in region i and period t, respectively, and u_{it} denotes a stochastic error term.[8] The error terms are in general allowed to be serially correlated. Time series like GDP are often modelled as so-called integrated processes. A stochastic process is called integrated (or 'has a unit root'), if it is not stationary itself but its first difference is. An important assumption necessary for many methods for panels containing integrated variables is that both the errors u_{it} and the regressor $\ln(y_{it})$ are cross-sectionally independent. This implies that if (8) corresponds to the data generating process, also the e_{it} are cross-sectionally independent, up to random θ_t. These independence assumptions, needed for so-called first generation panel unit root and cointegration analysis, are rather strong, and it is not granted at all that they hold in practice. In an increasingly interdependent world with large trade volumes it is e.g. not clear why the individual countries' GDP series should be independent. The issue of cross-sectional dependence and its implications for econometric analysis in a nonstationary environment is discussed in detail in Section 3.

The general formulation as displayed in Eq. (8) includes also country specific effects, α_i, and time effects, θ_t.[9] We model the country and time effects as fixed effects in this paper, whereas of course also random effects specifications are prominent in the literature. The shape of the functional relation is determined by β_1 and β_2, which depend neither on a specific region nor date. This homogeneity assumption is central to the standard panel analysis of the EKC: apart from the fixed effects α_i, and a stochastic error term u_{it}, all regions exhibit the same GDP-emissions pattern.[10] In particular, they all share the same turning point (if $\beta_2<0$), though the peak emission levels may differ across countries (see Fig. 3) via different country specific effects α_i. The turning point is located at $y^* = \exp\left(-\frac{\beta_1}{2\beta_2}\right)$.

The first econometric analysis of the CKC is due to Holtz-Eakin and Selden (1995), who use an annual panel of 130 countries over the period 1951–1986 and estimate their equation in levels as opposed to log-levels as illustrated in Eq. (8). They find support for an inverse U-shape, but the turning point is out of sample. Schmalensee, Stoker and Judson (1998) extend the data of Holtz-Eakin and Selden (1995) and use a 10-segment linear spline formulation and also find an inversely U-shaped relation. In the context of a small open economy, Friedl and Getzner (2003) estimate the CKC for Austria. They reject a quadratic formulation but find that an

[7] To be precise, Grossman and Krueger actually used a third order polynomial in GDP, whereas the quadratic specification seems to have been initiated by Holtz-Eakin and Selden (1995).

[8] In the literature also equations in levels instead of logarithms are popular. Note that all econometric problems discussed here apply equally to both formulations, in levels and in logarithms.

[9] In our implementation in the subsequent section, as is common in the panel unit root literature, we also investigate specifications including individual specific linear time trends.

[10] A fully homogeneous EKC supposes $\alpha_i = \alpha$ and identical distributions of u_{it} for all i.

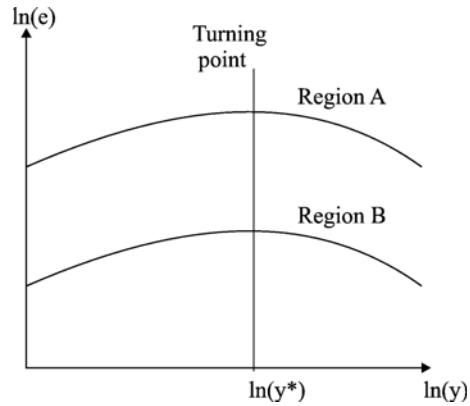

Fig. 3 – An EKC for two regions A and B. Though emission levels can differ among regions (via different country effects α_i), turning point income $y^* = \exp\left(-\frac{\beta_1}{2\beta_2}\right)$ is equal among all regions.

N-shaped cubic formulation proves to be an adequate choice for Austria. The problem is that the discussion in Section 3 will clarify that all these studies are subject to several important econometric problems that have up to now gone largely unnoticed in the EKC literature.

Before discussing these problems, we first 'estimate' the CKC for the data generated by the scenario 'TP', which clearly exhibits an inverse U-shape. We perform panel estimation of a quadratic CKC in both level and log-level terms.[11] This gives the following results in levels:

$$e_{it} = \hat{\alpha}_i + \underset{(0.747)}{1.786} y_{it} - \underset{(0.013)}{0.030} y_{it}^2 + \hat{u}_{it} \tag{9}$$

with $\hat{\alpha}_N = -14.220$ and $\hat{\alpha}_S = -0.620$ and 'standard errors' in brackets. Thus, estimation in levels generates an inverse U-shape, even with 'significant' coefficients. The fitted curve is depicted in Fig. 1 as a solid line. Estimation in log-levels, however, results in a U-shape with 'significant' coefficients:

$$\ln(e_{it}) = \hat{\alpha}_i - \underset{(0.245)}{0.749} \ln(y_{it}) - \underset{(0.048)}{0.132} (\ln(y_{it}))^2 + \hat{u}_{it} \tag{10}$$

with $\hat{\alpha}_N = 3.546$ and $\hat{\alpha}_S = 1.062$ and 'standard errors' in brackets.

Note for completeness that the inclusion of time effects θ_t leads to 'significant' coefficients with proper signs ($\beta_1 > 0$, $\beta_2 < 0$) in *both* the level specification and the logarithm specification. Thus, despite the clear evidence for an inverse U-shape in the 'TP' scenario (compare Fig. 1), pooled estimation of a homogeneous CKC is very sensitive to the specification of the relationship. This sensitivity obviously generalizes to real world econometric analysis.

If the estimated relationship (given that the underlying model, i.e. the data generating process, is generally unknown) is interpreted in the spirit of the Kuznets hypothesis, i.e. as a causal relationship between economic development and environmental quality, then inappropriate policy recommendations may be drawn from such a reduced form relationship. In our example, where the inverse U-shape occurs because of exogenous decarbonization at a certain rate, growth enhancing policies in order to 'pass the peak' of the CKC can eventually lead to *higher* emission paths. In order to see this consider e.g. a permanent increase in labor productivity in the South. This increases the marginal product of both capital and energy. Hence, for a CES production function this increases the input of energy and thus emissions in each period in the South. Therefore, this policy – based on the Kuznets curve observation – is seen to be counter-productive.

3. Problems in the econometric analysis of the environmental Kuznets curve hypothesis

In this section we turn to a discussion of major issues and problems arising in the econometric analysis of the environmental Kuznets hypothesis. Up to now most empirical studies, including those mentioned in Section 2.3, suffer from serious methodological shortcomings, many of which arise in particular in the presence of unit root type nonstationarity. Two main problems in this respect are: First, regressions involving integrated regressors and nonlinear transformations thereof (like (the logarithm of) per capita GDP and its square) require different asymptotic theory than the 'usual' asymptotic theory provided by the standard unit root and cointegration theory. Second, in a nonstationary panel context the unit root and cointegration methods applied so far (which, see item one, are not appropriate due to the nonlinear transformation) are all designed for cross-sectionally independent panels. Although such methods for cross-sectionally independent panels (so called first generation methods) are easy to use, also due to their increased availability in software packages, hardly any panel of economic data satisfies the cross-sectional independence assumption. This assumption, which requires GDP and emissions series to be independent across countries, is of course highly restrictive and unlikely to hold (compare the discussion in Section 2.3). Thus, the mentioned major problems lead us to question a large part of the existing literature. For a detailed discussion of these problems and one possible solution see Wagner (2005).[12]

In the present paper our goal is a bit more modest: We want to show that a careful application of 'standard' methods should lead a cautious researcher to question her findings and to become aware of potential problems. For our empirical illustrations of these problems we use the same data set as used in Section 2 for the CGE model, the difference being that we now do not aggregate the 107 countries into two groups and that we use all 13 years of observations.[13] Before we discuss the two main econometric

[11] Obviously this is just intended for illustrative purposes, as the data used for estimation are in fact deterministically generated from the calibrated model. The misspecification of the equation would be immediately visible from looking at the residuals $\hat{u}_{it}$ (which exhibit quadratic shape over time). From the deterministic behavior of the data it is also clear that the meaning of significance is nothing but a mere statement that standard *t*-values suggest significance. Of course, they are conceptually wrong.

[12] There are of course other solutions to these problems. E.g. Bradford, Fender, Shore and Wagner (2005) overcome these two main problems by using a formulation that only uses period averages of income and thus circumvents nonstationarity issues altogether.

[13] The example is only performed for illustrative purposes and is not intended to be a fully fledged econometric CKC analysis, since for e.g. the problem of parameter homogeneity as discussed in Section 3.1 is neglected.

issues in detail in Sections 3.3 and 3.4 we discuss for completeness in Section 3.1 the issue of parameter homogeneity and in Section 3.2 we discuss parametric versus non-parametric analysis. Also note here that all problems discussed in this section in general occur jointly in an empirical application. Thus, we separate the different problems here only to facilitate the discussion, whereas they have to be addressed jointly in empirical practice. Note that in empirical analysis also further problems may occur, e.g. structural stability of the data generating process or of estimated relationships may not be given. Such problems are in principle well understood and do not add anything to our discussion. For an investigation of structural breaks in the context of GDP and CO_2 emissions series see e.g. Heil and Selden (1999).

3.1. *Homogeneity across countries*

Let us first turn to the homogeneity assumption, which refers to β_1 and β_2 in Eq. (8) being identical for all countries. These parameters determine the shape of the EKC, which is identical for all countries contained in the panel if these two parameters are homogeneous. In Eq. (8) country specific elements of the EKC are only modelled by fixed or random effects and potentially by individual specific linear time trends, whereas the potential turning point is identical across countries. Dijkgraaf and Vollebergh (2005) test this assumption with reference to the results by Schmalensee et al. (1998). They restrict their panel to only 24 OECD countries from 1960–1997 with GDP measured in $(1995) purchasing power parities. Even a cursory comparison of GDP-CO_2 plots for Japan and the USA, they argue, casts serious doubts on the homogeneity assumption. They use a cubic extension of Eq. (8) and test the null hypothesis that the linear and quadratic coefficients are the same for all countries, i.e. $\beta_{ik} = \beta_k$, for $k = 1, 2$ and for all i. Like Schmalensee et al. (1998) they find within-sample turning points for all nations in their panel. An F-test, however, rejects the null hypothesis of equal coefficients at a 99% level of significance. This holds true even for most sub-panels. (They checked 380,000 combinations). The homogeneity assumption is decisively rejected, raising doubts on both the homogeneous polynomial (Eq. (8)) and the spline version. The conclusion of their work is that homogeneous panel estimates of the CKC may be inappropriate. However, when estimating the CKC for each country separately, they find support for the CKC hypothesis for 11 out of the 24 countries in their sample. This shows that a careful composition of the panel and a careful investigation of the homogeneity assumption by means of a specification search are both important for empirical analysis.

3.2. *Non-parametric approaches*

The second, 'lesser' methodological critique of the EKC concerns the parametric approach. Millimet, List and Stengos (2003) compare several modelling strategies, including semi-parametric techniques. In particular, they contrast the standard parametric framework with the more flexible semi-parametric approach for EKCs of nitrogen oxide and sulphur dioxide emissions in the United States. They clearly reject the parametric EKC approach for both pollutants. Especially in the case of sulphur dioxide, they find significant differences between parametric and semi-parametric estimates. Bertinelli and Strobl (2005) employ a semi-parametric estimator in a cross-country analysis for sulphur dioxide and CO_2 emissions.

Their panel comprises 108 countries over the period 1950–1990 on an annual basis. They show that emissions increase monotonically at low levels of per capita GDP. On higher levels, the relation is almost flat, i.e. it does not exhibit a turning point. They contrast their results with a parametric regression based on Eq. (8), which indicates for their sample again an inverse U-shape. This result, however, is mainly driven by data for the poorest countries. Hence they conclude that historical evidence about an inverse U-shaped EKC is not robust.

These examples show that the restriction to a simple polynomial relationship should be subjected to specification analysis more thoroughly than what appears to be common practice. Note already here, however, that the fundamental problems discussed in the following two subsections are equally relevant for both parametric and non-parametric approaches and are not resolved in either case.

3.3. *Unit roots, cointegration and nonlinear transformations of integrated regressors*

Environmental Kuznets curves involve a potentially integrated variable (like GDP or its logarithm) and its square as regressors, compare Eq. (8). This implies that if e.g. $\ln(y_{it})$ is a unit root nonstationary process its square is not an integrated process. This can be seen most easily as follows: Let $x_t = \Sigma_{j=1}^{t} \varepsilon_j$, where ε_t is a stationary process with positive spectrum at frequency 0. Hence, x_t is an $I(1)$ process and by construction $\Delta x_t = \varepsilon_t$. What about the first difference of x_t^2? Straightforward computations give that $\Delta x_t^2 = \Delta(\Sigma_{j=1}^{t} \varepsilon_j)^2$ is equal to $\Delta x_t^2 = \varepsilon_t^2 + 2\varepsilon_t \Sigma_{j=1}^{t-1} \varepsilon_j$, where Δ denotes the first difference operator. This shows that the first difference of the square of an integrated process is in general not stationary and hence the square of an $I(1)$ process is not an $I(1)$ process.

The relationship of the above example to the EKC (Eq. (8)) is clear: Both the logarithm of per capita GDP and its square are contained as regressors. However, as has been illustrated, at most one of them can be an integrated process. This fact has been overlooked in the EKC literature up to now. Several authors, e g. Perman and Stern (2003), nonetheless present unit root test results on log per capita GDP and its square. Furthermore, they even present 'cointegration' test results and estimates of the EKC.[14] This does not have a sound econometric basis. Consistent panel estimation techniques for this type of estimation problem have to be established first, as well as tests that are appropriate for such a 'nonlinear' cointegration problem.

Only recently there has been a series of papers by Peter Phillips and coauthors that addresses this problem for time series observations. This literature shows that the asymptotic theory

[14] Although Stern (2004) in his survey paper notes that it is very easy to do bad econometrics, his co-authored Perman and Stern (2003) paper is itself an example of falling into several pitfalls and we will therefore refer to it throughout our discussion. Other papers could serve the same purpose, for CO_2 emissions e.g. Friedl and Getzner (2003).

required, as well as the asymptotic properties obtained, generally differ fundamentally from the standard integrated case.[15]

Thus, in both a time series or panel context, the presence of nonstationary GDP or its logarithm (where it can be shown that generally again at most one of the two can be an $I(1)$ process) invalidates the use of standard unit root and cointegration techniques. Consequently, the findings obtained in studies applying such techniques are highly questionable.

3.4. *Unit roots, cointegration and cross-sectional dependence*

If variables are integrated but stochastically independent, the so-called 'spurious regression problem' occurs, when they are regressed on each other. Seemingly significant (with respect to standard t-statistics) coefficients may emerge from regressions of stochastically independent variables on each other, hence the name 'spurious'. This phenomenon was first observed by Yule (1926), and analyzed analytically in Phillips (1986). In order to obtain meaningful regression results from a regression containing integrated variables, it is necessary that these variables are *cointegrated*, i.e. share a common stochastic trend. Thus, the first step in a cointegration analysis is to test for unit root type nonstationarity and, if this is confirmed, a cointegration test will be the second step. Note for later reference that the problem discussed in the previous subsection, namely the nonlinear transformation of the regressor, will reappear later in this subsection in the cointegration testing step.

The short time span of our data with only 13 years necessitates the application of panel unit root tests. Let x_{it} denote the variable we want to test for a unit root (in our case these are the logarithm of per capita CO_2 emissions and the logarithm of per capita GDP) in an equation of the following form

$$x_{it} = \rho_i x_{it-1} + \alpha_i + \gamma_i t + u_{it}, \tag{11}$$

where u_{it} is a stationary process.[16] All so-called 'first generation tests' used in the EKC literature up to now assume cross-sectional independence of u_{it}. This is an unrealistic assumption, given the large degree of economic interactions across countries.[17] The null hypothesis of the panel unit root tests is given by H_0: $\rho_i=1$ for all i, against either the *homogeneous* alternative H_1^1: $\rho_i=\rho<1$ for all i, or against the *heterogeneous* alternative H_1^2: $\rho_i<1$, $i=1,\ldots, N_1$ and $\rho_i=1$, $i=N_1+1,\ldots, N$, for some N_1 such that $\lim_{N\to\infty} N_1/N>0$. The homogeneous alternative requires that under the alternative hypothesis all

Table 2 – Results of Im, Pesaran and Shin panel unit root tests for the logarithm of CO_2 emissions and the logarithm of per capita GDP including only fixed effects in the upper block-rows and fixed effects and time trends in the lower block-rows

Unit root testing		
Variable	IPS	IPS–LM
Fixed effects		
CO_2	0.230 (−1.628)	−1.291 (1.343)
GDP	**−1.590** (2.070)	0.070 (0.361)
Fixed effects and trends		
CO_2	**−2.094***(−2.485)	**0.259** (0.182)
GDP	**−3.423** (−1.346)	**0.456** (0.201)

The asymptotic 5% critical value is given by −1.645 for the IPS test and by 1.645 for the IPS-LM test. In brackets the bootstrap critical values are displayed. **Bold** indicates rejection based upon the bootstrap critical values and **bold*** indicates rejection based upon the asymptotic critical values but no rejection according to the bootstrap critical values.

cross-section members are stationary with furthermore identical first order serial correlation coefficient ρ. This restriction stems from the fact that such tests are constructed in a pooled fashion, where at some stage of the test procedure the coefficient ρ is estimated pooled for all observations together. The heterogeneous alternative allows for more flexibility in two ways: First, it allows for some cross-section members to be integrated also under the alternative and second it does not restrict the serial correlation coefficient to be identical under the alternative. For heterogeneous panels this alternative may be more relevant, hence we apply in this paper the two tests against the heterogeneous alternative developed by Im, Pesaran and Shin (1997, 2003). One of these two tests is essentially the group-mean of individual ADF t-statistics (labelled IPS) and the other is a group-mean Lagrange multiplier test (labelled IPS-LM). Group-mean refers here to the fact that for such tests unit root test statistics are computed for each individual cross-section member (i.e. country) which are afterwards combined appropriately. Both IPS tests are asymptotically standard normally distributed.

In addition to potential cross-sectional dependence there is another problem: The short time span of the panels may render asymptotic inference a bad guide for panel unit root testing (see Hlouskova and Wagner, 2006, for ample simulation evidence in this respect). Therefore we resort here to bootstrap inference and in particular we use the non-parametric bootstrap described next. Denote with $x_{it}\in\mathbb{R}$ the panel data observed for $i=1,\ldots, N$ and $t=1,\ldots, T$ (i.e. both the logarithms of per capita GDP and emissions). Then for each country the following equation is estimated by OLS:

$$\Delta x_{it} = \gamma_{i0} + \sum_{j=1}^{p_i} \gamma_{ij}\Delta x_{it-j} + u_{it} \tag{12}$$

The lag lengths p_i are allowed to vary across the individual countries in order to whiten the residuals u_{it}. Denote with $\hat{u}_{it}$

[15] Relevant papers are Chang, Park and Phillips (2001) and Park and Phillips (1999, 2001). Current research of the second author is concerned with an application of these theoretical results to the EKC/CKC hypothesis.

[16] Also time effects θ_t as contained in Eq. (8) can be included in the test procedure. Usually the processes u_{it} will exhibit serial correlation, which has to be taken into account appropriately in the test procedure.

[17] The results in Wagner (2005) show that for the present data set in fact for both variables common nonstationary factors appear to be present. Thus, the hypothesis of cross-sectional independence is most likely not fulfilled for the data at hand.

the residuals of Eq. (12). Then the non-parametric bootstrap procedure is based on the autoregression residuals as follows: Denote with $\hat{u}_t=[\hat{u}_{1t},\ldots,\hat{u}_{Nt}]'$ and generate the bootstrap residuals u_t^* by re-sampling $\hat{u}_t$, $t=p+2,\ldots,T$ with replacement. By re-sampling the whole vector, contemporaneous correlation across units is preserved in the bootstrap series. Given u_{it}^* the bootstrap data themselves are generated from

$$x_{it}^* = \begin{cases} x_{it} & t=1,\ldots,p_i+1 \\ \hat{\gamma}_{i0}+x_{it-1}^*+\sum_{j=1}^{pi}\hat{\gamma}_{ij}\Delta x_{it-j}^*+u_{it}^* & t=p_i+2,\ldots,T \end{cases} \quad (13)$$

We generate 5000 bootstrap replications of the data to obtain the bootstrap distribution (and hence critical values) of the test statistics by computing the test statistics for each bootstrap sample.

As mentioned above, resorting to the non-parametric bootstrap not only mitigates the problems with small sample inference based on asymptotic critical values (when all assumptions are fulfilled), it also robustifies inference with respect to contemporaneous short-run cross-sectional correlation due to re-sampling with an identical scheme for all cross-section members to a certain extent.[18]

The unit root test results in Table 2 carry two messages. Let us start in the specification with only fixed effects. Here the null hypothesis is only rejected for GDP when using the IPS test. When intercepts and trends are included, the null hypothesis is rejected (when resorting to bootstrap inference) with the exception of CO_2 and the IPS test. For this variable a rejection occurs only when resorting to the asymptotic critical values. However, the second (important) message is that the bootstrap and asymptotic critical values partly differ by a large extent. This may inter alia arise because of cross-sectional dependence. Thus, in the present situation such results may indicate misspecification of the test equation and should lead the researcher to perform further investigations, compare Wagner (2005).

Let us, however, for the moment continue with a 'standard' analysis and let us therefore (given the occurrence of non-rejections of the unit root hypothesis in some cases) proceed to panel cointegration testing, not without noting again that the discussion in the previous subsection has shown that doing so lacks theoretical foundations in the quadratic formulation (Eq. (8)).

We perform here two group-mean tests developed by Pedroni (2004), which we apply to the quadratic formulation (Eq. (8)) but also (to avoid the nonlinear transformation problem) on the equation without the squared logarithm of per capita GDP (which we refer to as linear specification). These tests are residual based cointegration tests, i.e. unit root tests performed on the residuals of Eq. (8) (respectively the linear specification) estimated by some appropriate method. If the variables are cointegrated, the residuals are stationary (hence the unit root hypothesis is to be rejected) and if they are not cointegrated the residuals are integrated (hence the unit root hypothesis is not to be rejected). Thus, the null hypothesis of these tests is that of *no cointegration*.

Table 3 – Results of panel cointegration tests, linear specification in the upper block-row and quadratic specification in the lower block-row

Cointegration test	PG_ρ	PG_{df}
Linear Specification		
FE	**−1.595** (−0.364)	**−10.956** (−6.424)
FE and Tr.	**1.477** (3.027)	**−14.352** (−9.018)
Quadratic Specification		
FE	**2.336** (2.647)	**−9.718** (−8.405)
FE and Tr.	**4.592** (5.586)	**−14.596** (−10.562)

The asymptotic 5% critical value is given by −1.645. In brackets the bootstrap critical values are displayed. **Bold** indicates rejection based upon the bootstrap critical values.

The two group-mean tests we report are the test based on the estimated first order serial correlation coefficient ρ (labelled PG_ρ) with serial correlation correction factors and the test where the serial correlation correction is performed by an augmented Dickey–Fuller type correction (labelled PG_{df}).[19] As for the unit root tests again the non-parametric bootstrap is implemented. Here the data are generated similarly as described above, with the difference that now bivariate bootstrap data vectors comprising the logarithms of per capita GDP and of CO_2 emissions are generated.[20] The number of replications is again 5000.

The results contained in Table 3 are subject to the cross-sectional dependence problem and in the quadratic formulation in addition subject to the nonlinear transformation problem.

At first sight, the evidence for a CKC might appear strong, since in both formulations both tests do reject the null hypothesis of no cointegration. However, the bootstrap critical values are now *very far* from the asymptotic critical values. Such large differences, especially with the quadratic formulation, are a rather clear indication that the behavior of the test statistic is not at all according to the theory. One reason for that can be misspecification, which is present by construction in the quadratic formulation. However, even when unaware of the mentioned fundamental econometric problems, these large differences should clearly alert researchers.

Note for completeness that estimation of the CKC on the present data set by using panel cointegration estimators (which are not suitable due to cross-sectional dependence and the nonlinear transformation problem, for the same reasons as discussed for the panel unit root and cointegration tests) show that any result can be supported by 'strategic' choice of the estimation procedure. Thus, also these results should lead to some doubts about the findings. As mentioned, Wagner (2005) finds strong evidence for cross-sectional dependence in the form of nonstationary common factors. He furthermore finds that the de-factored

[18] This claim can be easily verified by simulation, compare also Hlouskova and Wagner (2006). Corresponding results are available from the authors upon request.

[19] PG is just used as shorthand notation for Pedroni group-mean.

[20] For details see Appendix B of Wagner (2005).

per capita GDP series are stationary and thus performs a CKC regression on the de-factored observations that is not subject to the main problems discussed in this paper.[21] His results do not support the CKC hypothesis.

4. Summary and conclusions

The paper has highlighted several important problems that arise in both theoretical and econometric modelling of EKCs. Our discussion has been exemplified using CO_2 emissions, which are the most important pollutant on a global scale. However, the problems arise similarly for other pollutants as well.

With respect to the theoretical modelling approach we have identified several important problems arising in the dominant CGE approach, which we have exemplified within a small prototypical top-down model: First, computational tractability necessitates strong assumptions concerning functional forms, market structure and equilibrium concept. Often, in dynamic CGE models perfect competition is assumed. Of course, this assumption is highly unrealistic especially when modelling less developed countries. Second, the usual calibration procedures put a high weight on the observations of a single base-year. We have shown with our model that the results differ fundamentally when using different base-years (or to be precise different growth rates observed in different years in the data). Given the long-range projections based on results of such models this appears to be a major limitation of current practice. In general a more thorough understanding of the critical choices and the associated uncertainties in CGE modelling (e.g. with respect to damage estimates) should rank high on the research agenda of the community. In this respect, we see the clear need for combining CGE modelling with (structural) econometric analysis to allow for a (likelihood based) model specification and uncertainty analysis. Important input in this direction can be expected from the dynamic stochastic general equilibrium literature where estimation of dynamic models becomes more and more widespread.

In the discussion of the (single-equation) econometric problems arising in estimating EKCs, we have discussed in detail two (related) major problems that plague the bulk of the existing literature. First, the literature up to now ignores the econometric implications of the fact that Kuznets curve regressions involve nonlinear transformation of integrated regressors (GDP or the logarithm of GDP). This necessitates a different asymptotic theory than in the standard linear unit root and cointegration case, which has by and large gone unnoticed in the EKC literature. Second, in case of panels the applied methods in addition ignore the fact that almost all panels of economic time series are cross-sectionally dependent. This invalidates the use of so called first generation panel unit root and cointegration methods used in the EKC literature up to now that rely upon cross-sectional independence. These two problems together imply that much of the existing literature to date has to be regarded as questionable. In this paper we try to show that a careful and thoughtful application of standard tools should lead cautious researchers to be more skeptical about their findings than is commonly observed, which might have sufficed to avoid some pitfalls. The companion paper Wagner (2005) presents an econometric analysis that takes into account the two mentioned major problems and finds no support for a CKC (using the same data as here).

Summing up we conclude that both on the theoretical as well as on the econometric side many problems remain unresolved in EKC modelling. Given the importance of understanding GDP-emissions relationships this clearly indicates the need for research to overcome present limitations.

[21] De-factored observations denotes the observations minus the estimated common factors, often referred to as idiosyncratic component in the factor model literature.

Table 4 – Country list with members of the OECD in 1990 in bold face

Albania	Ecuador	Liberia	Seychelles
Algeria	Egypt	**Luxembourg**	Singapore
Antigua Barbuda	El Salvador	Macao	Solomon Islands
Argentina	Fiji	Malaysia	South Africa
Australia	**Finland**	Malta	**Spain**
Austria	**France**	Mauritania	Sri Lanka
Bahamas	French Guiana	Mauritius	St. Lucia
Bahrain	Gabon	**Mexico**	St. Vincent and Grenadines
Barbados	**Germany**	Mongolia	Suriname
Belgium	**Greece**	Morocco	Swaziland
Belize	Grenada	**Netherlands**	**Sweden**
Bolivia	Guatemala	New Caledonia	**Switzerland**
Botswana	Guyana	**New Zealand**	Syrian Arab. Rep.
Brazil	Honduras	Nicaragua	Thailand
Brunei	Hong Kong	Nigeria	Tonga
Bulgaria	Hungary	**Norway**	Trinidad and Tobago
Cameroon	**Iceland**	Oman	Tunisia
Canada	India	Pakistan	**Turkey**
Chile	Indonesia	Panama	United Arab. Emirates
China	Iran	Papua New Guinea	**United Kingdom**
Colombia	**Ireland**	Paraguay	**United States**
Costa Rica	Israel	Peru	Uruguay
Cyprus	**Italy**	Philippines	Venezuela
Denmark	Jamaica	**Portugal**	Vietnam
Djibouti	**Japan**	Puerto Rico	Zambia
Dominica	Jordan	Romania	Zimbabwe
Dominican Rep.	Korea Rep	Saudi Arabia	

Appendix A. Data and Sources

Our analysis is based on balanced panel data for 107 countries for the period 1986–1998. The list of countries is given in Table 4. The former Soviet Union and some eastern European countries are omitted from the sample because of a lack of data. Other countries like Kuwait are omitted because of large jumps in the emissions data. Member countries of the OECD in 1990 are in bold.

Per-capita CO_2 emissions are taken from the Carbon Dioxide Information Analysis Center (CDIAC) data set (see http://cidia.eds.ornl.gov/trends/emis/emcont.html). They are measured in metric tons of CO_2. Per capita GDP is measured in constant 1995 US$ and taken from the World Bank Development Indicators 2003.

REFERENCES

Bernard, A., Vielle, M., 1998. La Structure du Modèle GEMINI-E3. Economie et Prévision 5.

Bertinelli, L., Strobl, E., 2005. The environmental Kuznets curve semi-parametrically revisited. Economics Letters 88, 350–357.

Bradford, D., Fender, R., Shore, S.H., Wagner, M., 2005. The environmental Kuznets curve: Exploring a fresh specification. Contributions to Economic Analysis and Policy, vol. 4, No. 1. Berkeley Electronic Press. Article 5.

Brock, W.A., Taylor, M.S., 2005. Economic growth and the environment: a review of theory and empirics. In: Aghion, P., Durlauf, S. (Eds.), Handbook of Economic Growth. North-Holland, Amsterdam.

Chang, Y., Park, J.Y., Phillips, P.C.B., 2001. Nonlinear econometric models with cointegrated and deterministically trending regressors. Econometrics Journal 4, 1–36.

Dijkgraaf, E., Vollebergh, H.R.J., 2005. A note on testing for environmental Kuznets curves with panel data. Environmental and Resource Economics 32, 229–239.

Dumas, P., Ha-Duong, M., 2004. An abrupt stochastic damage function to analyse climate policy benefits. In: Hauri, A., Viguier, L. (Eds.), The Coupling of Economic and Climate Dynamics. Essays on Integrated Assessment. Kluwer, Dordrecht.

Energy Information Administration, 2004. International Energy Annual 2002.

Friedl, B., Getzner, M., 2003. Determinants of CO_2 emissions in a small open economy. Ecological Economics 45, 133–148.

Grossman, G.M., Krueger, A.B., 1991. Environmental impacts of a North American Free Trade Agreement. NBER Working Paper No. 3914.

Grossman, G.M., Krueger, A.B., 1993. Environmental impacts of a North American Free Trade Agreement. In: Garber, P. (Ed.), The Mexico–US Free Trade Agreement. MIT Press, Cambridge, pp. 13–56.

Grossman, G.M., Krueger, A.B., 1995. Economic growth and the environment. Quarterly Journal of Economics 110, 353–377.

Heil, M.T., Selden, T.M., 1999. Panel stationarity with structural breaks: carbon emissions and GDP. Applied Economics Letters 6, 223–225.

Hlouskova, J., Wagner, M., 2006. The performance of panel unit root and stationarity tests: results from a large scale simulation study. Econometric Reviews 25, 85–116.

Holtz-Eakin, D., Selden, T.M., 1995. Stoking the fires? CO_2 emissions and economic growth. Journal of Public Economics 57, 85–101.

Im, K.S., Pesaran, M.H., Shin, Y. (1997). Testing for unit roots in heterogeneous panels. Mimeo.

Im, K.S., Pesaran, M.H., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. Journal of Econometrics 115, 53–74.

Intergovernmental Panel on Climate Change, 2001. Climate Change Working Group III: Mitigation, Section 9.4.2.1.

Joos, F., Müller-Fürstenberger, G., Stephan, G., 1999. Correcting the carbon cycle representation: how important is it for the economics of climate change. Environmental Modeling and Assessment 4, 133–140.

Kemfert, C., 2002. An integrated assessment model of economy–energy–climate — the model WIAGEM. Integrated Assessment — An International Journal 4, 281–289.

Kopp, R.J., 2004. Near-term greenhouse gas emission targets. Discussion Papers, vol. 04-41. Resources for the Future.

Kuznets, S., 1955. Economic growth and income inequality. American Economic Review 45, 1–28.

Magill, M., Quinzii, M., 1996. Theory of Incomplete Markets. MIT Press, Cambridge.

Manne, A.S., 1977. A model of energy–economy interactions. In: Hitch, C.J. (Ed.), Modeling Energy–Economy Interactions: Five Approaches. Resources for the Future, Washington, DC.

Manne, A.S., 1999. Discounting and intergenerational equity. In: Weyant, J.P., Portney, P.R. (Eds.), Discounting and Intergenerational Equity. Resources for the Future, Washington, DC, pp. 111–129.

Manne, A.S., Mendelsohn, R.O., Richels, R.G., 1995. MERGE: a model for evaluating regional and global effects of GHG reduction policies. Energy Policy 23, 17–34.

McKibbin, W.J., Ross, M.T., Shackleton, R., Wilcoxen, P.J., 1999. Emissions trading, capital flows and the Kyoto protocol. In: Weyant, J. (Ed.), The Energy Journal Special Issue: the Costs of the Kyoto Protocoll, A Multi-Model Evaluation.

Millimet, D.L., List, J.A., Stengos, T., 2003. The environmental Kuznets curve: real progress or misspecified models? Review of Economics and Statistics 85, 1038–1047.

Nordhaus, W., 1991. To slow or not to slow: the economics of the greenhouse effect. The Economic Journal 101, 920–937.

Nordhaus, W., 1992. An optimal transition path for controlling greenhouse gases. Science 228, 1315–1319.

Park, J.Y., Phillips, P.C.B., 1999. Asymptotics for nonlinear transformations of integrated time series. Econometric Theory 15, 269–298.

Park, J.Y., Phillips, P.C.B., 2001. Nonlinear regressions with integrated time series. Econometrica 69, 117–161.

Pedroni, P., 2004. Panel cointegration. asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory 20, 597–625.

Perman, R., Stern, D.I., 2003. Evidence from panel unit root and cointegration tests that the environmental Kuznets curve does not exist. Australian Journal of Agricultural and Resource Economics 47, 325–347.

Phillips, P.C.B., 1986. Understanding spurious regressions in econometrics. Journal of Econometrics 33, 311–340.

Prinn, R., et al., 1999. Integrated global system model for climate policy assessment: feedbacks and sensitivity studies. Climatic Change 41, 469–546.

Schmalensee, R., Stoker, T.M., Judson, R.A., 1998. World carbon dioxide emissions: 1950–2050. Review of Economics and Statistics 80, 15–27.

Shoven, J.B., Whalley, J., 1992. Applying General Equilibrium. Cambridge Surveys of Economic Literature, Cambridge, UK.

Stern, D.I., 2004. The rise and fall of the environmental Kuznets Curve. World Development 32, 1419–1439.

Tol, R.S.J., Fankhauser, S., 1998. On the representation of impact in integrated assessment models of climate change. Environmental Modeling and Assessment 3, 63–74.

Wagner, M., (2005). The carbon Kuznets curve: a cloudy picture emitted by bad econometrics? Mimeo.

Yandle, B., Bjattarai, M., Vijayaraghavan, M., 2004. Environmental Kuznets curves: a review of findings, methods, and policy implications. PERC. Research Study 02.1 update.

Yule, G.U., 1926. Why do we sometimes get nonsense correlations between time series?—A study in sampling and the nature of time series. Journal of the Royal Statistical Society 89 (2–9), 30–41.