Economic Growth and Air Pollution: Three Empirical Essays Based on Nonparametric Methods

Thèse présentée à la Faculté des sciences économiques
Université de Neuchâtel

Pour l'obtention du grade de docteur ès science économiques

Par

Carlos ORDÁS CRIADO

Acceptée sur proposition du jury de thèse:

Prof. Jean-Marie GRETHER, Université de Neuchâtel, directeur de thèse.
Prof. Jaime DE MELO, Université de Genève, co-directeur de thèse
Prof. Thanasis STENGOS, University of Guelph
Prof. Philippe THALMANN, École Polytechnique Fédérale de Lausanne
Prof. Milad ZARIN-NEJADAN, Université de Neuchâtel, président du jury

Soutenue le 9 mars 2009

Neuchâtel, 2009

Université
de Neuchâtel

IMPRIMATUR POUR LA THESE

Economic Growth and Air Pollution: three empirical essays based on
nonparametric methods

Carlos Ordas Criado

UNIVERSITE DE NEUCHATEL

FACULTE DES SCIENCES ECONOMIQUES

La Faculté des sciences économiques,
sur le rapport des membres du jury

M. Jean-Marie Grether (directeur de thèse, Université de Neuchâtel)
M. Milad Zarin-Nejadan (président du jury de thèse, Université de Neuchâtel)
M. Jaime de Melo (Université de Genève)
M. Philippe Thalmann (EPFL, Lausanne)
M. Thanasis Stengos (University of Guelph, Canada)

Autorise l'impression de la présente thèse.

Neuchâtel, 30 mars 2009

Le doyen

Kilian Stoffel

■ Téléphone : +41 32 718 15 00 ■ Fax : +41 32 718 15 01 ■ E-mail : secretariat.seco@unine.ch ■ www.unine.ch

English abstract

Abstract: this dissertation includes 3 research papers, which explore empirically the link between the level of economic activity and air pollution at the macroeconomic level. A special emphasis is given to the application of recent tools developed in the nonparametric field as they allow for a better control of potential misspecification biases and they give more flexibility to the underlying relationships. Chapter 1 tests whether a sustainable link between per capita GDP levels and the environment exists for a variety of air pollutants' emissions in a panel of 48 Spanish provinces over the period 1990-2002. Chapter 2 investigates how cross-country gaps in per capita CO_2 emissions evolved over the 1960-2002 period for a panel of 166 world areas as well as for several country sub-groupings (rich/poor countries, specific geographic regions, economically integrated areas). An analysis of the dynamic of the cross-sectional distributions is conducted with robust scale and shape measures and formal shape and multimodality tests are applied. The latter approach is contrasted with a stochastic convergence analysis *à la Evans (1998).* Chapter 3 makes use of a broad range of regression techniques to fit a reduced form function, where growth rates in per capita CO_2 emissions are explained with past pollution levels, past per capita GDP levels and per capita GDP growth rates. Panel models are estimated with standard linear and nonlinear least squares and these specifications are tested against their nonparametric counterpart. This framework also allows exploring beta-convergence in per capita emissions conditional on GDP as well as beta-convergence in GDP conditional on pollution.

Keywords: Air pollution, carbon dioxide emissions, Environmental Kuznets Curve, convergence, distributional dynamics, mixed nonparametric and semiparametric regressions, panel poolability test, unit roots.

Résumé en français

Résumé : cette thèse se compose de 3 recherches dont l'objectif est l'étude empirique du lien entre croissance économique et pollution atmosphérique au niveau macroéconomique. Elle privilégie l'utilisation d'outils récents issus de l'analyse non-paramétrique car ils permettent une meilleure prise en compte de biais de mauvaise spécification et introduisent plus de flexibilité dans la forme fonctionnelle étudiée. Le premier chapitre vérifie l'existence d'une relation soutenable entre le niveau de PIB par tête et l'environnement dans 48 provinces espagnoles pour différents polluants de l'air durant les années 1990-2002. Le second chapitre s'intéresse à l'évolution des différences de niveau d'émissions de CO_2 par tête entre 1960 et 2002 dans 166 pays, ainsi que dans divers sous-ensembles de pays (riches/pauvres, appartenant à une même zone géographique, membres d'une zone d'intégration régionale. . .). La dynamique des distributions transversales est étudiée à l'aide de mesures robustes d'échelle et de forme fonctionnelle. Des tests d'égalité distributionnelle et de multi-modalité sont également appliqués pour tester la stationnarité des densités et l'émergence de différents modes. Ces résultats sont comparés à ceux obtenus à l'aide d'une analyse de convergence stochastique à la Evans (1998). Le dernier chapitre applique un large spectre de techniques de régressions à l'estimation d'une forme fonctionnelle réduite, dans laquelle la croissance des émissions de CO_2 par tête est fonction du niveau passé d'émissions par tête, du niveau passé de PIB par tête et du taux de croissance du PIB par tête. Des modèles de panel sont estimés par les moindres carrés ordinaires et les moindres carrés non-linéaires et ces spécifications sont comparées à des modèles non-paramétriques alternatifs. Ce cadre permet également d'explorer la notion de beta-convergence dans la pollution, conditionnelle au PIB, ainsi que la beta-convergence dans le PIB, conditionnelle à la pollution.

Mots clés : Pollution atmosphérique, émissions de CO2, Courbe de Kuznets Environnementale, convergence, dynamique distributionnelle, régressions non et semi-paramétrique mélangées, test d'empilement de panel, racine unitaire.

Acknowledgements

I would like to express my deepest gratitude to Prof. Jean-Marie Grether and Prof. Jaime de Melo who inspired me in the field of economics and allowed me to undertake the present research under their supervision and financial support. Many thanks for your trust, your encouragements in all aspects of my work, your critical feedbacks and the time spent reading the numerous drafts of my research papers, under tight time constraints sometimes.

I am warmly grateful to Prof. Milad Zarin-Nejadan (University of Neuchâtel) and to Prof. Thanasis Stengos (University of Guelph, Canada) for co-supervising my work as well as for their kindness. Prof. Thalmann (University of Lausanne) has also accepted to provide his expertise to this work, which I am sincerely indebted.

The Institute for Research in Economics from the University of Neuchâtel (Irene) has been the memorable place where I have shared my joys, doubts, efforts and my desk with stimulating colleagues, who have become friends. Thank you to Martine, Françoise and Kira for their help and smiles in all circumstances, and for making our institute feeling like a home. Mathieu, Johanne, Sonia, Gilles and Moez have very much contributed to make that time unforgettable, pushing me to find the right balance between work and refreshing breaks. My sincere acknowledgments also go to our IT scientists, Abdel, David and Florian, as well as to our library staff, Denis, Sandra and René for always providing the quick and flexible support needed in my task.

I owe a great debt to my friend David Ardia, for the passionate talks on many technical aspects of my work, going from statistical/econometric issues to a variety of computational problematics (in the R and Latex environments in particular). His expertise has been a great asset.

I had the chance to spend a year of my PhD program visiting several Economic Departments in Canada (at the University of Guelph, the University of Toronto and the University of Calgary) under a grant kindly provided by the Swiss National Science Foundation. Thanks a million to the many people who made my visits possible and for having provided all the facilities for my researches.

Least but not last, my family has always been by my side in this academic adventure. Thanks Dad, Javier and Iván for being there when I was absorbed in my dissertation. Javier has been kind enough to revise the English drafts of my papers. Thank you Joëlle for your love, support and optimism.

Carlos Ordás Criado

A la memoria de mi madre, Rosalía,
a mi hermanito Iván.

Contents

Introduction

While the environmental impacts of economic activity has become a major source of concern for policy makers, economists try to shed some light on the forces which shape the link between economic growth and pollution, both theoretically and empirically. This dissertation explores the latter side of the problematic at the macroeconomic level. It includes three applied essays which make use of recent econometric tools, mainly developed in the nonparametric field, to examine the relationship between economic growth and air pollution both at the national level and in an international context.

My first essay investigates the link between the level of economic activity and air pollutants' emissions, by focusing exclusively on per capita GDP and emissions levels. This simple bivariate model has generated a huge amount of interest in the empirical literature. Theoretically speaking, it may be seen as representing a reduced form function in which income influences technology, the scale and composition of output, and the demand for environmental quality. Changes in these factors in turn influence environmental pressure. This single equation implicitly assumes no feedback effect of environmental damages on the level of economic activity, it is static and it captures only instantaneous or short run effects of income on pollution. Indeed, the model is essentially descriptive and it does not provide any answer on whether the expected positive impact of a wealth increase on environmental quality is achieved by more stringent environmental policies or by autonomous structural and technological changes that are related to economic growth. Despite these limitations and difficulties in disentangling causal effects, this formulation tests ultimately the existence of a sustainable dynamics between economic growth and pollution, *i.e.* a steady increase in per capita income and in environmental quality. Among the many possible sustainable patterns, the U-inverted shape, also known as the Environmental Kuznets Curve (EKC henceforth), has been the most investigated and debated pattern. It posits an increasing pressure on natural resources at early stages of economic expansion which later stabilizes and ultimately declines as people get richer and the economy becomes more efficient. This is the so called 'EKC hypothesis'. My first paper proposes a strategy to check if the most common model used with panel data to test the EKC hypothesis provides consistent estimates of the income-pollution relationship for most of the individual countries/regions included in the panel. I employ nonparametric regressions to avoid a misspecification bias when fitting the data. Indeed, given the complexity of the income-pollution relationship, it is difficult to specify *a priori* the correct function which uncovers that link. Nonparametric methods let the data dictate the shape of the latter function by optimizing some fit criteria. With the help of a panel of 48 Spanish provinces on four air pollutants that covers the years 1990 to 2002, I explicitly show that EKC patterns can be found with non, semiparametric or cubic fixed effects models even when the income-pollution pattern is decreasing, stabilized or even increasing over time for most of the provinces. I argue that the time and spatial homogeneity of the panel should be checked before making any inference on the shapes estimated with pooled or fixed effects regressions. The U-inverted functions that I find in the Spanish dataset for all the pollutants appear to be consistent estimates of cross-sectional regressions for all the years of the panel. But these estimates do not depict time-series fits for the regions; they do not reflect at all the income-pollution trends in the regions for the investigated period.

The other two essays analyze whether carbon dioxide per capita emissions' levels are converging across countries. We focus on pollution convergence (in air pollution per capita levels) for two essential reasons. First, among the many policy measures put forward to mitigate global warming in the post-Kyoto effort, a significant number of proposals rely on per capita emissions targets. The supporters of that approach defend the fairness of this allocation scheme ('each individual should have the same right to pollute'). This principle ignores specific structural characteristics of the countries and it is debatable whether or not it constitutes an efficient approach. However, its operational simplicity and ability to set a 'unifying principle that facilitates an international greenhouse warming agreement' (Rose and Stevens, 1996, p.2-3) between governments has attracted institutional support. Finding convergence in per capita emissions worldwide or between groups of countries may thus be of particular interest in policy circles. Second, from a more theoretical perspective, several authors have amended standard macroeconomic growth models with pollution components and links have been established between income and pollution convergence. Investigating pollution growth with macroeconomic reduced form functions allow to analyze the relevance of these models, and to explore pollution growth and convergence at the same time within theoretically derived specifications.

Different tools exist to measure convergence in macroeconomic series. My second essay explores convergence in per capita CO_2 emissions, at the world-wide level and also within different subsets of countries, by focusing exclusively on the three main univariate convergence measures: sigma, stochastic and distributional convergence. My first contribution in that paper is to widely expand the number of countries used so far in the empirical literature on per capita carbon emissions' convergence, in order to generate systematic groupings of countries, based on income, geographic and economic integration criteria. Then recent exploratory methods are applied in the distributional as well as the stochastic convergence approach. The former analysis is essentially descriptive and relies on robust scale and shape indicators as well as distributional tests to evaluate changes in the cross-sectional distributions over time. The latter stochastic analysis is carried out essentially based on unit root tests combined with the concept of pair-wise convergence introduced by Evans (1998). Overall, significant differences emerge in the cross-sectional distributions over time, mainly between those from the pre and post-70s oil shocks period, for the world as well as for different groupings of countries. The distributional analysis provides little evidence of a strong polarization in national carbon series. Moreover, the evolution of the distributional patterns are difficult to interpret in terms of conditional convergence. By contrast, the stochastic convergence analysis identifies converging economies at the world-wide levels as well as for many country groupings.

Finally, the last essay employs the database on per capita CO_2 emissions constructed in the second essay to explore growth and β-convergence in carbon dioxide per capita emissions in a multivariate setting, *i.e.* with the help of the growth model with pollution of Alvarez et al. (2005). These authors derive in a simple way a reduced form function from a model *à la Ramsey* that allows to examine growth as well as β-convergence in pollution, conditional on income levels and growth rates. Moreover, by simply reverting the correlation scheme and accounting for potential simultaneity bias with instrumental variables, I also explore growth and β-convergence in GDP, conditional on pollution levels and growth rates. The analysis is carried out in a panel framework, with a variety of regression techniques: ordinary least squares, nonlinear least squares, semi and nonparametric regressions. The original test equation is augmented with time and OECD dummies for the empirical treatment. A recent nonparametric specification test is applied to check whether the functional constraints imposed by the theoretical model is supported by the data. I find that parametric models are in general misspecified and that nonlinearities and interactions between the variables are better captured with non or semiparametric regressions. Fully nonparametric estimates, which involve discrete and continuous explanatory factors, show interesting interactions between the OECD status and the main explanatory continuous variables. They also indicate that convergence in per capita pollution levels across countries may happen between countries which experience increasing income disparities.

The structure of this dissertation is as follows. The three essays discussed in the above paragraphs are presented in chapters 1, 2 and 3 respectively. Special contributions to the essays are acknowledged after the conclusion of each chapter. Most Tables and Figures are included in the text but some additional information is provided in appendices located at the end of each chapter. The complete list of all Tables and Figures can be found at the end of the dissertation. We end the dissertation with a brief general conclusion, where we suggest further extensions. The bibliography includes all the papers cited along the three chapters. Finally, most of the computations have been carried out in the R Development Core Team (2007) statistical environment. The code is available upon request.

Chapter 1

Temporal and Spatial Homogeneity in Air Pollutants Panel EKC Estimations

Two Nonparametric Tests Applied to Spanish Provinces

Editorial note: A short version of this paper has been published in *Environmental and Resource Economics*[1].

Abstract: Although panel data have been used intensively by a wealth of studies investigating the GDP-pollution relationship, the poolability assumption used to model these data is almost never addressed. This paper applies a strategy to test the poolability assumption with methods robust to functional misspecification. Nonparametric poolability tests are performed to check the temporal and spatial homogeneity of the panel and their results are compared with the conventional F-tests for a balanced panel of 48 Spanish provinces on four air pollutant emissions (CH_4, CO, CO_2 and NMVOC) over the 1990-2002 period. We show that temporal homogeneity may allow the pooling of the data and drive to well-defined nonparametric and parametric cross-sectional U-inverted shapes for all air pollutants. However, the presence of spatial heterogeneity makes this shape compatible with different time-series patterns in every province - mainly increasing or decreasing depending on the pollutant. These results highlight the extreme sensitivity of the income-pollution relationship to region-specific factors.

JEL classification: C14 · C23 · O40 · Q53

Keywords: Environmental Kuznets Curve, Air pollutants, Non/Semiparametric estimations, Poolability tests

---

[1] See http://www.springerlink.com/content/f456761736487wt5/?p=f6f8f73015844ddfbd51f0b30921e0d2&pi=0

1.1 Introduction

In the last fifteen years the relationship between economic growth and environmental quality has been one of the most investigated issues in the empirical literature. Air, water or land pollution, global warming or resources depletion are clearly related to human activities but the nature of that link remains highly controversial. The most famous example is probably the Environmental Kuznets Curve (EKC), which posits an U-inverted relationship between some measure of economic activity and environmental damage. The existence of that hump-shaped pattern has been challenged by a plethora of empirical research, particularly for atmospheric pollutants.

Two main caveats affect the empirical estimation of the income-pollution relationship. Firstly, economic theory suggests that the reduced form function postulated by the EKC hypothesis may not have a simple and unique functional shape. Secondly, even if a single function were to exist, it would be very sensitive to country or region specific factors, such as : factor endowments, sources of growth, differences in technology, social sensitivity to environmental damages, etc. These two characteristics have oriented the current empirical investigations on the income-pollution relationship in two directions: (i) parametric specifications have been replaced by nonparametric fitting methods to avoid functional misspecification; and (ii) controlling for heterogeneity in panel data has become a fundamental issue in obtaining unbiased estimates.

The vast majority of EKC's empirical papers use panel data structures (*i.e.* data on individual countries/regions observed over time). These papers make use of all the data points to get estimates of a common functional form to all countries/regions up to some deterministic vertical shift specific to every country/region or year of the panel. These panel data models are referred to as fixed effects and their estimates are said to be pooled because a unique function is assumed to hold for all countries or regions or years up to some intercept term. In most cases, and whether the functional form is parametrically specified or not[2], no formal check of the homogeneity assumption is provided on the time (*i.e.* stability of the cross-sectional regressions over time) and the spatial (*i.e.* equality of the time-series regressions across countries/regions) dimensions of the panel. Yet,

---

[2] For parametric specifications, see among others Selden and Song (1994), Grossman and Krueger (1995), Holtz-Eakin and Selden (1995), Schmalensee et al. (1998), Heil and Selden (2001), De Groot et al. (2004) or Aldy (2005); for non- or semiparametric ones, see Taskin and Zaim (2000), Millimet et al. (2003), Bertinelli and Strobl (2005) or Azomahou et al. (2006).

this assumption is crucial to get robust and unbiased estimates. Moreover, among the few authors who have tackled this issue[3] for different kinds of environmental damage, conflicting results have been reached for CO_2 emissions data. Dijkgraaf and Vollebergh (2005), for the 24 OECD countries, overwhelmingly reject the hypothesis of homogeneous income-pollution relationship between regions/countries made in the fixed-effects panel data models commonly used in the literature. Pooled estimates are consequently rejected. Azomahou et al. (2006) reach the opposite conclusion when checking the temporal poolability on a much larger panel of 100 countries with a poolability test robust to functional misspecification. This discrepancy may be attributed to the different procedures used; but it also raises a more fundamental question: to what extent is temporal homogeneity compatible with spatial heterogeneity?

This research contributes to the recent empirical literature on the EKC curve by testing for the first time the adequacy of the homogeneity assumption on both the temporal and the spatial dimensions with nonparametric tests robust to functional misspecification. Following Azomahou et al. (2006), we make use of Baltagi et al. (1996)'s nonparametric poolability test to check the temporal homogeneity of a panel on anthropogenic emissions of four air pollutants (CH_4, CO, CO_2 and NMVOC) for the Spanish provinces over the 1990-2002 period. These pollutants are particularly interesting as they display different growth aggregate patterns over the investigated period. Furthermore, we apply the simple procedures of Yatchew (2003) to check the equality of non- and semiparametric estimations of the income-emissions relationship (IER) at the regional level. This allows us to verify the spatial homogeneity hypothesis with a method robust to functional misspecification. We compare the results provided by the standard F-tests procedures applied to the quadratic and cubic models to our nonparametric tests. We are able to confirm the existence of robust and stable cross-sectional EKCs over time for most of the air pollutants investigated. However, this does not mean that every province displays the same IER for a given pollutant; for all of them, we find that the spatial homogeneity hypothesis is overwhelmingly rejected. We show explicitly that stable cross-sectional EKCs are perfectly compatible with either increasing or decreasing emissions in most of the regions depending on the pollutant. Consequently, pooled EKC estimates are compatible with all kinds of IERs at the most aggregated level. These results confirm the warnings made by de Bruyn et al. (1998) regarding the interpretation of the EKC shapes found with pooled panel data

---

[3] See List and Gallet (1999), Koop and Tole (1999), Dijkgraaf and Vollebergh (2005) or Aldy (2005) or Azomahou et al. (2006)

models.

The structure of this paper is as follows. Section 1.2 offers a brief survey of the main theoretical determinants of the income-pollution relationship. It includes a review of empirical literature focused on CO_2-IER encapsulating the main econometric issues which are linked to EKC estimates for other pollutants. The main findings for IER estimations on air pollutants with panel data at low level of geographical aggregation are also provided. Section 1.3 presents the econometric strategy. The Spanish data are described in Section 1.4 and Section 1.5 shows the econometric results. We present our conclusions in Section 1.6.

1.2 Income-pollution relationship: from theory to empirics

Most of the empirical studies[4] investigating the relationship between the level of economic activity and some pollution indicator have faced two main issues: defining the functional shape to be estimated; and getting robust estimates despite the short time series available.

Theoretical background. As Copeland and Taylor (2003) point out, in the absence of change in the structure and technology of the economy, increasing economic activity would result in an equiproportionate growth in pollution or other environmental impacts. This 'scale' effect suggests a monotonically increasing relationship between real GDP and pollution and makes economic growth and sustainable development two conflicting goals. However, economic growth generates technological progress; polluting inputs are used more efficiently in the production process or through abatment technologies. If the 'technical' effect is strong enough to offset the scale effect, economic growth is compatible with less pollution and the link may become locally decreasing. Three other mechanisms also lead to changes in the output composition of countries: unbalanced growth processes of production factors; biased technological progress between industries or variations in relative world prices. These specialisation patterns between unequally pollution-intensive sectors are usually referred to as 'composition' effects. The sources-of-growth explanation of the

---

[4] See Brock and Taylor (2004a) for an empirical and theoretical review of the literature on the relationship between economic growth and the environment or Stern (2003) for the EKC literature.

EKC shape relies on that particular argument. If economic growth is first induced by accumulation of a production factor (capital) used relatively more intensively in a polluting sector but then shifts toward accumulation of a factor (labor or human capital) more intensively used in a less or non polluting sector, a straightforward application of Rybczinsky's theorem leads pollution to follow the same path as the production of the polluting good, an U-inverted pattern. A similar argument can be used to explain why capital abundant economies (rich countries) are expected to pollute more than labor-abundant ones (poor countries). All these supply side arguments have two major implications. Firstly, economic growth may not require any environmental policy measure to be compatible with a more efficient use of polluting inputs or natural resources. Secondly, as Copeland and Taylor (2003, Ch. 3.1) indicate, we can have a stable relationship between pollution and technology and primary factors, and between income and these same variables, without having a simple and stable relationship between pollution and income. In plain words, the same level of income may be linked to different levels of pollution, depending on the factor which generated this income level.

From a social point of view, the willingness to tolerate the inconveniences of pollution in order to increase income plays a major role in determining the strength of policy responses to environmental damages. Consequently a pure scale effect generated by neutral growth could be overcome by environmental policy measures if, at some level of income, the relative willingness to pay for pollution reduction exceeds the relative growth in income[5]. The income-pollution relationship is also sensitive to the way pollution is measured (*i.e.* in levels, *per capita* or intensity terms), as well as to the level of spatial aggregation of the data. In this paper, we focus on *per capita* levels of pollution as it represents the most common specification of the dependent variable in the IER literature on air pollutants.

Empirical estimations. Given the variety of theoretical foundations, no single functional form can be advocated *a priori* to link indicators of environmental degradation with measures of economic activity. As the income-pollution relationship is a reduced form function, all the underlying forces which determine its shape for a particular geographical area are subsumed, *i.e.* they remain unexplained. The early empirical IER literature has addressed the functional uncertainty by retaining three main parametric flexible specifications: quadratic and cubic functions which capture nonlinearities and spline linear

---

[5] This is usually referred to as an income elasticity of marginal damage greater than one in the literature.

functions which gauge thresholds effects. More recently, researchers have turned to nonparametric and semiparametric regressions which leave the functional form unspecified and avoid the risk of choosing an inadequate parametric function. Moreover, the lack of long time series on pollutants at the country level has made authors favour cross-country/region panel data. The absence of a range of explanatory variables which consistently capture the differences between countries may lead to biased estimates. This heterogeneity issue has been neglected in most of the parametric and nonparametric analysis of IER panels. Moreover, when it has been investigated, the F-tests used were not robust to functional misspecification. Consequently, the estimated IER appears to be highly sensitive to the pollutant or environmental damage considered, to changes in the sample composition (size or/and time periods considered) and to differences in econometric specifications.

The case of air pollutants is suggestive, particularly the one for CO_2 emissions. Many authors make use of different versions of the database from the Carbon Dioxide Information Analysis Center (CDIAC) to test the EKC hypothesis with a panel of world countries. Holtz-Eakin and Selden (1995) (HES95), Heil and Selden (2001) (HS01) and Schmalensee et al. (1998) (SSJ98) use similar countries' panel data sets including over 120 countries and covering roughly 40 years[6]; they estimate time- and country-fixed effects quadratic functions (HES95 and HE01) and a spline-regression model with the same fixed effects (SSJ98). HES95 and HE01 find U-inverted shapes with very different turning points, ranging from US$35,000 to several millions depending on whether *per capita* income and emissions are measured in levels or in logarithms. SSJ98 get a within sample maximum of US$10,000 with a 10-segment regression. A nonparametric pooled regression is used by Taskin and Zaim (2000) to investigate the link between a CO_2 environmental efficiency index and GDP *per capita* for 52 countries over 1975-1990. Their results point towards a third order polynomial specification. A semiparametric version of the time- and country-fixed effects models used by HES95, HS01, and SSJ98 is estimated by Bertinelli and Strobl (2005) for a panel[7] of 122 countries over the 1950-1990 period. They find that the pooled regression are monotonically increasing.

Recently, Dijkgraaf and Vollebergh (2005) and Azomahou et al. (2006) tackle the fundamental assumption of poolability for CO_2-IER panels in parametric or nonparametric

---

[6] HES95, HE01 and SSJ98 make use of respectively 130, 135 and 141 countries and the time span is 1951-1986, 1951-1992 and 1950-1990.

[7] In that case, the data come from the World Resource Institute.

frameworks respectively. Focusing on the sample of 24 OECD countries mainly responsible for the U-inverted shape found in HES95, HS01 and SSJ98, Dijkgraaf and Vollebergh (2005) compare directly different versions of fixed-effects models to country-specific time-series regressions (with and without trends) and conclude that less than half (11) of the OECD countries display the U-inverted shape depicted by the pooled fixed-effects estimates. Azomahou et al. (2006) check the structural stability of the *per capita* IER with a nonparametric poolability test for a panel of 100 countries over the 1960-1996 period. They conclude that there is a stable cross-sectional relationship through time which allows the pooling of the data. The pooled country-fixed effects nonparametric regression displays a monotonically increasing pattern. In addition, nonparametric estimates are shown to be preferred to parametric ones.

Some authors have carried IER estimates with panels at low level of spatial aggregation. List and Gallet (1999) use state levels of SO_2 and NO_x emissions for the US spanning from 1929 to 1994. They estimate IERs with *per capita* data and a linear trend. The state-fixed effects models produce global EKCs for all states; quadratic and cubic state-specific regressions also yield a majority of respectively 79% and 98% hump-shaped functions for SO_2 emissions and a rough 80% EKCs for NO_x with both specifications. However, the vast majority of the state-specific turning points fall outside the confidence interval for the peak produced by the fixed-effects models. With the same data, Millimet et al. (2003) compare pooled time- and individual-fixed effects cubic models and spline regressions with time- and state-fixed effects semiparametric specifications[8]. They show that while the EKC obtained for *per capita* NO_x emissions is robust to the estimation strategy, the functional forms for SO_2 vary substantially. However, the null hypothesis of equality between the spline or cubic models and the partial linear models is rejected for both pollutants. These authors also compute specific semiparametric estimates for selected US states[9] and they conclude that the EKC shape remains robust at the state level for NO_x, but the results for SO_2 are mixed. De Groot et al. (2004) utilise a panel dataset on Chinese provinces covering the period 1982-1997. They investigate the IER for wastewater, waste gas (aggregate emissions of CO_2, NO_x and SO_2) and solid waste from the industrial sector with the pooled region-fixed effects model. They contrast the results obtained when expressing the dependent variable in levels, *per capita* and intensity terms.

---

[8] The linear trend from state-fixed effects cubic models of List and Gallet (1999) are here replaced by time-fixed effects.

[9] The time-fixed effects are replaced by state-specific linear time trends.

The relationship is shown as being monotonically decreasing for wastewater regardless of the dependent variable, increasing (respectively decreasing) for waste gas with the explained variable in levels or *per capita* (respectively intensity) terms and very versatile for solid waste depending on the dependent variable used. More recently, Aldy (2005) tests the EKC hypothesis for production as well as consumption-based *per capita* CO_2 emissions in the US at the state level. The author globally validates the EKC shape with the state- and year-fixed effects quadratic models as well as with the spline regressions. He provides evidence of significant different peaks for both CO_2 series. When state-specific quadratic models are fitted, the equality of the estimated functions and EKC peaks between states is rejected despite the fact that the vast majority of the states does depict EKC-type relationships. Since the data span over a long time period, Aldy (2005) also controls for common stochastic trends in the time-series and concludes that only about 20% of the state-specific relationships were cointegrated[10].

1.3 The nonparametric approach

The previous EKC literature has not tested the appropriateness of the homogeneity assumption on both the cross-section and the time dimensions of panel data sets in a nonparametric framework. This section proposes a simple strategy to fill this gap.

Let us define a very general functional relationship between one pollutant and an income indicator in a panel framework:

$$p_{it} = g_{it}(y_{it}) + \epsilon_{it} \quad \text{with } i = 1, \ldots, N\,;\, t = 1, \ldots, T \tag{1.1}$$

where p_{it} represents *per capita* emissions for some pollutant in state i at time t, y_{it} and $g_{it}()$ are respectively the *per capita* income and an unspecified heterogeneous function for state i and time t and ϵ_{it} is an iid(0, σ_ε^2) error term. As reported by Vollebergh et al. 2005, equation (1.1) cannot be identified without further restrictions, since for each (i,t) combination one single observation (y_{it}, p_{it}) is available. Following Hsiao's F-test strategy (2003, Ch.2) for the parametric case, we can identify $g_{it}()$ by imposing some general homogeneity assumptions on the cross-sectional and time dimensions. We can assume that $g_{it}()$ is constant over time but varies across states, thus $g_{it}() = g_i()$. Alternatively,

---

[10] This result confirms the concerns raised by Perman and Stern (2003).

we can make the assumption that $g_{it}()$ is constant across states but varies over time, thus $g_{it}() = g_t()$. Therefore, two tests can be formulated :

$$H_0 : g_i(y_{it}) = g_j(y_{it}), \forall \text{i}, \forall \text{j} \qquad H_0^* : g_t(y_{it}) = g_s(y_{it}), \forall \text{s}, \forall \text{t}$$

$$H_1 : g_i(y_{it}) \neq g_j(y_{it}), \text{ for some } i \neq j \qquad H_1^* : g_t(y_{it}) \neq g_s(y_{it}), \text{ for some } t \neq s$$

H_0 is the individual or spatial homogeneity hypothesis and H_0^* is the temporal homogeneity hypothesis. Given that H_0^* is assumed to hold when testing H_0 (and *vice-versa*), accepting either H_0 or H_0^* yield to the same pooled regression $p_{it} = g(y_{it}) + \epsilon_{it}$. A number of procedures exist for testing equality of nonparametric regressions functions. Yatchew (2003) suggests a simple nonparametric test which compares the weighted sum of the residual variance of every individual nonparametric regressions ($i.e$ the unrestricted residual variance s^2_{unr}) with the residual variance of the nonparametric pooled estimate ($i.e$ the restricted residual variance s^2_{res}).

Under H_0 or H_0^*, the pooled estimates ($\hat{p}_{it}^{NPpool}$) at some *per capita* income level y_0 can be computed by the Nadaraya-Watson estimator:

$$\hat{g}(y_0) = \sum_{i,t} w_{it}(y_0) p_{it} = \frac{\sum_1^{NT} K(\frac{y_{it}-y_0}{\lambda}) p_{it}}{\sum_1^{NT} K(\frac{y_{it}-y_0}{\lambda})} \tag{1.2}$$

where $K()$ is a kernel function and λ is the bandwidth. We estimate the pooled nonparametric[11] regression by using a cross-validation[12] bandwidth and a gaussian kernel and we calculate its residual variance (s^2_{res}) by simply averaging the sum of squared residuals.

Under H_1 (H_1^*), there exist $Q = T$ cross-sectional ($Q = N$ time-series) distinct non-

---

[11] Equation (1.2) shows explicitly the intuition behind nonparametric regressions. The estimated conditional mean at the local point y_0, $\widehat{E(p_{it}|y_0)} = \hat{g}(y_0)$, is a weighted average of all NT p_{it} values of the panel, with weights inversly proportional to the distance between each of the NT y_{it} observations of the independent variable and the local value y_0. The kernel function $K()$ is a density-shaped function which defines the weights while the λ term simply determines how many of the NT y_{it} points are included in the neighborhood of y_0 to compute the local conditional mean. The larger the bandwidth λ, the closer each local conditional mean to the unconditional mean and the smoother the estimate.

[12] In large samples, selecting λ through cross-validation is the same as computing the bandwidth that minimizes the integrated mean-squared error. This method balances optimally the bias and the variance of the estimate.

parametric regressions. Let $q = 1, \cdots, Q$ be the q^{th} subpopulation of size $n_q = N$ ($n_q = T$). The weighted sum of unrestricted residual variances (s^2_{unr}) can be computed by making use of mth order differencing estimators[13]. Yatchew (2003, Ch.4) shows that if we make use of the optimal bandwidth for pooled estimates, optimal differencing weights in s^2_{unr} and under the classical assumptions that the errors are iid(0,σ^2_ε) and independent between and within subpopulations, H_0 and H_0^* can be tested with the following statistic:

$$V = (mn)^{\frac{1}{2}} \frac{(s^2_{res} - s^2_{unr})}{s^2_{unr}} \xrightarrow{D} N(0,1) \tag{1.3}$$

where:

m is the order of differencing,

$$n = NT = \sum_{q=1}^{Q} n_q, \; q = 1, \dots, Q \text{ and where } \begin{cases} Q = N \text{ and } n_q = T \text{ if we test } H_0 \\ Q = T \text{ and } n_q = N \text{ if we test } H_0^*, \end{cases}$$

$$s^2_{res} = \frac{1}{n} \sum_{i=1}^{N} \sum_{t=1}^{T} (p_{it} - \hat{p}_{it}^{NPpool})^2,$$

$$s^2_{unr} = \sum_{q=1}^{Q} \frac{n_q}{n} s^2_{diff,q},$$

$$s^2_{diff,q} = \frac{1}{n_q} \sum_{r=1}^{n_q - m} (d_0 p_{q,r} + d_1 p_{q,r+1} + d_2 p_{q,r+2} + \cdots + d_m p_{q,r+m})^2,$$

$d_0, d_1, d_2, \cdots, d_m$ are differencing weights that satisfy $\sum_{k=0}^{m} d_k = 0, \sum_{k=0}^{m} d_k^2 = 1.$

This test[14] is one-sided, so we do not accept H_0 (or H_0^*) at the 95% confidence level if the empirical V is greater than 1.645. An important advantage of this test procedure is that it can easily be modified to check different kinds of null hypotheses. If the poolability assumption (H_0 or H_0^*) is accepted, we can verify the pertinence of conditioning $E(p_{it})$ on

---

[13] Note that the data must be previously reordered so that within each subpopulation the $(y_{q,1}, p_{q,1}), (y_{q,2}, p_{q,2}), \cdots, (y_{q,n_q}, p_{q,n_q})$ observations are in increasing order relative to the y's.

[14] When the residuals are heteroscedastic with unknown covariance matrix Ω, the denominator in equation (1.3) can be replaced, without modifying the asymptotic properties of the V statistic, by $\xi = \frac{1}{m}(\frac{1}{n}\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}_{-\mathbf{1}} + \cdots + \frac{1}{n}\hat{\boldsymbol{\epsilon}}'\hat{\boldsymbol{\epsilon}}_{-\boldsymbol{m}})$, where $\hat{\boldsymbol{\epsilon}}$ is the vector of the pooled nonparametric regression residuals and the subscript $-i$ stands for the lag order of $\hat{\boldsymbol{\epsilon}}$. Note also that, under the null hypothesis, s^2_{unr} in equation (1.3) can be replaced by s^2_{res} because both estimators of the residual variance are consistent, see Yatchew (2003, p.64).

y_{it} by replacing in equation (1.3) $\hat{p}_{it}^{NPpool}$ by $\widehat{E(p_{it})}$ in s^2_{res} and s^2_{unr} by s^2_{diff}, where s^2_{diff} is simply the residual variance differencing estimator applied to the p_{it} data as a whole. The same idea can be followed to compare parametric and nonparametric specifications[15].

Given the strong independence assumption imposed on the residuals, we also tested H_0^* by computing the Baltagi et al. (1996) J statistic[16], which allows the error term to have an arbitrary form of serial correlation and/or conditional heteroscedasticity on the time dimension or to include individual effects. As for the V statistic, the J statistic follows a N(0,1) distribution and the test is one-sided.

Panel structures rarely display enough homogeneity to allow estimations under H_0 or H_0^*. Therefore, the vast majority of the IER literature attempts to capture the time and spatial nonhomogeneities by assuming isomorphic functions through time and individuals up to some vertical deterministic shifts or intercept term (the so-called 'fixed effects'). This makes $g_{it}()$ becomes a semiparametric specification of the form $g_{it}() = \varphi_{it} + z(y_{it})$. Taking it further, the latter model becomes fully parametric by imposing $z(x_{it}) = \sum_{k=1}^{K} \alpha_k x_{it}^k$. Consequently, the fixed-effects assumption transforms equation (1.1) into the following two standard fixed-effects models:

$$p_{it} = \varphi_{it} + z(y_{it}) + \upsilon_{it} \tag{1.4a}$$

$$P_{it} = \alpha_{0it} + \sum_{k=1}^{K} \alpha_k y_{it}^k + \eta_{it}, \; k = 1, \cdots, K \tag{1.4b}$$

where the intercepts φ_{it} and α_{0it} in equations (1.4a) and (1.4b) are linear nonstochastic fixed effects which gauge unobserved state-specific factors that affects the differences in *per capita* emissions as well as time-specific factors which capture macroeconomic effects, changes in environmental legislation, etc; $z(y_{it})$ and $\sum_{k=1}^{K} \alpha_k x_{it}^k$ respectively in models

---

[15] *Ibid.* The null hypothesis that a known parametric regression function estimated by Least Squares $h(y_{it}, \gamma^{LS})$ is similar to some pooled pure nonparametric alternative $f(y_{it})$ can checked by replacing $\hat{p}_{it}^{NPpool}$ by $\hat{p}_{it}^{LS}$ in s^2_{res} and applying s^2_{diff} to p_{it} in equation (1.3).

[16] This statistic must be computed ensuring that some specific conditions on arbitrary parameters are satisfied, cf. Baltagi et al. (1996, p.349, condition C3). Note that the asymptotic properties of the J statistic relies on convergence properties of the residuals and not on differences between sum of squares of the pooled and unpooled nonparametric regressions. We would like to thank P. Nguyen Van for providing the Gauss code that we adapted to R.2.4.1 to compute this test. All errors are my own.

(1.4a) and (1.4b) are the unrestricted and restricted[17] common functional forms to each year as well as to each state of the panel; υ_{it} and η_{it} are stochastic error terms, both assumed iid over t and i and of mean 0 and constant variance (σ^2_υ and σ^2_η).

Model (1.4a) is a partial linear model which can be consistently estimated in three ways: (i) by Robinson (1988)'s double residuals as in Millimet et al. (2003), Bertinelli and Strobl (2005) or Nguyen Van and Azomahou (2007); (ii) by differencing as in Yatchew (2003, Ch. 4.5); or (iii) by replacing $z()$ by a consistent nonparametric estimate (some spline smoother of order r) and minimizing a penalised residual sum of squares. The latter method has been preferred because of its operational simplicity in R's statistical environment[18]. Equation (1.3) can be applied in the spirit of a specification test to assess if the semiparametric model consistently captures the temporal or spatial heterogeneity. When the partial linear regression (1.4a) is not rejected, the pertinence of including its linear term φ_{it} can be tested with a slightly modified version of the V-stat procedure, which is equivalent[19] to the standard linear restrictions test $R\beta = r$.

---

[17] The polynomial function is usually limited to K=3 when checking the EKC hypothesis. When the coefficient of its linear component is positive and significant, the coefficient of the quadratic component is negative and significant and the slope of the cubic component is nonsignificant, the EKC hypothesis is validated.

[18] This procedure consist in minimizing

$$\min_{\beta,r,\lambda} \sum_{i=1}^{n} (y_i - X_i\beta - \theta(Z_i - z, r))^2 + \lambda \int_{z_{min}}^{z_{max}} [\theta''(z)]^2 \mathrm{d}x,$$

where $\theta()$ is a r^{th}-order polynomial function and the integrated term is a roughness penalty. The *gam* function in the *mgcv* package proposes a consistent procedure to fit Generalized Additive Models that can be used to estimate semiparametric specifications. See Wood (2006) for further details.

[19] Yatchew (2003, p.179) shows that

$$(R\hat{\beta} - r)'(R\hat{\Sigma}_\beta R')(R\hat{\beta} - r) = \frac{n(s^2_{res} - s^2_{unr})}{s^2_{unr}(1 + \frac{1}{2m})} \xrightarrow{D} \chi^2_{rank(R)}$$

where the right-hand side ratio correspond to the modified V-stat. This equality is directly linked to the differencing estimation method for the semiparametric model. Following Yatchew (2003, Ch. 4.5), we can rewrite the SP model (1.4a) in matrix notation as $p = F\varphi + z(y) + \upsilon$. The nonlinear component $z(y)$ can be removed by differencing, i.e. $Dp = DF\varphi + Dz(y) + D\upsilon \approx DF\varphi + D\upsilon$, where D is a $(n \,\mathrm{x}\, n)$ differencing matrix. The OLS estimator of φ is therefore given by $\hat{\varphi}_{ols} = [(DF)'(DF)]^{-1}(DF)'Dp$. With these notations at hand, the components of the modified V-stat can be defined as $s^2_{unr} = \frac{1}{n}(Dp - DF\hat{\varphi}_{ols})'(Dp - DF\hat{\varphi}_{ols}))$, $s^2_{res} = \frac{1}{n}(Dp)'(Dp)$ and D is the differencing matrix of order m computed with optimal weights. Note that the p's can then be purged from its parametric effects $(p - F\hat{\varphi}_{ols})$ and a standard nonparametric method can be

Finally, model (1.4b) is the standard parametric model used to check the EKC hypothesis. Most authors control for fixed effects by applying the F-test that involves the sum of squared residuals from the pooled (SSR_p) and within (SSR_w) versions of model (1.4b). However, they omit a comparison of these magnitudes with the unrestricted[20] sum of squared residuals (SSR_u). We apply in section 1.5 the full F-test strategy on the spatial and time dimension.

1.4 Data description

Our database is a balanced panel of 48 Spanish provinces over the 1990-2002 period. The series come from two different sources. Spanish provinces' statistics for population and GDP, in constant 1996 USD and adjusted to PPP, are taken from Herrero et al. (2004). We focus on 48 provinces[21] whose air pollutant emissions are included in the inventory provided by Spain to the Convention on Long-Range Transboundary Air Pollution (CLRTAP). The annual emissions data on atmospheric pollutants have been supplied to us by the Spanish Ministry of the Environment and are extracted from the European Corinair 1990 inventory[22]. These data contain the anthropogenic and natural emissions of eight pollutants, split at the most aggregated level into eleven source groups[23]. To be consistent with our purpose, we excluded the natural emissions category and considered only the anthropogenic ones.

The pollutants included in the Corinair 1990 inventory are methane (CH_4), carbon monoxyde (CO) and dioxyde (CO_2), nitrous oxide (N_2O), ammonia (NH_3), non-methanic volatile organic compounds (NMVOC), nitrogen (NO_X) and sulphur oxydes (SO_X). In order to keep our analysis manageable, we focus on four of them, CH_4, CO_2, CO and NMVOC, which present very different evolution patterns at the aggregate level. The first two (CH_4, CO_2) are greenhouse gases for which Spain has commited, under the Kyoto

---

applied to get the estimated nonlinear portion of the semiparametric model ($\hat{z}(x)$).

[20] This term is contructed from either the cross-sectional parametric regressions for all years or the time-series parametric regressions for all regions/countries, see Hsiao (2003, Ch.2).

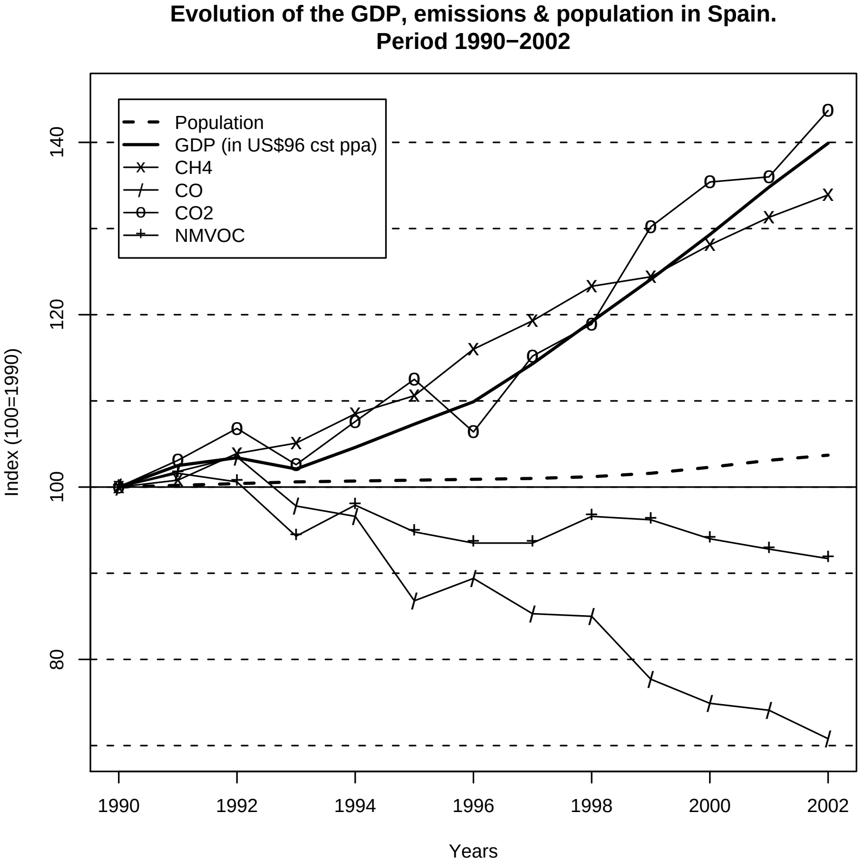
[21] See Appendix 1.7. Spain comprises 50 provinces. We excluded the overseas provinces of Las Palmas and Tenerife.

[22] Note that Roca et al. (2001) used the same database at the national level for different periods in a parametric context.

[23] These eleven categories are the first level of the Selected Nomenclature for Air Pollution (SNAP) and can be further divided into 57 sub-sectors, which include 277 detailed activities.

Protocol, not to increase emissions by more than 15% over the 1990 level by 2012. CO is a poisonous gas and NMVOC is a ground level ozone precursor. In 1990, three main sectors were the source for the majority of emissions: power generation (SNAP-group 1) for CO_2; road transport (SNAP-group 7) for CO and NMVOC; and agriculture (SNAP-group 10) for CH_4. Note that, according to this inventory, nature rarely accounts for more than 5% of global emissions in Spain, except for NMCOV where it represents a roughly stable 45% share between 1990 and 2002.

Figure 1.1: Spanish GDP, emissions and population. Period 1990-2002.



Source: Spanish Ministry of Environment (MMA) for air pollutants and Herrero et al. (2004) for GDP and population.

Figure 1 shows the evolution of aggregate anthropogenic Spanish emissions for the retained air pollutants. This figure also shows the changes of the Spanish population and

GDP over the sample period. CO_2 emissions clearly follow the exponential upward trend of GDP, while CH_4 emissions grow along a fairly linear path since 1990. NMVOC and CO emissions have been declining at different rates over the period, 8.3% and 29.2% respectively.

Table 1.1 presents some descriptive statistics on *per capita* emissions and real GDP for the whole panel. We can observe that the mean of the variables is always higher than the median, suggesting the presence of extreme values at the right tail of the data distributions. The standard deviation remains close to, or below, the median for most of the variables except for CO_2. A more accurate picture of the variability of the panel on its temporal and spatial dimensions is given by a one-way analysis of variance.

Table 1.1: Descriptive statistics. Provincial GDP, emissions and population in Spain. Period 1990-2002.

Variables	Median	Mean	Std. dev.	Min.	Max
CH_4	48.4	68.50	53.6	7.3	263.0
CO	98.3	107.6	49.4	17.8	317.8
CO_2	6146.7	8818.3	9071.3	836.0	68013.4
NMVOC	47.2	57.8	31.0	13.5	158.1
GDP	1.5	1.6	0.4	0.9	2.7
Obs.			624		

Note: All figures are *per capita*. Spanish provinces anthropogenic air pollutant emissions are in kg and real GDP in 10'000 USD1990 corrected by PPP.

Table 1.2 summarizes the data inter- and intra-variation for provinces and years. Variation here is predominantly 'between' provinces, ranging from 80.9% for *per capita* GDP to 98.5% for NMVOC *per capita* emissions, while it is higher 'within' than 'between' years and it varies from 84.5% for GDP *per capita* to 99.6% for NMVOC *per capita* emissions. Note that an ANOVA analysis (F-tests) always reject strongly the equality of the regional means for all the variables while the equality of the temporal means is accepted for *per capita* CH_4, CO_2 and NMVOC emissions. These results indicate that between-region variation is a major source of variation in our panel.

Table 1.2: One-way analysis of variance. Provincial GDP, emissions and population in Spain. Period 1990-2002.

Variables	σ^2_{tot}	$\sigma^2_{b,i}$	$\sigma^2_{w,i}$	$\sigma^2_{b,t}$	$\sigma^2_{w,t}$
CH_4	2865.9 (100%)	*96.0%	*4.0%	1.4%	98.6%
CO	2442.6 (100%)	*90.7%	*9.3%	*5.2%	*94.8%
CO_2	82.3 (100%)	*96.1%	*3.9%	1.2%	98.8%
NMVOC	963.1 (100%)	*98.5%	*1.5%	0.4%	99.6%
GDP	14182 (100%)	*80.9%	*19.1%	*15.5%	*84.5%

Note: *: significant at the 5% level. All figures are in *per capita* terms. CH_4, CO, and NMVOC are in kg, CO_2 in tonnes and GDP in USD90 and PPP-corrected. Total, between and within variances are given by σ^2_{tot}, σ^2_b, σ^2_w. The ratios of the mean squares are F-distributed with (47;576) and (12;611) degrees of freedom respectively. The corresponding critical F-values are 1.384 and 1.768.

1.5 Econometric analysis

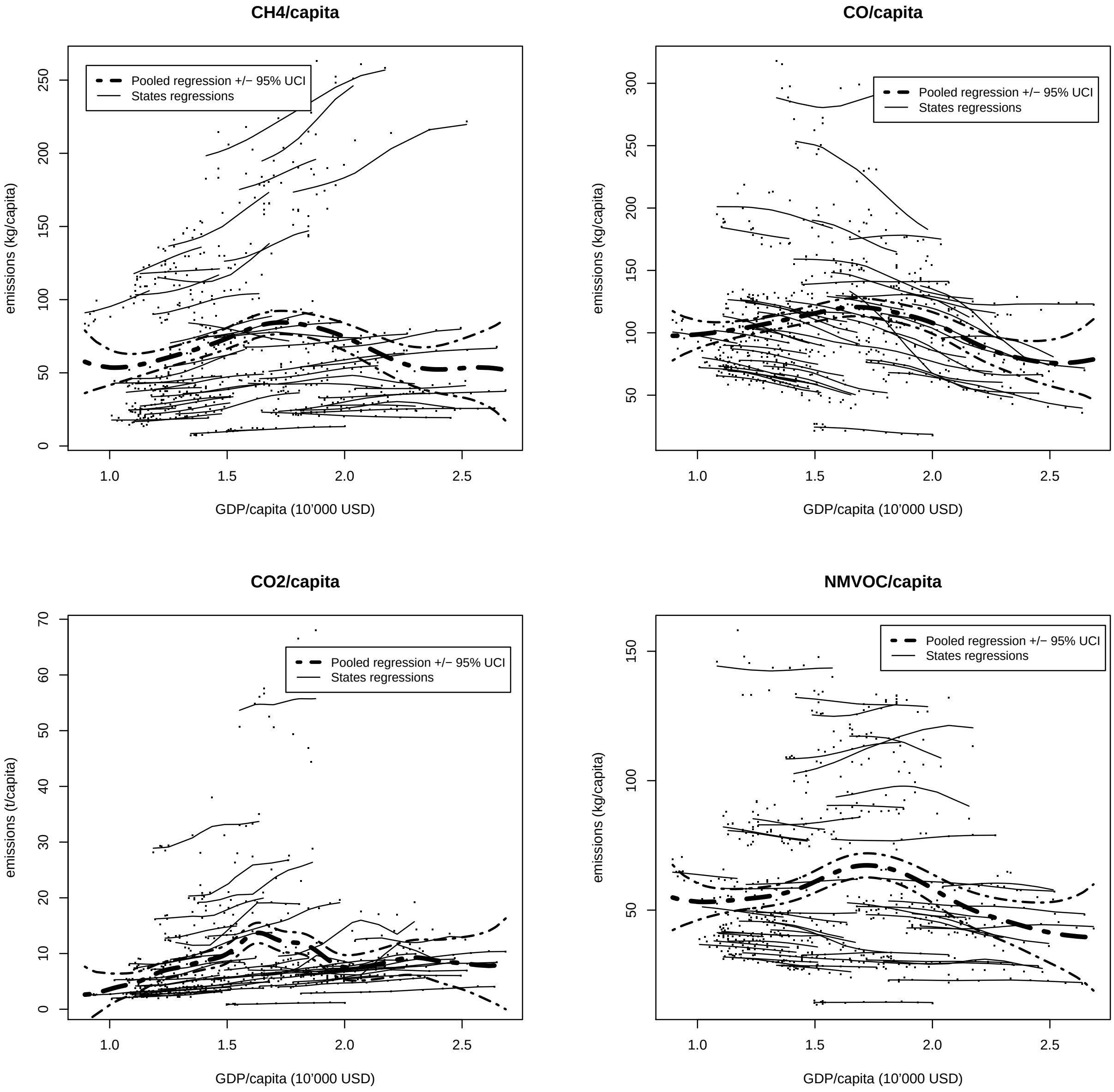
Nonparametric regressions are usually investigated through graphical devices. For each pollutant, Figure 2 compares the nonparametric pooled regression with nonparametric time-series regressions for each province; it roughly checks the equality of the IER between regions, *i.e.* the spatial homogeneity hypothesis. Figure 3 compares, for each pollutant, the pooled regression with nonparametric cross-sectional regressions for selected years and aims at investigating the structural stability of the relationship through time, *i.e.* the time homogeneity hypothesis. In all graphs, the nonparametric pooled regression is surrounded by the 95% uniform confidence band[24] suggested by Yatchew (2003, p.36). It contrasts graphically the equality between different pooled nonparametric and parametric functions by controlling whether the parametric shape falls within the whole confidence band.

Spatial heterogeneity. It is clear from a visual inspection of the four panels in Figure 2 that the pooled model with a single constant should be rejected as almost none of the region-specific regressions lie within the 95% confidence band. The existence of a common function for every province up to a vertical shift is neither strongly supported. Table 1.3 reports the results of the statistical tests described in section 1.3. In lines 1 and 2 we can see the V-tests strongly reject the H_0 hypothesis for all pollutants as well as the semiparametric specification. Consequently, the pooled nonparametric and partial linear

---

[24] This interval is more interesting than the pointwise one as 95% of the estimated confidence intervals contain the entire true function.

Figure 1.2: Spatial heterogeneity. Time series GDP-emissions fits for Spanish provinces. Period 1990-2002.



Note: Nadaraya-Watson nonparametric regressions with gaussian kernel. Pooled estimates computed with cross-validation bandwidth.

estimates do not capture consistently the state-specific IERs[25]. Poolability is therefore

---

[25] It is apparent in the panels of Figure 2 that clusters of regions with close income-emissions

Table 1.3: Spatial homogeneity tests for the GDP-emissions panel fits in Spain. Period 1990-2002.

Test type	Null Hypothesis	Df. n.	Df. d.	5% cutoff	CH_4	CO	CO_2	NMVOC
Pooled nonparametric and semiparametric regressions								
V-test	$g_i(y_{it}) = g_j(y_{it})$, ∀i, ∀j	-	-	1.65	20.20*	16.99*	16.15*	18.64*
V-test	$\varphi_i + z(y_{it}) = g_i(y_{it})$, ∀i	-	-	1.65	10.52*	10.16*	2.03*	9.78*
Pooled parametric (cubic) regressions with and without individual-fixed effects								
F-test	$\alpha_{0i} = \alpha_0$; $\alpha_{ki} = \alpha_k$, ∀i	188	432	1.22	278*	125*	105*	290*
F-test	$\alpha_{ki} = \alpha_k$, ∀i	141	432	1.24	7.81*	5.90*	1.21	2.65*
F-test	$\alpha_{0i} = \alpha_0 \mid \alpha_{ki} = \alpha_k$, ∀i	47	573	1.38	-	-	394*	-

Note: *: significant at the 5% level. The value of the V-statistic can vary depending on the order of differencing m used to compute the variance differencing estimator. We took the conservative option to fix $m = 1$ for all pollutants. Increasing m tends to increase the empirical V-stat. This latter statistic is always the version robust to heteroskedasticity (see Yatchew (2003)) and uses optimal differencing weights. The semiparametric regressions are estimated with the *gam* function from the *mgcv* package. All computations have been implemented on R.2.4.1.

rejected with tests robust to functional misspecification. The standard F-tests applied to the cubic[26] parametric models yield similar results for most of the air pollutants. In lines 3 and 4 of Table 1.3 we clearly reject the joint hypothesis of equality of intercepts and slopes in all cases, as well as the common slopes assumption for almost all the pollutants. The only exception concerns CO_2 emissions, for which state-fixed effects should be included in the cubic[27] model. However the latter results are not supported by the nonparametric test.

These findings confirm those reported in section 1.2 by List and Gallet (1999), Millimet et al. (2003) and Aldy (2005) for the SO_X and CO_2 emissions in the US states. We reject the common IER in all Spanish provinces and for all the investigated air pollutants. This also corroborates the main message of the theoretical body presented in section 1.2: the shape of the IER is very sensitive to regional/country-specific factors. As these differences are expected to be lower within regions pertaining to the same country than between countries, our results highlight the potential bias introduced by the lack of variables which pick up the regional or country differences when investigating the IER with fixed-effects panel data. Another interesting point in Figure 2 is that global pollutants (CH_4 and CO_2) are

---

patterns could be investigated and may show spatial homogeneity. However, the information at hand do not allow a systematic grouping of the provinces according to existing theories. By its very nature, using the reduced form model suggested by the EKC hypothesis render any structural interpretation arbitrary. That is why no attempt is made here to find spatial homogeneous clusters.

[26] The results for the quadratic specifications are similar and available upon request.

[27] For CO_2 emissions, the empirical F for the quadratic model is 146.6 for the joint equality of intercepts and slopes and 1.67 for the common slopes. Compared to $F_{(5\%;141;480)} = 1.24$ and to $F_{(5\%;94;480)} = 1.28$ respectively, we reject both null hypotheses.

increasing with GDP in most of the provinces while local pollutants (CO and NMVOC) are stabilised or decreasing. This is consistent with the political economy of environmental protection, which points toward more stringent policies when the environmental damage is local[28].

Temporal homogeneity. Time poolability would require that the cross-regional regressions for most of the 13 years[29] lie close to the pooled nonparametric estimates shown in Figure 3, as stated by H_0^*, or that most of the vertical shifts of the yearly cross-sectional regressions consist of parallel translations of a common function. The pooled nonparametric and cubic estimated functions describe well-defined U-inverted shapes for all the pollutants and so do the cross-sectional estimates for the first, middle and last year of the panel. However, the latter appear to be shifted horizontally and vertically, preserving approximately their shape but leading to different turning points each year. Clearly, the abscissa of the turning points always increases through time while the location of the ordinate depends on the underlying dynamic of the specific air pollutant. When *per capita* emissions (CH_4, CO_2) for most of the Spanish provinces are increasing with *per capita* income on Figure 2, the turning points in Figure 3 move to the north-east. When the estimated state-specific functions are mainly decreasing (constant), as for *per capita* CO (NMVOC) emissions, the turning points move to the south-east (east). Note, however, that the selected year-specific cross-sectional regressions in Figure 3 lie close to the 95% uniform confidence band, whatever the pollutant investigated.

Table 1.4 reports the results for the homogeneity tests applied to the time dimension. Lines 1 and 2 examine H_0^* and compare the J and V statistics. We accept the temporal homogeneity with both methods for three out of four air pollutants (CH_4, CO, CO_2). We reject H_0^* for NMVOC emissions with both J and V-stat at the 5% significance level. Consequently, the two nonparametric procedures converge to the same conclusion. We conclude that the horizontal and vertical shifts of the yearly regressions for the CH_4, CO and CO_2 panels in Figure 3 are not statistically significant. However, the horizontal translation over time for the cross-sectional NMVOC-IER is significant.

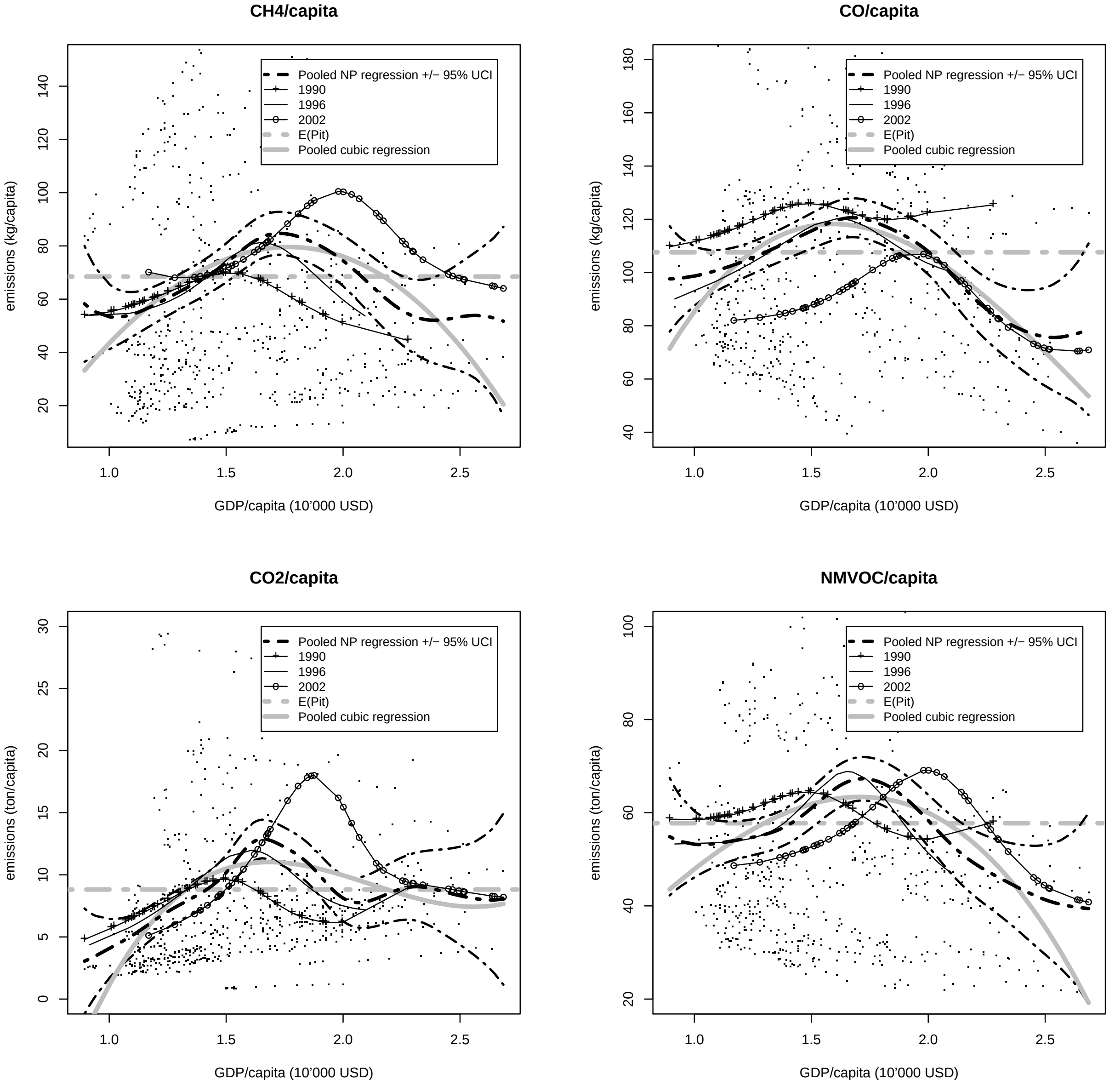
In line 3 of Table 1.4, we go a step further and contrast the partial linear models with year dummies with the cross-sectional nonparametric estimates for each year. We accept

---

[28] We thank an anonymous referee for pointing this out.

[29] In Figure 3, we only show years 1990, 1996 and 2002 to keep the graphs readable.

Figure 1.3: Cross-sectional GDP-emissions fits for Spanish provinces. Period 1990-2002.



Note: Nadaraya-Watson nonparametric regressions with gaussian kernel. Pooled estimates computed with cross-validation bandwidth.

the equality of both specifications for the same previous group of pollutants and reject it for NMVOC. For the latter pollutant, time poolability is therefore rejected. Line 4 indicates that the coefficients for the time-fixed effects are jointly equal to zero for CH_4 and

CO_2. Consequently, the pooled nonparametric regressions consistently capture the cross-sectional regressions over the whole period. For CO, even if neither of the pooled and semiparametric specifications is rejected versus the unrestricted regressions, the pooled regression is rejected in favour of the semiparametric one in line 4. Time-fixed effects appear to be appropriate in the CO case. Line 5 in Table 1.4 explores the relevance of conditioning the annual *per capita* emissions on *per capita* real GDP when the data are poolable. We reject the simple mean in favour of the conditional mean for two out of three cases at the 5% significance level[30]. The equality between the two means for CO_2 emissions would have been rejected at the slightly relaxed cutoff of 10%. In line 6, the pooled estimates of the cubic models are compared with the nonparametric ones. Nonparametric regressions do perform generally better[31] than cubic OLS models but the differences are not always significant as only two out of the four parametric specifications are rejected.

Lines 7 to 9 from Table 1.4 contain the results of Hsiao's poolabilty test strategy. The first two F-tests compare cross-sectional regressions for each year with respectively a pooled cubic regression (SSRp vs SSRu) and a cubic regression with time-fixed effects (SSRw vs SSRu). The third F-test verifies the adequacy of including time-fixed effects in the pooled data (SSRw vs SSRp).

The results are similar to the nonparametric ones for CH_4, CO and CO_2. We accept the poolabilty of the data for all pollutants and reject the time-fixed effects specification for all pollutants with the exception of CO *per capita* emissions. Contrary to the V-tests, the F-tests do not reject the poolabililty for the NMVOC-IER, suggesting a misspecification bias in the parametric procedure[32]. The equality between the unconditional and conditional mean is also rejected for all the pooled parametric estimates in line 10. Finally, line 11 and 12 compare the turning points for the nonparametric and parametric pooled regressions. The ordinates of the turning points are systematically larger for nonparametric specifications.

---

[30] On Figure 3, we notice that the emissions' unconditional means (dashed grey lines) lie close to or above the upper uniform confidence intervals for low GDP levels, below it in the turning point proximity and close to or above the confidence band for high GDP level for most of the pollutants. This suggests that the pooled relationship is concave.

[31] A positive V-stat indicates that the residual variance for the nonparametric regression is lower than the parametric one.

[32] A wrong specification of the parametric function can lead to a false acceptance of the poolability assumption with the F-test. This seems to be the case here as the equality of the nonparametric and parametric pooled estimates are rejected in line 6 of Table 1.4 for NMVOC.

Table 1.4: Temporal homogeneity tests for GDP-emissions panel fits in Spain.

Test type	Null Hypothesis	Df. n.	Df. d.	5% cutoff	CH_4	CO	CO_2	NMVOC
Pooled nonparametric and semiparametric regressions								
J-test	$g_t(y_{it}) = g_s(y_{it}),\ \forall$t, $\forall$s	-	-	1.65	-1.65	-0.38	0.89	1.66^{*}
V-test	$g_t(y_{it}) = g_s(y_{it}),\ \forall$t, $\forall$s	-	-	1.65	-0.93	-0.18	0.68	1.88^{*}
V-test	$\varphi_t + z(y_{it}) = g_t(y_{it}),\ \forall$t	-	-	1.65	-1.32	-1.16	0.78	1.84^{*}
V-test$^{(a)}$	$\varphi_t + z(y_{it}) = g(y_{it}),\ \forall$t	-	-	21.03	7.22	26.11^{*}	4.80	-
V-test	$E(P_{it}) = E(P_{it} \mid y_{it}) = g(y_{it})$	-	-	1.65	4.27^{*}	1.79^{*}	1.64	-
V-test	$\alpha_0 + \sum_{k=1}^{3} \alpha_k y_{it}^k = g(y_{it})$	-	-	1.65	3.80^{*}	0.83	0.55	3.02^{*}
Pooled parametric (cubic) regressions with and without time-fixed effects								
F-test	$\alpha_{kt} = \alpha_k,\ \forall$t	48	572	1.38	0.87	0.90	0.72	0.68
F-test	$\alpha_{0t} = \alpha_0\,;\ \alpha_{kt} = \alpha_k,\ \forall$t	36	572	1.44	0.91	0.29	0.89	0.85
F-test	$\alpha_{0t} = \alpha_0 \mid \alpha_{kt} = \alpha_k,\ \forall$t	12	608	1.77	0.75	2.85^{*}	0.23	0.17
F-test	$\alpha_0 \neq 0\,;\ \alpha_k = 0$	3	620	2.6	8.53^{*}	11.70^{*}	14.80^{*}	8.55^{*}
Turning points of the pooled nonpapametric and parametric(cubic) regressions								
Max$(\hat{p}_{it} = \hat{g}(y_{it}))^{(b)}$					[1.77;79.6]	[1.59;118.1]	[1.68;11.0]	[1.71;63.4]
Max$(\hat{\alpha}_0 + \sum_{k=1}^{3} \hat{\alpha}_k y_{it}^k)^{(b)}$					[1.73;84.4]	[1.68;120.6]	[1.64;13.7]	[1.72;67.3]

Note: *: significant at the 5% level. The J-statistic has been computed with $c = 1$, $\alpha = 5$ and $\alpha^{'} = 2$ (cf. footnote 15). These results are robust for almost all combinations of $c = (0.8, 1, 1.2)$ with $(\alpha, \alpha^{'}) = (5, 2)$. The value of the V-statistic can vary depending on the order of differencing m used to compute the variance differencing estimator. We took the conservative option to fix $m = 1$ for all pollutants. Increasing m tends to increase the empirical V-stat. This latter statistic is always the version robust to heteroskedasticity and uses optimal differencing weights (see Yatchew (2003)). The semiparametric regressions are estimated with the *gam* function from the *mgcv* package. All computations have been implemented on R.2.4.1. $^{(a)}$: This V-test is a slightly modified version of equation (1.3) which follows a $\chi^2_{rank(R)}$ distribution, see *ibid*. $^{(b)}$: [a;b] represent the maximum's abscissa and ordinate for the pooled regressions.

In sum, these results show that a U-inverted pooled regression is compatible with different income-emissions dynamics at the regional level. Structural stability for CH_4, CO and CO_2 suggests that the underlying data generation process in all regions appears to be stable over the 1990-2002 period, *i.e.* the regions tend to keep their relative position when cross-sectional IER are estimated for different years. This may be good news if the income-pollution relation in the regions is mainly decreasing. In Spain, for the period investigated, this is the case for CO emissions. But when the underlying dynamic is increasing, as for the greenhouse gases CH_4 and CO_2, structural stability indicates that no offseting force is at work to change the underlying dynamic of the IER.

1.6 Conclusion

In this paper, we use a balanced panel of 48 Spanish provinces on four pollutant emissions (CH_4, CO, CO_2 and NMVOC) covering the 1990-2002 period to investigate systematically the time and spatial heterogeneity which characterizes the relationship between *per capita* air pollutant emissions and *per capita* income.

In order to avoid functional specification bias, we follow many authors who turned to nonparametric estimation techniques to model this reduced form function. Most of them made the implicit assumption that every region or country included in the panel shares the same pollution-income relationship, up to some specific fixed temporal and/or individual effects. Our findings show that the temporal poolability assumption holds in the Spanish provinces for three (CH_4, CO, CO_2) out of four air pollutants when poolability tests robust to functional misspecification are employed. The pooled nonparametric regressions give rise to U-inverted income-pollution relations. However, these hump-shaped functions only reflect relatively short-run cross-sectional regressions for different periods. Our parametric and nonparametric tests reject overwhelmingly the null hypothesis of spatial homogeneity as well as the goodness-of-fit of the time- or individual-fixed effects semiparametric models. Investigating the reduced form function of the pollution-income relationship for *per capita* air pollutants emissions with fixed-effects panel data models, be it parametric or semiparametric, failed to account for differences in functional shapes between regions. It is likely that heterogeneity would be even greater when applied to cross-country rather than cross-regional panels.

Having established that spatial heterogeneity matters in panels and blurs the EKC picture, one may consider three avenues of future research when assessing the income-pollution reduced form function. From an econometric point of view, the evidence points toward the use of estimation methods which better account for differences between countries/regions such as non- or parametric quantile regressions, parametric random coefficients estimators, varying smooth coefficients models or country/region specific regressions when large individual series are available. From an economic point of view, the persistent heterogeneity in patterns between regions/countries support the extreme sensitivity of the income-pollution relation to differences in regional/country-specific factors, such as factors endowment, sources of growth, differences in production and abatment technologies or local sensitivity to environmental damages. At the same time, structural stability through time points toward the existence of some stable structural determinants which shape the income-pollution relation. The identification of these determinants certainly would deserve more research effort.

Acknowledgments

I am grateful to Ángeles Cristóbal and Natalia Alonso Sopeña from the Spanish Ministry of Environment (MMA) and to Francisco Perez and Rodrigo Aragón Rodríguez from the Instituto Valenciano de Investigaciones Económica (Ivie) for providing the data used in this paper. The final draft has been written while visiting the University of Guelph and the University of Toronto and has benefited from helpful comments when presented in courses 6170, 6800 and 2403 respectively. I sincerely acknowledge the hospitality of both Pr. McKitrick and Pr. Yatchew and the kindness of the secretaries. I am grateful to the Swiss National Research Foundation (under grant FN PBNE1-114686) and to Irene for their financial support. I am also indebted to an anonymous referee whose criticism helped significantly to improve the paper during the revision process for *Environmental and Resource Economics*. All remaining errors are the author's sole responsibility.

1.7 Appendix

1.7.1 List of Spanish Provinces

Alava, Albacete, Alicante, Almería, Asturias, Ávila, Badajoz, Barcelona, Burgos, Cáceres, Cádiz, Cantabria, Castellón, Cuidad Real, Córdoba, La Coruña, Cuenca, Gerona, Granada, Guadalajara, Guipúzcoa, Huelva, Huesca, Jaén, León, Lérida, Lugo, Madrid, Málaga, Múrcia, Navarra, Orense, Palencia, Pontevedra, La Rioja, Salamanca, Segovia, Sevilla, Soria, Tarragona, Teruel, Toledo, Valladolid, Valencia, Vizcaya, Zamora, Zaragoza.

1.7.2 mth-order differencing estimator and optimal differencing weights

Yatchew (2003, p.58) shows that, if the observations $(y_1, x_1)\ldots(y_\aleph, x_\aleph)$ of the nonparametric model $y = \phi(x) + \epsilon$ have been reordered in increasing order relative to the x's, the nonparametric effect of $y = \phi(x) + \epsilon$ can be removed by differencing. Let us define the following differencing matrix:

$$D_{\aleph\times\aleph} = \begin{bmatrix} d_0,\ d_1,\ d_2, \ldots d_m,\ 0, \ldots\ldots\ldots\ldots\ldots\ldots, 0 \\ 0,\ d_0,\ d_1,\ d_2, \ldots,\ d_m,\ 0, \ldots\ldots\ldots\ldots\ldots, 0 \\ :\ \ : \\ 0, \ldots\ldots\ldots\ldots\ldots\ldots 0,\ d_0,\ d_1,\ d_2, \ldots d_m, 0 \\ 0, \ldots\ldots\ldots\ldots\ldots\ldots 0,\ 0,\ d_0,\ d_1,\ d_2, \ldots d_m \\ 0, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, 0 \\ :\ \ : \\ 0, \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, 0 \end{bmatrix}$$

The last m rows have been filled with zeros so that D is square. Applying the differencing matrix to $y = f(x) + \epsilon$, we get $Dy = Df(x) + D\epsilon$. A typical element of the vector Dy is of the form :

$$d_0 y_i + \ldots + d_m y_{i+m} = d_0 f(x_i) + \ldots + d_m f(x_{i+m}) + d_0 \epsilon_i + \ldots + d_m \epsilon_{i+m}$$

Then the role of the constraints $\sum_{k=0}^{m} d_k = 0$ and $\sum_{k=0}^{m} d_k^2 = 1$ on the differencing weights $d_0, d_1, d_2, \cdots, d_m$ in equation (1.3) becomes more explicit. The first condition ensures that, as the x's become close, the nonparametric effect is removed. The second condition ensures that the variance of the weighted sum of residuals remains equal to σ_ϵ^2. The mth-order differencing estimator of the residual variance is then defined by:

$$s_{diff}^2 = \frac{1}{\aleph}\sum_{i=1}^{\aleph-m}(d_0 y_i + d_1 y_{i+1} + d_2 y_{i+2} + \cdots + d_m y_{i+m})^2 = \frac{1}{\aleph} y' D' D y$$

It can be shown that in large samples $E(s_{diff}^2) \cong \sigma_\epsilon^2$ and that in order to minimise $Var(s_{diff}^2)$, we can use optimal values for the differencing weights. We used the values given by Yatchew (2003, p.61).

1.7.3 Baltagi et al. (1996)'s nonparametric poolability test

These authors assume that $y_{it} = g_{it}() = g_t()$ and set $g(x_{it}) = \frac{1}{T}\sum_s g_s(x_{it})$ and $u_{it} = y_{it} - g(x_{it})$. Under H_0^*, $g() = g_t()$ and we can write as $y_{it} = g(x_{it}) + u_{it}$. Then they show that $E(u_{it}) = 0$ under H_0^* and that $E(u_{it}) \neq 0$ under H_1^*. The test statistic requires nonparametric estimates of $E(y_{it} \mid x_{it})$ and of the x_{it}'s density $\hat{p}(x_{it})$. These estimations are given by:

$$\hat{y}_{it} = \frac{1}{N \cdot T \cdot a \cdot \hat{p}(x_{it})} \sum_{j=1}^{N}\sum_{s=1}^{T}(y_{js} \cdot K(\frac{x_{it} - x_0}{a}))$$

$$\hat{p}_{it} = \frac{1}{N \cdot T \cdot a} \sum_{j=1}^{N}\sum_{s=1}^{T} K(\frac{x_{it} - x_{js}}{a})$$

Then the residuals $\hat{u}_{it} = y_{it} - \hat{y}_{it}$ and $\hat{p}_{it}$ are used to compute the empirical statistic:

$$J = \frac{N \cdot h^{0.5} \cdot J_N}{\sqrt{2 \cdot \hat{\sigma}_0^2}} \xrightarrow{D} N(0,1)$$

where:

$$J_N = \frac{1}{N \cdot (N-1) \cdot T \cdot h} \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j \neq i}^{N}((\hat{u}_{it} \cdot \hat{p}_{it})(\hat{u}_{jt} \cdot \hat{p}_{jt}) \cdot K(\frac{x_{it} - x_{jt}}{h}))$$

$$\hat{\sigma}_0^2 = \frac{1}{N \cdot (N-1) \cdot T^2 \cdot h} \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j \neq i}^{N}((\hat{u}_{it} \cdot \hat{p}_{it})^2(\hat{u}_{jt} \cdot \hat{p}_{jt})^2 \cdot K^2(\frac{x_{it} - x_{jt}}{h}))$$

$$a = c \cdot x_{sd} \cdot N^{\frac{-1}{\alpha}}$$

$$h = c \cdot x_{sd} \cdot N^{\frac{-1}{\alpha'}}$$

Note that x_{sd} is the standard deviation of x_{it} and c, α and α' are constants. The parameters a and h must respect the following conditions in order to ensure that the kernel estimator is consistent Baltagi et al. (1996, p. 349, condition C3):

(i) $Nh^p \to \infty$, **(ii)** $a \to 0$, **(iii)** $Na^{2v}h^{p/2} \to 0$, **(iv)** $h/a^2 \to \infty$

where p is the number of variables and v is the order of the kernel estimator. In our case, given that $p = 1$ and $v = 2$, replacing a and h into the condition yields to:

(i) $\alpha' > 1$, **(ii)** $1/\alpha < 0$, **(iii)** $1 - 4/\alpha - 1/2\alpha' < 0$, **(iv)** $2\alpha' < \alpha$

We can check that these conditions hold for $alpha = 5$ and $alpha' = 2$ or for $alpha = 7$ and $alpha' = 3$. The c parameter in h and a was arbitrarily chosen by Baltagi et al. (1996) between 0.8 and 1.2. We followed the same method.

Chapter 2

Convergence-clubs in per capita CO_2 emissions.

Who's converging, who's diverging?

Editorial note: This paper is co-authored with Prof. Jean-Marie Grether[33].

Abstract: This paper investigates the convergence club hypothesis for CO_2 per capita emissions with a panel of 166 non-overlapping world areas covering the period 1960-2002. The distributional dynamics of the data is explored based on the evolution of the cross-sectional distributions over time. Robust measures of dispersion, asymmetry and peakedness recently introduced by Brys et al. (2006) are applied to different world samples and groupings of countries based on their level of economic development, geographic proximity, initial emissions' level and membership to a regional integration area. The global nonparametric shape equality test of Li (1996) is applied to assess distributional time differences. Multimodality is also checked with the Dip test by Hartigan and Hartigan (1985). Moreover, nonparametric density estimates provide a graphical support which allows to visualize the cross-sectional distributional dynamics. Our results point toward a structural break in the world distributions after the 70s oil-shocks and stable, flatter, unimodal and right-skewed distributions afterwards. We find evidence that this break does not hold within every grouping of countries and that the group-specific densities display heterogeneous behaviors. The existence of converging economies at the world level and within groups is formally checked with unit root tests and the pair-wise convergence specification introduced by Evans (1998). This approach suggests that convergence forces are at work, with stable or decreasing per capita emissions' gaps between countries for the world panel over the whole 1960-2002 period and in particular after the oil shocks of the 70s. The latter result also holds within several country groupings such as most or least intensive initial polluters, countries achieving similar per capita income levels, geographic neighbors, as well as North-South and North-North economically integrated partners.

JEL Classification: C10 · C14 · Q54

Key Words: carbon dioxide emissions, air pollution, convergence, distribution dynamics, unit roots.

---

[33] Institute for Economic Research, University of Neuchâtel, Switzerland.

2.1 Introduction

While the Kyoto protocol commitments to prevent global warming expire in 2012, it becomes clear that most countries are heading towards overshooting their greenhouse gases emissions targets. A range of new policy measures have been put forward to pursue the international climate change effort. Among the many options, Aldy (2006) reports that 25% of the over forty proposals surveyed in Bodansky et al. (2004) are based on per capita emissions allocation schemes. The Global Common Institute[34] promotes an approach, dubbed ‘Contraction & Convergence’ (C&C), which consists in setting a long term sustainable emissions budget and sharing this budget between countries so that per capita levels of pollution are equalized in the long run. The fundamental principal of allocating world emissions according to the same individual "right to pollute" is appealing from an equity point of view. However, Stegman (2005) points out that this approach may not correspond to the ‘efficient’ distribution, *i.e* the allocation that maximizes the value of resources (where marginal costs and benefits of pollution are equal). This allocation scheme also ignores specific structural characteristics of countries, such as colder climate, natural resource endowments or comparative advantages in production. Nevertheless, its operational simplicity and its ability to set a "unifying principle that facilitates an international greenhouse warming agreement" (see Rose and Stevens 1996) between governments has attracted institutional support.

In the background of this policy debate, an important empirical fact calls for a better understanding of the dynamics of per capita carbon emissions: while total emissions keep increasing in most countries, per capita levels appear to have stabilized at the world level[35]. If national series show some evidence of converging trends, per capita targets may represent a more acceptable basis for political compromises than absolute levels. Our work contributes to the existing empirical literature on the topic by exploring two important questions: has the world cross-sectional distribution of per capita carbon emissions stabilized over the last 40 years? Are we able to identify groups of economies with converging per capita emissions based on simple criteria such as geographic proximity, similar income or initial emissions’ level, or economically integrated areas?

The economics of the C&C proposal has been explored by Boehringer and Welsch

---

[34] See Meyer (2001).

[35] See McKitrick and Strazicich (2005).

(1999) in a multi-region computable general equilibrium model of the world economy. These authors compare a regime of tradable *vs.* non-tradable emission rights for implementing C&C and find that the former allows a substantial reduction of long term costs of abatement. Another related framework is the neoclassical growth model by Brock and Taylor (2004b), called the 'Green Solow model', which predicts (conditional) convergence of per capita pollution levels between countries which share similar (different) structural economic characteristics. The latter model also shows that if technological progress in abatement activities exceeds that in goods production, the pollution trend is sustainable in the sense that pollution depicts an U-inverted relationship with GDP, with possibly country-specific turning points. The empirical evidence gathered so far regarding pollution convergence is contrasted. On the one hand, methods which analyze how the cross-sectional distribution of per capita CO_2 emissions (PCE henceforth) evolve over time conclude to divergence (or 'persistence') worldwide but convergence between industrial countries or large initial polluters. On the other hand, the stochastic convergence empirics presents contrasted results for OECD members[36] but convergence within groups of *a priori* more heterogeneous countries[37].

This paper investigates the convergence issue for per capita carbon dioxide emissions with a distributional approach as well as stochastic convergence methods in a novel fashion. Firstly, if we want to assess to what extent convergence occurs worldwide or within country groupings (called convergence-clubs), we should aim at working with the largest possible sample of world countries. Our first contribution is to widely expand the number of countries analyzed until now within the database used by many authors. By simply accounting for changes in borders, we identify a balanced panel of 166 non-overlapping world areas, spanning over the years 1960-2002, which represents 88% of the countries defined by the World Bank, and almost doubles the number of considered areas with respect to the previous literature. Secondly, we provide a simple descriptive analysis on how cross-sectional distributions evolve over time, without putting constraints on the dynamics of the process. The distributional changes are evaluated with robust scale and shape statistics as well as graphical devices of kernel density estimates. The results point clearly toward a structural break in the distributional world PCE dynamics coinciding with the 70s oils shocks and a stabilized worldwide cross-sectional distribution afterwards. This break is not found in some subgroups of countries, based on their level of economic

---

[36] See Strazicich and List (2003), Aldy (2006), Romero-Ávila (2007) and Barrasi et al. (2008).

[37] See Westerlund and Basher (2008).

development, geographical location, initial level of PCE and membership to a regional integration agreement. Thirdly, we abandon the standard use of (the log of) relative PCE series in the stochastic convergence framework and focus on the notion of pair-wise convergence introduced by Evans (1998). Unit root tests are computed for different groupings of countries and the country specific information is combined with the method proposed by Maddala and Wu (1999) to produce panel results. This method clearly points toward conditional convergence at the world level as well as for some specific groups of countries.

We begin this research with a survey of the indicators used to explore convergence in per capita pollution levels between countries as well as a review of the empirical findings obtained on the topic for carbon dioxide emissions. In section 3, we present our CO_2 data and carry out the distributional analysis. Section 4 presents the econometric approach and results and section 5 concludes.

2.2 Empirical literature

The empirical literature on convergence in cross-country pollution levels has grown exponentially in the last 5 years, in particular concerning per capita carbon dioxide emissions. Most studies borrow their tools from the income convergence literature. In section 2.2.1, we first present the main convergence measures that have been employed to analyze carbon convergence at the international level. Then, a review of the empirical literature is proposed in section 2.2.2, starting with the time-series properties of the CO_2 PCE series and ending with the convergence problematic.

2.2.1 Convergence measures

The evolution of the gaps in per capita CO_2 emissions between countries/regions is explored mainly with four measures of convergence. The first one, called $\underline{\beta\text{-convergence}}$, captures the idea that countries with lower initial levels of pollution per capita should experience higher pollution growth, and therefore eventually 'catch-up' the most polluting countries. In the presence of absolute β-convergence, regressing the subsequent period pollution average growth rate for each country on its initial pollution level should result in a negative and significant relationship. In practice, the 'catch up' phenomenon is expected to occur basically between similar countries whose economic activity took off at different points in time. Thus, economies with structural differences will tend to grow toward their own pollution level, and convergence becomes conditional upon countries characteristics. Conditional β-convergence is naturally investigated by adding a set of exogenous explanatory factors to the absolute β-convergence regression. The second empirical measure of convergence, dubbed σ-convergence, simply requires that dispersion across a group of countries decreases over time. Barro and Sala-i Martin (2004, Ch.11.1) show that β-convergence is a necessary but not sufficient condition for σ-convergence to occur. The strength of the β-convergence measure lies in its explicit link with the theoretical growth models[38] and its faculty to easily accommodate conditioning factors of convergence in the empirical analysis through standard regression methods. Its main drawback is that the method focuses on the conditional mean growth rate and that it does not guarantee any precise inference in terms of gap' dynamics between countries[39]. Note that σ-convergence would fail to capture polarisation phenomena in case of tendencies toward multimodality. In order to address the weaknesses of the latter two indicators, several authors have suggested different approaches based on univariate time-series analysis, mainly unit root tests, and usually referred to as stochastic convergence. Inspired by the work of Carlino and Mills (1993) on income convergence, the first method aims at testing to what extent initial departures from some hypothesized long run equilibrium in per capita pollution tend to vanish over time. Two main specifications have been used to test stationarity under the latter perspective:

$$z_{i,t} = a_{i,0} + \phi_i \, z_{i,t-1} + e_{i,t} \tag{2.1}$$

$$z_{i,t} = a_{i,0} + \phi_i \, z_{i,t-1} + a_{i,1} t_i + e_{i,t} \tag{2.2}$$

where $z_{i,t}$ is defined as the natural logarithm of $p_{i,t}/\bar{p}_{.,t}$, $p_{i,t}$ is per capita emissions of country i at time t, $\bar{p}_{.,t}$ represents a world/group cross-section arithmetic average of $p_{i,t}$ at time t and $e_{i,t}$ is a iid disturbance term. The constant $a_{i,0}$ in both equations is generally supposed to embody two unidentified terms $a_{i,0} = f(d_i^*, c_{i,0})$, where d_i^* is a time-invariant country-specific equilibrium differential, and c_{i_0} stands for an initial deviation from the

---

[38] It is important to note that the Solow-Swan model *per se*, as reported in Barro and Sala-i Martin (2004, Ch.1.2.10), derives no β-convergence formulation for pollution. Refer to Brock and Taylor (2004b) or Alvarez et al. (2005) for such a link.

[39] The shortcomings of β-convergence have been widely discussed in the income literature. See, among others, Friedman (1992), Quah (1993b), Evans (1996), Durlauf et al. 1996, Bernard and Durlauf (1996), Bliss (1998, 2000), Cannon and Duck (2000) and Durlauf et al. (2005) for a full econometric review on income convergence.

long-run equilibrium. In this framework, rejecting the unit root hypothesis ($\phi_i = 1$) indicates that a random shock to the series $z_{i,t}$ reverts toward a constant in equation (2.1) or toward the constant and a trend in equation (2.2). Absolute convergence holds when $a_{i,0}$ and $a_{i,1}$ are equal to zero. Conditional convergence requires either $a_{i,0} \neq 0$ and $a_{i,1} = 0$ or $\text{sign}(c_{i,0}) \neq \text{sign}(a_{i,1})$(40). Note that both unit root equations can be estimated with a variety of techniques, designed for individual countries or panel data, and accounting or not for structural breaks in the data generation process. Other variants of stochastic convergence, called pairwise convergence, have also been proposed by Bernard and Durlauf (1995), Evans (1998) and Pesaran (2007). These authors investigate the dynamic of income differences between pairs of countries and they propose different methodologies to extend the results of the pair-wise approach to all the members of a set of economies. Finally, the distributional convergence analysis initiated by Quah (1993a, 1996 and 1997) puts the emphasis on the dynamics characterizing the whole cross-sectional distributions. It consists in conditioning future cross-section distributions of per capita emissions of CO_2 on their past counterparts by assuming that current levels map into the future ones according to time-invariant transitional probabilities[41]. This is similar in the spirit to a first order autoregression, with distributions as argument instead of scalar or vectors. As noted by Quah (1997, p.36-37), if the mapping of the emission levels from the initial period to the next results in the same level, the cross-sectional distribution converges toward a point limit and collapses in the long run. Distributional persistence (divergence) holds when past pollution levels remain stable (increase) over time. The graphical representation of the process also allows to detect polarization phenomena and limiting (long run or ergodic) distributions can be computed.

2.2.2 Convergence in carbon emissions

The rest of this section is devoted to the empirical findings on CO_2 convergence. Many studies have used (early) series from Marland et al. (2006) on CO_2 to explore convergence in per capita emissions[42], based on (log) series expressed relative to or in deviation from

---

[40] See Carlino and Mills (1993, p.337).

[41] As noted by Quah (1993a, p.429), there is "no reason why the law of motion of the cross-sectional distribution need be first order, or why the relation need be time-invariant".

[42] These data are often referred to as data from Oak Ridge National Laboratory, Carbon Dioxide Information Analysis Center or US Department of Energy. All these sources are the same. Note that the CO_2 data from the World Bank's World Development Indicators also rely on this source.

some group mean (*e.g.* world or OECD cross-sectional average (log) levels). We first present three major papers which focus solely on the time-series properties of the (log) level of the data. Then we turn to the carbon convergence empirics.

Time-series properties of per capita CO_2 emissions. Heil and Selden (1999) are among the first to test for unit roots in PCE carbon dioxide series with level as well as logarithmic data[43]. The null of a unit root is checked for an unbalanced world panel of 135 countries over the pre and post-oil shock periods 1950-1973 and 1974-1992 with a preliminary version of Im et al. (2003) (IPS henceforth) and a specification of type (2.2) (*i.e.* with a constant and a trend). Note that the IPS methodology allows to draw inference for the whole panel on the basis of country-by-country unit root tests[44]. The null unit root hypothesis is never accepted against the alternative of trend stationarity, except for the level series over the pre-oil shock period 1950-1973. Therefore, level and logarithmic data yield different results as a structural break in 1973 is found only in levels. These authors do not verify if the trend coefficients for each country of the panel remain significant in trend stationary series. Lanne and Liski (2003) carry out a similar investigation for longer series in logarithms, spanning the years 1870-1998, for 16 OECD countries and with an endogenous break unit root test[45]. Among the ten series not supporting the presence of a unit root, four of them appear to have a significant break in the 70s, five display a break located early in the century (between 1900-1908) and three countries are trend stationary without significant break. The only countries to depict a downward sloping trend after the downturn are UK and Sweden. In the same vein, with shorter time series but a widely extended number of countries, McKitrick and Strazicich (2005) use the endogenous two-break LM test of Lee and Strazicich (2003) with world series of CO_2 PCE over the years 1950-2000. They find no evidence of a unit root in the global world PCE and they identify two structural breaks in 1968 and 1981. Moreover, after 1981, the trend is small in magnitude, negative and not significant. The test is also applied to 121 individual series as well

---

[43] In a previous study, Holtz-Eakin and Selden (1995) report stationarity for per capita CO_2 emissions when estimating the income-pollution relationship for a large set of world countries with a fixed-effects panel model and a quadratic function. However, the detailed unit root results are not published.

[44] See section 2.4.1 for further details.

[45] The additive outlier model is used here. Perron and Volgesang (1992) and Volgesang and Perron (1998) have proposed a class of test statistics which allows for two different forms of a structural break in unit root processes: the additive outlier model, which is more relevant for series exhibiting a sudden change, and the innovational outlier model, which captures gradual changes through time.

as to the 121-countries global world: in the latter case, a single downturn is identified in 1978 with a negative but not significant trend afterwards. Regarding the country specific results, 48% of the 95 stationary series have significant positive trends, 19% have significant negative trends and 33% are trendless[46]. Note that 60% of the breaks occur during the 1973-1982 period. Overall, these researches highlight three main characteristics of the national time-series on CO_2 per capita emissions : (i) most of them are not stochastically trended, (ii) they display significant structural breaks, mainly located around the 70s, (iii) world PCE level as well as a large portion of the national series depict null or decreasing deterministic trends after the last structural break identified.

Convergence in per capita CO_2 emissions. We now turn to the CO_2 convergence studies, mainly carried out with the same database. We begin with the β and σ-convergence results, we proceed with the stochastic approach and end with the distributional analysis. To our knowledge, Strazicich and List (2003) (SL2003 henceforth) are the first to explore convergence for CO_2 emissions[47]. They study both absolute and conditional β-convergence for a sample of twenty one OECD countries over the period 1960-1997. The conditional β-convergence analysis is carried out based on a set of regressors which capture country-specific characteristics: GDP, GDP squared, gasoline price, population density and a temperature indicator[48]. The regression results indicate that absolute β-convergence holds and that the convergence coefficient remains significant and negative for all investigated combinations of control variables. Among the conditioning factors, only the gas price and temperature appear to be significant[49]. Brock and Taylor (2004b) test absolute as well as conditional β-convergence based on their Green Solow model, by progressively augmenting the simple cross-sectional β-regression with time-averaged country-specific (estimations of) technological progress in abatement, saving rate, abatement level and effective depreciation rate of capital. The model is tested for OECD countries over the period 1960-1998 and the fits indicate that most of the explanatory power comes from the initial level of pollution, which display a significant negative effect. Alvarez et al. (2005) explore β and σ-convergence for sets of European

---

[46] Among the nonstationary series, we find Argentina, Canada, Finland, Germany, the Netherlands, Saudi Arabia, Spain, United States and Venezuela.

[47] A previous work by List (1999) examine convergence for series on NO_x and SO_2 emissions in US EPA regions.

[48] Note that this mix of factors is not formally derived from any particular theoretical model

[49] Both variables have negative impact on per capita CO_2 emissions' growth rates.

countries and for several air pollutants over the period 1990-2002[50]. The σ-convergence analysis simply compares the standard deviation of the log of CO_2 PCE between 1990 and 2000. Beta-convergence is tested by computing simple and augmented cross-sectional as well as panel regressions. Conditional β-convergence relies on a theoretical reduced form specification which includes initial pollution levels, initial GDP levels and GDP growth as explanatory factors. Their findings for carbon dioxide support the existence of σ, absolute and conditional β-convergence for most of the country groupings. Per capita GDP growth has a significant positive effect on per capita emissions growth but only through an interaction dummy identifying middle-income countries. Nguyen Van (2005) analyses absolute cross-sectional β-convergence for 100 world countries over the period 1966-1996 and finds a significant negative relationship. Finally, Aldy (2006) provides estimations of σ-convergence for OECD series as well as for a 88-country world sample over the period 1960-2000[51]. Based on two distinct measures of dispersion, he shows that the standard deviation of the cross-sectional PCE distributions decreases steadily for the OECD panel but increases slightly for the world. Then comparing the interquartile ranges in 1960 with those for the latter decades, he confirms the latter result for the OECD countries but without formally rejecting the null of dispersion's equality between the base year 1960 and the latter decades[52]. Significant differences are found for the world sample in 1990 and 2000. Overall, these studies find β-convergence for the OECD samples as well as at the world level. However, σ-divergence is prevalent for the world while the contrary holds in most of the OECD country groups.

Regarding stochastic convergence, Strazicich and List (2003) make use of the trend specification (2.2) in the panel IPS framework and they conclude that the pooled statistic derived from country-specific unit root tests leads to stochastic convergence for the OECD panel as a whole. However, these authors only base their conclusion of the rejection of the unit root null against trend stationarity, without formally testing the convergence conditions in equation (2.2). Aldy (2006) also employs specification (2.2) with a more powerful unit root test, the generalized least squares Dickey-Fuller test by Elliott et al.

---

[50] These authors consider data from the European Environment Agency on NO_2, SO_2 and CO_2 for three country groupings : the EU14 (EU15 without Luxembourg), the EU14 without the 'middle income economies' (Greece, Ireland Portugal, Spain) and the EU14 plus new Eastern entrants (Czech Republik, Hungary, Poland, Slovakia and Slovenia).

[51] See Aldy (2007) for a study on CO_2 convergence between US states.

[52] Aldy computes the standard deviation of the log of the CO_2 PCE's cross-sectional distributions and the interquartile range for relative series (*i.e.* cross-sectional level series divided by the corresponding cross-sectional arithmetic mean).

1996. He finds that only 13/88 and 3/23 countries reject the null of a unit root in his world and OECD sample respectively at the 10% level but he does not provide panel results. Nguyen Van (2005) analyses stochastic convergence at the world level with specification (2.1), the dynamic panel approach of Arrellano and Bond (1991) and the log of relative per capita CO_2 emissions taken every 5-year as well as 10-year periods for each country. Stochastic convergence is accepted with the 5-year data only. More recently, Barrasi et al. (2008) focus on the OECD countries used in SL2003 for the period 1950-2002 and criticize the latter research for not checking the sign constraints in specification (2.2) before making the convergence statement. Accounting for the sign restrictions and making use of more recent unit root techniques (with trend stationarity under the null), these authors report pollution divergence across the OECD members, even when the Dickey-Fuller and the IPS approaches from SL2003 are employed with recent methods improving the size and power of the latter test[53]. Romero-Ávila (2007) also study a similar OECD sample of 23 countries over the years 1960-2002 with both equations (2.1) and (2.2), allowing for an unknown number of endogenous breaks in both specifications and correcting for cross-correlation in the panel statistics[54]. Under the null of stationarity, he shows that convergence is widely rejected when structural breaks are ignored but overwhelmingly accepted when they are allowed. In the latter case, only four countries do not reject trend stationarity. A maximum of four breaks is found for four countries but most of them display two breaks. Among the 53 breaks identified, 17 are detected before the first oil shock and 18 happen within the 1973-1981 period. Finally, when the trend component is excluded, the same broad results are achieved, confirming stochastic convergence for OECD countries when structural breaks and cross-sectional dependencies are taken into account. While all these studies focus on the ability of (relative) individual series to revert toward a hypothesized 'steady state' (relative= equilibrium, they do not look directly into the pollution gap between countries. Using the pair-wise notion of convergence *à la* Evans (1998)[55] in conjunction with panel tests for specification (2.2) which account

---

[53] Given that the term $c_{i,0}$ is not identifiable in equation (2.2), these authors first check $\text{sign}(a_{i,0}) \neq \text{sign}(a_{i,1})$ in the spirit of Tomljanovic and Vogelsang (2002) but without allowing for structural breaks. Then unit root tests are carried out for the countries matching the latter sign constraint.

[54] The method of Carron-i Silvestre et al. (2005) is used to detect the breaks. Note also that, in reference to Li and Papell (1999), this author calls 'deterministic' convergence the unit root tests without trend while the term 'stochastic' convergence is reserved for the trended specification (2.2).

[55] This notion of stochastic convergence simply replaces the z_{it} variable in specifications (2.1) by $z_{i,t} = log(p_{i,t}) - 1/n \sum_{i=1}^{n} log(p_{i,t})$. Evans (1998) shows that all pairs of countries included in a

for cross-sectional dependencies in the data, Westerlund and Basher (2008) conclude to the existence of group-wise convergence within a world panel of twenty-eight developed and developing countries covering the period 1870-2002. They stress that their results should be robust to structural breaks as ignoring them favors accepting their null hypothesis of a unit root. Finally, two main methodological conclusions can be drawn from the empirical literature on stochastic convergence: accounting for structural breaks as well as cross-sectional dependencies is important for making consistent inference in the panel framework. So far, OECD countries are found to converge stochastically, except in one research. The evidence at the world level is mixed and needs more research effort.

Finally, four papers study the intra-distributional dynamics of per capita CO_2 emissions for large panels of world countries by assuming cross-sectional distributions evolving according to a stable time-invariant (Markovian) first order process (or stable transition probabilities). In a discrete framework, Aldy (2006) warns against the sensitivity of the long run distributions to the reference period (initial conditions) but he finds no 'meaningful' distributional convergence at the world level for relative per capita emission, whatever the chosen reference period. Nguyen Van (2005) argues that the time-invariant hypothesis of the transition process is quite robust in a continuous setting and that distributional convergence holds between countries with relative per capita emission levels above three, *i.e* initial most intensive polluters as well as for a subset of 26 industrial countries. Stegman (2005) confirms Nguyen's findings to some extent[56] for a world panel of 97 countries spanning the years 1950-1999. However, she stresses that centering the data instead of dividing them by the cross-sectional mean would result in no evidence of distributional convergence over the entire initial distribution's support. Finally, Ezcurra (2007) employs a panel of 87 countries over the period 1960-1999 and proposes distributional polarization measures based on exogenous partitioning of the sample to evaluate the intra-distributional mobility over time[57]. He shows that cross-country polarization decreases steadily when two groups are considered and that it remains steady during the 60s but decreases afterwards with a three groups partition. He also estimates ergodic densities and conditional den-

---

group converge toward another if each individual series, centered on the cross-sectional arithmetic mean of the group, is stationary; see section 2.4.1 for further details. Note that, by contrast to the original formulation by Evans, Westerlund and Basher (2008) include a trend in the unit root equation.

[56] Stegman (2005, p.17-18) outlines that the few observations available at the upper relative PCE levels may bias severely the stochastic kernel estimates.

[57] The method, based on the work of Esteban et al. (1999), relies on a grouping criteria that minimizes inequality due to within group dispersion.

sities on income, trade openness and climatic conditions (average annual temperature). The long run distribution appears to be unimodal and does not collapse. This suggests a high degree of persistence in world differences in relative per capita CO_2 emissions or low intra-distributional mobility. Regarding the conditional distributions, the density mass seems to be more concentrated around the average when the data are conditioned upon the per capita income and climatic conditions' variables, while trade openness does not seem to affect the original distributions. However, no formal tests are provided about the latter conclusions. Overall, the distributional analysis appear to have several methodological shortcomings that still need to be addressed. At the moment, its results indicate convergence between industrial economies and the most intensive initial polluters, but 'persistence' for the rest of the world.

Given the variety of country groupings used in the empirical convergence literature on per capita carbon emissions, it is fundamental to study to what extent the convergence results are sensitive to groupings criteria or to differences in sample size. Toward this end, our paper explores convergence across three world samples as well as groupings of countries based on simple and systematic classification criteria. We focus on stochastic as well as distributional convergence. The latter analysis employs unscaled PCE levels and remains fundamentally descriptive to avoid imposing excessive constraints on the dynamic of the distributional process and to highlight the most significant scale and shape characteristics of the cross-sectional distributions. Moreover, given the significant structural breaks found in most individual (relative) time series by most researchers, we explicitly test the equality of the cross-sectional distributions for successive decades. To our knowledge, this has not been done until now. We explore stochastic convergence relying on the notion of pairwise convergence of Evans (1998) as it looks directly into the stationarity of the pollution gaps between all pairs of countries included in a panel dataset. In that framework, we test convergence for the whole 1960-2002 period as well as for several sub-periods, mainly the pre and post-oils shocks. We carry out the panel tests accounting for cross-sectional dependencies in a simple way.

2.3 Descriptive analysis

After presenting the data in section 2.3.1, we report briefly in section 2.3.2 how total and per capita CO_2 emissions evolve over the period 1960-2002. Section 2.3.3 is dedicated to the distributional analysis. It first introduces the distributional tests and the robust scale and shape measures and then applies these approaches to cross-sectional PCE distributions for the world as well as for group-specific panels formed according to income, geographic, initial PCE levels and RIA membership criteria.

2.3.1 Data

As in other studies in the field, data on CO_2 emissions come from the Carbon Dioxide Information Analysis Centre[58] and reflect anthropogenic emissions from fossil fuel consumption, cement manufacturing and gas flaring, ignoring fuels supplied to ships and aircrafts. These series are available at several aggregation levels, such as individual countries, geographic regions or the world as a whole. National series capture the time pattern of more than 250 non-overlapping geographic areas for periods ranging from ten years to over two centuries[59]. When all these data are aggregated for all available years, we get total carbon emissions for what we call the CDIAC World. Once we account for changes in borders, this large database allows to build up a balanced panel of 166 non-overlapping national series[60] covering the period 1960-2002. The latter sample, called CDIAC166 henceforth, represents 183 out of the 208 countries (88%) reported by the World Bank in its CDrom World Development Indicator 2004 (WDI henceforth). Note that seven of the 166 areas included in the CDIAC166 sample are not sovereign territories[61] according to WDI. Compared to the World sample from the recent study by Aldy (2006) (ALDY2006 henceforth), this represents a significant increase in sample size (+ 78 countries) which should allow to capture consistently the convergence club phenomenon worldwide. Finally, note that 36 series included in the CDIAC166 sample depict pretty erratic time patterns, with annual PCE growth rates over 100% in absolute value for at least one of the years of the 1960-2002 period. We refer to the latter subset as 'outliers' as their impact on any measure based on statistical moments may be significant. Removing the 36 outlying series from CDIAC166 results in a World sample of 130 countries, CDIAC130 henceforth[62].

Regarding the different groupings of countries, the criteria used to build up income

---

[58] See Marland et al. (2006).

[59] See Westerlund and Basher (2008) for a treatment of the longest series available.

[60] Among the reconstructed series are countries like Germany, USSR, Yugoslavia, Czechoslovakia, Malaysia, Vietnam or Yemen. See Tables 2.14 and 2.15 in the Appendix for full details.

[61] These areas are French Guiana, Gibraltar, Guadeloupe, Martinique, Réunion, St. Pierre & Miquelon and Taiwan.

[62] See Table 2.7 in the Appendix for a rough view of the time patterns of the CDIAC130 and outliers' PCE series.

and geographic groups are borrowed from WDI. The World Bank defines the following four income categories on the basis of per capita Gross National Income levels in 2002: low or LI (<735$), lower-middle or LMI (736-2935), upper-middle or UMI (2936-9075), high income or HI (>9075$). The latter income class is further divided into OECD and non-OECD members. Countries are also classified into seven World geographic regions: East Asia & Pacific (EPA), Europe & Central Asia (ECA), Latin America & Caribbean (LAC), Middle East & North Africa (MENA), North America (NA), South Asia (SA) and Sub-Saharan Africa (SSA). Most of the CDIAC166 areas are equivalent to the world countries listed in WDI. However, some CDIAC166 areas encompass several WDI countries which pertain to the same geographic area but which may possess different income per capita levels. Therefore some sample adjustments are necessary to match the WDI groupings criteria with the CDIAC166 areas[63].

As mentioned in the previous section, countries are also grouped together according to their initial level of PCE in 1960. Here, we follow the methodology adopted by Ben-David (2000) in the income convergence literature. We first sort the CDIAC166 countries by increasing PCE level in 1960 and then split the ordered sample into subgroups of similar size. The first group contains the most intensive polluters and the last one the least.

Finally, we also put together the members of thirteen main regional integration agreements (RIA henceforth). Eight RIA are South-South initiatives (ACN, AFTA, CARICOM, CEMAC, COMESA, ECOWAS, LAIA, MERCOSUR), two of them are North-South (APEC12, NAFTA) agreements and three involve exclusively Northern or the most developed countries (EU6, EU15 and OECD)[64]. As membership of the RIA varies over time, specific RIA subsamples (founding members, longest memberships) are constructed.

---

[63] See Appendix, in particular tables 2.16 to 2.20.

[64] These acronyms stand for: ACN = Andean Community of Nations, AFTA = Asean Free Trade Association, APEC = Asia-Pacific Economic Cooperation, CARICOM = Caribbean Community and Common Market, CEMAC = Communauté Économique et Monétaire d'Afrique Centrale, COMESA = Common Market of Eastern and Southern Africa, ECOWAS = Economic Community of West African States, EFTA = European Free Trade Association, EU = European Union, LAIA = Latin American Integration Association, MERCOSUR = Mercado Común del Sur, NAFTA = North American Free Trade Association, OECD = Organization for Economic Co-operation and Development. See Tables 2.21 and 2.22 for full details on the included countries.

Table 2.1: World carbon dioxide emissions. Descriptive statistics for the period 1960-2002

	Levels (in thousand mt.)			Per capita (in mt.)							
Country / group	1960 (%)	2002 (%)	growth 60-02	1960 (rk.)	2002 (rk.)	growth 60-02	median	mean	std. dev	max	min
World											
CDIAC World	2577.0 (103.2%)	6973.0 (105.3%)	2.3%	0.85 (36)	1.12 (70)	0.6%	1.12	1.09	0.10	1.23	0.84
CDIAC166	2497.9 (100.0%)	6621.2 (100.0%)	2.3%	0.82 (39)	1.07 (72)	0.6%	1.07	1.04	0.09	1.17	0.81
CDIAC130	2462.3 (98.6%)	6226.5 (94.0%)	2.2%	0.83 (38)	1.06 (72)	0.6%	1.07	1.04	0.09	1.17	0.82
Selected countries											
Australia	24.1 (1.0%)	97.1 (1.5%)	3.3%	2.32 (11)	4.97 (11)	1.8%	3.79	3.71	0.78	5.17	2.32
Canada	52.6 (2.1%)	140.9 (2.1%)	2.3%	2.88 (8)	4.41 (14)	1.0%	4.11	4.03	0.48	4.72	2.84
China	212.9 (8.1%)	957.2 (14.5%)	3.6%	0.33 (64)	0.75 (86)	1.9%	0.42	0.44	0.19	0.75	0.16
India	32.9 (1.3%)	332.7 (5%)	5.5%	0.07 (160)	0.32 (109)	3.5%	0.16	0.17	0.08	0.32	0.07
Germany	222.2 (8.9%)	219.3 (3.3%)	-0.0%	3.07 (5)	2.66 (25)	-0.3%	3.46	3.31	0.36	3.84	2.64
Luxembourg	3.1 (0.1%)	2.6 (0.03%)	-9.8%	10.0 (2)	5.71 (9)	-1.3%	7.90	7.92	1.98	10.98	4.93
Qatar	>0.1 (0.0%)	9.9 (0.1%)	13.2%	1.06 (28)	12.52 (2)	5.9%	14.23	14.49	6.43	28.79	0.90
U. Arab. Emir.	>0.1 (0.0%)	25.6 (0.4%)	23.4%	0.03 (139)	10.49 (3)	14.7%	8.89	9.21	6.50	25.69	0.03
UK	159.3 (6.4%)	148.1 (2.2%)	-0.2%	3.04 (6)	2.47 (31)	-0.5%	2.82	2.83	0.23	3.18	2.47
USA	798.6 (32.0%)	1592.4 (24.0%)	1.6%	4.42 (4)	5.53 (10)	0.5%	5.36	5.29	0.39	5.96	4.36
Income											
High	1554.4 (62.2%)	3290.0 (49.7%)	1.8%	2.33 (11)	3.50 (19)	1.0%	3.23	3.14	0.32	3.50	2.33
Middle	878.2 (35.2%)	2703.1 (40.8%)	2.6%	0.62 (43)	0.96 (73)	1.0%	0.96	0.89	0.16	1.11	0.59
Low	61.9 (2.5%)	564.0 (8.5%)	5.3%	0.06 (118)	0.23 (116)	3.0%	0.14	0.15	0.05	0.23	0.06
Geogr. areas											
EAP & SA	366.2 (14.7%)	2189.5 (33.1%)	4.2%	0.23 (76)	0.63 (92)	2.4%	0.40	0.41	0.15	0.63	0.18
ECA	1131.0 (45.3%)	1732.1 (26.2%)	1.0%	1.67 (20)	1.99 (47)	0.4%	2.33	2.30	0.32	2.72	1.66
LAC	83.3 (3.3%)	354.2 (5.3%)	3.4%	0.38 (61)	0.66 (92)	1.3%	0.63	0.59	0.11	0.71	0.38
MENA	27.9 (1.1%)	412.0 (6.2%)	6.4%	0.27 (70)	1.27 (65)	3.7%	0.90	0.84	0.29	1.27	0.27
NA	851.2 (34.1%)	1733.6 (26.2%)	1.7%	4.27 (5)	5.42 (11)	0.6%	5.21	5.17	0.39	5.81	4.22
SSA	34.9 (1.4%)	135.7 (2.1%)	3.2%	0.16 (88)	0.21 (120)	0.6%	0.23	0.22	0.03	0.27	0.16
Init. polluters											
High	2171.3 (86.9%)	4299.4 (64.9%)	1.6%	2.0 (13)	2.70 (25)	0.7%	2.79	2.73	0.28	3.11	2.01
Middle	274.5 (11.0%)	1714.5 (25.9%)	4.4%	0.27 (70)	0.78 (85)	2.5%	0.47	0.47	0.19	0.78	0.19
Low	52.1 (2.1%)	607.3 (9.2%)	5.9%	0.05 (123)	0.25 (114)	3.6%	0.13	0.14	0.06	0.25	0.05
Other groups											
EU15	588.6 (23.6%)	853.3 (12.9%)	0.8%	1.86 (16)	2.25 (39)	0.4%	2.29	2.31	0.18	2.69	1.86
OECD	1543.1 (61.8%)	3171.5 (47.9%)	1.7%	2.35 (10)	3.48 (19)	0.9%	3.24	3.15	0.31	3.50	2.35
Ex-USSR	395.1 (15.8%)	587.3 (8.9%)	0.9%	1.74 (20)	2.06 (45)	0.4%	2.65	2.66	0.58	3.64	1.74

Source: Author's own calculations with CO_2 data from Marland et al. (2006) and population series from U.S. Census Bureau (2006). The sample CDIAC130 corresponds to the sample CDIAC166 once erratic countries are removed. The growth rates are geom. averages between the initial and final period. (%) corresponds to emissions' share relative to total emissions in sample CDIAC166. (rk) is the rank with respect to the national series included in CDIAC166. The income and geographical grouping criteria are based on World Development Indicators 2004. HI, MI and LI stand for High, Middle and Low Income. For the geographic areas: EAP = East Asia & Pacific, ECA = Europe & Central Asia (ECA), LAC = Latin America & Caribbean, MENA = Middle East & North Africa, NA = North America, SA = South Asia and SSA= Sub-Saharan Africa. Initial polluters categories are based on the 1/3 and 2/3 percentiles of the CO2 PCE's cross-sectional distribution in 1960.

2.3.2 Historical trends

In order to get an overview of the main trends for the CO_2 series in levels as well as in per capita terms, Table 2.1 provides some descriptive statistics for the whole 1960-2002 period. Columns 2 to 4 concern total emissions and report for each country/area initial and final total emissions, with the corresponding emissions' share for each country/aggregate group relative to the reference sample CDIAC166 in brackets. The emission's average growth rate over the 43 investigated years is provided in column 4. Summary statistics for carbon emissions in per capita terms are presented in columns 5 to 12. They include initial and final per capita levels with their related rank in the CDIAC166 individual sample, as well as location and variance statistics related to the time dimension of the series over the 43 years. Figure 2.1 displays graphically time patterns of total versus per capita CO_2 emissions for the world (CDIAC World, CDIAC166 and CDIAC 130) samples as well as for the main country groupings (essentially income and geographical groups).

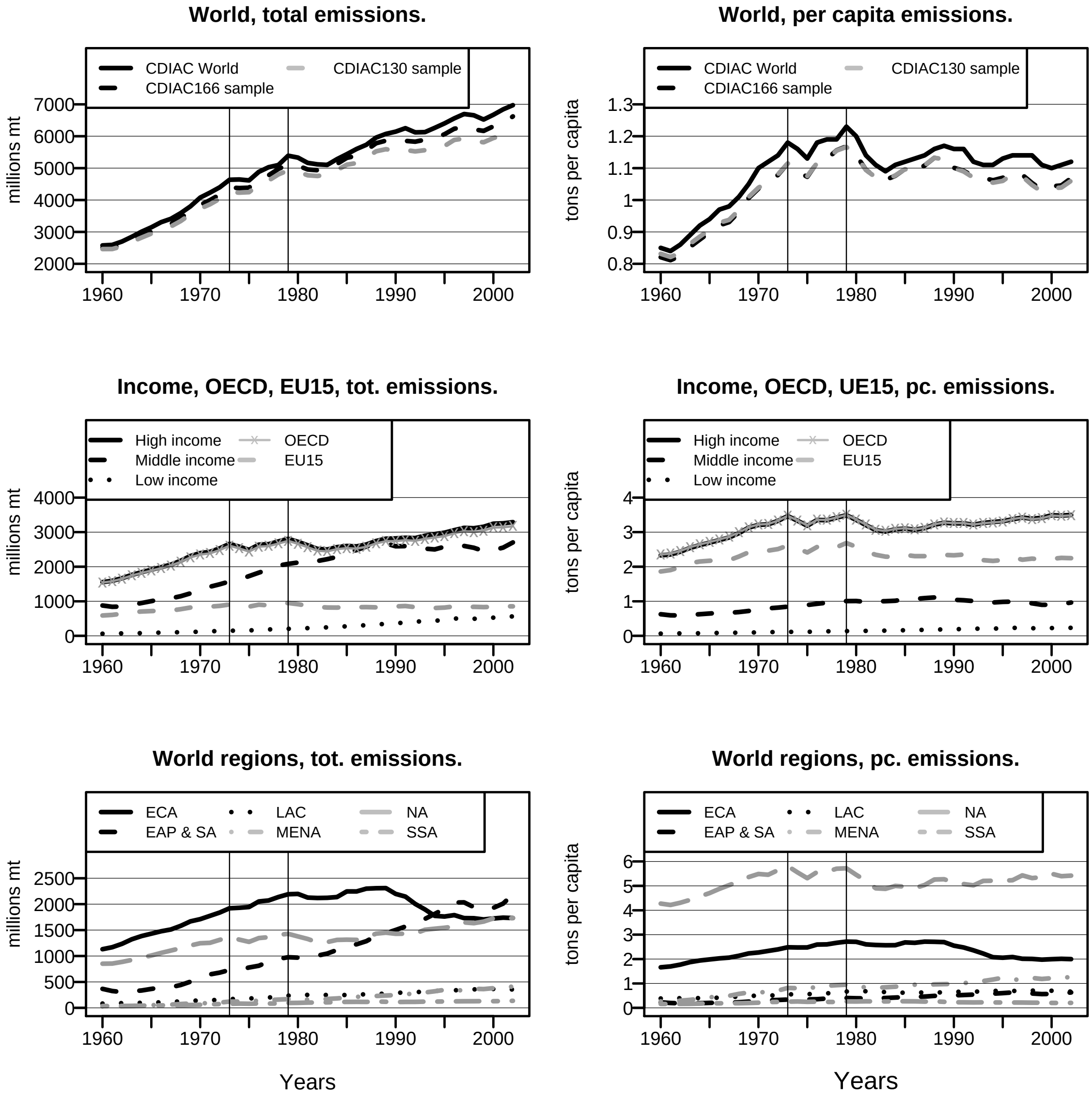
In lines 1 to 3 in Table 2.1, we can see that the CDIAC166 sample we built for the empirical investigation accounts for almost 97% (95%) of the total emissions of the CDIAC World in 1960 (2002), while the world sample without outliers CDIAC130 represents a fair 95% (90%) share of total emissions. This means that the most erratic series are not large carbon emitters. The three world samples display very similar trends. CDIAC World emissions went up from 2577 to 6973 thousands millions of metric tons during the 43 years, which represents an average growth rate of 2.4% per year. In per capita terms, the increase over the same period is significantly lower, 0.6% per year for the three world samples. The initial per capita emissions for CDIAC World is 0.85 metric tons of per capita emissions (mtPCE henceforth), which ranks CDIAC World at the 36th position among the CDIAC166 individual countries. By 2002, the per capita level and rank were 1.12 and 70th respectively. If we focus on selected individual countries from lines 4 to 14, we can see that the USA are by far the largest individual total carbon emitter in 1960 and 2002, with respectively 32% and 24% of CDIAC166 world's emissions. In 1960, Germany and China were at the second and third position, with respective shares of 8.9% and 8.1%. By 2002, while Germany managed to stabilize total emissions at its 1960 level, emerging economies such as China and India strongly increased their total emissions (by 3.6% and 5.5% per year respectively) and emissions' share (+6% and +3.7%). The largest increases in total emissions since 1960 took place in oil-producing countries (such as United Arab Emirates, Qatar, etc). It is important to highlight that in per capita terms, Chinese and

Indian emissions are far below the ones from more advanced economies. Luxembourg, UK, Germany and USA were the only rich and diversified economies producing over 3 mtPCE in 1960. These levels decreased over the period for the three European countries while US emissions have kept increasing on average by 0.5%. By 2002, Luxembourg, USA, Australia and Canada were the only diversified advanced economies producing more than 4 mtPCE, *i.e.* four times the world average level in 2002.

The aggregate series for the country groupings are shown in lines 14 to 28 of Table 2.1. In the Income groups, we notice that economies which have become the richest by 2002 (HI group) have increased their emissions by 1.8% per year on average. Despite the fact that the emissions' share of the latter group has declined in favor of the low and middle income economies, the wealthiest countries in 2002 are still the most intensive polluters in total as well as in per capita terms, and the emissions gap between the richest and the poorest has widened. The latter comment also holds for most intensive polluters in 1960 with respect to middle or less intensive ones. From the geographical point of view, East & Pacific Asia and South Asia has become the largest carbon emitter, followed by North America and Europe & Central Asia. In per capita terms, North America is by far the major polluter since 1960 and the pollution gap with respect to all other world regions has deepened.

Regarding the time trends in Figure 2.1, it is very clear on the left graphs that total emissions for all samples (world, income, regions, etc) have been strongly raising before the oil shocks, and kept increasing afterward at a lower pace. The aggregate emissions for the CDIAC world raised by roughly 2750 millions (+106%) of metric tons from 1960 until 1980 and 1340 millions metric tons (+25%) during the 1980-2002 period. The growth rates for the CDIAC166 and CDIAC 130 samples are fairly close. During the latter two sub-periods, total emissions have increased by +75% and +20% in the High Income as well as the OECD countries, by +140% and +16% in the Middle Income and +243% and +148% in the Low Income category. Regarding the geographic groups, the European & Central Asia group is the only one that has reduced its total emissions, by -25%, after the second oil shock. However, the most striking empirical fact comes from the series of per capita carbon emissions: most of them clearly stabilize either after the first or the second oil shock. In that respect, our data are in line with McKitrick and Strazicich (2005). In the presence of structural breaks in the individual and aggregate series, we may expect to find the same feature in the distributional dynamics of per capita emissions, *i.e.* non-stationary cross-sectional distributions that stabilize at some point. This is the question we address in the next section.

Figure 2.1: Time trends in world CO_2 emissions. Period 1960-2002.



Source: Figures on CO_2 emissions and population come from Marland et al. (2006) and U.S. Census Bureau (2006) respectively. CO_2 emissions are exclusively originated by fossil fuel consumption and cement production. The income and geographical countries' groupings are based on World Bank (2004). HI, MI and LI stand for High, Middle and Low Income. For the geographic areas: EAP = East Asia & Pacific, ECA = Europe & Central Asia (ECA), LAC = Latin America & Caribbean, MENA = Middle East & North Africa, NA = North America, SA = South Asia and SSA= Sub-Saharan Africa. The vertical lines plotted correspond to 1973 and 1979 oil shocks.

2.3.3 Evolution of cross-section distributions

Distributional dynamics has been extensively used in the income convergence literature, as well as in the empirics of PCE convergence with discrete and continuous approaches. As noted in section 2.2.2, this approach points towards 'relative persistence' worldwide, *i.e.* countries tend to keep their position relative to the world average but convergence is detected for a group of industrial economies. However, these conclusions have several methodological shortcomings[65]. For example, lack of consistency of the estimates when few observations are available and sensitivity of the results to initial conditions as well as to arbitrary scaling of the data. Our approach intends to address in a simple way these last two issues by proposing estimates which require neither assumptions on initial conditions nor scaling of the data.

Robust scale and shape statistics and distributional tests

Instead of adopting the Markovian process framework, we propose a simple exploratory distributional analysis. Firstly, we compute annual cross-sectional univariate kernel densities for unscaled PCE levels and report the results in 3D plots by stacking the annual estimates. The density estimates at time t are given by

$$\hat{f}_t(\mathrm{x}_{.t}) = \frac{1}{n_t h_t} \sum_{i=1}^{n_t} K\left(\frac{\mathrm{x}_{it} - \mathrm{x}_{.t}}{h_t}\right) \text{ for } t = 1960, \ldots, 2002 \tag{2.3}$$

where n_t is the sample size in year t, h_t is the smoothing parameter (bandwidth) for $\hat{f}_t$ and $K(.)$ represents a kernel function. We use Silverman's rule of thumb to determine h_t and a gaussian kernel for all years. Histograms with unequal bins are also presented to provide empirical frequency estimates for the first, middle and last years of the panel.

Secondly, we use robust scale and shape distributional measures to characterize how cross-sectional distribution evolve over time. Indeed, a more peaked density over time would suggest absolute convergence while a flattening shape with heavier tails provide hints for divergence. In between lies conditional convergence, which is compatible with a variety of distributional dynamics, *i.e.* tendencies to bi-polarity or multimodality that may stabilize once the balanced growth path is achieved for the 'converging' economies. Many authors emphasize that scale and shape distributional measures based on statistical

---

[65] See Stegman (2005).

moments display excessive sensitivity to outliers[66]. Moreover, it is now clear that the standard Kurtosis coefficient is heavily influenced by the tail weights. In that context, measures distinguishing peakedness from asymmetry and/or tail weight are to be preferred. The presence of highly volatile series also calls for a robust approach, which avoids arbitrary deletion of countries.

Brys et al. (2006) have recently compared the performance of different robust measures of peakedness and skewness by contaminating a large spectrum of symmetric distributions. They show that their asymmetry measure, called medcouple, as well as the peakedness estimator proposed by Schmid and Trede (2003) perform well and tolerate respectively up to 12.5% and 25% outliers before the estimator breaks down[67]. The robust peakedness statistic is defined as

$$P = \frac{x_{(1-p)} - x_{(p)}}{x_{(1-q)} - x_{(q)}} \tag{2.4}$$

where $0 < p < q < 0.5$ and $x_{(p)}$ is the pth quantile of a univariate sample[68]. The choice of p and q is arbitrary but following the afore-mentioned authors, we set $p = 0.125$ and $q = 0.25$, which is a good compromise between robustness and variance of the estimator. The peakedness indicator from equation (2.4) has several useful properties: in addition to its resistance to outliers, it is scale and location invariant[69] and it exists for any distribution. Given that this measure is not familiar, the following distributional benchmarks

---

[66] Remember that the standard asymmetry and kurtosis measures based on the rth central empirical moments m_r, are given by respectively $b_1 = m_3(X_n)/m_2(X_n)^{3/2}$ and $b_2 = m_4(X_n)/m_2(X_n)^2$, where X_n is a sample of size n. Brys et al. (2004, p.996) report that "One single outlier in the left tail of a symmetric or right-tailed sample can cause b_1 to become negative, whereas an outlier in the right tail of such a sample can unduly increase the classical skewness coefficient, making it hard to interpret". An explicit example is provided by Schmid and Trede (2003, p.2): "Generating $n = 10000$ standard normally distributed random observations, we obtain $b_2 = 3.0211$ which is very close to 3 (...). Replacing one observation by an outlier, say 10, we obtain $b_2 = 3.9155$ which is outside the critical value at the 1% level."

[67] The breakdown value is defined in Brys et al. (2006) as the amount of observations that need to be replaced in the sample to make the estimator worthless. For an univariate location estimator, the absolute value of the estimate breaks down when it becomes arbitrary large. For a scale estimator, *e.g.* the variance, the estimate breaks down when it becomes arbitrarily large or close to zero.

[68] Formally speaking, the n observations $X_n = \{x_1, x_2, \ldots, x_n\}$ are assumed independently sampled from a univariate distribution.

[69] $P(F_{aX+b}) = P(F_X)$.

could be useful for interpreting the results: for 10000 draws of the $N(0,1)$, $U(0,1)$, and $\frac{1}{2}N(-2,1)+\frac{1}{2}N(2,1)$ densities, P is respectively 1.70, 1.47, 1.33.

As robust indicator of skewness, Brys et al. (2006) advocate the medcouple. Let $X_n = x_1, x_2, ..., x_n$ be a i.i.d. univariate sample, ordered such that $x_1 \leq x_2 \leq \ldots \leq x_n$, the medcouple is given by:

$$MC = \underset{x_i \leq x_{(0.5)} \leq x_j}{\text{med}} h(x_i, x_j) \tag{2.5}$$

For $x_i \neq x_j$, the kernel $h()$ is defined as

$$h(x_i, x_j) = \frac{(x_j - x_{(0.5)}) - (x_{(0.5)} - x_i)}{x_j - x_i} \tag{2.6}$$

In the special case where $x_i = x_j = x_{(0.5)}$, $h()$ is set to

$$h(x_i, x_j) = \begin{cases} +1 & \text{for } i > j \\ 0 & \text{for } i = j \\ -1 & \text{for } i < j \end{cases} \tag{2.7}$$

The kernel function $h(.)$ is a standardized difference between the x_js and the x_is to the median and lies between $+1$ and -1. It is positive (resp. negative) when x_j (x_i) lies further from the median and it equals zero in case of a perfectly symmetric distribution. When the median coincides with observations included in the sample, the function $h(x_i, x_j)$ takes the values of 1 if $i > j$, -1 if $i < j$ and 0 if $i = j$. The median of all these kernel values gives the medcouple. A positive (negative) value of MC indicates a right-tailed (or left-tailed) distribution while 0 means symmetry. Note that MC is scale and location invariant, and it is robust to the presence of 25% outliers. Finally, given that the concept of σ-convergence is nothing but the evolution of cross-sectional distributions' dispersion over time, we also report the inter-quartile range (IQR henceforth) as a robust measure of spread.

Clearly, identifying increasing/decreasing peakedness and/or asymmetry for a distribution is not enough to capture a potential convergence club phenomenon. The literature on income convergence also employs multimodality tests. Among the many procedures available, Hartigan and Hartigan (1985) propose the so-called 'dip statistic', which measures departure from unimodality in the same spirit as the Kolmogorov-Smirnov test. Therefore we also control for multimodality in the cross-section distributions.

Finally, rather than computing differences between specific scale and shape statistics over time, we employ pair-wise comparisons of distributions for successive decades. The closeness between two distributions f and g,

$$\begin{aligned} H_0 :\ & f(x) = g(y) \\ & vs. \\ H_1 :\ & f(x) \neq g(y) \end{aligned} \tag{2.8}$$

can be checked by using the standard Kolmogorov-Smirnov test (KS test). Noting X_n and Y_m two samples of size n and m, recall that the KS procedure relies on the maximum distance (called D_{nm}) between the two empirical CDFs F_n and G_m. This test statistic follows a Kolmogorov-Smirnov distribution[70], noted $\sqrt{\frac{nm}{(n+m)}} D_{nm} \xrightarrow{d} K$. We also apply a procedure proposed by Li (1996) which accounts for dependencies between x and y. The latter test relies on the integrated square difference, its empirical statistic $J_n \xrightarrow{d} N(0,1)$ and the test is one-sided[71].

---

[70] Noting $D_{mn} = \sup |F_n(x) - G_m(x)|$, $d = \sqrt{\frac{nm}{(n+m)}}$, and when nm is large enough, the critical significance cutoffs for $\alpha = 10\%$, 5% and 1% levels can be approximated by $1.22/d$, $1.36/d$ and $1.63/d$ respectively.

[71] As indicated by Li (1996), the integrated square difference (noted I) is given by

$$I = \int [f(x) - g(x)]^2 \mathrm{d}x = \int [(f(x)\mathrm{d}F(x) + g(x)\mathrm{d}G(x) - 2g(x)\mathrm{d}G(x)].$$

When $m = n$, as in our balanced panel, the latter expression becomes feasible by replacing $f(x)$, $F(x)$, $g(x)$ and $G(x)$ by $f_n(x)$, $F_n(x)$, $g_n(x)$ and $G_n(x)$. Then I becomes I_n and the densities can be consistently estimated by kernel estimators. Adopting the notations $K^a_{ij} = K(\frac{a_i - a_j}{h})$, $K^{a,b}_{ii} = K(\frac{a_i - b_i}{h})$, $K^{a,b}_{ij} = K(\frac{a_i - b_j}{h})$ and $c = \frac{1}{(n^2 h^p)}$, I_n, it can be shown that:

$$I_n = I_{n1} + I_{n2} = c \sum_{i=1}^{n} \{2K(0) - 2K^{x,y}_{i,i}\} + c \sum_{i=1}^{n} \sum_{j \neq i, j=1}^{n} \{K^x_{ij} + K^y_{ij} - K^{x,y}_{ij} - K^{y,x}_{ij}\}$$

Li (1996) demonstrates that only the I_{n2} term is needed for the test. Under the null and when $h \rightarrow$

Note that these global distributional tests can be carried out to either compare two distributions over time (time poolability or homogeneity of the data over time) or the homogeneity of two groups for the same year (spatial poolability or homogeneity of two groups which compose the whole sample). In the empirical results, we check for both homogeneity concepts. With these distributional measures at hand, we can now turn to the analysis of the per capita carbon distributions.

Distributional results

World distributions. The evolution of the PCE cross-sectional densities for the CDIAC166 World panel is presented on the left panel of Figure 2.2 and a discretized counterpart is shown on the right hand side. Clearly, the mass of the highly peaked and right-skewed density (or the modal class of the histogram) in 1960 tends to migrate toward higher PCE levels over time. This dynamic is particularly strong during the 1960-1970 period but the PCE distribution stabilizes after the first oil-shock. The annual cross-sectional percentiles, drawn on the floor of the 3-dimensional device, show that the process is essentially driven by upper quantiles, *i.e.* the most intensive polluters at the beginning of the period. The cross-sectional mean is influenced by high values while the median and the lower 25th quantile keep increasing at relatively constant pace over the whole period. This points toward differentiated dynamics linked to initial PCE levels.

Table 2.2 provides the distributional statistics and tests for cross-sectional densities of the CDIAC166 World panel for the years 1960, 1970, 1980 1990 and 2000. Firstly, most of the scale and shape statistics based on moments appear to behave erratically and do not allow any precise inference. By contrast, the measures based on quantiles are much more stable and behave in accordance with the plotted density estimates.

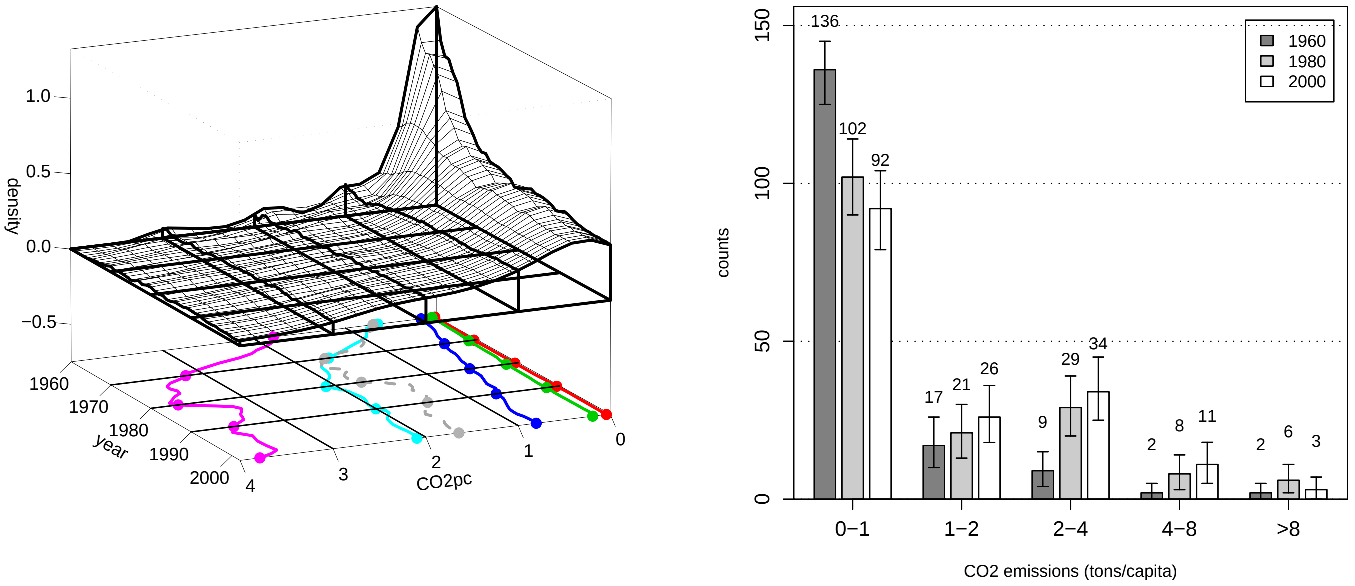
---

0 and $nh^p \to \infty$, he shows that $\frac{1}{nh^{p/2}} I_{n2} \xrightarrow{d} N(0, \sigma_0)$, where $\sigma_0 = 2\{\int[f(x) - g(x)]^2 \mathrm{d}x\}[\int K^2(u)\mathrm{d}u]$. This variance term can then be estimated in the same way as before:

$$\hat{\sigma}_0 = 2c \sum_{i=1}^{n} \sum_{j=1}^{n} [K_{ij}^x + K_{ij}^y - 2K_{ij}^{x,y}][\int K^2(u)\mathrm{d}u]$$

With I_{2n} and $\hat{\sigma}_0$ at hand , the test-statistic is provided by $J_n = \frac{1}{nh^{p/2}} I_{n2}/\hat{\sigma}_0 \xrightarrow{d} N(0,1)$. Note that no optimal bandwidth is needed in this test and that the asymptotic result holds when x_i and y_j are dependent or not for $i = j$. However, independence when $i \neq j$ is required.

Figure 2.2: Cross-sectional densities and histograms of the World CO_2 per capita emissions. Period 1960-2002.



Left graph: the red, green, navy blue, sky blue and magenta lines drawn on the floor represent respectively the 10%, 25%, 50%, 75% and 90% cross-sectional quantiles over time and are computed with a locally linear nonparametric quantile regressions, see Koenker (2007). The grey dashed line is the cross-sectional mean, slightly smoothed with a kernel regression (function *ksmooth*). The univariate cross-sectional kernel densities are estimated with a gaussian kernel and Silverman's rule of thumb for the bandwidth (function *density* and *bw=nrd0* in *R.2.6.0*). **Right graph:** 95% bootstrapped confidence bounds are computed for the frequencies.

Table 2.2: World cross-sectional densities of CO$_2$ per capita emissions. Scale and shape indicators. Period 1960-2000.

Statistic/Test	1960		1970		1980		1990		2000
Scale and shape statistics based on moments									
Spread (variance)	2.7		19.3		14.6		5.2		9.8
	[0.6,6.0]		[4.3,42.7]		[3.3,34.1]		[1.9,10.6]		[2.4,22.4]
Asymmetry (skewness)	6.0		6.0		7.0		5.1		6.7
	[4.2,9.8]		[3.9,8.7]		[5.5,11.8]		[4.0,8.8]		[5.4,11.5]
Peakedness (kurtosis)	48.0		47.4		67.4		43.1		63.6
	[14.8,87.8]		[12.0,79.9]		[42.4,126.2]		[28.3,81.4]		[42.6,120.0]
Scale and shape statistics based on quantiles									
Spread (IQR)	0.59		1.60		1.97		1.84		1.92
	[0.25,0.79]		[1.38,2.15]		[1.51,2.57]		[1.30,2.17]		[1.59,2.46]
Asymmetry (MC)	0.69		0.70		0.63		0.45		0.41
	[0.62,0.90]		[0.59,0.86]		[0.51,0.77]		[0.17,0.62]		[0.25,0.62]
Peakedness	2.60		1.74		1.72		1.62		1.43
	[1.64,3.60]		[1.09,2.08]		[1.20,1.98]		[1.21,1.91]		[0.85,1.59]
Global shape tests									
Unimod. vs multimod. (Dip test)	0.015		0.022		0.014		0.020		0.022
Shape equality (KS test)		0.17^{**}		0.10		0.06		0.06	
Shape equality (Li (1996) test)		2.10^{**}		0.56		0.28		0.26	

Note: In brackets are 95% basic bootstrap confidence limits, see Davison and Hinkley (1997, p.28-29). To avoid negative values for the variance's lower confidence bound, we used the basic percentile method (see *ibid*, p. 202-203). Note that we resampled blocks of full length in the time dimension. IQR and MC stand for interquartile range and medcouple respectively. * and ** indicate 5% and 10% significance levels for the Dip statistic, Kolmogorov-Smirnov (KS) and Li (1996) tests. For the DIP test, spline interpolations from the finite sample tabled critical values have been used (see Hartigan and Hartigan (1985)). Large samples' asymptotic values are applied in the other tests. All computations have been implemented in *R.2.6.0.*

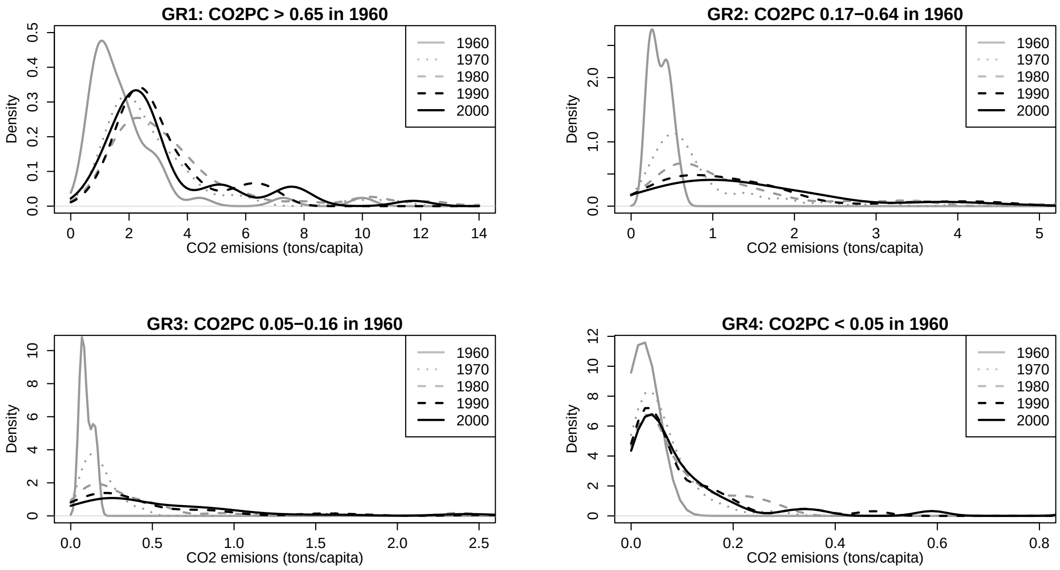
Therefore, we interpret only robust measures. The IQR statistic appears to increase strongly during the 1960-1980 period and remains steady afterwards. In terms of σ-convergence, this points toward divergence during the period 1960-1980 and relatively stable worldwide PCE differences since 1980.

The MC indicates that the cross-sectional densities are right-tailed over the whole period and that asymmetry reduces after the 70s. Peakedness clearly decreases over the whole period, with a large drop between 1960 and 1970. No significant departure from unimodality is detected in the reported cross-sectional distributions with the Dip test. Finally, both KS and Li's tests show a significant difference between the distributions in 1960 and 1970 but none for the next successive pairs of decades. Given that the Kolmogorov-Smirnov statistic provides very close results to Li's test, we report only the latter in the subsets' analysis.

The distributional analysis does not support unconditional convergence of PCE worldwide over the whole period. The strong flattening of the distribution in the first 10 years points toward divergence, driven essentially by large polluters. During the 70's oil-shocks, the distributional dynamics clearly stabilizes. The latter pattern leaves the door open for conditional convergence to occur. However, the multimodality test rejects any significant polarization phenomenon in the cross-sectional world distributions. Carrying the same analysis for groupings of countries will allow to detect to what extent the world pattern holds for countries which share some common structural characteristics.

Initial polluters. To better capture the internal dynamics of the cross-section world PCE distributions, we explore how the densities evolve when close initial polluters are put together. Four groups ranging from the most to the least intensive initial polluters are constructed according to the methodology described in section 2.3.1. Figure 2.3 and Table 2.3 report respectively the density fits and the distributional statistics for these groups. We see in Table 2.3 that the structural shock identified at the world level after the 70s holds for the four groups as pair-wise time poolability is rejected between 1960 and 1970 in all cases. We also strongly reject the equality between each group's distribution and the rest of the data distributions for all periods (see the 'Spatial pool.' line). Thus, different dynamics seem to affect each group. Most intensive initial polluters have highly peaked cross-sectional densities according to our robust peakedness indicator while least ones are characterized by strongly right-skewed distributions since 1970. Note that the peakedness

Figure 2.3: Initial polluters' cross-sectional densities of CO_2 per capita emissions. Period 1960-2000.



statistic does not always match the pattern depicted in the graphical analysis. An example is given for the lower-middle initial polluters, whose peakedness indicator increases steadily over time while the cross-sectional kernel densities flatten. Indeed, larger dispersion in the data may result in larger peakedness as well. Recall that the peakedness formula (2.4) accounts for a pretty large domain in the center of the distribution and that its numerator and denominator are both positive functions of dispersion. Regarding the *IQR* statistic, we notice on Table 2.3 that the spread keeps increasing over the decades for the middle initial polluters but it stabilizes for the most and least intensive initial polluters at some point. Therefore, while the latter subgroups may depict some (conditional) convergence after the 70s with more formal tests, the other two are not expected to do so. Note also that no multimodality is detected within these subgroups.

Income groups. Figure 2.4 depicts the annual cross-sectional densities for the income groupings of countries. First note the similarity between the large/low initial polluters patterns in Figure 2.3 and those for high/low income groups. Given that the richest/poorest countries are among the largest/lowest emitters of carbon PCE, both groups tend to produce similar distributional tests and statistics. Regarding the middle income countries,

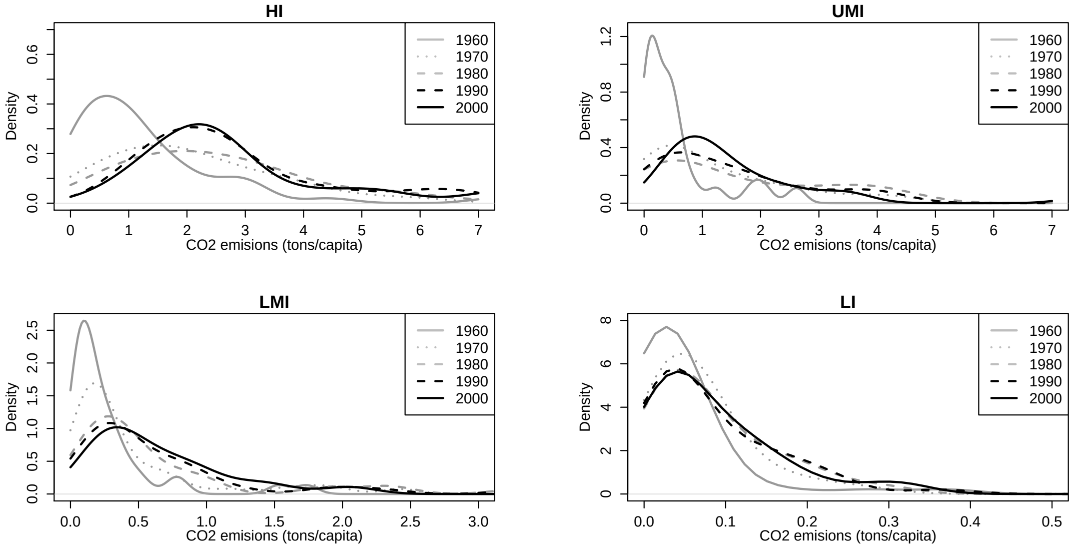
Table 2.3: Initial polluters' cross-sectional densities of CO_2 per capita emissions. Scale and shape indicators. Period 1960-2002.

Sample (size)	Statistic / Test	1960	1970	1980	1990	2000
PCE > 0.65 (42)	Spread (IQR)	1.26	1.97	2.21	1.79	1.97
	Asymmetry (MC)	0.17	0.45	0.30	0.31	0.44
	Peakedness	1.75	3.88	2.71	2.22	2.89
	Multimodality test	0.057	0.036	0.036	0.035	0.044
	Spatial pool.	79.6**	56.1**	46.2**	46.1**	35.8**
	Pair-wise time pool.	4.32**	0.39	0.24	0.37	
PCE $\in$ 0.17-0.64 (41)	Spread (IQR)	0.20	0.50	0.90	1.08	1.42
	Asymmetry (MC)	0.19	0.48	0.39	0.09	0.18
	Peakedness	1.60	2.34	2.42	2.04	1.51
	Multimodality test	0.046	0.044	0.040	0.039	0.044
	Spatial pool.	131**	49.5**	26.7**	18.6**	10.3**
	Pair-wise time pool.	4.07**	0.41	0.34	-0.08	
PCE $\in$ 0.05-0.16 (42)	Spread (IQR)	0.07	0.17	0.33	0.51	0.55
	Asymmetry (MC)	0.52	0.18	0.47	0.63	0.43
	Peakedness	1.40	1.66	1.61	1.62	1.72
	Multimodality test	0.050	0.059	0.038	0.040	0.039
	Spatial pool.	164**	46.6**	19.1**	13.6**	13.0**
	Pair-wise time pool.	4.42**	0.75	0.34	0.39	
PCE < 0.05 (41)	Spread (IQR)	0.02	0.05	0.10	0.11	0.11
	Asymmetry (MC)	0.05	0.60	0.57	0.55	0.54
	Peakedness	1.33	1.91	2.11	1.74	1.80
	Multimodality test	0.061	0.038	0.040	0.036	0.050
	Spatial pool.	241**	126**	87.8**	94.5**	92.9**
	Pair-wise time pool.	2.07**	0.17	-0.09	-0.10	

Note: IQR and MC stand for interquartile range and medcouple respectively. * and ** indicate 5% and 10% significance levels for the multimodality test, *i.e.* Dip statistic by Hartigan and Hartigan (1985), as well as for the shape equality test, *i.e.* J_n statistic by Li (1996). Spline interpolations from the tabulated critical values by Hartigan and Hartigan (1985, Table 1) have been used for the multimodality test while the statistic for the shape equality test is asymptotically standard normal.

we can see in Table 2.4 that no structural break is detected with the time poolability test. However, differences exist between upper-middle and lower-middle income subsets in the spatial dimension: while upper-middle income countries do not evolve differently than the rest of the world for most years, lower-middle income economies do. As for the previous grouping, no multimodality is detected within income groups. In terms of convergence, we expect rich/poor economies to mimic the large/low initial polluters. Convergence is not likely to occur in the middle income classes, at least before the end of the 70s.

Figure 2.4: Income groups' cross-sectional densities of CO_2 per capita emissions. Period 1960-2000.



Note: Kernel densities are computed with Silverman's optimal bandwidth and a gaussian kernel.

Geographic groups. Figure 2.5 reports the smoothed distributions for the geographic regions. Europe and Central Asia (ECA)'s densities become more symmetric over time and their dispersion is stable over the whole period according to the results in Table 2.5. The EAP, MENA and LAC kernel densities remain strongly asymmetric over time and experience increasing dispersion until at least the beginning of the 80s. Time and spatial poolability is mainly accepted for the EAP and MENA regions.

The LAC subgroup departs from the world pattern since 1980. Regarding the SA-SSA group, its members behave like the low income economies. In none of the regions we find significant polarization in PCE. Here, absolute convergence is more likely to occur within the ECA zone as a slight σ-convergence tendency is detected.

Regional Integration Agreements' groups. Finally, Figure 2.6 reports the distributional dynamics within some selected RIA, classified according to their North-North, North-South or South-South dimension. Clearly, the graphical devices for the North-North integration areas OECD/EU15 display increasing/stabilized peakedness and decreasing dispersion. The North-South patterns for APEC12 depict a strong increase in dispersion during the 60s which stabilize since then. The latter results are globally corroborated by

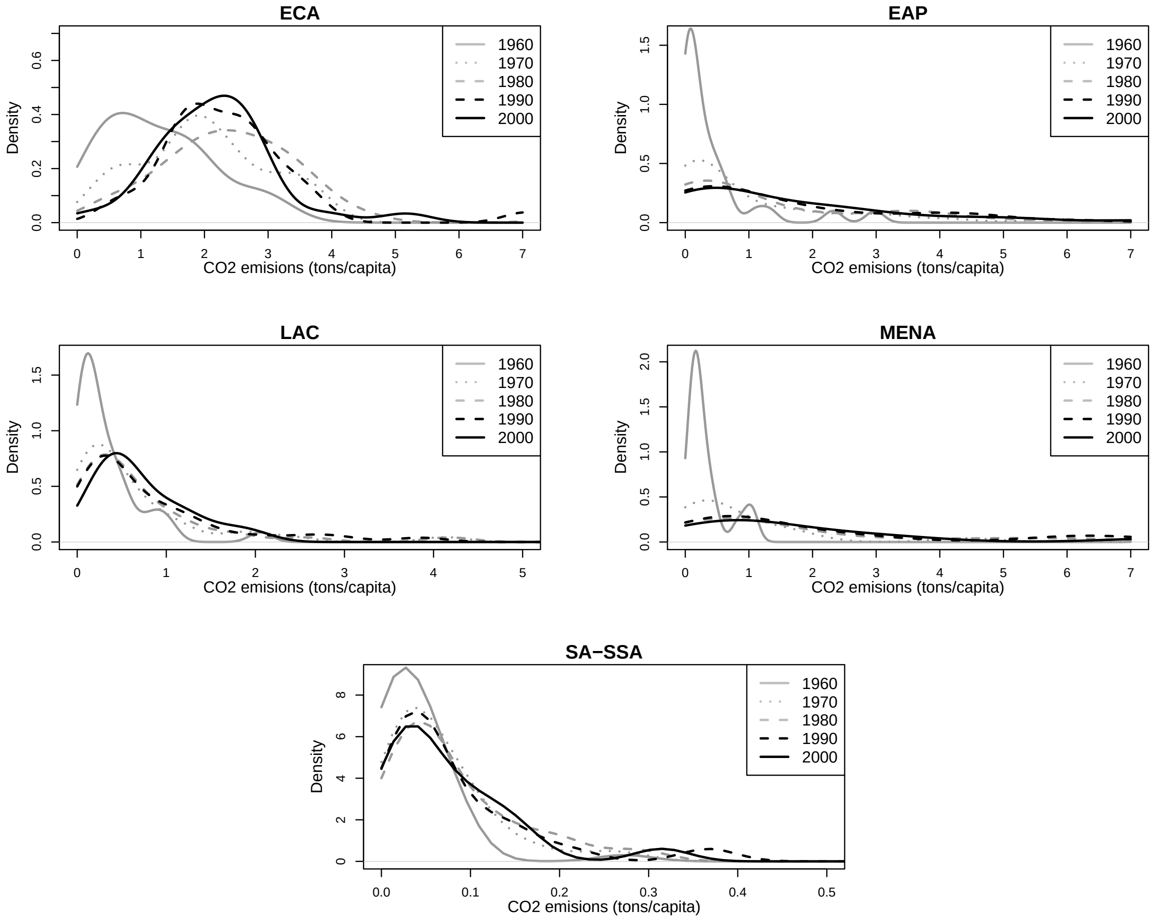
Table 2.4: Income groups' densities of CO_2 per capita emissions. Scale and shape distributional indicators. Period 1960-2000.

Sample (size)	Statistic / Test	1960	1970	1980	1990	2000
High Income (48)	Spread (IQR)	1.40	2.51	2.60	1.93	2.02
	Asymmetry (MC)	0.33	0.61	0.32	0.18	0.52
	Peakedness	1.97	3.69	2.83	2.34	2.84
	Multimodality test	0.063	0.042	0.044	0.045	0.043
	Spatial pool.	2.58^{**}	3.51^{**}	4.37^{**}	6.86^{**}	6.78^{**}
	Pair-wise time pool.		3.25^{**}	0.17	0.72	-0.04
Upper-Middle Income (24)	Spread (IQR)	0.46	1.43	2.14	1.56	1.24
	Asymmetry (MC)	-0.10	0.39	0.44	0.35	0.42
	Peakedness	2.88	1.64	1.58	1.79	1.77
	Multimodality test	0.071	0.059	0.038	0.040	0.039
	Spatial pool.	1.82^{**}	0.75	0.39	0.86	6.25^{**}
	Pair-wise time pool.		0.71	-0.08	0.05	0.31
Lower-Middle Income (38)	Spread (IQR)	0.24	0.43	0.56	0.59	0.61
	Asymmetry (MC)	0.38	0.44	0.32	0.35	0.33
	Peakedness	1.84	2.03	2.17	1.62	1.64
	Multimodality test	0.043	0.047	0.047	0.040	0.039
	Spatial pool.	7.62^{**}	8.01^{**}	10.2^{**}	13.1^{**}	16.6^{**}
	Pair-wise time pool.		0.54	1.14	.0.10	-0.04
Low Income (49)	Spread (IQR)	0.05	0.06	0.09	0.11	0.11
	Asymmetry (MC)	0.49	0.28	0.45	0.51	0.52
	Peakedness	1.61	2.44	2.03	1.58	1.69
	Multimodality test	0.033	0.033	0.033	0.030	0.042
	Spatial pool.	155^{**}	125^{**}	116^{**}	118^{**}	126^{**}
	Pair-wise time pool.		2.49^{**}	-0.09	-0.24	-0.04

Note: IQR and MC stand for interquartile range and medcouple respectively. * and ** indicate 5% and 10% significance levels for the multimodality test, *i.e.* Dip statistic by Hartigan and Hartigan (1985), as well as for the shape equality test, *i.e.* J_n statistic by Li (1996). Spline interpolations from the tabulated critical values by Hartigan and Hartigan (1985, Table 1) have been used for the multimodality test while the statistic for the shape equality test is asymptotically standard normal.

the scale and shape statistics in Table 2.6. We notice that time poolability is accepted for all groups and all successive pairs of decades while spatial poolability is rejected for OECD, EU15, suggesting a different data generation process from the rest of the world. Regarding the South-South RIA, CARICOM10, COMESA12 and LAIA have quite diverse cross-sectional densities, suggesting potential bipolarity in the CARICOM10 case. ECOWAS densities display a surprising increase in peakedness in 1990 which decreases in 2000. The scale and shape measures in Table 2.7 confirm these graphical patterns and indicate that time poolability is accepted for all South-South RIA while spatial poolability

Figure 2.5: Geographic groups' cross-sectional densities of CO_2 per capita emissions. Period 1960-2000.



Note: Kernel densities are computed with Silverman's optimal bandwidth and a gaussian kernel.

is rejected in most cases (COMESA12, ECOWAS and LAIA). The results regarding time poolability may be surprising as quite different patterns are depicted by the kernel estimates, specially when comparing 1960 *vs.* 1970 distributions. However, stronger contrasts may be needed for such small groups to conclude to distributional differences. This sample size issue does not affect the spatial poolability tests as the density of each subgroup is compared to the density of the rest of the world, which includes a large enough amount of data. Moreover comparing successive decades may be less effective in detecting differences than choosing a fixed benchmark in 1960. The most salient result in Table 2.6 is that while peakedness and dispersion for the OECD and EU15 groups are rather steady during

Table 2.5: Geographic groups' cross-sectional densities of CO_2 per capita emissions. Scale and shape indicators. Period 1960-2000.

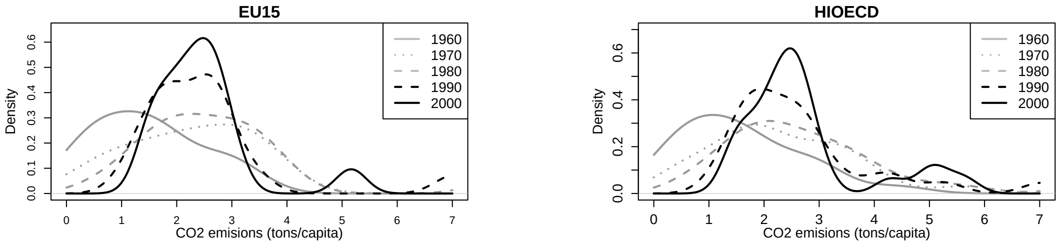
Sample (size)	Statistic / Test	1960	1970	1980	1990	2000
Europe & Central Asia (ECA) (30)	Spread (IQR)	1.29	1.15	1.43	1.03	1.17
	Asymmetry (MC)	0.28	0.28	0.11	0.03	0.02
	Peakedness	1.76	2.36	1.52	1.74	1.21
	Multimodality test	0.078	0.052	0.045	0.050	0.065
	Spatial pool.	22.2**	30.6**	29.0**	32.0**	26.5**
	Pair-wise time pool.		1.46*	0.09	-0.01	-0.01
East Asia & Pacific (EAP) (29)	Spread (IQR)	0.40	1.36	1.98	2.17	2.17
	Asymmetry (MC)	0.65	0.73	0.72	0.51	0.46
	Peakedness	2.12	2.22	1.80	1.86	1.63
	Multimodality test	0.051	0.047	0.050	0.044	0.052
	Spatial pool.	-1.88	-1.43	-0.96	-0.76	-1.08
	Pair-wise time pool.		0.18	0.10	0.05	-0.01
Middle-East & North Africa MENA (19)	Spread (IQR)	0.28	1.35	2.67	2.19	2.52
	Asymmetry (MC)	0.68	0.60	0.77	0.82	0.71
	Peakedness	2.96	6.05	2.11	2.72	2.96
	Multimodality test	0.049	0.073	0.055	0.073	0.050
	Spatial pool.	5.46**	-0.66	-0.39	0.03	-0.24
	Pair-wise time pool.		0.79	0.19	0.13	0.03
Latin America & Caribbean (LAC) (38)	Spread (IQR)	0.38	0.77	0.84	0.88	0.76
	Asymmetry (MC)	0.41	0.51	0.52	0.66	0.49
	Peakedness	2.04	2.29	2.57	1.96	2.00
	Multimodality test	0.036	0.040	0.031	0.045	0.047
	Spatial pool.	0.64	1.38	3.22**	4.86**	11.88**
	Pair-wise time pool.		0.44	-0.06	-0.08	0.57
Sou. Asia & Sub Sah. Africa (SA-SSA) (42)	Spread (IQR)	0.05	0.06	0.09	0.09	0.10
	Asymmetry (MC)	0.39	0.57	0.47	0.41	0.43
	Peakedness	1.44	2.26	2.08	1.86	1.49
	Multimodality test	0.041	0.043	0.055	0.045	0.042
	Spatial pool.	115**	95.8**	89.5**	102**	97.6**
	Pair-wise time pool.		1.57*	0.06	-0.25	0.08

Note: IQR and MC stand for interquartile range and medcouple respectively. * and ** indicate 5% and 10% significance levels for the multimodality test, *i.e.* Dip statistic by Hartigan and Hartigan (1985), as well as for the shape equality test, *i.e.* J_n statistic by Li (1996). Spline interpolations from the tabulated critical values by Hartigan and Hartigan (1985, Table 1) have been used for the multimodality test while the statistic for the shape equality test is asymptotically standard normal.

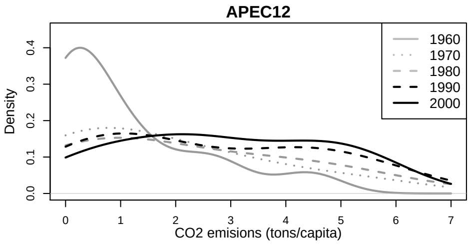
the 1960-1980 period, peakedness increases and dispersion strongly decreases afterwards. This is the only clear distributional absolute convergence profile among the groupings considered until now.

Figure 2.6: North-North and North-South RIA's cross-sectional densities of CO_2 per capita emissions. Period 1960-2000.

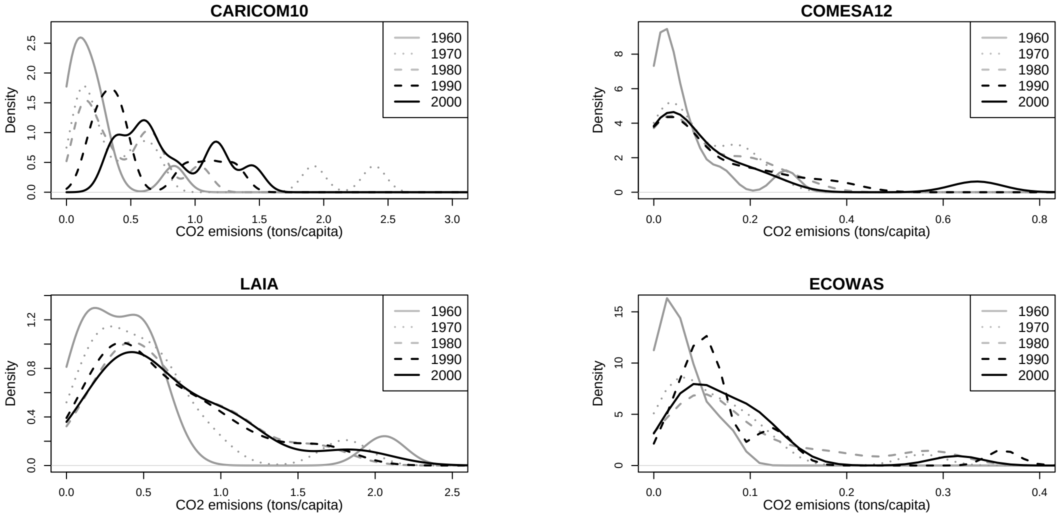
North-North RIA



North-South RIA



South-South RIA



Note: Kernel densities are computed with Silverman's optimal bandwidth and a gaussian kernel.

Table 2.6: North-North & North-South RIA's cross-sectional densities of CO_2 per capita emissions. Scale and shape indicators. Period 1960-2000.

Sample (size)	Statistic / Test	1960	1970	1980	1990	2000
NN-OECD (24)	Spread (IQR)	1.56	1.56	1.64	1.02	0.76
	Asymmetry (MC)	0.13	0.35	0.27	0.10	0.30
	Peakedness	1.69	1.77	1.49	2.51	3.61
	Multimodality test	0.057	0.069	0.045	0.050	0.053
	Spatial pool.	17.3**	20.3**	21.2**	26.1**	28.0**
	Pair-wise time pool.	0.69	-0.11	0.16	0.13	
NN-EU15 (15)	Spread (IQR)	1.50	1.67	1.46	0.92	0.71
	Asymmetry (MC)	0.00	0.06	0.14	-0.22	-0.10
	Peakedness	1.76	1.62	1.51	1.52	1.78
	Multimodality test	0.074	0.065	0.058	0.077	0.072
	Spatial pool.	8.56**	10.1**	12.2**	14.4**	15.6**
	Pair-wise time pool.	0.52	-0.04	0.25	-0.17	
NS-APEC12 (12)	Spread (IQR)	1.45	3.02	3.54	3.40	2.97
	Asymmetry (MC)	0.60	0.23	0.28	0.23	0.17
	Peakedness	1.79	1.63	1.42	1.36	1.54
	Multimodality test	0.070	0.060	0.064	0.010	0.085
	Spatial pool.	-0.71	-0.05	-0.01	0.96	2.73**
	Pair-wise time pool.	-0.18	0.01	0.02	0.04	

Note: IQR and MC stand for interquartile range and medcouple respectively. * and ** indicate 5% and 10% significance levels for the multimodality test, *i.e.* Dip statistic by Hartigan and Hartigan (1985), as well as for the shape equality test, *i.e.* J_n statistic by Li (1996). Spline interpolations from the tabulated critical values by Hartigan and Hartigan (1985, Table 1) have been used for the multimodality test while the statistic for the shape equality test is asymptotically standard normal. 'NN' and 'NS' stand for respectively 'North-North' and 'North-South' regional integration agreement.

Globally, this descriptive analysis indicates that the significant distributional differences prevalent between the pre and post oil-shocks period at the world level also affects some subsets of countries (the initial polluters subgroups, the high/low income economies as well as the European & Central Asian or Sub-Saharan African regions). Note that it may be inaccurate to use asymptotic cutoffs in comparing distributional shapes for small samples. In our case, this may be an issue with the time poolability results, but not for spatial poolability ones. Many subgroups possess significantly different cross-sectional distributions compared to the rest of the world over the whole period. However, only two of them, OECD and UE15 countries, display distributional patterns clearly compatible with absolute convergence in the sense that the cross-sectional distributions become more peaked with a decreasing trend in dispersion over the whole period. The vast majority of the other groups produce flatter distributions, at least during the first decades, with

Table 2.7: South-South RIA's cross-sectional densities of CO_2 per capita emissions. Scale and shape indicators. Period 1960-2000.

Sample (size)	Statistic / Test	1960	1970	1980	1990	2000
SS-CARICOM10 (10)	Spread (IQR)	1.81	0.54	0.854	0.74	0.59
	Asymmetry (MC)	0.00	0.24	0.10	0.72	0.33
	Peakedness	1.44	3.10	1.60	1.42	1.67
	Multimodality test	0.036	0.040	0.031	0.045	0.047
	Spatial pool.	-0.93	-1.54	-0.44	0.55	3.84^{**}
	Pair-wise time pool.	0.08	-0.08	0.11	0.42	
SS-COMESA12 (12)	Spread (IQR)	0.06	0.13	0.15	0.14	0.13
	Asymmetry (MC)	0.44	0.52	0.62	0.64	0.67
	Peakedness	1.83	1.17	1.32	1.55	1.56
	Multimodality test	0.079	0.103	0.082	0.055	0.083
	Spatial pool.	9.28^{**}	8.15^{**}	9.85^{**}	10.6^{**}	12.0^{**}
	Pair-wise time pool.	0.17	-0.18	-0.11	-0.02	
SS-ECOWAS (16)	Spread (IQR)	0.03	0.06	0.08	0.04	0.06
	Asymmetry (MC)	0.20	0.30	0.54	0.30	0.25
	Peakedness	1.76	1.87	2.33	2.53	1.32
	Multimodality test	0.057	0.070	0.058	0.060	0.068
	Spatial pool.	34.1^{**}	22.4^{**}	20.1^{**}	37.5^{**}	31.3^{**}
	Pair-wise time pool.	0.39	-0.09	-0.09	0.34	
SS-LAIA (12)	Spread (IQR)	0.34	0.39	0.51	0.52	0.63
	Asymmetry (MC)	-0.07	-0.01	0.49	0.52	0.52
	Peakedness	1.53	1.75	1.60	1.62	1.22
	Multimodality test	0.103	0.098	0.054	0.075	0.105
	Spatial pool.	2.45^{**}	2.09^{**}	3.30^{**}	2.47^{**}	3.20^{**}
	Pair-wise time pool.	-0.05	0.05	0.02	-0.03	

Note: IQR and MC stand for interquartile range and medcouple respectively. * and ** indicate 5% and 10% significance levels for the multimodality test, *i.e.* Dip statistic by Hartigan and Hartigan (1985), as well as for the shape equality test, *i.e.* J_n statistic by Li (1996). Spline interpolations from the tabulated critical values by Hartigan and Hartigan (1985, Table 1) have been used for the multimodality test while the statistic for the shape equality test is asymptotically standard normal. 'SS' stands for 'South-South' regional integration agreement.

variable asymmetries (mainly right skewed) and increasing dispersion, which stabilizes at some point in time. The latter patterns are more difficult to interpret in terms of convergence as they are characteristic of heterogeneous dynamics at the individual level. Therefore, putting together countries on the basis of their initial pollution levels, income, geographic location or membership of a regional integration agreement may allow to reduce the intra-group heterogeneity to some extend but does not necessarily avoid unambiguous conclusions regarding convergence. The use of alternative disparities distributional indicators, going from the simple Gini index to more involved indicators such as the entropy measures of Maasoumi et al. (2007) or the polarization measures used in Ezcurra (2007) would allow clearer inference. Another option is to employ a methodology which explicitly accounts for the heterogeneity characterizing the dynamic of each individual series within the group and provide results for the panel as a whole. Recent time-series techniques propose several methods to carry out the latter analysis. One of them is developed in the next section.

2.4 Time series analysis

In this section, we follow Westerlund and Basher (2008) and apply a the pair-wise convergence method of Evans (1998) to different world panel datasets as well as to the aforementioned grouping of countries. The panel unit root approach to pair-wise convergence is presented in section 2.4.1 and the results of the econometric estimates are detailed in subsection 2.4.2, and compared with those obtained from the distributional analysis.

2.4.1 Panel based tests for convergence

Following the presentation in Pedroni and Yao (2006), suppose that $y_{i,t}$ and $y_{j,t}$, the per capita CO_2 emissions for countries i and j at time t, exhibit nonstationarities individually and that both series are difference-stationary (then they have a unit root). If the difference $y_{it} - y_{jt}$ is stationary, the series are said to converge pair-wise so that they are cointegrated. In a group of N countries, if $y_{it} - y_{jt}$ is stationary for every pair i and j within the group, we have group-wise convergence between the N members. If the stationary difference has nonzero mean, the notion corresponds to conditional convergence, $i.e$ conditional on the country-specific fixed effects. Instead of computing $P_r^N = N!/(N-2)!$ unit root tests, one for each distinct pair of the group, Evans (1998) shows that pair-wise convergence between all members of a group reduces to the single criterion that the difference between the individual series, y_{it}, and the cross-sectional mean of the group at each point in time, $\bar{y}_t = (1/\tilde{N}) \sum_{i=1}^{\tilde{N}} y_{it}$, is stationary. Consequently, if $y_{it} - \bar{y}_t$ is stationary for each member of the panel, all members converge pair-wise. As Evans (1998) indicates, the null of nonconvergence can be interpreted as the unit root null of the following regression:

$$z_{i,t} = \mu_i + \phi_i z_{i,t-1} + \sum_{p=1}^{P} \beta_{i,p} \, \Delta z_{i,t-p} + \varepsilon_{i,t} \tag{2.9}$$

where $i = 1, \ldots, N$ cross-section series that are observed over periods $t = 1, \ldots, T$; $z_{it} = y_{it} - \bar{y}_t$, μ_i are fixed individual effects and ϕ_i is the autoregressive parameter for the PCE differential or convergence coefficient. The summation term of lagged differences is typical in Augmented Dickey Fuller (or ADF) type specifications and it captures the high order serial correlation by choosing the number of lags P which purge ε_{it} of the time correlation. However, the latter term can be dropped if a Philips-Perron (or PP) method is used to account for individual time dependencies. Specification (2.9) allows to test the following hypotheses:

$$H_0 : \ \phi_i = 1 \ \ \forall i \quad \text{vs.} \quad H_1 : \ \phi_i < 1 \ \text{ for some } i$$

Consequently, the null assumes nonstationarity for all the individual series while the alternative hypothesis only requires that some non empty subset of panel/group members converge toward one another. Note that, under the alternative, distributional convergence will not occur, as converging and diverging subsets of countries are allowed to coexist within the same group. Under cross-sectional independence of the stochastic term $\varepsilon_{i,t}$, several standard procedures allow the combining of individual unit root specifications to derive a panel result. The IPS[72] procedure relies on the mean of the empirical t-statistics ($\bar{t} = (1/N) \sum_{i=1}^{N} t_i$) of the autoregressive coefficients from the N individual ADF unit root tests. The alternative method suggested by Maddala and Wu (1999) and Choi (2001), called the Fisher-ADF or Fisher-PP tests, consists in combining the significance levels (p-values noted π_i) of the convergence coefficient on the basis of the Fisher statistic $P_\lambda = -2 \sum_{i=1}^{N} ln\, \pi_i \sim \chi^2_{2N}$. The drawback of all these methods is that they are not strictly valid under cross-sectional dependencies. Sophisticated tests exist to control for correlations in the error term[73] but a simple alternative is proposed by Maddala and Wu (1999) in the context of the Fisher-type test: Dufour and Torres (1996) exploit the Bonferroni inequality to produce a conservative cutoff robust to cross-correlated errors. The idea of this test is to break up the H_0 hypothesis for the panel as a whole into hypotheses at the individual level (H_{0i}) and to reject the global null if and only if any of its components H_{0i} is wrong, *i.e* rejected at the chosen significance level π_i. If we follow this rule, the Bonferroni inequality says that the significance level π for H_0 is given by $\pi \leq \sum_{i=1}^{N} \pi_i$. Dufour and Torres (1996) suggest simply to choose $\pi_i = \pi / N^{(74)}$. The lat-

---

[72] See Im et al. (2003).

[73] See the tests applied in Westerlund and Basher (2008).

[74] A simple intuition about the statistic is given by Pedroni and Yao (2006) as follows. If we are

ter statistic/test will be referred to as the DT-statistic or DT-test. Our empirical analysis applies the latter procedure in the context of the Fisher-PP estimation. Recall that the PP unit root estimation employs equation (2.9) without the summation term and controls for serial correlation by modifying the t-ratio and ϕ_i coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic. Our estimates are computed with the routine implemented in Eviews 5.1 and use a Barlett kernel and a Newey-West (1994) automated bandwidth selection method. Note also that the p-values associated with the individual PP unit root's t-statistics are not available in tabulated format, they need to be estimated by Monte-Carlo simulation. In this paper, we make use of approximated distributions computed by MacKinnon (1996). Remember that the purpose of the Fisher-PP procedure we apply is to provide a global statistic when the evidence, taken individually, is not strong enough. If a sufficiently large number of π_i's lies in the tail of the distribution that supports rejection over nonrejection of the unit root, the signal may be strong enough to reject the null for the whole panel.

Inherent to any pooled estimate based on individual statistics is the risk of rejecting the null when the result is driven by a few influential observations. As recommended in Maddala (1999), our empirical application determines the number of country-specific regressions which obtain marginal p-values for the convergence coefficient inferior to 10% or within the 10%-20% probability range.

Finally, for some groupings of countries, testing equation (2.9) under the restriction that $\mu_i = 0$ or even that $\mu_i = 0 \cap \phi_i = \phi$ for all i may be a reasonable hypothesis. For the former case, the Fisher-PP procedure is readily applicable while for the pooled assumption, the Levin et al. (2002) test[75] (LLC herefater) is an adequate estimation method. Recall that the modified t-statistic ($\tilde{t}$) from the LLC test is asymptotically standard normal. We also provide the half/double life based on the pooled convergence coefficient, *i.e.* $log(0.5)/log(\hat{\phi})$.

---

interested in knowing if we should reject H_0 for a panel as a whole at an arbitrary level π, let's say 10%, it would not be appropriate to do so because a single member of a panel of N individuals produces a p-value inferior to 10%. Indeed, without even knowing if H_0 is wrong for any of the individual series, we would expect to reject at the 10% level $0.1N$ members. Then, without any compelling reason not to do so, it would be reasonable to reject H_0 if a single individual display an empirical p-value below $\pi_i = \pi/N$, which corresponds to the criterion suggested by Dufour and Torres (1996).

[75] See the reference book of Eviews 5.1 for a detailed applied explanation.

2.4.2 Econometric results

This section presents and discusses the results of the formal convergence tests described in the previous section. Given that the Fisher-PP procedure requires the check of several criteria to evaluate the convergence assumption over the whole panel, the empirical analysis can be cumbersome. To avoid ambiguities, we adopt a conservative strategy in interpreting the results. The nonstationarity of PCE gaps within a specific subgroup of countries is not accepted *if and only if* the following two conditions are satisfied:

C.1: both Fisher-PP test and DT bound reject the presence of a unit root at the 10% level;

C.2: a sufficiently large number of members in the panel display stationary or vanishing differences individually[76].

Our first step consists in checking if our panel approach leads to PCE stochastic nonconvergence worldwide, as found in Aldy (2006). Then, we explore the convergence-clubs phenomenon by applying the unit root tests to income, geographic, RIA and initial PCE groupings of countries over the whole period 1960-2002 as well for the two pre- and post-oil shocks sub-periods, *i.e.* 60-73, 60-79 and 74-02, 80-02 respectively. We first interpret the less restrictive version of the unit root test and consider the Fisher-PP approach with fixed effects. Then, we turn to the absolute convergence version of equation (2.9), which assumes $\mu_i = 0$. Finally, we proceed with the results of the pooled estimates, *i.e.* the LLC procedure, when the group can reasonably be considered as homogeneous.

Worldwide convergence. Table 2.8 summarizes the results for different World samples (CDIAC166, ALDY2006 and CDIAC130). In the Fisher PP fixed effect model, nonconvergence is overwhelmingly rejected for all world samples over the whole period 1960-2002. At least 20% (37 countries out of 166) of the individual series reject the unit root hypothesis at the 10% level. However, noticeable differences exist between the pre and post-oil

---

[76] Obviously, we could tighten condition C.2 and fix arbitrarily a proportion of countries which should reject the null of the individual unit root tests at the level of 20% (recall that Tables 2.8 to 2.13 report the number of individual unit root tests which reject nonconvergence for levels within the 0-10% and 10-20% ranges). Retaining the proportion of, say, one out of five would imply that, once condition C.1 is satisfied, we still require rejection of the null for at least 33 countries of the largest world panel of 166 countries, while rejection for only one country is enough within the smaller (geographic) subgroup of five countries considered in our paper (South Asia). Moreover, condition C.2 is set only to detect if condition C.1 is met because of a few influential individual p-values. Therefore, no additional cutoff for C.2. is necessary.

shocks periods. While nonconvergence with fixed effects worldwide is accepted before the oil-shocks (periods 60-73 and 60-79) in most world samples, the contrary holds in all cases for the period 1980-2002. Note also that the DT bound may be oversensitive to extreme observations even when nonconvergence is rejected for only a few series. When we turn to the unit root results without fixed effects, we find evidence of a vanishing PCE gap for a significant set of countries, but only for the post oil-shocks periods (during 74-02 and 80-02 for the CDIAC166 sample, 74-02 for ALDY2006 and 80-02 for CDIAC130). Between 9 and 30 individual series reject the unit root assumption at the 10% level after the oil-shocks period in the world sample under investigation. No interpretation is made for the pooled estimates (LLC tests) as the CDIAC166 panel has been shown to be strongly heterogeneous with the distributional spatial poolability tests in section 2.3.3.

Overall, the world convergence picture is of conditional type for the period 1960-2002, with significant absolute convergence detected after the oil-shocks for smaller number of countries. These results support the distributional analysis which pointed to a stabilization of the flattening cross-sectional PCE distributions for the world during the oils-shocks period.

Convergence within initial polluters. The convergence results for the initial polluters' groupings are shown in Table 2.9. According to condition C.1, no conditional or absolute convergence is found over the whole period within the initial polluters subgroups. For the most intensive initial polluters, conditional as well as absolute convergence takes place after both oil-shocks periods (74-02 and 80-02). The same result holds for the lowest initial polluters but only over the years 1974-2002. If we further assume homogeneous behaviors in the latter two subgroups during the post oil-shocks periods, the LLC estimates suggest that the intra-group pollution gap is cut by half in approximately 20 years. By contrast, no convergence is found for the middle-range initial polluters. Recall that these two subgroups depict increasing dispersion (IQR indicator) over time in the distributional analysis.

Globally, these results confirm part of Nguyen Van (2005) and Stegman (2005) findings on distributional convergence. We also find convergence in the upper tail of the initial cross-sectional PCE distribution but only after the oil-shocks. By contrast, we report pair-wise convergence in the lowest range of the initial cross-sectional PCE distribution as well.

Table 2.8: World samples' unit root tests. Period 1960-2002.

					Fisher PP's test$^{(b)}$							
		LLC test$^{(a)}$			No fixed effects				With fixed effects			
Sample	period	ϕ	$\tilde{t}$-stat	life	$\sharp$10%	$\sharp$20%	χ^2_{emp}	DT-test	$\sharp$10%	$\sharp$20%	χ^2_{emp}	DT-test
World CDIAC166 (166)	60-02	1.002	1.548	418	13	13	250.1	0.0027	37	54	542.0***	0.0000***
	60-73	1.050	23.24	14	8	12	162.9	0.0003**	6	2	135.6	0.0001**
	60-79	1.012	6.205	59	9	7	164.3	0.0096	4	4	307.2	0.0002**
	74-02	0.992***	-7.561	-83	30	15	522.0***	0.0000***	23	9	475.1***	0.0000***
	80-02	0.992***	-5.945	-90	23	14	456.3***	0.0000***	48	21	824.7***	0.0000***
ALDY2006 (88)	60-02	1.004	2.487	189	6	5	134.9	0.0001***	39	7	426.1***	0.0001***
	60-73	1.025	7.060	28	10	2	138.1	0.0000***	11	8	174.0	0.0196
	60-79	1.013	4.634	55	7	2	109.4	0.0019	44	13	407.9***	0.0003**
	74-02	0.998	-1.586	-383	12	9	223.0***	0.0000***	11	10	205.6*	0.0033
	80-02	1.006	5.413	122	9	7	170.6	0.0023	18	8	236.2***	0.0000***
World CDIAC130 (130)	60-02	1.009	14.343	81	4	5	165.7	0.0000***	45	14	462.2***	0.0003**
	60-73	1.039	44.856	18	2	1	61.9	0.0066	10	5	150.0	0.0000***
	60-79	1.025	26.188	28	2	1	72.1	0.0033	14	8	267.9	0.0000***
	74-02	1.005	7.561	140	19	7	269.4	0.0002**	7	6	229.0	0.0007**
	80-02	1.003	4.737	218	24	7	308.1**	0.0002**	17	7	303.6**	0.0000***

Note: In parenthesis are the number of countries. $^{(a)}$: ϕ, the pooled conv. coeff., is centered in 1. $\tilde{t} \sim N(0,1)$ and 'life´ is the double(+) and half(-) life based on ϕ. $^{(b)}$: Maddala & Wu (1999)'s Fisher test makes use of marginal probabilities based on Philips-Perron (1988)'s individual unit roots regressions and McKinnon (1996)'s p-values. The empirical statistic χ^2_{emp} follows a $\chi^2_{2\tilde{N}}$, where $\tilde{N}$ is the number of cross-sectional units in the investigated sample. $\sharp$10% and $\sharp$20% represent the number of countries with p-values within the 0-10% and $\in$ 10% -20% range respectively. DT-test refers to the lower bound suggested by Dufour and Torres (1996) and based on Bonferroni's inequality.

Table 2.9: Initial polluters' unit root tests. Period 1960-2002.

					Fisher PP's test$^{(b)}$							
		LLC test$^{(a)}$			No fixed effects				With fixed effects			
Sample	period	ϕ	$\tilde{t}$-stat	life	♯10%	♯20%	χ^2_{emp}	DT-test	♯10%	♯20%	χ^2_{emp}	DT-test
PCE $>$ 0.65 (42)	60-02	0.990***	-2.682	-66	5	1	90.5	0.0079	4	10	117.5***	0.0014
	60-73	1.052	7.429	528	3	0	46.5	0.0005**	2	1	30.4	0.0640
	60-79	1.001	0.589	510	4	0	68.5	0.0043	3	2	95.1	0.0118
	74-02	0.961***	-8.693	-17	17	13	218.8***	0.0000***	11	9	273.9***	0.0000***
	80-02	0.968***	-5.754	-21	9	4	146.2***	0.0001**	18	5	273.9***	0.0000***
PCE 0.17-0.64 (41)	60-02	1.016	4.864	43	5	1	64.1	0.0116	4	4	94.2	0.0102
	60-73	1.045	4.153	16	3	6	50.6	0.0330	1	2	41.6	0.0904
	60-79	1.033	4.960	21	4	2	52.8	0.0230	4	3	70.8	0.0007**
	73-02	1.013	3.635	55	4	4	70.9	0.0099	6	3	85.8	0.0145
	80-02	1.007	1.709	106	6	3	86.0***	0.0018	5	2	76.6	0.0049
PCE 0.05-0.16 (42)	60-02	1.020	4.864	35	2	1	42.2	0.0398	3	7	85.6	0.0678
	60-73	1.080	6.073	9	0	2	35.6	0.1287	0	1	21.3	0.1008
	60-79	1.042	5.757	17	1	4	45.0	0.0986	1	1	41.6	0.0436
	74-02	1.019	7.246	36	3	0	47.7	0.0268	6	4	76.8	0.0153
	80-02	1.016	5.951	43	3	5	53.8	0.0401	9	2	96.1	0.0036
PCE $<$ 0.05 (41)	60-02	0.967***	-3.592	-21	0	7	97.3	0.1024	2	33	142.9***	0.0615
	60-73	0.939***	-2.197	-11	0	0	38.2	0.1024	0	0	19.9	0.5909
	60-79	0.920	-5.084	-8	0	2	77.9	0.1203	0	0	64.0	0.3095
	74-02	0.950***	-10.83	-14	31	8	233.0***	0.0020*	33	3	246.7***	0.0000***
	80-02	0.996	-0.618	-194	2	0	46.6	0.0131	2	12	107.5**	0.0448

Note: PCE stands for per capita CO_2 emissions. In parenthesis are the number of countries. $^{(a)}$: ϕ, the pooled conv. coeff., is centered in 1. $\tilde{t} \sim N(0,1)$ and 'life´ is the double(+) and half(-) life based on ϕ. $^{(b)}$: Maddala & Wu (1999)'s Fisher test makes use of marginal probabilities based on Philipps-Perron (1988)'s individual unit roots regressions and McKinnon (1996)'s p-values. The empirical statistic χ^2_{emp} follows a $\chi^2_{2\tilde{N}}$, where $\tilde{N}$ is the number of cross-sectional units in the investigated sample. ♯10% and ♯20% represent the number of countries with p-values within the 0-10% and $\in$ 10% -20% range respectively. DT-test refers to the lower bound suggested by Dufour and Torres (1996) and based on Bonferroni's inequality.

Convergence within income groups. Regarding the income groups, Table 2.10 shows that high income countries behave like the most intensive initial polluters; *i.e.* conditional as well as absolute convergence occurs after the 70s oil-shocks. For the upper-middle income class, we reject conditional convergence for all periods but accept absolute convergence from 1980 to 2002. Regarding the lower-middle income economies, significant conditional convergence is at work from 1960 to 2002 as well as from 1960 to 1979. However, for the latter period, very few countries are indeed converging (only 2/38 and 4/38 reject the individual null at the 10% level or within the >10%-20% probability range respectively). We also reject absolute nonconvergence for the years 1980 to 2002 for that income class. Finally, the poorest countries display significant absolute convergence, essentially after the oil-shocks. The main result that emerges when countries achieve similar per capita levels of income by 2000 is that a non negligible subset of them also experience pair-wise convergence in their per capita level of CO_2 emissions. However, this is true only for the post oil-shocks period as nonconvergence holds before as well as from 1960 to 2002 for most income groups.

Convergence within geographic areas. To keep Table 2.11 tractable, the results for periods 60-79 and 80-02 are omitted but this does not alter our conclusions. First, two areas depict clearly conditional pair-wise convergence over the whole period, Europe & Central Asia (ECA) and Sub-Saharan Africa (SSA), and one zone, East Asia and Pacific (EAP), shows weakly significant stable difference, but only for very few countries. Indeed, for the latter two regions, conditional convergence is stronger after the first oil-shock. Second, absolute convergence also holds within the ECA region for the whole as well as for the post oil-shock period. We also find absolute convergence across most of the other geographic groups since 1974, with the exception of the South Asia group that does not converge at all. These results are in accordance with the stabilized dispersion detected in the distributional graphs for most of the geographic subgroups after the 70s. We conclude that geographic proximity favor pair-wise convergence in per capita CO_2 emissions but mainly after the oil-shocks period.

Convergence within RIA members. For the RIA groupings, we compute the unit root regressions for the whole period 1960-2002 as well as for the subperiod since the integration area was created.

Table 2.10: Income groups' unit root tests. Period 1960-2002.

					Fisher PP's test$^{(b)}$							
		LLC test$^{(a)}$			No fixed effects				With fixed effects			
Sample	period	ϕ	$\tilde{t}$-stat	life	$\sharp$ 10%	$\sharp$ 20%	χ^2_{emp}	DT-test	$\sharp$ 10%	$\sharp$ 20%	χ^2_{emp}	DT-test
HI (48)	60-02	0.987***	-3.767	-51	6	5	114.9*	0.0046	3	10	122.8**	0.0050
	60-73	1.050	7.321	14	3	2	60.6	0.0001***	2	1	31.1	0.0726
	60-79	0.994	-1.149	-110	3	3	85.3	0.0047	1	3	89.8	0.0081
	74-02	0.965***	-8.849	-19	22	5	243.9***	0.0001***	9	8	168.7.3***	0.0002***
	80-02	0.969***	-6.263	-22	15	6	208.8***	0.0000***	22	6	337.3***	0.0000***
UMI (24)	60-02	1.001	0.309	673	2	2	45.2	0.0072	2	6	60.1	0.0194
	60-73	1.044	3.928	24	1	2	27.4	0.0032*	1	0	20.7	0.0477
	60-79	1.026	4.608	27	2	1	30.1	0.0036*	0	1	25.7	0.1491
	74-02	0.996	-1.139	-168	2	4	48.1	0.0689	3	1	51.6	0.0038*
	80-02	0.978***	-4.705	-32	7	3	86.4***	0.0006**	4	1	52.7	0.0031*
LMI (38)	60-02	1.001	0.712	495	1	3	49.9	0.0888	7	3	99.2**	0.0005**
	60-73	1.044	16.267	16	1	1	19.3	0.0592	1	2	35.2	0.0026*
	60-79	1.036	13.774	20	2	1	27.8	0.0482	2	4	99.6**	0.0000***
	74-02	0.995**	-2.250	-141	1	6	73.5	0.0709	5	4	83.1	0.0047
	80-02	0.991***	-3.700	-77	5	9	108.0***	0.0021***	7	3	81.7	0.0011**
LI(49)	60-02	1.003	1.484	231	6	4	118.7*	0.0005**	7	8	143.7***	0.0022
	60-73	1.075	10.82	10	3	2	43.0	0.0319	2	5	70.8	0.0000***
	60-79	1.037	9.551	19	4	2	60.7	0.0098	4	3	71.6	0.0002***
	74-02	0.999	-0.667	-529	8	4	131.9**	0.0003**	5	3	99.9	0.0010**
	80-02	0.997*	-1.519	-201	8	3	120.5*	0.0018*	6	1	106.4	0.0000***

Note: In parenthesis are the number of countries. Income groups based on WDI (2004), *i.e.* Gross National Income per capita in 2002 calculated with the World Bank Atlas methodology with Low (<735$), Lower Middle (736$ -2935$), Upper Middle (2936$ -9075$) and High (>9075$) per capita income. $^{(a)}$: ϕ, the pooled conv. coeff., is centered in 1. $\tilde{t} \sim N(0,1)$ and 'life´ is the double(+) and half(-) life based on ϕ. $^{(b)}$: Maddala & Wu (1999)'s Fisher test makes use of marginal probabilities based on Philipps-Perron (1988)'s individual unit roots regressions and McKinnon (1996)'s p-values. The empirical statistic χ^2_{emp} follows a $\chi^2_{2\tilde{N}}$, where $\tilde{N}$ is the number of cross-sectional units in the investigated sample. $\sharp$ 10% and $\sharp$ 20% represent the number of countries with p-values within the 0-10% and $\in$ 10% -20% range respectively. DT-test refers to the lower bound suggested by Dufour and Torres (1996) and based on Bonferroni's inequality.

Table 2.11: Geographic groups' unit root tests. Period 1960-2002.

					Fisher PP's test$^{(b)}$							
		LLC test$^{(a)}$			No fixed effects				With fixed effects			
Sample	period	ϕ	$\tilde{t}$-stat	life	$\sharp$10%	$\sharp$20%	χ^2_{emp}	DT-test	$\sharp$10%	$\sharp$20%	χ^2_{emp}	DT-test
ECA (30)	60-02	0.988***	-3.824	-57	7	4	119.9***	0.0000***	5	2	82.0**	0.0000***
	60-73	1.010	2.970	67	3	5	69.9	0.0000***	6	4	88.2**	0.0078
	74-02	0.977***	-5.311	-29	8	4	130.6***	0.0000**	5	1	86.5**	0.0000***
EAP (29)	60-02	1.001	0.233	625	1	2	41.2	0.0467	2	1	74.5*	0.0030*
	60-73	1.057	3.668	12	0	2	19.0	0.1219	1	1	21.8	0.0282
	74-02	0.990	-2.656	-67	6	2	95.4***	0.0015**	7	5	112.8***	0.0000***
SA (5)	60-02	1.036	7.041	19	0	2	7.8	0.1081	0	0	2.5	0.5050
	60-73	1.013	1.194	55	0	0	5.3	0.3753	1	2	15.2	0.0957
	74-02	1.040	4.715	18	0	0	5.4	0.2238	0	0	4.2	0.3561
MENA (19)	60-02	0.991*	-1.393	-75	1	1	31.2	0.0403	12	2	99.6***	0.0094
	60-73	1.022	1.099	32	0	2	18.2	0.0837	1	1	37.8	0.0568
	74-02	0.985***	-3.114	-46	4	1	75.4***	0.0000***	3	2	61.1**	0.0268
SSA (37)	60-02	1.002	0.596	383	5	1	89.6	0.0001***	6	5	127.1***	0.0002***
	60-73	1.074	11.294	10	3	2	32.9	0.0039	1	2	21.1	0.0856
	74-02	0.989***	-3.322	-61	18	7	190.0***	0.0000***	10	6	127.6***	0.0017*
LAC (36)	60-02	0.992**	-1.761	-83	2	4	67.6	0.3433	1	1	89.3*	0.0331
	60-73	1.066	7.216	11	1	1	24.2	0.0104	0	2	17.5	0.1122
	74-02	0.967***	-5.894	-20	6	8	117.5***	0.0000***	3	0	86.5	0.0220

Note: In parenthesis are the number of countries. Geographic groupings made according to criteria from WDI (2004), *i.e.* Europe and Central Africa (ECA), East Asia and Pacific (EAP), Latin America and Caribbean (LAC), Middle-East and North Africa (MENA), North America (NA), South Asia (SA) and Sub-Saharan Africa (SSA). $^{(a)}$: ϕ, the pooled conv. coeff., is centered in 1. $\tilde{t} \sim N(0,1)$ and ‘life´ is the double(+) and half(-) life based on ϕ. $^{(b)}$: Maddala & Wu (1999)'s Fisher test makes use of marginal probabilities based on Philipps-Perron (1988)'s individual unit roots regressions and McKinnon (1996)'s p-values. The empirical statistic χ^2_{emp} follows a $\chi^2_{2\tilde{N}}$, where $\tilde{N}$ is the number of cross-sectional units in the investigated sample. $\sharp$10% and $\sharp$20% represent the number of countries with p-values within the 0-10% and $\in$ 10% -20% range respectively. DT-test refers to the lower bound suggested by Dufour and Torres (1996) and based on Bonferroni's inequality.

The goal here is to check if economic integration enhances the convergence process of PCE between countries. We present the results in Tables 2.12 and 2.13 according to the South-South, North-South, North-North dimension of the RIA. It is very clear in Table 2.12 that the Economic Community of Eastern and Southern Africa (ECOWAS) display significant convergence, both conditional and absolute. However, the convergence process is not spectacularly boosted after the creation of the integration zone in 1975. Moreover the latter period coincides roughly with the first oil-shock, making both effects difficult to separate. Conditional convergence is also detected within the Asean Free Trade Association (AFTA5) and the Latin American Integration Association (LAIA) after the RIA takes place. However, PCE levels are not converging for the remaining five South-South RIA.

For the North-South integration area APEC12, conditional pair-wise convergence holds for a large number of countries since the RIA creation and some countries converge absolutely as well. For NAFTA, the results reflect the fact that Canada and USA have stable PCE differences since 1994 while Mexico diverges. Finally, the EU and OECD entities correspond to the only two North-North RIA retained. We reject conditional convergence within these areas but accept overwhelmingly absolute convergence for almost all the periods investigated. Economic integration, measured by membership to some integration agreement, seems to favor stochastic convergence *à la* Evans in PCE levels within the members of North-South as well as North-North partnerships. The results for South-South RIA point rather toward divergence, with some cases of conditional convergence.

Globally, the unit root tests outline the major role played by the oil-shocks in breaking the worldwide increasing gaps in carbon dioxide PCE. For most of the groupings of countries, we find that the 60s is a period of strong nonconvergence between groups' members. Second and compared to the previous literature, we highlight the fact that a significant portion of the most and the least intensive initial polluters exhibit convergence mainly after the oil-shock. Third, the countries which reach the highest income per capita levels tend to converge pair-wise toward similar PCE levels. However, this does not apply for countries converging to lower income levels. Fourth, the only geographical area that depicts convergence over the 1960-2002 and the post oil-shocks periods is the ECA zone. Finally, joining an integrated economic area may favor pair-wise convergence in PCE levels, but mainly for North-South and North-North countries.

Table 2.12: South-South RIA's unit root tests. Period 1960-2002.

					Fisher PP's test$^{(b)}$							
		LLC test$^{(a)}$			No fixed effects				With fixed effects			
Sample	period	ϕ	$\tilde{t}$-stat	life	♯10%	♯20%	χ^2_{emp}	DT-test	♯10%	♯20%	χ^2_{emp}	DT-test
CEMAC5	60-02	0.975^{*}	-1.365	-28	0	0	10.1	0.2272	0	0	9.5	0.3703
	69-02	0.971^{*}	-1.445	-24	0	0	9.2	0.2914	0	0	9.3	0.3494
COMESA12	60-02	1.044	5.862	16	0	0	6.8	0.2522	2	0	17.1	0.0506
	81-02	1.073	7.743	10	0	1	6.9	0.1634	2	0	15.1	0.0335
ECOWAS (16)	60-02	0.985^{*}	-1.631	-47	4	1	53.2^{**}	0.0085	4	4	74.1^{***}	0.0004^{***}
	75-02	0.968^{***}	-2.914	-21	5	1	57.2^{***}	0.0064	6	2	67.0^{***}	0.0060^{*}
AFTA5	60-02	1.008	0.723	84	0	0	5.2	0.2157	0	0	7.2	0.2679
	69-02	1.016	1.391	44	0	0	3.9	0.3837	2	1	33.3^{***}	0.0010^{***}
CARICOM10	60-02	1.039	3.762	18	1	1	14.0	0.0354	0	1	10.3	0.1245
	73-02	1.037	3.833	19	1	1	11.8	0.0461	0	0	5.3	0.2449
ACN4	60-02	0.957^{**}	-1.945	-16	0	1	10.7	0.1842	0	0	6.5	0.2045
	69-02	0.954^{*}	-1.645	-15	0	2	11.7	0.1341	1	0	11.3	0.0402
LAIA (12)	60-02	0.995	-1.193	-127	0	2	21.0	0.1193	0	2	21.6	0.1341
	80-02	0.988^{**}	-1.843	-56	1	0	24.6	0.0692	3	0	40.0^{**}	0.0004^{***}
MERCOSUR (4)	60-02	0.999	-0.156	-700	0	1	7.1	0.1719	0	1	6.7	0.1069
	91-02	0.996	-0.351	-170	0	1	6.6	0.1522	1	0	7.9	0.0599

Note: In parenthesis are the number of countries. $^{(a)}$: ϕ, the pooled conv. coeff., is centered in 1. $\tilde{t} \sim N(0,1)$ and 'life´ is the double(+) and half(-) life based on ϕ. $^{(b)}$: Maddala & Wu (1999)'s Fisher test makes use of marginal probabilities based on Philipps-Perron (1988)'s individual unit roots regressions and McKinnon (1996)'s p-values. The empirical statistic χ^2_{emp} follows a $\chi^2_{2\tilde{N}}$, where $\tilde{N}$ is the number of cross-sectional units in the investigated sample. ♯10% and ♯20% represent the number of countries with p-values within the 0-10% and ∈ 10% -20% range respectively. DT-test refers to the lower bound suggested by Dufour and Torres (1996) and based on Bonferroni's inequality.

Table 2.13: North-South and North-North RIA's unit root tests. Period 1960-2002.

					Fisher PP's test$^{(b)}$								
		LLC test$^{(a)}$			No fixed effects					With fixed effects			
Sample	period	ϕ	$\tilde{t}$-stat	life	♯10%	♯20%	χ^2_{emp}	DT-test		♯10%	♯20%	χ^2_{emp}	DT-test
							North-South RIAs						
APEC12	60-02	0.989*	-1.436	-62	1	1	25.7	0.0670		2	1	34.0*	0.0703
	89-02	1.004	0.982	158	4	0	45.9***	0.0001***		7	2	61.9***	0.0051*
NAFTA (3)	60-02	1.001	0.193	722	0	0	1.9	0.6371		0	1	7.8	0.1814
	94-02	1.005	0.523	141	0	0	1.3	0.6006		2	0	12.7**	0.0361
							North-North RIAs						
EU6	60-02	0.975***	-3.977	-27	3	2	38.9***	0.0010***	0	0	7.3	0.2387	
	73-02	0.962***	-3.761	-18	4	2	52.5***	0.0000***		2	3	27.4***	0.0232
EU15	60-02	0.979***	-4.422	-33	5	6	82.5***	0.0001***		3	0	28.0	0.0087
	73-02	0.959***	-5.533	-16	8	3	123.8***	0.0000**		3	4	43.8**	0.0197
OECD (24)	60-02	0.990***	-3.184	-72	7	4	93.4***	0.0002***		3	1	44.3	0.0140
	60-73	1.010	0.999	69	4	1	55.7	0.0015**		4	3	51.5	0.0099
	60-79	0.998	-0.616	-330	6	0	71.0**	0.0012**		3	2	60.2	0.0020**
	74-02	0.982***	-4.178	-38	10	4	134.9***	0.0000**		4	2	59.6	0.0025**
	80-02	0.986***	-2.685	-50	9	4	108.8***	0.0017**		6	4	90.9***	0.0012**

Note: In parenthesis are the number of countries. (a): ϕ, the pooled conv. coeff., is centered in 1. $\tilde{t} \sim N(0,1)$ and 'life´ is the double(+) and half(-) life based on ϕ. (b): Maddala & Wu (1999)'s Fisher test makes use of marginal probabilities based on Philipps-Perron (1988)'s individual unit roots regressions and McKinnon (1996)'s p-values. The empirical statistic χ^2_{emp} follows a $\chi^2_{2\tilde{N}}$, where $\tilde{N}$ is the number of cross-sectional units in the investigated sample. ♯10% and ♯20% represent the number of countries with p-values within the 0-10% and $\in$ 10% -20% range respectively. DT-test refers to the lower bound suggested by Dufour and Torres (1996) and based on Bonferroni's inequality.

2.5 Conclusion

In this paper we investigate the PCE convergence hypothesis for a large world panel of countries over the period 1960-2002 with two main tools: a robust descriptive analysis of the cross-sectional PCE distributions and unit root tests linked to the notion of pair-wise convergence. From both perspectives, the oil-shocks seem to play a major role in stabilizing the diverging dynamics or the within-groups pollution gaps. The first approach depicts mainly flattening cross-sectional distributions during the pre 70s oil-shocks and more stable ones afterwards. No significant polarization phenomenon is detected worldwide or across the investigated country groupings when multimodality is formally checked. We also report distributional heterogeneity in the world panel as some country subgroups depart significantly from the world picture. Among them, the OECD and EU15 economies are the only ones to exhibit strong decreasing PCE gaps over time with increasing distributional peakedness for the former group. Therefore, there is no strong evidence of distributional convergence, neither at the world level, nor within most of the investigated subsets.

By contrasts, unit root tests, combined with the pair-wise convergence criterion of Evans (1998), identify clearly the existence of (conditionally) converging economies for the world panel as well as within selected groupings of countries, *i.e* most/least intensive initial polluters, income and geographic subgroups, North-South or North-North integrated areas. Despite the fact that a rather low proportion (10%-20%) of 'significantly' converging economies is enough to reject nonconvergence for a panel as a whole, these results stress that decreasing per capita pollution levels of CO_2 emissions exist within a variety of country groupings formed on the basis of simple criteria. Moreover, as pollution convergence within most subgroups occurs mainly after the 70s, it is probable that the oil-shocks have played an important role.

While the distributional and time series approaches seem to drive to conflicting results, it is important to note that distributional nonconvergence or stabilized cross-sectional distributions may be fully compatible with stochastic convergence *à la Evans*. In the latter analysis, we took the option of testing the null of nonconvergence against the alternative of existence of some converging economies. When converging and nonconverging economies are present in a (sub)panel, the distributional analysis may result in increasing/stabilized spread without significant departures from unimodality, while a relatively small number of countries converging pair-wise may be enough to reject the null of panel nonconvergence. Indeed both the distributional and stochastic convergence points toward the role played by the 70s oils shocks in stabilizing or breaking the increasing pollution gap between countries that is clearly prevalent during the sixties.

Overall, given that CO_2 emissions have not been penalized by stringent policy measures during the time span covered by the panel, these results indicate that, despite the structural differences between countries, technical progress and price mechanisms favor a more efficient use of fossil fuels that curve down the scale effect. Obviously, this does not mean that the stabilized (or 'steady state') per capita emissions levels reached by some countries are optimal from an environmental point of view. Moreover, the existence of converging economies make reasonable per capita emissions' targets to emerge for groups of economies and may help to find acceptable compromises for the the post-Kyoto agreements. The absence of distributional convergence or significant multiple modes in the cross-sectional distributions rules out any significant polarization phenomenon at the world level.

To our knowledge, this paper is the first to explore the convergence club phenomenon for per capita carbon dioxide emissions with a systematic grouping of countries. Many technical improvements or extensions could be considered. While the geographic/income criteria we employ have the advantage of being clear and simple, our measure of economic integration is rather crude as it is based on formal agreements. Using an intensity measures of economic integration, such as bilateral trade as in Ben-David (2000), could be an interesting and straightforward extension. Secondly, our distributional analysis would benefit from being explored with more recent tools, with better finite-sample properties, developed in nonparametric econometrics, such as the (conditional) distributions for mixed data of Li and Racine (2003) and Hall et al. (2004) along the lines followed by Quah (1997) for income convergence. Note also that we applied the very standard version of the multimodality test by Hartigan and Hartigan (1985); a better calibration of it could be investigated to detect smaller departure from unimodality. Clearly, alternative and more direct disparity indicators could provide a clearer picture of distributional convergence, such as simple Gini index to more sophisticated indicators such as the entropy measures of Maasoumi et al. (2007) or the polarization measures used in Ezcurra (2007). Finally, unit root tests setting as the null convergence instead of nonconvergence could help in reconciling the distributional analysis with the stochastic con the time series approach. Allowing endogeneous breaks in the unit root tests also constitute a straightforward complementary application.

Acknowledgments

We thank the Swiss National Science Foundation for financial support under grant No100012-109926. This paper has been accepted for presentation at the SURED 2008 conference, in Ascona (Switzerland). Preliminary results have also benefited from comments of the participants to the Conference on Growth and Human Capital, May 2007, Guelph, Canada, where I held a poster session.

2.6 Appendix

This Appendix provides a full description of the database used in the empirical sections. Carbon emissions and population data come from respectively the US department of Energy (Marland et al. 2006) and the US Census Bureau 2006. The income and geographical groupings of countries are based on the CDrom World Development Indicators 2004, WDI2004 henceforth. Most of the tables and figures presented in this section possess self-explanatory footnotes.

Table 2.14 shows all the reconstructed individual series of CO_2 per capita emissions. The reconstruction consists simply in merging series of former sovereign entities that split off or (re)unified at some point during the 1960-2002 period.

Table 2.15 presents the countries included in the CDIAC166 samples, sorted in decreasing order according to their PCE level in 1960. Note that the countries in bold correspond to those included in Aldy (2006). We underlined countries which doubled its CO_2 per capita emissions for at one year over the period 1960-2002. The latter countries/areas represent the so-called outliers. Recall that the CDIAC130 sample correspond to the CDIAC166 sample once the 36 identified outliers are removed. The table also shows for every country the absolute and relative PCE levels in 1960, the mean PCE growth rate over the period and the rank of the PCE level in 2002. Figure 2.7 provide a rough picture of the series included in CDIAC130 as well as the 36 outliers. Note that the Y-axis on these graphs are in logarithmic scale (each unit implies doubling the initial level.).

Tables 2.16 to 2.18 refer to income groupings. Table 2.16 shows the income memberships of the CDIAC166 areas, Table 2.17 explains the matching process between the CDIAC areas and the WDI2004's income categories and Table 2.18 reports sample sizes. The latter table deserves some comments. In the first line, we can see that the low income group contains 49 CDIAC166 areas (37 when outliers are omitted). This figure corresponds to 53 (respectively 41) WDI countries out of the 58 maximal sample size for this category of income and represents 91% (respectively 71%) of the maximal size for this specific Word Bank group. Note that two CDIAC166 areas, USSR and Yugoslavia, contain sovereign states which are classified into different income classes in the WDI list. In order to determine a single income category for USSR and Yugoslavia, we computed the per capita Gross National Income for both former blocks and applied the WDI's criteria. With respectively 1519$ and 2948$, USSR and Yugoslavia appear to be lower-middle income and upper-middle income entities respectively. Relative to Aldy (2006), the increase in sample size is particularly large for the Low Income countries. By contrast, High Income economies are fully represented in Aldy (2006)'s. Overall, our data cover at least 80% of the sub-populations' income groups. Note also that 24 out of the 36 outliers belong to either the low income or the high income non-OECD groups.

Table 2.19 displays the geographic memberships of the CDIAC166 areas and Table 2.20 shows the sample size for the seven geographic groups. The number of countries in each group represents at least 75% of the WDI subpopulation size. Compared to ALDY2006's sample, our data account for significantly more countries from Sub-Saharan Africa, Latin America and Caribbean and East Asia and Pacific. We notice that 20 out of the 36 outliers are located in the Middle East & North Africa and Sub-Saharan Africa regions.

Finally, regarding the RIA groupings reported in Tables 2.21 and 2.22, we tried to find the right balance between the availability of data and the length of the membership of each individual country to the Regional Integration Agreement. No formal and unique criteria has been applied to all RIA groupings. We tried to maximize the final number of countries included in the groups but dropping out the too recent memberships.

Table 2.14: Reconstructed CDIAC areas for the period 1960-2002.

CDIAC area	CO2 emissions 1950-2002	Population 1950-2002
Czechoslovakia	Czechoslovakia (1950-1991)	
	Czech Republic (1992-2002)	Czech Republic (1950-2002)
	Slovakia (1992-2002)	Slovakia (1950-2002)
Bangladesh & Pakistan	East & West Pakistan (1950-1971)	
	Bangladesh (1972-2002)	Bangladesh (1950-2002)
	Pakistan (1972-2002)	Pakistan (1950-2002)
Germany	Federal Republic of Germany (1950-1990)	
	German Democratic Republic (1950-1990)	
	Germany (1991-2002)	Germany (1950-2002)
Malaysia	Peninsular Malaysia (1957-1969)	
	Sabah (1957-1969)	
	Sarawak (1957-1969)	
	Malaysia (1970-2002)	Malaysia (1950-2002)
Netherland Antille & Aruba	Netherland Antilles & Aruba (1950-1985)	
	Netherland Antilles (1986-2002)	Netherland Antilles (1950-2002)
	Aruba (1986-2002)	Aruba (1950-2002)
Rhodesia-Nyasaland	Rhodesia-Nyasaland (1950-1963)	
	Malawi (1964-2002)	Malawi (1950-2002)
	Zambia (1964-2002)	Zambia (1950-2002)
	Zimbabwe (1964-2002)	Zimbabwe (1950-2002)
Rwanda-Urundi	Rwanda-Urundi (1950-1961)	
	Rwanda (1962-2002)	Rwanda (1950-2002)
	Burundi (1962-2002)	Burundi (1950-2002)
USSR	USSR (1950-2002)	
	Armenia (1992-2002)	Armenia (1950-2002)
	Azerbaijan (1992-2002)	Azerbaijan (1950-2002)
	Belarus (1992-2002)	Belarus (1950-2002)
	Estonia (1992-2002)	Estonia (1950-2002)
	Georgia (1992-2002)	Georgia (1950-2002)
	Kazakhstan (1992-2002)	Karakhstan (1950-2002)
	Kirghizstan (1992-2002)	Kizgizstan (1950-2002)
	Latvia (1992-2002)	Latvia (1950-2002)
	Lithuania (1970-2002)	Lithuania (1950-2002)
	Moldavia (1970-2002)	Moldavia (1950-2002)
	Russia (1970-2002)	Russia (1950-2002)
	Tajikistan (1970-2002)	Tajikistan (1950-2002)
	Turkmenistan (1970-2002)	Turkmenistan (1950-2002)
	Ukraine (1970-2002)	Ukraine (1950-2002)
	Uzbekistan (1970-2002)	Uzbekistan (1950-2002)
Vietnam	Republik of South Vietnam (1955-1969)	
	Democratic Republic of Vietnam (1955-1969)	
	Vietnam (1970-2002)	Vietnam (1950-2002)
Yemen	Democratic Yemen (1955-1990)	
	Former Yemen (1955-1990)	
	Yemen (1991-2002)	Yemen (1950-2002)
Yugoslavia	Yugoslavia (1950-1991)	
	Bosnia-Herzegovina (1992-2002)	Bosnia-Herzegovina (1950-2002)
	Croatia (1992-2002)	Croatia (1950-2002)
	Macedonia (1992-2002)	Macedonia (1950-2002)
	Serbia-Montenegro (1970-2002)	Serbia-Montenegro (1950-2002)
	Slovenia (1992-2002)	Slovenia (1950-2002)

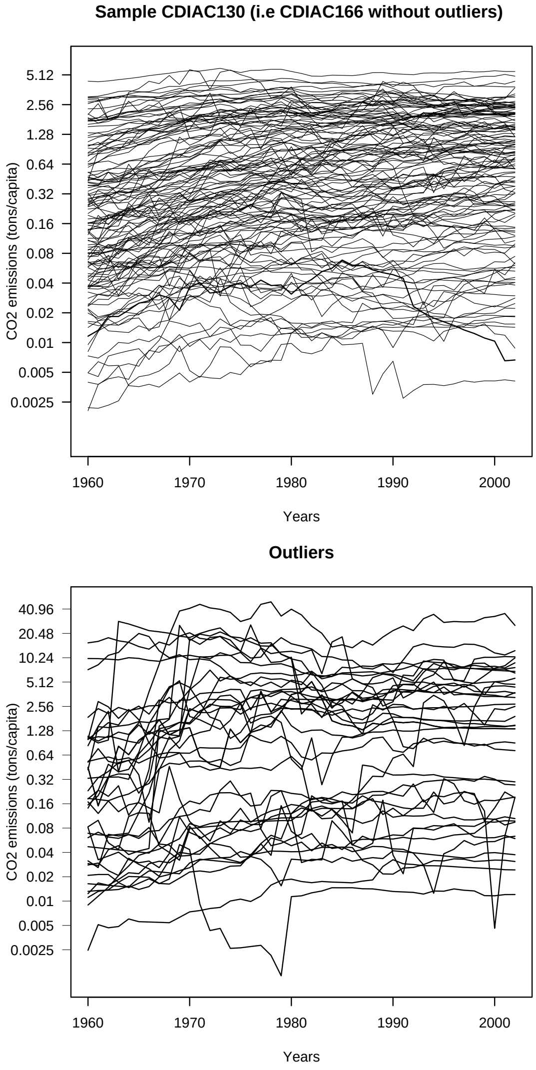
Source: Marland et al. (2006) and US Census Bureau (2006).

Table 2.15: World sample. Initial, final levels and growth rates of per capita CO_2 emissions. Period 1960-2000.

Rank 1960	Country	CO2pc 1960	Ratio to CDIAC 166	Growth rate 60-02	Rank 2002	Rank 1960	Country	CO2pc 1960	Ratio to CDIAC 166	Growth rate 60-02	Rank 2002	Rank 1960	Country	CO2pc 1960	Ratio to CDIAC 166	Growth rate 60-02	Rank 2002	Rank 1960	Country	CO2pc 1960	Ratio to CDIAC 166	Growth rate 60-02	Rank 2002
36	*World CDIAC*	*0.85*	*1.03*	*0.64%*	*70*	41	**Japan**	0.67	0.82	3.17%	27	83	Yemen	0.17	0.21	0.27%	125	125	Tonga	0.05	19.07	4.17%	111
39	*CDIAC166*	*0.82*	*1.00*	*0.62%*	*72*	42	**Argentina**	0.65	0.79	0.90%	77	84	Martinique	0.16	0.20	5.21%	61	126	St. Lucia	0.05	19.07	6.35%	92
1	Neth. Antilles & Arub	15.63	19.07	-1.98%	8	43	Korea (D.R.)	0.62	0.76	2.43%	51	85	**Egypt**	0.16	0.20	2.80%	97	127	Paraguay	0.04	19.07	3.50%	124
2	**Luxembourg**	10.00	12.20	-1.30%	9	44	**Italy & San Marino**	0.59	0.72	2.91%	47	86	**Turkey**	0.16	0.20	3.91%	82	128	Liberia	0.04	19.07	0.04%	148
3	**Kuwait**	7.28	8.88	0.14%	6	45	Gibraltar	0.53	0.65	5.05%	13	87	Guadeloupe	0.16	0.20	4.54%	71	129	Zaire	0.04	19.07	-3.36%	164
4	**USA**	4.42	5.39	0.52%	10	46	**Cuba**	0.53	0.65	0.18%	95	88	**Saudi Arabia**	0.15	0.19	7.72%	17	130	St. Vincent & Grenad	0.04	19.07	5.87%	102
5	**Germany**	3.07	3.74	-0.33%	25	47	**Former Yugoslavia**	0.53	0.64	2.12%	65	89	**Algeria**	0.15	0.19	3.91%	84	131	Guinea	0.04	19.07	0.21%	150
6	**United Kingdom**	3.04	3.71	-0.48%	31	48	Guam	0.52	0.64	6.19%	7	90	Libya	0.14	0.17	6.97%	28	132	**Thailand**	0.04	19.07	8.00%	75
7	New Caledonia	2.98	3.64	-0.51%	35	49	**Chile**	0.48	0.59	1.72%	74	91	Djibouti	0.14	0.17	1.04%	118	133	Bangladesh & Pakista	0.04	19.07	3.07%	128
8	**Canada**	2.88	3.51	1.00%	14	50	**Iran**	0.47	0.58	2.67%	60	92	**Korea (R.)**	0.14	0.17	7.00%	30	134	Samoa	0.04	19.07	4.26%	119
9	**Belgium**	2.72	3.32	-0.27%	32	51	**Uruguay**	0.47	0.57	-0.79%	107	93	Malaysia	0.14	0.17	6.22%	50	135	**Côte d'Ivoire**	0.04	19.07	2.63%	132
10	**Czechoslovakia**	2.61	3.19	0.02%	26	52	American Samoa	0.45	0.55	2.59%	64	94	Fiji	0.13	0.16	2.74%	101	136	**Sudan**	0.04	19.07	1.42%	140
11	**Australia**	2.32	2.83	1.78%	11	53	**Mexico**	0.45	0.54	1.94%	73	95	Belize	0.13	0.16	4.37%	83	137	**Myanmar**	0.03	19.07	0.79%	147
12	**Venezuela**	2.06	2.51	-1.22%	67	54	**Spain**	0.44	0.53	3.69%	44	96	French Polynesia	0.12	0.15	4.26%	87	138	**Angola**	0.03	19.07	4.31%	122
13	St. Pierre & Miquelo	2.03	2.48	0.29%	38	55	Faeroe Islands	0.43	0.53	5.23%	15	97	**Tunisia**	0.11	0.14	4.01%	93	139	**U. Arab Emirates**	0.03	19.07	14.67%	3
14	Iceland	1.89	2.30	0.24%	42	56	**Cyprus**	0.42	0.52	4.09%	36	98	**Ecuador**	0.11	0.13	3.72%	98	140	Papua New Guinea	0.03	19.07	3.64%	129
15	Greenland	1.88	2.29	0.88%	24	57	Palau	0.42	0.51	4.90%	19	99	**Costa Rica**	0.11	0.13	3.19%	103	141	Cape Verde	0.03	19.07	2.96%	135
16	**Poland**	1.84	2.25	0.30%	43	58	Suriname	0.41	0.51	2.90%	63	100	**Nicaragua**	0.10	0.12	1.74%	120	142	Equatorial Guinea	0.02	19.07	3.14%	138
17	**Sweden**	1.79	2.19	-0.30%	54	59	Cayman Islands	0.39	0.48	3.79%	48	101	**Guatemala**	0.09	0.11	2.35%	114	143	Solomon Islands	0.02	19.07	3.28%	136
18	**Denmark**	1.77	2.16	0.72%	34	60	**Lebanon**	0.39	0.47	2.69%	66	102	**Dominican R.**	0.09	0.11	4.87%	91	144	**Nigeria**	0.02	19.07	3.85%	131
19	**Netherlands**	1.75	2.13	0.89%	29	61	**Mongolia**	0.37	0.45	1.94%	81	103	**Honduras**	0.09	0.10	2.44%	115	145	Tanzania	0.02	19.07	0.59%	155
20	**USSR**	1.74	2.12	0.39%	45	62	Albania	0.34	0.41	-1.21%	121	104	Gabon	0.08	0.10	5.17%	88	146	Benin	0.02	19.07	2.94%	139
21	**France & Monaco**	1.62	1.97	0.08%	52	63	**Iraq**	0.33	0.40	2.36%	79	105	Sierra Leone	0.08	0.10	-2.23%	153	147	Haiti	0.02	19.07	2.62%	141
22	**South Africa**	1.53	1.86	0.76%	41	64	**China**	0.33	0.40	1.93%	86	106	**Kenya**	0.08	0.10	-0.61%	142	148	Madagascar	0.02	19.07	1.53%	151
23	**New Zealand**	1.33	1.62	1.36%	37	65	Guyana	0.32	0.38	1.42%	94	107	**Philippines**	0.08	0.10	2.59%	116	149	Comoros	0.02	19.07	1.94%	152
24	**Hungary**	1.24	1.51	0.49%	55	66	**Greece**	0.31	0.38	4.91%	33	108	**Morocco**	0.08	0.10	3.70%	104	150	Central African R.	0.02	19.07	0.29%	159
25	**Austria**	1.19	1.45	1.36%	39	67	**Taiwan**	0.29	0.35	5.39%	23	109	**Bolivia**	0.08	0.10	3.32%	108	151	Uganda	0.02	19.07	0.33%	160
26	Brunei	1.10	1.34	3.50%	12	68	**Colombia**	0.28	0.34	0.72%	105	110	Macau	0.08	0.09	6.49%	69	152	Gambia	0.01	19.07	3.10%	146
27	**Ireland**	1.08	1.31	2.44%	22	69	Malta	0.28	0.34	4.73%	46	111	Mauritius	0.07	0.09	5.39%	89	153	**Cameroon**	0.01	19.07	3.56%	144
28	**Qatar**	1.06	1.30	5.90%	2	70	Rhodesia-Nyasaland	0.27	0.33	-1.87%	130	112	**India**	0.07	0.09	3.48%	109	154	Togo	0.01	19.07	4.85%	137
29	Virgin Islands (US)	1.02	1.24	7.80%	1	71	**Hong Kong**	0.26	0.32	4.02%	62	113	Senegal	0.07	0.08	0.97%	133	155	Afghanistan	0.01	19.07	-1.28%	165
30	**Bahrain**	1.00	1.22	5.20%	5	72	**Portugal**	0.25	0.30	4.47%	53	114	Mozambique	0.07	0.08	-2.60%	158	156	Cambodia	0.01	19.07	0.19%	163
31	Bahamas	1.00	1.22	1.54%	49	73	**Jamaica**	0.25	0.30	3.56%	70	115	Grenada	0.07	0.08	5.65%	90	157	Laos	0.01	19.07	4.39%	143
32	**Norway**	1.00	1.22	3.16%	16	74	**Jordan**	0.24	0.29	3.01%	80	116	**El Salvador**	0.07	0.08	3.33%	112	158	Mauritania	0.01	19.07	8.50%	110
33	**Switzerland**	0.99	1.21	0.98%	56	75	**Panama**	0.24	0.29	2.06%	96	117	Vietnam	0.06	0.08	2.93%	117	159	Guinea Bissau	0.01	19.07	4.64%	145
34	Bermuda	0.97	1.18	1.84%	40	76	**Singapore**	0.23	0.28	6.68%	18	118	**Sri Lanka**	0.06	0.08	1.97%	127	160	Mali	0.01	19.07	1.56%	162
35	**Finland**	0.93	1.13	2.98%	20	77	**Peru**	0.22	0.27	0.34%	113	119	Congo (Rep.)	0.06	0.07	2.71%	123	161	Rwanda-Urundi	0.005	19.07	2.67%	161
36	**Puerto Rico**	0.84	1.02	0.32%	76	78	Barbados	0.20	0.25	4.24%	68	120	Reunion	0.06	0.07	6.55%	78	162	Chad	0.005	19.07	-0.44%	166
37	**Trinidad & Tobago**	0.84	1.02	5.99%	4	79	**Syria**	0.19	0.24	3.29%	85	121	**Indonesia**	0.06	0.07	4.33%	106	163	**Ethiopia**	0.004	19.07	4.36%	156
38	**Israel**	0.82	1.00	3.17%	21	80	French Guiana	0.19	0.23	4.94%	59	122	**Ghana**	0.06	0.07	1.29%	134	164	Burkina Faso	0.002	19.07	5.47%	157
39	**Romania**	0.79	0.97	0.67%	72	81	Antigua & Barbuda	0.18	0.22	5.00%	57	123	Dominica	0.05	0.06	5.36%	100	165	Nepal	0.002	19.07	7.02%	149
40	**Bulgaria**	0.77	0.94	1.54%	58	82	**Brazil**	0.18	0.22	2.30%	99	124	Sao Tome & Princip	0.05	0.06	2.67%	126	166	Niger	0.002	19.07	6.43%	154

Source: Author's own calculations with CO_2 data from Marland et al. (2006) and population series from U.S. Census Bureau (2006). This table reports all areas included in the CDIAC166 sample. In bold are the countries used in Aldy (2006). The underlined countries represent relatively erratic series, *i.e.* countries with relative changes in per capita CO_2 emissions over 100% for at least one year of the 1960-2002 period. * indicates non sovereign areas according to the World countries listed by the World Bank. The grey cells highlight countries whose CO_2 per capita series have been reconstructed according to the available information (see Table 2.14). The sample CDIAC130 corresponds to the sample CDIAC166 once underlined countries are removed.

Figure 2.7: CDIAC130 sample and outliers. Per capita CO_2 emissions. Period 1960-2002.



Source: Marland et al. (2006) and U.S. Census Bureau (2006). Figures on CO_2 take into account emissions originated by fossil fuel consumption. Outliers are countries with relative changes in per capita CO_2 emissions over 100% for at least one year of the 1960-2002 period.

Table 2.16: Income groups of CDIAC countries.

Low Income (LI)		Lower Middle Income (LMI)		Upper Middle Inc. (UMI)	High Income (HI)	
					non-OECD	OECD
Afghanistan	Madagascar	Albania	**Peru**	American Samoa	Anti.&Barb.	**Australia**
Angola	Mali	**Algeria**	**Philippines**	**Argentina**	Bahamas	**Austria**
Bangl.&Pak.	Mauritania	**Bolivia**	**Romania**	Belize	**Bahrain**	**Belgium**
Benin	**Mongolia**	**Brazil**	Samoa	**Chile**	Barbados	**Canada**
Burkina Faso	Mozambique	**Bulgaria**	**South Africa**	**Costa Rica**	Bermuda	**Denmark**
Cambodia	**Myanmar**	Cape Verde	**Sri Lanka**	**Czechosl.**	Brunei	**Finland**
Cameroon	Nepal	**China**	St. Vin. & Grenad.	Dominica	Cayman Isl.	**France& Mon.**
Ctr. Afric. R.	**Nicaragua**	**Colombia**	Suriname	**Former Yugos.**	**Cyprus**	**Germany**
Chad	Niger	**Cuba**	**Syria**	Gabon	Faeroe Isl.	**Greece**
Comoros	**Nigeria**	Djibouti	**Thailand**	Grenada	French Poly.	Iceland
Congo (Rep.)	Pap. N.Gui.	**Domin. Rep.**	Tonga	**Hungary**	Greenland	**Ireland**
Côte d'Ivoire	Rhod-Nyas.	**Ecuador**	**Tunisia**	**Lebanon**	Guam	**Italy & SM.**
Equat. Guinea	Rwan-Uru.	**Egypt**	**Turkey**	Libya	**Hong Kong**	**Japan**
Ethiopia	S. Tome & Prin.	**El Salvador**	**USSR**	Malaysia	**Israel**	**Korea (R.)**
Gambia	Senegal	Fiji		Mauritius	**Kuwait**	**Luxembourg**
Ghana	Sierra Leone	**Guatemala**		**Mexico**	Macau	**Netherlands**
Guinea	Solom. Isl.	Guyana		Palau	Malta	**New Zealand**
Guinea Bissau	**Sudan**	**Honduras**		**Panama**	Nth.Ant.&Aruba	**Norway**
Haiti	Tanzania	**Iran**		**Poland**	New Caledonia	**Portugal**
India	Togo	**Iraq**		**Sau. Arabia**	**Puerto Rico**	**Spain**
Indonesia	Uganda	**Jamaica**		St. Lucia	**Qatar**	**Sweden**
Kenya	Vietnam	**Jordan**		**Tri.&Tobago**	**Singapore**	**Switzerland**
Korea (D.R.)	Yemen	**Morocco**		**Uruguay**	**U.Arab Emir.**	**United Kingdom**
Laos	Zaire	Paraguay		**Venezuela**	Virgin Isl. (US)	**USA**
Liberia						

Note: Classification made according to criteria from World Development Indicator 2004 (WDI), *i.e.* Gross National Income per capita in 2002 . calculated with the World Bank Atlas methodology. The groups are: Low (<735$), Lower Middle (736$ -2935$), Upper Middle (2936$ -9075$) and High (>9075$) per capita income. In bold is the sample used by Aldy (2006). The underlined countries represent outliers, *i.e.* countries with relative changes in CO_2 per capita emissions over 100% for at least one of the years included in the investigated period. In the grey cells are countries which CO_2 and population series have been reconstructed according to the available information in Marland et al. (2006) and US Census Bureau (2006).

Table 2.17: Matching World Bank income groups with CDIAC areas.

		Original WDI income groups					
		LI	LMI	UMI	HInOECD	HIOECD	max.
Adj. groups	LI	- 6 (USSR)					58
	LMI	+ 6 (USSR)	- 3 (former YUG)	+ 3 (USSR)			60
	UMI		+ 3 (former YUG)	- 3 (USSR)	+ 1 (former YUG)		35
	HInOECD				- 1 (former YUG)		31
	HIOECD						24
	max.	64	54	34	32	24	208

Note : LI, LMI, UMI and HI stand for respectively Low (<735$), Lower Middle (736$ -2935$), Upper Middle (2936$ - 9075$)and High (>9075$) per capita income in 2002, see the World Development indicators CDrom for 2004. The 'max.' line reports the original number of countries in each income category. The 'max.' *column* reports the *adjusted* number of countries in each income category. Focusing on an example, the table reads as follows. The LI category includes 64 countries according to World Development Indicator 2004, among which we find 6 former USSR territories. USSR being an LMI entity as a whole, the 6 former Soviet Republic have been moved from the LI (-6) to the LMI (+6) class. Given that the LI class does not require other adjustments of that kind, it ends up with 58 WDI countries in column 'max.'.

Table 2.18: Income groups of countries. Samples' size.

	World samples					ALDY2006 sample	
	CDIAC areas		WDI countries			CDIAC areas	WDI countries
WDI income groups	♯ full	♯ ~~out.~~	max.	♯ full (%)	♯ ~~out.~~ (%)	♯ full	♯ full (%)
Low income	49	37	58	53 (91%)	41 (71%)	13	13 (22%)
Lower middle income	38	35	60	52 (87%)	49 (82%)	28	42 (70%)
Upper middle income	24	18	35	29 (83%)	23 (66%)	14	19 (54%)
High income nonOECD	24	12	31	25 (81%)	12 (42%)	9	9 (29%)
High income OECD	24	23	24	24 (100%)	23 (96%)	23	23 (96%)
Unclassified	7	5	-	-	-	1	-
Total	166	130	208	183 (88%)	148 (71%)	88	106 (51%)

Note: This table matches the CDIAC areas from Marland et al. (2006) with the countries included in the income groups from the World Development Indicators 2004 (WDI). '♯ full' indicates the number of areas included in the sample and '♯ ~~out.~~' represents the latter size once 'outliers' (see Table 2.15) are removed. 'max.' is the WDI subpopulation size and '(%)' indicates the percentage relative to the latter size. Seven CDIAC areas are not classified in WDI: French Guiana, Gibraltar, Guadeloupe, Martinique, Réunion, St. Pierre & Miquelon and Taiwan.

Table 2.19: Geographic groups of CDIAC countries.

East Asia & Pacific (EPA)	Europe & Central Asia (ECA)	Latin America & Caribbean (LAC)	Middle East & North Africa (MENA)	North America (NA)	South Asia (SA)	Sub-Saharan Africa (SSA)
American Samoa	Albania	Ant.&Barb.	**Algeria**	Bermuda	Afghanistan	**Angola**
Australia	**Austria**	**Argentina**	**Bahrain**	**Canada**	Bangl.& Pak.	Benin
Brunei	**Belgium**	Bahamas	Djibouti	**USA**	**India**	Burkina Faso
Cambodia	**Bulgaria**	Barbados	**Egypt**		Nepal	**Cameroon**
China	**Cyprus**	Belize	**Iran**		**Sri Lanka**	Cape Verde
Fiji	**Czechoslovakia**	**Bolivia**	**Iraq**			Central African R.
French Polynesia	**Denmark**	**Brazil**	**Israel**			Chad
Guam	Faeroe Islands	Cayman Islands	**Jordan**			Comoros
Hong Kong	**Finland**	**Chile**	**Kuwait**			Congo (Rep.)
Indonesia	**Former Yugoslavia**	**Colombia**	**Lebanon**			**Côte d'Ivoire**
Japan	**France & Monaco**	**Costa Rica**	Libya			Equatorial Guinea
Korea (D.R.)	**Germany**	**Cuba**	Malta			**Ethiopia**
Korea (R.)	**Greece**	Dominica	**Morocco**			Gabon
Laos	Greenland	**Dominican R.**	**Qatar**			Gambia
Macau	**Hungary**	**Ecuador**	**Saudi Arabia**			**Ghana**
Malaysia	Iceland	**El Salvador**	**Syria**			Guinea
Mongolia	**Ireland**	Grenada	**Tunisia**			Guinea Bissau
Myanmar	**Italy & San Marino**	**Guatemala**	**U. Arab Em.**			**Kenya**
New Caledonia	**Luxembourg**	Guyana	Yemen			Liberia
New Zealand	**Netherlands**	Haiti				Madagascar
Palau	**Norway**	**Honduras**				Mali
Papua New Guinea	**Poland**	**Jamaica**				Mauritania
Philippines	**Portugal**	**Mexico**				Mauritius
Samoa	**Romania**	Neth. Ant.&Aruba				Mozambique
Singapore	**Spain**	**Nicaragua**				Niger
Solomon Islands	**Sweden**	**Panama**				**Nigeria**
Thailand	**Switzerland**	Paraguay				Rhodesia-Nyasaland
Tonga	**Turkey**	**Peru**				Rwanda-Urundi
Vietnam	**United Kingdom**	**Puerto Rico**				Sao Tome & Principe
	USSR	St. Lucia				Senegal
		S. Vin.&Grenad.				Sierra Leone
		Suriname				**South Africa**
		Tri.&Tobago				**Sudan**
		Uruguay				Tanzania
		Venezuela				Togo
		Virgin Isl. (US)				Uganda
						Zaire

Note: This classification corresponds to the one used in World Development Indicators 2004 (WDI). In bold is the sample used by Aldy (2006). . The underlined countries represent outliers, *i.e.* countries with relative changes in CO$_2$ per capita emissions over 100% for at least one of the years included in the investigated period. In the grey cells are countries which CO$_2$ and population series have been reconstructed according to the available information in Marland et al. (2006) and US Census Bureau (2006), see Appendix 2.14.

Table 2.20: Geographic groups of countries. Samples' size.

	World sample					ALDY2006 sample	
	CDIAC areas		WDI countries			CDIAC areas	WDI countries
WDI geo. groups	♯ full	♯ ~~out.~~	max.	♯ full (%)	♯ ~~out.~~ (%)	♯ full	♯ full (%)
East Asia & Pacific	29	22	35	29 (83%)	22 (63%)	12	12 (34%)
Europe & Central Asia	30	28	55	49 (89%)	47 (85%)	26	45 (82%)
Latin America & Caribbean	36	30	38	37 (97%)	30 (79%)	21	21 (55%)
Middle East & North Africa	19	11	21	19 (90%)	11 (52%)	15	15 (71%)
North America	3	3	3	3 (100%)	3 (100%)	2	2 (67%)
South Asia	5	5	8	6 (75%)	6 (75%)	2	2 (25%)
Sub-Saharan Africa	37	26	48	40 (83%)	29 (60%)	9	9 (19%)
Unclassified	7	5	-	-	-	1	-
Total	166	130	208	183 (88%)	148 (71%)	88	106 (51%)

Note: This table matches the CDIAC areas from Marland et al. (2006) with the countries included in the geographic groups from the World Development Indicators 2004 (WDI). '♯ full' indicates the number of areas included in the sample and '♯ ~~out.~~' represents the latter size once 'outliers' (see Table 2.15) are removed. 'max.' is the WDI subpopulation size and '(%)' indicates the percentage relative to the latter size. Seven CDIAC areas are not classified in WDI: French Guiana, Gibraltar, Guadeloupe, Martinique, Réunion, St. Pierre & Miquelon and Taiwan.

Table 2.21: South-South Regional International Agreements.

South-South RIA
ACN/ANDEAN (Andean Community of Nations) Bolivia*(1969-), Chile (1969-1976), Colombia*(1969-), Ecuador*(1969-), Peru*(1969-), Venezuela (1973-2006).
AFTA/ASEAN (Asean Free Trade Association) Brunei (1984-), Indonesia*(1967-), Laos P.D.R. (1997-), Malaysia*(1967-), Myanmar (1997-), Philippines*(1967-), Singapore*(1967-), Thailand*(1967-), Vietnam (1995-), Cambodia (1999-), Timor Oriental (2002-).
CARICOM (Caribbean Community and Common Market) Antigua & Barbuda*(1974-), Bahamas (1983), Barbados*(1973-), Belize*(1974-), Dominica*(1974-), Grenada*(1974-), Guyana*(1973-), Haiti (2002-), Jamaica*(1973-), Montserrat† (1974), St. Kitts and Nevis & Anguilla† (1974), St. Lucia*(1974-), St. Vincent & The Grenadines*(1974-), Suriname (1995-), Trinidad & Tobago*(1973-).
CEMAC (Communauté Économique et Monétaire de l'Afrique Centrale) Cameroon*(1964-), Central African Republic*(1964-), Chad*(1964-), Congo*(1964-), Equatorial Guinea (1984-), Gabon*(1964-).
COMESA (Common Market for Eastern and Southern Africa) Angola*(1981-), Burundi & Rwanda*(Rwanda-Urundi) (1981-), Comoros*(1981-), Djibouti*(1981-), Egypt (1999-), Eritrea† (1994-), Ethiopia*(1981-), Kenya*(1981-), Lesotho (1981-1997), Libya (2005-), Madagascar*(1981-), Mauritius*(1981-), Mozambique (1981-1997), Namibia† (1981-2004), Sudan*(1981-), Swaziland† (1981-), Tanzania(1981-2000), Uganda*(1981-), Rhodesia-Nyasaland*(Zimbabwe & Zambia & Malawi, 1981-), Seychelles (2001-), Zaire*(Congo D.R) (1981-).
ECOWAS (Economic Community of West African States) Benin (1975-), Burkina Faso (1975-), Cape Verde (1976-), Côte d'Ivoire (1975-), Gambia (1975-), Ghana (1975-), Guinea (1975-), Guinea-Bissau (1975-), Liberia (1975-), Mali (1975-), Niger (1975-), Nigeria (1975-), Senegal (1975-), Sierra Leone (1975-), Togo (1975-), Mauritania (1975-2002).
LAIA/ALADI (Latin American Integration Association) Argentina (1980-), Bolivia (1980-), Brazil (1980-), Cuba (1980-), Chile (1980-), Colombia (1980-), Ecuador (1980-), Mexico (1980-), Paraguay (1980-), Peru (1980-), Uruguay (1980-), Venezuela (1980-).
MERCOSUR (Mercado Común del Sur) Brazil*(1991-), Argentina*(1991-), Uruguay*(1991-), Paraguay*(1991-), Venezuela (2006-)

Note: In brackets is the year the country officially joined the RIA. * denotes the countries included in the final RIA grouping because (i) they were member of the RIA for most of the period since the RIA was created and (ii) data on CO$_2$ and population were available in Marland et al. (2006) and US Census Bureau (2006) respectively. † indicates that data on CO$_2$ were missing for some years.

Table 2.22: North-South and North-North Regional International Agreements.

North-South RIA
APEC (Asia-Pacific Economic Cooperation) Australia*(1989-), Brunei*(1989-), Canada*(1989-), Chile (1994-), China (1991-), Hong-Kong (1991-), Indonesia*(1989-), Japan*(1989-), Malaysia*(1989-), Mexico (1993-), New Zealand*(1989-), Papua New Guinea (1993-), Peru (1998-), Philippines*(1989-), Russia (1998-), Singapore*(1989-), South Korea*(1989-), Taiwan (1991-), Thailand*(1989-), USA*(1989-), Vietnam(1998-).
NAFTA (North American Free Trade Association) Canada (1994-), Mexico (1994-), United States of America (1994-).
North-North RIA
EU (European Union) Austria*(1995-), Belgium*(1957-), Bulgaria(2007-), Cyprus (2004-), Czech Republic (2004-), Denmark*(1973-), Estonia (2004-), Finland*(1995-), France*(1957-), Germany R.F.*(1957-), Germany*R.D. (1990-), Greece*(1981-), Hungary(2004-), Ireland*(1973-), Italy*(1957-), Latvia (2004-), Lithuania (2004-), Luxembourg*(1957-), Malta (2004-), Netherlands*(1957-), Poland (2004-), Portugal(1986-)*, Romania(2007-), Spain*(1986-), Slovakia(2004-), Slovenia (2004-), Sweden*(1995-), United Kingdom*(1973-).

Note: In brackets is the year the country officially joined the RIA. * denotes the countries included in the final RIA grouping because (i) they were member of the RIA for most of the period since the RIA was created and (ii) data on CO$_2$ and population were available in Marland et al. (2006) and US Census Bureau (2006) respectively. † indicates that data on CO$_2$ were missing for some years.

Chapter 3

Growth and convergence in air pollution.

Evidence from a reduced form nonparametric approach.

Editorial note: This paper is co-authored with Prof. Thanasis Stengos[77].

Abstract: This paper proposes regression estimates of the Ramsey-type growth model with pollution developed by Alvarez et al. (2005). We use a panel of 97 countries spanning over years 1950-2002 to fit a reduced form function where growth rates in per capita CO_2 emissions are explained with initial pollution levels, initial GDP levels and GDP growth rates. Panel models with 10-year and 5-year data are estimated with standard linear and nonlinear least squares and these specifications are tested with the method recently proposed by Hsiao et al. (2007). Semiparametric as well as nonparametric regressions for mixed data are also computed with recent methods introduced by Wood (2000) and Racine and Li (2004). Parametric models are rejected in favor of their nonparametric counterparts and CO_2 growth rates display significant relationships with the explanatory factors as well as nonlinearities. We find evidence of pollution convergence, conditional on GDP levels and growth rates, for OECD and non-OECD countries, even in the presence of divergence in income levels between non-OECD economies. The results also highlight important interactions between the OECD status and the effects of explanatory factors.

JEL Classification: C14 · C23 · Q53

Key Words: Air pollution, carbon dioxide emissions, convergence, economic growth, mixed nonparametric regressions, panel.

---

[77] Department of Economics, University of Guelph, Canada.

3.1 Introduction

Access to natural resources has been for ages the cornerstone of economic prosperity and ensuring abundance in raw materials is perceived as the promise of long run growth. The other side of the coin is that depletion of limited resources put increasing pressure on nature and on the production process. Output expansion generates human waste that hurts the environment and penalizes the ability of nature to regenerate. This conflictive dynamic of economic growth calls for the use of more efficient technologies. From the normative side, it raises the question of what policy measures should be applied to achieve sustainable growth, *i.e* "a balanced growth path with increasing environmental quality and ongoing growth in per capita income[78]". From a positive point of view, the ability of market forces to freely balance the pros and cons of an extensive development is called into question by natural scientists. Capturing the underlying dynamic of the pollution-GDP relationship with data is essential from both perspectives. This paper focuses on the per capita CO_2 emissions and GDP dynamic and falls within the scope of the positive approach. It estimates with recent nonparametric regression methods a reduced form function, derived from a classical optimal growth model amended with a pollution stock component, that includes level as well as growth determinants of pollution growth.

Economic analysis has applied numerous techniques to tackle the GDP-pollution relationship and to explain the most prominent stylized facts. The most controversial empirical finding, known as the 'Environmental Kuznets Curve (EKC)', states that a U-inverted relationship exists between GDP and some specific pollutants. This particular shape being the simplest pattern consistent with sustainable growth, it has been replicated in a variety of theoretical frameworks[79]. Two growth models are of direct interest here. The 'Green Solow Model' of Brock and Taylor (2004b) is a neoclassical growth model which considers pollution as a by-product of the production sector and the existence of an abatement technology with constant returns to scale. A hump-shaped GDP-pollution pattern is found around the balanced growth path when technological progress in abatement exceeds that in goods production. The other testable prediction of this model is that, under equality of the saving rates, abatement intensities, rates of technological progress, but differing initial conditions across countries, absolute convergence in per capita pollution levels holds. As it is standard in Solow-type models, disparate countries have different steady state pollution

---

[78] Brock and Taylor (2004a, p.2-3).

[79] See Brock and Taylor (2004a), Andreoni and Levinson (2001) or Stockey (1998) among others.

levels and transitional paths, functions of both initial conditions and country characteristics. In the latter case, conditional convergence leads the countries with the closest structural parameters to converge toward similar (but not equal) per capita pollution levels and convergence clubs may emerge. The Green Solow model provides a theoretical basis for testing pollution convergence between countries in the well-known β-convergence framework. The major drawback of the model is that its fundamental "sustainability" parameter, technological abatement progress, is not directly observable and must be estimated with ad-hoc techniques.

Alvarez et al. (2005) propose an alternative analysis conducted in a maximizing utility framework with infinitely lived agents. These authors follow Stockey (1998) and Brock and Taylor (2004b) in the way pollution is modeled: emissions are proportional to the level of production and increasingly clean techniques reduce the pollution intensity over time. Using standard derivations from the growth literature, in particular from the Ramsey-Cass-Koopman[80] (RCK henceforth) framework, they produce a reduced form function around the steady state which makes the pollution growth rate of a country depend on its initial pollution level, its initial output level and its output growth rate. The appeal of that equation is twofold: (i) accounting for level and growth magnitudes in GDP allows to distinguish clearly between levels and growth effects of GDP on pollution growth and (ii) all variables are directly observable from available data. Our paper makes use of the latter specification to investigate the convergence hypothesis for per capita CO_2 emissions.

During the late 90s, the empirical analysis of the GDP-pollution relationship has been dominated by the EKC literature. This approach analyses the link between pollution and GDP by focusing on both variables in levels (total, per capita or intensity units) and geographically (countries or regions) aggregated data. The evidence gathered with this approach is rather mixed, in particular for per capita CO_2 emissions. Making use of panel data, the well-known papers by Holtz-Eakin and Selden (1995) and Heil and Selden (2001) employ parametric models with pooled data and conclude to EKC (or U-inverted) shapes for CO_2 per capita emissions. More recent works by Taskin and Zaim (2000), Bertinelli and Strobl (2005) and Azomahou et al. (2006) estimate non or semiparametric pooled regressions and get nonlinear increasing patterns. Univariate approaches with level data, inspired from the income convergence literature, have also been employed to explore the convergence of CO_2 per capita emissions between countries/regions. Strazicich and List

---

[80] See Barro and Sala-i Martin (2004, Ch.2) for the textbook version of the model.

(2003), Barrasi et al. (2008) and Westerlund and Basher (2008) use unit root tests to investigate stochastic convergence for different sets of world countries. The results are debated but a large body of that literature points rather toward convergence[81]. Distributional dynamics based on the seminal work by Quah (1993a; 1997) is applied by Nguyen Van (2005), Stegman (2005) and Aldy (2006) and the results indicate 'persistence' (neither convergence nor divergence) in the relative per capita pollution gaps between countries. Finally, regarding β-convergence regressions, Strazicich and List (2003) investigate absolute as well as conditional β-convergence for a sample of twenty one OECD countries with pollution growth rates covering the entire 1960-1997 period. The multivariate regression includes a set of ad hoc regressors which capture country-specific characteristics: GDP, GDP squared, gasoline price, population density and a temperature indicator. The convergence coefficient appears to be significant and negative in the simple regression as well as for different combinations of control variables. Among the conditioning factors, only gas price and temperature are found significant[82]. Nguyen Van (2005) also tests absolute β-convergence with a cross-sectional regression and a sample of 100 countries spanning over the period 1966-1996. He does not reject the absolute convergence hypothesis. Brock and Taylor (2004b) compute growth regressions for per capita CO_2 emissions based on their Green Solow model. The absolute cross-sectional $\beta-$type regression is progressively augmented with time-averaged country-specific (estimations of) technological progress in abatement, saving rate, abatement level and effective depreciation rate of capital. The model is tested for OECD countries over the period 1960-1998 and the fits indicate that most of the explanatory power comes from the initial level of pollution. Its associated coefficient has the expected negative sign suggesting β-convergence but none of the other explanatory factors are significant. Finally, Alvarez et al. (2005) estimate cross-sectional as well as panel nonlinear least square regressions based on their reduced form specification for sets of European countries and for several pollutants over the period 1990-2002. They find significant spatial convergence for all groups. GDP growth has a significant positive effect on per capita SO_2 emissions but it affects CO_2 and NO_2 emissions only through an interaction dummy identifying middle-income countries.

---

[81] Strazicich and List (2003) conclude to stochastic convergence for the OECD countries but Barrasi et al. (2008) challenge this finding by accounting for cross-sectional dependencies in the convergence test and providing a different interpretation of the unit root test employed by the latter authors. Westerlund and Basher (2008) use a longer time span and a different sample of world countries and they find stochastic convergence with the method introduced by Evans (1998).

[82] Both variables have a negative impact on per capita CO_2 emissions' growth rates.

Our paper builds on the theoretical formulation proposed by Alvarez et al. (2005) but departs from their econometric estimation in several ways. First, we focus solely on CO_2 emissions and expand widely both the time span and the number of countries included in the panel. Second, the evidence is based on a larger set of empirical methods. Parametric constrained and unconstrained fits are compared and a misspecification test is applied. A flexible regression approach is also proposed based on a semiparametric specification as well as fully nonparametric regressions that better captures nonlinearities and interactions between the regressors. Third, we account for potential endogeneity bias by providing instrumental variables estimates.

The structure of this paper is as follows. In section 3.2, we present the main features of the theoretical growth model of Alvarez et al. (2005). As these authors have proposed different versions of their model, our presentation follows their early 2004 version that we adapt to match[83] their latest version from 2005. The empirical part is presented in section 3.3 and includes the data description in section 3.3.1, the econometric strategy in 3.3.2 and the results in section 3.3.3. We conclude in section 3.4.

3.2 Theoretical model

As Stockey (1998) points out, when exploring some types of environmental degradations, like global warming, it is the cumulative effect of past actions that affects utility at each point date. The model proposed by Alvarez et al. (2005) is a growth model which incorporates the cumulated pollution stock as a component of the utility function of a representative and inter-temporal maximizing-utility household. Therefore, their framework appears to be a good candidate to investigate CO_2 emissions.

Production side: Matching Alvarez et al. (2005)' notations to those used in Barro and Sala-i Martin (2004), for any relevant variable (*e.g.* output, pollution, capital stock, etc …), Q, q and $\hat{q}$ designate level, per capita and effective per capita units while $\dot{Q}$ is the derivative of Q with respect to time. More precisely, let $K(t)$, $L(t)$ and x be respectively the capital stock of an economy at time t, the total amount of labor force available at time t (assumed equal to the population) and an exogenous and constant rate of labor-augmenting technological progress such that $x \in (0, 1)$; then, $k(t) = K(t)/L(t)$

---

[83] Any error in that exercise is our own.

and $\hat{k}(t) = K(t)/\hat{L}(t) = K/Le^{xt}$ are respectively capital stock per capita and capital stock in effective per capita units. Note that when not ambiguous, we drop the "time function" after a variable that depends on t, *i.e.* $L(t) \equiv L$. As usual in this setting, output $Y(t)$ is produced with a neoclassical production function $F(K(t), \hat{L}(t))$ that exhibits constant returns to scale in physical capital $K(t)$ and labor $\hat{L}(t)$, decreasing returns with respect to both factors and fulfills Inada conditions[84]. This paper makes use of a Cobb-Douglas production function that takes the form $Y = A_0 K^\alpha \hat{L}^{1-\alpha}$, where $A_0 > 0$ is a technological and exogenous factor and $\alpha \in (0,1)$ is the capital-output elasticity. Because of the constant return to scale assumption, the Cobb-Douglas function can be expressed in effective per capita units as

$$\hat{y} = A_0 \hat{k}^\alpha \tag{3.1}$$

The inputs $K(t)$ and $L(t)$ hired by the production sector have a fixed unitary cost of $R(t)$ and $w(t)$ while technological progress is free. Physical capital depreciates at the constant rate δ. In a competitive market, the price of each unit of K must cover at least the risk free investment rate $r(t)$ plus its depreciation rate δ, which implies $R(t) \geq r(t) + \delta$. The representative firm sells its output at a unit price to households and maximizes at any point in time a flow of net receipts given by $\pi(t) = F(K(t), \hat{L}(t)) - [R(t)K(t) + w(t)L(t)]$. Using the production function (3.1) and omitting the function of t, total profits can be similarly written $\pi = \hat{L}[A_0 \hat{k}^\alpha - R\hat{k} - we^{-xt}]$. Imposing a maximizing zero-profit condition implies that the marginal product of each factor equals its cost:

$$\frac{\partial F}{\partial K} = \alpha A_0 \hat{k}^{\alpha - 1} = r + \delta \tag{3.2}$$

$$\frac{\partial F}{\partial L} = (1-\alpha) A_0 \hat{k}^\alpha e^{xt} = w \tag{3.3}$$

Pollution: Economic activity is assumed to generate an unregulated pollution flow every period, noted $P(t)$, that cumulates in the atmosphere and dissipates at a constant rate $\delta_z \in (0,1)$. As pointed out by Brock and Taylor (2004a), assuming a dissipation rate independent of the pollution level may be a strong assumption depending on the context. A fixed exogenous factor ($\eta \geq 0$), such as a country-specific time-invariant structural factor,

---

[84] See Barro and Sala-i Martin (2004, Ch.1.2.1) for a detailed presentation of all the hypotheses underlying the neoclassical production function.

is assumed to affect pollution over time. Aggregate pollution and the law of motion of its stock, noted $\dot{Z}$, are given by

$$Z(t) = \eta \int_{-\infty}^{t} P(s) e^{-(\delta_z)(t-s)} \mathrm{d}s \tag{3.4}$$

$$\dot{Z}(t) = \eta P(t) - \delta_z Z(t) \tag{3.5}$$

Equation (3.5) follows directly from differentiating equation (3.4) with respect to time and applying the Leibniz's rule. The fundamental equation that characterizes the pollution flow in t is modeled as an intensity and scale effects of $\hat{y}(t)$ that are aggregated with a Cobb-Douglas technology. This relationship can be formulated as

$$\hat{p} = \tilde{B} \left(\frac{\hat{z}}{\hat{y}} \right)^{\phi} \hat{y} = \tilde{B} \hat{z}^{\phi} \hat{y}^{1-\phi} \tag{3.6}$$

$$\hat{p} = B \hat{z}^{\phi} \hat{k}^{\alpha(1-\phi)} \tag{3.7}$$

where $B = \tilde{B} A_0^{1-\phi}$; the efficiency factor $\tilde{B} := \tilde{B}_0 e^{-x_b t} > 0$, with $\tilde{B}_0 \geq 0$ and $x_b \geq 0$, allows exogenous improvements in the emission process and $\phi \in (0, 1($ is an elasticity parameter of the pollution technology.

Households: Identical households, with the same utility function and facing the same market and environmental conditions, provide labor services paid at rate $w(t)$, earn an interest $r(t)$ on the assets they possess, purchase goods for consumption and save by accumulating assets over time. Given the same asset endowment at the beginning of the period and a common natural and positive growth rate n shared by all households, the analysis can be carried out in a representative-agent framework. Each household wishes to maximize its utility that depends on current per capita consumption $c(t)$ of the only good produced and the cumulated pollution level $Z(t)$ present in the atmosphere, over an infinite planing horizon. The infinite horizon does not imply necessarily that agents live forever but that each household optimizes its behavior accounting for the welfare and resources constraints of its descendants. The overall utility is given by

$$\int_{0}^{\infty} u(c(t), h[Z(t)]) e^{nt} e^{-\rho t} \tag{3.8}$$

where $\rho > 0$ is the rate of time preference and the multiplication of $u(c)$ by the family size[85] $L = e^{nt}$ is interpreted as the utility of all members of the household alive at time t. To ensure a bounded utility when $c(t)$ and $Z(t)$ are constant, it is standard to set $\rho > n$. The function $u()$ is assumed to be concave and increasing in $c(t)$ and increasing in $h()$, with $h_Z < 0$. The budget constraint for the whole economy (aggregate households) can be specified dynamically as

$$\dot{\mathcal{A}}(t) \equiv \frac{\partial \mathcal{A}(t)}{\partial t} = r(t)\mathcal{A}(t) + w(t)L(t) - C(t) \tag{3.9}$$

The above equation simply states that savings (assets returns plus salaries minus consumption) is used to accumulate more assets. Denoting $a = \frac{\mathcal{A}(t)}{L(t)}$ the budget constraint in per capita terms becomes

$$\dot{a} = (r - n)a + w - c \tag{3.10}$$

With equations (3.8) and (3.10) at hand and provided that Z follows an exogenous time path[86], we are ready to develop the maximizing utility problem for households as in the RCK framework:

$$\max_{\{c(t)\}_{t \geq 0}} \int_0^\infty u(c(t), h[Z(t)])e^{(\rho - n)t} \tag{3.11}$$

s.t.

$$\text{(a) } \dot{a} = (r - n)a + w - c$$

$$\text{(b) } a(0) = a_0$$

$$\text{(c) } c(t) \geq 0$$

$$\text{(d) } \lim_{t \to \infty} [a(t)e^{-\int_0^t (r(s) - n)ds}] \geq 0$$

Because of the form of the utility function, condition (3.11.c) never binds, while con-

---

[85] The model assumes that the family size of $L(0) = 1$ by normalization.

[86] Variable $Z(t)$ does not represent a choice variable for consumers. They can influence the level of $Z(t)$ only through consumption. Note that this assumption of "exogeneous time path" is also adopted in Barro and Sala-i Martin (2004, Ch.3.1.3, p.149) for deriving an Euler equation when government purchases are introduced in a consumer's utility function like the one from equation (3.8).

dition (3.11.d) avoids households to play a Ponzi game[87]. This dynamic optimization problem can be solved with optimal control theory techniques. The Hamiltonian and its maximum principle conditions are

$$H = u(c(t), h[Z(t)])e^{(\rho-n)t} + \lambda(t)((r-n)a + w - c) \tag{3.12}$$

$$\begin{aligned}
&\text{(a) } \frac{\partial H}{\partial c} = 0 \Rightarrow \lambda = u_c e^{-(\rho-n)t} \\
&\text{(b) } \dot{\lambda}(t) = -\frac{\partial H}{\partial a} \Rightarrow \dot{\lambda}(t) = -(r-n)\lambda(t) \\
&\text{(c) } \lim_{t\to\infty}[\lambda(t)a(t)] = 0
\end{aligned}$$

Condition (3.12.c) is the transversality condition. Deriving $\lambda(t)$ with respect to time in (3.12.a) and combining with (3.12.b) allows to derive the Euler equation

$$\begin{aligned}
r &= \rho - \frac{u_{cc} \cdot c}{u_c} \cdot \frac{\dot{c}}{c} - \frac{u_{cZ} \cdot Z}{u_c} \cdot \frac{\dot{Z}}{Z} \quad \text{or} \\
\frac{\dot{c}}{c} &= -\frac{u_c}{u_{cc} \cdot c}(r - \rho + \frac{u_{cZ} \cdot Z}{u_c} \cdot \frac{\dot{Z}}{Z})
\end{aligned} \tag{3.13}$$

Using the constant elasticity utility function

$$u(c, Z) = \frac{(c^{\nu} h(Z)^{1-\nu})^{1-\sigma} - 1}{1-\sigma} \tag{3.14}$$

where $\sigma > 0$ is the inverse intertemporal elastictiy of substitution, $\nu \in [0, 1]$ gives the relative importance of consumption in welfare and the health (or damage) function of the pollution stock is given by $h(Z) = \frac{Z^{-\varepsilon}}{\varepsilon}$, where $\varepsilon \geq 0$ is the constant elasticity of the health function with respect to $Z(t)$. It can be shown that the Euler equation becomes

$$\frac{\dot{c}}{c} = \frac{1}{\xi}\left(r - \rho - \bar{\xi}\frac{\dot{Z}}{Z}\right) \tag{3.15}$$

where $\xi = 1 - \nu(1-\sigma)$ and $\bar{\xi} = \varepsilon(1-\nu)(1-\sigma)$.

---

[87] See Barro and Sala-i Martin (2004, p.89).

Now that the production, pollution and households sides of the economy have been presented, we can state the equilibrium conditions of Alvarez et al. (2005)'s model.

Equilibrium: setting the equilibrium on the assets' market to $a = k$ and replacing r and w from the asset constraint (3.10) by their level at the zero profit condition (3.2) and (3.3), we get the well-known Solow-type dynamics of capital[88]

$$\dot{\hat{k}} = A_0 \hat{k}^{\alpha} - \hat{c} - (x + \delta + n)\hat{k} \tag{3.16}$$

The transversality condition can also be restated at equilibrium: integrate between 0 and t expression $\frac{\dot{\lambda}(t)}{\lambda(t)}$ from condition (3.12.b) with respect to time and replacing the result into the transversality condition (3.12.c), the credit market constraint (3.11.d) appears to emerge naturally from the transversality condition. Substituting r in the latter expression for its zero-profit counterpart from (3.2) and using $\hat{k} = ke^{-xt}$ as well as the clearing market condition $a = k$, we can get a transversality condition that is identical to the standard RCK model without pollution

$$\lim_{t \to \infty} \hat{k}(t) e^{-\int_0^t [\alpha A_0 \hat{k}(s)^{\alpha - 1} - (n + \delta + x)] ds} = 0 \tag{3.17}$$

As $\hat{k}$ will converge toward a constant steady state value in the long run (see Steady state section in the next page), Barro and Sala-i Martin (2004, p.98) note that the transversality condition (3.17) requires that the steady state rate of return $f'(\hat{k}_s) - \delta$ exceeds $x + n$, the steady state growth rate of Y.

At that stage, it is useful to rewrite the system of differential equations that characterizes the whole economy in terms of effective per capita units. Note that the growth equation for consumption (3.15) involves a term $\dot{Z}/Z$ that needs to be expressed in efficient per capita units and the term r that can be replaced by its related non-profit condition

---

[88] Note that this equation departs slightly from Alvarez et al. (2005)'s one as it does not include explicitly a constant term, denoted ς, multiplying $\hat{y} = A_0 \hat{k}^{\alpha}$ in equation (3.16). That constant is intended to represent the fraction of output devoted to consumption or investment, the rest being allocated to pollution abatement activities. As this constant is kept fixed and exogenous by these authors, it can be subsumed within the A_0 term. Therefore the implicit assumption is that the fraction of output allocated to abatement is constant over time.

(3.2). The differential equation for capital (3.16) is already expressed in a convenient way. Regarding the pollution dynamics at equilibrium, note that equation (3.5) can be rewritten as $\dot{\hat{z}}(t) = \eta\hat{p}(t) - (\delta_z + n + x)\hat{z}(t)$. Then replacing $\hat{p}$ in the latter expression by its expression in (3.7), we get the required formulation. All this drives to

$$\frac{\dot{\hat{c}}}{\hat{c}} = \frac{1}{\xi}\left(A_0\alpha\hat{k}^{\alpha-1} - (\delta + \rho + x\xi) - \bar{\xi}(n + x + \frac{\dot{\hat{z}}}{\hat{z}})\right) \tag{3.18}$$

$$\frac{\dot{\hat{k}}}{\hat{k}} = A_0\hat{k}^{\alpha-1} - \frac{\hat{c}}{\hat{k}} - (x + \delta + n) \tag{3.19}$$

$$\frac{\dot{\hat{z}}}{\hat{z}} = \eta B\hat{z}^{\phi-1}\hat{k}^{\alpha(1-\phi)} - (\delta_z + n + x) \tag{3.20}$$

Steady state: The steady state condition for $\hat{c}$ and $\hat{k}$ in Alvarez et al. (2005) is identical to the one described in Barro and Sala-i Martin (2004, Ch.2.5), and yields to $\dot{\hat{c}}/\hat{c} = \dot{\hat{k}}/\hat{k} = \dot{\hat{y}}/\hat{y} = 0$. Regarding pollution stock and flow variables, note first that equation (3.5) implies that a stabilization of the growth rate of the pollution stock occurs when $\dot{\hat{p}}/\hat{p} = \dot{\hat{z}}/\hat{z}$. Second, standard logarithmic differentiation of (3.6) leads to $\frac{\dot{\hat{p}}}{\hat{p}} = -x_b + \phi\frac{\dot{\hat{z}}}{\hat{z}} + (1-\phi)\frac{\dot{\hat{y}}}{\hat{y}}$. Combining these two relationships with the steady state condition $\dot{\hat{y}}/\hat{y} = 0$, we get that $\dot{\hat{p}}/\hat{p} = -x_b/(1-\phi)$. Hence, along the balanced growth path $\tilde{z}(t) = \hat{z}(t)e^{[x_b/(1-\phi)]t}$ and $\tilde{p}(t) = \hat{p}(t)e^{[x_b/(1-\phi)]t}$ are constant. The steady state condition also implies that while consumption and the stock of capital in per capita terms grow at rate x, the (stock and) flow of per capita pollution grow at rate $x - x_b/(1-\phi)$, which may be null or negative. This is a modified version of the well-known "sustainability" statement which says that the stability of per capita emissions in a growing economy depends on the difference between the (polluting) output-augmenting technical progress and the rate of emissions-reduction technical progress. Bounded pollution per capita requires imposing $x \leq x_b/(1-\phi)$. Note that the pollution/output ratio at the steady state can be shown to be decreasing[89].

As it is standard in growth theory, local dynamics are analyzed by log-linearizing the nonlinear system (3.18) to (3.20) at its steady state. This leads in particular to an equation of the growth rate of per capita pollution that is the foundation of our empirical analysis. However, as argued by Alvarez et al. (2005), the nonseparability between consumption and health services in the utility function (3.14) yields to complicated dynamics that hinder

---

[89] See Alvarez et al. (2005, p.11).

obtaining a simple analytical local solution. Thus, these authors opt for the following separable utility function:

$$v(c,Z) = \frac{C^{1-\sigma}-1}{1-\sigma} + \frac{Z^{-\varepsilon}}{\varepsilon} \tag{3.21}$$

Given that $v_{c,Z}() = 0$, the Euler equation (3.15) becomes the one from the standard RCK model, $i.e$ $\frac{\dot{\hat{c}}}{\hat{c}} = \frac{(r-\rho-x\sigma)}{\sigma}$. This leads to a simpler expression describing the dynamics of $\hat{c}$

$$\frac{\dot{\hat{c}}}{\hat{c}} = \frac{1}{\sigma}\left(A_0\alpha\hat{k}^{\alpha-1} - (\delta+\rho+x\sigma)\right) \tag{3.22}$$

which replaces the former equation (3.18). The nonlinear system formed by equations (3.19), (3.20) and (3.22) can now be log-linearized around the steady-state, which drives to the following result in matrix representation

$$\begin{pmatrix} \dot{\log\hat{c}} \\ \dot{\log\hat{k}} \\ \dot{\log\hat{z}} \end{pmatrix} = \begin{pmatrix} 0 & -\beta_{ck} & 0 \\ -\beta_{kc} & \beta_{kk} & 0 \\ 0 & \beta_{zk} & -\beta_{zz} \end{pmatrix} \begin{pmatrix} \log\hat{c} - \log\hat{c}_s \\ \log\hat{k} - \log\hat{k}_s \\ \log\hat{z} - \log\hat{z}_s \end{pmatrix} \tag{3.23}$$

where $\hat{c}_s$, $\hat{k}_s$, $\hat{z}_s$ represent the steady state values[90] while the β_{ab}'s represent the log-transformation of the differential equation for variable $\log\hat{a}$ with respect to $\log\hat{b}$. These βs are:

---

[90] It is straightforward to show that

$$\hat{c}_s = A_0\hat{k}_s^{\alpha} - (\delta+n+x)\hat{k}_s$$
$$\hat{k}_s = \left(\frac{\alpha A_0}{(\eta+\delta+x\sigma)}\right)^{\frac{1}{1-\alpha}}$$
$$\hat{z}_s = \left(\frac{\eta B}{\delta_z + x + n - \frac{x_b}{1-\phi}}\right)^{\frac{1}{1-\phi}} \hat{k}_s^{\alpha}$$

$$
\begin{aligned}
\beta_{ck} &= \frac{(1-\alpha)(\rho+\delta+x\sigma)}{\sigma} \\
\beta_{kc} &= \frac{\rho+\delta+x\sigma}{\alpha} - (\delta+n+x) \\
\beta_{kk} &= \rho - n + x(\sigma-1) \\
\beta_{zk} &= \alpha(1-\phi)(\delta_z + x + n - \frac{x_b}{1-\phi}) \\
\beta_{zz} &= (1-\phi)(\delta_z + x + n - \frac{x_b}{1-\phi})
\end{aligned}
$$

The βs' matrix represents the Jacobian of the differential system (3.23). The 3×3 log-linearized differential system becomes recursive and it can be treated in a two step procedure involving two 2×2 differential subsystems: one in $\dot{\log\hat{c}}$ and $\dot{\log\hat{k}}$ (which corresponds to the usual RCK model); the other one in $\dot{\log\hat{k}}$ and $\dot{\log\hat{z}}$. In the first step, subsystem

$$
\begin{pmatrix} \dot{\log\hat{c}} \\ \dot{\log\hat{k}} \end{pmatrix} = \begin{pmatrix} 0 & -\beta_{ck} \\ -\beta_{kc} & \beta_{kk} \end{pmatrix} \begin{pmatrix} \log\hat{c} - \log\hat{c}_s \\ \log\hat{k} - \log\hat{k}_s \end{pmatrix} \tag{3.24}
$$

is solved. Its Jacobian can be shown to have two real roots (one positive and one negative), which drive to a saddle path equilibrium and to the well-known solution[91] for capital's differential equation:

$$
\log\hat{k}(t) - \log\hat{k}_s = (\log\hat{k}(0) - \log\hat{k}_s)e^{-\tilde{\beta}t} \tag{3.25}
$$

where $\tilde{\beta} = \frac{\sqrt{\beta_{kk}^2 + 4\beta_{ck}\beta_{kc}} - \beta_{kk}}{2} > 0$. The latter result is used in the second subsystem

$$
\begin{pmatrix} \dot{\log\hat{k}} \\ \dot{\log\hat{z}} \end{pmatrix} = \begin{pmatrix} \tilde{\beta} & 0 \\ \beta_{zk} & -\beta_{zz} \end{pmatrix} \begin{pmatrix} \log\hat{k} - \log\hat{k}_s \\ \log\hat{z} - \log\hat{z}_s \end{pmatrix} \tag{3.26}
$$

It can be shown that the two roots of the Jacobian being negative (*i.e.* $-\tilde{\beta}$ and $-\beta_{zz}$),

---

[91] See, among others, Barro and Sala-i Martin (2004, p.111).

subsystem (3.26) is stable. The solution for $\tilde{z}$ can be obtained in the standard way[92]

$$\begin{aligned}\log\tilde{z}(t) - \log\tilde{z}_s &= (\log\tilde{z}(0) - \log\tilde{z}_s)e^{-\beta_{zz}t} \\ &\quad + \frac{\beta_{zk}}{\beta_{zz} - \tilde{\beta}}(\log\hat{k}(0) - \log\hat{k}_s)(e^{-\tilde{\beta}t} - e^{-\beta_{zz}t})\end{aligned} \tag{3.27}$$

Solutions (3.25) and (3.27) can be used to obtain expressions for $\hat{y}$ and $\tilde{p}$. In the former case, by rewriting the production function (3.1) in logs and recalling that the relationship holds at $t = 0$ as well as at the steady state, we get

$$\log\hat{y}(t) - \log\hat{y}_s = (\log\hat{y}(0) - \log\hat{y}_s)e^{-\tilde{\beta}t} \tag{3.28}$$

Following the same reasoning for the pollution flow equation (3.6), substituting $\hat{p}(t)$ for $\tilde{p}(t)$ and using equations (3.27) and (3.28) yields

$$\begin{aligned}\log\tilde{p}(t) - \log\tilde{p}_s &= (\log\tilde{p}(0) - \log\tilde{p}_s)e^{-\beta_{zz}t} \\ &\quad + \lambda(\log\hat{y}(0) - \log\hat{y}_s)(e^{-\tilde{\beta}t} - e^{-\beta_{zz}t})\end{aligned} \tag{3.29}$$

where $\lambda = (1 - \phi) + \frac{\beta_{zk}}{\beta_{zz} - \tilde{\beta}}\frac{\phi}{\alpha}$. Subtracting $\log\tilde{p}(0)$ from both sides of equation (3.29), adding and subtracting $\lambda\log\hat{y}(0)$ at the right-hand side and rearranging, we obtain

$$\begin{aligned}\log\tilde{p}(t) - \log\tilde{p}(0) &= [\log\tilde{p}(0) - \log\tilde{p}_s - \lambda(\log\hat{y}(0) - \log\hat{y}_s)](e^{-\beta_{zz}t} - 1) \\ &\quad + \lambda(\log\hat{y}(t) - \log\hat{y}(0))\end{aligned} \tag{3.30}$$

Result (3.30) constitutes the penultimate step toward a growth equation that links pollution growth to past/initial pollution levels, past/initial output levels and output growth. To make the latter equation estimable, Alvarez et al. (2005) transform the key variables in per capita units, discretize the equation between $t - T$ and t, and obtain the

---

[92] Note that the eigenvectors $\begin{pmatrix} v_{11} \\ v_{21} \end{pmatrix}$ and $\begin{pmatrix} v_{12} \\ v_{22} \end{pmatrix}$ associated with the eigenvalues $-\tilde{\beta}$ and $-\beta_{zz}$ of subsystem (3.26) are normalized by setting $v_{2,1} = 1$ and $v_{2,2} = 1$, leading to the matrix of eigenvectors $\begin{pmatrix} \frac{\beta_{zz}-\tilde{\beta}}{\beta_{zk}} & 0 \\ 1 & 1 \end{pmatrix}$. By defining the arbitrary constants at $t = 0$, we get solutions (3.27) and (3.25).

following reduced form function[93]

$$
\begin{aligned}
GP(t) = x(1-\lambda) - \frac{x_b}{1-\phi} + \bar{\beta}\left(x(t-T)(1-\lambda) - \frac{x_b}{1-\alpha}(t-T)\right) \\
+ \bar{\beta}(\log\tilde{p}_s - \lambda\log\hat{y}_s) - \bar{\beta}\log p(t-T) + \bar{\beta}\lambda\log y(t-T) + \lambda GY(t)
\end{aligned}
\tag{3.31}
$$

where T is the length of the time period; $GP(t) := \frac{1}{T}\log[p(t)/(p(t-T)]$ and $GY(t) := \frac{1}{T}\log[y(t)/(y(t-T)]$ are per capita pollution and output growth rates between time $t-T$ and t; $p(t-T)$ and $y(t-T)$ are initial levels of pollution and output in $t-T$; $\bar{\beta} = (1-e^{-\beta_{zz}T})/T$. Equation (3.31) is similar to the usual β-type growth equations: the first two terms at the right-hand side are technical progress parameters linked to labor and pollution efficiency; the third term represents initial technical levels; the forth term involves steady state equilibria for pollution and output; the fifth expression in the second row measure β-type convergence with respect to past pollution levels while the last two terms capture how pollution growth is linked to past output levels and the output dynamic. Putting together the technical and steady-state expressions under parameter γ and using time subscripts, a simple estimable form arises.

$$
GP_t = \gamma - \bar{\beta}\log p_{t-T} + \bar{\beta}\lambda\log y_{t-T} + \lambda GY_t \tag{3.32a}
$$

$$
GP_t = \gamma + \beta\log p_{t-T} + \delta_1\log y_{t-T} + \delta_2 GY_t \tag{3.32b}
$$

Equation (3.32a) constitutes the reduced form that can be estimated empirically under the β-convergence framework[94]. Setting $\beta = -\bar{\beta}$, $\delta_1 = -\beta\lambda$ and $\delta_2 = \lambda$ in (3.32a), we get the equivalent equation (3.32b) that we use in the empirical section. Note that the theoretical model imposes a nonlinear restriction between the coefficients which multiply p_{t-T}, y_{t-T} and GY_t

---

[93] To make the discretization more explicit, first note that $\log\tilde{p}(t) = \log[p(t)e^{\frac{x_b}{1-\phi}t}e^{-xt}] = \log p(t) + \frac{x_b}{1-\phi}t - xt$. Hence, discretization of the left-hand side expression of equation (3.30) drives to $[\log\tilde{p}(t) - \log\tilde{p}(t-T)]/T = \frac{1}{T}\log\left(\frac{p(t)}{p(t-T)}\right) + \frac{x_b}{1-\phi} - x$. Applying the same procedure to $\log\tilde{p}(0)$, $\log\hat{y}(0)$ and $\lambda(\log\hat{y}(t) - \log\hat{y}(0))$ and rearranging yields equation (3.31). Note that our result differ slightly from Alvarez et al. (2005) as the third term from equation (3.31) is not present in their paper.

[94] See Barro and Sala-i Martin (2004, Ch.11 and 12) or de la Fuente (2000, Ch.11.4.a).

3.3 Empirical analysis

3.3.1 Data

Many recent contributions employ CO_2 emissions data from Marland et al. (2006) to investigate the convergence hypothesis for carbon emissions with univariate methods. We use these data as well as population series from U.S. Census Bureau (2006) and GDP series of Maddison (2007) to construct the variables entering the reduced form function (3.32a). Our panel database consists in yearly series for 97 countries spanning from 1950 until 2000. According to the World Bank definition, CO_2 (carbon dioxide) emissions are those stemming from the burning of fossil fuels and the manufacture of cement. They include contributions to the carbon dioxide produced during consumption of solid, liquid, and gas fuels and gas flaring. CO_2 is a stable gas which is not transformed chemically in the atmosphere. However, some CO_2 is removed from the atmosphere by a natural process that includes the effect of vegetation, soils and oceans. Moreover, human activities such as reforestation, deforestation or land management may increase or decrease the amount of CO_2 removed from the atmosphere. For example, the global natural CO_2 removal rate has been estimated to be around 60 percent for the period 1980 to 1989 and 52 percent for the 1989 to 1998 period, see IPCC, 2000. If we were to account for human-induced changes in land use and forestry we could derive country-specific values on the basis of CO_2 emission data provided in the website of the United Nations Framework Convention on Climate Change (UNFCCC)[95]. As part of their obligation, countries report to the UNFCCC their annual emissions of greenhouse gases, with data currently spanning the period 1990-2004. For countries in our sample, emissions are provided with and without taking into account CO_2 removal resulting from direct human-induced land use, land use change and forestry (LULUCF). The ratio of emissions with LULUCF over emissions without LULUCF gives the rate of CO_2 removal because of human activities. The fact that there is a natural and a human induced removal rate for each individual country is important, since it points to different possibilities for convergence.

3.3.2 Econometric methods

Parametric approach. As it is common in recent papers, the theoretical model is investigated with panel regressions. Rather than making use of yearly growth rates for pollution and income, our panel estimates are based on four 10-year ($T = 10$) as well as

---

[95] See http://unfccc.int/ghg_emissions_data/predifined_qeuries/items/3814.php

eight 5-year ($T = 5$) periods, starting in year 1960. As pointed out by Barro and Sala-i Martin (2004, Ch. 11.10), taking shorter periods carries the risk of missing long run adjustments. More precisely, short run growth rates tend to capture short term adjustments around the trend rather than long run convergence. In the presence of business cycles, this leads to an upward bias of the estimates of the convergence speed, see Shioji (1997). Including country-specific effects in our case would result in adding 96 dummies to the panel regressions. To avoid this inflation of regressors we simply discriminate for OECD membership with a single dummy. Time dummies are also included to capture potential structural breaks in the relationship. Therefore the panel model is given by

$$GP_{i,t} = \alpha_1 + \alpha_2 D_i + \alpha_{3,t} D_t + \beta \log P_{i,t-T} + \delta_1 \log Y_{i,t-T} + \delta_2 GY_{i,t} + \varepsilon_{i,t} \tag{3.33}$$

where[96]:

$GP_{i,t}$: is the growth rate of CO_2 per capita emissions in the i-th country, measured by the average log changes $(1/T)\log(P_{i,t}/P_{i,t-T})$ over the time span $t-T$ to t;

D_i : is a dummy equal to 1 if the i-th country is an OECD member and equal to 0 if not;

D_t : are dummy variables set to 1 for each period t of the panel;

$P_{i,t-T}$: is the level of CO_2 per capita emissions (tons/capita) in the i-th country at time $t-T$;

$GY_{i,t}$: is the growth rate of GDP per capita in the i-th country, measured by the average log changes $(1/T)\log(Y_{i,t}/Y_{i,t-T})$ over the time span $t-T$ to t;

$Y_{i,t-T}$: is the level of GDP per capita (in 1990 International Geary-Khamis dollars) in the i-th country at time $t-T$;

$\varepsilon_{i,t}$: is an iid error term.

When T is set to the entire length of the time dimension, the latter specification becomes a cross-sectional model where the dynamic component is captured by growth rates

---

[96] Usually, econometric specification use small letters to designate parameters and capital letters for variables. We follow this habit in the econometric sections and change slightly the notations with respect to the theoretical section: in the empirical part P and Y denote per capita terms and small caps are parameters.

over the whole period. Therefore, equation (3.33) can be naturally estimated either with cross-sectional or panel regressions. The latter framework has the advantage of better capturing nonlinearities in the relationship.

Given the potential feedback effect of pollution on GDP, regression's coefficients can suffer from endogeneity bias. We address this issue by providing Instrumental Variables (IV henceforth) estimates. Following Barro and Sala-i Martin (2004, Ch12.2.2), we keep the first 10 years of observations out of the sample to build instruments. We retain $Y_{i,t-(T+10)}$ as instrument for $Y_{i,t-T}$ and $GY_{i,t-(T+10)}$ for $GY_{i,t}$.

Recall also that $\delta_1 = -\beta\delta_2$ in the theoretical model. As in Alvarez et al. (2005), we use nonlinear least squares[97] (NLS) to account for this restriction. This estimation procedure requires specifying initial values for the regression parameters. We provide unconstrained OLS estimates as starting values and verify convergence toward the same optimum for other random starting values uniformly distributed in the $(-10, 10)$ interval. Once $\hat{\beta}^{nls}$ and $\hat{\delta}_2^{nls}$ are available, variance for $\hat{\delta}_1^{nls}$ can be estimated with the delta method[98]. With this information at hand, the significance of $\hat{\delta}_1^{nls} \neq 0$ can be checked by using a standardized statistic, asymptotically distributed as a standard normal (not a student).

In addition to the constrained fits, we investigate unconstrained estimations with OLS regressions and we check that the nonlinear constraint holds with a Wald test[99]. We also apply the specification test by Hsiao et al. (2007) to check if the parametric linear models provide consistent estimates. More exactly, if $E(y_i, x_i)$ is the true but unknown conditional mean that is approximated by some parametric model $E(y_i, x_i; \varphi)$, this test contrasts the following two hypotheses: $H_0 : E(y_i, x_i) = E(y_i, x_i; \varphi)$ *vs.* $H_1 : E(y_i, x_i) \neq E(y_i, x_i; \varphi)$,

---

[97] See the 'nls' function in R, the 'nl' command in Stata or 'ls' in Eviews 5.1. We used the latter software.

[98] The variance of the first order Taylor approximation of $\delta_1 = c(\beta, \delta_2) = -\beta\,\delta_2$ drives to the approximate formula: $\sigma_c^2 = c_\beta^2\,\hat{\sigma}_\beta^2 + c_{\delta_2}^2\,\hat{\sigma}_{\delta_2}^2 + 2\,c_\beta c_{\delta_2}\,\hat{\sigma}_{\beta,\delta_2}$. The variance and covariance terms are taken from the nls variance-covariance matrix and are White-corrected when appropriate. We applied the robust version of the Breusch and Pagan (1979)'s heteroscedasticity test recommended in Greene (2003, Ch.11.4.3) to check the homoscedasticity null of the LS estimates. This version of the test is resistant to departures from residuals' normality.

[99] See Greene (2003, Ch.6.5) for a short presentation on Wald tests for nonlinear restrictions. Note that the Wald statistic is not invariant to mathematically equivalent formulations of the constraint, in particular for small samples. Nevertheless, the test has the advantage of not requiring distributional normality.

almost everywhere[100]. If H_0 is not accepted, more flexible specifications can be explored. This paper considers two alternatives to the linear parametric model (3.33) : a semi-parametric additive model which gives full flexibility to the continuous explanatory components as well as a fully nonparametric regression which allows all kinds of interactions between the independent variables.

Nonparametric approach. We further proceed to estimate a more flexible version of equation (3.33) whereby we relax certain functional restrictions that allow for some of the variables to enter parametrically and the others nonparametrically but with a separable structure. This is the partially linear (PLR) additively separable regression model which can be written as

$$GP_{i,t} = \alpha_1 + \alpha_2 D_i + \alpha_{3,t} D_t + \sum_{j=1}^{3} f_j(x_j^c) + \varepsilon_{i,t} \tag{3.34}$$

where $f_j(x_j^c)$ are three unknown nonlinear functions, one for each jth continuous factors from model (3.33), *i.e.* $x_1^c = \log P_{i,t-T}, x_2^c = \log Y_{i,t-T}, x_3^c = GY_{i,t}$. The first three terms in specification (3.34) constitute the linear part of the PLR model while the last term $\sum_{j=1}^{3} f_j(x_j^c)$ is the additive nonparametric component. Compared to the parametric model (3.33), the PLR setting imposes no restriction on the flexibility of the additive nonparametric factors and it allows a straightforward graphical representation along all its dimensions. The additive block is a quite restrictive special case of the general smooth function $f(x_1^c, x_2^c, x_3^c)$ but it can be estimated more efficiently than a fully nonparametric setting when it represents the true relationship. Linton and Nielsen (1995), Fan et al. (1998) and Fan and Li (2004) use marginal integration to estimate the components of the additive semiparametric PLR model in equation (3.34). Application of marginal integration to a general PLR model implies that the asymptotic distribution of $(\widehat{f}_j(z_j) - f_j(z_j))$

---

[100] In the context of mixed data, the test is based on the following three key magnitudes:

$$T_n^a = n(\hat{h}_1 \dots \hat{h}_q)^{1/2} I_n^a(\hat{h}.\hat{\lambda})/\hat{\sigma}_a \xrightarrow{d} N(0,1) \text{ under the null}$$

$$I_n^a = \frac{1}{n(n-1)} \sum_i \sum_{i \neq j} \hat{u}_i \hat{u}_j K_{\hat{h}}(X_i^c, X_j^c) L(X_i^d, X_j^d, \hat{\lambda})$$

$$\hat{\sigma}_a^2 = \frac{2\hat{h}_1 \dots \hat{h}_q}{n(n-1)} \sum_i \sum_{i \neq j} \hat{u}_i^2 \hat{u}_j^2 K_{\hat{h}}(X_i^c, X_j^c)^2 L(X_i^d, X_j^d, \hat{\lambda})^2$$

For further detail, see Li and Racine (2007, Ch.12.1).

is the same as if the other components $f_s(z_s)$ for $j \neq s$ were known. In other words, $\widehat{f}_s(z_s)$ behaves in the same way as if it were a one-dimensional local nonparametric estimator and avoids the curse of dimensionality that plagues many nonparametric and semiparametric applications. From another perspective, PLR specifications can also be fitted with generalized additive models's techniques[101]. Wood (2000; 2006) proposes to decompose the flexible additive components in a finite sum of spline terms and to apply penalized least squares combined with a cross-validation to control for the smoothness of the functions. More precisely, let each $f_j(.)$ component be represented by $f_j(z_j) = \sum_{k=1}^{K} \beta_k g_k(.)$, where $g_k(.)$ is a family of K spline basis functions, and let the penalizing roughness term be $\int [f''(z_j)]^2 dz_j$, therefore equation (3.34) can be estimated by minimizing the following expression:

$$\min_{\alpha_r, \beta_k, \lambda_j} \sum_i \left(y_i - \sum_{r=1}^{R} \alpha_r x_{r,i} - \sum_{j=1}^{J} f_j(z_{j,i}) \right)^2 + \sum_{j=1}^{J} \lambda_j \int f_j''(z_j)^2 dz_j$$

where the $f_j(.)$ terms are replaced by their spline decomposition and the λ_js are determined by generalized cross-validation. In this paper, specification (3.34) is explored with both marginal integration and penalized smoothing splines. Given that both techniques drive to similar patterns, only the spline results are reported for each estimated function $\widehat{f}_j(x_j^c)$, with $j = 1, 2, 3$, and 95-percent bayesian[102] confidence intervals.

Finally, the functional restrictions in the parametric model (3.33) and the nonparametric additive hypothesis of the PLR equation (3.34) are fully relaxed by estimating

$$GP_{i,t} = f(\mathbf{x}^d, \mathbf{x}^c) + \epsilon_{i,t} \tag{3.35}$$

where $\mathbf{x}^d = [D_i, \tilde{D}_t]$ are the usual discrete regressors but with $\tilde{D}_t$ defined as a single discrete trend factor and $\mathbf{x}^c = [\log P_{i,t-T}, \log Y_{i,t-T}, GY_{i,t}]$ are the continuous explanatory factors. Racine and Li (2004) have recently proposed a new method to estimate nonpara-

---

[101] See Stone (1985) or the monographs by Hastie and Tibshirani (1990) or Gu (2002) among others.

[102] See Wood (2006, Ch.4.8 and 4.9) for further details.

metric regressions with mixed independent variables[103]. These authors also emphasize that using least squares cross-validation to determine the bandwidths allows to automatically discriminate between relevant and irrelevant regressors[104].

The relationship between the continuous predictors and the response in non or semiparametric regressions is usually reported graphically. Consequently, the results for specifications (3.34) and (3.35) are presented with partial regression plots. For the semiparametric model, given that the shapes for the nonparametric additive terms are similar for the pooled, OECD or non-OECD countries up to an additive constant, only the pooled relationships are reported with the parametric benchmarks. For the fully nonparametric setting, interactions may drive to specific patterns, depending on the OECD or non-OECD status. For the latter model, we follow Maasoumi et al. (2007): if we wish to present the nonparametric regression of $GP_{i,t}$ on the continuous regressors $\mathbf{x}^c$ for OECD countries, we plot

- $GP_{i,t}$ versus $E(GP_{i,t} \mid D_i = 1, \tilde{D}_t^\star, \log P_{i,t-T}, \log Y_{i,t-T}^\star, GY_{i,t}^\star)$,
- $GP_{i,t}$ versus $E(GP_{i,t} \mid D_i = 1, \tilde{D}_t^\star, \log P_{i,t-T}^\star, \log Y_{i,t-T}, GY_{i,t}^\star)$ and
- $GP_{i,t}$ versus $E(GP_{i,t} \mid D_i = 1, \tilde{D}_t^\star, \log P_{i,t-T}^\star, \log Y_{i,t-T}^\star, GY_{i,t})$,

where the upper star '$\star$' indicates that the variable is kept at its median level. The same method is used for non-OECD countries. Finally, note that the regression tables include estimates of the flexible components' gradients of the pooled non/semiparametric regressions at their median level[105].

3.3.3 Growth regressions

Table 3.1 contains the regression results for the panel regression with 10-year series. The results with 5-year series are fairly similar and are reported in the appendix (see Table 3.2).

---

[103] The estimator is given by:

$$\hat{g} = \frac{n^{-1}\sum_n Y_i K_\gamma(\mathbf{x}, \mathbf{X}_i)}{n^{-1}\sum_n K_\gamma(\mathbf{x}, \mathbf{X}_i)}, \text{ where } \mathbf{x} = [x^c, x^d] \text{ and } \mathbf{X}_i = [X_i^c, X_i^d]$$

$$K_\gamma(\mathbf{x}, \mathbf{X}_i) = W_h(x^c, X_i^d) L(x^d, X_i^d, \lambda), \text{ where } \gamma = (h, \lambda)$$

See Li and Racine (2007, Ch.4.4) for further details.

[104] Irrelevant regressors will be oversmoothed and the relationship will become flat.

[105] The median of $\{\tilde{D}_t^\star, \log P_{i,t-T}^\star, \log Y_{i,t-T}^\star, GY_{i,t}^\star\}$ are equal to $\{1980, -0.772, 7.878, 0.018\}$ for the 5-year series and $\{1980, -0.824, 7.784, 0.019\}$ for the 10-year series.

Table 3.1: Pollution growth regressions. Panel results with 10-year data. Period 1960-2000.

	Parametric models					Non/semipa. models	
	Ordinary LS			Nonlinear LS		PLR fit$^{(a)}$	NP fit$^{(b)}$
Variables	(A)	(B)	(C)	(D)	(E)	(F)	(G)
constant	0.040***	-0.082**	-0.044***	-0.026	-0.027	0.044***	-
d1970	-0.014**	-0.011**	-0.015***	-0.010**	-0.015**	-0.014**	-0.002***
d1980	-0.047***	-0.035***	-0.045***	-0.032***	-0.045***	-0.045***	-0.030***
d1990	-0.036***	-0.028***	-0.027***	-0.025***	-0.025***	-0.025***	-0.020***
OECD	0.005	-0.010**	-0.005	-0.002	-0.003	-0.002	0.029
$P_{i,t-T}$ (β)	-0.005***	-0.001*	-0.011***	-0.009***	-0.010***	-0.009***	-0.014**
$Y_{i,t-T}$ (δ_1)		0.013***		0.006***			
$Y_{i,(t-10)-T}$ (δ_1)			0.009*		0.006***	0.008	0.011*
$GY_{i,t-T}$ (δ_2)		0.611***		0.647***			
$GY_{i,(t-10)-T}$ (δ_2)			0.638***		0.643***	0.733***	0.853***
N	388	388	388	388	388	388	388
R2 adj	0.24	0.37	0.31	0.37	0.31	0.35	0.46
F-stat	26***	34***	26***				
Heterosced.$^{(c)}$	27***	39***	38***	40***	38***		
Wald $\delta_1 = -\beta\delta_2$$^{(d)}$		1.58	0.53				
P(Specific.)$^{(e)}$	0.000	0.000	0.010				

Notes: ***, ** and * denote the 1%, 5% and 10% significance levels. All computations made in *R.2.6.0.* (a): The pooled PLR estimation (1F) is computed with the *gam* function from the *mgcv, v.1.3-29* package with the default options. Approximate p-values are given for the smooth functions under the null that each smooth term is zero. Gradients for the nonparametric continuous components are computed at their median level. (b): The pooled nonparametric mixed fit (1G) is estimated with the *npbwreg* function from the *np, v.0.14-2* package with options *bwmethod="ls.cv", regtype="ll", ukertype="liracine", nmulti=50.* Significance levels are based on wild bootstrap from function *npsigtest.* The time variable is here a single discrete trend factor. Gradients for the nonparametric components are computed at their median level. (c) : 'Heterosced.' is the heteroscedasticity LM-test by Breusch and Pagan (1979), computed with the variance estimator proposed by Koenker (1981), robust to departure from normality, see Greene (2003, Ch.6.5.) for a short presentation. The latter statistic is χ^2-distributed, with d.f. = nb. of regressors (constant excluded). (d): The test-statistic for $\delta_1 = -\beta\delta_2$ follows a standard normal and its square corresponds to the Wald statistic. (e): 'P(Specific.)' stands for the probability associated to the nonparametric specification test by Hsiao et al. (2007) for continuous and discrete data models (see the function *npcmstest*, package *np, v.0.14-2*). The latter probability is based on 200 bootstrap's replications.

Column (A) tests absolute β-convergence within the linear specification. Columns (B) and (C) are OLS parametric fits which represent β-convergence regressions, conditional upon the levels and growth rates of GDP, the latter column being Instrumental Variables (IV henceforth) estimation which controls for potential endogeneity bias. Note that columns (B) and (C) also represent unconstrained linear estimates of specification (3.33) while columns (D) and (E) are their constrained counterparts, estimated with nonlinear least squares. Columns (F) and (G) show the pooled IV semiparametric and nonparametric estimates. Graphical devices complete the latter results, the PLR patterns are drawn on Figure (3.1) while Figure (3.2) depicts the partial regression plots for the nonparametric mixed models.

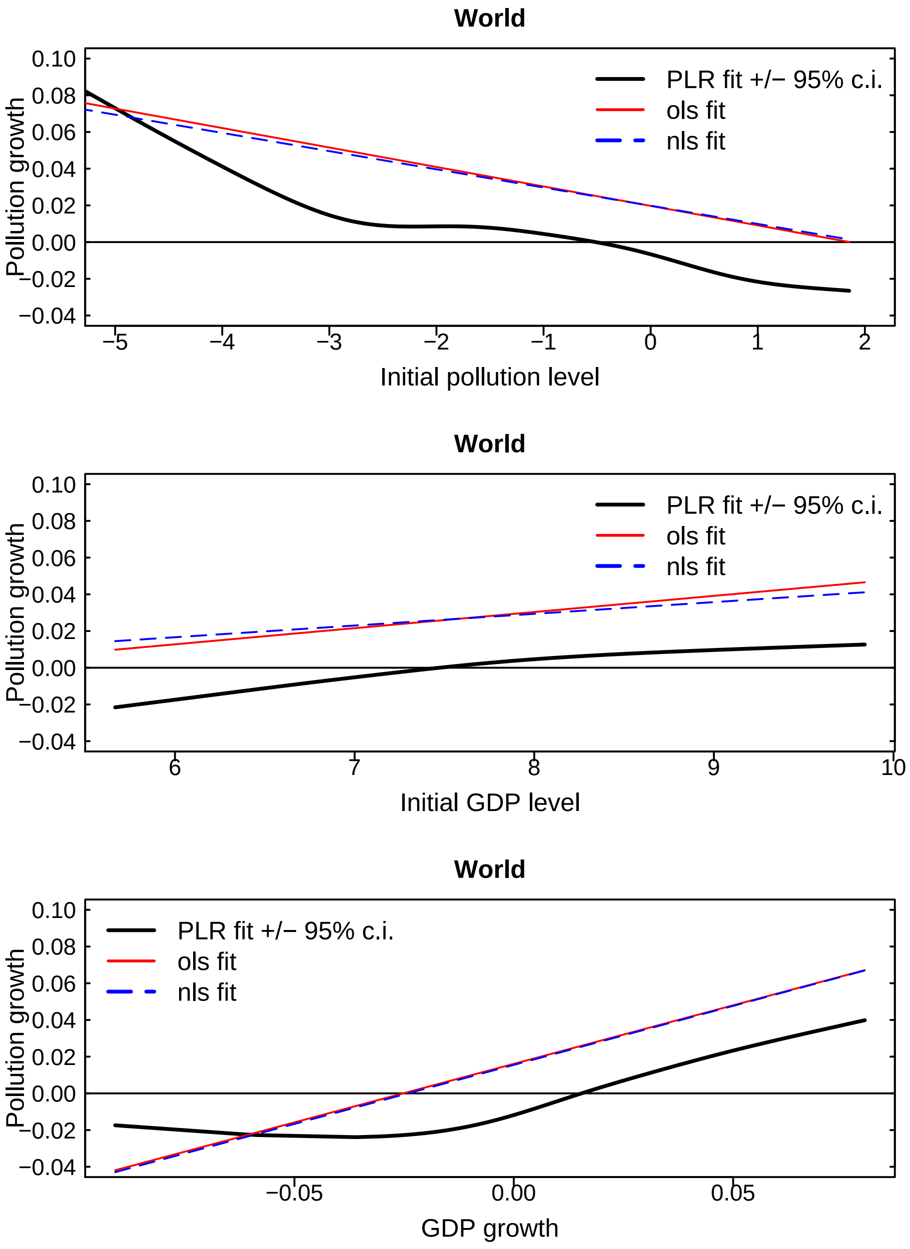
We notice in Table 3.1 that all parametric regressions display heteroscedastic errors as the null of homoscedasticity is overwhelmingly rejected with a robust version of the LM-test by Breusch and Pagan (1979). The coefficients' standard deviation for the parametric fits are White-corrected while the significance test for the semiparametric and nonparametric models relies on a bayesian approach and wild bootstrapping[106] respectively. We can also see that taking into account the GDP variables (Y and GY) improves the explanatory power of the absolute β-convergence regression (A) as its adj. R2 increases from 0.24 to over 0.30 in all the multivariate models. For the conditional specifications (B) and (D), the OLS and NLS estimates yield similar results and the same holds for their IV counterparts (C) and (E). Indeed, we cannot reject the null hypothesis that the constraint $\delta_1 = -\beta\delta_2$ is verified in the OLS models (B) and (C) at conventional cutoffs and coefficient $\hat{\delta}_1$ is positive and significant. Note also that $\delta_1^{NLS} = -\beta^{NLS}\delta_2^{NLS} = 0$ is rejected in the nonlinear fits (D) and (E) as the coefficient associated with the $Y_{i,t-T}$ variable is significant and positive. Therefore, the parameters' constraint imposed by the theoretical model is verified in the parametric fits. However, when we apply the specification test by Hsiao et al. (2007) to the OLS fits, we can see at the bottom of columns (A) to (C) that the parametric models are rejected at the 5% level, which means that they are misspecified. Consequently, the nonparametric approaches (F) and (G) are expected to depict nonlinearities as well as different patterns, specific to OECD membership for the fully nonparametric model. From now on, we focus our analysis on the PLR and the nonparametric fits.

The first striking result from the PLR estimates in column (F), Table 3.1, is that time dummies matter. Emissions' growth is affected by significant structural shocks over time. In particular, the pollution dynamics experiences a strong and significant downward adjustment during the first 10-year period following the oil-shocks. Secondly, there is evidence of conditional β-type convergence as the negative partial derivative at the median level of the $P_{i,t-T}$ variable suggests. Figure 3.1 confirms the downward sloping total effect

---

[106] See Racine (1997) or the function *npsigtest* from the *np.v.0-14-2* package in *R*.

Figure 3.1: Pollution growth partially linear IV regressions with 10-year data. Partial plots. Period 1960-2000.



Note: PLR regression estimates generated with the function *gam* from library *mgcv, v.1.3-29* in R. Marginal integration estimates drive to similar shapes.

of initial PCE levels on pollution growth over the whole range of the predictor. Note that the convergence speed is higher for low and large initial polluters than for intermediate ones. This result holds even when the PLR model omits the GDP regressors[107]. Regarding the effect of past (or initial) GDP levels on posterior pollution growth rates, the partial relationship is linear and positive but not significant[108]. Finally, we can also see on the bottom plot on Figure 3.1 that the relationship between economic growth and pollution growth has a flat portion when GDP growth is negative and increases afterwards. Note that the same scale has been deliberately used on all partial plots in order to ease the comparison of the relative contribution of each explanatory variable to pollution growth.

We now turn to the fully nonparametric estimates from model (3.35). Recall that the 'coefficients' reported in column (G), Table 3.1, are indeed derivatives evaluated at the median level of the explanatory factor, keeping the other regressors steady at their median level. The gradients over the whole range of each continuous predictor are depicted in the OECD *vs* non-OECD[109] partial plots on Figure 3.2. We notice that the time effects remain relevant in the nonparametric setting and that the derivatives for the significant regressors in column (G) are quite close to those reported in column (F) for the PLR model. However, Figure 3.2 highlights important differences between the OECD *vs.* non-OECD economies. Firstly, β-convergence in PCE is strong at low initial emissions levels for the non-OECD countries and the convergence speed decreases as $\log P_{i,t-T}$ increases. The contrary holds for OECD countries, *i.e.* we go from a flat relationship toward larger convergence as initial per capita emissions increase. Few observations are available at low initial pollution levels for the latter group, so the confidence interval is large in that range. Overall, given that conditional convergence is expected to occur between similar economies once we control for their economic performance, our results indicate that convergence is more likely to occur at larger pollution levels in the OECD subset and at lower ones for the rest of the world. Secondly, similar past GDP levels drive globally to much larger posterior pollution growth rates in the OECD group. More important is the fact that the partial $GP_{i,t}$ *vs.* $\log Y_{i,t-T}$ relationship is decreasing in the OECD group while it is upward sloping over the whole range of income for the non-OECD countries. Several factors may explain that OECD countries, which pollute more in per capita terms, tend to

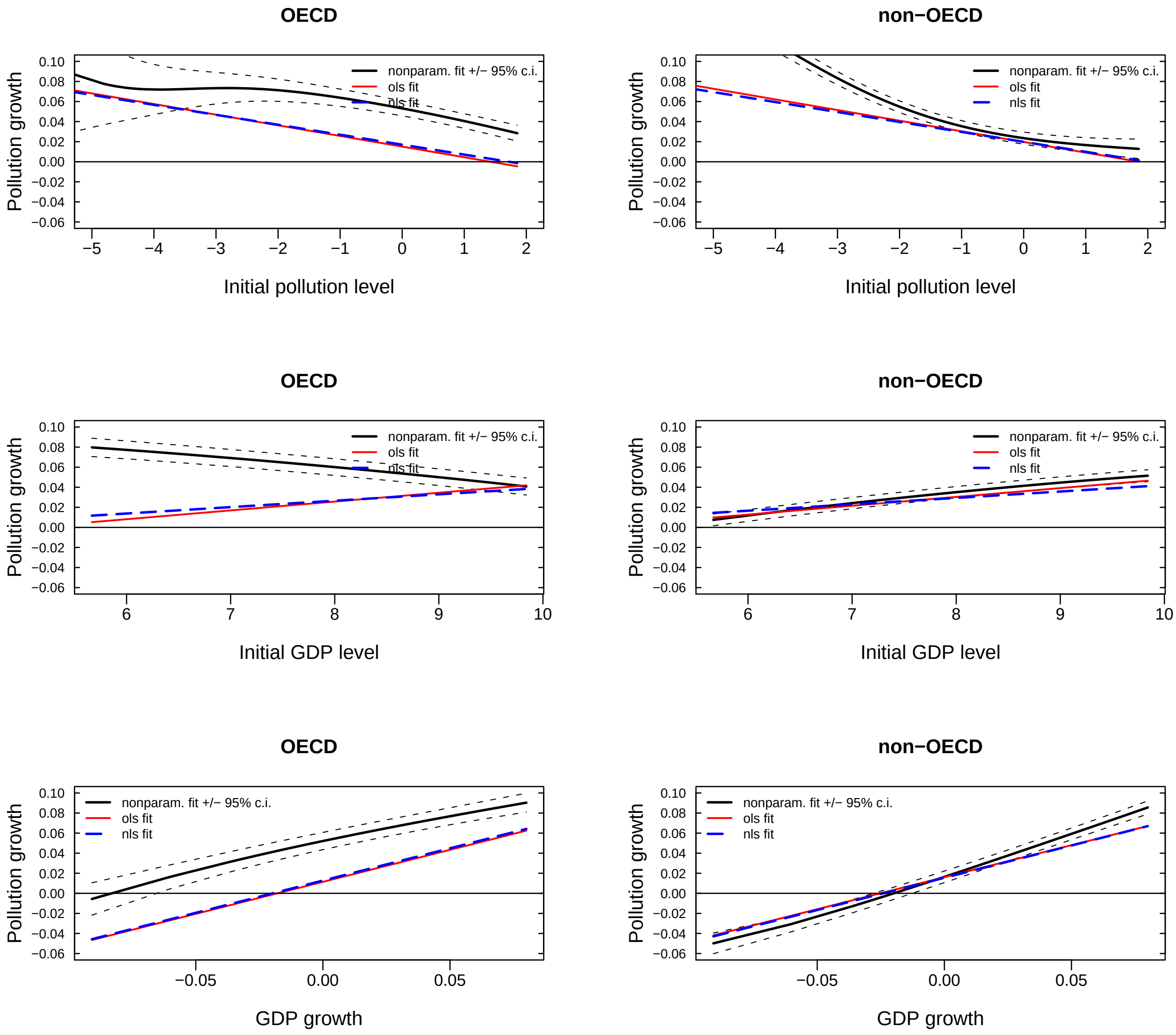
---

[107] Here, we refer to a model of the form $GP_{i,t} = \alpha_1 + \alpha_2 D_i + \alpha_{3,t} D_t + f(\log P_{i,t-T}) + \varepsilon_{i,t}$. The adj. R2 for the latter regression drops to 0.29.

[108] The associated p-value is 0.147.

[109] The median for the OECD dummy is zero, which correspond to the non-OECD group.

Figure 3.2: Pollution growth mixed nonparametric IV regressions with 10-year data. Partial plots. Period 1960-2000 period. OECD *vs* non-OECD countries.



Notes: Nonparametric regressions based on Racine and Li (2004). Partial regression estimates computed with the function *npregbw* from library *np, v.0-14-2* with *bwmethod="ls.cv", regtype="ll", ukertype="liracine", nmulti=50.* Confidence intervals are based on asymptotic standard errors.

reduce their pollution growth when they get richer: larger polluting inputs' productivity, more efficient abatement technologies, economic growth based on the accumulation of less polluting factors or more stringent environmental rules. The reduced form estimated in this paper is too general to allow the precise identification of the origin of the structural differences; it merely brings them into light. Finally, GDP growth and pollution growth remain positively linked for both OECD status. The fully nonparametric fits do not reproduce the flat portion found for this relationship in the PLR model. The last plot in Figure 3.2 also shows that for identical GDP growth rates, higher pollution growth prevails in the OECD countries. However, non-OECD countries tend to 'catch-up' as GDP growth rises. The reader can check in the Appendix, in Table 3.2 and Figure 3.6, that the regression results with 5-year series are pretty similar. Hence, the main patterns described previously remain valid even in shorter horizons.

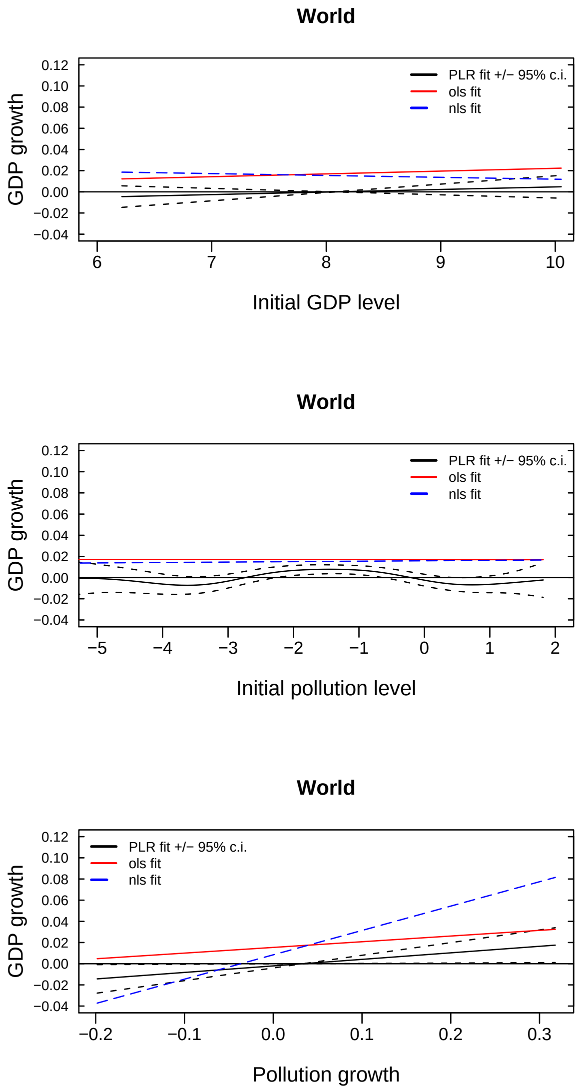
Additional exploratory analysis is carried out with the reduced form function investigated in this section by reversing the correlation structure between the dependent variable $GP_{i,t}$ and its GDP counterpart $GY_{i,t}$ in equation (3.32a). In this case, we get a reduced form that explores how GDP growth rates are affected by initial GDP levels, initial pollution levels and pollution growth[110]. The interest of that formulation is to check empirically to what extent β-convergence in per capita pollution is compatible with β-convergence in per capita GDP and if pollution variables have any impact on GDP growth. Indeed, exactly the same analysis conducted for pollution growth can be applied to GDP growth. The OLS estimates for the parametric models[111] being rejected at the 5% cutoff with the specification test by Hsiao et al. (2007), we concentrate on the PLR estimates as well as the fully nonparametric regressions displayed with their usual parametric benchmarks on Figures 3.3 and 3.4 respectively. We can see on the upper panel of Figure 3.3 that there is little evidence of β-convergence in per capita GDP for the pooled PLR estimates and a formal test indicates that the additive $f(\log Y_{i,t-T})$ term is not significant. This result is consistent with the income convergence literature. Regarding the effect of initial pollution levels on GDP growth, $f(\log P_{i,t-T})$ is significant but very flat around zero. We find no evidence that past pollution levels of PCE penalize/favor economic growth, whatever their level. This is not surprising as carbon dioxide is a global pollutant, which does not favor pollution control's initiatives from individual countries. Moreover, the only 'binding' international agreement on greenhouse gases, the Kyoto protocol, came into force after the time period covered by our panel. Regarding the partial GDP growth - pollution growth

---

[110] This is simply $GY_t = \bar{\bar{\gamma}} - \bar{\bar{\beta}}\log Y_{t-T} + \bar{\bar{\beta}}\bar{\bar{\lambda}}\log P_{t-T} + \bar{\bar{\lambda}} GP_t$, where the double bar symbol indicate different parameters than those in equation (3.32a). Notice that the constraint on the Y_{t-T} term in equation (3.32a) still holds, but for the P_{t-T} variable instead.

[111] We refer to estimates as in columns (A), (B) and (C) from Table 3.1.

Figure 3.3: GDP growth partially linear IV regressions with 10-year data. Partial plots. Period the 1960-2000.



Note: PLR regression estimates generated with the function *gam* from library *mgcv, v.1.3-29* in R. Marginal integration estimates drive to similar shapes.

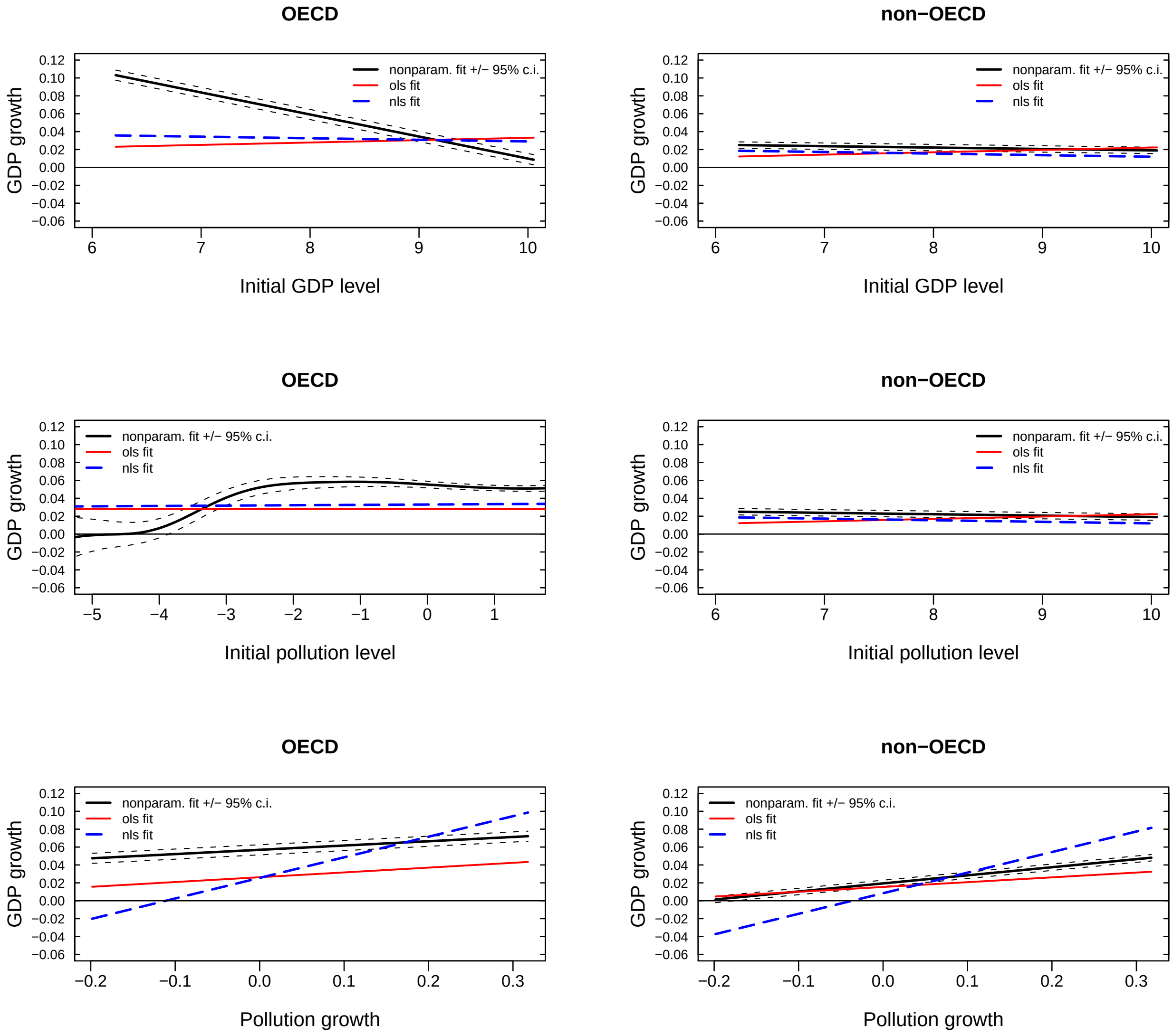
link, the increasing linear pattern found in the previous analysis with the PLR model also holds here and it is significant at the 5% level. Figure 3.4 carries out the same analysis, with fully nonparametric estimates distinguishing between OECD and non-OECD countries. While the results for the non-OECD results economies are identical to the PLR ones, the OECD group appears to depict (conditional) β-convergence in per capita GDP. This result is fully consistent with the growth literature. Regarding the other predictors, initial pollution levels are linked to larger GDP growth in the OECD countries in the center of the GDP support and flat away from it. No relationship is apparent for the non-OECD countries. Finally, a monotonic increasing pattern is found between GDP growth and pollution growth for both the OECD and non-OECD countries but larger GDP growth rates and a flatter slope is found for the OECD group.

Overall, these results corroborate the view that sustained economic growth yields sustained growth in per capita emissions of carbon dioxide. However, offsetting shocks captured by time dummies curve down the pollution dynamics. For the OECD countries, past larger GDP levels also favor lower pollution growth rates but the theoretical model does not allow the identification of any causal scheme. Beta-type convergence in carbon dioxide PCE, both absolute and conditional to the economic performance, is found and should favor the emergence of groups of countries heading toward similar PCE levels. Note that the latter result is compatible with β-convergence in per capita GDP for the OECD countries as well as β-divergence for the rest of the economies. This seems to indicate that, empirically, β-convergence in per capita pollution is compatible with either convergence or divergence in per capita income levels between countries. We find little evidence of any effect of pollution on economic growth. Large past levels in per capita CO_2 emissions do not affect GDP growth. However, this issue could be further investigated by employing a framework that allows pollution to enter as a factor of production using a total factor productivity approach as in Kalaitzidakis et al. (2008).

3.4 Conclusion

This paper investigates the income-pollution relationship with a β-type reduced form function derived from a growth model. Estimates based on parametric, semiparametric as well as fully nonparametric regressions are proposed for a large panel of world countries on per capita CO_2 emissions over the period 1960-2000. Our results, based on panel estimates with 10-years and 5-years growth rates, show that parametric forms are misspecified and do not account properly for nonlinearities and interactions between the explanatory factors.

Figure 3.4: GDP growth mixed nonparametric IV regressions with 10-year data. OECD *vs* non-OECD countries. Period 1960-2000.



Notes: Nonparametric regressions based on Racine and Li (2004). Partial regression estimates computed with the function *npregbw* from library *np, v.0-14-2* with *bwmethod="ls.cv", regtype="ll", ukertype="liracine", nmulti=50.* Confidence intervals are based on asymptotic standard errors.

We provide flexible estimates that outline negative time shocks on pollution growth, especially after the 70s oil-crisis, as well as important differences in the pollution dynamics between OECD and non-OECD countries. We also show that GDP growth goes hand in hand with emissions growth as they display a monotone increasing relationship in both the OECD and non-OECD groups. However, regarding the level effects, larger past GDP levels favor lower subsequent pollution growth only in the OECD countries. Note that this result does not imply a negative income-pollution relationship as pollution growth rates remain positive over the entire GDP range covered by our OECD sample. It seems that OECD countries have not reached the negative branch of the EKC curve for per capita carbon dioxide emissions as per capita carbon emissions still increase at a decreasing rate.

We find empirical evidence of β-convergence in per capita pollution levels, both absolute and conditional on GDP levels and growth, for OECD and non-OECD countries, with varying speed of convergence depending on the initial level of pollution. It appears that this configuration is empirically fully compatible with either β-type convergence or divergence in per capita GDP levels as the OECD countries display β-convergence in per capita income while the rest of the world diverges. Therefore, reducing/stabilizing income disparities is not a prerequisite to decreasing/stabilizing pollution gaps between the world countries and it gives support to the notion that achieving pollution targets can take place independently of any hypothetical equalization of income levels. This is important from a policy point of view and calls for further research with other pollutants, such as SO_2 and NO_2 for example, as data become available to ascertain the generality of these empirical results.

Finally, it is important to note that our paper focuses arbitrarily on the reduced form function proposed by Alvarez et al. (2005), as it represents the first attempt to provide a reduced form function from a Ramsey model with pollution stock that results in estimable β-regression for pollution. Incorporating damages from pollution flow and/or stock in the consumers' utility function within the Ramsey-Koopman-Cass framework has been addressed by Van der Ploeg and Withagen (1991). The link with the β-convergence framework is a matter of present research.

Acknowledgments

The draft of this paper has been written while visiting the University of Calgary during the summer semester 2007. I sincerely acknowledge the hospitality of both Prof. Taylor and Prof. Boyce as well as Pearl's kindness. I am also grateful to the Swiss National Research Foundation (under grant FN PBNE1-114686).

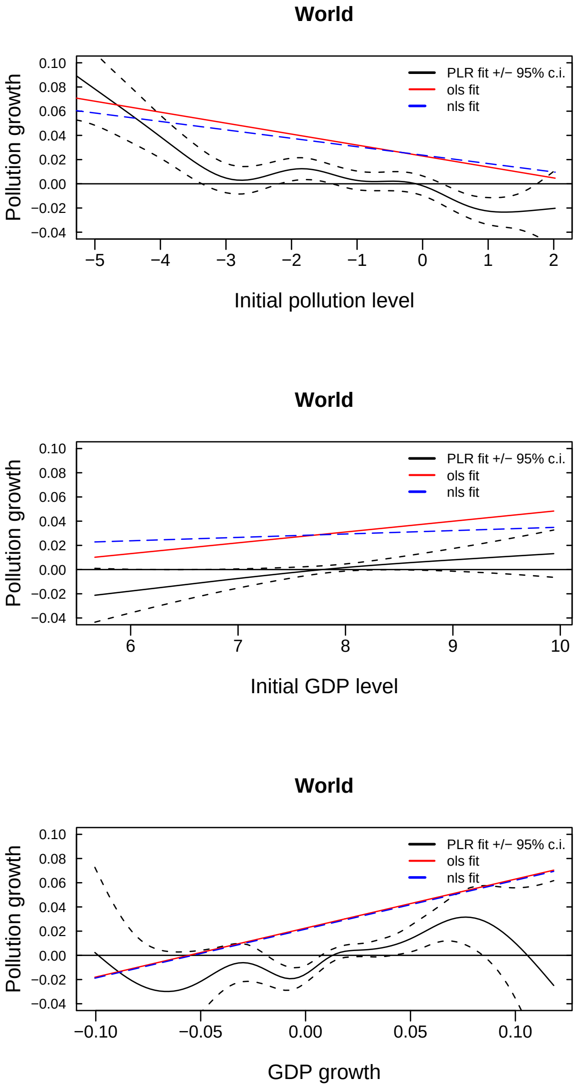
3.5 Appendix

Table 3.2: Pollution growth regressions. Panel results with 5-year data. Period 1960-2000.

	Parametric models					Non/semipa. models	
	Ordinary LS			Nonlinear LS		PLR fit$^{(a)}$	NP fit$^{(b)}$
Variables	(1A)	(1B)	(1C)	(1D)	(1E)	(1F)	(1G)
constant	0.038***	-0.105***	-0.044	-0.032	0.004	0.040***	-
d1965	0.000	0.000	0.003	- 0.003	0.003	0.004	-0.004**
d1970	-0.009	-0.010	-0.009	-0.009	-0.009	-0.008	-0.001**
d1975	-0.011	-0.009	-0.011	-0.007	-0.010	-0.009	-0.002**
d1980	-0.057***	-0.043***	-0.057***	-0.040***	-0.055***	-0.055***	-0.034**
d1985	-0.033***	-0.024***	-0.033***	-0.022***	-0.030***	-0.031***	-0.016**
d1990	-0.039***	-0.025***	-0.033***	-0.022***	-0.030***	-0.031***	-0.019**
d1995	-0.033***	-0.030***	-0.030***	-0.027***	-0.027***	-0.028***	-0.019**
OECD	0.004	-0.013***	-0.004	-0.005	0.001	-0.000	0.023*
$P_{i,t-T}$ (β)	-0.004***	-0.012***	-0.009***	-0.009***	-0.007***	-0.001***	-0.005**
$Y_{i,t-T}$ (δ_1)		0.016**		0.006***			
$Y_{i,(t-10)-T}$ (δ_1)			0.009**		0.003**	0.008	0.006***
$GY_{i,t-T}$ (δ_2)		0.658***		0.701***			
$GY_{i,(t-10)-T}$ (δ_2)			0.405***		0.403***	0.375***	0.289***
N	776	776	776	776	776	776	776
R2 adj	0.14	0.27	0.16	0.26	0.16	0.20	0.29
F-stat	15***	27***	17***				
Heterosce.$^{(c)}$	49***	67***	54***	68***	54***		
Wald $\delta_1 = -\beta\delta_2$$^{(d)}$		‡2.33**	1.42				
P(Specific.)$^{(e)}$	0.000	0.000	0.005				

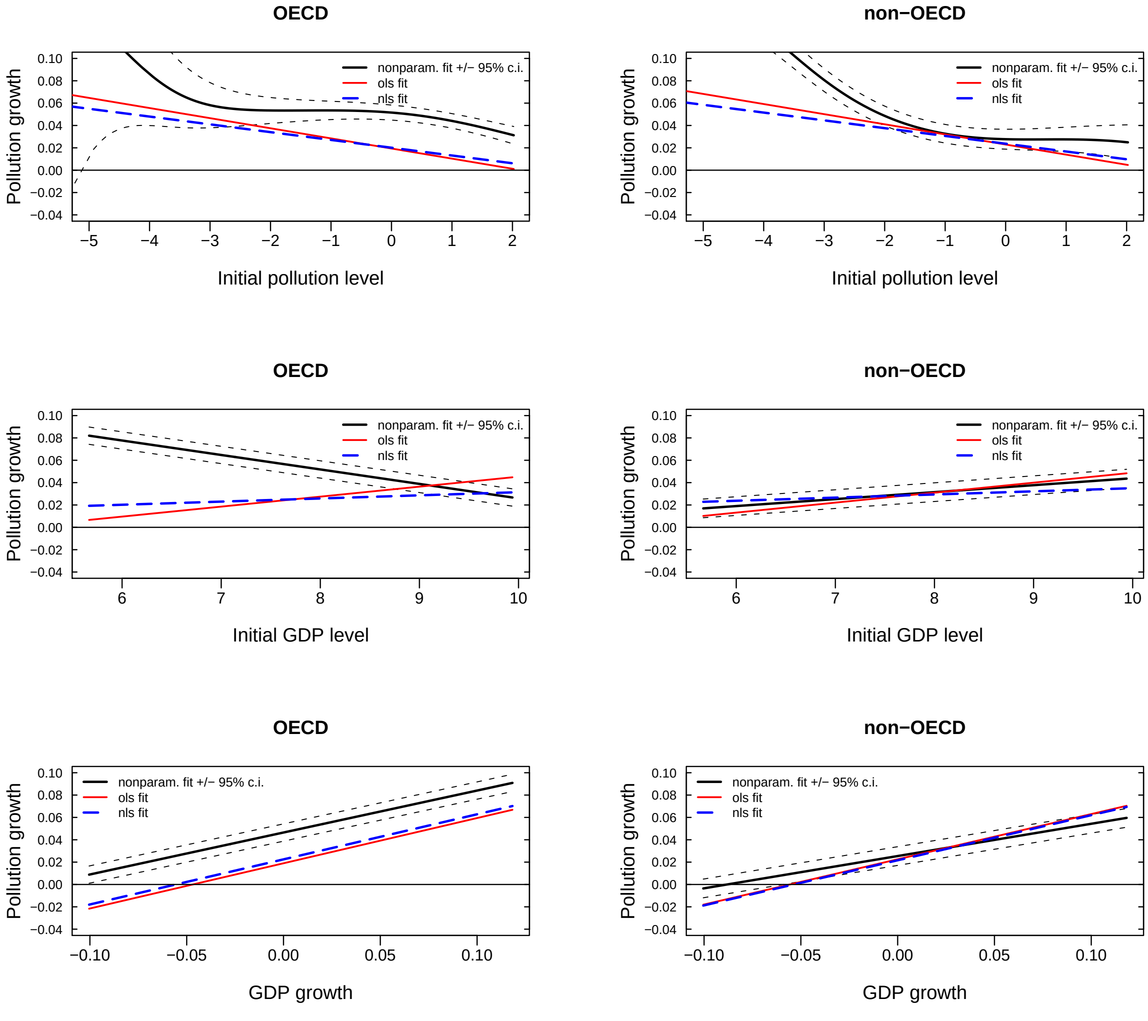
Notes: ***, ** and * denote the 1%, 5% and 10% significance levels. All computations made in *R.2.6.0.* (a): The pooled PLR estimation (1F) is computed with the *gam* function from the *mgcv, v.1.3-29* package with the default options. Approximate p-values are given for the smooth functions under the null that each smooth term is zero. Gradients for the nonparametric continuous components are computed at their median level. (b): The pooled nonparametric mixed fit (1G) is estimated with the *npbwreg* function from the *np, v.0.14-2* package with options *bwmethod="ls.cv", regtype="ll", ukertype="liracine", nmulti=50.* Significance levels are based on wild bootstrap from function *npsigtest.* The time variable is here a single discrete trend factor. Gradients for the nonparametric components are computed at their median level. (c) : 'Heterosced.' is the heteroscedasticity LM-test by Breusch and Pagan (1979), computed with the variance estimator proposed by Koenker (1981), robust to departure from normality, see Greene (2003, Ch.6.5.) for a short presentation. The latter statistic is χ^2-distributed, with d.f. = nb. of regressors (constant excluded). (d): The test-statistic for $\delta_1 = -\beta\delta_2$ follows a standard normal and its square corresponds to the Wald statistic. (e): 'P(Specific.)' stands for the probability associated to the nonparametric specification test by Hsiao et al. (2007) for continuous and discrete data models (see the function *npcmstest*, package *np, v.0.14-2.* The latter probability is based on 200 bootstrap's replications.

Figure 3.5: Pollution growth partially linear IV regressions with 5-year data. Partial plots Period 1960-2000.



Note: PLR regression estimates generated with the function *gam* from library *mgcv, v.1.3-29* in R. Marginal integration estimates drive to similar shapes.

Figure 3.6: Pollution growth mixed nonparametric IV regressions with 5-year data. Partial plots. OECD *vs* non-OECD countries. Period 1960-2000.



Notes: Nonparametric regressions based on Racine and Li (2004). Partial regression estimates computed with the function *npregbw* from library *np, v.0-14-2* with *bwmethod="ls.cv", regtype="ll", ukertype="liracine", nmulti=50.* Confidence intervals are based on asymptotic standard errors.

General conclusion

Global warming is probably one of the biggest challenges that will be faced by mankind in the future. The scope of the problem is truly global, as it involves all countries in the world and almost every dimension of human activity. Finding appropriate and implementable solutions will probably take decades and require huge efforts by politicians, academics and entrepreneurs. In the face of these considerations, any attempt to make a contribution to the domain may appear as a rash and presumptuous exercise.

More precisely and modestly, this dissertation is an attempt to better understand the lack of robustness of empirical findings in the economic literature. The focus is on the link between economic activity and polluting air emissions, which has long been recognized as both crucial and very complex. It is fair to say that today many issues such as the existence of an environmental Kuznets curve or the putative converge of emissions per capita remain largely unsettled. In part, this lack of robustness may be attributable to poor data or missing control variables, and the empirical essays collected here do their best to take these factors into account. However, what I have been trying to illustrate is the importance of a third and often neglected factor: misspecification of the underlying relationship. In letting the data dictate the relationship rather than imposing it, recent advances in nonparametric techniques have offered me a powerful framework, not only to check for this bias, but also to correct it.

The results of this revised evidence are summarized below, first in detail and chapter by chapter, to allow for a precise discussion of the remaining caveats, then broadly presenting the main lessons and future avenues of research.

Chapter 1 proposes a simple test procedure to check the consistency of the income-pollution link found in the previous empirical literature with static fixed effects panel data models. I show that some implicit homogeneity assumptions made on the time and/or spatial dimension of the panel may drive to misleading conclusions regarding sustainability. The income-pollution equation is apparently straightforward to estimate. However, in a panel context, it is fundamental to check to what extent the fixed effects or pooled estimates hold in the dimension of interest. For that purpose, nonparametric methods are of help to avoid misspecification bias, but they do not prevent the researcher from checking some fundamental homogeneity assumptions on the cross-sectional or time dimensions of the panel. In my empirical application with a regionwide panel for Spain, the income-pollution relationship is found to vary widely across provinces for the four air pollutants and period investigated (CH_4, CO, CO_2 and NMVOC emissions over 1990 to 2002) and the U-inverted shapes depicted by static fixed effects or pooled models, estimated with parametric or semi/nonparametric regressions techniques, may be in total contradiction with the income-pollution provincial trends. This methodological issue is key to get meaningful and consistent estimates in the EKC context: firstly because as it is likely to affect even more cross-country rather than cross-regional studies, and secondly because it can results in mistaken policy recommendations or sustainability conclusions. Toward this end, any method allowing greater functional flexibility in the individual dimension of panels should be preferred to pooled estimators. From an economic perspective, the analysis uncovers the presence of sustainable forces (pollution-reducing technical progress, protective/defensive environmental policy measures, . . .) that help to curve down the positive scale effect of economic growth on atmospheric pollution in Spain, but only for local air pollutants, such as CO and NMVOC, which have direct and immediate effects on health and/or economic activities. There is no empirical evidence that economic growth is sustainable regarding human-induced emissions of GHG gases (CH_4, CO_2).

Chapter 2 employs distributional and univariate time-series methods to examine convergence in carbon dioxide per capita emissions. Contrary to total emissions, which are strongly upward trended in most countries and at the world level, per capita emissions have stabilized since the 70s in many countries and for the world. This empirical feature opens the door to the existence of clubs of countries, converging toward similar stable per capita levels of carbon emissions, that may serve as acceptable targets for most of the world economies and therefore ease a large consensus to achieve international commitments on carbon emissions. I find that comparing cross-sectional distributions of per capita emissions over time, with robust scale and shape distributional measures as well as distributional tests, helps to better understand the overall dynamics of per capita emissions' gaps. I identify flattening, unimodal and right skewed cross-sectional world distributions of per capita CO_2 emissions until the 70s oil shocks and stabilized distributions afterwards. A variety of specific patterns for different subsets of countries are also found, on the basis of simple grouping criteria (countries with similar initial pollution levels or final income per capita, geographic neighbors, economically integrated areas). Only very few groups (OECD and UE15 countries) exhibit decreasing dispersion and increasing peakedness over time, which would validate distributional convergence. I neither find significant departures from unimodality over time with the selected multimodality test. The latter test rejects strong multipolarization phenomena, and therefore the existence of (multiple and) well-defined convergence clubs. By contrast, unit root tests combined with the notion of pairwise convergence of Evans (1998) do suggest significant convergence at the world level as well as within many country groupings, mainly after the 70s oil-shocks. The conclusions of the distributional and the stochastic convergence approaches may appear conflicting, but they are not. Indeed, my time series analysis tests nonconvergence across all the countries against the alternative that some countries converge. It allows converging and nonconverging economies to exist within the same (sub)sample, without giving rise to significant multimodality in the cross-section distributions. Overall, these results suggest that, although no strong and global convergence mechanisms are at work for a majority of countries, a significant proportion of them experiences decreasing differences in their per capita levels of carbon emissions. From a policy perspective, this may help the emergence of an acceptance of reasonable emissions' targets in per capita terms, as several major developing countries perceive the latter objectives fairer than the grandfathering scheme adopted under Kyoto Protocol. From the methodological side, further research is needed to reconcile unit root tests with the distributional approach. Among the possible extensions, the distributional analysis could gain from: improving or extending the grouping criteria; exploring the (conditional) distributions with the nonparametric methods for mixed data of Li and Racine (2003) and Hall et al. (2004); considering alternative multimodality tests; measuring more directly per capita pollution disparities through a simple Gini index, entropy measures such as those used in Maasoumi et al. (2007) or the polarization measures used in Ezcurra (2007). Unit root tests setting convergence as the null hypothesis instead of nonconvergence could help in reconciling the distributional analysis with the stochastic convergence approach while allowing endogenous breaks in the unit root tests could also shed a new light on our results. Finally, convergence in total emissions or emissions per unit of output/value remains a virgin field that needs to be explored.

Chapter 3 uses different regression techniques to estimate a relationship that links pollution growth in CO_2 per capita emissions to past levels of CO_2 emissions, past GDP levels and GDP growth. This particular specification is a reduced form function derived from a theoretical growth model of capital accumulation, known as the "Ramsey-Koopman-Cass" model, amended with pollution by Alvarez et al. (2005). I revisit the pollution convergence in carbon emissions in this β-regression framework and, by reverting the correlation scheme, I also explore income convergence, conditional on initial levels of pollution, GDP and pollution growth. Given the panel nature of the data, the empirical specification includes time and OECD dummies. The main innovation of the paper is to explore for the first time the income-pollution link with fully mixed nonparametric regressions (*i.e.* fully nonparametric regressions which include discrete and continuous explanatory variables). The advantage of the latter formulation is that it captures nonlinearities and interactions between the explanatory factors, without the need of specifying them *a priori*. The results are compared with standard parametric and semiparametric specifications and indicate that parametric models are in general misspecified. I highlight important interactions between the main explanatory factors and OECD status, that are implicitly captured by the fully nonparametric specification. The major economic result is that even if positive GDP growth induces positive growth in per capita carbon emissions, defensive effects are also present in that greenhouse gas' dynamics. One is captured through beta-type pollution convergence component (past levels of emissions) and affects both the OECD and the non-OECD countries. The other one is linked to past GDP levels and is only detected in OECD countries. As these results are found in a context of β-convergence in per capita income for OECD countries but β-divergence for non-OECD countries, the empirical evidence suggests that decreasing/stabilizing pollution gaps between countries may be empirically compatible with narrowing, stabilizing or even increasing gaps in per capita income between nations. Furthermore, note that no adverse effect is found of past carbon emissions' levels or pollution growth on GDP growth. Regarding theory, this paper focuses arbitrarily on the reduced form function proposed by Alvarez et al. (2005) as it represents the first attempt to provide a reduced form from a Ramsey model with pollution stock that results in an estimable β-regression for pollution. Incorporating damages from pollution flow and/or stock in the consumers' utility function within the Ramsey-Koopman-Cass framework has been addressed by Van der Ploeg and Withagen (1991). However, a satisfactory link with the β-convergence framework remains to be established.

Final considerations. Although clearly perfectible, the empirical work collected in these three chapters sends a double warning to the applied scientists interested in the economy-environment links.

On the one hand, misspecification biases do matter. A common feature of my three essays is to show that nonparametric methods in general clearly outperform traditional investigation techniques. Many of the relationships found in that field would undoubtedly gain from being revisited with non or semiparametric methods, as the great flexibility they allow, compared to parametric approaches, often sheds a new light on apparently well-established results.

On the other hand, apart from methodological issues, the broad picture that emerges in terms of empirical findings is that there are both reasons to worry and ground for hope regarding the sustainability of air pollution patterns. The different chapters illustrate which dimensions are critical in shaping the final outcome. My first chapter shows that air pollution patterns are sustainable across Spanish regions as long as the specific pollutant is a local or a regional one. However, if the pollutant is global, as in the case of CO_2, the dynamics is more worrying, and exhibits a strong positive relationship between pollution and economic growth. It is thus legitimate to pursue the analysis further on this particular global pollutant, and at the world-wide level, as done in the later chapters. After careful examination, three empirical regularities deserve a particular mention, inasmuch they may shape further research efforts. First, the oil shocks changed the distribution of per capita CO_2 emissions across countries, which became flatter and more stable over time. This suggests that further shocks in oil prices in future decades will be a critical determinant of world-wide emissions. Second, although no clear-cut pattern emerges at the world-wide level, there seems to be some groups of countries within which per capita emissions clearly tend to converge towards common levels. These convergence forces are probably linked with income per capita convergence, as in the case of OECD countries. However, and this is the third stylized fact, other forces are also at work. It has been shown that the dynamics of per capita emissions is independent of the dynamics of income per head, and that other factors matter, such as trade, geography and historical conditions. A better understanding of these additional dimensions may contribute to improve the efficiency of global policies aiming at reducing global CO_2 emissions.

List of Tables

List of Figures

Bibliography

Aldy, J. (2005). An environmental kuznets curve analysis of u.s. state-level carbon dioxide emissions. The Journal of Environment Development, 14:48–72.

Aldy, J. (2006). Per capita carbon dioxyde emissions: Convergence or divergence. Environmental and Resource Economics, 33:533–555.

Aldy, J. (2007). Divergence in state-level per capita carbon dioxide emissions,. Land Economics, 83(3):353–369.

Alvarez, F., Marrero, G., and Puch, L. (2005). Air pollution and the macroeconomics across european countries,. Documento de Trabajo 2005-10, Fundación de Estudios de Economía Aplicada, Madrid.

Andreoni, J. and Levinson, A. (2001). The simple analytics of the environmental kuznets curve. Journal of Public Economics, 80:269–286.

Arrellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte-carlo evidence and an application to employment equations. Review of Economic Studies, 58:277–297.

Azomahou, T., Laisney, F., and Nguyen Van, P. (2006). Economic development and co$_2$ emissions: A nonparametric panel approach. Journal of Public Economics, 90:1347–1363.

Baltagi, B., Hidalgo, J., and Li, Q. (1996). A nonparametric test for poolability using panel data. Journal of Econometrics, 75:345–367.

Barrasi, M., Cole, M., and Elliott, R. (2008). Stochastic divergence or convergence of per capita carbon dioxide emissions: Re-examining the evidence. Environmental and Resource Economics, 40(1):121–137.

Barro, R. and Sala-i Martin, X. (2004). Economic Growth. The MIT Press, Massachusetts, second edition.

Ben-David, D. (2000). Trade, growth and disparities among nations. Special Study 5, World Trade Organization, Geneva.

Bernard, A. and Durlauf, S. (1995). Convergence in international output. Journal of Applied Econometrics, 10:97–108.

Bernard, A. and Durlauf, S. (1996). Interpreting tests of the convergence hypothesis. Journal of Econometrics, 71:161–173.

Bertinelli, L. and Strobl, E. (2005). The environmental kuznets curve semi-parametrically revisited. Economics Letters, 88:350–357.

Bliss, C. (1998). Galton's fallacy and economic convergence. Oxford Economic Papers, 51:4–14.

Bodansky, D., Chou, S., and Jorge-Tresoni, C. (2004). International climate efforts beyond 2012: a survey of approaches. Working paper, Pew Center on Global Climate Change.

Boehringer, C. and Welsch, H. (1999). C&c - contraction and convergence of carbon emissions: The economic implications of permit trading. Discussion paper 99-13, ZEW.

Breusch, T. and Pagan, A. (1979). A simple test for heteroscedasticity and random coefficient variation. Econometrica, 47:1287–1294.

Brock, W. and Taylor, M. (2004a). Economic growth and the environment: a review of theory and empirics. Working paper 10854, NBER.

Brock, W. and Taylor, M. (2004b). The green solow model. Working paper 10557, NBER.

Brys, G., Hubert, M., and Struyf, A. (2004). A robust measure of skewness. Journal of Computational and Graphical Statistics, 53(4):996–1017.

Brys, G., Hubert, M., and Struyf, A. (2006). Robust measures of tail weight. Computational Statistics and Data Analysis, 50:733–759.

Cannon, E. and Duck, N. (2000). Galton's fallacy and economic convergence. Oxford Economic Papers, 52:415–419.

Carlino, G. and Mills, L. (1993). Are us regional incomes converging? Journal of Monetary Economics, 32:335–346.

Carron-i Silvestre, J.-L., del Barrio-Castro, T., and Lopez-Bazo, E. (2005). Breaking the panels: an application to the gdp per capita. The Econometrics Journal, 8:159–175.

Choi, I. (2001). Unit root tests for panel data. Journal of International Money and Finance, 20:249–272.

Copeland, B. and Taylor, M. (2003). Trade and the Environment. Princeton University Press, New Jersey.

Davison, A. and Hinkley, D. (1997). Bootstrap Methods and their Applications. Cambridge Series in Statistical and Probabilistic Mathematics, 1st edition.

de Bruyn, J., van den Bergh, J., and Opschoor, J. (1998). Economic growth and emissions: reconsidering the empirical basis of environmental kuznets curves. Ecological Economics, 25:161–175.

De Groot, H., Withagen, C., and Minliang, Z. (2004). Dynamics of china's regional development and pollution: an investigation into the environmental kuznets curve. Environment and Development Economics, 9:507–537.

de la Fuente, A. (2000). Mathematical Methods and Models for Economists. Cambridge University Press, 1st edition.

Dijkgraaf, E. and Vollebergh, R. (2005). A test for parameter homogeneity in co_2 panel ekc estimations. Environmental and Resource Economics, 32:229–239.

Dufour, J. and Torres, O. (1996). Union-intersection and sample-split methods in econometrics with application to ma and sur models. Working paper, Summer meetings of the Econometric Society, Iowa City.

Durlauf, S., Johnson, P., and Temple, J. (2005). Growth econometrics. In Durlauf, S. and Aghion, P., editors, Handbook of Economic Growth, page Ch.8. Elsevier.

Durlauf, S., Sala-i Martin, X., Bernard, A., Jones, C., Quah, D., and Galor, O. (1996). Controversies (on growth and convergence). Economic Journal, 106(437):1016–1069.

Elliott, G., Rothenberg, T., and Stock, J. (1996). Efficient tests for for an autoregressive unit root. Econometrica, 64(4):813–836.

Esteban, J., Gradin, C., and Ray, D. (1999). Extension of a measure of polarization with an application to the income distribution of five oecd countries. Luxembourg Income Study Working Paper Series 218, Syracuse University, Maxwell School of Citizenship and Public Affairs.

Evans, P. (1996). Using cross-country variances to evaluate growth theories. Journal of Economic Dynamics and Control, 20:1027–1049.

Evans, P. (1998). Using panel data to evaluate growth theories. International Economic Review, 39(2):295–306.

Ezcurra, R. (2007). Is there cross-country convergence in carbon dioxide emissions. Energy Policy, 35:1363–1372.

Fan, J., Haerdle, W., and Mammen, E. (1998). Direct estimation of additive and linear components for high-dimensional data. The Annals of Statistics, 26:943–971.

Fan, J. and Li, R. (2004). New estimation and model selection procedure for semiparametric modeling in longitudinal data analysis. Journal of American Statistical Association, 99:710–723.

Friedman, M. (1992). Do old fallacies ever die. Journal of Economic Literature, 30(4):2129–2132.

Greene, W. (2003). Econometric Analysis. Prentice Hall International, 5th edition.

Grossman, G. M. and Krueger, A. (1995). Economic growth and the environment. Quarterly Journal of Economics, 112:353–378.

Gu, C. (2002). Smoothing Spline ANOVA Models. Springer, New York.

Hall, P., Racine, J., and Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. Journal of the American Statistical Association, 99:1015–1026.

Hartigan, J. and Hartigan, P. (1985). The dip test of unimodality. The Annals of Statistics, 13(1):70–84.

Hastie, T. and Tibshirani, R. (1990). Generalized Additive Models. Chapman and Hall.

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

Heil, M. and Selden, T. (2001). Carbon emissions and economic development: future trajectories based on historical experience. Environment and Development Economics, 6:63–83.

Herrero, C., Soler, A., and Villar, A. (2004). Capital humano y desarrollo humano en españa, sus comunidades autónomas y provincias. 1980-2000. Programa de apoyo al capital humano, Fundación Bancaja y Instituto Valenciano de Investigaciones Económicas (Ivie).

Holtz-Eakin, D. and Selden, T. (1995). Stocking the fires? co_2 emissions and economic growth. Journal of Public Economics, 57:85–101.

Hsiao, C. (2003). Analysis of Panel Data. Cambridge University Press, Cambridge, 2nd edition.

Hsiao, C., Li, Q., and Racine, J. (2007). A consistent model specification test with mixed categorical and continuous data. Journal of Econometrics.

Im, K., Pesaran, M., and Shin, Y. (2003). Testing for unit root in heterogeneous panels. Journal of Econometrics, 115:53–74.

Kalaitzidakis, P., Mamuneas, T., and Stengos, T. (2008). The contribution of greenhouse pollution to economic growth. Discussion paper 2008-2, University of Guelph, Department of Economics.

Koenker, R. (1981). A note on studentizing a test for heteroscedasticity. Journal of Econometrics, 17(1):107–112.

Koenker, R. (2007). quantreg: Quantile Regression. R package version 4.06, http://www.r-project.org.

Koop, G. and Tole, L. (1999). Is there an the environmental kuznets curve for deforestation. Journal of Development Economics, 58:231–244.

Lanne, M. and Liski, M. (2003). Trends and breaks in per capita carbon dioxide emissions 1870-2028. working paper 03-002, Center for Energy and Environmental Policy Research.

Lee, J. and Strazicich, M. (2003). Minimum lagrange multiplier unit root test with two structural breaks. The Review of Economics and Statistics, 85(4):1082–1089.

Levin, A., Lin, C. F., and Chu, C. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. Journal of Econometrics, 108:1–24.

Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. Econometric Reviews, 15(3):261–274.

Li, Q. and Papell, D. (1999). Convergence of international output. time series evidence for 16 oecd countries. International Review of Economics and Finance, 8:267–280.

Li, Q. and Racine, J. (2003). Nonparametric estimation of distributions with categorical and continuous data. Journal of Multivariate Analysis, 86:266–292.

Li, Q. and Racine, J. (2007). Nonparametric Econometrics. Princeton University Press, New Jersey.

Linton, O. and Nielsen, J. (1995). Estimating structured nonparametric regression by the kernel method. Biometrika, 82:93–101.

List, J. (1999). Have air pollutant emissions converged among u.s. regions? evidence from unit root tests. Southern Economic Journal, 66(1):144–155.

List, J. and Gallet, C. (1999). The environmental kuznets curve: does one size fit all? Ecological Economics, 31:409–423.

Maasoumi, E., Racine, J., and Stengos, T. (2007). Growth and convergence: A profile of distribution dynamics and mobility. Journal of Econometrics, 136:483–508.

MacKinnon, J. (1996). Numerical distribution functions for unit root and cointegration tests. Journal of Applied Econometrics, 11:601–618.

Maddala, G. (1999). On the use of panel data methods with cross-country data. Annales d'économie et de statistique, 55-56:429–449.

Maddala, G. and Wu, S. (1999). A comparative study of unit root tests with panel daza and new simple test. Oxford Bulletin of Economics and Statistics, 61(s1):631–652.

Maddison, A. (2007). Historical statistics. Online data, Groningen Growth and Development Center, University of Groningen,. http://www.ggdc.net/maddison.

Marland, G., Boden, T., and Andres, R. (2006). Global, regional, and national fossil fuel co_2 emissions. In: Trends: A compendium of data on global change, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Dpt. of Energy, Tenn., U.S.A. http://cdiac.ornl.gov/trends/emis/tre_coun.htm, June.

McKitrick, R. and Strazicich, M. (2005). Stationarity of global per capita carbon dioxide emissions: Implications for global warming scenarios. Working Paper 2005-03, University of Guelph.

Meyer, A. (2001). Contraction and Convergence: The Global Solution to Climate Change. Schumacher Briefings. Green Books, The Schumacher Society. Global Commons Institute: http://www.gci.org.uk/.

Millimet, D., List, J., and Stengos, T. (2003). The environmental kuznets curve: real progress or misspecified models? The Review of Economics and Statistics, 85(1):1038–1047.

Nguyen Van, P. (2005). Distribution dynamics of co_2 emissions. Environmental and Resource Economics, 32.

Nguyen Van, P. and Azomahou, T. (2007). Heterogeneity in environmental quality: An empirical analysis of deforestation. Journal of Development Economics, 84(1).

Pedroni, P. and Yao, J. (2006). Regional income divergence in china. Journal of Asian Economics, 17.

Perman, R. and Stern, D. (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(3).

Perron, P. and Volgesang, T. (1992). Nonstationarity and level shifts with an application to purchasing power parity. Journal of Business and Economic Statistics, 10(3).

Pesaran, M. (2007). A pair-wise approach to testing output and growth convergence. Journal of Econometrics, 138.

Quah, D. (1993a). Empirical cross-sectional dynamics in economic growth. European Economic Review, 37:426–434.

Quah, D. (1993b). Galton's fallacy and tests of the convergence hypothesis. Scandinavian Journal of Economics, 95(4):427–443.

Quah, D. (1997). Empirics for growth and distribution: Stratification, polarization, and convergence clubs. Journal of Economic Growth, 2:27–59.

R Development Core Team (2007). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Racine, J. (1997). Consistent significance testing for nonparametric regression. Journal of Business and Economic Statistics, 15:369–379.

Racine, J. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. Journal of Econometrics, 119:99–130.

Robinson, P. (1988). Root-n-consistent semiparametric regression. Econometrica, 56:931–954.

Roca, J., Padilla, E., Farre, M., and Galletto, V. (2001). Economic growth and atmospheric pollution in spain: discussing the environmental kuznets curve hypothesis. Ecological Economics, 39:85–99.

Romero-Ávila, D. (2007). Convergence in carbon dioxide emissions among industrialised countries revisited. Energy Economics, forthcoming.

Rose, A. and Stevens, B. (1996). Equity aspects of marketable permits approach to global warming policy. Working paper, 7^{th} Annual Conference of the European Association of Environmental and Resource Economics in Lisbon, Portugal,.

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

Schmid, F. and Trede, M. (2003). Simple test for peakedness, fat tails and leptokurtosis based on quantiles. Computational and Data Analysis, 43:1–12.

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

Shioji, E. (1997). It's still 2%: Evidence on convergence from 116 years of the us states panel data. Working paper, Universitat Pompeu Fabra.

Stegman, A. (2005). Convergence in carbon emissions per capita. Working paper 8, CAMA, The Australian National University.

Stern, D. (2003). The Environmental Kuznets Curve. Internet Encyclopaedia of Ecological Economics.

Stockey, N. (1998). Are there limits to growth. International Economic Review, 39(1):1–31.

Stone, C. (1985). Additive regression and other nonparametric models. Annals of Statistics, 13:689–705.

Strazicich, M. and List, J. (2003). Are co$_2$ emissions levels converging among industrial countries? Environmental and Resource Economics, 24:263–271.

Taskin, F. and Zaim, O. (2000). Searching for a kuznets curve in environmental efficiency using kernel estimation. Economics letters, 68:217–223.

Tomljanovic, M. and Vogelsang, T. (2002). Are u.s. regions converging? using new econometric methods to examine old issues. Empirical Economics, 27:49–62.

U.S. Census Bureau (2006). International data base (idb). Online data, Population Division, International Program Center,. http://www.census.gov/ipc/www/idb/.

Van der Ploeg, F. and Withagen, C. (1991). Pollution control and the ramsey problem. Environmental and Resource Economics, 1(2):215–236.

Volgesang, T. and Perron, P. (1998). Additional tests for a unit root allowing for a break in the trend function at an unknown time. International Economic Review, 39(4).

Vollebergh, H., Dijkgraaf, E., and Melenberg, B. (2005). Environmental kuznets curves for co$_2$: Heterogeneity versus homogeneity. Discussion paper 25, CentER, Tilburg University.

Westerlund, J. and Basher, S. (2008). Testing the convergence in carbon dioxide emissions using a century of panel data. Environmental and Resource Economics, 40:109–120.

Wood, S. (2000). Modeling and smoothing parameter estimation with multiple quadratic penalties. Journal of Royal Statistical Aociety, Series B 65:95–114.

Wood, S. (2006). Generalized Additive Models: An Introduction with R. Chapman & Hall, Texts in Statistical Sciences.

World Bank (2004). World development indicators. Technical report, Washington DC, USA. CD-Rom.

Yatchew, A. (2003). Semiparametric Regression for the Applied Econometrician. Cambridge University Press, Cambridge.