SEEDS
Sustainability Environmental Economics
and Dynamics Studies

Working Paper Series

Nonlinearity, heterogeneity and unobserved effects in the carbon dioxide emissions-economic development relation for advanced countries

by

Massimiliano Mazzanti, Antonio Musolesi

22/2014

SEEDS Working Paper 22/2014
August 2014
by Massimiliano Mazzanti, Antonio Musolesi

Nonlinearity, heterogeneity and unobserved effects in the carbon dioxide emissions-economic development relation for advanced countries

Massimiliano Mazzanti*& Antonio Musolesi†‡

November 16, 2013

Abstract

We study long run carbon dioxide emissions-economic development relationships for advanced countries grouped in policy relevant groups: North America and Oceania, South Europe, North Europe. By relying on recent advances on Generalized Additive Mixed Models (GAMMs) and adopting interaction models, we handle simultaneously three main econometric issues, named here as *functional form bias*, *heterogeneity bias* and *omitted time related factors bias*, which have been proved to be relevant but have been addressed separately in previous papers. The model incorporates nonlinear effects, eventually heterogeneous across countries, for both income and time. We also handle serial correlation by using autoregressive moving average (ARMA) processes. We find that country-specific time related factors weight more than income in driving the northern EU Environmental Kuznets. Overall, the countries differ more on their carbon-time relation than on the carbon-income relation which is in almost all cases monotonic positive. Once serial correlation and (heterogeneous) time effects have been accounted for, only three Scandinavian countries – Denmark, Finland and Sweden – present some threshold effect on the CO2-development relation.

Keywords: *Environmental Kuznets curve; Semiparametric models; Generalized Additive Mixed models; Interaction models.*

JEL classification: *C14, C23, Q53*

*Univ. Ferrara, Department of Economics, Ferrara, Italy

†INRA, UMR 1215 GAEL, F-38000 Grenoble, France

‡Univ. Grenoble Alpes, UMR 1215 GAEL, F-38000 Grenoble, France

1 Introduction

Many diversified stylised facts have been proposed on the relationship between pollution and economic development. An extensive overview of the main theoretical matters involved can be found in Borghesi (2001), who discusses the Kuznets conceptual framework, which touches on inequality in relation to sustainable development issues.[1] More recently Brock and Taylor (2010) explain how that environmental Kuznets curves (EKC) framework[2] is coherent with a reformulated "green Solow model" where emission per capita growth is driven by GDP per capita growth and technology. Moreover, the relationships between environmental performance, growth and innovation patterns have received increasing attention in the policy agenda of advanced economies, (OECD, 2002, 2010, 2011), and in particular within the European Union (EU), inside the general debate that has followed the Stern Review (Dietz, 2011) around climate change adaptation and mitigation actions. The economic and policy debate today largely revolves around the chances to boost a more competitive, greener economy, and the issue of climate change constitutes a substantial part of this (EEA, 2013).

This paper analyses the long-term CO2-income relationship for more developed countries. The relevance of carbon dioxide depends on the fact that even more advanced economies have not meaningfully reduced CO2 emissions so far, though an overall picture can conceal heterogeneous facts (Musolesi et al., 2010; Mazzanti and Musolesi, 2013, UNDP, 2009). Carbon dioxide is a global public good whose features touch upon both country specifities and interdependencies; the relevance of international issues and country interdependecies is, in our view, more consistent with pooled analyses rather than country-based studies. The key role of energy issues (energy mix, efficiency), that share similarities across countries with notable exceptions and outliers due to exogenous/structural factors and investment decisions, explains the necessity of adopting an intermediate perspective: aggregation with country based insights.

Specifically, we aim to investigate which (groups of) advanced countries have succeeded in reducing CO2 emissions while growing in income, achieving a negative elasticity of greenhouse gas emissions with respect to their

[1] Kijima et al (2010) provide a recent survey of theoretically oriented papers with a view to dynamic issues.

[2] See e.g. Andreoni and Levinson, 2001; Millimet et al., 2003; Grossman and Krueger, 1994, as seminal benchmarks. An interesting recent theoretical paper on ECK is Figueroa and Pasten (2013).

GDP[3]. We use the same group classification adopted by Mazzanti and Musolesi (2013), who focus on advanced countries, by subdividing them into the Umbrella group[4], Northern Europe (EU-North) and Southern Europe (EU-South), which witness quite different economic, policy and institutional features. This classification, which is aimed at providing original insights for policy, is strictly integrated with methodological development. Indeed, from both a methodological and a policy oriented perspective, it is interesting to investigate whether countries that belong to groups sharing structural similarities may eventually present different income effects and/or tend to differentiate with respect to unobservable time-related factors. This because some countries might lead the way to become 'policy' and 'technological' leaders in the green economy. This also because CO2 is a global public good and contrary to local pollutants such as particulate matters and regional emissions as acidificants, CO2 calls into question the role of external global factors/global spillover effects.

Other unobserved global factors possibly affecting the level of CO2 emissions might include, for instance, aggregate technological shocks, global environmental policies or oil price shocks that may influence the production of CO2 through their effects on production costs. These factors may affect eventually heterogeneously different countries under study. Moreover, since the countries under study may differ with respect to some unobservable time varying variable such as technology, policy, institutions, culture, etc which linked to the production of CO2, it seems to be especially worth emphasising the role of unobservable time-related factors and try to answer to the following questions: how do time-related factors (perhaps heterogeneously) affect long-term CO2 dynamics? And how does their inclusion in the model affect the estimation of income effects? In particular, we try to simultaneously handle three main econometric issues, named here as *functional form bias*, *heterogeneity bias* and *omitted time-related factors bias*.

[3]Future studies may well analyse emerging countries as well. In the case of eastern EU, for example, emerging countries pose problems given the collapse in emissions and GDP in the 90's. India and China are historically less interesting since they have clearly followed a growth-oriented path with very limited environmental factors in terms of innovation and policy. The Economist (August 10th 2013, pp 17-21) has recently debated over the Chinese environmental challenge. We here take a long term ex post view on advanced countries that have been mostly responsible for global CO2 emissions so far.

[4]The Umbrella group refers to a loose coalition of non-EU developed countries formed after the Kyoto protocol that have sustained a mild approach to climate policy, predominantly North America and Australia. See Barrett (2003) for further insights.

Early EKC literature focused on very constrained specifications, such as parametric formulations (typically polynomial functions) imposing common slopes across countries and not accounting for the (possibly heterogeneous) effect of time-related factors. One strand of the empirical literature has relaxed parametric formulation to adopt non parametric methods (Azomahou et al.2006; Azomahou and Mishra, 2008). This may help, for instance, to avoid the false inference that the CO2-income relation is not monotonic if the true relation has a threshold. Another strand has focused on the heterogeneity bias associated to the estimation of models with common slopes. As Hsiao (2003) points out, if the true relation is characterised by heterogeneous intercepts and slopes, estimating a model with individual intercepts but common slopes could produce the false inference that the estimated relation is curvilinear. Empirically, this situation is more likely when the range of the explanatory variables diverges across cross-sections. This situation generally corresponds to the estimation of EKC for groups of countries because: i) per capita GDP presents high variation across countries, ii) the different groups of countries cannot be characterised by a common slope and, consequently, there is a high risk of estimating a false curvilinear relation when using homogeneous estimators. Only very recently, both strands have recognised the relevance of taking into account unobserved time effects, which may eventually explain a large part of the evolution of CO2. This has been supported for instance by Melenberg et al. (2009), who use a nonparametric setting and by Musolesi and Mazzanti (2013) who adopt, among other estimators, the Common Correlated Effect (CCE) approach developed by Pesaran (2006) in a parametric framework. A major limitation of the former is that it does not allows for the effects of income or time to vary across cross-sections; while the latter does not allow for non parametric effects.

In order to achieve our goal, namely disentangling income and time-related effects (which are possibly heterogeneous across countries) in the study of greenhouse gas dynamics, while allowing for possible residual serial correlation at the same time, we use Generalized Additive Mixed Models (GAMMs, Ruppert et al., 2003; Augustin et al., 2009; Wood, 2006). GAMMs contain Generalized Additive Models (GAMs) as a special case, introduced by Hastie and Tibshirani (1990) and more recently developed both in theoretical and computational directions. The estimation of GAMs relies on the decomposition of the smooth functions on a spline basis; then a penalty term is added into the log-likelihood (Wood, 2003, 2006). Wood (2004) in particular provides an optimally stable smoothness selection method which presents some advantages when compared to previous approaches, such as modified backfitting (Hastie and Tibshirani, 1990) or Smoothing Spline ANOVA (e.g. Gu and Wahba, 1993). Smoothing parameter estimation and reliable confidence interval calculation is difficult to obtain with modified backfitting, whereas Smoothing Spline ANOVA provides well-founded smoothing parameter selection methods and confidence intervals with good coverage probabilities but at high computational costs. To circumvent these problems, Wood (among others) suggests the use of penalized regression splines (2000): this nevertheless leaves a number of practical problems concerning convergence and numerical stability unsolved. Wood (2004) further developed the model by providing an optimally stable smoothness selection method and subsequently provided a computationally efficient method for direct generalized additive model smoothness selection (2008). A very appealing feature of the method proposed by Wood in 2004 with respect to other approaches is that it has been shown to perform very well even in the case of almost co-incident covariates. The mixed model approach adopted here provides a consistent a computationally manageable way to simultaneously handle smoothing and serial correlation. Wood also provides some useful details on how GAMs are represented as mixed models (2004, 2006a).

Despite their appeal, GAMs and GAMMs also require some caveats worth mention. A first possible limitation of GAMs/ GAMMs is that they are less general than fully non-separable models. These models, however, present difficulties which are not present with GAMs in terms of interpretation, statistical feasibility (e.g. the curse of dimensionality) and identification (see e.g. Hoderlein and White, 2012 and Evdokimov, 2010). A second caution to consider concerning GAMs is that a fully developed asymptotic theory has yet to exist. However, some asymptotic results have recently been provided by Yoshida and Nato, (2012, 2013). In particular, they have shown the asymptotic normality of the penalized spline estimator in a GAM framework (2013), thus generalizing the results of Kauermann et al. (2009), who focuses on the penalized spline estimator in generalized linear models (GLMs). They also show the asymptotical normality of the penalized quasi likelihood approach by Breslow and Clayton (1993) which is an efficient method when applied to GAMMs.

An interesting feature of the proposed approach, compared to that of related literature, is that it allows for the estimation of the nonparametric time effect rather than considering it as nuisance term. This is very important in an economic and policy oriented analysis because it allows not only (i) to obtain a *proper* income effect, but it also allows (ii) to nonparametrically investigate how time-related factors may drive long-term CO2 evolution. Moreover, (iii) the adoption of interaction models (see e.g. Ruppert et al., 2003) allows us to consider a specification allowing for both country-specific nonparametric time effects and perhaps even country-specific nonparametric income effects. This is possible in practice given the large time series dimension of our data set and permits a maximum level of country-specific heterogeneity. This allows for handling the heterogeneity bias if the true relation is characterised by country-specific effects. This is also important from an economically oriented angle. It can be expected that even countries belonging to similar geographical/economic groups may have different income effects and may also tend to 'specialize' with respect to time-related unobservable factors which in turn may heterogeneously affect CO2 emissions such as innovation and technological progress, energy and also policy. Finally, (iv) we handle serial correlation by using autoregressive moving average (ARMA) processes.

The remainder of paper is structured as follows. Section 2 presents the data, the country groups, and some descriptive statistics. Section 3 debates around issues of econometric specification, estimation and identification. Section 4 presents the model building; Section 5 comments on the main results and Section 6 concludes.

2 Data

Data on emissions is taken from the database on global, regional, and national fossil fuel CO2 emissions prepared for the US Department of Energy's Carbon Dioxide Information Analysis Centre (CDIAC). For our study, we use the subset of emission data that matches the available time series on GDP per capita. Data on GDP per capita in 1990 International 'Geary-Khamis' dollars is taken from the database managed by the OECD.

We use the subset of emission data that matches the available time series on GDP per capita on the basis of joint availability, series continuity, and country definitions. This resulted in a sample which covers a long period (1960-2001). Table 1 below summarises the main variables used and the descriptive statistics.

The Umbrella group presents the highest average level of both CO2 per capita (expressed in terms of tonnes per capita) and GDP per capita (3.14 and 15,143, respectively), while southern European countries are characterised by the lowest average levels of these variables (1.48 and 10,215). The northern European countries have a similar average level of GDP per capita (14,203) compared to the Umbrella group but are characterised by lower levels of emissions (2.61).

Figures 1-3 depict the relationship between CO2 and income for the three samples. We provide real data, and the curve fitted (non-parametrically)

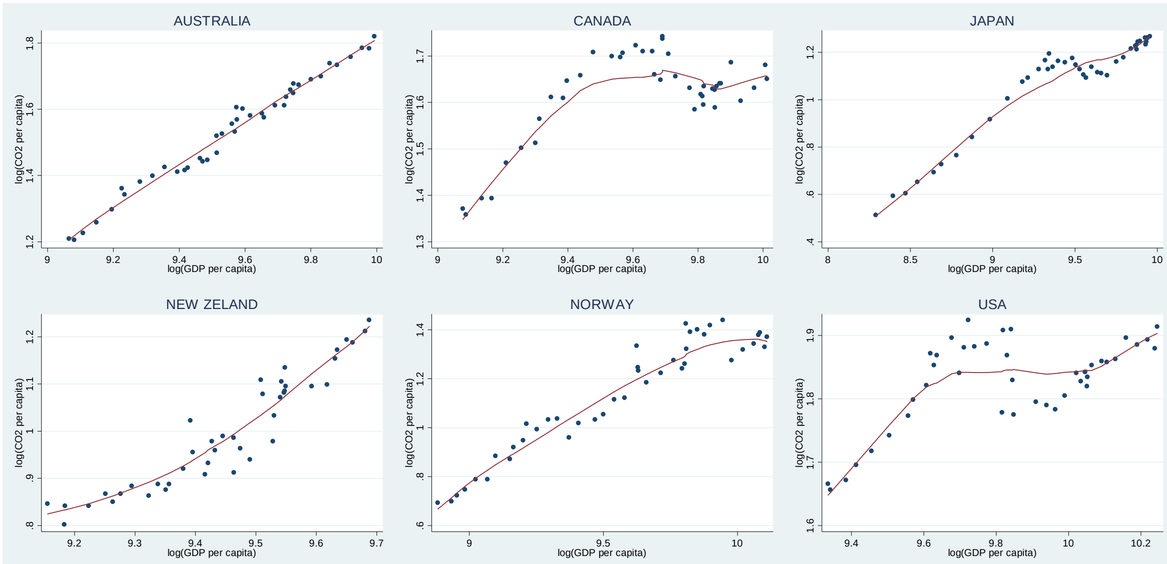


Figure 1: UMBRELLA countries (scatter: real values. Line: robust locally weighted scatterplot smoothing)

by robust locally weighted scatter plot smoothing (lowness). The CO2-GDP relationship is quite homogeneous within each group: it is clearly monotonic for the Umbrella group and for the southern EU-but shows an inverted U shape for -northern EU countries.

Table 1: Descriptive statistics

	Mean	S.D.	Min	Max
Umbrella group				
CO2 per capita	3.144921	1.393584	0.67	5.85
GDP per capita (GDPpc)	15,143.21	4,763.547	3,986.417	28,129.23
EU North				
CO2 per capita	2.60875	0.5630643	0.91	3.88
GDP per capita (GDPpc)	14,203.73	3,759.392	6,230.359	23,160
EU South				
CO2 per capita	1.488294	0.6085014	0.25	3.05
GDP per capita (GDPpc)	10,215.44	4,265.277	2,955.836	23,201.45

T= 1960-2001; CO2 per capita in t/pc; GDP per capita in 1990 International 'Geary-Khamis' dollars

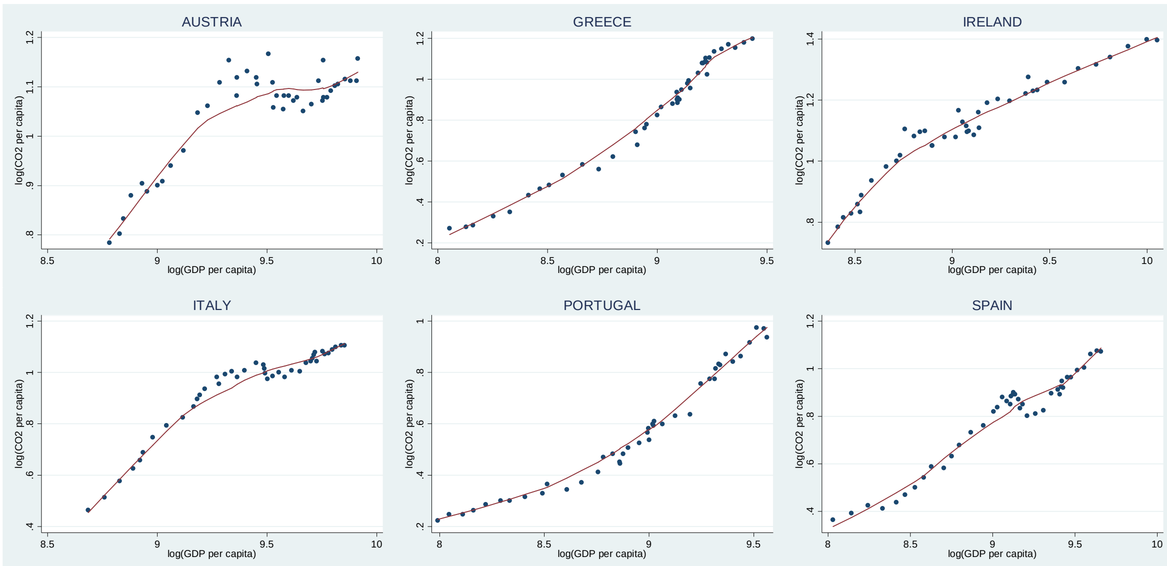


Figure 2: EU-SOUTH countries (scatter: real values. Line: robust locally weighted scatterplot smoothing)

3 Econometric specification, identification and estimation

3.1 Econometric specification: various misspecification biases

Let us suppose that the researcher observes panel data (y_{it}, x_{it}), where y is the logarithm of CO_2 emissions per capita, x is the logarithm of per capita GDP; $i \in \Gamma$, and Γ is the set of cross-section units $\Gamma = \{1, 2, ..., N\}$ and $t \in \Lambda$=$\{1, 2, ..., T\}$ indicates time series observations. A very general specification is obtained by adopting a fully non separable model such as

$$y_{it} = f\left(x_{it}, c_i, t, \varepsilon_{it}\right)$$

where f is real unknown function, c_i are individual effects capturing time invariant heterogeneity, t capture the effect of time-related omitted factors, and ε_{it} is the the idiosyncratic term.

To date, there is an increasing amount of theoretical literature on non parametric panel data estimators aiming to provide very general econometric set-ups such as the non-parametric panel data model, i.e. a model of the following kind $y_{it} = f(x^1_{it}, ..., x^k_{it}) + c_i + \varepsilon_{it}$, the partially or fully non-separable

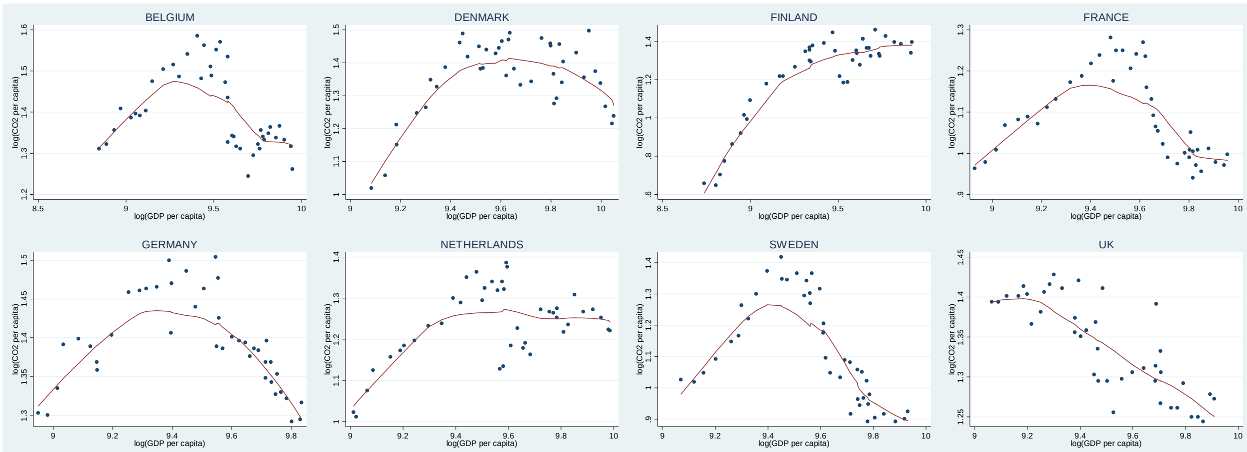


Figure 3: EU-NORTH countries (scatter: real values. Line: robust locally weighted scatterplot smoothing)

models, i.e. $y_{it} = f(x^1_{it}, ..., x^k_{it}, c_i) + \varepsilon_{it}$ and $y_{it} = f(x^1_{it}, ..., x^k_{it}, c_i, \varepsilon_{it})$ (see e.g. Henderson et al. 2008; Su and Ullah, 2010).[5]

Despite their appeal, fully or partially non-separable models present theoretical and computational difficulties and the identification conditions arising in such models can be difficult to hold (see Hoderlein and White, 2012 and Evdokimov, 2010).[6]

These considerations allow us to focus on additive models (Stone, 1985, Hastie and Tibshirani 1990). They avoid the curse of dimensionality since each of the individual additive terms is estimated using a univariate smoother. They are also easily interpretable, while fully non-separable models present problems of interpretability, and they do not present big identification problems. Finally,

[5] An approach which has been proposed to estimate models in cases where explanatory variables do not enter additively, differently from individual effects and the error term, is resorting to a local linear approximation of the model and then using the profile least square method (Su and Ullah, 2006, 2010). This allows for estimating the model without using a transformation to eliminate the fixed effects. Another widely adopted approach has been to first take differences to eliminate the individual effects. At this point, the differenced equation can be estimated, after a local linear approximation, for example by using local linear least squares (Li and Stengos, 1996) or through the iterative kernel estimator (Henderson et al, 2008).

[6] Hoderline and White (2012) focus on the identification of fully non separable models and even though their main result is that a generalised version of differencing identifies local average responses, they also find that such a result is confined to the subpopulation of "stayers" (Chamberlain, 1982), i.e. the population for which the explanatory variables do not change over time; a case which does not correspond to our empirical framework.

and very importantly, additive models fit perfectly with the purpose of this paper to disentangle (possibly heterogeneous) income and time effects. Therefore, we more specifically assume that the income effect, the effect of (time invariant) unobserved heterogeneity, the effect of time and the idiosyncratic effect are separable:

$$y_{it} = c_i + f_i\left(x_{it}\right) + g_i\left(t\right) + \varepsilon_{it} \tag{1}$$

where f_i captures the effect of income on CO_2 emissions while the effect of time is measured through the function g_i. Both effects are eventually heterogeneous across countries.

It is worth noting, however, that the early literature on income environment long run relationships has focused on very constrained specifications, as for instance setting $g_i\left(t\right) = 0$; $f_i\left(x_{it}\right) = p\left(x_{it}, \beta\right)$, where $p\left(x_{it}, \beta\right)$ is a polynomial function, and obtaining the additive fixed effects specification

$$y_{it} = c_i + p\left(x_{it}, \beta\right) + \varepsilon_{it}$$

Compared to (1), such a specification may suffer of different kinds of misspecification bias, and in particular:

- *functional form bias* if the true relation between CO2 and GDP cannot be approximated with a polynomial function $p\left(x_{it}, \beta\right)$. This has been largely recognised in the literature, which stresses the need for non-constrained functional specifications (Azomahou et al.2006; Azomahou and Mishra, 2008; Azomahou et al. 2009);
- *heterogeneity bias* since it is possible that the effect of GDP on CO2 can be heterogeneous across countries. A more realistic assumption would allow for individual income effects, $f_i\left(x_{it}\right)$ (Musolesi et al., 2010; Mazzanti and Musolesi, 2013);
- *omitted time-related factors bias* dues to the omission of a (eventually heterogeneous) relevant time effect. The literature has widely adopted the restriction $g_i\left(t\right) = 0$. This is motivated by the following reasons: it allows for greater comparability with existing studies and, perhaps more importantly, this kind of econometric specification is useful if the researcher is interested in capturing the total effects of GDP on CO2 including the indirect effects linked to omitted (or unobserved) variables, such as energy prices, technological changes, environmental policies, etc, which are correlated with both GDP and time. However, if the goal is to measure the ceteris paribus impact of GDP on CO2 emissions, imposing $g_i\left(t\right) = 0$ might be not appropriate because it leads to an omitted

time-related factor bias. To the best of our knowledge very few studies to date have focused on such an issue. For instance, among other panel data estimators, Mazzanti and Musolesi (2013) applied the CCE approach proposed by Pesaran (2006). Such a method allows for unobserved common factors heterogeneously affecting the dependent variable. However, such factors are viewed as nuisance variables while the main focus rests on the estimation of the heterogeneous (but parametric) effect of income on CO2. Melenberg et al. (2009) and Ordás Criado et al. (2011) have provided nonparametric analyses. Both studies estimated eq. 1 without imposing a parametric formulation while imposing that $f_i\left(x_{it}\right)$ is homogeneous across countries. A difference between these two works is that, while Melenberg et al. (2009) assumes that the unobserved time-relatedfactor $g_i\left(t\right)$ is common to specific groups of countries within the sample and considered the function $g_i\left(t\right)$ as a nuisance term, Ordás Criado et al. (2011) introduced a common time effect by means of time fixed effects.

The main advantage of using recent developments in GAM theory is that it allows for tackling the three issues of unconstrained functional form, omitted time-related factors and heterogeneity bias, simultaneously.

3.2 Identification

A fully additive model as in eq. (1) can be dealt with GAMMs. This does not require a local linear approximation and, also in this case, the individual fixed effects can be treated as nuisance terms to be eliminated with a transformation, or as dummy variables. Both approaches present some relative drawbacks and benefits. In particular, the latter approach may be computationally complicated but does not suffer like the former from the possible (partial) lack of identification arising with the adoption of a transformation approach such as first differencing to eliminate the individual effects. Indeed, by differencing Equation (1) we get (see also Azomahou and Mishra, 2008; Azomahou et al. 2009, Su and Ullah, 2010):

$$\left(y_{it}-y_{it}\right)=f_i\left(x_{it}\right)-f_i\left(x_{it-1}\right)+g_i\left(t\right)-g_i\left(t-1\right)+\left(\varepsilon_{it}-\varepsilon_{it-1}\right),$$

and some components of the functions f anf g may not fully identified because as argued by Su and Ullah (2010), if, for example,

$$f(x_{it}) = a + m(x_{it}),$$

then differencing does not allow the identification of $f(x_{it})$, and eventually only $m(x_{it})$ can be identified. Secondly, such an approach doubles the non-parametric functions to be estimated. In our empirical framework this problem becomes extremely important because estimating eq. (1), after differencing and without imposing the constraint that $f_i\left(x_{it}\right) = f\left(x_{it}\right) \forall i$ or that $g_i\left(t\right) = g\left(t\right) \forall i$, requires the estimation of $N * 4$ nonparametric functions. Thus, first differencing may be useful in practice when N is large compared to T, as usual in micro data, or to estimate a 'feedback effect' through the function $f\left(x_{it-1}\right)$ as in Azomahou and Mishra (2008).

Given the structure of our panel data set (small N and large T), it is not computationally costly to estimate the model directly without eliminating the individual effects. We follow thus such an approach by including the individual intercepts in the parametric part of the level equation as in Mammen *et al.* (2009) and Ordás Criado *et al.* (2011).

3.3 Estimation

The estimation is carried out by using the gamm *( )* function of the mgcv R package (Wood, 2013). In the identity link-normality case, the mgcv routine performs the estimation by using general linear mixed effects modelling software, lme, while in the generalized case only approximate inference is available, and relies on the Penalized Quasi-Likelihood approach by Breslow and Clayton (1993). It allows correlated errors by calling the nlme R package (Pinheiro et al., 2013). Penalised Regression Splines are adopted as a basis for representing the smooth terms (Wood, 2003, 2006ab). The smoothing parameter values are selected by the GCV (Generalised Cross Validation) criterion, [7] and statistical inference is made by computing 'Bayesian p-values' (Wood, 2013). These appear to have better frequentist performance (in terms of power and distribution under the null) than the alternative strictly frequentist approximation (Wood, 2006a,b).

[7]Since the GCV may present a tendency towards over fitting, we have increased the amount of smoothing by correcting the GCV score by a factor $\delta = 1.4$ which can correct the over fitting without compromising model fit (Kim and Gu, 2004).

4 A semiparametric model for CO_2 emissions

4.1 Alternative specifications for $f_i(x_{it})$ and $g_i(t)$

In the following, we provide an empirical strategy to choose among alternative specifications for both $f_i(x_{it})$ and $g_i(t)$ as well as for the covariance structure. Concerning $f_i(x_{it})$ and $g_i(t)$ we will focus on the following models.

We first consider a specification imposing a common income effect and without accounting for any kind of time effect. This model has been largely adopted in previous works (Azomahou et al., 2006; Azomahou and Mishra, 2008; or Azomahou et al., 2009) and allows for obtaining results on the total effect of income on CO2, including indirect effects linked to omitted variables. Next, we introduce time-related factors into the model. This is the main focus of the paper. We first assume that both these time factors and the income homogeneously affect CO2 evolution. This should allow both to obtain a proper income effect and to examine the effect of time-related factors which may drive CO2 evolution. This specification, however, may suffer from a heterogeneity bias. We thus relax the hypothesis of a homogeneous time effect and estimate a semiparametric model allowing for country-specific nonparametric time effects, $g_i(t)$. Such a kind of specification has been already proved to be very useful in a parametric framework. In a policy evaluation framework, Heckman and Hotz's (1989) proposed the so called random growth model allowing for individual specific trend, i.e. a model of the kind $y_{it} = c_i + \gamma_i t + \beta x_{it} + \varepsilon_{it}$. Wooldridge (2005) provides very useful methodological insights, while Papke (1994) and Friedberg (1998) are examples showing empirically how important can be to allow for individual specific trends. A motivation of such specification is that it allows (c_i, γ_i) to be arbitrarily correlated with x_{it}. This can certainly relevant when x_{it} is an indicator of program evaluation as in Heckman and Hotz (1989) but could also be a key issue in our framework since both CO2 emissions and per capita GDP can plausibly depend on individual-specific trends, in addition to the level effect, c_i. More recently in a more macroeconometric oriented framework, Pesaran (2006) has proposed, the CCE approach, which makes use of a factor model representation to allow a finite number of unobservable (and/or observed) common factors to have an heterogeneous effect on the dependent variable. One main reason supporting a modelisation allowing for country-specific nonparametric time effects, $g_i(t)$ is that even for countries that belong to similar geographical/economic realities, we note that the effect of unobservable time-related factors on CO2 can be expected to be heterogeneous across countries. This is because countries tend to 'specialise' with respect to unobservable time-related factors such as innovation, energy and also policy. Such a modelisation may also be motivated in cases with common time effects, e.g. the case of a common policy, but with country-specific reactions. Moreover, there are not well-established (theoretical or empirical) reasons to impose linearity. More specifically, innovation specialization is due to both market characteristics and willingness to create comparative advantages. Ennvironmental Policy ‘specialisation’ largely depends on the belief in policy-induced innovation effects (Costantini and Mazzanti, 2012), upon which some world areas might construct green technology competitive advantages. Energy issues depend on both policy frameworks and structural country features.

Alternatively, we allow for heterogeneous income effects (for a detailed discussion of this issue see e.g. Musolesi et al. 2010 and Mazzanti and Musolesi, 2013). Finally we fully exploit the time dimension of our data and consider an "unconstrained" model with both heterogeneous time effects and heterogeneous income effects.

In summary, the five alternative specifications for $f_i(x_{it})$ and $g_i(t)$ can be therefore written as:

M1 Individual fixed effects specification. $g_i(t) = 0$ (no time effect), $f_i(x_{it}) = f(x_{it})\ \forall i$ (homogeneous income effect), so that:

$$y_{it} = c_i + f(x_{it}) + \varepsilon_{it} \tag{M1}$$

M2 Individual fixed effects and common time effect. $g_i(t) = g(t)\ \forall i$ (homogeneous time effect), $f_i(x_{it}) = f(x_{it})\ \forall i$ (homogeneous income effect):

$$y_{it} = c_i + f(x_{it}) + g(t) + \varepsilon_{it} \tag{M2}$$

M3 Individual fixed effects, individual time effects. We hold the constraint $f_i(x_{it}) = f(x_{it})\ \forall i$ (homogeneous income effect):

$$y_{it} = c_i + f(x_{it}) + g_i(t) + \varepsilon_{it} \tag{M3}$$

M4 Individual fixed effects, individual income effects. We hold the constraint $g_i(t) = g(t)\ \forall i$ (homogeneous time effect):

$$y_{it} = c_i + f_i(x_{it}) + g(t) + \varepsilon_{it} \tag{M4}$$

M5 Individual fixed effects, individual time and individual GDP effect:

$$y_{it} = c_i + f_i(x_{it}) + g_i(t) + \varepsilon_{it}. \tag{M5}$$

4.2 Covariance structure

The error vector ε is distributed as $N(\mathbf{0}, \sigma^2\Lambda)$, where Λ is block diagonal with ε_i having covariance matrix Λ_i.The ε_i reflect the serial error correlation, which is modeled by a mixed autoregressive and moving average (ARMA) process using the approach given by Pinheiro and Bates (2000). An ARMA(p,q) can be written as:

$$\varepsilon_{it} = \sum_{j=1}^{p} \phi_j \varepsilon_{it-j} + \sum_{l=1}^{q} \theta_l v_{it-l} + v_{it}$$

where the ϕs and θs are the autoregressive an moving average's parameters and v_{it} is a random Gaussian white noise.

4.3 Model selection

We will now outline the two steps of our selection procedure.

Step one: *Selection of the serial correlation structure of the error term.*

First, for each specification M1-M5, we use the ACF function of the nlme R package (Pinheiro et al., 2013) to deduct the appropriate error structure (e.g. Hamilton, 1994). Second, since the estimated autocorrelation pattern does not generally provide a unique indication being possibly consistent with different processes, we also use model selection and testing procedures to choose the most appropriate error process (see e.g. Pinheiro and Bates, 2000, p. 239-244). Details are provided below.

Step two: *Selection of the appropriate level of heterogeneity (with respect to time and income).*

Step one has allowed to choose the appropriate error structure for each specification M1-M5. We now compare the five selected models. In our context, i.e. the identity link-normality case, the mgcv routine performs the estimation by using general linear mixed effects modelling software, lme, while in the generalized case only approximate inference is available, relying on the Penalized Quasi-Likelihood approach by Breslow and Clayton (1993). In our restricted case, thus, the inferential framework for linear mixed models applies and can be used for model comparison. Therefore we extensively use both information criteria (AIC and BIC) and, for nested models, the likelihood ratio test (Wood, 2006a, 2013. Also see Augustin et al., 2009 for the generalized case).

5 Results

5.1 Model selection

Step one. The plot of the empirical autocorrelation function of standardized residuals with 5% level two-sided critical bounds is displayed in Figure 4. The plot concerns the M1-M5 models for the Umbrella group. For the other groups, rather similar plots are obtained and the detailed results are available upon request. For M1 and M2 the plot in Figure 4 is consistent with an AR(1) process with positive autoregressive parameter. This process has a correlation function which decreases exponentially with lag: $h(\phi, k) = \phi^k$, where ϕ is the autoregressive parameter and k is the lag. Such a plot could also be consistent with an AR model of order greater than one, for which the autocorrelation does not admit a simple representation, being defined recursively through a difference equation detailed in Hamilton (1994) among others. The introduction of heterogeneous effects (M3-M5) make the memory of the process decrease substantially especially when individual time effects are introduced (M3 and M5). For such models, the autocorrelation function has lower values but at the same time more complex dynamics. This suggests some MA processes but could also be the result of some ARMA processes.

We then use AIC and BIC to select the preferred model for each specification M1-M5 (see Table 2). They do not suggest the same model in only a few cases. In such cases when the competitive models are nested, we use the likelihood ratio test since we have access to the full likelihood. This is the case of M2 and M4 for EU-North. For the former, we contrast an ARMA(2,1) chosen with AIC with an ARMA (1,1) resulting from BIC while for the latter we compare an ARMA(1,2) with and ARMA (1,1). In both cases, the likelihood ratio test does not provide a very clear indication (p values equal to 0.14 and 0.07, respectively). In practice, due to the smaller penalty term, the AIC

	M1	M2	M3	M4	M5
Umbrella group					
AIC	AR(2)	AR(3)	MA(3)	AR(2)	MA(2)
BIC	AR(2)	AR(2)	WH	AR(2)	WH
EU North					
AIC	AR(1)	ARMA(2,1)	ARMA(1,1)	ARMA(1,2)	ARMA(1,2)
BIC	AR(1)	ARMA(1,1)	ARMA(1,1)	ARMA(1,1)	AR(2)
EU South					
AIC	ARMA(1,1)	ARMA(1,1)	MA(2)	ARMA(1,1)	MA(1)
BIC	ARMA(1,1)	ARMA(1,1)	MA(2)	ARMA(1,1)	MA(1)

ARMA: autoregressive moving average process,
AR: autoregressive process
MA: moving average process
WH: white noise process

Table 2: Model selection: step one

tends to keep more terms in the model than the BIC. Aiming at whitening residuals, we use the AIC to choose the preferred model.

Step two. Table 3 next compares the five selected specifications M1-M5 using the AIC. It also contrasts the serially uncorrelated model with the selected correlated model for each specification. In general, the AIC decreases remarkably moving from an uncorrelated to a correlated model. Only when individual time effects are introduced (M3 and M5) is the decrease weaker. Such an indication is also confirmed by approximate hypothesis testing. In all cases the likelihood ratio test rejects the constrained uncorrelated model, with p-values which are slightly higher for M3 and M5.

Comparing the AIC of the five selected correlated models allows us to choose the final model. For the Umbrella group and the EU-South group, we choose the M3 "random growth" specification (common income effect and heterogeneous time effect) with an MA(3) and MA(2) error structure respectively. For the EU-North group, the model selection procedure suggests M2 (common income, common time) with an ARMA(2,1) error structure.

Some relevant remarks are in order. First, having allowed for serially correlated errors has decreased the AIC (and BIC) in all five models M1-M5. However, such a decrease proved more important for rather constrained models (M1 and M2) and if we had contrasted the five models without introducing serial correlation, we would have erroneously selected the "unconstrained" M5

	M1	M2	M3	M4	M5
Umbrella group					
Uncorrelated	-623.4803	-642.3798	-974.1013	-885.1518	-964.7689
Correlated	-968.0401	-964.3868	-981.0451	-953.2796	-970.5999
	AR2	AR3	MA(3)	AR(2)	MA(2)
L.Ratio test: p-value	0.0001	0.0001	0.0048	0.0001	0.0073
EU North					
Uncorrelated	-444.2861	-447.1753	-703.5995	-1002.913	-1028.414
Correlated	-1083.0346	-1106.4520	-1083.9701	-1089.168	-1075.803
	AR(1)	ARMA(2,1)	ARMA(1,1)	ARMA(1,2)	ARMA(1,2)
L.Ratio: p-value	0.0001	0.0001	0.0001	0.0001	0.0001
EU South					
Uncorrelated	-660.5092	-711.0474	-1078.990	-1027.111	-1071.355
Correlated	-1056.1951	-1056.1757	-1086.611	-1062.452	-1081.073
	ARMA(1,1)	ARMA(1,1)	MA(2)	ARMA(1,1)	MA(1)
L.Ratio: p-value	0.0001	0.0001	0.003	0.0001	0.0006

The reported values are the corresponding AIC
ARMA: autoregressive moving average process
AR: autoregressive process
MA: moving average process

Table 3: Model selection: step two

model for the EU-North group. Second, once serial correlation and (heterogeneous) time effect have been accounted for, the *proper* effect of income appears to be homogeneous across countries. This result complements Mazzanti and Musolesi (2013). The specifications allowing heterogeneous income effect perform relatively well only for EU-North.

Next, we focus our attention on the selected models, the M3_MA(3) model for Umbrella, the M2_ARMA(2,1) for EU-North and the M3_MA(2) model for EU-South.

5.2 The selected models' estimates

Concerning the selected Umbrella's M3 "semiparametric random growth" specification with MA(3) errors, the resulting plots of the smooth terms are depicted in fig. 5. The component smooths are shown with confidence intervals that include the uncertainty about the overall mean (Marra and Wood, 2012).

All the smooths are highly significant (detailed approximate significance of the smooth terms available upon request).

For the Umbrella, while there is evidence of a homogenous and monotonic positive CO2-GDP relation, the CO2-time relation is overall roughly an inverted U for the USA, Canada and Japan, while it is positive for the other countries (Australia, Norway and New Zealand). Time related components drive the 'relative delinking' of such countries [8]. Specifically, though they have not achieved absolute reduction in CO2, the decrease in the CO2/GDP ratio is driven by factors that pertain to and are contained in the 'time-related black box'. Different 'Innovation intensities' (especially patented innovation) which have historically favoured Japan and the US (see for example Johnstone et al., 2012, 2010; Dechezlepretre et al., 2011), and which characterise the first set of countries, as well as the energy structure of the economy (namely endowments of carbon-intense sources that have penalised Oceania), could well explain group differences related to 'time factors'. This heterogeneity is hidden by common time factor specification; in fact, previous studies that highlighted the existence of only relative delinking for the Umbrella group did not unveil this existing heterogeneity in time effects (Mazzanti and Musolesi, 2013) [9]. Countries possessing larger stocks of (fossil fuel) resources have comparatively less incentives to increase efficiency through innovation and apply policies that reshape the energy structure towards coal-free sources. They are also less exposed to international energy shocks; the plots arguably show that the 1970's oil shocks supported significant a decrease in GHG emissions through energy mix changes and clean innovation diffusion (OECD, 2011). Future works might aim to discover which policy and market 'events' worked to break the long term CO2 trend.

Similar outcomes are revealed by the analysis of southern EU countries for which a M3 with MA(2) error structure for the EU-South (Fig. 6) has been selected. Figure 6 shows how once again we find a monotonic nonlinear postive CO2-income relation. We also note that Italy and Spain present an inverted U CO2-time component relation, while Portugal and Greece show a positive and monotonic relationship [10]. This is interesting since though the overall

[8]Relative delinking associates to a positive and lower than one elasticity between GDP and CO2 (taking the overall effect), while absolute delinking is related to a negative elasticity.

[9]Even heterogeneous panel models such as Swamy, mean group, CCEMG and Bayesian estimators unveiled the flaws of homogeneous estimators that might present EKC even in the presence of real monotonic GDP-CO2 trends on the one hand, but on the other they are not useful in detecting the specificity of idiosyncratic country effects.

[10]We note that we assess performances in terms of CO2 reductions compared to GDP and Time. Performances under the Kyoto Protocol also reflect 'distance from the target' assess-

EU-South performance has been deficient with respect to GHG reductions, a relevant country such as Italy (around 12% of the EU GDP) is at least compensating the GDP effect with a bell shaped time-CO2 link. More than to the intensity of clean innovation adoption and patents (Gilli et al., 2013; Johnstone et al., 2010), this is associated to the high energy efficiency and relevant share of renewables (hydroelectric) that Italy continued to present in the 1980-90's, and to the good long dynamic GHG performance of some of its industrial sectors (Marin and Mazzanti, 2013). Though Spain and Italy present rather different sector compositions of the economy (manufacturing-service shares) their time-related factors are on track to compensate the GDP scale effect in the future. Greece and Portugal are still on a development-oriented path that does not include CO2 reductions by policy and innovation factors. In fact, they were exempted from cutting CO2 by the Kyoto protocol as was Spain, a rare case in the EU. This (economically motivated) lack of stringent targets might have reduced the intensity of efforts made towards the achievement of joint economic-environmental goals.

Thus, the monotonic CO2-common time factor relation we drew out from other econometric models in the case of southern EU countries (Mazzanti and Musolesi, 2013) appears specifically driven by the poorest within the poorer set of countries in the EU. Though they do not massively impact on the overall EU GHG picture, the evidence signals risks of unsustainable convergence (assuming economic convergence was in place before the 2008-9 downturn) on the side of some peripheric countries [11].

Finally, we focus on the selected M2 "common time effects" model with ARMA(2,1) errors for EU-North countries. The results are reported in Figure 7. For the EU-North group, the CO2-time component relation is found to be homogeneous across countries and clearly negative. This indicates that the significant, unobserved time-related factors have negatively and primarily impacted CO2 emissions, as to more than compensate the GDP scale effect in some cases (the UK, Germany, Sweden and Finland are noteworthy examples of countries that have succeeded in reducing CO2 emissions). This evidence is coherent with recent information on the average EU performance (EEA, 2008). The factors explaining this evidence are largely linked to the way northern EU countries reacted to oil shocks, some as far back as the mid 1980's, mainly through energy saving and innovation actions. Such countries were then later

ment. Targets were bargained as reductions or increases with respect the 1990 benchmark on the basis of economic and 'political' considerations.

[11] We stick to production related emissions that are relevant for policy targets. The inclusion of Trade might change one country's performance (Marin et al., 2013; Levinson, 2009) depending upon the CO2 that is embodied in exports and imports.

characterised by a more stringent adoption of environmental policy (Johnstone et al., 2012), including a relatively larger use of market-based instruments such as carbon taxes in the 1990's . Among the EU-North member states, Denmark has historically had the highest environmental taxation as share of GDP figure in the EU according to Eurostat data (higher than 5%), while Sweden established itself as a prominent implementer of green fiscal reforms well back in the early 1990's, and still holds the highest carbon tax level in the EU - now more than 150$ per tonne on average (Andersen and Ekins, 2009). Shifting the tax burden to environmental taxation might promote long term welfare by enhancing elements not accounted for in the GDP (OECD; 2013), while it may also increase or not be detrimental to growth (Costantini and Mazzanti, 2012).

As additional corollary evidence, while it is true that Japan, Germany and the US rank in the first three positions in terms of climate change oriented patents, thus providing contents to the time component related evidence we commented on for Umbrella and the EU-North; Germany ranks first as far as the value of such patents is concerned (Dechelezpretre et al., 2011). Germany is a clear key player in EU-North performance given its weight and leadership in environmental technological development. The energy intensity of its GDP also reflects this: in 2009 Germany was slightly better than Japan and significantly better than the US (World Bank data[12]).

For EU_north, the specifications allowing heterogeneous income effect (M4 and M5) perform relatively well (see table 3) and thus we also provide in fig. 8 the plots of the smooth terms $f_i(x_{it})$ ffor the M4 model with ARMA (1,2) error structure. Fig. 8 presents a somewhat expected but interesting figure. Countries that are showing some signs of potential inverted U EKC even for the income-carbon relationship[13] are the 'usual' Scandinavian countries, often at the top of Human Development Indexes and competitiveness rankings: Denmark, Sweden and Finland.

It is worth noting that the evidence is absolutely coherent with the results of the EU-funded COMETR project that ex post evaluated the impact of carbon taxation associated to diverse assumptions on the recycling of

[12] The Economist, 10th August 2013, page 19.

[13] The finding of a bell shape for 'net' CO2-GDP relationships would be a radical result that, notwithstanding the mentioned role of trade patterns, relates to a strong decarbonization. We here note that, all things being equal, those 3 countries sum a robust time related component to a better than average GDP-CO2 effect. EKCs are thus explained by (1) a low elasticity concerning the net GDP-CO2 link (2) a robust negative link between time effects and CO2. Though simple, this argument has not really been touched upon in the literature by disentangling income and time effects by areas and country.

revenue though a general equilibrium model(Andersen, 2007). Results shows that Finland and Sweden integrate environmental and economic performances more than others (Gilli et al., 2013 presents sector evidence on Swedish economic - environmental performances). Sweden and Finland are among the countries (Germany and the UK as well) that have reduced CO2 compared to 1990 Kyoto benchmark levels. Though performing less than the other two countries in terms of GHG performance, Denmark has historically been the best performing EU country in terms of the energy intensity of its GDP.

The aforementioned results provide evidence on the fallacy of some simplistic EKC interpretations, and on the biased evidence that homogeneous and parametric settings may present. At least for Umbrella and the EU-South, it clearly shows that the relation CO2-time is also heterogeneous across countries; while income effect heterogeneity somewhat matters for the EU-North. We commented on diversified evidence of how carbon-time (and carbon-income) dynamics present highly idiosyncratic contents that deserve specific attention and can differentiate potentially similar countries. Our study and results partly refer to Melenberg et al. (2009) but provide new and more specific insights. In fact, we find consistent positive income effects for both cases [SO2, CO2] and time effect estimates with a clear U-shaped trend for SO2-emissions but only slightly so for CO2-emissions. The 'slight' time effect Melenberg et al. (2009) holistically treat is here investigated in depth by different models, across countries and groups of them.

Summing up, the main evidence we find is that only for northern EU countries the time effect nevertheless outweighs the GDP scale effect, which drives up CO2 emissions in any case, to different levels. This is the explanation behind the inverted U shaped curve, namely reduction of CO2 occurring while GDP grows[14], that recent studies found (Galeotti et al., 2009; Mazzanti and Musolesi, 2013) also for some EU areas, and the GHG accounting has shown over the past[15].

[14]In the IPAT (Impact=Population*affluence*technology) framework, this means that the factor T outweighs P and A. IPAT can be extended to include energy intensity issues, another effect that compensates for the enlarging scale of the economy.

[15]ICGC (2012) notes that while the final Kyoto phase is assessed: 'Luxembourg and Canada are the farthest from the emission levels they agreed to keep, by 29% and 27% respectively. Other countries that emitted more than the emissions budget agreed for this period are Austria, Iceland, New Zealand, Australia, the United States, Lichtenstein, Spain, Denmark, Switzerland, Slovenia, Norway, Italy, Japan, the Netherlands and Ireland'. Greece and Portugal are in more comfortable positions but they bargained substantial increases with respect to 1990.

6 Summary and conclusions

We examine long term carbon emissions-income relationships for the most advanced economies within OECD, given their role as leaders in the current climate change policy agenda. They are grouped into relevant policy groups: North America and Oceania, South Europe, North Europe.

Within the adopted groups' classification, however, the countries under study may differ with respect to some unobservable time varying variables such as technology, institutions, culture, etc., which linked to the production of CO2. Moreover, some unobserved global effects (such as those linked to the global public good nature of CO2) may affect heterogeneously different countries under study. Given this, an interesting feature of the proposed approach, compared with the related literature, is thus that it allows the estimation of the nonparametric time effect rather than considering it as nuisance term. This feature is very important from the point of view of an economic and policy oriented analysis because it allows for (i) obtaining a *proper* income effect; (ii) nonparametrically investigating how the time-related factors may drive the CO2 long-run evolution, perhaps heterogeneously across countries. The proposed model also allows for iii) handling serial correlation by using autoregressive moving average (ARMA) error processes. We rely on recent advances on Generalized Additive Mixed Models (GAMMs).

As preliminary analysis we propose a two step model selection procedure allowing us to choose both the degree of heterogeneity with respect to time and income, as well as the covariance structure. We emphasized that having allowed for serially correlated errors improves the statistical quality of the model expressed in terms of a trade off between goodness of fit and model complexity. We also pointed out that once serial correlation and (heterogeneous) time effect have been accounted for, the *proper* effect of income appears to be homogeneous across countries. Moreover, the introduction of heterogeneous time effects makes the memory of the process substantially decrease. For such models, the autocorrelation function has lower values but also more complex dynamics compared to models without time effects or which constrain such effects to be homogeneous across countries. In the end, our proposed model selection procedure allows us to choose a "semiparametric random growth" specification (common income effect and heterogeneous time effect) with an MA(3) and MA(2) error structure for the Umbrella group and the EU-South, respectively, while for the EU-North group, it suggests a common income and common time model with an ARMA(2,1) error structure. The specifications allowing heterogeneous income effect perform relatively well only for EU-North.

Concerning the main results of estimation, we find that when introducing a nonparametric time effect into the model, even if homogeneous across countries, EKC evidence is importantly affected if compared to previous studies (and to non reported results as well). We provide strong empirical support that the negative *total* pollution-development relationship (which includes indirect effects linked to the omitted or unobserved variables) that appears for some advanced countries is explained to a large extent by country-specific time-related factors that can outweigh the inevitable CO2 increasing GDP scale effect. More precisely, only three Scandinavian countries – Denmark, Finland and Sweden – present some threshold effect on the CO2-development relation, whereas for all other countries this relation appears to be monotonic and positive.

In other words, these results indicate that time-related factors were actually behind the reduction of CO2 in northern Europe - even during growth periods: for these countries, time related factors have been able to more than counterbalance GDP scale effects. At the same time, however, the CO2-time relationship is negative in the northern EU and in areas of OECD (North America, Japan, Italy) that stand out in terms of environmental invention and innovation, policy commitment and/or energy efficiency. In summary, the time effect nevertheless outweighs the GDP scale effect only for northern EU countries, which drives up CO2 emissions in any case, though with different intensities.

This strongly suggests the fallacy of the simplistic 'environmental Kuznets curve' argument when it does not account for and model specific time effects. Idiosyncratic elements related to energy, policy and innovation characterise heterogeneity, both across groups of countries and within groups of homogeneous countries as well.

Our analysis suggests some new research directions. First, taking a policy oriented perspective, future works should investigate the contents of the country specific time factors we here touched upon, for example by testing if and what policy and innovation adoption might break the EKC GDP driven ascending path. We have to learn from our past to inform the future post-Kyoto era agenda, which will witness a considerably different environment, with advanced countries still leading policy and technological domains, but main emerging countries hitting their environmental turning point by emitting around three-quarters of global CO2. Given inertia, structural breaks drivne by policy - technological dynamics are eventually necessary to cut emissions by 70-90% by 2050. Secondly, from a methodological perspective, we clearly presented evidence of dependence in the error structure, which is probably due to the imposed separability in the regression function. A natural step further would be to specifically test the separability of income and time effects and estimate a non separable model (with that respect we carried out some preliminary analyses, that are available upon request). Alternatively, within the separable framework, it is worth noting that the identified error structure implies eventual omitted dynamics that could be modeled directly as part of the regression function. This strategy is possibly providing complementary insights.

References

[1] Andreoni J., Levinson A. (2001), The simple analytics of the environmental Kuznets curve, Journal of Public Economics, 80, 269-86.

[2] Andersen M.S., Ekins P. (2009), Carbon Taxation: Lessons from Europe, Oxford University press, Oxford/NY.

[3] Andersen, M.S. et al. (2007), Competitiveness Effects of Environmental Tax Reforms (COMETR): Publishable Final Report to the European Commission, http://www2.dmu.dk/cometr

[4] Augustin NH, Musio M, von Wilpert K, Kublin E, Wood SN, Schumacher M (2009), Modeling spatiotemporal forest health monitoring data. J Am Stat Assoc 104(487), 899–911.

[5] Azomahou T., Laisney F., Phu Ngayen V. (2006), Economic development and CO2 emissions: a non parametric panel approach, Journal of Public Economics, 90, 1347-63.

[6] Azomahou, T., T. Mishra, 2008. Age dynamics and economic growth: Revisiting the nexus in a nonparametric setting, Economics Letters vol. 99(1), pages 67-71.

[7] Azomahou T., Goedhuys M., Phu Ngayen V. (2009), A structural nonparametric reappraisal of the CO2 missions-income relationship, ONU-MERIT working paper.

[8] Barrett S. (2003), Environment And Statecraft: The Strategy of Environmental Treaty-making, Oxford University Press.

[9] Borghesi S. (2001), The environmental Kuznets curve: a critical survey. In: Franzini, M., Nicita, A. (Eds.), Economic Institutions and Environmental Policy. Ashgate Publishing, Farnham, UK, pp. 201–224.

[10] Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. Journal of the American Statistical Association, 88, 9–25.

[11] Brock W., Taylor S. (2010), The Green Solow Model, Journal of Economic Growth, 15, 127-53.

[12] Chamberlain, G. (1982), Multivariate regression models for panel data. Journal of Econometrics 18 (1), 5–46.

[13] Costantini V. Mazzanti M. (2012), On the Green side of Trade competitiveness? Research Policy, 41.

[14] Dechezlepretre, A., Glachant, M., Hascic, I., Johnstone, N. and Meniere, N. (2011) Invention and transfer of climate change mitigation technologies on a global scale: a study drawing on patent data,

[15] Dietz S. (2011), The Stern Review, in Stephen H. Schneider (ed.) The Encyclopedia of Climate and Weather, Oxford: Oxford University Press.

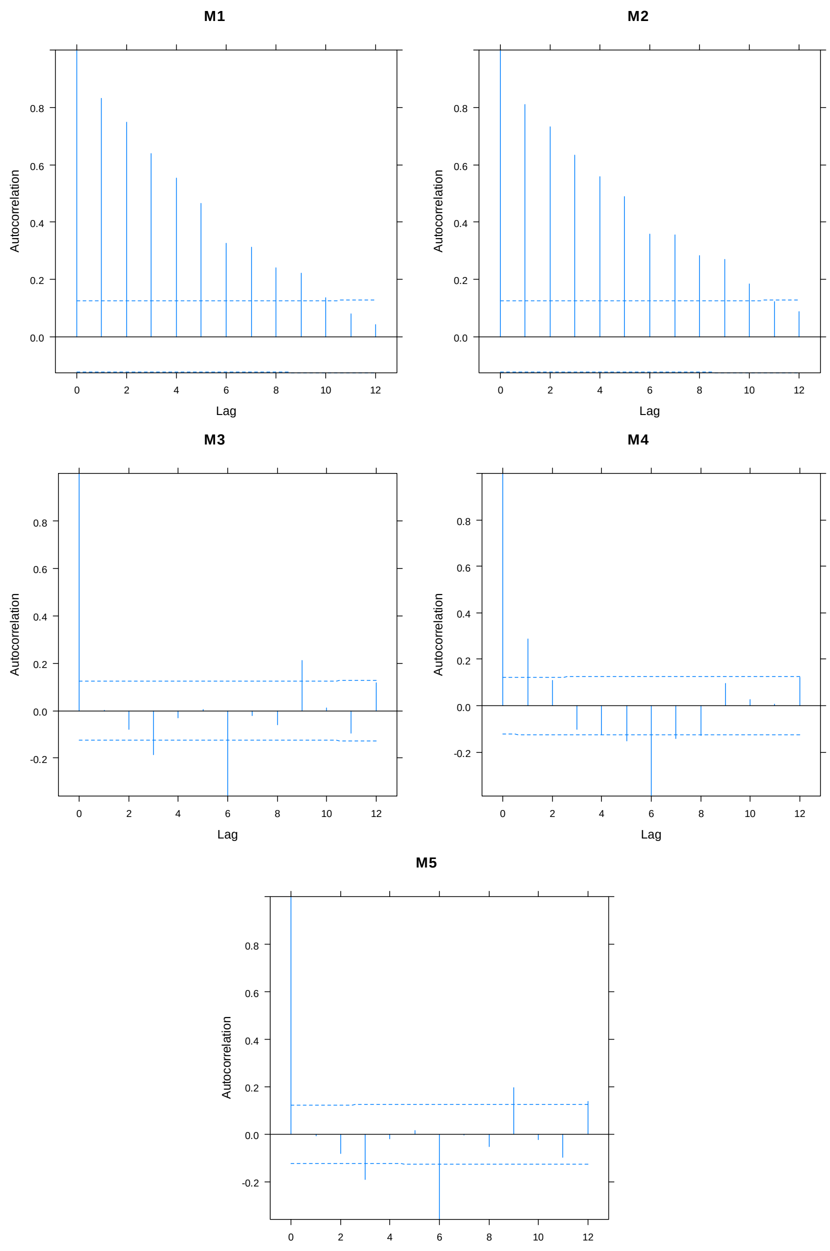
[16] EEA (2013), Towards a Green Economy in Europe, Copenhagen: European Environment Agency.
[17] EEA (2008), Greenhouse gas emission trends and projections in Europe 2008, Copenhagen: European Environment Agency.
[18] Egli H. Steger T. (2007), A Dynamic model of the environmental Kuznets curve: turning point and public policy, Environmental & Resource Economics, 36, 15-34.
[19] Evdokimov, K. (2010) Identification and Estimation of a Nonparametric Panel Data Model with Unobserved Heterogeneity, mimeo
[20] Figueroa E. Pasten R. (2013), A tale of two elasticities: a general theoretical framework for the environmental Kuznets curve analysis, Economics Letters, 119, 85-88.
[21] Friedberg, L. (1998), Did Unilateral Divorce Raise Divorce Rates?, American Economic Review 88, 608-627.
[22] Galeotti, M., Manera, M. and Lanza, A. (2009) On the robustness of robustness checks of the EKC hypothesis, Environmental and Resource Economics, 43, 369–90
[23] Gilli M. Mazzanti M. Nicolli F. (2013), Sustainability, Environmental Innovations and competitiveness in evolutionary perspectives. Evidence from the EU, Journal of Socio Economics, June.
[24] Grossman G.M., Krueger A.B. (1995), Economic Growth and the Environment, Quarterly Journal of Economics May, 353-357.
[25] Gu, C and Wahba, G. (1993) Semiparametric analysis of variance with tensor product thin-plate splines. J. R. Statist. Soc. B 55(2), 353-368.
[26] Hamilton, James D. (1994). Time Series Analysis. Princeton: Princeton University Press.
[27] Hastie, T. Tibshirani R. (1990), Generalized Additive Models, Chapman and Hall.
[28] Heckman, J.J., Hotz V.J. (1989), Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training, Journal of The American Statistical Association, 84(408), 862-874.
[29] Henderson, D. J., R. J. Carroll, and Q. Li. 2008. Nonparametric estimation and testing of fixed effects panel data models. Journal of Econometrics 144: 257-275.
[30] Hoderline, S. and H. White. (2012). Nonparametric identification in nonseparable panel data models with generalized fixed effects. Working Paper, Dept. of Economics, Brown University.
[31] Hsiao C. (2003), Analysis of panel data. Cambridge University Press, Cambridge

[32] Kijima M., Nishide K., Ohyama A. (2010), Economic models for the EKC: a survey, Journal of Economic Dynamics and Control, 34, 1187-1201.
[33] Kauermann,G., Krivobokova,T., and Fahrmeir,L. (2009), Some asymptotic results on generalized penalized spline smoothing. J. R. Statist. Soc.B 71, 487–503.
[34] Johnstone N., Hascic I., Popp, D. (2010), Renewable energy policies and technological innovation: evidence based on patent counts, Environmental & Resource Economics, 45, 133-55
[35] Johnstone N. Haščič I. Poirier J. Hemar M. Michel C., (2012). Environmental policy stringency and technological innovation: evidence from survey data and patent counts, Applied Economics, vol. 44(17), pages 2157-2170
[36] ICCG (2012), The state of compliance in the Kyoto Protocol, ICCG Reflection 12, ICCG.
[37] Levinson, A. (2009). Technology, International Trade, and Pollution from US Manufacturing, American Economic Review, 99(5): 2177-92
[38] Li, Q. and T. Stengos. 1996. Semiparametric estimation of partially linear panel data models. Journal of Econometrics 71: 289-397.
[39] Mammen, E., Stove, B., and Tjostheim, D., 2009, Nonparametric additive models for panels of time series, Econometric Theory, 25, 442-481.
[40] Marin G., Mazzanti M., (2013), The evolution of environmental and labour productivity dynamics, Journal of Evolutionary Economics, April.
[41] Marra, G and S.N. Wood (2012) Coverage Properties of Confidence Intervals for Generalized Additive Model Components. Scandinavian Journal of Statistics, 39(1), 53-74.
[42] Melenberg, B. Dijkgraaf, E. Vollebergh, H., (2009), Identifying reduced-form relations with panel data: The case of pollution and income, Journal of Environmental Economics and Management, 58, 27-42.
[43] Millimet D., List J., Stengos T. (2003), The EKC: real progress or misspecified models? The Review of Economics and Statistics, 85, 1038-1047.
[44] Mazzanti M. Musolesi A. (2013), The heterogeneity of carbon Kuznets curves for advanced countries: comparing homogeneous, heterogeneous and shrinkage/Bayesian estimators, Applied Economics, 45, 3827-42.
[45] Musolesi A., Mazzanti M., Zoboli R. (2010), A Bayesian Approach to the estimation of EKC for CO2, Applied Economics, 42, 2275-87.
[46] OECD (2002), Indicators to Measure Decoupling of Environmental Pressure from Economic Growth, Paris, OECD.
[47] OECD, (2010), Taxation, Innovation and the Environment, OECD Green Growth Studies, OECD Publishing, Paris.
[48] OECD, (2011), Fostering Innovation for Green Growth, OECD Green

Growth Studies, OECD Publishing, Paris
[49] OECD (2013), Choosing fiscal consolidation instruments compatible with growth and equity, OECD Publishing, Paris
[50] Ordas Criado, C., Valente, S. and Stengos, T. (2011), Growth and the Pollution Convergence Hypothesis: : Theory and Evidence, Journal of Environmental Economics and Management, 2011, 62 , 199—214.
[51] Papke, L.E. (1994), Tax Policy and Urban Development: Evidence from the Indiana Enterprise Zone Program, Journal of Public Economics 54, 37-49.
[52] Pesaran, M. H. (2006). Estimation and inference in large heterogenous panels with multifactor error structure. Econometrica 74, 967–1012.
[53] Pinheiro, J. and Bates, D. (2000). Mixed-effects models in S and S-PLUS. Springer-Verlag, New York.
[54] Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and the R Core team (2013). nlme: Linear and Nonlinear Mixed Effects Models.
[55] Ruppert, D., Wand, M. P., and Carroll, R. J., (2003), Semiparametric Regression. Cambridge: Cambridge University Press.
[56] Su, L., and A. Ullah (2006). Profile likelihood estimation of partially linear panel data models with fixed effects, Economics Letters 92: 75-81.
[57] Su, L., and A. Ullah. (2010), Nonparametric and Semiparametric Panel Econometric Models: Estimation and Testing, mimeo.
[58] Stone, C.J. (1985),"Additive Regression and Other Nonparametric Models", Annals of Statistics, 13, 689-705.
[59] UNDP (2009), Human Development Report 2007/2008. Fighting climate change: Human solidarity in a divided world, UNDP.
[60] Wood, S.N. (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. J.R. Statist. Soc. B 62, 413-428.
[61] Wood, S.N. (2003) Thin plate regression splines. J.R. Statist.Soc.B 65(1), 95-114.
[62] Wood, S.N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models, Journal of the American Statistical Association. 99:673-686.
[63] Wood S.N. (2006a), Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC Press.
[64] Wood, S.N., (2006b), On confidence intervals for generalized additive models based on penalized regression splines, Australian and New Zealand Journal of Statistics. 48(4): 445-464.
[65] Wood, S.N. (2008), Fast stable direct fitting and smoothness selection for generalized additive models. J.R.Statist.Soc.B 70(3):495-518.
[66] Wood, S.N. (2011) Fast stable restricted maximum likelihood and

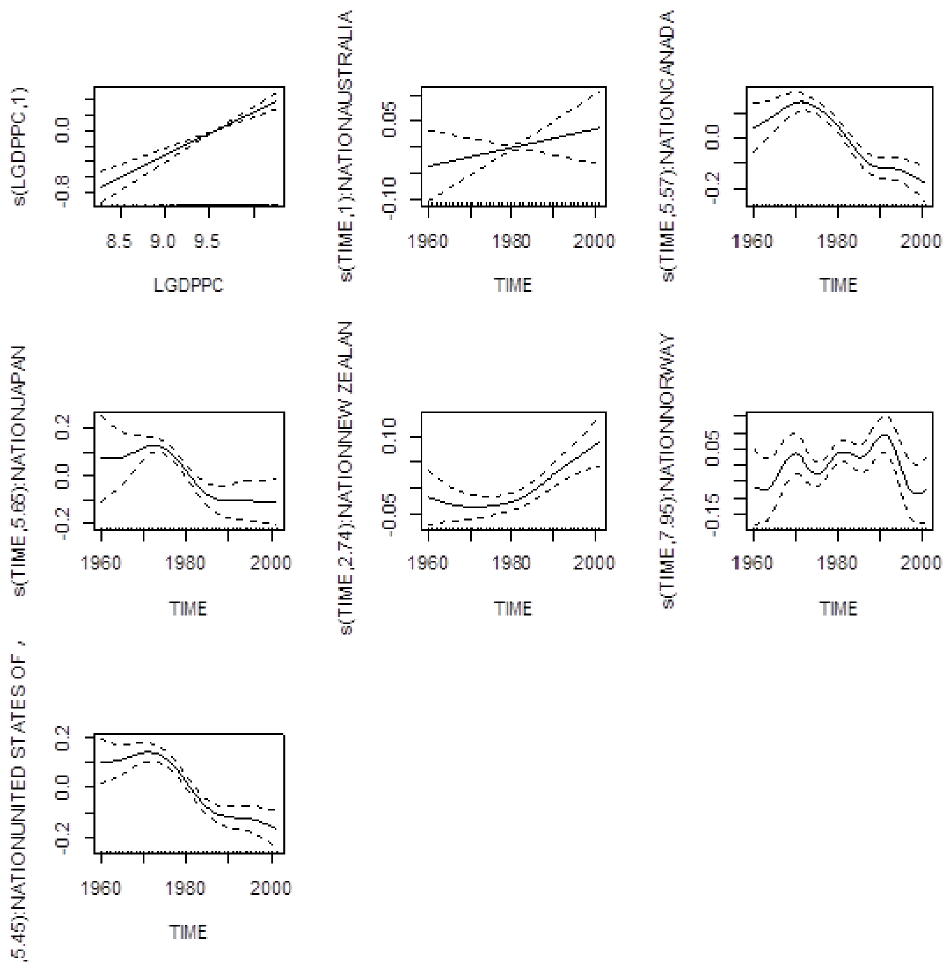
marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36.
[67] Wood, S.N. (2013), mgcv R package, CRAN.
[68] Wooldridge J.M, (2005), Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models," The Review of Economics and Statistics, MIT Press, vol. 87(2), pages 385-390.
[69] Yoshida,T. and Naito, K. (2012), Asymptotics for penalized additive B-spline regression, Journal of the Japan Statistical Society, 42 (1), 81–107.
[70] Yoshida,T. and Naito, K. (2013), Asymptotics for penalized splines in generalized additive models, mimeo.

Figure 4: Empirical autocorrelation function. Umbrella.



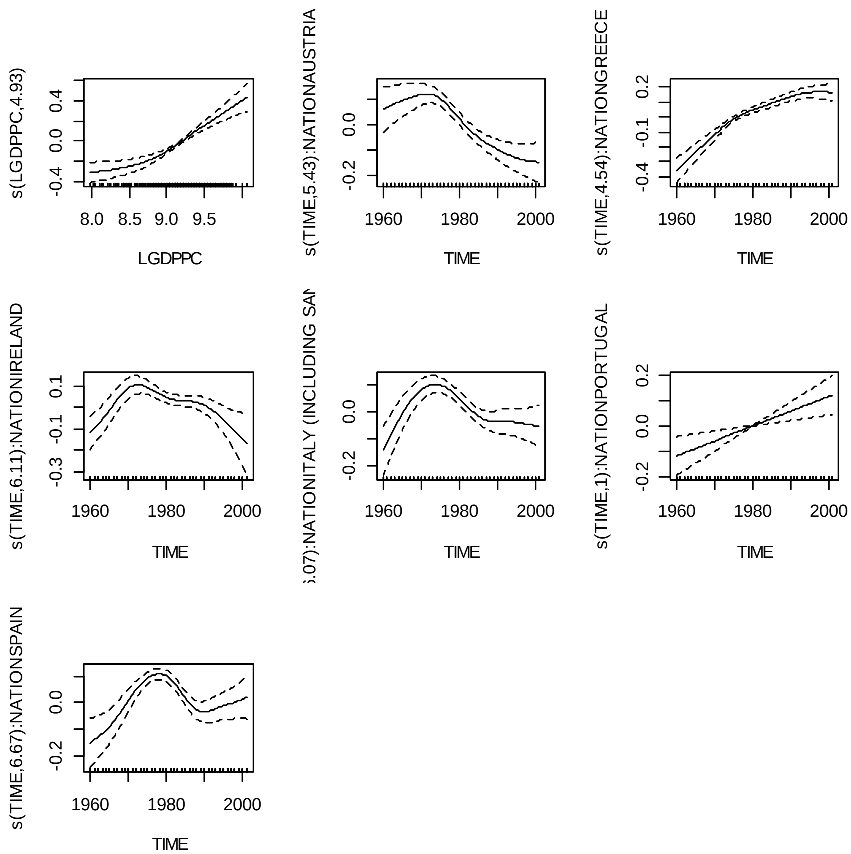
Standardized residuals: raw residuals divided by the standard errors. Approximate two-sided critical bounds (0.05 level).

Figure 5: Umbrella (M3, MA(3))



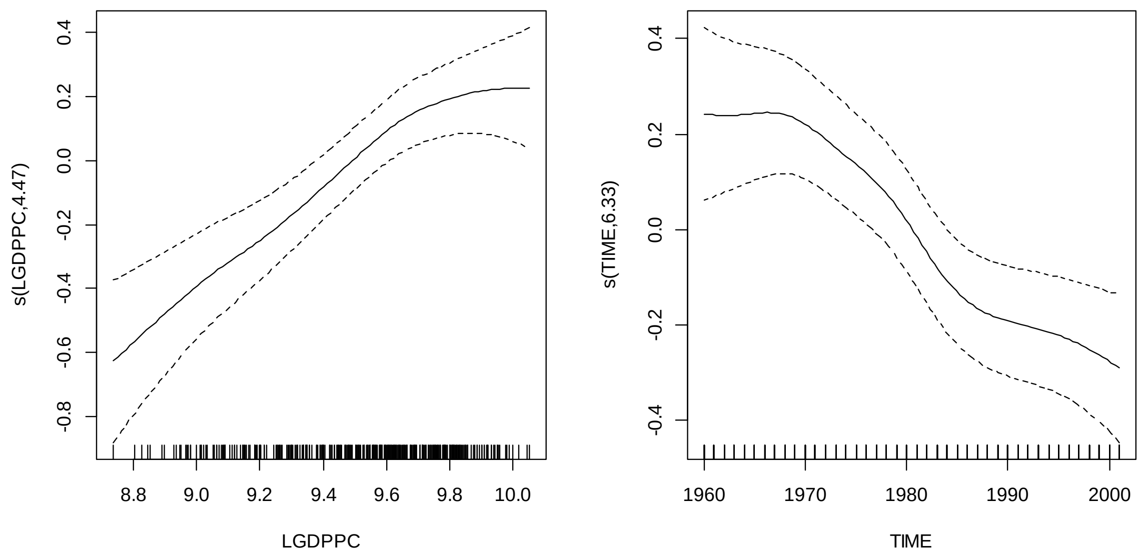
Notes. s(LGDPPC, edf) indicates the estimated smooth function (and its 95% confidence interval) of log (GDP per capita) and edf represents the estimated degrees of freedom.
s(TIME, edf)NATION "NAME OF THE COUNTRY" indicates the estimated smooth function (and its 95% confidence interval) of the "factor-by-curve interaction" (interaction between the common trend and the country's indicator variable) and edf represents the estimated degrees of freedom.

Figure 6: EU_sud (M3, MA(2))



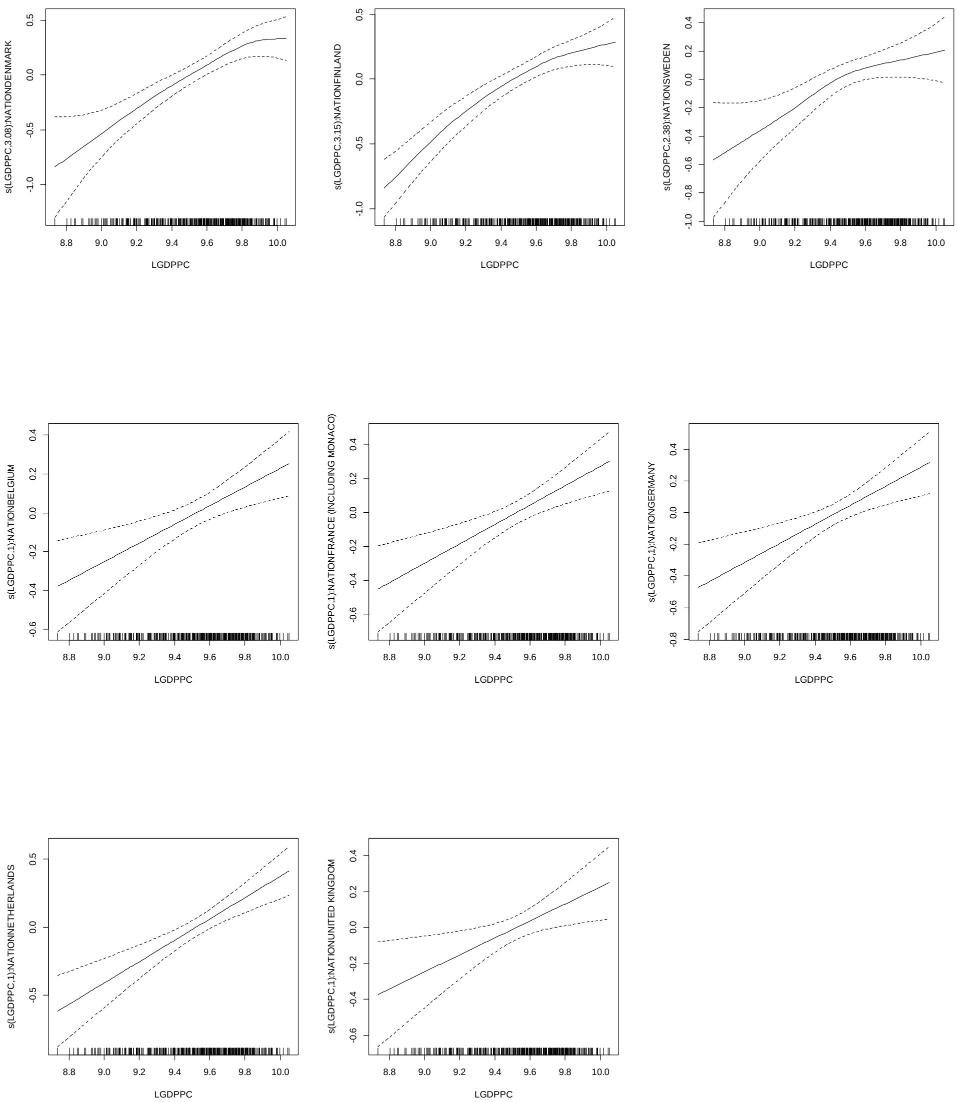
Notes. s(LGDPPC, edf) indicates the estimated smooth function (and its 95% confidence interval) of log (GDP per capita) and edf represents the estimated degrees of freedom.
s(TIME, edf)NATION “NAME OF THE COUNTRY” indicates the estimated smooth function (and its 95% confidence interval) of the “factor-by-curve interaction” (interaction between the common trend and the country’s indicator variable) and edf represents the estimated degrees of freedom.

Figure 7: EU_nord (M2, ARMA(2,1))



Notes. s(LGDPPC, edf) indicates the estimated smooth function (and its 95% confidence interval) of log (GDP per capita) and edf represents the estimated degrees of freedom. $s(TIME, edf)$ indicates the estimated smooth function (and its 95% confidence interval) of time and edf represents the estimated degrees of freedom.

Figure 8: EU_nord (M4, ARMA(1,2)): heterogeneous income effect



Notes. s(LGDPPC, edf) NATION "NAME OF THE COUNTRY" indicates the estimated smooth function (and its 95% confidence interval) of the "factor-by-curve interaction" (interaction between log(GDP per capita) and and the country's indicator variable) and edf represents the estimated degrees of freedom.