Growth and Pollution Convergence: Theory and Evidence

C. Ordás Criado[*] [†] [‡] S. Valente[§] T. Stengos[¶]

Final version: March 14, 2011

Running title: Growth and Pollution Convergence

[*]Corresponding author. Email: cordas@ethz.ch; fax: +41 44 632 1622.
[†]ETH Zurich, Center for Energy Policy and Economics (CEPE), 8032 Zurich, Switzerland.
[‡]Université Laval, Département d'économique, G1V0A6, Québec, Canada.
[§]ETH Zurich, Center of Economic Research (CER), 8032 Zurich, Switzerland.
[¶]University of Guelph, Department of Economics, N1G2W1, Ontario, Canada.

Growth and Pollution Convergence: Theory and Evidence

Abstract

Stabilizing pollution levels in the long run is a pre-requisite for sustainable growth. We develop a neoclassical growth model with endogenous emission reduction predicting that, along optimal sustainable paths, pollution growth rates are (i) positively related to output growth (scale effect) and (ii) negatively related to emission levels (defensive effect). This dynamic law reduces to a convergence equation that is empirically tested for two major and regulated air pollutants - sulfur oxides and nitrogen oxides - with a panel of 25 European countries spanning the years 1980-2005. Traditional parametric models are rejected by the data. More flexible regression techniques confirm the existence of both the scale and the defensive effect, supporting the model predictions.

JEL Classification numbers: C14, C23, O13, Q53

Key Words: Air pollution, convergence, economic growth, nonparametric regressions.

1 Introduction

Making economic growth compatible with environmental preservation in the long run is a major challenge facing modern societies. The economic literature tackles this issue by analyzing the conditions under which an economy may achieve sustainable growth – that is, balanced growth paths characterized by growing per capita incomes and non-declining environmental quality (Brock and Taylor, 2005). In this framework, sustainability requires satisfying a general condition of *pollution convergence*: pollution must be bounded in the long run and approach a finite steady state level despite positive growth in GDP per capita. A more specific question is: how do pollution dynamics interact with output dynamics along sustainable growth paths? In this paper, we tackle this issue at both the theoretical and the empirical level. First, considering a growth model with endogenous pollution abatement, we show that the optimal path is characterized by a precise dynamic relationship between pollution growth rates, emission levels, and output growth rates, which induces pollution convergence in the long run. Second, we test this dynamic law empirically for two major air pollutants, using panel data from European countries.

Our analysis is based on a neoclassical growth model where pollution generated by the production process reduces private welfare. In this framework, sustainable growth requires the use of more efficient technologies as well as defensive expenditures to curb emissions (Van der Ploeg and Withagen, 1991; Bovenberg and Smulders, 1995; Brock and Taylor, 2010). We analyze the optimal path of an economy in which purposeful investment in clean technologies generates positive feedback effects on environmental quality: output growth is driven by capital accumulation and labor-augmenting technological progress whereas pollution growth is contrasted by emission-reducing technical change. Differently from Brock and Taylor (2010), we assume that both the propensities to consume and to invest in clean technologies are endogenously determined by utility maximization. We show that pollution stabilization in the long run is associated with a precise law: during the whole transition, the growth rate of emissions per capita is (i) negatively related to the level of emissions per capita and (ii) positively related to the growth rate of output per capita. Result (i) is called 'defensive effect', and reflects the effectiveness of abatement expenditures in limiting pollution growth. Result (ii) is a 'scale effect' implied by the positive relation between output and emission levels. By virtue of the defensive effect, pollution growth rates regress to zero and emissions per capita are bounded in the long run. The dynamic law derived in the theoretical model may be interpreted as a *β-convergence equation*, i.e., an intertemporal relation predicting an inverse relationship between the growth rate of the variable of interest (pollution, in our case) and its past level.

The notion of β-convergence is widely used in the growth literature, where the variable of interest is per capita income. Indeed, the neoclassical growth model predicts an inverse relationship between output levels and growth rates during the transition to the long run equilibrium. Accordingly, the empirical literature tests whether this prediction is observed across countries. Early studies focused on *absolute* β-convergence, i.e., the hypothesis that the negative growth-level relationship is characterized by the same parameters across the different economies considered in the sample[1]. More recent studies allow for structural heterogeneities among countries and test for *conditional* β-convergence, i.e., check the existence of a negative growth-level relationship in incomes after controlling for the effects of other determinants of the growth dynamics (Barro and Sala-i-Martin, 2004). In the present paper, we study pollution dynamics in a similar fashion but with an additional element: our model predicts not only a negative growth-level relationship in the variable of interest "pollution", but also a positive interaction between pollution growth and income growth. Our empirical analysis can thus be considered a convergence test in which β-convergence in pollution is conditional on country-specific output dynamics.

[1]The hypothesis of absolute beta-convergence seems to hold within selected groups of industrialized economies - namely, OECD countries in the post-war period - but it does not hold if the country sample is extended to include non-OECD economies: see, e.g., Acemoglu (2009).

We empirically test the existence of both the scale and the defensive effect for two major air pollutants, sulfur oxides (SOX) and nitrogen oxides (NOX), using panel data for 25 Eastern and Western European countries over the period 1980-2005. We consider different regression methods. The standard parametric approach confirms the existence of these effects for both pollutants but the linear models are rejected by the data. We address this issue by exploiting more flexible approaches - i.e., semi-parametric and nonparametric regressions - that better capture nonlinearities and heterogeneities across Eastern and Western European countries. Our results are consistent with the predictions of the theoretical model and confirm the existence of scale effects and defensive effects for SOX and NOX.

Our analysis differs from the previous literature on pollution dynamics in several respects. Most studies examine alternative notions of pollution convergence. A first body of contributions analyses *stochastic convergence*, that is, the time-series properties of pollution differentials between countries or regions. List (1999) and Bulte et al. (2007) find stationary gaps across the US states for NOX and SO2 emissions per capita especially since the start of the federal regulation period in the 1970s; the results of Nguyen Van (2005) and Aldy (2006) show worldwide divergence for CO2 emissions per capita since the 1960s, while Strazicich and List (2003) and Romero-Ávila (2008) obtain convergence among OECD countries over the same period. A second approach is followed by Aldy (2006), who analyses *σ-convergence* – i.e., the evolution of simple dispersion measures over time – and confirms the above results for CO2. A third notion of convergence, proposed by Quah (1993a), focuses on *distributional dynamics*[2] and examines the evolution of the spatial distribution of pollution levels over time. In this framework, Nguyen Van (2005), Aldy (2006) and Ordás Criado and Grether (forthcoming) reject the existence of polarization phenomena across countries except within the OECD and European areas.

Before our paper, the notion of pollution β-convergence has been analyzed in List (1999)

[2]This method is also referred to as Markov transition matrices or 'stochastic kernels'.

for NOX and SO2, and in Strazicich and List (2003), Nguyen Van (2005) and Brock and Taylor (2010) for CO2 emissions per capita. These authors test the slope of a log-linear relation between pollution growth and pollution levels and obtain a negative coefficient. Differently from these contributions, (i) we estimate a reduced-form equation derived from the optimality conditions of a dynamic model, (ii) we explicitly test for misspecification and employ nonparametric methods, and (iii) we consider NOX and SOX emissions at the European level. In particular, the choice of analyzing SOX and NOX is linked to the assumptions of the theoretical model, in which environmental quality is preserved through defensive expenditures, and pollution is represented as a local welfare-reducing flow.

As noted by Friedman (1992), Quah (1993b) and Evans and Karras (1996), among others, beta-convergence in income does not formally imply decreasing gaps in income across countries over time, nor it does guarantee stationarity for the income levels of the countries included in the sample. In line with these considerations, our results do not represent an estimate of the speed of convergence in pollution across countries, but rather an empirical test of the existence of the defensive effect – a necessary condition for sustainability in the long run.

2 A model of growth and optimal emission reduction

The relationships between economic growth and pollution dynamics are investigated by a growing body of theoretical literature. In the traditional approach – pioneered by Keeler et al. (1971), and extended by Van der Ploeg and Withagen (1991) and Bovenberg and Smulders (1995), amongst others - pollution is positively related to output levels according to an *emission function* representing the environmental damage caused by the production process. If the emission intensity is fixed, pollution increases linearly with production. However, if the economy invests resources in the development of cleaner technologies, the elasticity of emissions to output is reduced over time and economic growth induces positive feedback effects on environmental quality. This mechanism is emphasized by Brock and Taylor (2010) in a 'Green Solow Model' – i.e., a neoclassical growth model in which the economy spends fixed fractions of income in capital investment and in pollution-abatement activities – in which pollution converges to a finite steady state if the rate of emission-reducing technological progress is sufficiently high[3].

In this section, we build a model of optimal emission reduction that allows us to study the interactions between pollution convergence and capital accumulation in a more specific context. Differently from Brock and Taylor (2010), we assume that both the saving rate and the propensity to spend in abatement are endogenously determined by utility maximization. This allows us to derive the basic equation to be employed in the empirical analysis from the optimality conditions of a centralized social problem: the saddle path followed by the economy determines a precise dynamic relationship between pollution growth, pollution levels, and income growth rates, that can be tested using regression methods.

At the conceptual level, the difference with Brock and Taylor (2010) is twofold. First, the equation describing pollution convergence over time is explicitly micro-founded. Second, because we assume optimal control of the pollution flow at the economy level, our framework is particularly suited for applications to local air pollutants – e.g., NOX and SOX – but not to CO2, which is the pollutant examined in most related literature. At the formal level, our model shares the general features of Ramsey-Cass-Koopmans models with welfare-reducing emissions. Van der Ploeg and Withagen (1991) study several variants of this framework in the absence of technological progress - they consider, in particular, a 'flow-pollution model' where emission-reducing investment is excluded, and a 'clean-technologies model' where abatement is optimized but there is no capital accumulation. Our model can be interpreted either as an extension of the 'flow-pollution model' to include optimal investment

[3]The aim of Brock and Taylor (2010) is primarily to rationalize the existence of an Environmental Kuznets Curve (EKC) in a simple theoretical framework including capital accumulation and emission abatement. They exploit the work of Mankiw, Romer and Weil (1992) to derive a convergence equation for pollution.

in emission reduction, or equivalently, as an extension of the 'clean-technologies model' to include capital accumulation. In both cases, the value added of our analysis is the derivation of the joint dynamics of pollution, capital and output in an environment where the paths of these variables are optimally chosen and there are positive rates of labor-augmenting and emission-reducing technological progress.

2.1 The Ramsey setting

As our empirical analysis will focus on the dynamics of emissions per capita, we will treat pollution as a flow-variable that affects private utility in per capita terms. The model economy is characterized by the following assumptions:

$$Y(t) = F(K(t), B(t)N(t)), \quad B(t) = B_0 e^{\pi t}, \quad N(t) = N_0 e^{nt}, \tag{1}$$

$$\dot{K}(t) = Y(t) - C(t) - X(t) - \delta K(t), \tag{2}$$

$$P(t) = \tau(t) Y(t), \tag{3}$$

$$U(t) = U(\bar{c}(t), \bar{p}(t)), \quad U_{\bar{c}} > 0, \quad U_{\bar{c}\bar{c}} \leq 0, \quad U_{\bar{p}} < 0, \quad U_{\bar{p}\bar{p}} < 0, \tag{4}$$

where $t \in [0, \infty)$ is the time index.[4] Technology (1) assumes that aggregate output, Y, is produced by means of capital, K, and efficient labor, BN, according to a linearly homogeneous production function $F(.,.)$ displaying positive and strictly decreasing marginal productivities in both factors and satisfying the Inada conditions. Population N grows at the exogenous rate $n > 0$, and labor efficiency B grows at the given rate of labor-augmenting technical progress $\pi > 0$. Expression (2) is the accumulation constraint, where $\delta \geq 0$ is the rate of physical depreciation of capital: net investment equals output minus the sum of consumption, C, and defensive expenditures, X. By defensive expenditures we mean resources devoted to activities that reduce the emission intensity of the production sector –

[4]Using standard notation, we define $\dot{Z} \equiv dZ/dt$ as the time-derivative of the generic variable $Z(t)$, and $G_Q \equiv \partial G/\partial Q$ and $G_{QQ} \equiv \partial^2 G/\partial Q^2$ as the partial derivatives of the generic function $G(Q)$.

henceforth called ICT (investment in cleaner technology). The pollution function (3) asserts that aggregate emissions per unit of time, P, are proportional to aggregate output, and τ is the global emission intensity. Expression (4) defines private utility, U, as a function of consumption per capita, $\bar{c} \equiv C/N$, and emissions per capita, $\bar{p} \equiv P/N$, where $U_{\bar{p}} < 0$ and $U_{\bar{p}\bar{p}} < 0$ guarantee that the disutility from pollution is convex.

In order to obtain a full characterization of the optimal dynamics, we model (i) the process of emission reduction and (ii) the trade-off between consumption and environmental quality by means of two specifications often exploited in the literature. First, following Brock and Taylor (2010), we assume that the aggregate emission intensity is given by

$$\tau(t) \equiv \Omega(t) \cdot \left[1 - \frac{X(t)}{Y(t)}\right]^{\varepsilon}, \quad \varepsilon > 1, \tag{5}$$

where $\Omega(t)$ is the *baseline emission intensity*, and the second term is a function representing the effects of ICT. The share of output devoted to defensive expenditures, X/Y, will be called *ICT effort*, and corresponds to the propensity to invest in cleaner technologies, bounded between zero and unity. From (5), maximal ICT effort, $X = Y$, implies zero emissions whereas zero defensive expenditures, $X = 0$, imply that the emission intensity equals the baseline level. Also the baseline intensity varies over time as it is influenced by technological progress. As shown by Brock and Taylor (2010), explosive dynamics in pollution per capita can be avoided only if the rate of emission-reducing progress – i.e., the effects of technical improvements that reduce $\Omega(t)$ over time – is at least equal to the rate of output-augmenting technical progress. In the present model, this sustainability condition corresponds to $\dot{\Omega}(t)/\Omega(t) \leq -\pi$. In order to ensure that stabilizing per capita emissions in the long run is technically feasible, we assume symmetric rates of emission-reducing and labor-augmenting progress by setting $\Omega(t) = \Omega_0 e^{-\pi t}$. Alternative assumptions that satisfy the sustainability condition – e.g., $\Omega(t) = \Omega_0 e^{-\omega t}$ with $\omega > \pi$ – would complicate the analysis of steady-state equilibria without affecting the main results concerning pollution convergence.[5]

The second assumption is related to private preferences. In this respect, we specify a utility function that allows us to characterize optimal dynamics analytically:

$$U\left(\bar{c}\left(t\right),\bar{p}\left(t\right)\right)=\sigma\ln\bar{c}\left(t\right)-\varsigma\bar{p}\left(t\right)^{\theta},\quad \theta>1, \tag{6}$$

where $\sigma > 0$ and $\varsigma > 0$ are the weights on utility from consumption and disutility from pollution, respectively. Function (6) satisfies all the properties listed in (4), consistently with the conditions for a well-behaved problem – see Van der Ploeg and Withagen (1991). In particular, $\theta > 1$ ensures that the marginal damage from emissions is increasing.

As noted before, the distinguishing feature of our analysis is that we assume both the consumption and the ICT time paths to be chosen optimally. Denoting by $\rho > 0$ the social discount rate, the optimal path is defined as a sequence of consumption levels and defensive expenditures maximizing present-value welfare

$$\int_0^{\infty} U\left(\bar{c}\left(t\right),\bar{p}\left(t\right)\right)e^{-\rho t}dt, \tag{7}$$

subject to the accumulation constraint (2), the pollution function (3), and the non-negativity constraint $K\left(t\right) \geq 0$ in each t for a given initial stock $K\left(0\right) = K_0$. This problem can be solved more easily by denoting ICT effort as

$$\chi\left(t\right)\equiv X\left(t\right)/Y\left(t\right),$$

and normalizing the relevant variables in terms of labor-efficiency units. Setting $y \equiv Y/\left(BN\right)$

[5] Assuming more intense emission-reducing technical progress, $\dot{\Omega}/\Omega = -\omega < -\pi$, would not affect the main results: pollution per capita would tend to a particular steady-state level (i.e., zero) in the long run. We assume symmetric rates of emission-reducing and labor-augmenting progress because asymmetric rates would imply additional technical difficulties without any gain for the present analysis. This point is clarified in an Appendix available at JEEM's online archive of supplementary material, which can be accessed at http://aere.org/journals/.

and $k \equiv K/(BN)$, the homogeneous production function in (1) yields the intensive form $y = f(k) = F(k,1)$, where f_k coincides with the marginal product of capital. As a consequence, equations (2)-(3) can be written as

$$\dot{k}(t) = f(k(t))[1-\chi(t)] - c(t) - (\delta+n+\pi)k(t), \tag{8}$$

$$p(t) = \Omega(t)[1-\chi(t)]^{\varepsilon} f(k(t)), \tag{9}$$

where $c \equiv C/(BN)$ and $p \equiv P/(BN)$ are 'normalized' consumption and pollution, respectively. The arguments in the utility function respectively equal $\bar{c} = cB$ and $\bar{p} = pB$, and the optimal path can be found by maximizing (7) subject to (8)-(9), using $c(t)$, $p(t)$ and $\chi(t)$ as control variables. As shown in the Appendix, the necessary conditions for optimality yield

$$\dot{c}(t)/c(t) = \left[f_k(t)(1-\chi(t))\left(1-\varepsilon^{-1}\right)\right] - (\rho+\delta+n+\pi), \tag{10}$$

$$1-\chi(t) = \Gamma f(k(t))^{-\frac{\theta-1}{\varepsilon\theta-1}} c(t)^{-\frac{1}{\varepsilon\theta-1}}, \tag{11}$$

where $\Gamma > 0$ is an exogenous constant, and $f_k(t) \equiv f_k(k(t))$. Expression (10) is the growth rate of normalized consumption along the optimal path, which is different from the usual Keynes-Ramsey rule due to the presence of abatement effort, $\chi(t)$, and of the elasticity factor $1-\varepsilon^{-1}$ that quantifies the distortion in the marginal benefit from accumulation induced by welfare-reducing emissions.[6] Condition (11) determines the optimal propensity to spend in abatement, which exhibits a precise link with the time-paths of $k(t)$ and $c(t)$: if consumption and output increase (decrease) along the optimal path, the abatement effort $\chi(t)$ increases (decreases) as well. The dynamic properties of the optimal path can be analyzed as follows.

[6] In general, the Keynes-Ramsey rule asserts that the consumption growth rate is proportional to the difference between the marginal benefit from capital accumulation and the utility discount rate. In the neoclassical model without pollution, the marginal benefit from accumulation equals the marginal product of capital net of depreciation; in the current model, instead, it equals the term in square brackets in (10), which is strictly lower than f_k. The reason is that the optimal path takes into account the fact that higher capital implies *ceteris paribus* higher pollution, which induces a wedge between capital productivity and utility-benefits from accumulation.

Omitting time-arguments for simplicity, define the right hand side of (11) as a function

$$\Phi(k,c) \equiv \Gamma f(k)^{-\frac{\theta-1}{\varepsilon\theta-1}} c^{-\frac{1}{\varepsilon\theta-1}}, \tag{12}$$

where $\Phi_k < 0$ and $\Phi_c < 0$. Substituting $1-\chi = \Phi(k,c)$ in (8) and (10), the resulting dynamic system is

$$\dot{k} = f(k)\Phi(k,c) - c - (\tilde{\rho}-\rho)k, \tag{13}$$

$$\dot{c} = f_k(k)\Phi(k,c)\left(1-\varepsilon^{-1}\right)c - \tilde{\rho}c, \tag{14}$$

where we have defined $\tilde{\rho} \equiv \rho + \delta + n + \pi$. We denote by (c^{ss}, k^{ss}) the couple of values representing the simultaneous steady-state equilibrium of system (13)-(14). It can be shown that the steady state exists and is unique for any well-behaved neoclassical production function[7]. Also the stability properties of the steady state are quite general and do not require assuming specific technologies:

Lemma 1 *The simultaneous steady-state equilibrium (c^{ss}, k^{ss}) of system (13)-(14) is saddle-point stable. Given the initial condition $k(0) = k_0$, the optimal path is unique and implies convergence towards (c^{ss}, k^{ss}).*

Lemma 1 has three main implications. First, convergence towards (c^{ss}, k^{ss}) implies that the propensity to spend in clean technologies and the marginal product of capital are constant in the long run. The asymptotic value of ICT effort, χ^{ss}, is determined by condition (11). Imposing stationarity in (10), we obtain the long-run level of the marginal product of capital

$$f_k^{ss} = \frac{\rho+\delta+n+\pi}{(1-\chi^{ss})(1-\varepsilon^{-1})}. \tag{15}$$

Expression (15) implies that normalized capital k^{ss} will be lower than in the Ramsey model

[7]A proof of the existence and uniqueness is available at the online archive.

- where the modified golden rule $f_k^{ss} = \rho + \delta + n + \pi$ holds.[8] The second implication of Lemma 1 is that the economy displays balanced growth in the long run. Both $c = C/NB$ and $k = K/NB$, as well as ICT effort $\chi = X/Y$, achieve stationary values. Hence, aggregate output, capital, and expenditures grow asymptotically at the same balanced rate

$$\lim_{t\to\infty} \frac{\dot{Y}(t)}{Y(t)} = \lim_{t\to\infty} \frac{\dot{K}(t)}{K(t)} = \lim_{t\to\infty} \frac{\dot{C}(t)}{C(t)} = \lim_{t\to\infty} \frac{\dot{X}(t)}{X(t)} = n + \pi. \tag{16}$$

Result (16) implies that, in the long run, per capita variables grow at the rate of labor-augmenting technical progress, π. The third implication of Lemma 1 is that pollution per capita, $\bar{p}(t) = p(t)B(t)$, converges to a constant steady-state level. From (3) and (1), the dynamics of $\bar{p}(t)$ are governed by

$$\bar{p}(t) = \Omega(t)\,\Phi(k(t), c(t))^{\varepsilon} f(k(t))\,B(t) = \Omega_0 B_0 \Gamma^{\varepsilon} f(k(t))^{\frac{\varepsilon-1}{\varepsilon\theta-1}} c(t)^{-\frac{\varepsilon}{\varepsilon\theta-1}}, \tag{17}$$

which implies

$$\lim_{t\to\infty} \bar{p}(t) = \bar{p}^{ss} \equiv \Omega_0 B_0 \Gamma^{\varepsilon} f(k^{ss})^{\frac{\varepsilon-1}{\varepsilon\theta-1}} (c^{ss})^{-\frac{\varepsilon}{\varepsilon\theta-1}}. \tag{18}$$

The transitional dynamics of pollution per capita can be studied by re-introducing $\bar{p}(t)$ in the dynamic system (13)-(14). In fact, pollution per capita and consumption are two jump variables linked by an optimality relation: using (17), we can transform system (13)-(14) into an equivalent system describing the joint dynamics of $k(t)$ and $\bar{p}(t)$. As the dynamics of $(k(t), c(t))$ are saddle-point stable, we expect the same behavior to arise in the $(k(t), \bar{p}(t))$ plane. This result is formally proved below. In order to obtain an explicit relation between pollution per capita and other endogenous variables of empirical interest, we henceforth assume that the aggregate technology is Cobb-Douglas.

[8] As the planner takes into account the fact that higher capital implies *ceteris paribus* higher pollution, it is optimal to accumulate less capital with respect to the modified golden rule in the long run (see Van der Ploeg and Withagen, 1991).

2.2 Transitional dynamics of pollution per capita

Suppose that technology (1) takes the Cobb-Douglas form $Y = K^{\alpha} (BN)^{1-\alpha}$, and write normalized output $f(k) = k^{\alpha}$, where $\alpha \in (0, 1)$. In this case, the dynamics of pollution per capita and normalized capital are governed by the non-linear system (see Appendix A)

$$g(\bar{p}(t)) = \varphi_0 - \varphi_1 k(t)^{\alpha-1-\frac{\alpha}{\varepsilon}} \, \bar{p}(t)^{-\left(\theta-\frac{1}{\varepsilon}\right)}, \tag{19}$$

$$g(k(t)) = \varphi_2 k(t)^{-(1-\alpha+\frac{\alpha}{\varepsilon})} \, \bar{p}(t)^{\frac{1}{\varepsilon}} \left(1 - \varphi_3 \bar{p}(t)^{-\theta}\right) - \varphi_4, \tag{20}$$

where $g(\bar{p}(t)) \equiv (d\bar{p}(t)/dt)/\bar{p}(t)$ and $g(k(t)) \equiv \dot{k}(t)/k(t)$ are instantaneous growth rates and $(\varphi_0, \varphi_1, \varphi_2, \varphi_3, \varphi_4)$ are exogenous constants, all strictly positive. From (19)-(20), the stationary loci read

$$g(\bar{p}(t)) = 0 \quad \rightarrow \quad k(t) = \left[\frac{\varphi_1}{\varphi_0} \bar{p}(t)^{-\left(\theta-\frac{1}{\varepsilon}\right)}\right]^{\frac{1}{1-\alpha+\frac{\alpha}{\varepsilon}}}, \tag{21}$$

$$g(k(t)) = 0 \quad \rightarrow \quad k(t) = \left[\frac{\varphi_2}{\varphi_4} \bar{p}(t)^{\frac{1}{\varepsilon}} \left(1 - \varphi_3 \bar{p}(t)^{-\theta}\right)\right]^{\frac{1}{1-\alpha+\frac{\alpha}{\varepsilon}}}, \tag{22}$$

where locus (21) is strictly decreasing and (22) is strictly increasing in the $(k(t), \bar{p}(t))$ plane. Equations (19) and (20) respectively imply $\partial g(\bar{p})/\partial \bar{p} > 0$ and $\partial g(k)/\partial k < 0$ and thereby the phase diagram reported in Figure 1. The simultaneous steady-state $(k^{ss}, \bar{p}^{ss})$ is saddle-point stable. In particular, given an initial stock $k(0) = k_0$, the associated initial level of pollution per capita $\bar{p}(0)$ lies along a saddle path that is strictly decreasing in the $(k(t), \bar{p}(t))$ plane. If we want to reproduce the dynamics of an economy exhibiting positive accumulation during the transition, we must assume $k_0 < k_{ss}$. In this case, the strictly-decreasing saddle path implies that the initial level of pollution per capita is above the long-run value, $\bar{p}(0) > \bar{p}^{ss}$. Hence, the transitional dynamics are characterized by a *decreasing time path of pollution per capita*, as shown in Figure 1. These results can be formally proved by linearizing system

(19)-(20); see the Appendix for details. This gives

$$g\left(\bar{p}\left(t\right)\right)\approx m_{1}\left(\bar{p}\left(t\right)-\bar{p}_{ss}\right)+m_{2}\left(k\left(t\right)-k_{ss}\right), \tag{23}$$

$$g\left(k\left(t\right)\right)\approx m_{3}\left(\bar{p}\left(t\right)-\bar{p}_{ss}\right)+m_{4}\left(k\left(t\right)-k_{ss}\right), \tag{24}$$

where the coefficients are $m_1, m_2, m_3 > 0$ and $m_4 < 0$. The Jacobian matrix associated with (23)-(24) confirms saddle-point stability and, in particular, yields the equation of the stable arm

$$\left(k\left(t\right)-k_{ss}\right)=\phi\left(\bar{p}\left(t\right)-\bar{p}_{ss}\right), \qquad \phi<0, \tag{25}$$

where $\phi < 0$ implies a negatively-sloped saddle path. The stable-arm equation allows us to obtain an explicit relation between pollution per capita growth and the other endogenous variables of empirical interest. In fact, substituting (25) in (23), and using (24) to eliminate normalized capital $k(t)$ from the resulting expression, we obtain

$$g\left(\bar{p}\left(t\right)\right)\approx\frac{m_{1}}{\alpha\left(m_{3}+m_{4}\phi\right)}\left(g\left(\bar{y}\left(t\right)\right)-\pi\right)+\phi m_{2}\left(\bar{p}\left(t\right)-\bar{p}_{ss}\right), \tag{26}$$

where $g(\bar{y}(t))$ is the growth rate of output per capita $\bar{y}(t)$. Collecting the constant terms and checking the signs of the exogenous parameters appearing in (26), we obtain

Proposition 2 *Along the optimal path, the instantaneous growth rate of emissions per capita is (i) positively related to the growth rate of output per capita and (ii) negatively related to the level of emissions per capita:*

$$g\left(\bar{p}\left(t\right)\right)\approx H_{0}+H_{1}g\left(\bar{y}\left(t\right)\right)-H_{2}\bar{p}\left(t\right), \tag{27}$$

where $H_1 > 0$ *and* $H_2 > 0$.

Proposition 2 emphasizes the main prediction of our model: along the optimal path, pollution levels obey a precise dynamic relationship. First, the growth rate of emissions per capita is positively related to the growth rate of output per capita: this is a 'scale effect' implied by the positive relation between output and emission levels. Second, the growth rate of emissions per capita is negatively related to the level of emissions per capita: we label this a 'defensive effect' as it reflects the effectiveness of abatement expenditures in limiting pollution growth[9]. The difference with the reduced forms of Solow-type models employed, e.g., in Brock and Taylor (2010) and Bulte et al. (2007), is that (27) incorporates all the optimality conditions governing consumption and investment decisions.

It is worth stressing that result (27) can be interpreted as an equation predicting β-*convergence in pollution*: the defensive effect determines a negative relationship between pollution growth and pollution levels. This notion of β-convergence is however conditional on the existence of the scale effect, the positive relation between pollution growth and output growth. In the remainder of this paper, our aim is to verify whether the existence of both the defensive and the scale effects receives empirical support. In this respect, equation (27) suggests considering a model equation like

$$GP_t = \gamma_0 - \beta \log P_{t-T} + \gamma_1 GY_t, \tag{28}$$

where (GP_t, GY_t, P_{t-T}) represent $(g(\bar{p}(t)), g(\bar{y}(t)), \bar{p}(t))$ in a discrete-time setting with T-year periods growth rates, and testing empirically whether the coefficients β and γ_1 are strictly positive. An extended version of (28) is

$$GP_t = \gamma_0 - \beta \log P_{t-T} + \gamma_1 GY_t + \gamma_2 \log Y_{t-T}, \tag{29}$$

[9] Our definition of defensive effects may include – but is not limited to – the technique effect typically mentioned in the literature, e.g. Brock and Taylor (2010, eq.6), because it captures, in addition to the rate of emission-reducing technical progress, all the endogenous effects that contrast pollution growth over time - e.g., variations in the propensity to invest in clean technologies - after controlling for the effect of output growth.

which includes initial levels of output per capita, Y_{t-T}, as an additional explanatory variable. The reason for considering the extended version (29) is that (28) is directly obtained from a first-order approximation of the saddle path. The deviations arising between the exact non-linear saddle path in Figure 1 and the linearized saddle path (25) are essentially due to the dynamics of capital per efficient labor: in order to capture these high order effects cleaned out by the Taylor expansions (23)-(24), the natural hypothesis is to include the initial level of output per capita as an additional regressor without postulating a priori a definite sign for coefficient γ_2.

3 Empirical methodology

This section proposes an empirical methodology which investigates the existence of scale effects and defensive effects by testing equations (28) and (29). As is common in recent economic growth papers, the predictions of the theoretical model are explored with panel regressions. Our panel estimates are based on four 5-year periods starting in year $t = \{1980, 1985, 1990, 1995, 2000\}$. As pointed out by Barro and Sala-i-Martin (2004), 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).

We consider three regression approaches: parametric, semiparametric and fully non-parametric. All these specifications allow for structural dissimilarities within groupings of countries through a group-specific dichotomous variable. Time dummies are also included to account for potential structural breaks and to capture time-specific effects in the relationship.

Parametric model. The panel regression is given by

$$GP_{i,t} = \alpha_1 + \alpha_2 D_i + \alpha_{3,t} D_t + \beta \log P_{i,t-T} + \gamma_1 GY_{i,t} + \gamma_2 \log Y_{i,t-T} + \varepsilon_{i,t} \tag{30}$$

where $GP_{i,t}$ is the growth rate of emissions per capita 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 EU15 member and equal to 0 if not; D_t are dummy variables for each period t of the panel; $P_{i,t-T}$ is the level of emissions per capita (tons/capita) in the i-th country at time $t-T$; $Y_{i,t-T}$ is the level of GDP per capita (in 1990 International Geary-Khamis dollars) in the i-th country at time $t-T$; $GY_{i,t}$ 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; $\varepsilon_{i,t}$ is an *iid* error term. Model (30) encompasses the two test-equations proposed in the previous section: if we impose $\gamma_2 = 0$, the regression corresponds to a stochastic version of equation (28), while estimating the full regression amounts to testing equation (29). Model (30) can be naturally estimated either with cross-sectional or panel regressions[10]. The latter framework has the advantage of better capturing unobserved heterogeneity and nonlinearities in the relationship.

Given the potential feedback effect of pollution on GDP, the regression's coefficients can suffer from endogeneity bias. We address this issue by providing Instrumental Variables (henceforth, IV) estimates. Following Barro and Sala-i-Martin (2004), we keep the first 5-year period starting in 1980 out of the sample to build instruments. Therefore, IV versions of the panel regressions are proposed with $Y_{i,(t-1)-T}$ as instrument for $Y_{i,t-T}$ and $GY_{i,t-1}$ for $GY_{i,t}$.

Two fundamental hypotheses are tested regarding the OLS specifications. The null of homoscedasticity is checked with a robust version of the Breusch-Pagan LM heteroscedasticity test (Greene, 2008). We also apply the misspecification test of Hsiao et al. (2007) to check if the parametric linear models provide consistent estimates. This test contrasts the following two hypotheses: $H_0 : E(y_i, x_i) = E(y_i, x_i; \varphi)$ *vs.* $H_1 : E(y_i, x_i) \neq E(y_i, x_i; \varphi)$. If

[10]When T is set to the entire length of the time dimension, we can drop the dummies and specification (30) becomes a cross-sectional model where the dynamic component is captured by growth rates over the whole period.

H_0 is not accepted, more flexible specifications can be explored. This paper considers two alternatives to the linear parametric model (30) that we present below.

Semiparametric model. Our second specification is a semiparametric additive model which gives full flexibility to the continuous explanatory components. The dummy variables enter the equation parametrically whereas the other regressors enter nonparametrically with a separable structure. This is the partially linear (PLR) additively separable regression model which can be written as

$$GP_{i,t} = \alpha_1 + \alpha_2 D_i + \alpha_{3,t} D_t + \sum_{j=1}^{3} f_j(x_j^c) + \varepsilon_{i,t} \tag{31}$$

where $f_j(x_j^c)$ are three unknown nonlinear functions, one for each j-th continuous factor in (30), that is, $x_1^c = \log P_{i,t-T}$, $x_2^c = GY_{i,t}$, $x_3^c = \log Y_{i,t-T}$. The first three terms in specification (31) 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 (30), the PLR setting imposes no restriction on the flexibility of the additive nonparametric factors and has a straightforward graphical representation. The additive block is a special case of the general smooth function $f(x_1^c, x_2^c, x_3^c)$, which can be estimated more efficiently than a fully nonparametric setting when it represents the true relationship. Several approaches provide consistent fits of PLR models. Here we employ the procedure of Wood (2006) which decomposes the flexible additive components into a finite sum of spline terms and controls for the smoothness of the functions with cross-validation. The results are reported graphically for each estimated function $\widehat{f}_j(x_j^c)$, with $j = 1, 2, 3$.

Fully nonparametric model. We relax the functional restrictions in the parametric model (30) and the nonparametric additive hypothesis of the PLR equation (31) by estimating a nonparametric regression which allows all kind of interactions between the independent variables (in particular between the continuous regressors and the dummy variables). This fully flexible specification is given by

$$GP_{i,t} = f(\mathbf{x}^d, \mathbf{x}^c) + \epsilon_{i,t} \tag{32}$$

where $\mathbf{x}^d = [D_i, \tilde{D}_t]$ are the usual discrete regressors (but with D_t defined as a single discrete trend factor) and $\mathbf{x}^c = [\log P_{i,t-T}, GY_{i,t}, \log Y_{i,t-T}]$ are the continuous explanatory factors. We estimate (32) using the new kernel method proposed by Racine and Li (2004). This estimator allows us to compute nonparametric regressions with mixed independent variables (i.e., discrete and continuous regressors) and is consistent in panels with a small time dimension t relative to the individual dimension[11] i. We use least squares cross-validation in conjunction with a locally linear kernel estimator to determine the bandwidths: this allows us to correct for potential bias near the support's boundaries and to discriminate between linear and nonlinear regressors[12].

Graphical representations. The relationship between the continuous predictors and the response in non or semiparametric regressions is usually reported graphically. We show the results for specifications (31) and (32) with partial regression plots. In that respect, we follow Maasoumi et al. (2007): we present the nonparametric regression of $GP_{i,t}$ on the continuous regressors $\mathbf{x}^c$ for EU15 countries ($D_i = 1$) by plotting

- $GP_{i,t}$ versus $E(GP_{i,t} \mid D_i = 1, \tilde{D}_t = \bar{t}, P_{i,t-T}, Y^\star_{i,t-T}, GY^\star_{i,t})$,
- $GP_{i,t}$ versus $E(GP_{i,t} \mid D_i = 1, \tilde{D}_t = \bar{t}, P^\star_{i,t-T}, Y_{i,t-T}, GY^\star_{i,t})$ and
- $GP_{i,t}$ versus $E(GP_{i,t} \mid D_i = 1, \tilde{D}_t = \bar{t}, P^\star_{i,t-T}, Y^\star_{i,t-T}, GY_{i,t})$,

where the superscript '$\star$' indicates that the variable is kept at its median level and $\bar{t}$ is a

[11] Given the use of 5-year data, the time dimension is of length four, which is small compared to the 25 countries observed each year.

[12] As mentioned in Li and Racine (2007), 'the traditional local constant kernel estimator may have large bias when estimating a regression function near the boundary of support. The local linear estimator is one of the best known approaches for bias correction'. The authors also emphasize that local linear least squares cross-validation has the ability to select large values for the bandwidth when $g(x)$ is linear in the x regressor.

selected year. The same method is used for non-EU15 countries ($D_i = 0$). Note that for the PLR model (31), the shapes of the nonparametric additive terms are similar for the pooled sample, for each country grouping, and year, up to an additive constant. For the fully nonparametric setting, interactions may yield specific shapes depending on the levels of the discrete factors.

4 Data and Results

4.1 Data

A consistent empirical testing of the optimal pollution-GDP relationships (28) or (29) imposes two basic requirements regarding the pollution data: (i) the negative impact of pollution on welfare needs to be linked to the flow of emissions (and not to the pollution stock), and (ii) regulatory mechanisms must be at work to enforce (potentially optimal) defensive measures. Since the 1980s, the European states have been particularly pro-active in fighting atmospheric pollution, and more specifically two acidifying gases' emissions: sulphur oxides (SOX) and nitrogen oxides (NOX). The Helsinki 1985 and the later Oslo 1994 and Goteborg 1999 Protocols committed about twenty European countries to substantially reduce their sulfur emissions through a variety of mechanisms. NOX emissions experienced similar early control initiatives across Europe through the 1988 Sofia Protocol and the Large Combustion Plant European Directive (2001/80/EC)[13]. Therefore, exploring the presence of the defensive and scale effects for per capita SOX and NOX emissions with a European panel of countries covering the post-1980 period appears as a natural step to test the pollution convergence equation.

Our database is a balanced panel of 25 European (Eastern and Western) countries that covers the 1980-2005 period. We use GDP and population series from Maddison (2008). Data

[13] See Bratberg et al. (2005), Finus and Tjotta (2003) and Murdoch et al. (1997) for a discussion of the effectiveness of these Protocols in mitigating pollution.

for SOX and NOX emissions come from the EMEP-CEIP database WebDab and correspond to series used in the EMEP models[14]. These pollution data are based on officially reported emissions, but inconsistent/missing observations are corrected or gap-filled[15]. Our analysis focuses on SOX and NOX emissions derived from human activities and ignores those occurring in natural environments without human influence. SOX emissions are mainly linked to combustion processes at the industry and plant level. NOX particles are essentially emitted by road transportation, other mobile sources, and electricity generation. Sulfur and nitrogen oxides are well-known to have a large negative impact on human health and natural ecosystems[16]. Through the reaction with other substances, SOX and NOX emissions cause lung diseases; they modify land and water ecosystems and generate acid rains that affect nature as well as buildings, cars and historical monuments.

Figure 2 displays the national series on per capita SOX and NOX emissions as well as GDP per capita. Solid lines represent the historical EU members EU15, while dashed lines indicate the non-EU15 group, i.e., the most recent Eastern EU members and the non-EU members[17]. The left graph shows that all per capita GDP series are upward trended and

[14] The European Monitoring and Evaluation Programme (EMEP) is a protocol signed in 1984 under the Convention on Long-Range Transboundary Air Pollution (CLRTAP) which requires that parties report to the treaty secretariat on several air pollutant emissions. Since January 2008, the EMEP Centre on Emission Inventories and Projections (CEIP) operates the EMEP emission database (WebDab), which records anthropogenic and natural emissions for a large variety of air pollutants (acidifying/eutrophying compounds, ozone precursor, heavy metals and particulate matters).

[15] More details are provided in Mareckova et al. (2008). The reader interested in precise SOX and NOX definitions can refer to UN Economic Commission for Europe (2009, p.17). The definition of SOX comprises all sulphur compounds, expressed as sulphur dioxide (SO2), the major part of which is SO2. Similarly, nitrogen oxides (NOX) include nitric oxide and nitrogen dioxide, expressed as nitrogen dioxide (NO2).

[16] See the US Environmental Protection Agency at http://www.epa.gov/air/urbanair/.

[17] The EU15 and non-EU15 (*) countries are: Albania*, Austria, Belgium, Bulgaria*, Czechoslovia*, Denmark, Finland, France, Former Yugoslavia*, Former USSR*, Germany, Greece, Hungary*, Ireland, Italy, Netherlands, Norway, Poland*, Portugal, Romania*, Spain, Sweden, Switzerland*, Turkey*, United Kingdom. Note that consistent GDP, population and emission series over the whole period 1980-2005 were available for Germany and Czechoslovakia (Czech Republic + Slovakia) despite the changes in borders. Regarding Former USSR and Former Yugoslavia, GDP per capita is computed for both blocks by including the GDP of all the new Republics, while emissions series exclude (over the whole 1980-2005 period) some Republics for which emissions data were missing. More precisely, Bosnia & Herzogowina, Croatia, Macedonia and Slovenia were included in Yugoslavia's emissions but Montenegro & Serbia were excluded. Similarly, Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Latvia, Lithuania, Moldova, Russian Republic and Ukraine were included in the USSR emissions but Kyrgyzstan, Uzbekistan, Tajikistan and Turkmenistan

that all non-EU15 countries but Switzerland have substantially lower per capita GDP levels. There is no clear evidence of a decreasing gap in per capita GDP either within or between these two groups over the 1980-2005 period. This is confirmed by the descriptive statistics reported in Table 1 where the difference in median income between the EU15 and the non-EU15 countries strongly increases over time. By contrast, many of the downward sloping NOX and SOX series in Figure 2 seem to stabilize at some point and converge in the sigma sense across the whole sample. Both the global and the group-specific interquartile range of emissions decrease over time in Table 1. Note however that the median level of NOX emissions per capita remains typically higher in the EU15 while the reverse holds for SOX. These patterns suggest the presence of two distinct groupings of countries that may display different behaviors in the econometric analysis.

4.2 Regression results

Tables 2 and 3 contain the regression results for SOX and NOX respectively[18]. Columns (A) to (D) test the presence of the scale and defensive effects in the parametric specifications with a standard OLS fixed-effects estimator. These regressions also represent β-type convergence equations for pollution, conditional upon the levels and growth rates of per capita GDP. Columns (A) and (B) focus on specification (28) where column (B) is the IV counterpart of (A), which controls for potential endogeneity bias. Columns (C) and (D) display the results for the extended model (29) in the same manner. Column (E) shows the linear part of the semiparametric regression with some diagnostic statistics. Column (F) displays the R^2 of the nonparametric estimates. Graphical devices (Figures 3 and 4 for SOX and NOX respectively) complete the results by displaying linear, partially linear and fully nonparametric partial relationships, for year $\bar{t} = 1985$ and by EU15 status, keeping all other continuous factors at

were dropped.

[18]All the econometric results and figures presented in this paper are obtained with the software `R.2.12.1`.

their respective medians.

Looking at the results for SOX emissions, Table 2 shows that none of the parametric regressions display heteroscedasticity, as the null of homoscedasticity cannot be rejected with the Breusch-Pagan test at conventional significance levels[19]. Therefore, the usual coefficients' standard deviation can be safely used to assess significance. The SOX dynamics appear to be globally unaffected by time-shocks over the period under scrutiny except for the year 1995. The defensive effect in the linear models (A) and (B) is not always significant, while the expected scale effect exists and is highly significant in both the short and the extended parametric specifications. Taking into account the GDP variable in models (C) and (D) improves the explanatory power of the parsimonious specifications (A) and (B), i.e., the adjusted R^2 increases substantially. However, applying the specification test[20] to the OLS fits (last line in Table 2), all parametric models are rejected at the 5% level. Therefore a closer look at the flexible estimates is necessary and we expect them to depict nonlinearities as well as potentially different patterns, specific to the EU15 membership for the fully nonparametric regressions[21]. The PLR estimates in column (E) confirm that a structural shock affected the pollution dynamics in 1995, and that the greater flexibility introduced in the continuous regressors clearly increases the explanatory power of model (D). We also observe that the misspecification test does not reject the PLR regression (E), and that the fully nonparametric model (F) captures 12% of additional total variance with respect to its semiparametric counterpart. We now proceed to evaluate the fits obtained with the flexible models with partial regression plots[22]. Note that the linear estimates are also shown for comparison

[19] We used the routine `bptest` from package `lmtest-0.9-27` to perform the Breusch-Pagan LM test and the function `hccm` from package `car-2.0-9` to compute White-corrected covariance matrices when heteroscedasticity was detected (see NOX regressions).

[20] We used the routine `npcmstest` from package `n.p-0.40-4` to apply Hsiao et al. (2007)'s misspecification test.

[21] The PLR estimates have been computed with the `gam` function from package `mgcv-1.7-2`.

[22] The bandwidths of the nonparametric regressions presented in this paper are available in the supplementary material at the online archive. They are all computed with function `npregbw` from package `np-0.40-4`. The partial regression plots have been partly generated with the help of package `plotrix-3.0-9`.

purposes.

The upper graphs in Figure 3 show that the defensive effect linked to initial pollution levels is confirmed for the EU15 as well as for the non-EU15 countries with both the PLR and the fully nonparametric models. The least-square cross-validation methodology employed to determine the bandwidths does not detect departures from linearity for that partial relationship. The scale effect linked to GDP growth is shown in the middle plots: it is positive as expected, with larger partial elasticities for GDP growth in the EU15 countries. However, the scale effect for the non-EU15 group is hump-shaped with the fully nonparametric fit[23]. The initial GDP variable introduced to capture potential nonlinearities appear to have a negative but linear impact on pollution growth. Overall, despite the rejection of the parametric specifications by the data, the flexible estimates indicate that SOX emissions in Europe display a dynamic income-pollution relation that is consistent with our model's prediction.

The results for NOX depart from the SOX ones in several important aspects. First, as shown in Table 3, the explanatory power of the models is typically larger and all parametric regressions display heteroscedastic errors. Therefore, the coefficients' standard deviation for the parametric fits are White-corrected while those of the linear part of the semiparametric model rely on a Bayesian approach (Wood, 2006). We observe that the coefficients are significant across all models. Taking into account the GDP variable in models (C) and (D) does not significantly improve the explanatory power of the parsimonious specifications (A) and (B) as the R^2s remain very similar. Second, regressions (A) and (B) clearly establish the existence of both the defensive and scale effects, as well as β-convergence for NOX emissions across Europe, with significantly larger growth rates within the EU15 countries. Adding the GDP levels in specifications (C) and (D) does not change these results. Third, the misspecification tests conclude that most linear models are misspecified at the significance level of 10%. The short linearized equation (28) seems to match the data better than

[23]The partial relationships for SOX for alternative levels of the time factor $\bar{t} = \{1990, 1995, 2000\}$ remain robust, see the supplementary material at the online archive.

the expanded linear models (C) and (D): both (A) and (B) regressions would pass the specification test if we required stronger evidence of misspecification, say 5% or 1% cutoffs, to reject the null of correct specification. In order to check whether the rejection of the extended specification is due to a lack of flexibility in the parametric specification, we estimate the IV model (D) within a PLR structure. Column (E) shows that the PLR model is not rejected and confirms that significant time-shocks have affected the NOX dynamics over the observed period.

Fourth, given the uncertainty regarding the most appropriate specification for NOX – model (B) versus model (E) – we proceed in column (F) with a fully flexible approach, i.e., the IV model (D) estimated with a fully nonparametric regression. Before moving to the partial regression plots, note that the latter model explains 88% of the total variance (see column (F) in Table 3). The upper plots in Figure 4 corroborate the existence of a defensive effect for both country groupings with the PLR as well as the nonparametric models. The middle plots confirm the positive effect of GDP growth on pollution growth (again, with larger partial elasticities in the EU15 economies). The effect of the initial GDP on the subsequent pollution growth rates is nonlinear but ambiguous, as the confidence interval includes the zero over large portions of the support for both areas. Also, note that the confidence interval surrounding the non-EU15 fits are globally larger. In sum, the path followed since 1985 by the NOX emissions per capita is fully compatible with the convergence equation predicted by the theoretical model, but with a stronger evidence holding within the EU15 countries[24].

[24] The partial relationships for the fully nonparametric regressions are potentially different for alternative levels of the time factor $\bar{t} = \{1990, 1995, 2000\}$. The scale effect remains clearly positive over time for NOX while the defensive effect tends to become flatter. These results are available in the supplementary material at the online archive.

5 Conclusion

Growth theories are particularly useful to unveil transitional or long-run relationships between pollution, capital accumulation and other central determinants of economic growth (Xepapadeas, 2005). In this paper, we develop a growth model of a representative economy where the interplay between purposeful abatement of pollution, technological progress and diminishing return of capital generates an optimal growth path characterized by a precise dynamic law: the growth rate of emissions per capita is (i) negatively related to the level of emissions per capita and (ii) positively related to the growth rate of output per capita. Result (i) is a 'defensive effect' reflecting the effectiveness of abatement expenditures in limiting pollution growth. Result (ii) is a 'scale effect' implied by the positive relation between output and emission levels. This dynamic law can be interpreted as a β-convergence equation: by virtue of the defensive effect, pollution growth rates regress to zero and emissions per capita are bounded in the long run.

This theoretical prediction is tested for a panel of 25 European countries on per capita SOX and NOX emissions spanning the years 1980 to 2005. Regression estimates based on linear models as well as semi-parametric and fully nonparametric methods support the model predictions, identifying a clear scale effect linked to GDP growth and a negative effect captured through the impact of the past pollution level component.

Acknowledgments

Preliminary results for this paper were presented at the Fourth World Congress of Environmental and Resource Economists 2010 in Montreal (Canada), at SURED 2010 in Ascona (Switzerland) and at the 64th Econometric Society European meeting 2009 in Barcelona (Spain). We are grateful to Arik Levinson, Charles F. Mason and two anonymous Referees for insightful comments that significantly improved the paper. The usual disclaimer applies.

Appendix

Derivation of (10)-(11) Substituting $\bar{c} = cB$ and $\bar{p} = pB$ in the utility function (6), the current-value Hamiltonian associated with the optimal control problem is

$$H(c,p,\chi,k) = \sigma \ln(cB) - \varsigma (pB)^{\theta} + \lambda^{k}\left[f(k)(1-\chi) - c - (\delta+n+\pi)k\right] +$$
$$+\lambda^{p}\left[\Omega(1-\chi)^{\varepsilon} f(k) - p\right],$$

where λ^{k} is the dynamic multiplier for constraint (8) and λ^{p} is the Lagrange multiplier for constraint (9). The necessary conditions for optimality are

$$H_c = 0 \qquad \lambda^{k} = \sigma c^{-1}, \tag{A-1}$$
$$H_p = 0 \qquad \lambda^{p} = -\varsigma\theta p^{\theta-1}B^{\theta}, \tag{A-2}$$
$$H_{\chi} = 0 \qquad \lambda^{k} = -\lambda^{p}\Omega\varepsilon(1-\chi)^{\varepsilon-1}, \tag{A-3}$$

together with the co-state equation

$$H_k = \rho\lambda^{k} - \dot{\lambda}^{k} \rightarrow \qquad \dot{\lambda}^{k}/\lambda^{k} = \rho+\delta+n+\pi - f_k(1-\chi)\left(1-\varepsilon^{-1}\right), \tag{A-4}$$

where we have used $\lambda^{k}/\lambda^{p} = -\Omega\varepsilon(1-\chi)^{\varepsilon-1}$ from (A-3). Time-differentiation of (A-1) yields $\dot{c}/c = -\dot{\lambda}^{k}/\lambda^{k}$, which can be plugged into (A-4) to obtain (10). Combining (A-1), (A-2) and (A-3) to eliminate λ^{k} and λ^{p}, we obtain

$$\Omega\varepsilon(1-\chi)^{\varepsilon-1}\varsigma\theta p^{\theta-1}B^{\theta} = \sigma c^{-1},$$

where we can substitute $p = \Omega(1-\chi)^{\varepsilon} f(k)$ from (9) to obtain

$$(1-\chi)^{\varepsilon\theta-1} = \left[\frac{\sigma}{\varepsilon\varsigma\theta(\Omega B)^{\theta}}\right]\frac{f(k)^{1-\theta}}{c}. \tag{A-5}$$

Because $\Omega(t) B(t) = \Omega_0 B_0$ is constant, the term in square brackets in (A-5) is constant. Solving (A-5) for $(1-\chi)$, and defining $\Gamma \equiv \left[\varepsilon\varsigma\theta\left(\Omega_0 B_0\right)^{\theta}\sigma^{-1}\right]^{-1/(\varepsilon\theta-1)}$, we obtain (11).

Proof of Lemma 1 The existence, uniqueness and saddle-point stability of the steady state (c^{ss}, k^{ss}) are proved in detail in the Appendix available at the online archive. Note that saddle-point stability of (c^{ss}, k^{ss}) is directly connected to saddle-point stability of $(\bar{p}^{ss}, k^{ss})$; the latter result is proved below for the Cobb-Douglas case.

Derivation of system (19)-(20) From (17), we have

$$c = \left(\Omega_0 B_0 \Gamma^{\varepsilon}\right)^{\frac{\varepsilon\theta-1}{\varepsilon}} \bar{p}^{-\frac{\varepsilon\theta-1}{\varepsilon}} f(k)^{1-\frac{1}{\varepsilon}}. \tag{A-6}$$

Plugging (A-6) in (13), the growth rate of $k(t)$ equals

$$g(k) = \left(f(k)/k\right)\Phi(k,c) - \left(\Omega_0 B_0 \Gamma^{\varepsilon}\right)^{\frac{\varepsilon\theta-1}{\varepsilon}} \bar{p}^{-\frac{\varepsilon\theta-1}{\varepsilon}} \left(f(k)/k\right) f(k)^{-\frac{1}{\varepsilon}} - \left(\tilde{\rho}-\rho\right),$$

where we can substitute $\Phi(k,c) = \left(\Omega_0 B_0\right)^{-\frac{1}{\varepsilon}} \bar{p}^{\frac{1}{\varepsilon}} f(k)^{-\frac{1}{\varepsilon}}$ from (17) to obtain

$$g(k) = \left(f(k)/k\right) f(k)^{-\frac{1}{\varepsilon}} \left(\Omega_0 B_0\right)^{-\frac{1}{\varepsilon}} \bar{p}^{\frac{1}{\varepsilon}} \left[1 - \left(\Omega_0^{\theta} B_0^{\theta} \Gamma^{\varepsilon\theta-1}\right) \bar{p}(t)^{-\theta}\right] - \left(\tilde{\rho}-\rho\right). \tag{A-7}$$

When the technology is Cobb-Douglas, $f(k) = k^{\alpha}$, we have $\left(f(k)/k\right) f(k)^{-\frac{1}{\varepsilon}} = k^{\alpha-1-\alpha/\varepsilon}$. Plugging this result in (A-7), and defining the constants $\varphi_2 \equiv \left(\Omega_0 B_0\right)^{-\frac{1}{\varepsilon}} > 0$, $\varphi_3 \equiv \Omega_0^{\theta} B_0^{\theta} \Gamma^{\varepsilon\theta-1} = \sigma/\left(\varsigma\varepsilon\theta\right) > 0$ and $\varphi_4 \equiv \tilde{\rho}-\rho = \delta + n + \pi > 0$, we obtain (20). As regards (19), re-write (14) as

$$g(c) = \alpha\left(f(k)/k\right)\Phi(k,c)\left(\frac{\varepsilon-1}{\varepsilon}\right) - \tilde{\rho}, \tag{A-8}$$

where we have used $f_k = \alpha \left(f \left(k \right) / k \right)$ and $g \left(c \right) \equiv \dot{c}/c$ for the consumption growth rate. Next time-differentiate (17) to get

$$g \left(\bar{p} \right) = \frac{\varepsilon - 1}{\varepsilon \theta - 1} g \left(f \left(k \right) \right) - \frac{\varepsilon}{\varepsilon \theta - 1} g \left(c \right), \tag{A-9}$$

where, given $f \left(k \right) = k^{\alpha}$, the growth rate of normalized output equals $g \left(f \left(k \right) \right) = \alpha g \left(k \right)$. Plugging $g \left(f \left(k \right) \right) = \alpha g \left(k \right)$ and substituting $g \left(k \right)$ with (A-7), and substituting $g \left(c \right)$ by means of (A-8), we obtain

$$g \left(\bar{p} \right) = \frac{\varepsilon \tilde{\rho} - \alpha \left(\tilde{\rho} - \rho \right) \left(\varepsilon - 1 \right)}{\varepsilon \theta - 1} - \frac{\varepsilon - 1}{\varepsilon \theta - 1} \alpha \left(f \left(k \right) / k \right) \Phi \left(k, c \right) \left(\Omega_0^{\theta} B_0^{\theta} \Gamma^{\varepsilon \theta - 1} \right) \bar{p}^{-\theta}. \tag{A-10}$$

Substituting $\Phi \left(k, c \right) = \left(\Omega_0 B_0 \right)^{-\frac{1}{\varepsilon}} \bar{p}^{\frac{1}{\varepsilon}} f \left(k \right)^{-\frac{1}{\varepsilon}}$ from (17), and defining the constants $\varphi_0 \equiv \frac{\varepsilon \tilde{\rho} - \alpha (\tilde{\rho} - \rho)(\varepsilon - 1)}{\varepsilon \theta - 1} = \frac{\varepsilon \tilde{\rho} (1 - \alpha) + \alpha \tilde{\rho} + \alpha \rho (\varepsilon - 1)}{\varepsilon \theta - 1} > 0$ and $\varphi_1 \equiv \alpha \frac{\varepsilon - 1}{\varepsilon \theta - 1} \left(\Omega_0 B_0 \right)^{\theta - \frac{1}{\varepsilon}} \Gamma^{\varepsilon \theta - 1} > 0$, we obtain (19).

Derivation of (23), (24) and (25) From (19), we have $\partial g \left(\bar{p} \right) / \partial \bar{p} > 0$ and $\partial g \left(\bar{p} \right) / \partial k > 0$. From (20), we have , $\partial g \left(k \right) / \partial p > 0$ and $\partial g \left(k \right) / \partial k < 0$. Hence, the coefficient matrix of the linearized system is given by $m_1 \equiv \partial g \left(\bar{p} \right) / \partial \bar{p}|_{\bar{p}^{ss}} > 0$, $m_2 \equiv \partial g \left(\bar{p} \right) / \partial k|_{\bar{p}^{ss}} > 0$, $m_3 \equiv \partial g \left(k \right) / \partial \bar{p}|_{k^{ss}} > 0$, $m_4 \equiv \partial g \left(k \right) / \partial k|_{k^{ss}} < 0$. Given these signs, system (23)-(24) displays two real roots of opposite signs, the stable root being

$$\bar{\mu} \equiv (1/2) \left[\left(m_1 + m_4 \right) - \sqrt{\left(m_1 + m_4 \right)^2 - 4 \left(m_1 m_4 - m_2 m_3 \right)} \right] < 0.$$

The stable arm equation is given by $\frac{k(t) - k_{ss}}{\bar{p}(t) - \bar{p}_{ss}} = \frac{\bar{\mu} - m_1}{m_2}$, where $\bar{\mu} < 0$, $m_1 > 0$, and $m_2 > 0$ imply that the right hand side is a strictly negative constant, $\phi \equiv \frac{\bar{\mu} - m_1}{m_2} < 0$.

Derivation of (26) and proof of Proposition 2 Since output per capita equals $\bar{y} \left(t \right) = B \left(t \right) k \left(t \right)^{\alpha}$, its growth rate is given by $g \left(\bar{y} \left(t \right) \right) = \pi + \alpha g \left(k \left(t \right) \right)$. Plugging (24) in this expression, we have

$$g\left(\bar{y}\left(t\right)\right) = \pi + \alpha\left[m_3\left(\bar{p}\left(t\right) - \bar{p}_{ss}\right) + m_4\left(k\left(t\right) - k_{ss}\right)\right].$$

Eliminating $\left(k\left(t\right) - k_{ss}\right)$ by means of the stable-arm equation (25) and rearranging terms yields

$$\left(\bar{p}\left(t\right) - \bar{p}_{ss}\right) = \frac{g\left(\bar{y}\left(t\right)\right) - \pi}{\alpha\left(m_3 + m_4\phi\right)}.$$

Plugging this expression in (23), and using (25) to eliminate $\left(k\left(t\right) - k_{ss}\right)$, we obtain (26). Defining $H_1 \equiv \frac{m_1}{\alpha(m_3+m_4\phi)}$, $H_2 \equiv -\phi m_2$ and $H_0 \equiv H_2\bar{p}_{ss} - \pi H_1$, we obtain equation (26) in Proposition 2. Since $\alpha > 0$, $m_1 > 0$, $m_2 > 0$, $m_3 > 0$, $m_4 < 0$ and $\phi < 0$, coefficients H_1 and H_2 are both strictly positive, which completes the proof. $\square$

References

Acemoglu, D. (2009) *Introduction to Modern Economic Growth* (Princeton NJ: Princeton University Press)

Aldy, J.E. (2006) 'Per capita carbon dioxide emissions: Convergence or divergence.' *Environmental and Resource Economics* 33(4), 533–555

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

Bovenberg, A.L., and J. Smulders (1995) 'Environmental quality and pollution-augmenting technological change in a two-sector endogenous growth model.' *Journal of Public Economics* 57(3), 353–360

Bratberg, E., S. Tjotta, and T. Oines (2005) 'Do voluntary environmental agreements work?' *Journal of Environmental Economics and Managment* 50(3), 583–597

Breusch, T.S., and A.R. Pagan (1979) 'A simple test for heteroscedasticity and random coefficient variation.' *Econometrica* 47(5), 1287–1294

Brock, W.A., and M.S. Taylor (2005) 'Economic growth and the environment: a review of theory and empirics.' In *Handbook of Economic Growth,* ed. Philippe Aghion and Steven N. Durlauf (Elsevier) pp. 1750–1821. Volume 1B

— (2010) 'The green solow model.' *Journal of Economic Growth* 15(2), 127–153

Bulte, E., J. List, and M.C. Strazicich (2007) 'Regulatory federalism and the distribution of air pollutant emissions.' *Journal of Regional Science* 47(1), 155–178

Evans, P., and G. Karras (1996) 'Do economies converge? evidence from a panel of u.s. states.' *The Review if Economics and Statistics* 78(3), 384–388

Finus, M., and S. Tjotta (2003) 'The oslo protocol on sulfur reduction: the great leap forward?' *Journal of Public Economics* 87(9-10), 2031–2048

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

Greene, W.H. (2008) *Econometric Analysis,* sixth ed. (Prentice Hall International)

Hsiao, C., Q. Li, and J.S. Racine (2007) 'A consistent model specification test with mixed categorical and continuous data.' *Journal of Econometrics* 140(2), 802–826

Keeler, E., M. Spence, and R. Zeckhauser (1971) 'The optimal control of pollution.' *Journal of Economic Theory* 4(1), 19–34

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

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

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

Maasoumi, E., J. Racine, and T. Stengos (2007) 'Growth and convergence: A profile of distribution dynamics and mobility.' *Journal of Econometrics* 136(2), 483–508

Maddison, A. (2008) 'Historical statistics.' Online data, Groningen Growth and Development Center, University of Groningen,, October. http://www.ggdc.net/maddison

Mankiw, N Gregory, David Romer, and David N Weil (1992) 'A contribution to the empirics of economic growth.' *The Quarterly Journal of Economics* 107(2), 407–37

Mareckova, K, R. Wankmueller, M. Anderl, B. Muik, S. Poupa, and M. Wieser (2008) 'Inventory review 2008.' Technical Report, European Environment Agency and Centre on Emission Inventories and Projections

Murdoch, J.C., T. Sandler, and K. Sargent (1997) 'A tale of two collectives: Sulphur versus nitrogen oxides emission reduction in europe.' *Economica* 64(254), 281–301

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

Ordás Criado, C., and J.-M. Grether (forthcoming) 'Convergence in co_2 per capita emissions: a robust distributional approach.' *Resource and Energy Economics*

Quah, D. (1993a) 'Empirical cross-sectional dynamics in economic growth.' *European Economic Review* 37(2-3), 426–434

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

Racine, J., and Q. Li (2004) 'Nonparametric estimation of regression functions with both categorical and continuous data.' *Journal of Econometrics* 119(1), 99–130

Romero-Ávila, D. (2008) 'Convergence in carbon dioxide emissions among industrialised countries revisited.' *Energy Economics* 30(5), 2265–2282

Shioji, E. (1997) 'It's still 2%: Evidence on convergence from 116 years of the us states panel data.' Working paper 236, Universitat Pompeu Fabra

Strazicich, M.C., and J.A. List (2003) 'Are co_2 emissions levels converging among industrial countries?' *Environmental and Resource Economics* 24(3), 263–271

UN Economic Commission for Europe (2009) 'Guidelines for reporting emission data under the convention on long-range transboundary air pollution.' Technical Report: ECE/EB.AIR/97, United Nations - Economic Commission for Europe

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

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

Xepapadeas, A. (2005) 'Economic growth and the environment.' In *Handbook of Environmental Economics,* ed. K.-G. Maler and J.R. Vincent (Elsevier B.V.) pp. 1219–1271. Volume 3

Tables & Figures

Table 1: Descriptive statistics for GDP, SOX and NOX. Period 1985-2005.

	EU25			EU15			non-EU15		
	1985	1995	2005	1985	1995	2005	1985	1995	2005
GDP									
min	2.4	2.1	3.5	8.3	10.3	14.1	2.4	2.1	3.5
max	19.6	21.7	27.4	17.4	21.7	27.4	19.6	20.6	23.2
median	9.7	14.4	19.3	14.8	17.6	22.2	6.3	5.4	7.5
iqr	8.6	12.3	13.9	3.7	3.3	3.7	1.9	2.1	3.0
SOX									
min	11.6	3.9	2.3	17.8	5.8	3.2	11.6	3.9	2.3
max	258.8	178.9	120.8	99.5	51.3	50.3	258.8	178.9	120.8
median	48.8	25.8	13.2	39.5	23.1	8.5	52.6	34.9	22.7
iqr	39.9	28.0	18.9	39.5	23.8	11.0	95.6	42.8	17.3
NOX									
min	8.1	4.6	7.0	9.7	23.9	17.5	8.1	4.6	7.0
max	66.5	50.8	42.9	57.7	50.8	42.9	66.5	37.4	31.3
median	30.8	30.0	22.8	33.2	31.6	27.5	24.4	17.8	17.4
iqr	20.2	15.2	9.5	19.4	9.3	10.0	24.3	12.6	8.8

Notes: The pollution figures are in tons of emissions per capita. GDP is measured in 1000 Geary-Khamis 1990-US dollars per capita. Data for emissions come from the Center on Emission Inventories and Projections. GDP and population figures are taken from Maddison (2008).

Table 2: Regressions results. SOX pollution growth *vs.* initial pollution levels and GDP.

	Parametric models				Non/semipa. models	
	Ordinary LS				PLR fit$^{(a)}$	NP fit$^{(b)}$
Variables	(A)	(B)	(C)	(D)	(E)	(F)
constant	0.013	0.007	0.153***	0.145***	-0.078***	-
d1990	-0.023	-0.020	-0.025	-0.022	-0.020	-
d1995	-0.041**	-0.040**	-0.041**	-0.041**	-0.042**	-
d2000	-0.026	-0.024	-0.020	-0.020	-0.027	-
EU15	-0.032**	-0.030**	0.029	0.029	0.066**	-
$P_{i,t-T}$ (β)	-0.013*	-0.011	-0.021***	-0.019**	-	-
$GY_{i,t}$ (γ_1)	0.802***	-	0.653**	-	-	-
$GY_{i,t-1}$ ($\gamma_{1,IV}$)	-	0.719***	-	0.563**	-	-
$Y_{i,t-T}$ (γ_2)	-	-	-0.063***	-	-	-
$Y_{i,(t-1)-T}$ ($\gamma_{2,IV}$)	-	-	-	-0.062***	-	-
N	100	100	100	100	100	100
R2/R2 adj.	0.17/0.11	0.16/0.11	0.28/0.22	0.27/0.22	0.48/0.40	0.60/-
F-stat	3.12***	3.06***	5.02***	4.94***	-	-
Heterosced. LM-stat.$^{(a)}$	5.08	4.40	8.71	8.12	-	-
P(Correct. Specific.)$^{(b)}$	0.020	0.010	0.002	0.005	0.193	-

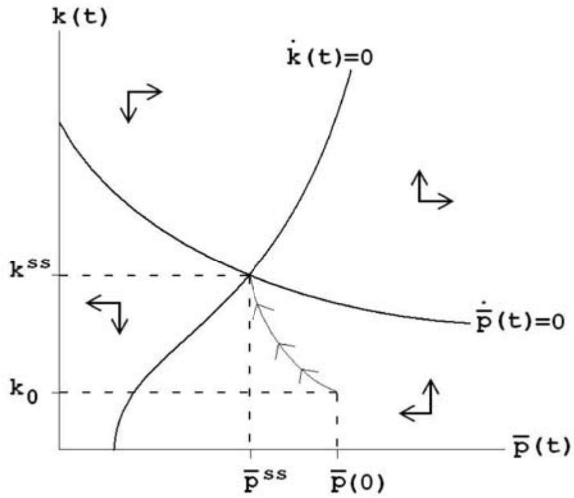
Notes: ***, ** and * denote the 1%, 5% and 10% significance levels. (a) : 'Heterosced. LM-stat.' is the heteroscedasticity LM-statistic of Breusch and Pagan (1979), computed with the variance estimator robust to departure from normality proposed by Koenker (1981). Under the null of homoscedasticity, the statistic is χ^2-distributed, with d.f. = nb. of regressors (constant excluded). (b): 'P(Correct. Specific.)' stands for the probability associated to the nonparametric specification test by Hsiao et al. (2007) for continuous and discrete data models under the null of correct specification. The latter probability is based on 399 *iid* bootstrap's replications.

Table 3: Regression results. NOX pollution growth *vs.* initial pollution levels and GDP.

	Parametric models				Non/semipa. models	
	Ordinary LS				PLR fit	NP fit
Variables	(A)	(B)	(C)	(D)	(E)	(F)
constant	0.125***	0.120***	0.130***	0.126***	-0.028***	-
d1990	-0.026***	-0.025***	-0.026***	-0.024***	-0.021***	-
d1995	-0.039***	-0.039***	-0.037***	-0.038***	-0.034***	-
d2000	-0.039***	-0.039***	-0.035***	-0.035***	-0.035***	-
EU15	0.019**	0.018**	0.028*	0.029***	0.062***	-
$P_{i,t-T}$ (β)	-0.041***	-0.039***	-0.036***	-0.033***	-	-
$GY_{i,t}$ (γ_1)	0.667***	-	0.642***	-	-	-
$GY_{i,t-1}$ ($\gamma_{1,IV}$)	-	0.613***	-	0.588***	-	-
$Y_{i,t-T}$ (γ_2)	-	-	-0.010*	-	-	-
$Y_{i,(t-1)-T}$ ($\gamma_{2,IV}$)	-	-	-	-0.015**	-	-
N	100	100	100	100	100	100
R2/R2 adj.	0.54/0.51	0.54/0.51	0.55/0.52	0.55/0.52	0.71/0.66	0.88/-
F-stat.	18.3***	18.2***	16.7***	16.4***	-	-
Heterosced. LM-stat.$^{(a)}$	16.4**	17.1***	17.6**	18.8***	-	-
P(Correct. Specific.)$^{(b)}$	0.073	0.155	0.000	0.003	0.882	-

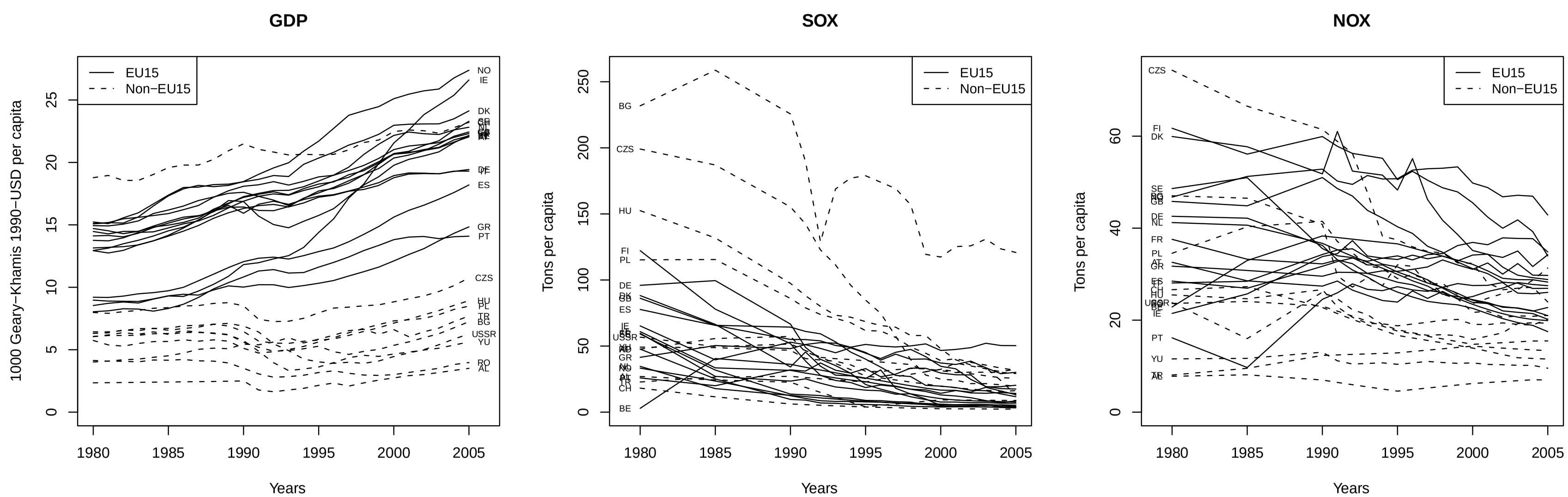
Notes: ***, ** and * denote the 1%, 5% and 10% significance levels. (a) : 'Heterosced. LM-stat.' is the heteroscedasticity LM-statistic of Breusch and Pagan (1979), computed with the variance estimator robust to departure from normality proposed by Koenker (1981). Under the null of homoscedasticity, the statistic is χ^2-distributed, with d.f. = nb. of regressors (constant excluded). (b): 'P(Correct. Specific.)' stands for the probability associated to the nonparametric specification test by Hsiao et al. (2007) for continuous and discrete data models under the null of correct specification. The latter probability is based on 399 *iid* bootstrap's replications.

Figure 1: Phase diagram of capital vs pollution.



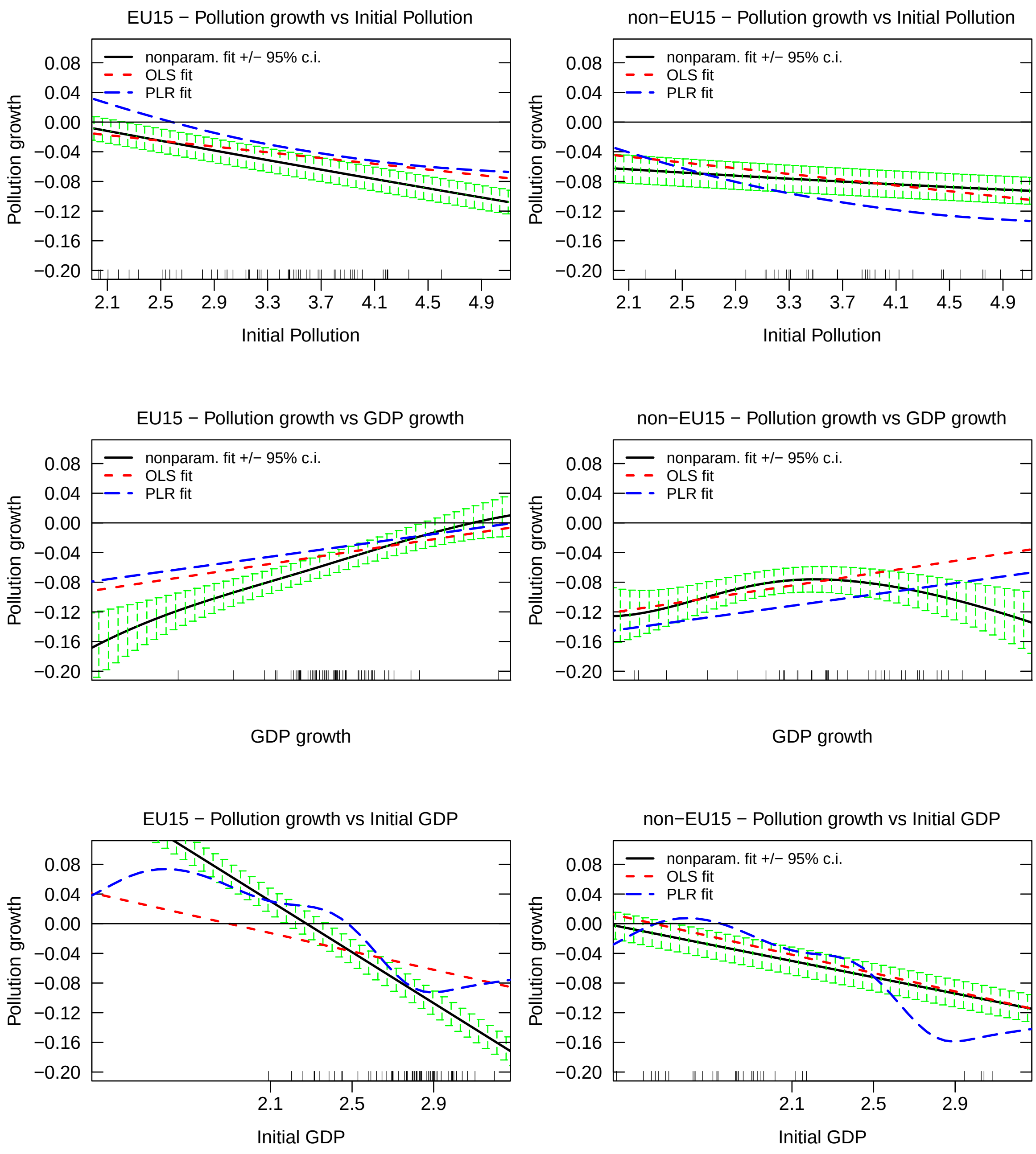
Notes: Phase diagram of system (19)-(20): saddle-point stability of the joint dynamics of capital per efficient labor, $k(t)$, and pollution per capita, $\bar{p}(t)$. The optimal trajectory of capital per efficient labor and pollution per capita with initial condition $k(0) = k_0 < k^{ss}$ is given by the arrowed line.

Figure 2: Per capita levels of GDP, SOX and NOX. European countries, national trends 1980-2005.



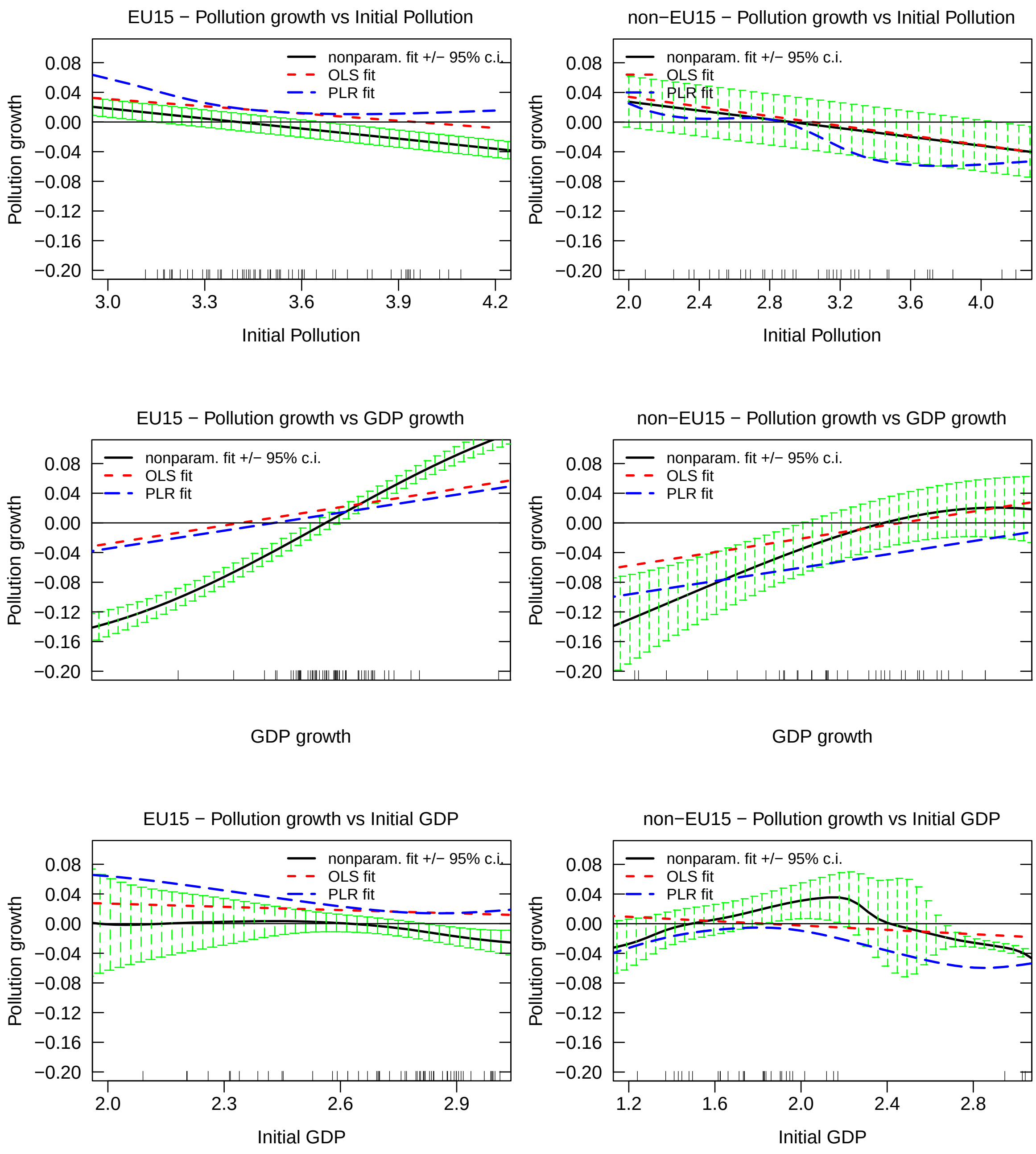
Notes: Yearly GDP and population figures come from Maddison (2008). Emissions data are taken from the Center on Emission Inventories and Projections and are available for the specific years 1980, 1985 and 1990 and on a yearly basis later on.

Figure 3: Nonparametric partial regressions by EU15 status : SOX pollution growth *vs.* initial pollution levels and GDP.



Notes: Nonparametric regressions based on Racine and Li (2004).

Figure 4: Nonparametric partial regressions by EU15 status : NOX pollution growth *vs.* initial pollution levels and GDP.



Notes: Nonparametric regressions based on Racine and Li (2004).