Convergence in CO2 emissions: A Spatial Economic Analysis

Vicente Rios and Lisa Gianmoena

Preliminary Draft, October 2016 (Do not cite)

Abstract

This paper analyzes the evolution of CO2 emissions per capita in a sample of 141 countries during the period 1970-2014. The study extends the neoclassical Green Solow Model to take into account technological externalities in the analysis of CO2 emissions per capita growth rates. Spatial externalities are used to model technological interdependence, which ultimately implies that the CO2 emissions rate of a particular country is affected not only by its own degree of emissions but also by the pollution generated by the remaining countries. In order to investigate the empirical validity of this result, convergence in CO2 emissions is examined by means of dynamic spatial panel econometric techniques. Estimates show the existence of a negative and statistically significant relationship between initial levels of CO2 emissions and subsequent growth rates. This finding is partly due to the role played by spatial spillovers induced by neighboring economies. The observed link is robust to the inclusion in the analysis of different explanatory variables that may affect CO2 emissions growth rates. In a second step, combining recently developed spatial-non parametric techniques with spatial bayesian model selection techniques we identify three distinct clubs in the distribution of CO2 emissions per capita. The estimation of the corresponding three-endogenous regime dynamic spatial model with parameter heterogeneity reveals that in the context of CO2 emissions per capita, the hypothesis of the spatial convergence clubs is more consistent with the data than that of conditional convergence.

Keywords: CO2, Emissions, Convergence, Spatial Panels, Spatial Clubs.

1 Introduction

The relationship between economic growth and the environment has always been controversial. On one side, optimistic researchers tend to highlight the progress made in urban sanitation, improved living standards and resource use efficiency resulting from technological change while others consider that economic growth leads to the emergence of pollution problems which may have a destabilizing effect on the climate. As a matter of fact, the limited natural resource base of the planet, viewed as the key source of limits to growth, has promoted a long and heated debate among economists and environmentalists (**?**; **?**; **?**). However, to a certain extent, there is now less concern over the exhaustion of resources such as oil or uranium and far more concern on the nature's limited ability to act as a sink for human wastes.

Indeed, in recent years, the collective awareness about air pollution caused by CO2 emissions, global warming and climate change have increased considerably. [1] As explained by **?**, if environment's ability to reduce and dissipate wastes is exceeded, environmental quality may fall and policy responses to this reduction consisting in more intensive clean up or abatement efforts could lower the return to investment. Others, focusing on the role of irreversible damage, have claimed that growth may be limited when the ecosystem deteriorates and settles on a newer lower and less productive steady state (**?**; Dechert, 2001).

This has been reflected in the literature analyzing the relationship between per capita income and pollution. This strand of analysis has focused in the so-called environmental Kuznets curve (EKC), which points to the existence of an inverted U-shaped pollution-income relation-ship (**?**). That is, in underdeveloped economies, pollutant emissions per capita tend to grow but once a threshold of income is reached they decrease leading to an improved environmental quality (**?**; **?**). A closely related strand of economic analysis linking growth and environment, which builds upon macroeconomic growth models, is that of environmental convergence (**?**; **?**). Importantly, this modeling approach, exploiting the typical convergence properties of the neoclassical model together with a natural regeneration function yield both (i) an EKC and (ii) a prediction of absolute/conditional environmental convergence.

Empirical studies are crucial in this regard, given that they provide a deeper

[1] Concerns on the effects of CO2 emissions and other greenhouse gases have led the United Nations held numerous conferences and summits aimed at signing international treaties to control emissions, most notably Kyoto-1997 and Paris-2015. The reason is that the emission of carbon dioxide (CO2) into the atmosphere as a result of human economic activities (IPCC, 2007, 2013) has been proved to have effects on climate.

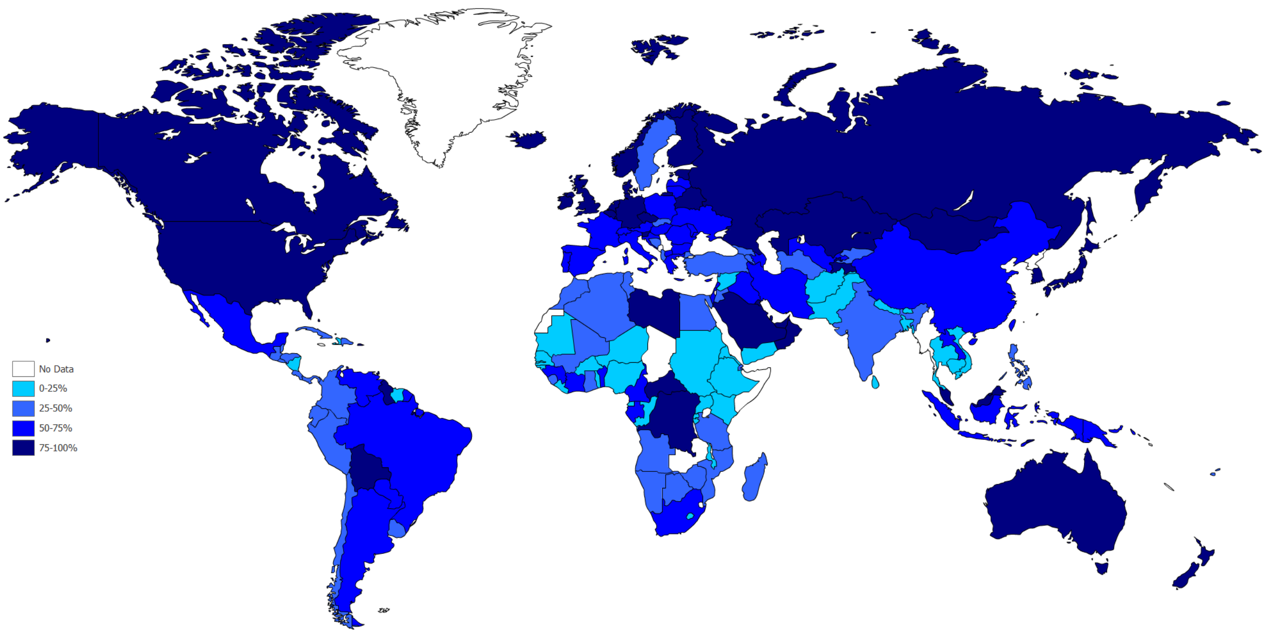
understanding of the phenomenon of CO2 emissions by confronting the plausibility of the theory and the explanatory power of the variables involved in it. The results emerging from the studies of the EKC for CO2 are mixed as there are studies finding an inverted U shape (Carson *et al.*, 1997; Carson and Lundstrom, 2001) and studies finding a monotonic relationship (Cole *et al.*, 1997; Heil and Selden, 2001). Thus, the issue of whether or not an EKC for CO2 exists is far from settled given that the results tend to be sensitive to (i) the sample units and the period considered and to (ii) the econometric methodology employed (see **?** or **?** for a more detailed review). On the other hand, the observation of convergence/divergence in the evolution of CO2 emissions across countries is not conclusive as empirical studies employing parametric and non parametric approaches virtually fit all possibilities. Using non-parametric econometric analysis Ezcurra (2007) finds a slow process of convergence while Aldy (2006) find global divergence. Similarly, while Nguyen Van (2005) using a panel data model finds no evidence of convergence Brook and Taylor (2010) using a cross section find evidence supporting conditional convergence.

As **?** points out, most of the EKC research focuses on time-series issues such as stationarity, co-integration, etc. Nevertheless, an important point that has been over-looked by most of the EKC literature is the fact that CO2 emissions are not only correlated in time but also in space. Likewise, both parametric and non-parametric empirical analysis focusing on the issue of convergence did not take into account the existence of spatial dependence. The omission of relevant spatial interaction terms in econometric analysis is of major importance as it could lead to bias/inconsistent and inefficient estimates **?**. From the theoretical point of view, spatial interactions in CO2 emissions among economies may arise as a consequence of countries strategic response to transboundary pollution flows as governments might strategically manipulate environmental standards in an attempt to attract capital, or for trade purposes. This, in turn, might result in countries mimicking each others' environmental policies which ultimately may lead to similar environmental quality along the spatial dimension. Another argument to consider spatial interactions in the analysis of CO2 emissions, which has been high-lightened by spatial growth models is that traditional growth models omitting technological interdependence might be seriously miss-specified (**?**; **?**; **?**; **?**).

Importantly, these observations regarding the relevance of space in the distribution of CO2 emissions can be corroborated when looking at Figures (1) and (2). Figure (1) provides a first insight on the role of space in the distribution of average CO2 emissions around the globe during 1990-2010. Direct observation of Figure (1) clearly suggests there is a geographical component behind the evolution of the distribution of

CO2 emissions. As a further check on the role played by spatial location of the various countries in explaining CO2 emissions, Figure (2) displays the estimated spatially conditioned stochastic kernel of relative CO2 emissions per capita following ?.[2] The results of the stochastic kernel in Figure (2) reveal that the probability mass tends to be located parallel to the axis corresponding to the original distribution. Accordingly, spatial effects are a relevant factor explaining the observed variability in CO2 emissions.

Figure 1: Spatial Distribution of CO2 Emissions per capita

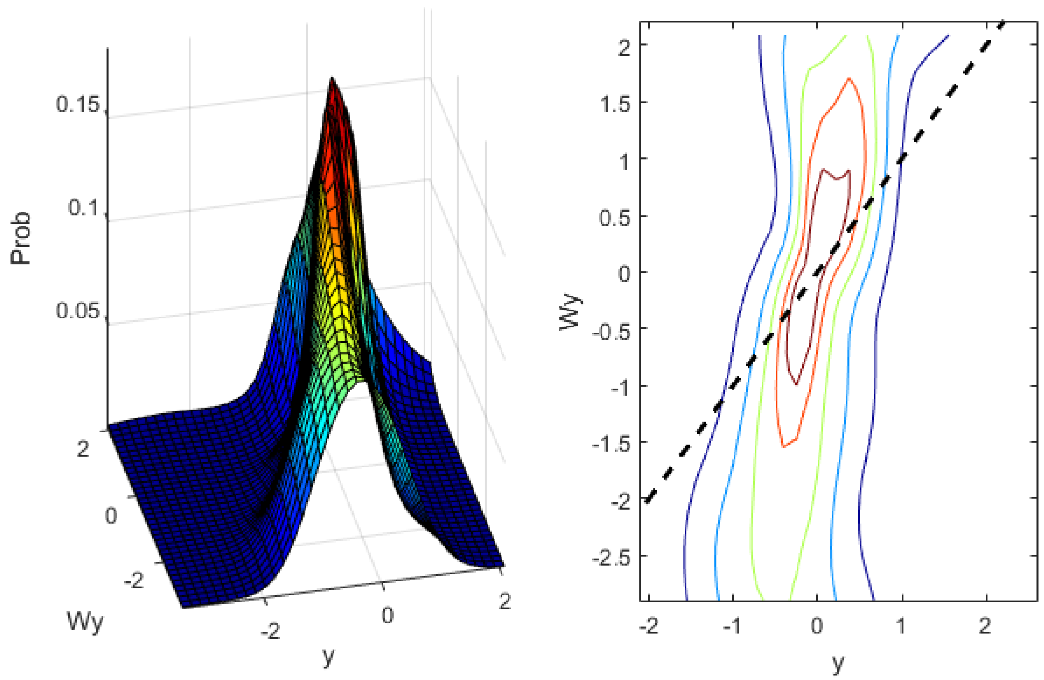


To extend our understanding of the patterns of CO2 emissions the paper makes several novel contributions to the literature.

First, following recent developments in spatial economics we expand the Green-Solow model in order to account for spatial interactions. To that end, a spatially augmented Green-Solow model with technological interdependence among economies is developed. Spatial externalities are used to model technological interdependence, which ultimately implies that the economic growth rate and the CO2 emissions of a particular country is affected not only by its own factors but also by those of neighbouring economies. Using numerical techniques we analyze the effects of different structural parameter changes.

[2]The estimation of the stochastic kernel relies in Gaussian kernel smoothing functions developed by ? and it is performed by employing the L-stage Direct Plug-In estimator with an adaptative bandwith that scales pilot estimates of the joint distribution by $\alpha = 0.5$, as suggested by ?.

Figure 2: Conditional Stochastic Kernel of CO2 Emissions per capita



Second, starting from the theoretical model, a Dynamic Spatial Durbin Model specification for CO2 emissions is derived and employed in the econometric exercise using annual data for the period 1990-2010 for a sample of 123 countries. This general spatial panel specification including country-fixed and time-period fixed effects is estimated by means of the Bias-Corrected-Maximum-Likelihood (BCQML) developed of **?** for dynamic spatial panels allowing us to test the different convergence hypothesis for CO2 emissions. In this regard, a variety of econometric tests regarding spatial co-integration, parameter identification and model selection which are relevant to perform inference in the context of dynamic spatial panels are carried out. The model selection in this context is particularly important as different models ultimately imply different spillover processes **?**. Therefore, instead of assuming a specific spatial specification the present study carries out a Bayesian comparison procedure to dynamic spatial panel models which helps to analyze jointly the probability of the different spatial models and the spatial interaction matrices.

Finally, we augment the base-line convergence spatial panel data model by including linear and quadratic GDP per capita terms in order to test the prediction of the EKC in CO2 emissions which represents, as such, a novel application in the field of environmental economics.

The paper is organized as follows. After this introduction, Section 2 presents a theoretical growth model to investigate the effect of spatial interactions on the path of CO2 emissions and derives the empirical specification. Section 3 describes the data and the econometric approach used in the analysis. The empirical findings of the paper are discussed in Section 4. The final section offers the main conclusions from this work and the policy implications of the research.

2 The Spatial Green Solow Model

This section develops a spatially augmented Green-Solow model which builds upon previous work of**?**. In this model economy, technological progress in the production of goods and technological progress in abatement are exogenous. The key distinct feature of the model with respect **?** is that includes technological externalities in the production of goods, which implies interdependence among the n countries denoted by $i = 1, \ldots, n$. These economies have the same production possibilities but they differ because of different savings rates, population growth rates, depreciation rates and spatial locations.

2.1 The Model

Consider the labor-augmenting Cobb-Douglas production function:

$$Q_{it} = K_{it}^{\alpha} \left(B_{it} L_{it}\right)^{1-\alpha}, 0 < \alpha < 1 \tag{1}$$

where Q is the level of output, K is the level of capital, L is the level of labor, B is the level of technology and the subscript i and t denote the value of the above variables for country i at period t. We further assume exogenous population growth and exogenous technological progress in abatement such that:

$$L_{it} = L_{i0} e^{pt} \rightarrow \frac{\dot{L}_{it}}{L_{it}} = p \tag{2}$$

$$\Omega_{it} = \Omega_{i0} e^{-g_a t} \rightarrow \frac{\dot{\Omega}_{it}}{\Omega_{it}} = -g_a \tag{3}$$

where p is the population growth and $g_a > 0$ is the technological progress in abatement. We introduce spatial correlation across economies by means of technological spillovers following **?**. Hence, technological advances in one country are allowed to have spillover effects on other economies. We specify the level of technology in the production of goods as:

$$B_{it} = B_{i0} e^{g_b t} \prod_{j \neq i}^{N} B_{jt}^{\lambda w_{ij}} \tag{4}$$

The technology level in economy i at period t, B_{it}, is determined not only by its own initial level B_{i0} but also by its neighbors B_{jt} which may spill over to economy i. The magnitude of the spillover effect is measured by λ and w_{ij} specifies the connectivity structures on whether and how much the technology is transmitted from j to i. We assume $W_n = \frac{w_{ij}}{\sum_{j \neq i}^{N} w_{ij}}$ so that all weights are between 0 and 1. Additionally we assume zero diagonal elements to exclude self-influence. Rewriting previous expression in log form and stacking over i we get:

$$\ln \mathbf{B}_t = \ln \mathbf{B}_0 + g_b t \iota_n + \lambda W_n \ln \mathbf{B}_t = [I_n - \lambda W_n]^{-1} \ln \mathbf{B}_0 + \frac{g_b t}{1-\lambda} \iota_n \tag{5}$$

where ι_n is an $N \times 1$ vector of ones and because of W_n is row-normalized $[I_n - \lambda W_n]^{-1} \iota_n = \frac{1}{1-\lambda}$. Therefore, the growth rate of technology in country i is given by $\frac{\dot{B}_{it}}{B_{it}} = \frac{g_b}{1-\lambda}$ which is greater than g_b due to the spillover effect if $0 < \lambda < 1$. Capital accumulates via investments and depreciates at rate δ such that:

$$\dot{K}_{it} = I_{it} - \delta K_{it} = s_i Q_{it} - \delta K_{it} \tag{6}$$

To model the effect of pollution we assume that every unit of economic output Q_{it} generates Ω_{it} units of pollution at every point in time if this pollution is unabated. However, the amount of pollution released to the atmosphere will differ from the amount produced if there is abatement. In this framework, each economy devotes a constant (and exogenous) fraction of output to abate pollution, $0 \leq \theta \leq 1$, where $\theta = Q^A/Q$. After abatement, a unit of output produces $a(\theta)\,\Omega_{it}$ units of pollution in period t. We further assume the abatement function $a(\theta)$ satisfies the following properties: (i) $a(0) = 1$, (ii) $a^{'}(\theta) < 0$ and (iii) $a^{''}(\theta) > 0$ which implies that abatement has a positive but diminishing marginal impact on pollution reduction. To combine our assumptions on pollution and abatement we follow **?** and specify output available for consumption or investment as $Y_{it} = (1-\theta) Q_{it}$. Therefore, pollution is defined as:[3]

$$E_{it} = Q_{it}\Omega_{it}a(\theta) \tag{7}$$

Equation (**??**) requires a brief comment. First, note that aggregate pollution emissions are determined by the scale of economic activity Q_{it} and by the techniques of production $\Omega_{it}a(\theta)$. The second point to high-light is that it is the production of output (and not the use of inputs) what determines pollution. Given that there is only one good, "composition effects", understood such as those that occur when the economy specializes in relatively less pollution intensive services or relatively less natural intensive industries, are zero.

Therefore, the main departures from the standard Solow model are: (i) the fact that pollution is co-produced with every unit of output, (ii) the assumption of some fraction of output devoted to abatement and (iii) the existence of technological interdependence in the production of goods. However, none of these assumptions fundamentally alters the dynamics of the standard Solow model. Note that, indeed, in the present framework, pollution does not feedback into the growth rate of output and that abatement affects the level of output but not its long run growth rate. The model can be solved like the standard Solow model by transforming our measures of disposable output, capital and pollution into effective units ($y_{it} = Y_{it}/B_{it}L_{it}$, $k_{it} = K_{it}/B_{it}L_{it}$, $e_{it} = E_{it}/B_{it}L_{it}$):

$$y_{it} = (1-\theta) f(k_{it}) \tag{8}$$

$$\dot{k}_{it} = s_i (1-\theta) f(k_{it}) - \left(\delta + p + \frac{g_b}{1-\lambda}\right) k_{it} \tag{9}$$

$$e_{it} = a(\theta)\Omega_{it} f(k_{it}) \tag{10}$$

As in the Solow model, starting for any $k_{i0} > 0$, the economy converges to a unique steady state capital per effective worker level k_i^* and a steady state income per

[3]Alternatively, we can write emissions at any time t as: $E_{it} = B_{i0}L_{i0}\Omega_{i0}a(\theta) e[g_E t] k^{\alpha}$ where B_{i0}, L_{i0}, and Ω_{i0} are the initial conditions.

effective worker level y_i^* which are given by Equations (**??**) and (**??**) below:

$$k_i^* = \left[\frac{s_i (1-\theta)}{\frac{g_b}{1-\lambda} + p + \delta} \right]^{\frac{1}{1-\alpha}} \tag{11}$$

$$y_i^* = \left[\frac{s_i (1-\theta)}{\frac{g_b}{1-\lambda} + p + \delta} \right]^{\frac{\alpha}{1-\alpha}} \tag{12}$$

Note that k_i^* and y_i^* will be the same for all economies if θ, λ, s, p and δ are assumed to be the same for all i. Importantly, when $\lambda = 0$ so that there are no spillover effects, the steady state level of capital and income will be the same as **?**. With a positive λ so that $0 < \lambda < 1$, the spillover effect will increase the overall growth rate of technology and hence will decrease the steady state value of k_i^* because the effective labor $B_{it}L_{it}$ increases. On the balanced growth path, aggregate GDP, consumption and capital all grow at rate $g_Y = g_K = g_C = \frac{g_b}{1-\lambda} + p$ while their corresponding per capita magnitudes grow at rate $g_y = g_k = g_c = \frac{g_b}{1-\lambda} > 0$. Finally, since k_{it} approaches the constant k_i^* we can infer from Equation (**??**) the aggregate level of pollution emissions grows at rate:

$$g_E = \frac{g_b}{1-\lambda} + p - g_A \tag{13}$$

which may be positive, negative or zero. Note that in the case where $g_b > 0$ and $g_A > \frac{g_b}{1-\lambda} + n$, the economy will display a sustainable growth path. Equation (**??**) clearly shows how technological progress in goods production has a very different environmental impact than does technological progress in abatement. Technological progress in goods production creates a "scale effect" that raises emissions which is captured in the first two terms in Equation (**??**) since aggregate output grows at rate $\frac{g_b}{1-\lambda} + p$ along the balanced growth path. Thus, technological progress in goods production is necessary to generate per capita income growth. On the other hand, technological progress in abatement creates a "pure technique effect" driving emissions downwards. Therefore, technological progress in abatement must exceed growth in aggregate output in order for pollution to fall and improve environmental quality.

Off the balanced growth path, the growth rate of economy and emissions depends on the level of capital stock. In particular we have that:

$$\frac{\dot{k}_{it}}{k_{it}} = s_i (1-\theta) k_{it}^{\alpha-1} - \left(\delta + p + \frac{g_b}{1-\lambda} \right) \tag{14}$$

$$\frac{\dot{E}_{it}}{E_{it}} = g_e + \alpha\frac{\dot{k}_{it}}{k_{it}} = \left(\frac{g_b}{1-\lambda} + p + \alpha\frac{\dot{k}_{it}}{k_{it}}\right) - g_a \tag{15}$$

Equation (**??**) implies that if the economy starts with a capital stock smaller than the steady state level of capital given k_i^* in Equation (**??**) such that $0 < k_{0,i} < k_i^*$, the economy will accumulate capital $\frac{\dot{k}_{it}}{k_{it}} > 0$ until it reaches the steady state ($\lim_{t\to\infty} k_{it} = k_i^*$) where it stops the accumulation $\left(\lim_{t\to\infty}\frac{\dot{k}_{it}}{k_{it}} = 0\right)$. If we assume there exists a sustainable balanced growth path, such that in the long run $g_{E,i} < 0$, then, with low enough initial level of capital, there exists a point in time t^* such that for $t < t^*$ $g_E > 0$ (i.e, total emissions rise because abatement is not enough to outweigh extra pollution caused by faster growth of GDP), for $t = t^*$ $g_E = 0$ (i.e, emissions are exactly offset by the rate at which they are abated) and for for $t > t^*$, $g_E < 0$ (i.e, improvements in emission intensity Ω outweigh additional production created by production), which ultimately implies an *Environmental Kuznets Curve* profile, with peak at time t^*. [4] The capital stock at this turning point (T) is given by $k(iT)$:

$$k(iT) = \left[\frac{s_i\,(1-\theta)}{p + \frac{g_b}{1-\lambda} + \delta - \frac{g_E}{\alpha}}\right]^{\frac{1}{1-\alpha}} \tag{16}$$

T is defined by:

$$T : k(iT) = \left[(k_i^*)^{1-\alpha}\left(1 - e^{-\phi t}\right) + (k_{i,0})^{1-\alpha}\, e^{-\phi t}\right] \tag{17}$$

and re-arranging we find that:

$$T = \frac{1}{\phi} ln\left[\frac{(k_i^*)^{1-\alpha} - (k_{i0})^{1-\alpha}}{(k_i^*)^{1-\alpha} - (k_i)^{T(1-\alpha)}}\right] \tag{18}$$

Thus, the calendar time to reach the peak of emissions declines with the speed of convergence of each economy: $\phi = (1-\alpha)\left(p + \frac{g_b}{1-\lambda} - \delta\right)$. In order to investigate the effect of introducing spatial interactions in the Green-Solow model we now conduct a steady state analysis (see Figure (**??**)). In the numerical example the simulations are carried using initial conditions $B_{i0} = 1$, $L_{i0} = 1$, and $\Omega_{i0} = 1$, population growth $p = 0.01$, capital share $\alpha = 0.4$, savings $s_i = 0.25$, capital depreciation $\delta = 0.025$, a

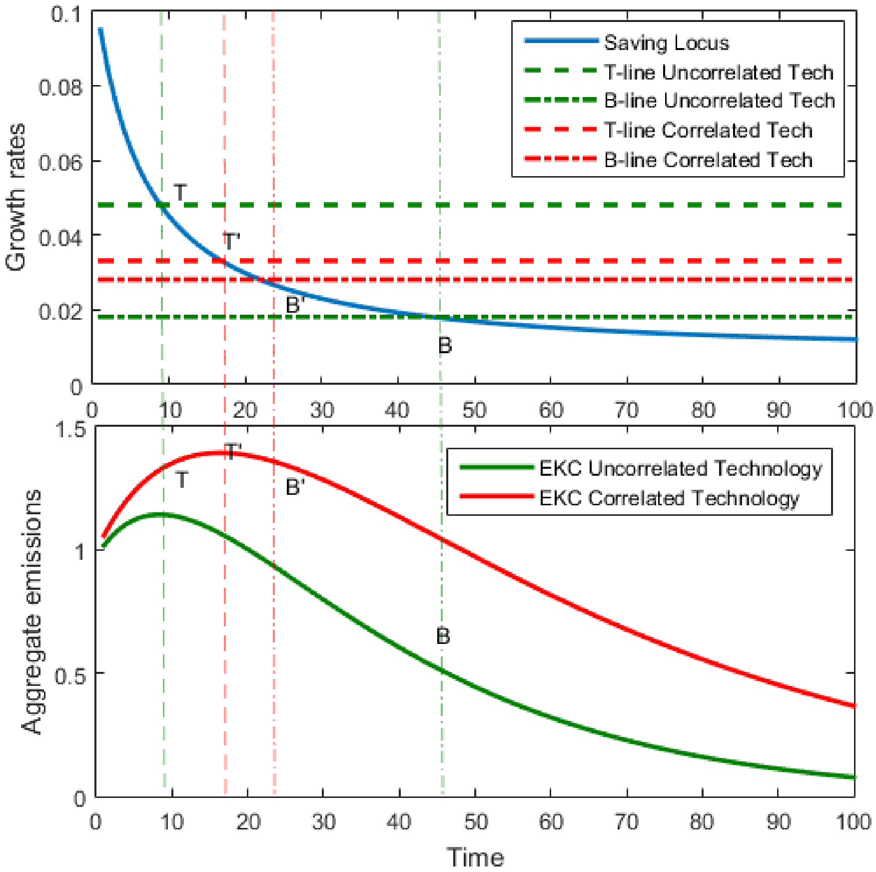
[4] The emergence of a EKC follows primarily from the mechanics of convergence coupled with the dynamics by a standard regeneration function.

fraction of output for abatement $\theta = 0.05$, technological growth in abatement $g_a = 0.05$, technological growth rate in goods production $g_b = 0.01$ and spatial interaction parameter $\lambda = 0.75$. The introduction of a spatially correlated technology to produce goods shifts the growth rate of technological progress from g_b to $\frac{g_b}{1-\lambda}$ which generates effects that are not clear cut and deserve some comments. First, spatial interactions shifts the the T-line $\alpha\left(p + g_b + \delta\right) - g_{E,i}$ downward (equilibrium change from T to T') since $\alpha\frac{g_b}{1-\lambda} - \frac{g_b}{1-\lambda} < 0$, which means it increases the effective capital per worker in the turning point $k(iT)$. This stems from the fact that $(\alpha - 1)\, g_b$ is higher than $(\alpha - 1)\,\frac{g_b}{1-\lambda}$ in Equation (**??**). Similarly, by Equation (**??**), increased spatial interactions raise the growth rate of emissions per capita at any k_{it} via a "scale effect". On the other hand, both the growth rate of capital per worker falls and the steady state value of capital per effective worker decreases (equilibrium change from B to B'). Although in the numerical simulation presented increased technological interactions reduces T, the effect on the calendar to achieve the peak of emissions is indeterminate as T could rise or fall due to the fact that differencing T with respect to technology yields a complex expression depending on a number of parameters. [5]

In order to investigate how the other structural parameters of the model affect the dynamics of pollutant emissions, Figure (**??**) displays a comparative steady state analysis when changing (i) the initial conditions from $B_{i0} = 1$, $L_{i0} = 1$, and $\Omega_{i0} = 1$ to $B_{i0} = 0.75$, $L_{i0} = 0.75$, and $\Omega_{i0} = 0.75$, (ii) the savings rate, from $s_i = 0.25$ to $s_i = 0.4$, (iii) the intensity in abatement from $\theta = 0.05$ to $\theta = 0.3$ and (iv) the growth rate of technological progress at abatement, from $g_a = 0.05$ to $g_a = 0.1$. As can be seen in Panel (a) lower initial conditions in B_{i0}, L_{i0}, Ω_{i0} and k_{i0} have a direct effect on E_{it} but have no impact on the steady state magnitudes of k_i^* nor on long run growth rates as the T-line and the B-line are not affected. Panel (b) shows the effect of increasing the savings rate s_i. This change accelerates the process of capital accumulation, increases the steady state values of k_i^* and the magnitudes $k(iT)$ needed for the turning point of the EKC. However, the steady state growth rate of emissions and income per capita remain unchanged (neither the B-line nor the T-line are shifted). Panel (c) shows the effect of increasing abatement intensity θ due to a tighter environmental policy. This type of policy slows down capital accumulation via smaller investment I_{it} which decreases the magnitudes of $k(iT)$ needed for the turning point of the EKC and leaves the steady state growth rate of emissions and income per capita unchanged (B-line and T-line does not move). Although increasing θ has an impact on the pollution path this type of policy does not affect g_a which implies that in this setting, a tighter

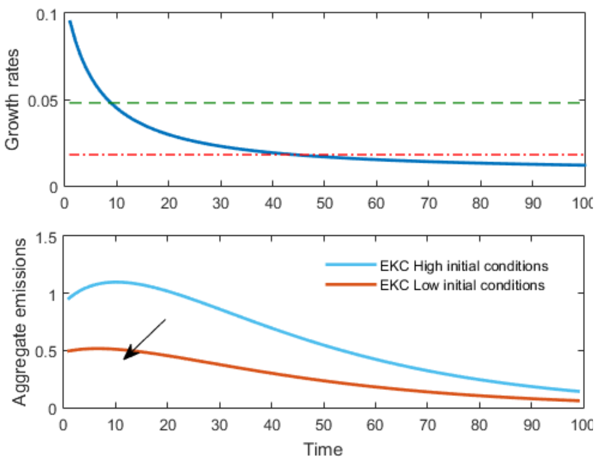
[5] A similar result emerges from changes in population growth. On one hand, population growth lowers the steady state capital per worker which lowers transitional growth for all k. On the other hand, population raises emissions, the growth rate of emissions and the point at which emissions start to fall.

Figure 3: Spatial vs Non Spatially Correlated Technology

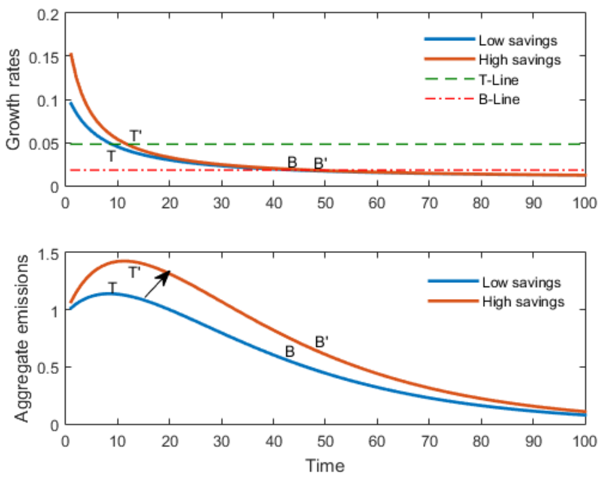


environmental policy cannot turn an unsustainable economy in a sustainable one. This is because of in this model, emission reduction is obtained by a decrease in k_{it} and in y_{it}, not because of increasingly effective abatement. Finally, Panel (d) plots the effects of an increase in the technological progress at abatement. As it can be observed, technological progress in abatement decreases the EKC turning point $k(iT)$ and the steady state growth rate of emissions while leaving the steady state levels of capital and income per effective worker unaltered.

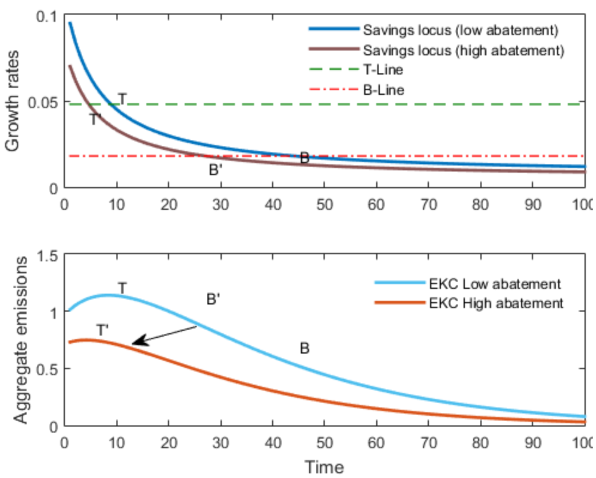
Figure 4: Spatial Green-Solow Model: Sensitivity Analysis



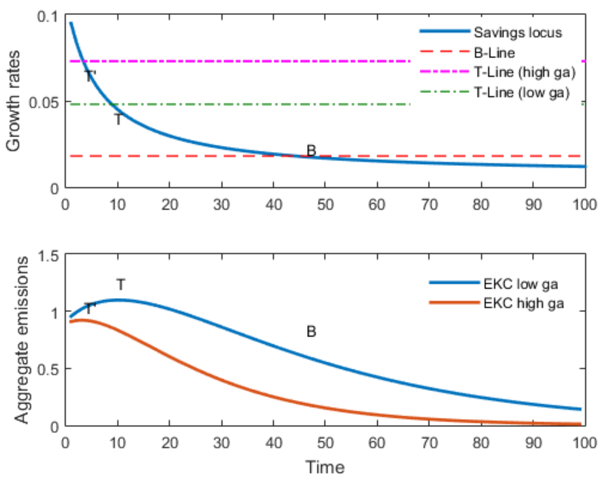
(a) Initial Conditions



(b) Savings



(c) Abatement



(d) Tech Growth Abatement

2.2 Derivation of the Estimation Equation

We now proceed to derive an estimation equation of the growth rate of emissions per capita. To that end, we first use the fact that the growth rate of income per effective worker can be expressed as: $\frac{d\ln y_{it}}{dt} = -\phi\left[\ln y_{it} - \ln y_i^*\right]$. Solving this first-order differential equation and substracting the income per worker at some initial date $\ln y_{it-\tau}$ we obtain:

$$\ln y_{it_2} = e^{-\phi\tau}\ln y_{it_1} + \left(1 - e^{-\phi\tau}\right)\ln y_i^* \tag{19}$$

where $\tau = t_2 - t_1$. Using $\ln y_{it} = \ln y_{it}^c + \ln A_{it}$ where y_{it}^c is the output per capita, we get:

$$\ln y_{it_2}^c = e^{-\phi\tau}\ln y_{it_1}^c + \left(\ln A_{it_2} - e^{-\phi\tau}\ln A_{it_1}\right) + \left(1 - e^{-\phi\tau}\right)\ln y_i^* \tag{20}$$

Stacking the i observations and substituting $\left(\ln \mathbf{A}_{t_2} - e^{-\phi\tau}\ln \mathbf{A}_{t_1}\right)$ by $\left(I_n - \lambda W\right)^{-1}\left(1 - e^{-\phi\tau}\right)\ln \mathbf{A}_0 + \frac{g_b}{1-\lambda}\left(t_2 - e^{-}\phi t_1\right)\iota_n$ in Equation (**??**) we get:

$$\begin{aligned}\ln \mathbf{y}_{t_2}^c\left(I_n - \lambda W\right) = & \quad \left(I_n - \lambda W\right)\left(e^{-\phi\tau}\right)\ln \mathbf{y}_{t_1}^c + \left(1 - e^{-\phi\tau}\right)\ln \mathbf{A}_0 \\ & + g\left(t_2 - e^{-\phi}t_1\right)\iota_n + \left(I_n - \lambda W\right)\left(1 - e^{-\phi\tau}\right)\ln \mathbf{y}^*\end{aligned} \tag{21}$$

Equation (**??** can be simplified to:

$$\ln \mathbf{y}_t^c = \lambda W \ln \mathbf{y}_t^c + \gamma \ln \mathbf{y}_{t-1}^c + \zeta W \ln \mathbf{y}_{t-1}^c + \mathbf{c}_i + \epsilon_{it} \tag{22}$$

where $\gamma = \left(-e^{-\phi\tau}\right)$, $\zeta = -\lambda\left(e^{-\phi\tau}\right)$, $c_i = \left(1 - e^{-\phi\tau}\right)\left(\ln A_{i0} + \frac{\alpha}{1-\alpha}\ln X_i + \frac{\lambda\alpha}{1-\alpha}W\ln X_i\right) + g\left(t_2 - e^{-\phi\tau}t_1\right)$ with $X_i = \left[\frac{s_i}{p+g_b+\delta}\right]$ and ϵ_{it} are added transitory error terms that are assumed to be i.i.d. Finally, we transform Equation (**??**) into the emissions per capita counterpart using $e_{it}^c = \Omega_{it} a\left(\tilde{\theta}\right) y_{it}$ where $a\left(\tilde{\theta}\right) = a\left(\theta\right)/\left[1 - \theta\right]$.

$$\ln \mathbf{e}_t^c = \tilde{\lambda} W \ln \mathbf{e}_t^c + \tilde{\gamma}\ln \mathbf{e}_{t-1}^c + \tilde{\zeta} W \ln \mathbf{e}_{t-1}^c + tilde\mathbf{c}_i + \tilde{\epsilon}_{it} \tag{23}$$

Equation (**??**) takes the form of a Dynamic Spatial Lag Model (DSLM) with coefficient heterogeneity for emissions and the initial levels of emissions. However, note that if elements in c_i such as the savings rate or the population growth rate are assumed to be time-varying which is more realistic, one can express Equation (**??**) as a Dynamic Spatial Durbin Model (DSDM):

$$\ln \mathbf{e}_t^c = \tilde{\lambda} W \ln \mathbf{e}_t^c + \tilde{\gamma} \ln \mathbf{e}_{t-1}^c + \tilde{\zeta} W \ln \mathbf{e}_{t-1}^c + \tilde{\beta} \ln \mathbf{X}_t + \tilde{\psi} W \ln \mathbf{X}_t + \tilde{\mathbf{c}}_i + \epsilon_{it} \tag{24}$$

where $X_t = \frac{s_{it}}{n_t + g_b + \delta_t}$. Furthermore, note that in the previous development we have assumed homogeneous parameters $(\alpha, p, g, \lambda, \delta)$ implying the convergence speed is homogeneous. Relaxation of the restrictions of $p = p_i$ and $\delta = delta_i$ for $i = 1, \ldots, n$ while assuming that $phi = phi_i$ for all $i = 1, \ldots, n$ produces the unconstrained law of motion estimated by **?** and by **?** in the context of growth regressions such that:[6]

$$\ln \mathbf{e}_t^c = \tilde{\lambda} W \ln \mathbf{e}_t^c + \tilde{\gamma} \ln \mathbf{e}_{t-1}^c + \tilde{\zeta} W \ln \mathbf{e}_{t-1}^c + \tilde{\beta} \ln \mathbf{X}_t + \tilde{\psi} W \ln \mathbf{X}_t + \tilde{\mathbf{c}}_i + \epsilon_{it} \tag{25}$$

Using this model it is possible to examine the convergence speed of CO2 emissions per capita given that if $\tilde{\gamma} = a\left(-e^{-\phi\tau}\right) > 0$ (< 0) we may have a positive convergence (divergence) process.[7] Importantly, estimation of different versions of the previous equation allows us to test different competing convergence hypothesis:

(i) *The absolute convergence hypothesis* claims per capita emissions of countries converge to one another in the long-run independently of their initial conditions.

(ii) *The conditional convergence hypothesis* suggests that per capita emissions of countries that are identical in their structural characteristics (i.e, savings, technologies, population growth rates, etc) converge to one another in the long-run independently of their initial conditions.

(iii) *The club convergence hypothesis* suggests that per capita incomes of countries that are identical in their structural characteristics converge to one another in the long run provided that their initial conditions are similar as well.

[6]In this context note that $X_t = (s_{it}, n_{it} + g_b + \delta_{it})$

[7]The transformation employed above should not have effects in our estimates of convergence as long as $a = \Omega_{i0} a\left(\tilde{\theta}\right) = \Omega_{i0} \frac{(1-\theta)^{\epsilon}}{1-\theta} \approx 1$ with $\Omega_{i0} = 1$ and ϵ very small as in **?**

As explained by **?** and **?**, **?** in addition to $\gamma > 0$ in Equation (**??**) the absolute convergence hypothesis constrains $c = \mathbf{c}_i$, $\tilde{\gamma} = \tilde{\gamma}_i \leftrightarrow \phi = \phi_i$ for all i and $\left(\tilde{\zeta}, \tilde{\beta}, \tilde{\psi} = 0\right)$. The conditional convergence hypothesis, relaxes the latter constraint but requires parameter homogeneity $c = \mathbf{c}_i$, $\gamma = \gamma_i$, $\tilde{\zeta} = \tilde{\zeta}_i$, $\tilde{\beta} = \tilde{\beta}_i$, $\tilde{\psi} = \tilde{\psi}_i$ while the club convergence hypothesis allows cross-country variation in $\mathbf{c}_i$, λ_i γ_i, $\tilde{\zeta}_i$, $\tilde{\beta}_i$ and $\tilde{\psi}_i$. [8]

3 Econometric Approach

The empirical counterpart to the implicit model in Equation (**??**) including country fixed is given by:

$$Y_t = \mu + \rho W Y_t + \tau Y_{t-1} + \eta W Y_{t-1} + X_t \beta + W X_t \theta + \epsilon_t \tag{26}$$

where Y_t is a $N \times 1$ vector consisting of observations for the average annual CO2 emissions per capita measured over 5 years windows for every country $i = 1, \ldots, N$ at a particular point in time $t = 1, \ldots, T$, X_t, is an $N \times K$ matrix of exogenous aggregate socioeconomic and economic covariates with associated response parameters β contained in a $K \times 1$ vector that are assumed to influence CO2 emissions per capita. τ, the response parameter of the lagged dependent variable Y_{t-1} is assumed to be restricted to the interval $(-1, 1)$ and $\epsilon_t = (\epsilon_{1t}, \ldots, \epsilon_{Nt})'$ is a $N \times 1$ vector that represents the corresponding disturbance term which is assumed to be i.i.d with zero mean and finite variance σ^2. The variables WY_t and WY_{t-1} denote contemporaneous and lagged endogenous interaction effects among the dependent variable. In turn, ρ is called the spatial auto-regressive coefficient. W is a $N \times N$ matrix of known constants describing the spatial arrangement of the countries in the sample. $\mu = (\mu_1, \ldots, \mu_N)'$ is a vector of country fixed effects. In this context country fixed effects control for all country-specific time invariant variables whose omission could bias the estimates (**?**, **?**). The control variables included in the analysis, the descriptive statistics and the data sources are presented in Table (**??**) below:

The estimator employed in this research to explore the relationship between the set of variables and CO2 emissions per capita is the BCQML developed by (**? ?**).

[8]The assumption of an heterogeneous λ suffices to generate diverse spatial regimes which allows for different intensities in the interaction among economies depending on their concrete spatial location. It is possible to generate multiple emission per capita spatial regimes allowing different degrees of technological connectivity that depend on the spatial allocation such that $\lambda = \lambda_i$ if country i belongs to the steady state basins of attraction defined by $B_r\left(c(i)\right) = \{i \in M | d\left(i, c(i)\right) \leq r\}$ where $d\left(i, c(i)\right)$ is a function of the distance between country i and the center of the club $c\left(i\right)$ and $\lambda = \lambda_j$ otherwise.

Table 1: Data: Descriptive Statistics

Variable	Mean	Standard Deviation	Unit	Source
Carbon Dioxide Emissions per capita	8.236	1.459	ln (mt/pop)	WB
GDP per capita	8.263	1.569	ln (GDP/pop)	PWT
GDP Squared per capita	70.742	26.130	ln $(GDP/pop)^2$	PWT
Investment Share	19.19	11.56	PWT	
Democracy	3.597	6.722	Index	Polity IV
Trade Openess to GDP	79.369	43.819	Percentage	WB
Industry VAB share in GDP	32.600	11.912	Percentage	WB

Notes: (1) GDP per capita is PPP constant prices of 2011. (2) WB denotes World Bank and (3) PWT denotes Penn World Tables.

As shown in **?** **?** the estimation of Equation (**??**) including both time effects and individual effects will yield a bias of the order $O\left(max\left(n^{-1}, T^{-1}\right)\right)$ for the common parameters. By providing an asymptotic theory on the distribution of this estimator, they show how to introduce a bias correction procedure that will yield consistent parameter estimates provided that the model is stable, (i.e, $\tau + \rho + \eta < 1$). As **?** explain, the estimation of a dynamic spatial panel becomes more complex in the case the condition $\tau + \rho + \eta < 1$ is not satisfied. If $\tau + \rho + \eta$ turns out to be significantly smaller than one the model is stable. On the contrary, if its greater than one, the model is explosive and if the hypothesis $\tau + \rho + \eta = 1$ cannot be statistically rejected, the model is said to be spatially co-integrated. Under explosive or spatially co-integration model scenarios, **?**, propose to transform the model in spatial first differences to get rid of possible unstable components in Y_t. This important condition is verified when the estimations are carried out.

Many empirical studies use point estimates of one or more spatial regression models to test the hypothesis as to whether or not spatial spillover effects exist. However, **?** have recently pointed out that this may lead to erroneous conclusions and that a partial derivative interpretation of the impact from changes to the variables of different model specifications provides a more valid basis for testing this hypothesis. Within the context of the DSDM of equation (**??**), the matrix of partial derivatives of Y_t with respect the *k-th* explanatory variable of X_t in country 1 up to country N at a particular point in time t is:

$$\frac{\partial Y_t}{\partial X_t^k} = \left[(I - \rho W)^{-1}\right]\left[\mu + \iota_N \alpha_t + \beta^{(k)} + \theta^{(k)} W\right] \tag{27}$$

Interestingly, in the previous model it is possible to compute own $\partial Y_{it+T}/\partial X_{it}^k$ and cross-partial derivatives $\partial Y_{it+T}/\partial X_{jt}^k$ that trace the effects through time and space.

Specifically, the cross-partial derivatives involving different time periods are referred as diffusion effects, since diffusion takes time. Conditioning on the initial period observation and assuming this period is only subject to spatial dependence (**?**) the data generating process can be expressed as:

$$Y_t = \sum_{k=1}^{K} Q^{-1} \left(\beta^{(k)} + \theta^{(k)} W\right) X_t^{(k)} + Q^{-1} \left(\mu + \iota_N \alpha_t + \epsilon_t\right) \tag{28}$$

where Q is a lower-triangular block matrix containing blocks with $N \times N$ matrixes of the form:

$$Q = \begin{bmatrix} B & 0 & \dots & & 0 \\ C & B & & & 0 \\ 0 & C & \ddots & & \vdots \\ \vdots & & \ddots & & \\ 0 & \dots & & C & B \end{bmatrix} \tag{29}$$

with $C = -(\tau + \eta W)$ and $B = (I_N - \rho W)$. One implication of this, is that by computing C and B^{-1} it is possible to analyze the -own and cross-partial derivative impacts for any time horizon T. Generally, the T-period ahead (cumulative) impact on CO2 emissions per capita from a permanent change at time t in k-th variable is given by:

$$\frac{\partial Y_{t+T}}{\partial X_t^k} = \sum_{s=1}^{T} \left[(-1)^s \left(B^{-1}C\right)^s B^{-1}\right] \left[\mu + \iota_N \alpha_t + \beta^{(k)} + \theta^{(k)} W\right] \tag{30}$$

When T goes to infinity, the previous expression collapses to the long run effect, which is given by:

$$\frac{\partial Y_t}{\partial X_t^k} = \left[(1-\tau) I - (\rho + \eta) W\right]^{-1} \left[\mu + \iota_N \alpha_t + +\beta^{(k)} + \theta^{(k)} W\right] \tag{31}$$

According to **?**, the properties of this partial derivatives are as follows. First, if a particular explanatory variable in a particular region changes, CO2 emissions per capita will change not only that country but also in other countries. Hence, a change in a particular explanatory variable in country i has a *direct* effect on that country, but also an *indirect* effect on the remaining countries. Finally, the *total* effect, which is object of main interest, is the sum of the direct and indirect impacts. Following **?** the direct effect are measured by the average of the diagonal entries whereas the indirect effect is measured by the average of non-diagonal elements.

The model in Equation (**??**) can be contrasted against alternative dynamic spatial panel data model specifications such as the *Dynamic Spatial Lag Model* (DSLM), the *Dynamic Spatial Error Model* (DSEM) and the *Dynamic Spatial Durbin Error Model* (DSDEM). As can be checked, the DSDM can be simplified to the DSLM by shutting down exogenous interactions $\theta = 0$:

$$Y_t = \mu + \tau Y_{t-1} + \rho W Y_t + \eta W Y_{t-1} + X_t \beta + \epsilon_t \tag{32}$$

to the DSDEM if $\eta = \rho\beta = 0$

$$\begin{aligned} Y_t = \mu + \tau Y_{t-1} + X_t \beta + \theta + \upsilon_t \\ \upsilon_t = \lambda W \upsilon_t + \epsilon_t \end{aligned} \tag{33}$$

where $\epsilon_t \sim i.i.d.$, and to the DSEM if $\eta = \theta + \rho\beta = 0$

$$\begin{aligned} Y_t = \mu + X_t \beta + W X_t \theta + \upsilon_t \\ \upsilon_t = \lambda W \upsilon_t + \epsilon_t \end{aligned} \tag{34}$$

In any case, the estimation of the above equations involves defining a spatial weights matrix. Given that this is a critical issue in spatial econometric modeling (**?**) a variety of row-standarized W geographical distance based matrices between the sample regions are considered. The use of geographical distance matrices ensures the exogeneity of the W, as recommended by **?** and avoids the identification problems raised by **?**. Several matrices based on the k-nearest neighbours ($k = 5, 6, \ldots, 15$) computed from the great circle distance between the centroids of the various regions are considered. Additionally, various inverse distance matrices with different cut-off values above which spatial interactions are assumed negligible are employed. As an alternative to these specifications, a set exponential distance decay matrices whose off-diagonal elements are defined by $w_{ij} = exp(-\theta d_{ij})$ for $\theta = 0.005, \ldots, 0.03$ are taken under consideration. The latter matrices, although assume spatial interactions are continuous are characterized by faster decays.

In order to choose between DSDM, DSAR, DSDEM and DSEM specifications of the CO2 emissions, and thus between a global-local, global, local or zero spillovers specifications as well as to choose between different potential specifications of the spatial weight matrix W, a Bayesian comparison approach is applied. Note that this exercise is relevant as it helps to validate whether or not the spillovers and the nature of interactions in the theoretical model are supported by the data. This approach determines the Bayesian posterior model probabilities (PMP) of the alternative specifications given a particular spatial weight matrix, as well as the PIP of different spatial weight matrices given a particular model specification. These probabilities are based on the log marginal likelihood of a model obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. If the log marginal likelihood value of one model or of one W is higher than that of another model or another W, the PMP is also higher. One advantage of Bayesian methods over Wald and/or Lagrange multiplier statistics is that instead of comparing the performance of one model against another model based on specific parameter estimates, the Bayesian approach compares the performance of one model against another model (in this case DSDM against DSDEM, DSLM and DSEM), on their entire parameter space. Moreover, inferences drawn on the log marginal likelihood function values for the models under consideration are further justified because they have the same set of explanatory variables, X and WX, and are based on the same uniform prior for ρ and λ. In this exercise, non-informative diffuse priors for the model parameters $(\tau, \eta, \beta, \theta, \sigma)$ are used following the recommendation of **?**. In particular, the normal-gamma conjugate prior for $\beta, \theta, \tau, \eta$ and σ and a beta prior for ρ:[9]

In order to choose between DSDM, DSAR, DSDEM and DSEM specifications of the CO2 emissions, and thus between a global-local, global, local or zero spillovers specifications as well as to choose between different potential specifications of the spatial weight matrix W, a Bayesian comparison approach is applied. Note that this exercise is relevant as it helps to validate whether or not the spillovers and the nature of interactions in the theoretical model are supported by the data. This approach determines the Bayesian posterior model probabilities (PMP) of the alternative specifications given a particular spatial weight matrix, as well as the PMP of different spatial weight matrices given a particular model specification. These probabilities are based on the log marginal likelihood of a model obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. If the log marginal likelihood value of one model or of one W is higher than that of another

[9]Parameter c are set to zero and T to a very large number $(1e + 12)$ which results in a diffuse prior for β, θ, τ, η while diffuse priors for σ are obtained by setting $d = 0$ and $v = 0$. Finally $a_0 = 1.01$. As noted by **?**, pp. 142, the $Beta\,(a_0, a_0)$ prior for ρ with $a_0 = 1.01$ is highly non-informative and diffuse as it takes the form of a relatively uniform distribution centered on a mean value of zero for the parameter ρ. For a graphical illustration on how ρ values map into densities see Figure 5.3 pp. 143. Also, notice that the expression of the Inverse Gamma distribution corresponds to that of Equation 5.13 pp.142.

model or another W, the PMP is also higher. One advantage of Bayesian methods over Wald and/or Lagrange multiplier statistics is that instead of comparing the performance of one model against another model based on specific parameter estimates, the Bayesian approach compares the performance of one model against another model (in this case DSDM against DSDEM, DSLM and DSEM), on their entire parameter space. Moreover, inferences drawn on the log marginal likelihood function values for the models under consideration are further justified because they have the same set of explanatory variables, X and WX, and are based on the same uniform prior for ρ and λ. In this exercise, non-informative diffuse priors for the model parameters $(\tau, \eta, \beta, \theta, \sigma)$ are used following the recommendation of LeSage (2014). In particular, the normal-gamma conjugate prior for $\beta, \theta, \tau, \eta$ and σ and a beta prior for ρ:[10]

$$\begin{aligned} &\pi(\beta) \sim N(c, T) \\ &\pi\left(\frac{1}{\sigma^2}\right) \sim \Gamma(d, v) \\ &\pi(\rho) \sim \frac{1}{Beta(a_0, a_0)} \frac{(1+\rho)^{a_0-1}(1-\rho)^{a_0-1}}{2^{2a_0-1}} \end{aligned} \tag{35}$$

Columns 1 to 4, in Table (**??**) report the PMP for the different spatial specifications including spatial fixed and time-period fixed effects given alternative specifications of W which allows the comparison of the different models for each W. In columns 5 to 8 for a given spatial specification, PMP across spatial weight matrices are reported. As shown in Table (**??**), for most of the spatial weight matrices the Spatial Durbin appears to be best specification and for the DSDM specification the W matrix with higher PMP is that of 15-nearest neighbors. Importantly, this finding supports the DSDM specification derived from the theoretical model including endogenous and exogenous interaction instead of other possible alternatives. The model comparison also reveals that the DSEM/DSDEM process are never the best candidate to describe CO_2 emissions outcomes.

[10]Parameter c are set to zero and T to a very large number $(1e+12)$ which results in a diffuse prior for β, θ, τ, η while diffuse priors for σ are obtained by setting $d = 0$ and $v = 0$. Finally $a_0 = 1.01$. As noted by LeSage and Pace (2009), pp. 142, the $Beta(a_0, a_0)$ prior for ρ with $a_0 = 1.01$ is highly non-informative and diffuse as it takes the form of a relatively uniform distribution centered on a mean value of zero for the parameter ρ. For a graphical illustration on how ρ values map into densities see Figure 5.3 pp. 143. Also, notice that the expression of the Inverse Gamma distribution corresponds to that of Equation 5.13 pp.142.

Table 2: Spatial Bayesian Model Selection.

	Posterior Probabilities Across Spatial Models				Posterior Probabilities Across Spatial Weight Matrices			
Spatial Weight Matrix	DSDM	DSLM	DSEM	DSDEM	DSDM	DSLM	DSEM	DSDEM
W Cut-off 1000 km	0.000	0.002	0.000	0.000	0.000	0.000	0.000	1.000
W Cut-off 1500 km	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
W Cut-off 2000 km	0.004	0.000	0.249	0.000	0.000	0.000	0.000	1.000
W Cut-off 2500 km	0.994	0.000	0.656	0.000	0.000	0.000	0.000	1.000
W Cut-off 3000 km	0.001	0.000	0.001	0.000	0.000	0.000	0.000	1.000
W $exp-(\theta d)$, $\theta = 0.005$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
W $exp-(\theta d)$, $\theta = 0.01$	0.000	0.000	0.000	0.000	0.000	0.013	0.000	0.987
W $exp-(\theta d)$, $\theta = 0.015$	0.000	0.000	0.000	0.000	0.000	0.932	0.000	0.068
W $exp-(\theta d)$, $\theta = 0.02$	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.995
W $exp-(\theta d)$, $\theta = 0.03$	0.000	0.000	0.000	0.000	0.000	0.097	0.000	0.903
W $exp-(\theta d)$, $\theta = 0.04$	0.000	0.000	0.000	0.000	0.000	0.999	0.000	0.001
W $exp-(\theta d)$, $\theta = 0.05$	0.000	0.000	0.000	0.000	0.000	0.952	0.000	0.048
K-Nearest neighbors ($K = 5$)	0.000	0.998	0.000	0.000	0.000	0.000	0.000	1.000
K-Nearest neighbors ($K = 6$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
K-Nearest neighbors ($K = 7$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
K-Nearest neighbors ($K = 8$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
K-Nearest neighbors ($K = 9$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
K-Nearest neighbors ($K = 10$)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000
K-Nearest neighbors ($K = 15$)	0.000	0.000	0.094	1.000	0.000	0.000	0.000	1.000
W Contiguity	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Notes: We develop Bayesian Markov Monte Carlo (MCMC) routines for spatial panels required to compute Bayesian posterior model probabilities replacing cross-sectional arguments of James LeSage routines by their spatial panel counterparts, for example a block-diagonal $NT \times NT$ matrix, $diag(W, ..., W)$ as argument for W. Note that this requires to transform the data using the orthogonal projector of Lee and Yu ( ) . the All W's are row-normalized.

4 Results

4.1 Baseline Results

Table (??) shows estimation results of different dynamic (A, C and E) and spatial-dynamic (B, D, F) panel data models explaining the evolution of CO2 emissions per capita. Models A and B consist of functional specifications with a constant term and where the only explanatory terms included in the regression are the level of CO2 emissions per capita in period t-1 and the CO2 emissions in period t-1 of neighboring economies. Therefore, these specifications provide the benchmark of the absolute convergence hypothesis. As can be seen, the time lag parameter of CO2 emissions per capita in specifications A and B is positive and significant, but it implies very low convergence rates of 0.6% and 0.7% respectively. Specifications C and D include fixed effects and control for the relevant structural characteristics of the Spatial Green Solow Model presented above (i.e, the level of investment and population growth). As can be seen, in specification C the investment has a positive effect at the 1 % level, while the population growth does not appear to be relevant. On the other hand, in specification D both the investment and the investment in neighboring economies are significant. The fact that the controls are meaningful explaining CO2 emissions per capita provides evidence against the hypothesis of absolute convergence in favor of the conditional convergence. However, it should be noted that the fixed effects are statistically significant for both the model C (LR = 398.88, p = 0.00) and for the D (LR = 471.25, p = 0.00) which suggests that the initial conditions of variables captured in the fixed effects such as the initial level of technological development, affect the evolution of CO2 emissions along the study period. This, in turn, provides evidence for the hypothesis of convergence clubs. To further control for other factors that literature has identified as possible determinants of the level of pollution, models E and F include the level of democratic depth of the country, the degree of trade openness and the share of the industrial sector in the productive structure. Given that the effect of the industry share appears to be statistically relevant, it is likely that prior specifications could be affected by the omitted variable bias. In this regard, it is important also to stress that different measures of goodness of fit point to the specification F as the best of the different alternatives. Finally, note that in this specification the spatio-temporal terms are significant. The estimated time lag is about 0.821, the space-time lag term is -0.293 and the spatial lag term is 0.044. This result confirms that the dynamic spatial panel data modeling framework used in this analysis is suitable for studying the evolution of CO2 emissions per capita and that

CO_2 emissions per capita in neighboring economies affect emissions per capita of any country.

As mentioned in the previous section, correct interpretation of the parameter estimates in the DSDM requires to take into account the direct, indirect and total effects associated with changes in the regressors. Table (**??**) shows this information. Considering the average direct impacts of Table (**??**), it is important to notice that there are some differences to the DSDM model coefficient estimates reported in Table (**??**). Differences between these two measures are due to feedback effects passing through the entire system and ultimately reaching the country of origin.

Focusing on the main aim of the paper, we now proceed to examine the issue of CO_2 emissions per capita convergence. To that end, we use the Error Correction Model representation following **?** to simulate the convergence direct, indirect and total effects. Results reveal that the relationship between initial levels of emissions and future emissions growth rates is negative and statistically significant, thus confirming the empirical evidence provided by the previous analysis of **?**. In particular, the estimates show that a 1% increase the initial level of per capita emissions is associated with a decrease in the average growth rate of around -0.25%. Nevertheless, this total convergence effect is the sum of the direct and indirect impact of the initial level on its growth rate. The direct effect, Table (**??**) indicates that an increase in the initial level of emissions registered by a specific country exerts a negative and statistically significant impact on its growth rate. In turn, the indirect effect shows that this increase also influences negative and significantly on the growth rates of neighboring countries. Overall, we find that the implied speed of convergence is 5.07% which is higher than the 1.6% obtained by **?** , which can be explained by the relevance of spatio-temporal interactions.

Direct impact estimates in Table (**??**), disp lay interesting features which are worth mentioning. First, as regards the investment there is evidence that an increase in the investment in country i exerts increases emissions per capita in i. Second, it is observed that higher population growth rates and higher shares of industry in the sectoral composition affect positively emissions in i. On the other hand, the effect of an increase in the democratic depth and in the trade openness in country i by itself does not affect emissions per capita in i. Short run indirect effects are significant at the 5% level for five out of six variables. Indirect effects amplify significantly direct effects in the case of investment and democracy while go in the opposite direction for the population growth rate and the industry share. The results show that the amplification phenomenon is particularly pronounced as it indirect effects account

Table 3: CO_2 Convergence Equation Estimation Results

Models	No Spatial Dynamic (A)	Spatial Dynamic(B)	No Spatial Dynamic (C)	Spatial Dynamic (D)	No Spatial Dynamic (E)	Spatial Dynamic (F)
Constant	0.292***	0.116**				
	(9.39)	(2.52)				
ln Emissions pc (t-1)	0.970***	0.965***	0.743***	0.861***	0.725***	0.821***
	(236.11)	(176.42)	(48.57)	(48.80)	(47.41)	(45.73)
Implied ϕ	0.006	0.007	0.059	0.030	0.064	0.039
ln Neighbor's emissions pc (t-1)		-0.504***		-0.333***		-0.293***
		(-10.56)		(-5.47)		(-4.84)
Investment (t)			0.011***	0.007***	0.011***	0.007***
			(8.14)	(4.89)	(8.30)	(5.24)
Neighbor's Investment (t)				0.031***		0.031***
				(6.59)		(6.73)
Population growth (t)			0.007	0.005	0.001	0.008
			(0.86)	(0.64)	(0.11)	(0.92)
Neighbor's Pop. growth (t)				-0.001		-0.018
				(-0.03)		(-0.72)
Democracy (t)					-0.001	0.000
					(-0.94)	(-0.19)
Neighbor's Democracy (t)						-0.004
						(-1.08)
Trade Openess (t)					0.000	0.000
					(1.15)	(0.92)
Neighbor's Trade Openess (t)						0.000
						0.01
Industry Share (t)					0.009***	0.008***
					(7.22)	(5.75)
Neighbor's Industry Share (t)						-0.010***
						(-2.77)
ln Neighbor's emissions pc (t)		0.528***		0.227***		0.044***
		(10.95)		(3.50)		(12.38)
Fixed Effects	No	No	Yes	Yes	Yes	Yes
$\tau + \rho + \eta$	0.970	0.989	0.743***	0.755***	0.725***	0.572**
R2	0.978	0.980	0.984	0.986	0.985	0.987
Log Like	-84.348	-12.146	140.950	170.666	169.591	209.452
Sige	0.067	0.060	0.047	0.047	0.045	0.045
PMP	0.00	0.00	0.00	0.00	0.00	1.00

Notes: The dependent variable is in all cases the level of emissions per capita for the various countries. t-statistics in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

for more than a half of the total effect. The interpretation of this result is that if all countries $j = 1, \ldots, N$ other than i experiment a change in X^k, this will have a stronger effect in i that if only i experiments a change in X^k even if i generate spillover effects that go back to i. This is due to the fact that the DSDM contains a global spillover multiplier. As mentioned above, the sum of direct and indirect effects allows one to quantify the total effect on CO2 emissions per capita of the different control variables. When direct and indirect effects are jointly taken into account, Table (**??**) indicates that the total effect is statistically significant exclusively in the case of investment, population growth and democracy.

Table 4: Effects Decomposition

Variables	Direct Effects	Indirect Effects	Total Effects
Convergence effect			
Initial Emissions	-0.146***	-0.108**	-0.253***
	(-7.44)	(-1.97)	(-4.79)
Implied ϕ	0.0291	0.0216	0.0507
Short term			
Investment	0.007***	0.032***	0.039***
	(4.13)	(5.14)	(6.02)
Population growth	0.031***	-0.132***	-0.101***
	(3.58)	(-3.67)	(-2.72)
Democracy	-0.001	-0.024***	-0.026***
	(-0.58)	(-4.82)	(-5.30)
Trade Openess	0.000	0.001	0.001
	(0.83)	(0.84)	(1.09)
Industry share	0.009**	-0.015***	-0.006
	(6.22)	(-2.77)	(-1.08)
Long term			
Investment	0.042***	0.118**	0.161***
	(3.17)	(2.53)	(3.34)
Population growth	0.252***	-0.658***	-0.406***
	(3.01)	(-4.04)	(-2.71)
Democracy	-0.004	-0.103***	-0.107***
	(-0.25)	(-2.75)	(-3.01)
Trade Openess	0.002	0.003	0.005
	(0.75)	(0.56)	(1.03)
Industry share	0.068***	-0.093***	-0.025
	(4.68)	(-3.26)	(-0.99)

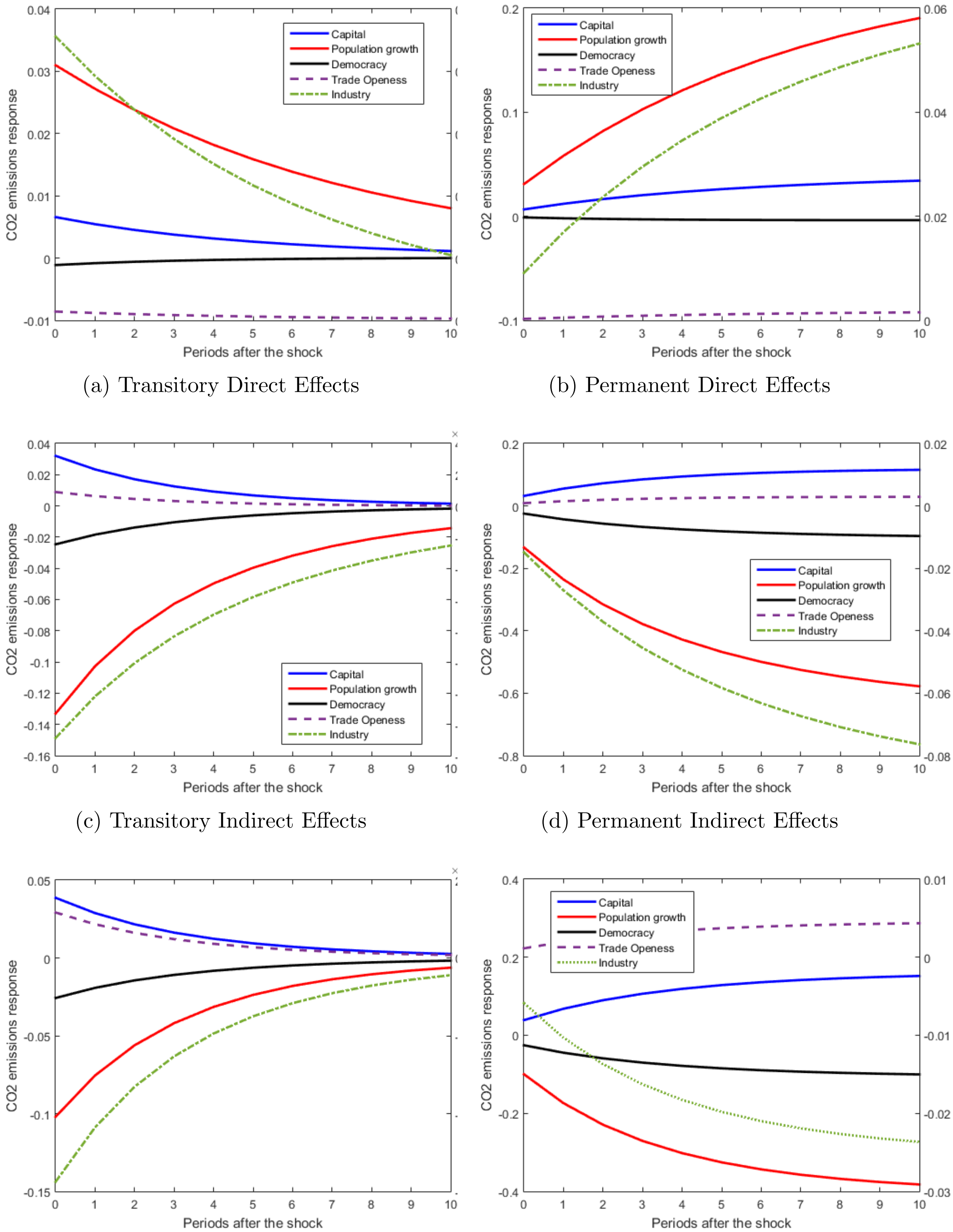
Notes: Inferences regarding the statistical significance of these effects are based on the variation of 1000 simulated parameter combinations drawn from the variance-covariance matrix implied by the BCML estimates of Equation (**??**). To compute the speed of convergence we use the error correction model (ECM) representation of Equation (**??**) following **?** pp 300. t-statistics in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

To study the dynamic responses of CO2 emissions per capita to changes in the different regressors, the model is used to perform impulse-response analysis using Equation (??). Impulse-response functions in a dynamic spatial panel context contain both, temporal dynamic effects and spatial diffusion effects which correspond to exogenous changes that propagate across space. Figure (??) decomposes the dynamic trajectory of CO2 emissions per capita after a transitory change in a regressor into direct (a), indirect (c) and total responses (e) and after a permanent change (subfigures b, d, and f) which in the infinite is exactly the long-run effect reported in the last rows of Table (??). In Figure (??) we plot the trade openness and the industry share with dashed lines and in the right y-axis to differenciate with respect investment, population and democracy which are statistically significant in both the short and the long run. We find that with the time, direct cumulative effects of investment increase its share with respect the total long run effect while on the contrary, democracy and population growth direct effects decrease its relevance which implies that spatio-temporal diffusion is particularly relevant for the later. Exploration of the propagation pattern reveals that simultaneous effects occurring in the period of impact of the shock are around the 23% of the long-run effect. Importantly, three periods after the shock, the cumulative effect accounted for a 65% of the long run impact. Focusing on the long run we find that after five periods (25 years) the figure is around the 80% and that ten periods later (50 years) the cumulative effect amounts to a 95%. These results suggest that, the full effect on CO2 emissions per capita resulting from changes in the model regressors takes time to materialize and the short run analysis may considerable under-estimate the final effects.

4.2 An analysis of the Convergence Club Dynamics

In the previous analysis we have seen that: (i) the economic surrounding of a country seems to influence the CO2 Emissions per capita perspectives for that country, which is also reflected in the fact that initial CO2 emissions of neighboring economies have a statistically significant effect in the process of convergence and that (ii) the fixed effects are significant, providing evidence supporting the hypothesis of club rather than that of conditional convergence. In this regard, it should be stressed that the club convergence hypothesis is consistent with heterogeneous parameter and multiple regime processes, while the analysis we carried out so far so only explored the heterogeneity in the parameters through a DSDM homogeneous model with fixed effects. Moreover, it is important to note that in the field of spatial econometrics and growth, some studies linked the concept of spatial heterogeneity with that of the convergence

Figure 5: CO2 Emissions Dynamic Diffusion Effects



(a) Transitory Direct Effects

(b) Permanent Direct Effects

(c) Transitory Indirect Effects

(d) Permanent Indirect Effects

(e) Transitory Total Effects

(f) Permanent Total Effects

spatial clubs. [11].

To investigate the presence of spatial clubs in CO2 emissions and the possible existence of heterogeneous dynamics due to different spatial regimes we employ a new methodology based on a dynamic version of the Moran Scatterplot. Therefore, in this context, spatial clubs should be meant as clusters of countries with similar levels both of CO2 and spatially lagged CO2 at the beginning of the sample. We follow the methodology applied in **?** to identify the spatial clubs. In particular, we apply the *k-median* algorithm for $k = 2, 3$ and 4 and compute Posterior Model Probabilities and perform likelihood ratio tests to identify which club classification is more consistent with the data. [12]

Table 5: Heterogeneous Regime Selection

	Homogenous DSDM	Two Clubs DSDM	Three Clubs DSDM	Four Clubs DSDM
Log Like	209.45	274.41	301.26	309.10
PMP	0.02	0.12	0.70	0.16

Notes: To compute the likelihood and the posterior of the various models we extend the variance-covariance matrix derived in Elhorst and Freret (2009) for the case of two-regime fixed effects spatial lag model, pp 940.

The results of Table (**??**) show that the model with highest probability is that of three spatial clubs (PMP = 0.7). This result is confirmed by iterative likelihood-ratio tests of against the homogeneous DSDM and the two-regime DSDM. While the model characterizing the process of CO2 emissions per capita can not be simplified from 4 clubs to 1 (LR = 141.88, p = 0.00) or to 4 to 2 clubs (LR = 69.37, p = 0.00) and while the model of 3 clubs can not be reduced to 1 (LR 126.21, p=0.00) or to 2 clubs (LR=53.69, p=0.00) the model with 4 clubs describes the data worse than that of 3 clubs (LR 15.67, p = 0.26). The evidence stemming from the Local Directional Moran Scatterplot suggests the persistence of three clubs where Club 2 tends to converge to Club 3 while Club 3 seems fairly stable in its relative position. For more details on the dynamic evolution of the identified clubs see the Appendix.

The estimation results of the heterogeneous DSDM are shown in Table ( ). The first three columns display the estimated parameters for regressors in each of the clubs while the last three columns report the differences. As can be seen the results

[11]Some studies have employed heterogeneous X parameter models (Baumont et al., 2004; Ertur and Koch, 2007) while others considered heterogeneous spatial regimes for Y (Elhorst and Freret, 2009) but not both

[12]The *k-median* algorithm is a variation of *k-means* algorithm where instead of calculating the mean for each cluster to determine its centroid, it is use its median. The use of median should minimize the impact of possible outliers, (see **?** for more details on *k-median* algorithm).

obtained suggest the need to consider heterogeneity in the modeling process as for many regressors disparities are highly relevant. Similarly, w

4.3 The Spatial Environmental Kuznets Curve

Finally we check one of the main regularities emerging from our theoretical model, the EKC. Table () reports the main results of the EKC. The first column of Table () presents the results obtained in the estimation of the DSDM when employing the BCML estimator. Importantly, the results regarding the direct effects of the GDP and the GDP squared seem to indicate the existence of EKC relationship as emissions increase with the GDP but decrease with the squared GDP. In turn, the indirect effect shows that changes in the GDP and the GDP squared also influence significantly emission levels of neighbouring countries. Indeed, the indirect effect accounts for more than half of the overall total impact caused by changes in GDP and GDP squared, thus corroborating the empirical relevance of spatial spillovers in this context. The analysis of the total effects displays interesting features that are consistent with the empirical literature of environmental economics (Yandle *et al.*, 2002; Dinda, 2004). Higher levels of democracy and industry are related to lower levels of CO2 emissions while a higher level of trade openness does not affect CO2 emissions.

5 Conclusions

This paper analyzes the evolution of CO2 emissions per capita in a sample of 141 countries during the period 1970-2014. The study extends the neoclassical Green Solow Model to take into account technological externalities in the analysis of CO2 emissions per capita growth rates. Spatial externalities are used to model technological interdependence, which ultimately implies that the CO2 emissions rate of a particular country is affected not only by its own degree of emissions but also by the pollution generated by the remaining countries. In order to investigate the empirical validity of this result, convergence in CO2 emissions is examined by means of dynamic spatial panel econometric techniques. Estimates show the existence of a negative and statistically significant relationship between initial levels of CO2 emissions and subsequent growth rates. This finding is partly due to the role played by spatial spillovers induced by neighboring economies. The observed link is robust to the inclusion in the analysis of different explanatory variables that may affect CO2 emissions growth rates. In a second step, combining recently developed spatial-non parametric techniques with spatial bayesian model selection techniques we identify three distinct clubs in the distribution of CO2 emissions per capita. The estimation of the corresponding three-endogenous regime dynamic spatial model with parameter

Table 6: Heterogeneous Dynamic Spatial Durbin Model

	Club Estimates			Club Disparities		
	Club 1	Club 2	Club 3	Differences Club 1 vs Club2	Differences Club 1 vs Club3	Differences Club 2 vs Club3
Emissions per capita (t-1)	0.707***	0.718***	0.538***	-0.011	0.169***	0.180***
	(27.86)	(26.00)	(14.56)	(-0.29)	(3.77)	(3.94)
Investment	0.010***	0.008***	0.004*	0.001	0.006	0.004
	(3.20)	(4.15)	(1.87)	(0.41)	(1.48)	(1.37)
Population growth	0.030**	0.036*	-0.001	-0.006	0.031*	0.037*
	(2.23)	(1.89)	(-0.08)	(-0.25)	(1.79)	(1.71)
Democracy	-0.002	-0.004	-0.004	0.001	0.001	0.000
	(-0.80)	(-1.45)	(-1.13)	(0.33)	(0.30)	(0.00)
Trade	0.003***	0.000	-0.001***	0.003***	0.004***	0.001
	(3.72)	(-0.82)	(-2.58)	(3.68)	(4.54)	(1.59)
Industry	0.003	0.008***	0.014***	-0.006*	-0.011***	-0.005*
	(1.22)	(3.88)	(6.47)	(-1.84)	(-3.51)	(-1.74)
Neighbor's Emissions per capita (t-1)	-0.414***	-0.044	-0.086	-0.371***	-0.329***	0.042
	(-4.41)	(-0.46)	(-0.90)	(-2.82)	(-2.66)	(0.34)
Neighbor's Investment	0.050***	0.021**	0.031***	0.028**	0.019	-0.010
	(5.05)	(2.45)	(3.73)	(2.17)	(1.47)	(-0.79)
Neighbor's Population growth	-0.048	0.040	0.010	-0.087	-0.058	0.030
	(-0.85)	(0.57)	(0.29)	(-0.94)	(-0.88)	(0.37)
Neighbor's Democracy	-0.023***	0.009	-0.019**	-0.032***	-0.004	0.028**
	(-3.06)	(1.13)	(-2.26)	(-2.80)	(-0.38)	(2.44)
Neighbor's Trade Openess	0.004	0.003	-0.001	0.001	0.005	0.004
	(1.20)	(1.45)	(-0.73)	(0.23)	(1.36)	(1.54)
Neighbor's Industry	0.006	-0.026***	-0.037***	0.032***	0.042***	0.011
	(0.68)	(-2.86)	(-4.74)	(2.60)	(3.70)	(0.87)
Neighbor's Emissions per capita (t)	0.195**	0.171**	-0.049	0.024	0.244**	0.220**
	(2.11)	(2.06)	(-0.57)	(0.27)	(2.54)	(2.56)

Notes: The t-test on the statistical significance of disparities in the effects among clubs i and j for each factor k is computed as $t_k = \frac{D_k}{\sqrt{\Sigma_k}} = \frac{R_{k(si)} - R_{k(sj)}}{\sigma^2_{k(si)} + \sigma^2_{k(sj)} - 2Cov_{k(si),k(sj)}}$ where $R_{k(s)}$ is the average coefficient in the club s and $\sigma^2_{k(s)}$ and $Cov_{k(si),k(sj)}$ denote the variances and covariance of the estimates for factor k. t-statistics in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 7: Propagation of Short Run Effects Across Convergence Clubs

Origin	Variable	Club1 Response	Club 2 Response	Club 3 Response
	Investment	0.016***	0.019***	0.026***
	Population growth	-0.002	0.000	-0.003
Club1	Democracy	-0.006**	-0.008**	-0.010**
Shock	Trade Openess	0.002**	0.002**	0.003
	Industry	0.002	0.003	0.004
	Investment	0.008**	0.010**	0.013*
	Population growth	0.024	0.031*	0.039
Club 2	Democracy	0.001	0.001	0.002
Shock	Trade Openess	0.001	0.001	0.001
	Industry	-0.004*	-0.004	-0.006*
	Investment	0.009**	0.011***	0.015***
	Population growth	0.002	0.002	0.003
Club 3	Democracy	-0.006**	-0.007**	-0.010**
Shock	Trade Openess	-0.001**	-0.001**	-0.001**
	Industry	-0.005*	-0.004	-0.008*

Notes: Inferences regarding the statistical significance of the total effects in CO2 emissions per capita are based on the variation of 1000 simulated parameter combinations drawn from the variance-covariance matrix implied by the BCML estimates of the Three-Regime DSDM. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

heterogeneity reveals that in the context of CO2 emissions per capita, the hypothesis of the spatial convergence clubs is more consistent with the data than that of conditional convergence. Finally, we estimated a model of the EKC finding the U shaped predicted by the Spatial Green Solow Model.

References

Aldy, J.E. (2006): Per Capita Carbon Dioxide Emissions: Convergence or Divergence?. *Environmental and Resource Economics*, 33, 4, 533-555.

Anselin, L. and Bera, A. (1998): *Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics.* In: A. Ullah and D.E.A. Giles (eds.), *Handbook of Applied Economic Statistics*, 237-289. Marcel Dekker, New York.

Baltagi, B.H. (2001): *Econometric Analysis of Panel Data.* Second Edition. John Wiley & Sons, New York.

Bardi, U. (2011): *The Limits to Growth Revisited.* Springer-Verlag New York, 1, XIII, 119.

Table 8: Environmental Kuznets Curve Results

			Short Run			Long Run		
	Non Spatial Model	Spatial Model	Direct Effects	Indirect Effects	Total Effects	Direct Effects	Indirect Effects	Total Effects
$\ln GDPpc$	1.055***	0.719***	0.743***	1.72***	2.463**	3.537***	3.864**	7.401***
	(10.80)	(6.32)	(-64.53)	(-35.24)	(-48.47)	(-53.07)	(-21.12)	(-41.54)
$(\ln GDPpc)^2$	-0.051***	-0.033***	-0.034***	-0.091***	-0.125***	-0.162***	-0.216**	-0.378***
	(-9.36)	(-5.03)	(-5.14)	(-3.32)	(-4.40)	(-4.35)	(-2.11)	(-3.79)
Democracy	-0.006***	-0.002	-0.002	-0.022***	-0.025***	-0.009	-0.067***	-0.076***
	(-4.01)	(-1.03)	(-1.14)	(-4.75)	(-5.52)	(-0.85)	(-2.98)	(-3.68)
Trade Openess	0.000	0.000	0.000	0.000	0.001	0.002	-0.000	0.001
	(0.17)	(1.02)	(1.01)	(0.09)	(0.30)	(0.99)	(-0.03)	(0.31)
Industry Share	0.008***	0.007***	0.007***	-0.015***	-0.008	0.035***	0.035***	-0.056***
	(6.43)	(4.63)	(-45.43)	(-2.63)	-(1.42)	(-42.95)	(-3.09)	(-1.35)
Neighbor's $\ln GDPpc$		1.030***						
		(2.83)						
Neighbor's $(\ln GDPpc)^2$		-0.056***						
		(-2.76)						
Neighbor's Democracy		-0.016***						
		(-3.88)						
Neighbor's Trade Openess		0.000						
		(0.00)						
Neighbor's Industry share		-0.012***						
		(-3.13)						
ln Emissions pc (t-1)	0.637***	0.799***						
	(35.03)	(41.42)						
ln Neighbor's emissions pc (t-1)	0.014	-0.327***						
	(0.44)	(-5.20)						
ln Neighbor's emissions pc (t)		0.286***						
		(4.29)						
Log Like								
Sige	0.041	0.043						
R2								
Model Probs								

Notes: Inferences regarding the statistical significance of these effects are based on the variation of 1000 simulated parameter combinations drawn from the variance-covariance matrix implied by the BCML estimates. t-statistics in parentheses. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Barro, R.J. and Sala-i-Martin, X. (1997): Technological Diffusion, Convergence, and Growth, *Journal of Economic Growth*, Springer, vol. 2(1), pages 1-26, March.

Brook, W.A. and Taylor, M.S.(2010): The Green Solow model. *Journal of Economic Growth*, 15, 2, 127153.

Carson, R.T., Jeon, Y., Mc Cubbin, D.R.(1997): The relationship between air pollution emissions and income: US data. *Environment and Development Economics*, 2, 433-450.

Cole, M.A., Rayner, A.J. and Bates, J.M. (1997): The Environmental Kuznets Curve: An Empirical Analysis.*Environment and Development Economics* 2, 4, 401-416.

Corrado L, Fingleton B (2012): Where is the Economics in Spatial Econometrics? *Journal of Regional Science*, 52, 210-239.

Dasgupta, P. and Mäler K.G.(2000): Net National Product, Wealth, and Social Well-Being,*Environment and Development Economics*, 5(2): 69-93.

Debarsy, N, Ertur, C, and LeSAGE J.P., (2010): Interpreting Dynamic Space-Time Panel Data Models, LEO Working Papers, Laboratoire d'Economie d'Orleans (LEO), University of Orleans.

Dechert, W.D.(2001): Growth Theory, nonlinear dynamics and Economic modeling,Edward Elgar: Cheltenham

Durlauf, S.N. and Johnson, P.A. (1995): Multiple Regimes and Cross-Country Growth Behaviour, *Journal of Applied Econometrics*, John Wiley and Sons, Ltd., vol. 10(4), pages 365-84, Oct.-Dec.

Elhorst, J.P. (2010): Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, 5, 9-28.

Elhorst, J.P, Zandberg, E., Jakob De Haan (2013): The Impact of Interaction Effects among Neighbouring Countries on Financial Liberalization and Reform: A Dynamic Spatial Panel Data Approach. *Spatial Economic Analysis*, 8:3, 293-313.

Elhorst, J.P.(2014): *Spatial Econometrics: From Cross-sectional Data to Spatial Panels.* Springer: Berlin New York Dordrecht London.

Ertur, C. and Koch, W. (2007): Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence. *Journal of Applied Econometrics*, 22, 1033-1062.

Ezcurra, R. and Rios, V. (2015):Volatility and Regional Growth in Europe: Does Space Matter? *Spatial Economic Analysis*, 10, 3, 344-368 .

Ezcurra, R. (2007): Is there cross-country convergence in carbon dioxide emissions?. *Energy Policy*, Volume 35, Issue 2, February 2007, Pages 13631372

Fiaschi, D. and Gianmoena, L and Parenti, A.(2014): Local Directional Moran Scatter Plot - Ldms, *Region et Developpement*, LEAD, Universite du Sud - Toulon Var, vol. 40, pages 97-112.

Fischer M (2011): A Spatial Mankiw-Romer-Weil Model: Theory and Evidence. *Annals of Regional Science*, 47, 419-436.

Fingleton, B. and López-Bazo E. (2006): Empirical growth models with spatial effects, *Papers in Regional Science*, 85 (2006), pp. 177198

Grossman, G.M. and Krueger, A.B.(1995): Economic Growth and the Environment. *The Quarterly Journal of Economics*, 110, 2, 353-377.

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

Johnson, P. and Takeyama, L.N. (2001): Convergence Among the U.S. States: Absolute, Conditional, or Club?, Vassar College Department of Economics Working Paper Series 50, Vassar College Department of Economics, revised Oct 2003.

Johnson, P. and Takeyama, L.N. (2001): Initial conditions and economic growth in the US states, *European Economic Review*, Elsevier, vol. 45(4-6), pages 919-927, May.

Kijima, M., Nishide, K. and Ohyama, A.(2010): Economic models for the environmental Kuznets curve: A survey. *Journal of Economic Dynamics and Control*, 34, 7, 11871201

Leisch, F. (2006): A toolbox for k-centroids cluster analysis. *Computational Statistics and Data Analysis*, 51(2):526544.

Lee, L.F. and Yu, J.(2010): Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects. *Journal of Econometrics*, 154, 165-185.

Lee, L.F. and Yu, J. (2010): A spatial dynamic panel data model with both time and individual effects. *Econometric Theory*, 26, 564-594.

LeSage, J.P. (2014): Spatial Econometric Panel Data Model Specification: A Bayesian approach. *Spatial Statistics*, 9, 122-145.

LeSage, J.P. and Pace, R.K.(2009): *An Introduction to Spatial Econometrics*. Chapman and Hall, Boca Raton, FL.

Lieb, CH.M.,(2003): The Environmental Kuznets Curve. A Survey of the Empirical Evidence and of Possible Causes. *University of Heidelberg, Department of Economics, Discussion Paper Series*, No. 391.

Maddison, D.(2006): Environmental Kuznets curves: A spatial econometric approach. *Journal of Environmental Economics and Management*, 51, 2, 218230.

Manski, C.F.(1993): Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60, 531-542.

Meadows, Donella H.(1972): The Limits to growth; a report for the Club of Rome's project on the predicament of mankind. New York: Universe Books.

Magrini, S.(2007): Analysing Convergence through the Distribution Dynamics Approach: Why and How?. *Working Paper Department of Economics, Ca Foscari University of Venice* No. 13/WP/2007.

Mankiw, N.G., Romer, D. and Weil, D.(1992): A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 107:407437, 1992

Nguyen Van (2005): Distribution dynamics of CO2 emissions. *Environmental and Resource Economics*, 32, 4, 495508.

Nordhaus, William D., (1993): Rolling the "DICE": an optimal transition path for controlling greenhouse gases, *Resource and Energy Economics*, Elsevier, vol. 15(1), pages 27-50, March.

Silverman, B.W.(1986): Density Estimation for Statistics and Data Analysis. Chapman and Hall, London, 1986.

Stern D.I.(2004): The Rise and Fall of the Environmental Kuznets Curve. *World Development*, 32, 8, 14191439.

Yu, J., Jong, R. and Lee, F. (2012): Estimation for spatial dynamic panel data with fixed effects: the case of spatial cointegration. *Journal of Econometrics*, 167: 16-37.

Appendix: Spatial Clubs and Convergence Dynamics

Figure (1) shows a clear spatial distribution of the CO2 emission among countries. A classical way to analysis the spatial clusterization of the economies for different level of CO2 emission is thought the *Moran Scatter Plot*. The advantage of Moran scatter plot is its easy interpretation giving a graphical representation of the relation of the variable (in our case the (relative) CO2 emission) in one country with respect the values of that variable in the neighbouring countries, i.e the spatial lag variable $W \times CO2$.

Figure (6) shows the Moran scatter plot for different sub-period, 1970-75, 1076-80, 1981-85, 1986-90, 1991-95, 1996-00, 2001-05, 2006-10 and 2011-15. In all the sub-periods the distribution of the CO2 appears clusterized in three different groups, suggesting the formation of three different clubs, indicated by three yellow circles: a first club **C1** characterized by countries with lower level of CO2, a second one **C2** with medium level and, **C3** which higher level of emission.

We follow the methodology applied in **?** to indentify the spatial clubs. Precisely, we apply the *k-median* algorithm. The *k-median* algorithm is a variation of *k-means* algorithm where instead of calculating the mean for each cluster to determine its centroid, it is use its median. The use of median should minimize the impact of possible outliers, (see **?** for more details on *k-median* algorithm). In this contest, spatial clubs should be meant as clusters of country with similar levels both of CO2 and spatially lagged CO2.

The evidence from the Moran scatter plot in different sub periods suggests the persistence of three clubs. However we observe some little movements among clubs. In particular we observe a process of clusterization inside all of them (moving closer to the bisector) and (i) a process of divergence for C1 (it tends to move toward lower level of CO2 along the bisector), and (ii) a convergence process between C2 and C3. The overall impression is that club C2 tends to converge to C3, while the club C1 seems fairly stable as (relative) position.

However, the comparison among the Moran scatter plot in different periods of time in Figures (6) does not provide any information on the dynamics of these three spatial clubs. To fill this gap, the evolution in time and space of the three spatial clubs is analyzed through the *Local Directional Moran Scatter Plot* developed by (**?**).[13] Precisely, the spatial dynamics of the clubs in the Moran space is represented by a random vector field (RVF), which measure for each point in the lattice (defined in the y and $W \times y$ space) the *expected movement* calculated on the base of the distribution of probabilities of the observed movement of the real observed data.

In particular, given a subset L of the possible realization of (y, Wy) (i.e. a lattice

[13]For a detailed explanation of the methodology see (**?**)

in the Moran space), a RVF is represented by a random variable $\Delta_\tau z_i$, where $\Delta_\tau z_i \equiv (\Delta_\tau y_i, \Delta_\tau W y_i) \equiv (y_{it+\tau} - y_{it}, W y_{it+\tau} - W y_{it})$, indicating the spatial dynamics (i.e. the dynamics from period t to period $t+\tau$ represented by a movement vector) at $z_i \equiv (y_i, W y_i) \in L$. For each point in the lattice z_i we estimate the τ-period ahead *expected movement* $\mu_{\Delta_\tau z_i} \equiv E\left[\Delta_\tau z_i | z_i\right]$ using a *local mean estimator*, where the observations are weighted by the probabilities $\omega\left(z_i, z_{jt}^{OBS}\right)$ derived from the kernel function, i.e.:

$$\widehat{\mu}_{\Delta_\tau z_i} = \sum_{t=1}^{T-\tau} \sum_{j=1}^{N} \omega\left(z_i, z_{jt}^{OBS}\right) \Delta_\tau z_{jt}^{OBS} = \Pr\widehat{(\Delta_\tau z | z_i)} \Delta_\tau z^{OBS}. \tag{36}$$

where,

$$\omega\left(z_i, z_{jt}^{OBS}\right) = \frac{K\left(\frac{(z_i - z_{jt}^{OBS})^T S^{-1} (z_i - z_{jt}^{OBS})}{h^2}\right) \frac{\det(S)^{-\frac{1}{2}}}{2h^2}}{\sum_{t=1}^{T-\tau} \sum_{j=1}^{N} K\left(\frac{(z_i - z_{jt}^{OBS})^T S^{-1} (z_i - z_{jt}^{OBS})}{h^2}\right) \frac{\det(S)^{-\frac{1}{2}}}{2h^2}} \tag{37}$$

The probabilities $\omega\left(z_i, z_{jt}^{OBS}\right)$ are estimate via a kernel function (K)[14] which allow us to measure the distance between z_i (a point in the lattice L) and z_{jt}^{OBS} (which is the corresponding point in the lattice of a given observation in the space y and Wy).

Graphically, Equation (**??**) is the RVF, represented by a red arrows. In the estimation we set $\tau = 45$ that corresponds to the last period available, so that, starting from the information of the period 1970 we obtain the expected dynamics τ-period ahead.

Figure (7) reports the estimated expected value of the 45-year ahead directions. The three spatial clubs present in 1970 are still there in 2015, but the expected dynamics of convergence between club C2 and C3 and the divergence for C1 is now much clear.

[14] In the estimation we use a multivariate Epanechnikov kernel (see **?**)

Figure 6: Convergence Clubs Dynamics

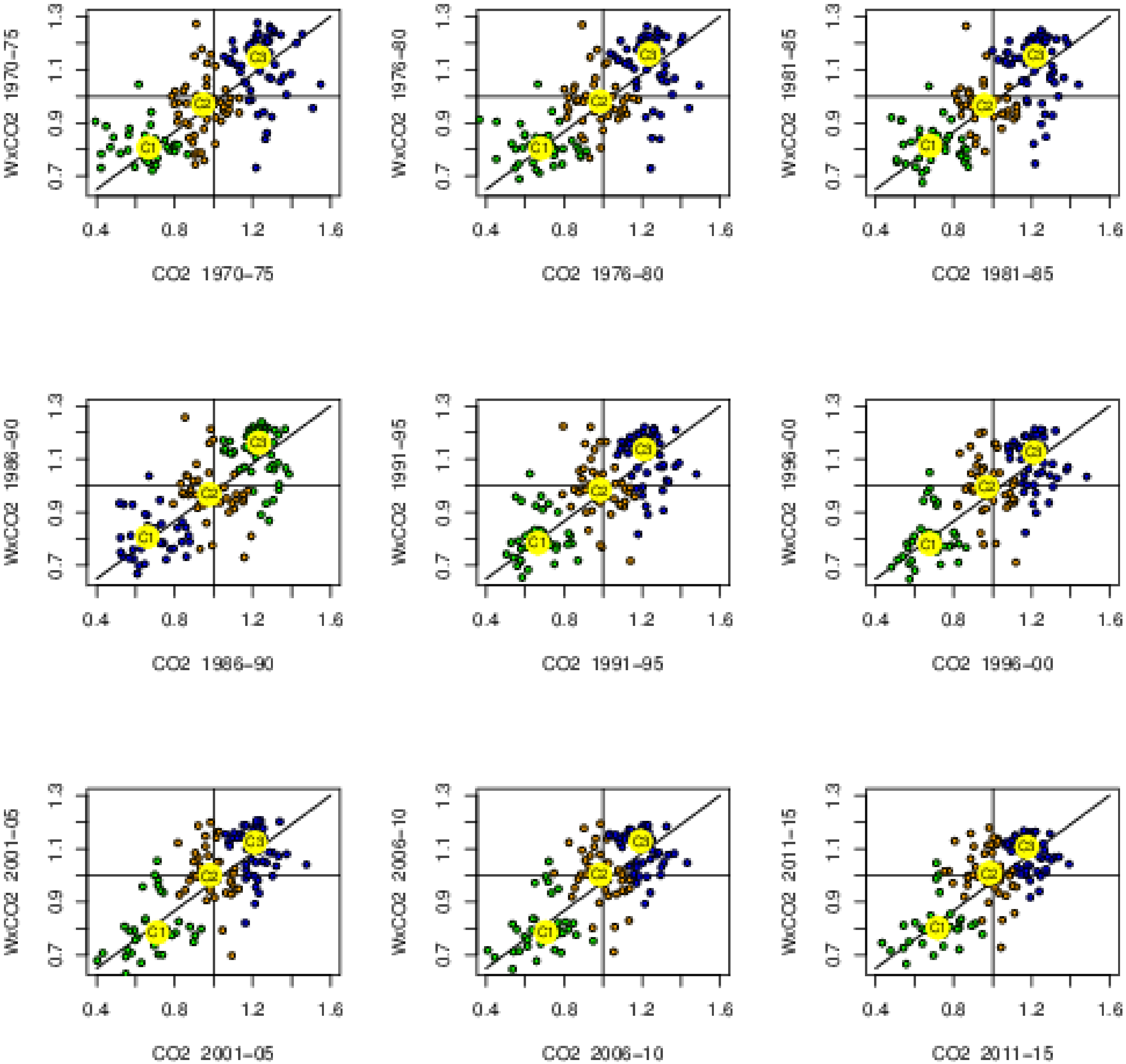


Figure 7: Local Directional Moran Scatter