Environmental Kuznets Curve for CO_2 Emissions in Middle-Income Countries: a Dynamic Spatial Panel Data Analysis

Hanen Ragoubi[1] | *University of Sousse, Sousse, Tunisia*
Zouheir Mighri[2] | *University of Jeddah, Jeddah, Saudi Arabia*

Abstract

This paper examines the carbon dioxide (CO_2) Environmental Kuznets curve (EKC) hypothesis of a balanced panel of 50 middle-income countries over the period 1996–2013 using a dynamic spatial panel data model with country and time-period fixed effects. Using a Bayesian comparison approach, we systematically searched for the most suitable spatial weights matrix describing the spatial arrangement of the countries in the sample. We found substantial spatial dependence effect in CO_2 emissions across the sample of middle-income countries, highlighting the influence these countries exert on their neighbors. Besides, the empirical results showed that the relationship between economic growth and CO_2 emissions shaped as an inverted-U trajectory. Furthermore, it has been found that trade openness and energy intensity are the main factors on slightly increasing CO_2 emissions, while the urbanization contributes to relative decrease in CO_2 emissions.

Keywords	***JEL code***
CO_2 emission, EKC hypothesis, dynamic spatial panel, Bayesian comparison, spillover effects	*C21, C23, P25, Q53, Q56*

INTRODUCTION

Over the three last decades, global warming, and particularly increasing temperatures, have a significant deep impact on economic productivity (Burke et al., 2015). Indeed, economic production has warmed the earth by releasing mass emissions of greenhouse gas in the atmosphere. In particular, the ever-increasing global emissions of CO_2 appear to be aggravating this issue. Accordingly, both the global environmental change and sustainable development become the critical challenge for human beings today (Roy Chowdhury and Moran, 2012). Exploring the potential relationship between economic growth

1 Research Laboratory for Economy, Management and Quantitative Finance (LaREMFiQ), University of Sousse, Rue Khalifa Karoui Sahloul 4, Sousse, Tunisia. E-mail: hragoubi2205@gmail.com.
2 College of Business, Department of Finance and Economics, University of Jeddah, Hamzah Ibn Al Qasim Street, Al Sharafeyah, Jeddah 23218, Kingdom of Saudi Arabia. Also Higher Institute of Finance and Taxation of Sousse, University of Sousse, Rue Khalifa Karoui Sahloul 4, Sousse, Tunisia. E-mail: zmighri@gmail.com.

and environmental degradation is also becoming a necessity in order to provide policy recommendations for taking a sustainable development trend in countries.

To explore the way of sustainable development, Grossman and Krueger (1991 and 1995) put forward the EKC theory to depict the relationship between economic growth and environmental degradation. Different econometric methodologies[3] have been used to investigate the CO_2 EKC hypothesis in different countries and regions. However, mixed empirical results are reported (Richmond and Kaufmann, 2006; Aldy, 2006; Galeotti et al., 2006; Kaika and Zervas, 2013a, 2013b, among others). Scholars have shown that the formulation of the EKC hypothesized multiple shaped EKC such as U, inverted-U, N, etc. For instance, Grossman and Krueger (1991) pointed out that economic growth can improve environmental quality after an economy has reached an adequate level of development. Furthermore, there were pieces of evidence that the testing results depended on the specific econometric models (Roy Chowdhury and Moran, 2012).

The mixed results further confirm that studies based on traditional cross-sectional panel data or time series techniques would provide incorrect inferences because of ignoring the spatial correlations dimension. Compared with traditional econometric methods, the spatial econometric techniques can be used to explore whether the local regional economic performances depend on the neighbors or not. While conventional econometric approaches have been used in most EKC studies, there is little evidence in the context of the nexus between economic growth and CO_2 emissions using spatial econometric techniques (Zhao et al., 2014; Kang et al., 2016; Meng et al., 2017; Meng and Huang, 2018; You and Lv, 2018).

As acknowledged by LeSage and Pace (2009), ignoring spatial dependence would lead to biased estimated parameters. Besides, Roy Chowdhury and Moran (2012) argued that spatial effects represent an important factor influencing the impact of economic growth on CO_2 emissions since several environmental problems, including CO_2 emissions, are inherently spatial. Furthermore, Anselin (2001) argues that spatial units (countries, states, counties, provinces, cities, etc.) can interact strongly with one another via channels such as trade, technological spillover, capital inflow, and common political, economic, and environmental policies. Recent research suggests that the closer the two countries are in terms of geographic distance, the more likely the economic activities and environmental degradation within each country will affect one another (You and Lv, 2018). In other words, economic growth and CO_2 emissions across countries are not independent. If such dependencies are not considered, some bias will be produced when estimating the EKC. As argued by Elhorst (2010a, 2010b), spatial econometric techniques provide ways to test and accommodate many forms of dependence among observations.

This study contributes to the empirical literature in several ways. First, it offers a more rigorous examination of the relationship between CO_2 emissions and economic growth for middle-income countries. The influence factors of CO_2 emissions are not only per capita real income but also other social, economic and industrial variables such as trade openness, urbanization, energy intensity and population which will be incorporated in the economic model to improve the accuracy of EKC fitting. Second, this paper uses the recently developed dynamic spatial panel models with controls for spatial and time-specific effects in order to capture the spatial interactions between explanatory variables and CO_2 emissions focusing on the middle-income countries. Specifically, this study seeks to explore the CO_2 emissions Kuznets curve in middle-income countries, and a comparative analysis between the non-spatial panel data model and the dynamic spatial panel data model is conducted to validate the spatial spillovers effects of variables in order to provide more rigorous references for policymakers. Third, using a Bayesian comparison approach developed by LeSage (2014, 2015), this study tests and compares simultaneously four frequently used dynamic spatial panel data models and twelve spatial weight matrices describing the mutual relationships among the middle-income countries, all within a common framework, which helps clarify the impact of neighboring countries on CO_2 emissions.

[3] A large strand of empirical literature is summarized in Table 1.

The remainder of this paper is organized as follows. Section 1 outlines the theoretical framework of the empirical model specification, the conventional spatial autocorrelation measures and the methodology of dynamic spatial panel data models. Section 2 provides a description of the data. Section 3 is devoted to the empirical estimation results and discussions. Final section concludes this paper and provides some policy suggestions.

Table 1 Summary of previous EKC studies on CO_2 emissions

Authors	Time period	Regions	Econometric methodology	Shaped EKC
Holtz-Eakin and Selden (1995)	1951–1986	130 countries	panel data	no EKC relationship
Carson et al. (1997)	1990	US states	cross-sectional data	inverted-U-shaped relationship
Roberts and Grimes (1997)	1962–1991	low-medium-high income countries	time series	inverted-U-shaped relationship for rich countries no EKC relationship for low/medium income countries
Lim (1997)	1980s onwards	South Korea	time series	no EKC relationship
Moomaw and Unruh (1997)	1950–1992	16 industrial OECD countries	panel data	N-shaped relationship
Schmalensee et al. (1998)	1950-1990	141 countries	panel data	inverted-U-shaped relationship
De Bruyn et al. (1998)	1960–1993 intervals	Netherlands, W. Germany, UK, USA	time series	no EKC relationship
Galeotti and Lanza (1999)	1970–1996	110 countries	panel data	inverted-U-shaped relationship
Agras and Chapman (1999)	various years	34 countries	panel data	no EKC relationship
Perrings and Ansuategi (2000)	1990	114 countries	panel data	no EKC relationship
Lindmark (2002)	1870–1997	Sweden	time series	inverted-U-shaped relationship
Friedl and Getzner (2003)	1960–1999	Austria	time series	N-shaped relationship
Cole (2004)	1980–1997	21 countries	panel data	inverted-U-shaped relationship
Dijkgraaf and Vollebergh (2005)	1960–1997	OECD countries	panel data	inverted-U-shaped relationship
Aldy (2005)	1960–1999	US states	panel data	inverted-U-shaped relationship in few states no EKC relationship (consumption model)
Azomahou et al. (2006)	1960–1996	100 countries	panel data	no EKC relationship
Richmond and Kaufmann (2006)	1973–1997	36 countries	panel data	no EKC relationship
Lantz and Feng (2006)	1970–2000	5 Canadian regions	panel data	no EKC relationship
Kunnas and Myllyntaous (2007)	1800–2003	Finland	time series	no EKC relationship
Coondoo and Dinda (2008)	1960–1990	88 countries	panel data	inverted-U-shaped relationship for Europe no EKC relationship for whole
Lee et al. (2009)	1960–2000	89 countries	panel data	N-shaped relationship for the whole panel inverted-U-shaped relationship in middle-income, American and European countries
Aslanidis and Iranzo (2009)	1971–1997	77 Non-OECD countries	panel data	no EKC relationship

Table 1 (continuation)

Authors	Time period	Regions	Econometric methodology	Shaped EKC
Dutt (2009)	1960–2002	124 countries	panel data	no EKC relationship (1960-1980) Inverted-U-shaped relationship (1984-2002)
Halicioglu (2009)	1960–2005	Turkey	time series	no EKC relationship
Jalil and Mahmud (2009)	1971–2005	China	time series	inverted-U-shaped relationship
Aslanidis and Iranzo (2009)	1971–1997	non-OECD countries	panel data	no EKC relationship
Narayan and Narayan (2010)	1980–2004	43 developing countries	panel data and time series	inverted-U-shaped relationship in 15 countries (time series) inverted-U-shaped relationship in Middle Eastern and South Asian countries (panel data)
Acaravci and Ozturk (2010)	1960–2005	19 European countries	time series	inverted-U-shaped relationship in 2 countries
Iwata et al. (2011)	1960–2003	28 countries (17 OECD, 11 non-OECD countries)	panel data	no EKC relationship
Wang et al. (2011)	1995–2007	28 China's provinces	panel data	U-shaped relationship
Jaunky (2011)	1980–2005	36 high-income countries	panel data	inverted-U-shaped relationship in 5 countries no EKC relationship for whole panel
Fosten et al. (2012)	1830–2003	United Kingdom	time series	inverted-U-shaped relationship
Esteve and Tamarit (2012)	1857–2007	Spain	time series	inverted-U-shaped relationship
Du et al. (2012)	1995–2009	29 China's provinces	panel data	no EKC relationship
Ahmed and Long (2012)	1971–2008	Pakistan	time series	inverted-U-shaped relationship
Saboori et al. (2012)	1980–2009	Malaysia	time series	inverted-U-shaped relationship
Saboori and Sulaiman (2013)	1980–2009	Malaysia	time series	no EKC relationship
Ozturk and Acaravci (2013)	1960–2007	Turkey	time series	inverted-U-shaped relationship
Burnett et al. (2013)	1970–2009	48 US states	spatial panel data	inverted-U-shaped relationship
Onafowora and Owoye (2014)	1970–2010	8 countries	time series	inverted-U-shaped relationship in two of the eight countries N-shaped relationship in six of the eight countries
Shahbaz et al. (2014a)	1971–2010	Tunisia	time series	inverted-U-shaped relationship
Farhani and Ozturk (2015)	1971–2012	Tunisia	time series	no EKC relationship
Apergis and Ozturk (2015)	1990–2011	14 Asian countries	panel data	inverted-U-shaped relationship
Yin et al. (2015)	1999–2011	China (29 provinces)	panel data	inverted-U-shaped relationship
Wang et al. (2016b)	1995–2011	30 China's provinces	spatial panel data	N-shaped relationship
Kang et al. (2016)	1997–2012	30 China's provinces	spatial panel data	inverted-N-shaped relationship
Li et al. (2016)	1996–2012	28 China's provinces	spatial panel data	inverted-U-shaped relationship
Wang and Liu (2017a)	1992–2013	341 China's cities	panel data and dynamic panel data	inverted-U-shaped relationship
Meng and Huang (2018)	1995–2012	331 China's cities	spatial panel data	no EKC relationship
You and Lv (2018)	1985–2013	83 developed and developing countries	spatial panel data	inverted-U-shaped relationship

Source: Created by the authors

1 THEORETICAL FRAMEWORK AND METHODOLOGY

1.1 EKC Hypothesis

Originally, EKC is an empirical hypothesis that characterizes an inversely U-shaped curve for the relationship between economic growth and environmental quality. Several indices of environmental quality degenerate with economic growth. As suggested by Grossman and Krueger (1995), the environment deterioration starts to decrease after reaching a threshold. Furthermore, Maddison (2006) pointed out that development may promote environmental quality as a result of economies of scale from pollution reduction, technological upgrade, industrial structure escalation, and public's demand for a clean environment. In this paper, the considered model for the EKC is a polynomial function type that is expressed as follows:

$$Y_{it} = \alpha_i + \beta_1 X_{it} + \beta_2 X_{it}^2 + \beta_3 Z_{it} + \varepsilon_{it} , \tag{1}$$

where Y stands for the indices of environmental degradation, while X refers to the economic growth level, usually measured by per capita Gross Domestic Product (GDP), and Z includes other influential factors for the environment. The polynomial function form of EKC offers to us an adequate tool to estimate the nonlinear relationship (if it exists) between economic growth and CO_2 emission.

1.2 STIRPAT Model

In this paper, we use the STIRPAT model (Dietz and Rosa, 1997; York et al., 2003) as our theoretical foundation to test the existence of an EKC for CO_2 emissions related to affluence. Ehrlich and Holdren (1971) first proposed the concept of IPAT (Influence, Population, Affluence, and Technology). The IPAT model relates environmental impact to population, affluence and technology. Nevertheless, this model is only an overly simplified function form and just indicates that the impact of human activities on the environment can fully be differentiated into population, affluence, and technology effects. Therefore, the IPAT model cannot estimate to what extent a specific factor affects the environment in such a framework, not to mention test any hypothesis. An additional limitation is that the IPAT model has been criticized as being primarily a mathematical equation which is not suitable for hypothesis testing, and also assuming a rigid proportionality between effects and factors.

To overcome these limitations, Dietz and Rosa (1997) proposed a stochastic version of IPAT, known as STIRPAT and later refined by York et al. (2003), expressed by the following equation:

$$I_{it} = \alpha_0 P_{it}^{\alpha_1} A_{it}^{\alpha_2} T_{it}^{\alpha_3} e_{it} , \tag{2}$$

where I denotes the environmental impact, P, A and T indicate human activities, i.e., respectively, population, affluence (per capita), and technological influences (per unit of economic activity). α_0, α_1, α_2 and α_3 are coefficients to be estimated and e denotes the random disturbance (the proportionality of IPAT model pre-assume $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 1$). The subscript i refers to the ith country and vary across observations.

The regression form of the STIRPAT model for estimation and hypothesis testing is obtained by logarithmic transformation of the variables in Formula (2). In this case, the coefficients α_1, α_2, and α_3 stand for the Ecological Elasticity (EE) which measures the sensitivity of environmental impacts to a change occurring in a driving force. It is defined as the proportion of change in environmental impacts due to its significant determinants. Using natural logarithms, the STRIPAT model can be converted to a convenient linear specification for panel estimation:

$$lnI_{it} = a_0 + \alpha_1 lnP_{it} + \alpha_2 lnA_{it} + \alpha_3 lnT_{it} + lne_{it} . \tag{3}$$

The above basic model analyses the impacts of population (P), economic development (A) and industrial structure (T) on the environmental impacts, but ignores other important factors influencing CO_2 emissions. According to the EKC hypothesis, CO_2 emissions is a function of par capita GDP and square of per capita GDP (Kasman and Duman, 2015; Kang et al., 2016; Meng and Huang, 2018; You and Lv, 2018, among others). Therefore, a quadratic or higher term of affluence can enter the STIRPAT specification. Besides, we further investigate the effects of additional factors on CO_2 emissions such as urbanization, energy intensity and trade openness (Martínez-Zarzoso et al., 2007; Pao and Tsai, 2011; Madlener and Sunak, 2011; Zhang et al.,2014; Al-Mulali et al., 2015; Kang et al., 2016; You and Lv, 2018; Lv and Xu, 2019, among others). Accordingly, we applied an augmented STIRPAT for our study purpose:

$$lnCO_{2it} = a_0 + \alpha_1 ln(POP_{it}) + \alpha_2 ln(RGDP_{it}) + \alpha_3 ln(RGDP_{it})^2 + \alpha_3 ln(TECH_{it}) + \alpha_4 ln(TRO_{it})$$
$$+ \alpha_5 ln(URBA_{it}) + \alpha_6 ln(EI_{it}) + \alpha_7 CV_{it} + \mu_i + \eta_t + \varepsilon_{it} , \qquad (4)$$

where CO_2 denotes per capita carbon dioxide emissions; *TRO* represents the trade openness; *POP* is the total population and measures the impact of demographic factors on CO_2 emissions; *RGDP* stands for per capita real GDP, which is seen as a proxy for economic factors; *URBA* denotes the urbanization level, which is typically associated with increased economic activity resulting in high energy consumption, and thus accelerating the emission of CO_2 (Martínez-Zarzoso and Maruotti, 2011; Adams and Klobodu, 2017); *TECH* is the technological improvement, measured by percentage of industrial activity with respect to total production, and represents a proxy for the level of environmentally damaging technology (Martínez-Zarzoso et al., 2007); *EI* refers to the energy intensity[4] per unit of GDP and can be considered as a proxy for energy consumption (Martínez-Zarzoso et al., 2007); μ_i is the individual fixed effect, which controls for all space-specific time-invariant variables that if omitted could potentially bias the coefficient estimates; η_t denotes the time period effects; ε is the standard error term; and CV_{it} stands for the potential control variables that could influence the CO_2 emissions.

In general, the estimation of the empirical model, i.e., Formula (4), tests the statistical significance of the coefficients α_2 and α_3. The following cases may occur (Dinda, 2004; Kaika and Zervas, 2013a):

i. If $\alpha_2 = \alpha_3 = 0$, then there is no relationship between economic growth and CO_2 emissions.
ii. If $\alpha_2 > 0$ and $\alpha_3 = 0$, then a monotonic increasing or linear relationship exists between economic growth and CO_2 emissions.
iii. If $\alpha_2 < 0$ and $\alpha_3 = 0$, then a monotonic decreasing or linear relationship exists between economic growth and CO_2 emissions.
iv. If $\alpha_2 > 0$ and $\alpha_3 < 0$, then an inverted-U-shaped relationship (EKC) exists between economic growth and CO_2 emissions.
v. If $\alpha_2 < 0$ and $\alpha_3 > 0$, then a U-shaped relationship exists between economic growth and CO_2 emissions.

Note that only the (iv) case indicates an EKC-relationship. Accordingly, the EKC is a specific form of the CO_2-income relationship. If the (iv) case holds, then the turning point is calculated as follows:

$$RGDP^* = exp(-(\alpha_2/2\alpha_3)) . \qquad (5)$$

[4] Energy intensity was measured as energy use divided by GDP at purchasing power parity (PPP) prices, where energy use refers to apparent consumption (production + imports – exports).

1.3 Spatial autocorrelationl

Spatial autocorrelation is a spatial data analysis method which is used to examine the degree of spatial dependence or autocorrelation in spatial data. It includes i) the global spatial autocorrelation which is used to estimate the overall degree of spatial dependence, and ii) the local indicators of spatial association (LISA) which is used to assess the impact of individual locations on the magnitude of the global statistic and to identify the locations and types of clusters. The spatial weights were created by rook contiguity rule and applied to describe the spatial relationships among countries. We explored the spatial distribution of per capita CO_2 emissions from 50 middle-income countries by calculating the Global Moran's I (Moran, 1950) and LISA (Anselin, 1995) using GeoDa software. The Global Moran's I statistic can be specified as follows:

$$I = \frac{n}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}} \cdot \frac{\sum_{j=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} (y_i - \bar{y})} \; ; i \neq j \, , \tag{6}$$

where $-1 \leq I \leq 1$; y_i and y_j are the values of the per capita CO_2 emissions of countries i and j, respectively; $\bar{y}$ is equal to the average of the per capita CO_2 emissions of all countries; w_{ij} is the element in row i column j of a spatial weights matrix and denotes the spatial weight between country i and country j; and n is the number of countries.

At a given level of statistical significance, $I > 0$ points to positive spatial autocorrelation, and the greater the value of I, the more obvious the spatial correlation. $I < 0$ refers to negative spatial autocorrelation, and the smaller the value of I, the greater the spatial difference. Otherwise, $I = 0$ points to a random spatial distribution. As argued by Anselin and Florax (1995), a significant positive Moran's I value indicates spatial clustering, while a significative negative Moran's I value indicates spatial dispersion across the sample of geographical units.

To evaluate the statistical significance of the Global Moran's I, both a z-score and p-value can be calculated. The z_I-score for the statistic I is computed as follows:

$$z_I = \frac{I - E(I)}{\sqrt{V(I)}} \longrightarrow N\,(0,1) \, , \tag{7}$$

where $E(I) = -1/(n-1)$; $V(I) = E(I^2) - E^2(I)$.

Alternatively, LISA is calculated as follows:

$$I_i = \frac{z_i}{\sum_{i=1}^{n} z_i^2} \times z_i^{\circ} \, , \tag{8}$$

where z_i denotes the observation for country i on per capita CO_2 emissions as a deviation from the mean, and z_i° is the spatial lag for location i, obtained as follows:

$$z_i^{\circ} = \sum_{j=1}^{n} w_{ij} z_j \, . \tag{9}$$

1.4 Dynamic spatial panel data models

A spatial econometric model is a linear regression model extended to include spatial interaction effects among the dependent variable, the explanatory variables, the error terms, or some combination thereof. Including all spatial lags yields a so-called general nesting spatial (GNS) model (Elhorst, 2014a, 2014b). When accounting for the dependent variable lagged one period, such a specification is known as a dynamic GNS model. The econometric counterpart of the dynamic GNS model reads, in vector form, as:

$$\boldsymbol{Y}_t = \tau \boldsymbol{Y}_{(t-1)} + \delta \boldsymbol{W}\boldsymbol{Y}_t + \eta \boldsymbol{W}\boldsymbol{Y}_{(t-1)} + X_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\,\boldsymbol{\theta} + \boldsymbol{\mu} + \lambda_t\,\boldsymbol{\iota}_N + \boldsymbol{\nu}_t\,, \tag{10}$$

$$\boldsymbol{\nu}_t = \lambda \boldsymbol{W}\boldsymbol{\nu}_t + \boldsymbol{\varepsilon}_t\,, \tag{11}$$

where $\boldsymbol{Y}_t$ is an $N \times 1$ vector consisting of one observation of the dependent variable for every spatial unit ($i = 1, \ldots, N$) in the sample at a particular point in time $t(t = 1, \ldots, T)$, which for this study is the CO_2 emissions ; $\boldsymbol{X}_t$ denotes an $N \times K$ matrix of exogenous or predetermined explanatory variables. Note that a vector or a matrix with subscript $t - 1$ stands for its serially lagged value, while a vector or a matrix premultiplied by $\boldsymbol{W}$ denotes its spatially lagged value. Moreover, the $N \times N$ matrix $\boldsymbol{W}$ denotes a non-negative matrix of known constants describing the spatial arrangement of the spatial units in the sample. It should be stressed that the diagonal elements of the matrix $\boldsymbol{W}$ are set to zero by assumption, since no spatial unit can be viewed as its own neighbor. Furthermore, the parameters τ, δ and η denote the response parameters of successively the dependent variable lagged in time, $\boldsymbol{Y}_{t-1}$, the dependent variable lagged in space, $\boldsymbol{W}\boldsymbol{Y}_t$, and the dependent variable lagged in both space and time, $\boldsymbol{W}\boldsymbol{Y}_{t-1}$. The variables $\boldsymbol{W}\boldsymbol{Y}_t$ and $\boldsymbol{W}\boldsymbol{Y}_{t-1}$ stand for contemporaneous and lagged endogenous interaction effects among the dependent variables. The symbols $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ stand for $K \times 1$ vectors of the response parameters of the exogenous explanatory variables. Furthermore, the error term specification consists of different components: the vector $\boldsymbol{\nu}_t$ that is assumed to be spatially correlated with autocorrelation coefficient λ; the $N \times 1$ vector $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ that consists of i.i.d. disturbance terms, which have zero mean and finite variance σ^2; the $N \times 1$ vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)'$ that contains spatial specific effects μ_i and is meant to control for all spatial-specific, time-invariant variables whose omission could bias the estimates in a typical cross-sectional study; and the time specific effects $\lambda_t(t = 1, \ldots, T)$, where $\boldsymbol{\iota}_N$ is a $N \times 1$ vector of ones meant to control for all time-specific, unit-invariant variables whose omission could bias the estimates in a typical time-series study.

It should be mentioned that spatial- and time period-specific effects can be treated as fixed or random effects. Otherwise, direct interpretation of the coefficients in the dynamic GNS model is not straightforward since they do not represent true partial derivatives (LeSage and Pace, 2009). Elhorst (2012, 2014a, 2014b) show that the matrix of (true) partial derivatives of the expected value of the dependent variable with respect to the kth independent variable for $i = 1, \ldots, N$ in year t for the long-term is given by the $N \times N$ matrix:

$$\left[\frac{\partial E(\boldsymbol{Y})}{\partial x_{1k}} \cdots \frac{\partial E(\boldsymbol{Y})}{\partial x_{Nk}}\right] = [(1-\tau)\boldsymbol{I} - (\delta + \eta)\boldsymbol{W}]^{-1}[\beta_k \boldsymbol{I}_N + \theta_k \boldsymbol{W}]\,, \tag{12}$$

whose diagonal elements represent long-term impacts on the dependent variable of unit 1 up to N if the kth explanatory variable in the own country changes, while its off-diagonal elements represent the long-term impacts on the dependent variable if the kth explanatory variable x_k in other countries changes. The average diagonal element of this matrix can be used as a summary indicator for the direct effect, whereas the average row sum of its off-diagonal elements represents a summary indicator of the spillover effect. Furthermore, these impacts are independent of t since the spatial weight matrix W is not time-varying, and error terms drop out due to the use of expectations.

As acknowledged by LeSage and Page (2009), the direct effect is defined as the average diagonal element of the full $N \times N$ matrix expression on the right-hand side of Formula (12); the indirect effect (i.e. country spillover effects) is the average row or column sum of the off-diagonal elements. Moreover, short-term direct and country spillover effects can be obtained by setting $\tau = \eta = 0$.

It should be stressed that the dynamic GNS model is problematic since is its parameters are not identified (Anselin et al., 2008; Elhorst, 2014a, 2014b). Indeed, the interaction effects among the dependent variable and the error terms cannot be distinguished formally, if the interaction effects among the explanatory variables are also included. Therefore, one of the two spatial interaction effects should be excluded. If the spatial interaction effects for the dependent variable are excluded ($\delta = \eta = 0$), the dynamic SDEM specification results, while the spatial multiplier matrix $[(1 - \tau)\boldsymbol{I} - (\delta + \eta)\boldsymbol{W}]^{-1}$ reduces to $1/(1 - \tau)\boldsymbol{I}$.

If the spatial interaction effects among the error terms is left aside ($\lambda = 0$), a dynamic spatial Durbin model (SDM) results. Although the SDM specification does not account for interaction effects among the error terms, which reduces the efficiency of the parameter estimates, it does not affect the consistency of the parameter estimates. Besides, it does not influence the direct or spillover effects derived from Formula (12).

As pointed out by Anselin et al. (2008), LeSage and Pace (2009), and Elhorst (2014a, 2014b), among others, an important difference between the SDEM and SDM specifications is that the country spillover effects in the first model are local, whereas in the second model they are global in nature. Local spillovers occur at other countries only if they are connected to each other. In other words, local spillovers occur when $\delta = 0$ and $\theta \neq 0$, and countries are connected. If two countries i and j are unconnected, such that $w_{ij} = 0$, a change in x_{ik} of country i cannot affect the dependent variable of country j, and vice versa. Global spillovers instead occur when $\delta \neq 0$ and $\theta = 0$, regardless of whether countries are connected, so a change to x_{ik} of country i due to the spatial multiplier matrix $(\boldsymbol{I} - \delta\boldsymbol{W})^{-1}$ gets transmitted to all other countries, even if the two countries are unconnected, i.e., $w_{ij} = 0$.

If CO_2 emissions at a local level can spread to other countries across the continent or around the world, even if they are not directly connected, then the SDM or SAR specifications make more sense, due to their ability to capture such global spillovers. If other countries are connected to each other, the SDEM specification may be more appropriate since it captures only local country spillovers. Otherwise, the choice between local and global spillovers depends on the specification of the spatial weight matrix W. It should be stressed that a sparse spatial weight matrix with only a limited number of non-zero elements, such as a binary contiguity matrix, is more likely to occur in combination with a global spillover model ($\delta \neq 0$, $\theta = 0$), while a dense spatial weight matrix in which many off-diagonal elements are non-zero (e.g. inverse distance matrix) is more likely in combination with a local spillover model ($\delta = 0$, $\theta \neq 0$). Therefore, the choice of spatial model and spatial weight matrix might be improved if they take place within a common framework.

In this paper, we employ a Bayesian comparison approach (LeSage, 2014; LeSage, 2015) in order to choose between a global spillover model, i.e., SDM, and a local spillover model, i.e., SDEM, as well as to choose between different potential specifications of the spatial weight matrix W. It should be noted that this approach allows determining the Bayesian posterior model probabilities of the SDM and SDEM specifications given a particular spatial weight matrix, as well as the Bayesian posterior model probabilities of different spatial weight matrices given a particular spatial panel model specification. These probabilities are based on the log marginal likelihood of a spatial panel 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 spatial panel model or of one spatial weight matrix W is higher than that of another model or another W, the Bayesian posterior model probability is also higher. It should be stressed that the classical LR, Wald and/or LM statistics compare the performance of one spatial model against another spatial model based on specific parameter estimates within the parameter space. However, the main strength of the Bayesian comparison approach is that it compares the performance of one spatial model against another spatial model on their whole parameter space (LeSage, 2014; LeSage, 2015). Furthermore, statistical inferences drawn on the log marginal likelihood function values for the SDM and SDEM models are further justified since they have the same set of explanatory variables, i.e., X_t and WX_t, and are based on the same uniform prior for δ and λ. This prior takes the following form:

$$p(\delta) = p(\lambda) = 1/D \,, \tag{13}$$

where:

$$D = 1/\omega_{max} - 1/\omega_{min} \,, \tag{14}$$

and ω_{max} and ω_{min} denote respectively the largest and the smallest (negative) eigenvalue of the spatial weight matrix W. Note that this prior requires no subjective information on the part of the practitioner since it relies on the parameter space $(1/\omega_{min}, 1/\omega_{max})$ on which δ and λ are defined, where $\omega_{max} = 1$ if W is row normalized. Finally, and depending on the outcomes of the Bayesian comparison approach, either the SDM or the SDEM model is estimated using maximum likelihood estimation (MLE). Then, the estimation results could serve to test the following null hypotheses:

$$H_0 : \theta = 0 \text{ and } = 0 \,, \tag{15}$$

$$H_0 : \theta + \delta\beta = 0 \text{ and } + \delta\tau = 0 \,. \tag{16}$$

That is, it is possible to test whether the dynamic SDM might be reduced to a dynamic SAR model or dynamic SEM. Both tests follow a chi-squared distribution with $K + 1$ degrees of freedom (i.e., the number of spatially lagged explanatory variables and the spatially lagged dependent variable) and take the form of a Wald test, since the simplified models have not been estimated.

2 DATA AND VARIABLES

In this paper, we use a balanced panel sample of 50 countries[5] over the period 1996–2013. In contrast to high income countries, time series data on energy use in many middle-income countries are very limited. Therefore, we limited our sample to 50 middle-income countries due to the availability of reliable data. Furthermore, the beginning of the sample period is motivated by the fact that the transition of several middle-income countries from socialism to capitalism has likely led to a structural break in environmental policy in general. The dependent variable is CO_2 emissions (metric tons of per capita carbon dioxide emissions), which are considered as the primary greenhouse gas responsible for global warming and proxies for overall environmental pollution in a country.

In our empirical analysis, affluence is the natural log of per capita real GDP (real GDP divided by population at the end of the year), population is the natural log of total population in a country, technology is the natural log of the weight of the industry in economic activity (the proportion of the added value of industry to GDP), energy intensity is the natural log of total energy use per dollar of GDP (kg of oil equivalent per capita), trade openness is the natural log of trade openness (exports plus imports as percent of GDP) and urbanization is the natural log of urbanization (% urban population in the total population).

All data except per capita real GDP are obtained from World Development Indicators (WDI) online database. The series of real GDP (at constant 2011 national prices in millions 2011 US$) is obtained

5 Table A1 in the Appendix provides the list of sample countries.

Table 2 Summary statistics

	ln CO_2	*ln POP*	*ln RGDP*	*ln EI*	*ln TECH*	*ln TRO*	*ln URBA*
Mean	0.5759	16.8724	11.9313	6.7971	3.4205	–0.4238	3.9542
Median	0.5483	16.8309	11.7698	6.6574	3.3771	–0.3901	4.0414
Maximum	2.7502	20.9690	15.7035	8.5501	4.3492	0.6021	4.4900
Minimum	–1.9926	13.9235	9.1564	4.8820	0.9909	–1.9393	2.8726
Std. Dev.	1.0152	1.4326	1.5232	0.7266	0.3087	0.5127	0.3680
Skewness	–0.2151	0.2950	0.2372	0.2305	–0.4442	–0.4432	–1.0374
Kurtosis	2.5662	2.6978	2.1477	2.5733	9.8599	2.9691	3.6196
Observations	900	900	900	900	900	900	900

Source: Own estimates

Table 3 Correlation coefficient matrix and VIF test

	VIF	*ln(CO_2)*	*ln(RGDP)*	*ln(TRO)*	*ln(URBA)*	*ln(POP)*	*ln(TECH)*	*ln(EI)*
ln(CO_2)		1.0000						
ln(RGDP)	1.65	0.2650***	1.0000					
		(0.0000)						
ln(TRO)	1.91	0.1807***	–0.4337***	1.0000				
		(0.0000)	(0.0000)					
ln(URBA)	1.92	0.6076***	–0.0255	0.1150***	1.0000			
		(0.0000)	(0.4441)	(0.0005)				
ln(POP)	1.48	0.0678**	0.7499***	–0.5907***	–0.2078***	1.0000		
		(0.0420)	(0.0000)	(0.0000)	(0.0000)			
ln(TECH)	2.15	0.3000***	–0.0110	0.2331***	0.2682***	–0.0040	1.0000	
		(0.0000)	(0.7411)	(0.0000)	(0.0000)	(0.9053)		
ln(EI)	1.94	0.9263***	0.2049***	0.1745***	0.5841***	0.0100	0.2767***	1.0000
		(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.7656)	(0.0000)	

Notes: * denotes p<0.1. ** denotes p<0.05. *** denotes p<0.01.
Source: Own estimates

from the Penn World Table version 9.1.[6] Table 2 summarizes the descriptive statistics of the above-mentioned variables.

The correlation coefficients of the variables are displayed in Table 3. CO_2 emissions have a relatively low and significant correlation with per capita real GDP and trade openness. While the correlation between CO_2 emissions and urbanization is moderate, it is rather strong and significant between CO_2 emissions and energy intensity. However, the correlation between CO_2 emissions and population is weak and statistically significant. To test for multi-collinearity issue, a variance inflation factor (VIF) test is used over a data range of 1.48–2.15, with a mean value of 1.842. As shown in Table 3, the VIF values are all less than the cut-off value of 10, indicating that there is no multi-collinearity.

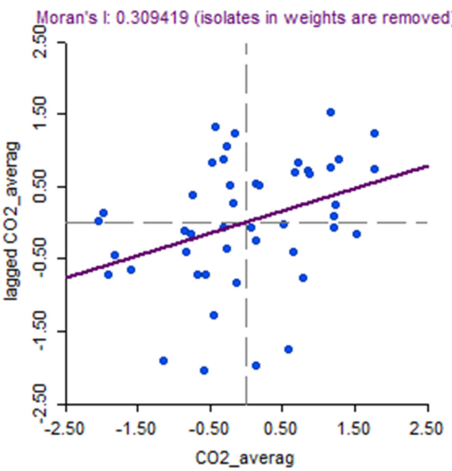
6 *<https://www.rug.nl/ggdc/productivity/pwt>.*

3 EMPIRICAL RESULTS AND DISCUSSIONS

3.1 Exploratory spatial data analysis

Following Abreu et al. (2005), among others, we examined the spatial dependence and spatial heterogeneity in our dataset using the exploratory spatial data analysis (ESDA) approach. To further test whether spatial dependence exists or not, we computed the global Moran's I to identify spatial dependence among the observations, where a significant positive Moran's I value indicates spatial clustering and a negative Moran's I value with statistical significance indicates spatial dispersion across the sample countries (Anselin and Florax, 1995; Anselin, 2006). Furthermore, global Moran's I is a measure of the geographical concentration of a distribution. Generally, the larger the global Moran's I index, the more significant the spatial dependence among countries. A trend of rapid spatial autocorrelation can be clearly seen in Figure 1.

Figure 1 Moran's I Scatter Plot for country-level CO_2 emissions in middle income countries, 1996–2013



Source: Own construction

The results of the global spatial autocorrelation for the CO_2 variable by using global Moran's I statistic are summarized in Table 4. Using both the z test and its corresponding p value, we test the statistical significance of the Moran's I values. As shown in Table 4, the Moran's I index values are positive and statistically significant at the 5% level or better. This means that air pollution in middle-income countries exhibits significant positive spatial autocorrelation, which ranges from 0.2627 to 0.3875. Note that the high positive values signal the occurrence of similar attribute values over space, and hence spatial clustering. This means that CO_2 emissions in middle income countries are spatially autocorrelated between 1996 and 2013. They also appear to be less spatially clustered in 2013 than in 1996.

Table 4 Statistical tests of global Moran's *I* of CO_2 emissions in middle-income countries

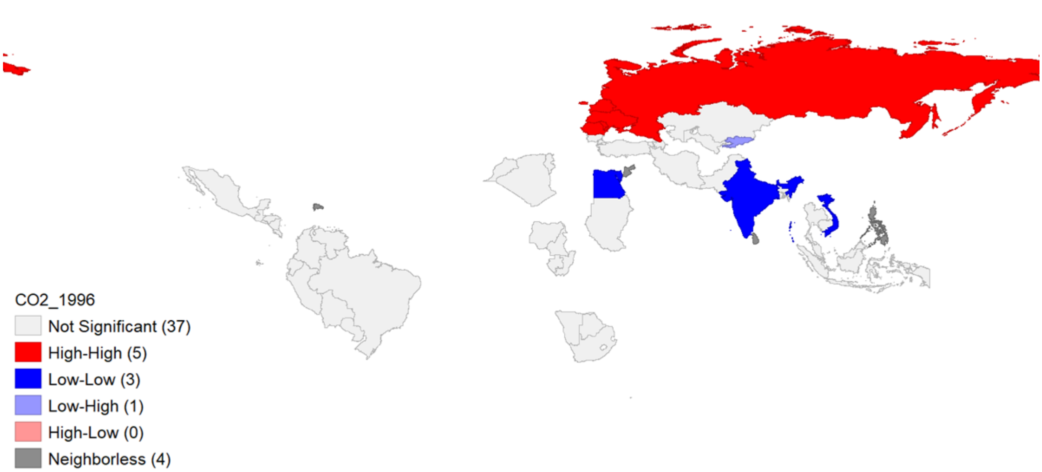
Year	Moran's *I*			Year	Moran's *I*		
	Statistic	*Z* score	*p*-value		Statistic	*Z* score	*p*-value
1996	0.3875***	2.7119	0.0090	2005	0.2871**	2.1404	0.0220
1997	0.3594**	2.5717	0.0120	2006	0.3087**	2.2915	0.0190
1998	0.3299**	2.3633	0.0160	2007	0.3106**	2.2977	0.0190
1999	0.2933**	2.1327	0.0220	2008	0.3375**	2.4963	0.0160
2000	0.2627**	1.9544	0.0280	2009	0.2929**	2.1792	0.0200
2001	0.2715**	2.0171	0.0230	2010	0.2956**	2.2065	0.0220
2002	0.2725**	2.0214	0.0280	2011	0.3166**	2.3489	0.0160
2003	0.2670**	1.9838	0.0260	2012	0.3396**	2.5056	0.0140
2004	0.2825**	2.0987	0.0250	2013	0.3159**	2.3438	0.0160
Average	0.3094**	2.2721	0.0180				

Notes: * denotes p<0.1. ** denotes p<0.05. *** denotes p<0.01. The null hypothesis is no global spatial autocorrelation.
Source: Own estimates

In a second step, we turn to spatiotemporal patterns of country level CO_2 emissions. To visually explore the spatial dependence of the middle-income countries' CO_2 emissions, we undertook a local LISA analysis with the aim of identifying local spatial autocorrelations. The results of the LISA allowed us to identify a detailed local pattern of spatial clustering in relation to changes in per capita CO_2 emission levels. The resulting LISA cluster maps of the countries for which the local Moran's I statistics are statistically significant at the 5% level are displayed in Figures 2, 3 and 4. These figures reveal characteristics of significant local spatial autocorrelation in the distribution of initial CO_2 level in 1996, CO_2 level in 2013 and the average annual CO_2 level over the study period. Spatially, countries with high levels of per capita CO_2 emissions are clustered with neighboring countries that have similar values. Besides, countries with low values of per capita CO_2 emissions clustered with neighboring countries with similar values. The red color denotes the High-High (H-H) clusters (i.e., high values surrounded by high values), while the blue represents Low-Low (L-L) clusters (i.e., low values surrounded by low values). Note that H-H and L-L clusters are the main types of spatial distribution. Furthermore, the pink areas indicate H-L associations and the blue-gray areas denote Low-High (L-H) correlations (i.e., low values surrounded by high values). The gray clusters represent countries that are not associated in a spatially significant manner.

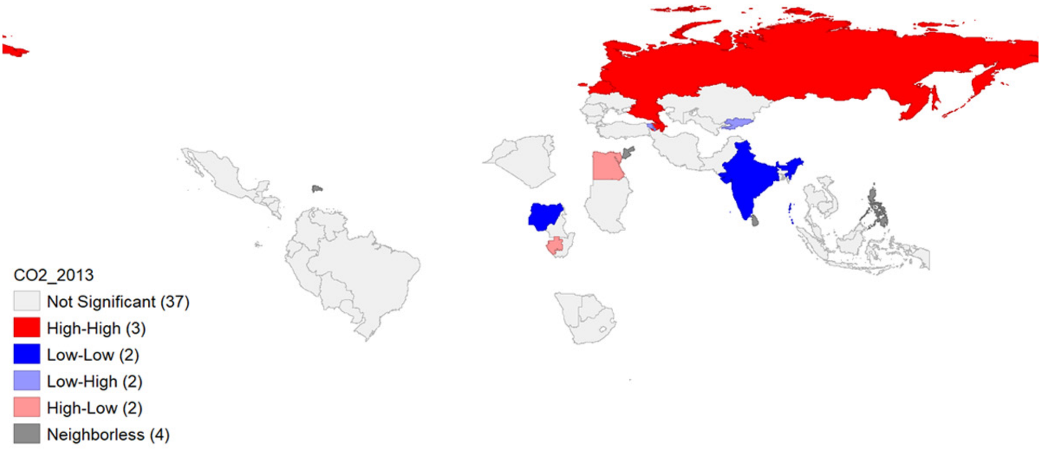
The number and the distribution of each cluster of countries also display regional dynamic characteristics. For instance, in 1996, the numbers of countries belonging to H-H and L-L cluster were 5 and 3 respectively, accounting for 16% of the sample of middle-income countries. This phenomenon is consistent with the situation revealed by a relatively large global Moran's I (0.3875). Correspondingly, only 2% of all countries conformed to the remaining High-Low (H-L) and L-H classifications. These results indicate the existence of a significant dual structure in the spatial distribution of country's per capita CO_2 emissions in 1996. However, by 2013, the number of H-H and L-L countries had decreased by 3 and 2, respectively, indicating that the spatial extent of dependence of per capita CO_2 emissions had weakened markedly between 1996 and 2013. The corresponding global Moran's I index also decreased (0.3159). These results imply that, for geographic data, it is almost inevitable that "close things are more related than distant things," a phenomenon that can be described in terms of "spatial dependence." In addition, the computed findings confirm our previous analysis of spatial dependence in per capita CO_2

Figure 2 Local Moran Scatter Plot map for *ln*(CO_2) in 1996



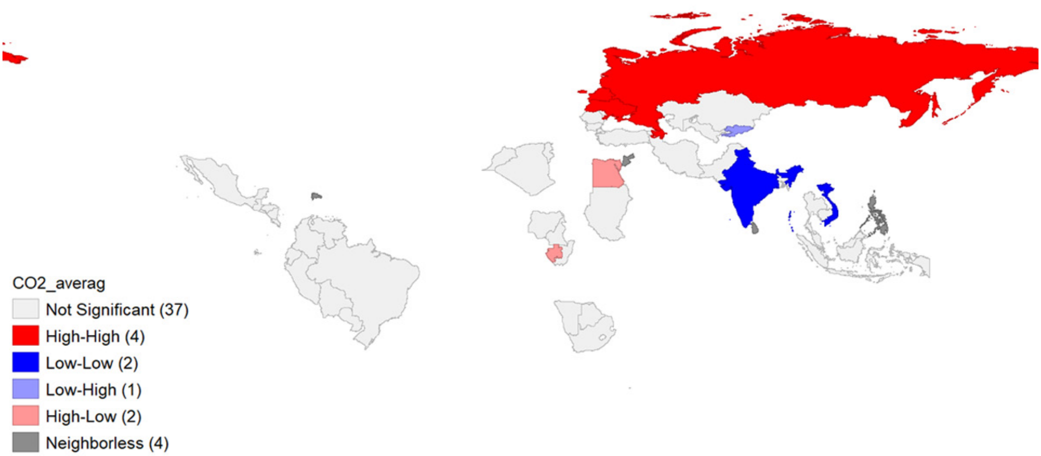
Source: Own construction

Figure 3 Local Moran Scatter Plot map for *ln*(CO_2) in 2013



Source: Own construction

Figure 4 Local Moran Scatter Plot map for *ln*(CO_2) over 1996 to 2013



Source: Own construction

emissions at the country level. Note that if such dependence is ignored, standard econometric models risk being biased in ways that conceal the impact of the determinants they purport to study – in our case, changes in per capita CO_2 emissions in middle income countries. Therefore, we empirically test whether the spatial panel econometrics models are better than conventional econometrics and chose the appropriate model to analyze the impact factors of per capita CO_2 emissions in middle income countries.

3.2 Spatial econometric regression results

To decide which type of model (spatial vs. non-spatial) best fits the data, we begin our investigation by testing several different model specifications. This testing procedure is a mixture of a specific-to-general approach and general-to-specific approach (Elhorst, 2012). Note that the procedure begins by testing the non-spatial panel model against the spatial lag and spatial error models. If the non-spatial panel models are rejected, then the spatial Durbin model (SDM) is tested to determine if it can be simplified to either the spatial lag or spatial error model. It should be stressed that this step seeks corroborating evidence from the first step.

Table 5 reports the estimation results for the non-spatial panel data models: pooled OLS only (no fixed or time-period effects), spatial fixed effects only (no time-period effects), time-period fixed effects only (no fixed effects) and both spatial fixed effects and time-period fixed effects, respectively.

To investigate the null hypothesis that the spatial fixed effects and time-period effects are jointly insignificant, we performed a likelihood ratio (LR) test. The null hypothesis that the spatial fixed effects are jointly insignificant is rejected at the 1% significance level (1 806.2527; 50 degrees of freedom; $P = 0.0000<0.01$). Likewise, the null hypothesis that the time-period fixed effects are jointly insignificant is rejected at the 1% significance level (41.7798; 18 degrees of freedom; $P = 0.0012<0.01$). These findings justify the extension of the model with fixed effects and time-period effects.

It should be stressed that if the country-level fixed effects term is correlated with the explanatory variables, but it is not controlled for within the model, then ordinary least squares (OLS) estimates will result in omitted variable bias (OVB). The pooled OLS estimates (column 2 in Table 5) for all the coefficients in the model are all highly statistically significant ($p < 0.01$), except for TECH variable, which arguably results from the OVB. Given the joint significance of the fixed and time-period effects from the LR test, we focus on the estimation results in column 5 in Table 5.

Table 5 Estimation results without spatial interaction effects

	Pooled OLS	Spatial fixed effects	Time-period fixed effects	Spatial and time-period fixed effects
lnRGDP	0.7598***	0.4199***	0.7857***	0.5241***
	(0.0000)	(0.0023)	(0.0000)	(0.0001)
lnRGDP^2	–0.0293***	–0.0095	–0.0300***	–0.0060
	(0.0000)	(0.1049)	(0.0000)	(0.2981)
lnTRO	0.2147***	0.1406***	0.2515***	0.1486***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
lnURBAN	0.4014***	0.1585	0.4256***	0.2485**
	(0.0000)	(0.1811)	(0.0000)	(0.0339)
lnPOP	0.0718***	–0.1250	0.0752***	0.0848
	(0.0000)	(0.1351)	(0.0000)	(0.3394)
lnTECH	0.0577	–0.0215	0.0530	–0.0417
	(0.1559)	(0.5124)	(0.1894)	(0.2093)
lnEI	1.1189***	0.6014***	1.1120***	0.6357***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Intercept	–14.7637***	-	-	-
	(0.0000)	-	-	-

Table 5 (continuation)

	Pooled OLS	Spatial fixed effects	Time-period fixed effects	Spatial and time-period fixed effects
R^2	0.8871	0.5120	0.8888	0.3327
$\bar{R}^2$	0.8862	0.5087	0.8881	0.3283
σ^2	0.1173	0.0161	0.1142	0.0154
FE R^2		0.9845	0.8899	0.9852
Log Likelihood	−308.6256	584.9733	−297.2632	605.8631
LM spatial lag	64.6114***	5.6159**	70.5684***	3.3388*
	(0.0000)	(0.0180)	(0.0000)	(0.0680)
Robust LM spatial lag	6.1237**	14.5566***	10.7820***	60.3123***
	(0.0130)	(0.0000)	(0.0010)	(0.0000)
LM spatial error	153.7638***	0.6179	139.7624***	4.2464**
	(0.0000)	(0.4320)	(0.0000)	(0.0390)
Robust LM spatial error	95.2760***	9.5585***	79.9760***	61.2198***
	(0.0000)	(0.0020)	(0.0000)	(0.0000)

Notes: All variables are in natural logarithms. Numbers in the parentheses represent P values. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.
Source: Own estimates

It should be mentioned that all the non-spatial panel data models may suffer from misspecification if spatial dependence exists within the data. To test for the presence of spatial dependence, we begin by conducting the classical Lagrange Multiplier (LM) tests and their robustness to examine whether non-spatial panel data models ignore the spatial interaction effects of data or not (Anselin et al., 2008; Burridge, 1980). These tests' results are presented in the bottom part in Table 5. For the classical LM test (labeled "LM spatial lag"), the hypothesis of no spatially lagged dependent variable is strongly rejected at the 5% significance level or better for each of the specifications. In addition, and for the classical LM test (labeled "LM spatial error"), the hypothesis of no spatially autocorrelated error term is rejected for each of the specifications except for spatial fixed effects model (although the hypothesis of no spatially lagged dependent variable is rejected at the 1% significance level with this specification). Regarding the results of their robustness tests (Debarsy and Ertur, 2010), both hypotheses are rejected at the 5% significance level or better for each of the specifications. These findings imply the existence of spatial dependence among the panel data, which is consistent with the results of Moran's I index (see Table 4). Besides, they imply that a model specification with a spatially lagged dependent variable may be favored over a non-spatial panel model since we find consistent rejection of the hypothesis of no spatially lagged dependence. However, if the robust LM tests reject a non-spatial panel data model in favour of the SAR model or SEM model, one of these models must be carefully endorsed.

To further test which spatial panel data model specification is more appropriate, LeSage and Pace (2009), and Elhorst (2014b) recommend estimating the SDM, and then conducting both LR and Wald tests to verify whether it can be simplified to the SAR model or to the SEM (see also Burridge, 1981).

In this paper, we take a broader view and apply a Bayesian comparison approach. First, the Bayesian posterior model probabilities of the SDM and SDEM specifications are calculated, as well as the simpler SAR and SEM specifications, to identify which model specification best describes the data. Second, this analysis is repeated for several specifications of the neighbourhood matrices, to find the specification of W that best describes the data.

For this empirical study, we use the following principles to construct twelve spatial weight matrices:

i. Sharing a common land or maritime border implies the first-order binary contiguity matrix, W_1. Maritime borders are based on the United Nations Convention on the Law of the Sea and additional sources further explaining this convention.

ii. The influence of a country might go beyond its immediate neighbors, as implied by the inverse distance matrix and the different cut-off points. Hence, we also consider a second order binary contiguity matrix, $W_2 = W_1 \times W_1$.

iii. A country may respond to the threat of even more distant countries, which is also the main reason that elements of the weight matrix within a certain radius of a country are not always set to 0. Therefore, we include a third-order binary contiguity matrix, $W_3 = W_2 \times W_1$.

iv. Except for the matrix based on the common border countries, the spatial weight matrix could be based on the calculation of distances using the spherical distance between geographic centroids of the countries. Therefore, we create a distance based spatial weight matrix, labeled as W_4, using latitude and longitude coordinates and the Great Circle distance formula.[7]

v. Inverse distance matrix based on the geographical distance between the centroids of every pair of countries. This matrix is labeled as W_5.

vi. k-nearest neighbours matrix for k = 5, 6, 7, 8, 9, 10 and k = 20: it is a binary matrix of the k-nearest neighbour, where the weight $w_{ij} = 1$ if the country j is within the k-nearest neighbour of the country i and $w_{ij} = 1$ if otherwise. Therefore, we create seven additional spatial weight matrices, which are labeled as W_6 for $k = 5$, W_7 for $k = 6$, W_8 for $k = 7$, W_9 for $k = 8$, W_{10} for $k = 9$, W_{11} for $k = 10$, and W_{12} for $k = 20$.

Finally, all the matrices are row normalized, which is standard in spatial econometrics literature when the elements of W have a binary (0/1) character.

Table 6 Simultaneous Bayesian comparison of dynamic spatial panel data model specifications and spatial weight matrices

W matrix	Statistics	SAR	SDM	SEM	SDEM
	Log marginal	538.7932	544.9566	541.0575	545.3341
W_1	Model probabilities	0.0008	0.4031	0.0082	0.5879
	Posterior model probabilities	0.0000	0.0001	0.0000	0.0002
	Log marginal	539.0006	549.7389	538.3043	549.4675
W_2	Model probabilities	0.0000	0.5674	0.0000	0.4326
	Posterior model probabilities	0.0000	0.1690	0.0000	0.0129
	Log marginal	538.8849	551.5624	539.2391	550.9163
W_3	Model probabilities	0.0000	0.6561	0.0000	0.3438
	Posterior model probabilities	0.0000	0.1695	0.0000	0.0550
	Log marginal	538.6144	539.9027	539.2829	539.8787
W_4	Model probabilities	0.0988	0.3584	0.1929	0.3499
	Posterior model probabilities	0.0000	0.0000	0.0000	0.0000

[7] Formally, the spherical distance (in kilometers) between the centroids of two countries is defined as follows: $d_{ij} = 6\,366.2 \times Arccos\{\{\cos|Y_i - Y_j| \times \cos X_i \times \cos X_j\}+\{\sin X_i \times \sin X_j\}\}$. X_i denotes the latitude of the centroid of country i, while Y_i is the longitude of the centroid of country i.

Table 6 (continuation)

W matrix	Statistics	SAR	SDM	SEM	SDEM
W_5	Log marginal	–861.2964	–2 079.6651	–1 280.7416	–2 277.1929
	Model probabilities	1.0000	0.0000	0.0000	0.0000
	Posterior model probabilities	0.0000	0.0000	0.0000	0.0000
W_6	Log marginal	538.7246	549.0644	538.7472	548.8082
	Model probabilities	0.0000	0.5637	0.0000	0.4363
	Posterior model probabilities	0.0086	0.0086	0.0000	0.0067
W_7	Log marginal	539.3243	549.1309	538.6578	548.9059
	Model probabilities	0.0000	0.5560	0.0000	0.4440
	Posterior model probabilities	0.0000	0.0092	0.0000	0.0074
W_8	Log marginal	540.2980	548.3807	538.6815	547.7051
	Model probabilities	0.0002	0.6626	0.0000	0.3372
	Posterior model probabilities	0.0000	0.0044	0.0000	0.0022
W_9	Log marginal	540.7850	542.3514	538.6779	542.1279
	Model probabilities	0.1027	0.4917	0.0125	0.3932
	Posterior model probabilities	0.0000	0.0000	0.0000	0.0000
W_{10}	Log marginal	539.9414	548.1252	538.6742	547.2536
	Model probabilities	0.0002	0.7049	0.0001	0.2949
	Posterior model probabilities	0.0000	0.0034	0.0000	0.0014
W_{11}	Log marginal	539.6121	543.9585	538.6546	543.6084
	Model probabilities	0.0075	0.5806	0.0029	0.4090
	Posterior model probabilities	0.0000	0.0001	0.0000	0.0000
W_{12}	Log marginal	538.8998	553.1569	540.5124	552.4275
	Model probabilities	0.0000	0.6747	0.0000	0.3253
	Posterior model probabilities	0.0000	0.1571	0.0000	0.2493

Notes: The highest posterior model probability in each row is highlighted in italics and the probabilities in each block sum to 1.
Source: Own estimates, based on LeSage (2014, 2015)

The results displayed in Table 6 show that both the dynamic SAR and SEM models are generally outperformed by either the dynamic SDM or dynamic SDEM specifications. In terms of the log marginal likelihood value, the worst-performing spatial neighbourhood matrix is the inverse distance matrix (W_5). This matrix corroborates the point that decomposing market potential variables into their underlying components and considering the spatially lagged values of these components creates a much greater degree of empirical flexibility. If the neighbourhood matrix is specified as a p-order binary contiguity matrices for $p = 2, 3$, as either a distance neighbourhood matrix, or as k-nearest neighbours matrices for $k = 5, 6, 7, 8, 9, 10, 20$, then the Bayesian posterior model probabilities point to the dynamic SDM specification. Conversely, if the neighbourhood matrix is specified as a first-order binary contiguity matrix, the Bayesian posterior model probabilities point to the dynamic SDEM specification. Alternatively, if neighbourhood matrix based on the inverse distance is adopted, the Bayesian posterior model probabilities provide further evidence in favour of the dynamic SAR specification.

Table 6 also contains the Bayesian posterior model probabilities of the different spatial models (SAR, SDM, SEM, SDEM), in combination with the twelve proposed spatial weight matrices. These probabilities are calculated for dynamic versions of the spatial panel data model specifications. With these probabilities, we can simultaneously identify the most likely spatial econometric model and the most likely spatial weight matrix. Note that the probabilities are based on the log-marginal likelihood obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. Furthermore, they are normalized such that the probabilities of all 48 combinations sum to 1. Following LeSage (2014, 2015), this normalization is based on the (non-linear) property that the Bayesian posterior model probability increases if the log-marginal likelihood value of one model or one W exceeds that of another model or W.

The results in Table 6 show that by considering the log-marginal values and Bayesian posterior model probabilities of the different specifications of the neighbourhood matrix, it is to be noted that the third-order binary contiguity matrix, i.e., W_3, and the SDM specification achieve the best performance of all 48 combinations, in line with the initial robust LM test statistics for the nonspatial panel data model, which pointed to a SAR rather than a SEM. Accordingly, spatially lagged explanatory variables (WX) are necessary and should be included in the empirical model.

Furthermore, we decided to estimate the dynamic SDM specification using the bias-corrected maximum likelihood (ML) estimator developed by Elhorst (2010a, 2010b), and Lee and Yu (2010a).[8] Note that the results without the bias correction are almost identical.[9] Nevertheless, since the dynamic SDM model produces global country spillover effects, it is more likely to occur in combination with a sparse spatial weight matrix. Therefore, we determined the average number of neighbors of each country in the sample based on these two spatial weight matrices. It equals 6.16 for the W_2 matrix, 9.46 for the W_3 matrix, and 6.00 for the W_4 matrix. Alternatively, the average number of adjacent neighbors based solely on land or maritime borders, i.e., W_1, is 3.080. Based on the principle of sparsity, the W_3 matrix thus seems to offer a better choice than W_1 , W_2 and W$_4$ matrices.

The estimation results of the dynamic SDM with fixed and time-period effects specification, based on the W_3 matrix, are reported in Table 7. Then, the results could serve to test whether the dynamic SDM might be simplified to a dynamic SAR model or to a dynamic SEM. The empirical findings reject both hypotheses and show that the dynamic SDM is preferred over the dynamic SAR model or the dynamic SEM. Otherwise, a necessary and sufficient condition for stationarity (stability), i.e., $\tau + \delta + \eta = 0.7923 < 1$, is also satisfied. This result is confirmed by the Wald test, across which the null hypothesis $\tau + \delta + \eta = 1$ is strongly rejected at the 1% level of significance.

3.3 Analysis of estimation results

Since the diagnostic results suggest that the dynamic SDM with spatial and time-period fixed effects in Table 6 is the best fitting, we will limit the interpretation of coefficient estimates on it. It should be mentioned that our results are in line with some of the results of previous empirical studies. As shown in Table 7, the CO_2 emissions strongly depend on their value in the previous year, or internal habit persistence (Korniotis, 2010); its coefficient amounts to 0.1595 and is highly significant at the 1% level.

[8] This bias correction is necessary since the dependent variables lagged in time and in both space and time on the right-hand side of Formula (10) are correlated with the spatial fixed effects , which is the spatial counterpart of the Nickell bias, as shown by Yu et al. (2008), and Lee and Yu (2010a) for a dynamic spatial panel data model with and without time-period fixed effects, respectively. In addition, the bias correction is needed because the demeaning procedure to wipe out the country and time-period fixed effects in a standard panel data model (Baltagi, 2005) produces a singularity among the transformed error terms if the model is augmented with a spatial lag in the dependent variable, causing the asymptotic distributions of the parameters not to be properly centered.

[9] To save space, the estimation results of the dynamic SAR model without the bias correction are not reported in this paper, but they are available upon request.

The significant positive estimated coefficient δ indicates that CO_2 emissions in neighboring countries have a positive effect on local CO_2 emissions. Besides, we find evidence of what Korniotis (2010) labels external habit persistence; the coefficient of the CO_2 emissions observed in neighboring countries in the previous period is negative and statistically insignificant (–0.1060, p-value <0.01). Otherwise, countries respond to the CO_2 emissions set in neighboring countries in the same year, such that the coefficient τ takes a positive value of 0.7388 and is highly significant (p-value <0.01), in line with the common feature of horizontal interaction among countries (Brueckner, 2003).

Focusing on the estimated coefficient of per capita real GDP, the elastic coefficient is 0.3319 and statistically significant at the 1% level, which indicates that per capita real income has a negative effect on CO_2 reduction. In addition, the estimated coefficients of the quadratic polynomial of real per capita GDP are highly significant indicating that the relationship between CO_2 emissions and economic growth validate the traditional EKC hypothesis. Our results corroborate the view of other authors (e.g., You and Lv, 2018). The turning point of EKC for CO_2 emissions in the dynamic SDM model is approximately \$ 1 849 516.4465. While it is difficult to estimate the specific year when the turning point has been occurred, governments should abandon the pattern of treatment after pollution, develop the economy and cure the environmental issues at the same time.

Concentrating on the estimated coefficient of trade openness, the elastic coefficient is 0.0509 and significant at the 5% level. All else being equal, higher trade openness increases CO_2 emissions. This result indicates that import and export trade have a negative effect on CO_2 reduction. This result accepts the pollution haven hypothesis (PHH), or pollution haven effect, that polluting countries will relocate to jurisdictions with less stringent environmental regulations. Our results are not consistent with the views of Kearsley and Riddel (2010), Dong et al. (2010), and Kang et al. (2016).

The estimated coefficient on both population and technology are respectively positive and negative but statistically insignificant. Accordingly, we can ignore their impact on per capita CO_2 emissions. Otherwise, the estimated coefficient on energy intensity is positive and highly significant. It indicates that a 1% increase on the total energy use per dollar of GDP will lead to a 0.2247 % increase in CO_2 emissions. This result implies that, all else equal, higher energy intensity, increased CO_2 emissions in a given country. This result is consistent with the view of Shahbaz et al. (2015).

Finally, urbanization has a negative and significant effect on CO_2 emissions in middle-income countries. This finding is not consistent with the views of You and Lv (2018). The elastic coefficient of urbanization is –0.2509, which means a 1% increase in urban population will result in a 0.2509% decrease in CO_2 emissions. In other words, urbanization has a positive impact on CO_2 reduction in middle-income countries. This result indicates that middle-income countries considered in this paper promote low-carbon urbanization progress and spread the application of green architecture technology with the topic of energy-saving and environmental protection to develop green city. Overall, the results of this study show that per capita real income, trade openness and energy intensity have significant positive effects on CO_2 emissions, while urbanization has a significant negative effect on CO_2 emissions. Considering these results, policymakers should realize an integrated policy with the aim at reducing CO_2 emissions based on the determinants.

3.4 Direct and spillover effects

It is noteworthy that the coefficients of the dynamic SDM do not directly reflect the marginal effects of the corresponding explanatory variables on the dependent variable. Therefore, we report both the short-term and long-term impacts of the direct and spillover effects of the explanatory variables. Table 7 displays both short and long-term estimates of the direct and spillover effects, derived from the parameter estimates using Formula (12). To draw inferences regarding the statistical significance of these effects, the variation of 1 000 simulated parameter combinations is used, drawn from the variance-covariance matrix implied by the ML estimates.

To get better fitting effects, we conduct a comparative analysis between the dynamic SDM with spatial and time-period fixed effects in Table 7 and non-spatial panel data model with two-way fixed effects in Table 5. The results indicate that most coefficients in non-spatial panel data model are larger than those in dynamic spatial panel data model. Two main reasons could explain this difference. The first one is mainly attributed to ignoring the spatial spillover effect of data. The second reason is due to the feedback

Table 7 Results of the dynamic SDM with $W = W_3$

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$W * ln(CO_2)_t$: δ	0.1595***	0.0000	-	-	-	-	-	-	-	-
$ln(CO_2)_{t-1}$: τ	0.7388***	0.0000	-	-	-	-	-	-	-	-
$W * ln(CO_2)_{t-1}$: η	–0.1060***	0.0000	-	-	-	-	-	-	-	-
$ln(RGDP)_t$	0.3319***	0.0017	0.1510	1.5542	–1.1431	–1.5933	0.8889	0.6858	–2.9471	–0.2954
$ln(RGDP)_t^2$	–0.0115**	0.0133	–0.0055	–1.2788	0.0384	1.3563	–0.0315	–0.7068	0.1000	0.2700
$ln(TRO)_t$	0.0509**	0.0165	0.0578***	2.7957	0.0405	0.3912	0.1976	0.9516	0.0292	0.0442
$ln(URBA)_t$	–0.2509**	0.0222	–0.2259**	–2.2668	0.1795	0.3599	–0.9323	–0.7773	0.2357	0.0390
$ln(POP)_t$	0.0495	0.5680	0.0336	0.4904	–0.1329	–0.4346	0.1741	0.3284	–0.2097	–0.1015
$ln(TECH)_t$	–0.0068	0.6391	–0.0102	–0.4476	–0.0218	–0.1849	–0.0290	–0.1567	–0.0506	–0.0810
$ln(EI)_t$	0.2247***	0.0000	0.2499***	8.1734	0.1548	0.9899	0.8927***	2.9320	0.1922	0.1439
$W * ln(RGDP)_t$	0.0719	0.2074	-	-	-	-	-	-	-	-
$W * ln(RGDP)_t^2$	–0.0023	0.3038	-	-	-	-	-	-	-	-
$W * ln(TRO)_t$	–0.0129	0.1559	-	-	-	-	-	-	-	-
$W * ln(URBA)_t$	0.0246	0.4998	-	-	-	-	-	-	-	-
$W * ln(POP)_t$	0.0043	0.9443	-	-	-	-	-	-	-	-
$W * ln(TECH)_t$	0.0040	0.7317	-	-	-	-	-	-	-	-
$W * ln(EI)_t$	–0.0530***	0.0001								
Observations	850	-	-	-	-	-	-	-	-	-
R^2	0.9940	-	-	-	-	-	-	-	-	-
σ^2	0.0081	-	-	-	-	-	-	-	-	-
Log-likelihood	943.4051	-	-	-	-	-	-	-	-	-
$\tau + \delta + \eta$	0.7923	-	-	-	-	-	-	-	-	-
Wald's stability test: $\tau + \delta + \eta = 1$	59.2319***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SAR	59.4501***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SEM	233.9187***	0.0000	-	-	-	-	-	-	-	-

Notes: Country and time-period fixed effects are included. All variables are in natural logarithms. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.
Source: Own estimates

effects that arise CO_2 emissions of local country as a result of influencing the CO_2 emissions of adjacent countries. In addition, one part of the feedback effects is from spatially lagged dependent variable, while the other part comes from the spatially lagged independent variables.

The coefficient estimates and short-term direct effects estimates derived from the parameter estimates using Formula (12) exhibit a plausible model structure. The direct effect of trade openness on CO_2 emissions is positive and highly significant but lesser than 1. A one percentage point increase of the trade openness has an adverse effect on CO_2 emissions, equal to 0.0578 percentage points. The impact of urbanization on a country's CO_2 emissions is negative and statistically significant at the 5% level. The direct effect of the energy intensity variable is positive and highly significant. This finding indicates that the CO_2 emissions increases with a higher level of energy intensity. However, only the energy intensity variable exhibits significant long-term direct effects, but its magnitude almost a fourth.

Spatial spillover effects are local in nature and cannot be observed directly from the estimated coefficients reported in Table 5. Alternatively, we report the average values of the short and long-term spillover effects of Formula (12) in Table 7. The observed spillover effects in the short term or in the long term are not statistically significant. Therefore, the considered explanatory variables observed in neighboring countries do not have impacts on CO_2 emissions.

3.5 Robustness checks

We report and discuss the results of two robustness checks, thereby focusing on short-term direct and country spillover effects. First, we re-estimate the dynamic SDM specification by replacing the spatial weight matrix by the second-order binary contiguity matrix W_2, in line with the results in Table 6. The results reported in Table 8 show that changes are somewhat tiny for almost the independent variables whether in terms of statistical significance or magnitude, which further confirms the robustness of our main findings with model specification.

With our second robustness check, we follow You and Lv (2018) by exploring whether the results are changed when ruling out population explanatory variable and expressing the main variables as population weighted values. As acknowledged by You and Lv (2018), the rationale behind of this model is that it factors out the impacts of population on each of these variables. To do so, we also repeated the Bayesian comparison approach and selected simultaneously the best spatial econometric model as well as the best spatial weight matrix. The Bayesian comparison approach allows selecting simultaneously both the dynamic SDM model and W_2 as the most likely spatial panel model and the most likely spatial weight matrix, respectively.[10] The results from Table 9 further support the robustness of the previous findings.

CONCLUSIONS AND POLICY IMPLICATIONS

In this paper, we contributed to the existing literature by performing a more rigorous analysis of the relationship between economic growth and CO_2 emissions in middle income countries. We firstly examined the EKC hypothesis for CO_2 emissions at the country level using a dynamic SDM model with country and time period fixed effects. We also computed the short- and long-term spillover effects of explanatory variables for CO_2 emissions in neighboring countries. Our results imply a positive, nonlinear relationship between economic growth and CO_2 emissions. In other words, we found evidence for the EKC hypothesized, inverted U-shaped relationship between CO_2 emissions and economic growth in middle-income countries. Moreover, trade openness and energy intensity were the major drivers of increasing CO_2 emissions, while urbanization effect plays a crucial role in carbon reduction. The results were generally hold when robustness checks were performed.

[10] To save space, the results of the Bayesian comparison approach are not reported but are available upon request.

Based on the empirical findings of this study, the following policy recommendations are put forward to further mitigate CO_2 emissions in middle-income countries. First, the results of this paper showed evidence of an inverted U-shaped relationship between CO_2 emissions and economic growth, suggesting that CO_2 emissions increases at the early stages of development, but goes down at later stage of development. In this vein, the policies should be device in a way to reduce CO_2 emissions at the later stages of economic development. The PHH stipulates that, when big industrialized countries seek to set up factories abroad, they will often search for the cheapest option in terms of resources and labor that offers the land

Table 8 First robustness check: results of the dynamic SDM with $W = W_2$

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$W * ln(CO_2)_t: \delta$	0.1753***	0.0000	-	-	-	-	-	-	-	-
$ln(CO_2)_{t-1}: \tau$	0.7317***	0.0000	-	-	-	-	-	-	-	-
$W * ln(CO_2)_{t-1}: \eta$	−0.1504***	0.0000	-	-	-	-	-	-	-	-
$ln(RGDP)_t$	0.2185**	0.0170	0.2551**	2.2990	0.3468	0.3230	0.7127	0.1053	−1.9641	−0.0257
$ln(RGDP)_t^2$	−0.0059	0.1788	−0.0086*	−1.7435	−0.0234	−0.5390	−0.0142	−0.0571	0.1294	0.0463
$ln(TRO)_t$	0.0451**	0.0253	0.0374	1.5106	−0.0712	−0.3850	0.1734	0.1384	0.0822	0.0059
$ln(URBA)_t$	−0.2229**	0.0142	−0.2325**	−2.3990	−0.0597	−0.0779	−0.5854	−0.1701	3.2875	0.0861
$ln(POP)_t$	0.0039	0.6216	−0.0232***	−0.2692	−0.2673	−0.6168	0.1450	0.0460	1.2978	0.0368
$ln(TECH)_t$	0.0073	0.9561	−0.0023	−0.0831	−0.0829	−0.3538	0.0870	0.0561	0.8105	0.0458
$ln(EI)_t$	0.2368***	0.0000	0.2375***	6.8674	−0.0041	−0.0145	0.6395	0.1838	−3.0480	−0.0764
$W * ln(RGDP)_t$	−0.0625	0.5022	-	-	-	-	-	-	-	-
$W * ln(RGDP)_t^2$	0.0027	0.4646	-	-	-	-	-	-	-	-
$W * ln(TRO)_t$	−0.0033	0.5496	-	-	-	-	-	-	-	-
$W * ln(URBA)_t$	0.0466	0.5393	-	-	-	-	-	-	-	-
$W * ln(POP)_t$	0.0162	0.2813	-	-	-	-	-	-	-	-
$W * ln(TECH)_t$	0.0050	0.6336	-	-	-	-	-	-	-	-
$W * ln(EI)_t$	−0.0417***	0.0022	-	-	-	-	-	-	-	-
Observations	850		-	-	-	-	-	-	-	-
R^2	0.9940		-	-	-	-	-	-	-	-
σ^2	0.0079		-	-	-	-	-	-	-	-
Log-likelihood	953.2061		-	-	-	-	-	-	-	-
$\tau + \delta + \eta$	0.7566		-	-	-	-	-	-	-	-
Wald's stability test: $\tau + \delta + \eta = 1$	88.0738***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SAR	66.4035***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SEM	280.4242***	0.0000	-	-	-	-	-	-	-	-

Notes: Country and time-period fixed effects are included. All variables are in natural logarithms. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.
Source: Own estimates

and material access they require. This hypothesis surpasses the income per capita that noticeably increases CO_2 emissions in the middle-income countries. The existence of EKC in the middle-income countries gives food-for-thought for the environmentalist to establish environmentally friendly and sustainable policies. Besides, the world is in fierce competition which can damage the natural flora of the world's resources that is considered the brazen growth for the economies. Thus, there is a strong need to set an optimistic target for economic growth that would easily be achieved without the cost of environmental degradation. Second, middle-income countries should decrease the amount of trade for lower pollution. However, this decision may deteriorate the economic situation of these countries. Although trade openness in conjunction with economic growth may cause environmental worsening, it is an important contributor to economic growth of several middle-income countries. Accordingly, policymakers should use trade openness to stimulate non-polluted industries by imposing taxes on polluted industries and creating incentives on non-polluted industries in order to encourage producers to shift toward cleaner and more environmentally friendly industries. Third, the positive impact of energy intensity on CO_2 emissions emphasizes the importance of re-structuring the energy use in middle-income countries such that increase in energy intensity does not necessarily translate into higher CO_2 emissions. As an adequate solution for these countries, governments should promote renewable energy technologies. Finally, urban planners should use efficient urbanization to curb the CO_2 emissions, especially for the countries with high density of population. Particularly, they should take thoughtful action on climate change by improving the public transportation systems and the energy efficiency of buildings and increasing the share of renewable energy sources in energy supplies.

Table 9 Second robustness check: results of the dynamic SDM with $W = W_2$

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
$W * ln(CO_2)_t: \delta$	0.1799***	0.0000	-	-	-	-	-	-	-	-
$ln(CO_2)_{t-1}: \tau$	0.7311***	0.0000	-	-	-	-	-	-	-	-
$W * ln(CO_2)_{t-1}: \eta$	−0.1579***	0.0000	-	-	-	-	-	-	-	-
$ln(RGDP)_t$	0.2268***	0.0082	0.2467**	2.4668	0.1996	0.3230	0.6230	0.1343	−3.5816	−0.0663
$ln(RGDP)_t^2$	−0.0063*	0.0965	−0.0080*	−1.8229	−0.0166	−0.5390	−0.0121	−0.0497	0.1664	0.0592
$ln(TRO)_t$	0.0429**	0.0216	0.0400*	1.8849	−0.0304	−0.3850	0.2199	0.1146	0.9094	0.0328
$ln(URBA)_t$	−0.2023**	0.0218	−0.2488***	−2.9100	−0.4525	−0.0779	−0.2300	−0.0279	6.3815	0.0699
$ln(POP)_t$	0.0089	0.9485	−0.0025	−0.1027	−0.0990	−0.6168	0.1867	0.0528	2.0634	0.0418
$ln(TECH)_t$	0.2356***	0.0000	0.2354***	7.6816	−0.0037	−0.3538	0.6551	0.1722	−2.6891	−0.0610
$ln(EI)_t$	−0.0542	0.7221	-	-	-	−0.0145	-	-	-	-
$W * ln(RGDP)_t$	0.0024	0.6364	-	-	-	-	-	-	-	-
$W * ln(RGDP)_t^2$	−0.0055	0.3288	-	-	-	-	-	-	-	-
$W * ln(TRO)_t$	0.0717**	0.0135	-	-	-	-	-	-	-	-
$W * ln(URBA)_t$	0.0068	0.5174	-	-	-	-	-	-	-	-
$W * ln(POP)_t$	−0.0421***	0.0017	-	-	-	-	-	-	-	-
$W * ln(TECH)_t$	850	-	-	-	-	-	-	-	-	-
$W * ln(EI)_t$	0.9940	-	-	-	-	-	-	-	-	-

Table 9 (continuation)

Variable	Estimates		Short-term effects				Long-term effects			
			Direct		Spillover		Direct		Spillover	
	Coefficient	p-value	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Observations	0.0080	-	-	-	-	-	-	-	-	-
R^2	952.56153	-	-	-	-	-	-	-	-	-
σ^2	0.7531	-	-	-	-	-	-	-	-	-
Log-likelihood	90.9290***	0.0000	-	-	-	-	-	-	-	-
$\tau + \delta + \eta$	13.3232**	0.0382	-	-	-	-	-	-	-	-
Wald's stability test: $\tau + \delta + \eta = 1$	14.9861**	0.0101	-	-	-	-	-	-	-	-
Wald test for dynamic SAR	66.4035***	0.0000	-	-	-	-	-	-	-	-
Wald test for dynamic SEM	280.4242***	0.0000	-	-	-	-	-	-	-	-

Notes: Country and time-period fixed effects are included. All variables are in natural logarithms. * denotes $p < 0.1$. ** denotes $p < 0.05$. *** denotes $p < 0.01$.
Source: Own estimates

References

ABREU, M., DE GROOT, H. L. F., FLORAX, R. J. G. M. Space and growth: a survey of empirical evidence and methods. *Region et Development*, 2005, 21, pp. 13–44.

ACARAVCI, A. AND OZTURK, I. On the relationship between energy consumption, CO_2 emissions and economic growth in Europe. *Energy*, 2010, 35, pp. 5412–5420.

ADAMS, S. AND KLOBODU, E. K. M. Urbanization, Democracy, Bureaucratic Quality, and Environmental Degradation. *Journal of Policy Modeling*, 2017, 39(6), pp. 1035–1051.

AGRAS, J. AND CHAPMAN, D. A dynamic approach to the Environmental Kuznets Curve hypothesis. *Ecological Economics*, 1999, 28, pp. 267–277.

AHMED, K. AND LONG, W. Environmental Kuznets curve and Pakistan: an empirical analysis. *Proc. Econ. Financ.*, 2012, 1, pp. 4–13.

ALDY, J. E. An Environmental Kuznets Curve analysis of U.S. state-level carbon dioxide emissions. *The Journal of Environment Development*, 2005, 14, pp. 48–72.

ALDY, J. E. Per capita carbon dioxide emissions: convergence or divergence? *Environ. Resour. Econ.*, 2006, 33(4), pp. 533–555.

AL-MULALI, U., OZTURK, I., LEAN, H. H. The Influence of Economic Growth, Urbanization, Trade Openness, Financial Development, and Renewable Energy on Pollution in Europe. *Natural Hazards*, 2015, 79(1), pp. 621–644.

ANSELIN, L. Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 1995, 27(2), pp. 93–115.

ANSELIN, L. Spatial econometrics. In: BALTAGI, BADI H. (eds.) *A companion to theoretical econometrics*, Massachusetts: Blackwell Publishers Lte., 2001.

ANSELIN, L. Spatial econometric. In: MILLS, T. C. AND PATTERSON, K. (eds.) *Palgrave Handbook of Econometrics Econometric Theory*, Vol. 1, pp. 901–969, Basingstoke: Palgrave Macmillan, 2006.

ANSELIN, L. AND FLORAX, R. J. *Small sample properties of tests for spatial dependence in regression models: some further results. New Directions in Spatial Econometrics.* Berlin, Heidelberg: Springer, 1995, pp. 21–74.

ANSELIN, L., LE GALLO, J., JAYET, H. Spatial Panel Econometrics. In: MATYAS, L. AND SEVESTRE, P. *The Econometrics of Panel Data. Fundamentals and Recent Developments in Theory and Practice.* 3rd Ed. Springer-Verlag, 2008, pp. 627–662.

APERGIS, N. AND OZTURK, I. Testing environmental Kuznets curve hypothesis in Asian countries. *Ecological Indicors*, 2015, 52, pp. 16–22.

ASLANIDIS, N. AND IRANZO, S. Environment and development: is there a Kuznets curve for CO_2 emissions? *Applied Economics*, 2009, 41, pp. 803–810.

AZOMAHOU, T., LAISNEY, F., VAN PHU, N. Economic development and CO_2 emissions: a non-parametric panel approach. *Journal of Public Economics*, 2006, 90, pp. 1347–1363.

BALTAGI, B. H. *Econometric analysis of panel data*. 3rd Ed. Wiley, Chichester, 2005.

BRUECKNER, J. K. Strategic Interaction Among Governments: an Overview of Empirical Studies. *International Regional Science Review*, 2003, 26(2), pp. 175–188.

BURKE, M., HSIANG, S. M., MIGUEL, E. Global non-linear effect of temperature on economic production. *Nature*, 2015, 527, pp. 235–239.

BURNETT, J. W., BERGSTROM, J. C., DORFMAN, J. H. A spatial panel data approach to estimating US state-level energy emissions. *Energy Economics*, 2013, 40, pp. 396–404.

BURRIDGE, P. On the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society. Series B-Statistical Methodology*, 1980, 42(1), pp. 107–108.

BURRIDGE, P. Testing for a common factor in a spatial autoregression model. *Environment and Planning A*, 1981, 13, pp. 795–800.

CARSON, R. T., JEON, Y., MCCUBBIN, D. R. The relationship between air pollution emissions and income: US Data. *Environment and Development Economics*, 1997, 2, pp. 433–450.

COLE, M. A. Trade, the pollution haven hypothesis and the Environmental Kuznets Curve: examining the linkages. *Ecological Economics*, 2004, 48, pp. 71–81.

COONDOO, D. AND DINDA, S. Carbon dioxide emissions and income: a temporal analysis of cross-country distributional patterns. *Ecological Economics*, 2008, 65, pp. 375–385.

DE BRUYN, S. M., VANDENBERGH, J. C. J. M., OPSCHOOR, J. B. Economic growth and emissions: reconsidering the empirical basis of Environmental Kuznets Curves. *Ecological Economics*, 1998, 25, pp. 161–175.

DEBARSY, N. AND ERTUR, C. Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, 2010, 40, pp. 453–470.

DIETZ, T. AND ROSA, E. A. Effects of Population and Affluence on CO_2 Emissions. *Proceedings of the National Academy of Sciences*, 1997, 94(1), pp. 175–179.

DIJKGRAAF, E. AND VOLLEBERGH, H. R. J. A test for parameter homogeneity in CO_2 panel EKC estimations. *Environ. Resour. Econ.*, 2005, 32, pp. 229–239.

DINDA, S. Environmental Kuznets Curve hypothesis: a survey. *Ecological Economics*, 2004, 49, pp. 431–455.

DONG, Y., ISHIKAWA, M., LIU, X., WANG, C. An analysis of the driving forces of CO_2 emissions embodied in Japan-China trade. *Energy Policy*, 2010, 38, pp. 6784–6792.

DU, L., WEI, C., CAI, S. Economic development and carbon dioxide emissions in China: Provincial panel data analysis. *China Economic Review*, 2012, 23, pp. 371–384.

DUTT, K. Governance, institutions and the environment-income relationship: a cross-country study. *Environment, Development and Sustainability*, 2009, 11, pp. 705–723.

EHRLICH, P. R. AND HOLDREN, J. P. Impact of Population Growth. *Science*, 1971, 171(3977), pp. 1212–1217.

ELHORST, J. P. Spatial panel data models. In: FISCHER, M. AND GETIS A. (eds.) *Handbook of Applied Spatial Analysis*, Berlin: Springer, 2010a, pp. 377–407.

ELHORST, J. P. Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 2010b, 5(1), pp. 9–28.

ELHORST, J. P. Dynamic spatial panels: models, methods, and inferences. *Journal of Geographical Systems*, 2012, 14(1), pp. 5–28.

ELHORST, J. P. Matlab Software for Spatial Panels. *International Regional Science Review*, 2014a, 37(3), pp. 389–405.

ELHORST, J. P. *Spatial Econometrics: from Cross-Sectional Data to Spatial Panels.* Heidelberg, New York, Dordrecht, London: Springer, 2014b.

ESTEVE, V. AND TAMARIT, C. Threshold cointegration and nonlinear adjustment between CO_2 and income: the Environmental Kuznets Curve in Spain, 1857–2007. *Energy Economics*, 2012, 34, pp. 2148–2156.

FARHANI, S. AND OZTURK, I. Causal relationship between CO_2 emissions, real GDP, energy consumption, financial development, trade openness, and urbanization in Tunisia. *Environmental Science and Pollution Research*, 2015, 22(20), pp. 15663–15676.

FOSTEN, J., MORLEY, B., TAYLOR, T. Dynamic misspecification in the environmental Kuznets curve: Evidence from CO_2 and SO_2 emissions in the United Kingdom. *Ecological Economics*, 2012, 76, pp. 25–33.

FRIEDL, B. AND GETZNER, M. Determinants of CO_2 emissions in a small open economy. *Ecological Economics*, 2003, 45, pp. 133–148.

GALEOTTI, M. AND LANZA, A. Richer and cleaner? A study on carbon dioxide emissions in developing countries. *Energy Policy*, 1999, 27, pp. 565–573.

GALEOTTI, M., LANZA, A., PAULI, F. Reassessing the environmental Kuznets curve for CO_2 emissions: a robustness exercise. *Ecol. Econ.*, 2006, 57(1), pp. 152–163.

GROSSMAN, G. M. AND KRUEGER, A. B. *Environmental impacts of a North American free trade agreement.* Cambridge, Mass., USA: National Bureau of Economic Research, 1991, pp. 1–39.

GROSSMAN, G. M. AND KRUEGER, A. B. Economic growth and the environment. *The Quarterly Journal of Economics*, 1995, 110(2), pp. 353–377.
HALICIOGLU, F. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy*, 2009, 37, pp. 1156–1164.
HOLTZ-EAKIN, D. AND SELDEN, T. M. Stoking the fires? CO_2 emissions and economic growth. *Journal of Public Economics*, 1995, 57, pp. 85–101.
IWATA, H., OKADA, K., SAMRETH, S. A note on the Environmental Kuznets Curve for CO_2: a pooled mean group approach. *Applied Energy*, 2011, 88, pp. 1986–1996.
JALIL, A. AND MAHMUD, S. F. Environment Kuznets curve for CO_2 emissions: a cointegration analysis for China. *Energy Policy*, 2009, 37, pp. 5167–5172.
JAUNKY, V. C. The CO_2 emissions-income nexus: evidence from rich countries. *Energy Policy*, 2011, 39, pp. 1228–1240.
KAIKA, D. AND ZERVAS, E. The environmental Kuznets curve (EKC) theory. Part A: Concept, causes and the CO_2 emissions case. *Energy Policy*, 2013a, 62, pp. 1392–1402.
KAIKA, D. AND ZERVAS, E. The environmental Kuznets curve (EKC) theory. Part B: Critical issues. *Energy Policy*, 2013b, 62, pp. 1403–1411.
KANG, Y.Q., ZHAO, T., YANG, Y. Y. Environmental Kuznets curve for CO_2 emissions in China: a spatial panel data approach. *Ecological Indicators*, 2016, 63, pp. 231–239.
KASMAN, A. AND DUMAN, Y. S. CO_2 emissions, economic growth, energy consumption, trade and urbanization in new EU member and candidate countries: a panel data analysis. *Econ. Model.*, 2015, 44, pp. 97–103.
KEARSLEY, A. AND RIDDEL, M. A further inquiry into the pollution haven hypothesis and the environmental Kuznets curve. *Ecological Economics*, 2010, 69(4), pp. 905–919.
KORNIOTIS, G. M. Estimating Panel Models with Internal and External Habit Formation. *Journal of Business & Economic Statistics*, 2010, 28(1), pp. 145–158.
KUNNAS, J. AND MYLLYNTAUS, T. The Environmental Kuznets Curve hypothesis and air pollution in Finland. *Scandinavian Economic History Review*, 2007, 55(2), pp. 101–127.
LANTZ, V. AND FENG, Q. Assessing income, population, and technology impacts on CO_2 emissions in Canada: Where's the EKC? *Ecological Economics*, 2006, 57, pp. 229–238.
LEE, C-C., CHIU, Y-B., SUN, C-H. Does one size fit all? Are examination of the Environmental Kuznets Curve using the dynamic panel data approach. *Review of Agricultural Economics*, 2009, 31(4), pp. 751–778.
LEE, L. F. AND YU, J. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 2010, 154(2), pp. 165–185.
LESAGE, J. P. Spatial econometric panel data model specification: a Bayesian approach. *Spatial Statistics*, 2014, 9, pp. 122–145.
LESAGE, J. P. Software for Bayesian cross section and panel spatial model comparison. *Journal of Geographical Systems*, 2015, 17(4), pp. 297–310.
LESAGE, J. P. AND PACE, R. K. *Introduction to Spatial Econometrics.* New York: CRC Press, 2009.
LI, T., WANG, Y., ZHAO, D. Environmental Kuznets Curve in China: New evidence from dynamic panel analysis. *Energy Policy*, 2016, 91, pp. 138–147.
LIM, J. Economic Growth and Environment: Some Empirical Evidences from South Korea. *Seoul Journal of Economics*, 1997, 10(3), pp. 273–292.
LINDMARK, M. An EKC-pattern in historical perspective: carbon dioxide emissions, technology, fuel prices and growth in Sweden 1870–1997. *Ecological Economics*, 2002, 42, pp. 333–347.
LV, Z. AND XU, T. Trade openness, urbanization and CO_2 emissions: Dynamic panel data analysis of middle-income countries. *The Journal of International Trade & Economic Development*, 2019, 28(3), pp. 317–330.
MADDISON, D. Environmental Kuznets curves: a spatial econometric approach. *J. Environ. Econ. Manag.*, 2006, 51(2), pp. 218–230.
MADLENER, R. AND SUNAK, Y. Impacts of urbanization on urban structures and energy demand: What can we learn for urban energy planning and urbanization management? *Sustainable Cities and Society*, 2011, 1, pp. 45–53.
MARTÍNEZ-ZARZOSO, I. AND MARUOTTI, A. The impact of urbanization on CO_2 emissions: evidence from developing countries. *Ecol. Econ.*, 2011, 70(7), pp. 1344–1353.
MARTÍNEZ-ZARZOSO, I., BENGOCHEA-MORANCHO, A., MORALES-LAGE, R. The impact of population on CO_2 emissions: evidence from European countries. *Environmental and Resource Economics*, 2007, 38, pp. 497–512.
MENG, B., WANG, J., ANDREW, R., XIAO, H., XUE, J., PETERS, G. P. Spatial spillover effects in determining China's regional CO_2 emissions growth: 2007–2010. *Energy Economics*, 2017, 63, pp. 161–173.
MENG, L. AND HUANG, B. Shaping the relationship between economic development and carbon dioxide emissions at the local level: Evidence from spatial econometric models. *Environmental and Resource Economics*, 2018, 71(1), pp. 127–156.
MOOMAW, W. R. AND UNRUH, G. C. Are Environmental Kuznets Curves misleading us? The case of CO_2 emissions. *Environment and Development Economics*, 1997, 2, pp. 451–463.

MORAN, P. A. P. Notes on Continuous Stochastic Phenomena. *Biometrika*, 1950, 37(1/2), pp. 17–23.
NARAYAN, P-K. AND NARAYAN, S. Carbon dioxide emissions and economic growth: panel data evidence from developing countries. *Energy Policy*, 2010, 38, pp. 661–666.
ONAFOWORA, O. A. AND OWOYE, O. Bounds testing approach to analysis of the environment Kuznets curve hypothesis. *Energy Economics*, 2014, 44, pp. 47–62.
OZTURK, I. AND ACARAVCI, A. The long-run and causal analysis of energy growth, openness and financial development on carbon emissions in Turkey. *Energy Economics*, 2013, 36, pp. 262–267.
PAO, H. T. AND TSAI, C. M. Multivariate Granger causality between CO_2 emissions, energy consumption, FDI (foreign direct investment) and GDP (gross domestic product): Evidence from a panel of BRIC (Brazil, Russian Federation, India, and China) countries. *Energy*, 2011, 36(1), pp. 685–693.
PERRINGS, C. AND ANSUATEGI, A. Sustainability, growth and development. *Journal of Economic Studies*, 2000, 27(1/2), pp. 19–54.
RICHMOND, A. K. AND KAUFMANN, R. K. Is there a turning point in the relationship between income and energy use and/or carbon emissions? *Ecological Economics*, 2006, 56(2), pp. 176–189.
ROBERTS, J. T. AND GRIMES, R. Carbon intensity and economic development 1962–1991: a brief exploration of the Environmental Kuznets Curve. *World Development*, 1997, 25(2), pp. 191–198.
ROY CHOWDHURY, R. AND MORAN, E. F. Turning the curve: a critical review of Kuznets approaches. *Appl. Geogr.*, 2012, 32(1), pp. 3–11.
SABOORI, B. AND SULAIMAN, J. Environmental degradation, economic growth and energy consumption: Evidence of the environmental Kuznets curve in Malaysia. *Energy Policy*, 2013, 60, pp. 892–905.
SABOORI, B., SULAIMAN, J., MOHD, S. Economic growth and CO_2 emissions in Malaysia: a cointegration analysis of the environmental Kuznets curve. *Energy Policy*, 2012, 51, pp. 184–191.
SCHMALENSEE, R., STOKER, T. M., JUDSON, R. A. World carbon dioxide emissions: 1950–2050. *The Review of Economics and Statistics*, 1998, 80, pp. 15–27.
SHAHBAZ, M., KHRAIEF, N., UDDIN, G. S., OZTURK, I. Environmental Kuznets curve in an open economy: a bounds testing and causality analysis for Tunisia. *Renewable and Sustainable Energy Reviews*, 2014, 34, pp. 325–336.
SHAHBAZ, M., SOLARIN, S. A., SBIA, R., BIBI, S. Does energy intensity contribute to CO_2 emissions? A trivariate analysis in selected African countries. *Ecological Indicators*, 2015, 50, pp. 215–224.
WANG, S. AND LIU, X. China's city-level energy-related CO_2 emissions: Spatiotemporal patterns and driving forces. *Applied Energy*, 2017, 200, pp. 204–214.
WANG, S., FANG, C., WANG, Y. Spatiotemporal variations of energy-related CO_2 emissions in China and its influencing factors: an empirical analysis based on provincial panel data. *Renewable and Sustainable Energy Reviews*, 2016, 55, pp. 505–515.
WANG, S. S., ZHOU, D. Q., ZHOU, P., WANG, Q. W. CO_2 emissions, energy consumption and economic growth in China: a panel data analysis. *Energy Policy*, 2011, 39, pp. 4870–4875.
YIN, J., ZHENG, M., CHEN, J. The effects of environmental regulation and technical progress on CO_2 Kuznets curve: an evidence from China. *Energy Policy*, 2015, 77, pp. 97–108.
YORK, R., ROSA, E. A., DIETZ, T. STIRPAT, IPAT and ImPACT: Analytic Tools for Unpacking the Driving Forces of Environmental Impacts. *Ecological Economic*s, 2003, 46(3), pp. 351–365.
YOU, W. AND LV, Z. Spillover effects of economic globalization on CO_2 emissions: a spatial panel approach. *Energy Economics*, 2018, 73, pp. 248–257.
YU, J., DE JONG, R., LEE, L. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 2008, 146(1), pp. 118–134.
ZHANG, Y. J., LIU, Z., ZHANG, H., TAN, T.-D. The impact of economic growth, industrial structure and urbanization on carbon emission intensity in China. *Nat Hazards*, 2014, 73, pp. 579–595.
ZHAO, X., BURNETT, J. W., FLETCHER, J. J. Spatial analysis of China province-level CO_2 emission intensity. *Renewable and Sustainable Energy Reviews*, 2014, 33, pp. 1–10.

APPENDIX

Table A1 Country list

Lower middle-income countries ($ 996 to $ 3 895)		Upper middle-income countries ($ 3 896 to $ 12 055)	
Country Name	**Country Code**	**Country Name**	**Country Code**
Bangladesh	BGD	Algeria	DZA
Bolivia	BOL	Armenia	ARM
Cambodia	KHM	Azerbaijan	AZE
Cameroon	CMR	Belarus	BLR
Congo, Rep.	COG	Botswana	BWA
Egypt, Arab Rep.	EGY	Brazil	BRA
El Salvador	SLV	Bulgaria	BGR
Honduras	HND	Colombia	COL
India	IND	Costa Rica	CRI
Indonesia	IDN	Dominican Republic	DOM
Kyrgyz Republic	KGZ	Ecuador	ECU
Moldova	MDA	Gabon	GAB
Morocco	MAR	Guatemala	GTM
Nicaragua	NIC	Iran, Islamic Rep.	IRN
Nigeria	NGA	Jordan	JOR
Pakistan	PAK	Kazakhstan	KAZ
Philippines	PHL	Malaysia	MYS
Sri Lanka	LKA	Mexico	MEX
Sudan	SDN	Namibia	NAM
Tunisia	TUN	Paraguay	PRY
Ukraine	UKR	Peru	PER
Uzbekistan	UZB	Romania	ROU
Vietnam	VNM	Russian Federation	RUS
		South Africa	ZAF
		Thailand	THA
		Turkey	TUR
		Venezuela, RB	VEN

Source: World Bank Country Classifications by income level (2018–2019)