Lett Spat Resour Sci
https://doi.org/10.1007/s12076-018-0203-8

ORIGINAL PAPER

Taming the SO_2 and NO_x emissions: evidence from a SUR model for the US

Michael Polemis[1] · **Thanasis Stengos**[2]

Received: 12 November 2017 / Accepted: 15 March 2018

Abstract We construct a Seemingly Unrelated Regression (SUR) model to investigate the link between local environmental pollution (sulfur dioxide-SO_2 and nitrogen oxides-NO_x emissions) and economic growth on a panel data set framework for the US over the period 1990–2012. The presence of different polynomials of GDP for each equation of SO_2 and NO_x respectively allows us to utilize a SUR model to estimate jointly the two equations in order to examine the total effect of environmental degradation. While we find evidence of a quartic relationship between SO_2 emissions and economic development in a single equation framework this outcome does not seem to hold when we utilize a SUR model controlling for cross section dependence.

Keywords Environmental Kuznets curve · SUR · Cross section dependence · Local pollutants

JEL Classification C33 · Q56 · Q43

1 Introduction

Within the last years there is a vast body of literature examining the existence of the Environmental Kuznets Curve (EKC) hypothesis using panel data techniques (parametric and semiparametric models) with controversial results (see for example Polemis

✉ Michael Polemis
mpolemis@unipi.gr

Thanasis Stengos
tstengos@uoguelph.ca

1 University of Guelph, Piraeus, Greece

2 University of Guelph, Guelph, ON, Canada

Published online: 19 March 2018

2018; Sephton and Mann 2016; Desbordes and Verardi 2012; Hong and Wagner 2008; Halkos 2006; Halkos 2003; Millimet et al. 2003).[1] The majority of the existing studies devoted on testing an EKC hypothesis estimate reduced-form equations that enter the model either in a parametric (piecewise linear, quadratic, cubic models) or in a nonparametric form (i.e. semiparametric, partially linear models, etc).[2]

Millimet et al. (2003) explore the significance of modeling policies when calculating the association between emissions-income. Similarly to our study, they use USA state-level panel data on two air pollutants (NO_x and SO_2) in order to estimate several EKCs by comparing parametric and semiparametric techniques. They argue in favor of the semiparametric approach confirming the hypothesis of an inverted U-shape between emissions and regional economic growth. On the contrary, Halkos and Tsionas (2001), empirically test the existence of an EKC, using switching regime models along with cross-sectional data and Bayesian Markov chain Monte Carlo methods. Their findings reject the existence of an EKC since a monotonic relationship between environmental degradation and income prevails.

Halkos and Tzeremes (2011) claim that some global pollutants such as CO_2 emissions exhibit an N shape, implying that the environmental damage starts rising again after a fall to a specific point. On the other hand, relatively few empirical studies adopt a simultaneous equations system in order to address the impact of economic growth on environmental degradation. Dean (2002) uses a panel simultaneous equations system drawn from a Heckscher–Ohlin model in order to capture certain effects of trade liberalization on the environmental quality (water pollution). The sample included 28 Chinese provinces over the period 1987–1995 and the empirical findings suggest that there is a direct negative trade effect on environmental damage which is fully reversed when the income growth is taken into account. In another study, Jayanthakumaran and Liu (2012) try to assess the relationship in China between trade, growth and emissions using provincial panel data for water and air pollution over the period 1990–2007. They use a variety of econometric techniques ranging from a quadratic log function specification to a simultaneous equations system similar to Dean's approach to provide little support in favor of the EKC hypothesis.

Rupasingha et al. (2004) investigate the relationship between US county per capita income and toxic pollutants using the EKC framework. The model incorporates *inter alia* ethnic diversity and spatial characteristics (i.e. distance, urbanization, ruralness, etc.). This study accepts that the relationship follows an inverted-U-shape. Moreover, ethnic diversity and spatial effects are found to be important in understanding toxic pollution in US counties uncovering an EKC-type relationship with respect to income inequality. Keene and Deller (2015) use US county data to examine patterns in ambient concentrations as a measure of air pollution within the framework of the EKC hypothesis. In this study social capital and aspects of ruralness enter the EKC equation as separate regressors. The empirical findings do indicate the validity of the EKC hypothesis, suggesting that higher levels of social capital places downward pressure

[1] EKC hypothesis implies a non linear relationship of an inverted 'U' type between environmental degradation and economic growth.

[2] For a survey of the EKCs on an empirical and theoretical perspective see the relevant studies of Dinda (2004) and Kijima et al. (2010) respectively.

on ambient concentrations, but that effect is weaker in more rural areas. Moreover, it is argued that the promotion of economic growth may harm the environment at lower levels of income but will improve the environment as income continues to grow.

Lastly, in a recent study Tzeremes (2018) investigates the relationship between global pollutants (CO_2 emissions), energy consumption, and economic growth by applying a time-varying causality approach for 50 US states spanning the period 1960–2010. This study, indicates eight bidirectional time-varying causalities between energy consumption and CO_2 emission, six cases of two-way time-varying causalities between economic growth and energy consumption, and five bidirectional time-varying causalities between economic growth and CO_2 emission. Similarly to Halkos and Tzeremes (2011), the empirical findings do not support the validity of the EKC, albeit the majority of states support an inverted N-shaped relationship.

Despite the rich body of literature, existing studies fail to deal with three major issues. First, they assume that the variables or the random disturbances are not correlated across the panel dimension justifying the existence of cross-sectional independence. In other words, the majority of the EKC studies assume that the growth-emissions nexus is space invariant. However, this is a rather strong assumption that has to be tested rather than assumed in order to avoid biased results. Moreover, it is common for macro-level data to violate this assumption which will result in low power and size distortions of tests that assume cross-section independence. The latter may arise due to common unobserved effects generated by changes in federal environmental legislation. Therefore, cross section independence is a strong assumption that has to be tested rather than assumed in order to avoid misleading results. Second and most importantly, nearly all of the existing studies ignore the interdependence between the two major local pollutants such as SO_2 and NO_x emissions which in our model act as a driving force to uncover the validity of the EKC hypothesis. Third but also significant much of the empirical literature has worked with aggregated data (country or block of countries level). These datasets have many drawbacks (i.e., less information, inconsistency), since the use of aggregated data takes away much price variation. We deal with this limitation by using regional data (state level) in our study. The main advantage of using regional data when examine EKC hypothesis enable us to capture spatial phenomena which cannot be captured with aggregated data.

The presence of different polynomials of GDP for each equation of SO_2 and NO_x respectively allows us to utilize a SUR system to estimate jointly the two equations in order to examine the joint effect of environmental degradation.[3] While we find evidence of a quartic relationship between SO_2 emissions and economic development when estimated separately, this outcome does not seem to hold when we utilize a SUR model controlling for cross section dependence. In that case we find a relationship consistent with an inverted U shape as in the case of NO_x.

[3] It is worth mentioning that a polynomial regression raises some challenging technical issues that must be addressed (Wagner and Hong 2016).

Table 1 Descriptive statistics

Variables	(1) Obs	(2) Mean	(3) St. dev	(4) Min	(5) Max
GDP	1173	40,511	17,922	17,392	172,917
GDP^2	1173	1.962e+09	2.856e+09	3.025e+08	2.990e+10
GDP^3	1173	1.271e+14	4.580e+14	5.261e+12	5.170e+15
GDP^4	1173	1.200e+19	7.473e+19	9.150e+16	8.940e+20
SO_2	1173	0.0474	0.0635	0	0.554
NO_X	1173	0.0279	0.0407	0	0.419

2 Data and methodology

Our empirical application is based on a panel of 51 US states (including the District of Columbia) covering the period from 1990 to 2012. The environmental variables (CO_2, SO_2 and NOx emissions per capita) are obtained by the Energy Information Administration (EIA), while logged values of per capita real GDP (in 2009 $US) by state is drawn from the Regional Economic Accounts of the Bureau of Economic Analysis. Table 1 depicts the main descriptive statistics from the model variables.

3 Results and discussion

3.1 Cross section dependence and cointegration testing

One of the additional complications that arise when dealing with panel data compared to the pure time-series case, is the possibility that the variables or the random disturbances are correlated across the panel dimension. The early literature on unit root and cointegration tests adopted the assumption of no cross-sectional dependence. However, as mentioned above it is common for macro-level data to violate this assumption which will result in low power and size distortions of tests that assume cross-section independence. Therefore, before proceeding to unit root and cointegration tests we test for cross-section dependence.

We carry out the first part of the empirical analysis by examining the presence of cross-section dependence. We use the cross-section dependence tests proposed by Breusch and Pagan (1980) and Pesaran (2004). Both tests strongly reject the null hypothesis (p value = 0.000) of cross-section independence for all the models.[4]

To examine the stationarity properties of the variables in our models we use the second generation unit root tests for panel-data proposed by Breitung and Das (2005) and Pesaran (2007). Both tests are based on OLS regressions, however the Breitung

[4] The CD tests were carried out in STATA using the "*xtcd*" and "xtcsd" routines, while we use two Breusch–Pagan LM tests. The first test allows for groupwise heteroskedasticity by using the command "*xttest2*" and the second is a Modified Wald test for groupwise heteroskedasticity in fixed effect regression model by using the command "*xttest3*".

Table 2 Panel unit root test results

Test	NO_x	SO_2	NO_x	GDP	GDP^2	GDP^3	GDP^4
Levels							
CIPS with an intercept	− 2.080	− 2.049	− 2.028	− 1.714	− 1.654	− 1.761	− 1.832
CIPS with an intercept and a trend	− 2.180	− 2.314	− 2.105	− 2.070	− 1.931	− 1.629	− 1.590
First differences							
CIPS with an intercept	− 4.915***	− 4.386***	− 4.915***	− 3.644***	− 3.545***	− 3.320***	− 3.115***
CIPS with an intercept and a trend	− 4.979***	− 4.620***	− 4.979***	− 3.889***	− 3.758***	− 3.393***	− 3.202***

CIPS stands for the cross-sectionally augmented IPS (CIPS) panel unit root test. The null hypothesis assumes stationarity
Significant at ***1%

and Das approach breaks down if it is assumed that cross-correlation is due to common factors while the Pesaran (2007) test, denoted as *CIPS*, remains valid. The *CIPS* test is based on the cross-section augmented Dickey–Fuller test as follows:

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + \beta_i \bar{y}_{i,t-1} + c_i \Delta \bar{y}_t + u_{it} \tag{1}$$

where $\bar{y}_{i,t-1} = \frac{1}{N}\sum_{i=1}^{N} y_{i,t-1}$ and $\Delta \bar{y}_t = \frac{1}{N}\sum_{i=1}^{N} \Delta y_{i,t}$ are used as a proxy for the effect of the unobserved common factor. The *CIPS* test statistic is simply the average t-statistic of the OLS estimate for ρ_i for the individual sections. The null hypothesis that $\rho_i = 0$ for all i is tested against the alternative that only fractions of the series are stationary. The test results suggest that no variables are integrated of an order greater than one (see Table 2).

In order to investigate whether a long-run equilibrium relationship exists among the variables in both models we implement two cointegration tests proposed by Westerlund (2007) that allow for cross-section dependence. In general, the tests are an error-correction approach to testing for cointegration that are based on the statistical significance of the error correction term. The intuition behind this approach is that if a long run relationship between the variables in our model, we can write a regression that allows us to estimate the error-correcting terms which reflect the response of the system to random shocks that "*pushes*" the system towards its long-run equilibrium point. If the error-correction terms are significantly different from zero across sections, then there is evidence in favor of the existence of a long-run relation. The null hypothesis in

Table 3 Westerlund ECM panel cointegration tests

Equation	Statistic			
	G_τ	G_α	P_τ	P_α
SO_2 = f (GDP)	– 3.049*** (0.000)	– 15.449*** (0.000)	– 17.690*** (0.001)	– 17.457*** (0.000)
SO_2 = f $(GDP)^2$	– 3.003*** (0.000)	– 14.973*** (0.001)	– 18.085*** (0.001)	– 18.331*** (0.000)
SO_2 = f $(GDP)^3$	– 2.997*** (0.000)	– 14.896*** (0.001)	– 17.672*** (0.001)	– 18.803*** (0.000)
SO_2 = f $(GDP)^4$	– 2.963*** (0.000)	– 14.734*** (0.001)	– 17.635*** (0.002)	– 19.328*** (0.000)
NO_X = f (GDP)	– 2.566** (0.031)	– 16.381*** (0.000)	– 25.759*** (0.000)	– 16.722*** (0.000)
NO_X = f $(GDP)^2$	– 2.552** (0.041)	– 17.221*** (0.000)	– 25.758*** (0.000)	– 21.524*** (0.000)

The test regression was fitted with a constant and trend and one lag and lead. The kernel bandwidth was set according to the rule $4(T/100)^{2/9}$. The null hypothesis assumes that there is no co-integration
The numbers in parentheses denote the *p* values. Significant at ***1 and **5% respectively

both tests is that of no cointegration. The test statistics of the first two test, denoted G_t, is general enough to allow for individual-specific intercepts and short-run dynamics and is constructed as a weighted average of the estimated error-correcting coefficients across each province in our model. The alternative hypothesis in this test of tests is that at least one section in the panel is cointegrated. The second test assumes that the intercept is the same across sections and tests against the alternative hypothesis that the panel is cointegrated as a whole. The test statistic is denoted by p_t. The results of the tests are presented in the next table; the critical values were created using a bootstrapping method. The results clearly indicate that the first and the second test reject the null hypothesis of no cointegration for both models (Table 3).[5]

3.2 Parametric results

In the previous section we found evidence in favor of cointegration. Hence, our next step is to estimate the long-run equilibrium relationships for each type of emissions (SO_2 and NO_x) separately. Similarly to other empirical studies (Millimet et al. 2003; Zarzoso and Morancho 2004), we estimate separately the following (parametric) OLS fixed-effects panel data models:[6]

[5] The tests were carried out in STATA using the "*xtwest*" routine. It should be noted that the results are sensitive to the selection of the lag structure of the model. Persyn and Westerlund (2008) point out that this sensitivity might occur in small datasets.

[6] The degree of the polynomial for each equation has been determined by the maximum number of statistically significant powers. For example in the case of NO_X third and higher degree polynomial specifications have the extra powers of GDP to be not statistically significant.

$$SO_{2it} = \alpha_i + \beta_t + b_0 + b_1 GDP_{it} + b_2 GDP_{it}^2 + b_3 GDP_{it}^3 + b_4 GDP_{it}^4 + e_{it} \quad (2)$$

$$NO_{Xit} = \alpha_i + \beta_t + b_0 + b_1 GDP_{it} + b_2 GDP_{it}^2 + u_{it} \quad (3)$$

i = 1, 2,…51 and t = 1, 2,…22 where SO_{2it} and NO_{Xit} are the per capita pollutant emissions in state i at time t; α_i and β_t are state and time fixed effects used in order to capture common factors across the cross-section element; GDP_{it} is real GDP per capita for state i at time t and finally e_{it} and u_{it} are zero mean i.i.d. innovations.

From Table 4, it is shown that all the coefficients on the GDP terms in Model 1 (SO_2 emissions) are statistically significant alternating their signs starting from negative to positive.[7] This clearly suggests the existence of a quartic (non linear) polynomial GDP form consisting of three "*turning*" points. On the contrary, in Model 2 (NO_x emissions) the EKC has a U-shaped form since the estimated coefficients of income and squared income are statistically significant alternating their signs starting from negative to positive (Halkos 2003).[8] In other words, regional income level decreases up to a certain point (estimated low) and then increases.

From the above findings, we conclude that each single equation has different RHS variables (i.e. polynomials of GDP). As a consequence, we can apply the SUR method. The SUR estimation results are also reported in Table 4 (Model 3). From the careful consideration of the estimations some interesting results emerge. First, the quartic relationship between SO_2 emissions and economic development does not seem to hold since the coefficient of the cubed and the quartet income (b_3 and b_4) are not significantly different from zero even at the $p < 0.10$ level. Second, the income and the squared income coefficients (b_1 and b_2) although statistically significant alternate their signs starting from negative to positive. This suggests the existence of a 'U' shaped relationship between regional SO_2 emissions and per capita income. The declining part of the EKC may be because of a shock, while the increasing part may be because of an equilibrium association. However, the U-shape EKC still holds when we account for the impact of the other local pollutant (NO_X) on the regional economic growth.

4 Conclusions

We construct a simple SUR model to investigate the link between local environmental pollution (sulfur dioxide-SO_2 and nitrogen oxides-NO_x emissions) and economic growth on a panel data set framework for the US over the period 1990–2012. The presence of different polynomials of GDP for each equation of SO_2 and NO_x respectively allows us to utilize a SUR system to estimate jointly the two equations in order to examine the joint effect of environmental degradation.

Our findings re-shape the impact of regional economic growth on local pollutants by employing techniques allowing for cross-section dependence. Using second generation unit root and cointegration testing we have found strong evidence rejecting the

[7] However, the Wooldridge tests for first order autocorrelation denote that the errors display serial dependence.

[8] If $b_1 > 0$, $b_2 < 0$ and $b_3 > 0$, then we come up with an N-shaped relationship or cubic polynomial (Kijima et al. 2010).

Table 4 Parametric results

Control variables	Model 1 dependent variable: SO_2	Model 2 dependent variable: NOx	Model 3 SUR dependent variable: SO_2	Model 3 SUR dependent variable: NOx
Gdp	− 7.69e−06****** (− 4.41)	− 1.65e–06*** (− 4.76)	− 4.93e−06*** (− 2.88)	− 1.25e–06*** (− 7.04)
Gdp^2	1.55e−10*** (4.97)	6.59e−12*** (5.07)	6.53e−11** (1.66)	5.31e−12*** (4.78)
Gdp^3	− 1.21e−15*** (− 4.74)	–	− 4.37e–16 (− 1.24)	–
Gdp^4	3.21e−21*** (4.40)	–	1.14e–21 (1.08)	–
Constant	0.179*** (6.22)	0.083*** (9.64)	0.160*** (6.52)	0.046*** (16.17)
Diagnostics				
Observations	1173	1173	1173	1173
Polynomial GDP	Quartic	Quadratic	Quadratic	Quadratic
R^2	0.29	0.29	0.09	0.05
F-test/X^2	98.10*** [0.00]	89.98*** [0.00]	6.11*** [0.01]	49.79*** [0.00]
W-T	125.78*** [0.00]	1813.61*** [0.00]	–	–
Root mean square error	–	–	0.039	0.060
Breusch–Pagan test	–	–	541.89*** [0.00]	

Robust z-statistics/t-statistics are in parentheses. The numbers in square brackets denote the *p* values. W-T is the Wooldridge F-test for first order autocorrelation in the error term. The cross-section dependence has been controlled in the estimations. To preserve space and for the sake of simplicity we do not report the estimates of the time dummies

Significant at ***1 and **5% respectively

validity of the EKC hypothesis. The main empirical findings reveal that all the model variables are cointegrated and each single equation has different RHS variables (polynomials of GDP) allowing for the implementation of a simple SUR model controlling for cross section dependence. In other words, we argue that even if the local pollutants seem not to be related, in fact they are and policy makers and government officials need to have a comprehensive approach that tackles everything.

The above findings could be important for policy makers and government officials. More specifically, they call for the need to strengthen the effectiveness of environmental degradation policies by ensuring sustainability of the electricity sector in order to drastically reduce emissions. Lastly, when interpreting our empirical findings, some limitations have to be taken into consideration. Although we used several polynomial powers of GDP to address the impact of growth on environmental degradation and test the EKC hypothesis, special attention should be given to the use of additional macroeconomic indicators (i,e unemployment rate, debt/GDP ratio, level of public deficit, etc.). Moreover, another alley for future research may be to include spatial or threshold aspects such as geographical proximity or terms of trade among the US states, in order to unveil possible state differences and the sources of these different patterns.

Acknowledgements The authors would like to thank the Editor-in-Chief Henrik Folmer for giving them the opportunity to revise their work. Special thanks also go to the fruitful comments and suggestions made by two anonymous reviewers of this journal that enhanced the merit of the paper. The authors also wish to express their gratitude to the Organizational Committee of the 4th Environmental Economics and Natural Resources Conference organised by the University of Thessaly, in Volos, Greece (November 4–5, 2016). Special acknowledgements should be given to Professor George Halkos and the participants of the conference for their fruitful comments and suggestions that enhance the merit of the paper. All remaining errors belong to the authors. The usual disclaimer applies.

References

Breitung, J., Das, S.: Panel unit root tests under cross-sectional dependence. Stat. Neerl. **59**, 414–433 (2005)

Breusch, T.S., Pagan, A.R.: The Lagrange multiplier test and its applications to model specification in econometrics. Rev. Econ. Stud. **47**(1), 239–253 (1980)

Dean, J.M.: Does trade liberalization harm the environment? A new test. Can. J. Econ. **35**(4), 819–842 (2002)

Desbordes, R., Verardi, V.: Refitting the Kuznets curve. Econ. Lett. **116**, 258–261 (2012)

Dinda, S.: Environmental Kuznets curve hypothesis: a survey. Ecol. Econ. **49**, 431–455 (2004)

Halkos, G.: Environmental Kuznets curve for sulfur: evidence using GMM estimation and random coefficient panel data models. Environ. Dev. Econ. **8**(04), 581–601 (2003)

Halkos, G.: Economic development and environmental degradation: testing the existence of an environmental Kuznets curve at regional level. In: ERSA Conference papers ersa06p527, European Regional Science Association (2006)

Halkos, G., Tsionas, E.G.: Environmental Kuznets curves: Bayesian evidence from switching regime models. Energy Econ. **23**(2), 191–210 (2001)

Halkos, G., Tzeremes, N.: Growth and environmental pollution: empirical evidence from China. J. Chin.Econ. Foreign Trade Stud. **4**(3), 144–157 (2011)

Hong, S.H., Wagner, M.: Nonlinear cointegration analysis and the Environmental Kuznets curve. Economics Series 224, Institute for Advanced Studies (2008)

Jayanthakumaran, K., Liu, Y.: Openness and the environmental Kuznets curve: evidence from China. Econ. Model. **29**, 566–576 (2012)

Keene, A., Deller, S.C.: Evidence of the environmental Kuznets' curve among US counties and the impact of social capital. Int. Reg. Sci. Rev. **38**(4), 358–387 (2015)

Kijima, M., Nishide, K., Ohyama, A.: Economic models for the environmental Kuznets curve: a survey. J. Econ. Dyn. Control **34**, 1187–1201 (2010)

Millimet, D., List, J., Stengos, T.: The Environmental Kuznets Curve: real progress or misspecified models? Rev. Econ. Stat. **85**(4), 1038–1047 (2003)

Persyn, D., Westerlund, J.: Error-correction based cointegration tests for panel data. STATA J. **8**(2), 232–241 (2008)

Pesaran, M.: A simple panel unit root test in the presence of cross-section dependence. J. Appl. Econ. **22**, 265–312 (2007)

Pesaran, M.H.: General diagnostic tests for cross section dependence in panels. In: Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge (2004)

Polemis, M.: Revisiting the Environmental Kuznets Curve: a semi-parametric analysis on the role of market structure on environmental pollution. Lett. Spatial Resour. Sci. **11**(1), 27–35 (2018)

Rupasingha, A., Goetz, S.J., Debertin, D.L., Pagoulatos, A.: The environmental Kuznets curve for US counties: a spatial econometric analysis with extensions. Papers Reg. Sci. **83**(2), 407–424 (2004)

Sephton, P., Mann, J.: Compelling evidence of an environmental Kuznets curve in the United Kingdom. Environ. Resour. Econ. **64**(2), 301–315 (2016)

Tzeremes, P.: Time-varying causality between energy consumption, CO_2 emissions, and economic growth: evidence from US states. Environ. Sci. Pollut. Res. **25**(6), 6044–6060 (2018)

Wagner, W., Hong S.H.: Cointegrating polynomial regressions: Fully modified OLS estimation and inference. Econom. Theory **32**(5), 1289–1315 (2016)
Westerlund, J.: Testing for error correction in panel data. Oxford Bull. Econ. Stat. **69**, 709–748 (2007)
Zarzoso, I.M., Morancho, A.B.: Pooled mean group estimation of an environmental Kuznets curve for CO_2. Econ. Lett. **82**, 121–126 (2004)