WORKING PAPER 2004-07

Resource and Environmental economics and Policy Analysis (REPA) Research Group

Department of Economics
University of Victoria

Another look at the income elasticity of non-point source air pollutants: A semiparametric approach

Nilanjana Roy and G. Cornelis van Kooten

REPA Working Papers:

2003-07 – Resolving Range Conflict in Nevada? The Potential for Compensation via Monetary Payouts and Grazing Alternatives (Hobby and van Kooten)
2003-08 – Social Dilemmas and Public Range Management: Results from the Nevada Ranch Survey (van Kooten, Thomsen, Hobby, and Eagle)
2004-01 – How Costly are Carbon Offsets? A Meta-Analysis of Forest Carbon Sinks (van Kooten, Eagle, Manley, and Smolak)
2004-02 – Managing Forests for Multiple Tradeoffs: Compromising on Timber, Carbon and Biodiversity Objectives (Krcmar, van Kooten, and Vertinsky)
2004-03 – Tests of the EKC Hypothesis using CO_2 Panel Data (Shi)
2004-04 – Are Log Markets Competitive? Empirical Evidence and Implications for Canada-U.S. Trade in Softwood Lumber (Niquidet and van Kooten)
2004-05 – Conservation Payments under Risk: A Stochastic Dominance Approach (Benítez, Kuosmanen, Olschewski and van Kooten)
2004-06 – Modeling Alternative Zoning Strategies in Forest Management (Krcmar, Vertinsky, and van Kooten)
2004-07 – Another look at the income elasticity of non-point source air pollutants: A semiparametric approach (Roy and van Kooten)

For copies of this or other REPA working papers contact:

REPA Research Group
Department of Economics
University of Victoria
PO Box 1700 STN CSC
Victoria, BC V8W 2Y2 CANADA
Ph: 250.472.4415
Fax: 250.721.6214
http://repa.econ.uvic.ca

Abstract: In this paper, a semiparametric model is used to examine the relationship between pollution and income for three non-point source pollutants. Statistical tests reject the quadratic specification in favor of the semiparametric model in all cases. However, the results do not support the inverted-U hypothesis for the pollution-income relationship.

Keywords: Environmental Kuznets curve; choice of functional form; nonparametric estimation

JEL classification: C13, Q40, Q25

Acknowledgments: We wish to thank A. Ullah for useful comments and suggestions. We are especially grateful to N. Khanna for making her data available to us. Any remaining errors are ours.

1. Introduction

The environmental Kuznets curve (EKC) hypothesis suggests an inverted-U relationship between economic growth and environmental degradation. The most common test of this hypothesis has been to regress measures of ambient air and water quality on various specifications of per capita income and other relevant regressors, generally using a quadratic or cubic functional form. Dasgupta et al. (2002) express concern about the appropriateness of functional forms in the empirical literature: "*in most cases, the implied relationship between income growth and pollution is sensitive to inclusion of higher-order polynomial terms in per capita income whose significance varies widely*". Our purpose is to investigate the issue of functional form, using the same data as Khanna (2002). Employing U.S. data for 1990, Khanna uses a quadratic specification, regressing the logarithm of the ambient concentration of a pollutant on logarithm of income and a number of control variables. Including a quadratic term in income implies that the relationship is constrained to be U-shaped or an inverted U shape, thus disallowing the possibility of two turning points, say. To address this, we revisit her results using a nonparametric approach to estimation.

To the best of our knowledge, only three papers in the EKC literature (Giles and Mosk (2003), Taskin and Zaim (2000), Millimet et al. (2002)) have used purely nonparametric models, but none have used a semiparametric model. Since Khanna's preferred model for each of the three pollutants has at least twenty-two regressors, pure nonparametric estimation is not feasible (due to the curse of dimensionality). Also, given that her dependent variable is based on different numbers of observations, the implied heteroskedasticity in the model has to be properly taken into account by appropriately adjusting the standard semiparametric estimation technique. We also test Khanna's parameteric specification against our semiparametric model using a test by Li and Wang (1998).

2. Econometric model

The semiparametric model employed here is given as:

$$y_i = m(x_i) + z_i'\delta + u_i, \qquad i = 1,\ldots,n \tag{1}$$

where y_i is the logarithm of ambient concentration of a particular pollutant in region i; x_i is the logarithm of median household income in region i; z_i is a p×1 vector of demographic, political, and other control variables; u_i is the random error term with $E(u_i|x_i, z_i)=0$. Khanna's (2002) model is a special case of (1) with $m(x_i) = \beta_1 x_i + \beta_2 x_i^2$. We estimate (1) for carbon monoxide (CO), ozone (O_3) and nitrogen oxides (NO_x), but unlike Khanna, we use a nonparametric approach.

While Khanna uses weighted least squares approach with the number of observations at each site as the weights, we modify the standard nonparametric estimation of model (1) to take into account the heteroskedasticity. Following Robinson (1988), Stock (1989), and Kniesner and Li (2202), we estimate $f(x_i)$ (the density function of x_i), and the conditional means, $E(y_i|x_i)$ and $E(z_i|x_i)$ by $\hat{f}_i = \frac{1}{nh}\sum_{j=1}^{n} K_{ij}$,

$\hat{y}_i = \frac{1}{nh}\sum_{j=1}^{n} y_i K_{ij} / \hat{f}_i$, and $\hat{z}_{ik} = \frac{1}{nh}\sum_{j=1}^{n} z_{jk} K_{ij} / \hat{f}_i$ respectively, where $K_{ij} = (K(x_i - x_j)/h)$ is the kernel function (we used a normal kernel), z_{ik} is the k^{th} component of the z_i vector, and h is the smoothing parameter.[1] Our density weighted, heteroskedasticity adjusted estimator of δ is given as

$\hat{\delta} = S^{-1}_{(z-\hat{z})\hat{f}} S_{(z-\hat{z})\hat{f},(y-\hat{y})\hat{f}}$ where $S_{A\hat{f},B\hat{f}} = n^{-1}\sum_{i=1}^{n} w_i A_i \hat{f}_i w_i B_i' \hat{f}_i$, $S_{A\hat{f}} = S_{A\hat{f},A\hat{f}}$, and w_i is the square root of the number of observations for the i^{th} region.

To obtain estimators of $m(x_i)$ and its derivative (which denotes the income elasticity of pollution), we first rewrite model (1) as: $y_i - z_i'\hat{\delta} = m(x_i) + v_i$, where $v_i = z_i'(\delta - \hat{\delta}) + u_i$, is the new error term. Then, using a Taylor series expansion, this equation can be rewritten as:

[1] We used h = c.stdx.$n^{-1/5}$ where stdx is the standard deviation of the variable x, and we used c=0.8, 1, 1.2, and 1.4. Since the results were very similar across the c values, to save space we present the results for c=1. The results for all c values can be obtained from the authors upon request.

$$y_i - z_i'\hat{\delta} = m(x) + (x_i - x)\beta(x) + error \quad (2)$$

where β(x) is the first derivative of $m(x_i)$ evaluated at x_i=x.

Nonparametric kernel estimators of m(x) and β(x) can be obtained by using a generalized form of the local linear least squares estimation approach. In particular, we minimize the objective function $(y - Z\hat{\delta} - X\gamma(x))'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}(y - Z\hat{\delta} - X\gamma(x))$ with respect to m(x) and β(x), where y is a n x 1 vector, Z is a n×k matrix, X is a n×2 matrix with X_i = [1 (x_i-x)] as a typical element, γ(x) = [m(x) β(x)]' is a 2×1 vector, $\sqrt{K(x)}$ is a n×n diagonal matrix with the square root of the kernel function K_{ij} as a typical element, and Ω^{-1} is the inverse of an n×n diagonal matrix with w_i as the i^{th} diagonal element. The resulting estimator of γ(x) is given as: $\hat{\gamma}(x) = \{X'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}X\}^{-1}\{X'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}(y - Z\hat{\delta})\}$.

For more details about the generalized local linear estimator but in the context of panel data models see Henderson and Ullah (2003).

3. Results

So that the results are comparable, the variables in z_i in model (1) are the same as in Khanna (2002) and are provided in Table 1. All of the variables in Table 1 are in logarithmic form, except the dummy variable indicating whether or not the region is urban (=1). We include but do not report the results for nine dummy variables for EPA regions, and two to three (depending on the pollutant) dummy variables to account for highly influential observations, exactly as in Khanna (2002).

We first estimated parametric specifications of model (1) with $m(x_i) = \beta_1 x_i$, and then $m(x_i)= \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$ to see how sensitive Khanna's specification is to alternative functional forms. The results are given in Table 1 along with Khanna's result. Note that, for both CO and NO_x, the statistical significance of the income variable is highly sensitive to specification of functional form. For example, in the CO model, income is a statistically significant for the linear specification but not for the quadratic and cubic specifications; on the other hand, for NO_x, the income variable is statistically significant for the quadratic specification but not for the others. Therefore, our preferred model is the semiparametric model that does not impose any functional form restriction on the income variable. We also tested Khanna's quadratic specification against the semiparametric alternative using Li and Wang's (1998) test. For c=1, the values of the test statistic for CO, O_3 and NO_x are 166.471, 10.883, and 89.111, respectively, with the bootstrapped critical values at 1% level of significance being 1.257, 1.205 and 1.334, respectively. The null hypothesis that the quadratic specification is appropriate is clearly rejected in each case. The results are not sensitive to the choice of c.

In Table 2, we provide various distributional statistics for the semiparametric income elasticity estimates, $\hat{\beta}(x_i)$ for i=1,…,n for each pollutant, and for the elasticity estimates from Khanna's parametric model (denoted k). [2] While the mean values are similar across the two models, the percentile distributions indicate that the elasticity estimates are quite different.

Plots of the estimates of the non-linear components of the logarithm of income (vertical axis) for CO, O_3 and NO are provided in Figure 1 for c=1. The results agree with Khanna's conclusion that none of the plots exhibit an inverted-U relationship between the pollutant concentration and income.

4. Conclusion

In this paper, we use a semiparametric framework to extend the work of Khanna (2002) to study the relationship between three pollutant concentrations and income using U.S. data. The advantage of the nonparametric approach adopted here is that it allows the data to determine the functional form with respect to the income variable, rather than imposing an ad hoc functional form a priori. This approach is useful in avoiding the problem of functional form misspecification.

[2] The results for the other c values are available from the authors upon request.

References

Dasgupta, S., B. Laplante, H. Wang and D. Wheeler, 2002, Confronting the Environmental Kuznets Curve, Journal of Economic Perspectives, 16 (1), 147-168.

Giles, D. and C. Mosk, 2003, Ruminant eructation and a long-run Environmental Kuznets' Curve for enteric methane in New Zealand: Conventional and Fuzzy Regression Analysis, Econometrics working paper, University of Victoria.

Henderson, D.J. and A. Ullah, 2003, A nonparametric random effects estimator, working paper, University of California, Riverside.

Khanna, N., 2002, The income elasticity of non-point source air pollutants: revisiting the environmental Kuznets curve, Economics Letters, 77, 387-392.

Kneisner, T. and Q. Li, 2002, Nonlinearity in dynamic adjustment: Semiparametric estimation of panel labor supply, Empirical Economics, 27,

Li, Q. and S.Wang, 1998, A simple consistent bootstrap test for a parametric regression function, Journal of Econometrics, 87, 145-165.

Millimet, D., J. List, and T. Stengos, 2002, The Environmental Kuznets Curve: real progress or misspecified models?, working paper, Southern Methodist University.

Robinson, P., 1988, Root-N-consistent semiparametric estimation, Econometrica, 56, 931-954.

Stock, J.H., 1989, Nonparametric policy analysis, Journal of American Statistical Association, 84, 567-575.

Taskin, F. and O. Zaim, 2000, Searching for a Kuznets curve in environmental efficiency using kernel estimation, Economics Letters, 68, 217-223.

Table 1: Comparison of Parametric and Semi-parametric Environmental Kuznets Curve Model Estimates, Carbon Monoxide, Ozone and Nitrous Oxides

Explanatory	CO Model				Ozone Model				NOx Model			
Variables	1	2	3	SP	1	2	3	SP	1	2	3	SP
income	-0.108**	-1.261	23.838		0.021	-0.537	0.805		0.004	-4.689*	16.585	
	(0.053)	(0.206)	(0.156)		(0.512)	(0.367)	(0.931)		(0.952)	(0.000)	(0.442)	
income squared		0.059	-2.489			0.028	-0.106			0.233*	-1.896	
		(0.247)	(0.144)			(0.348)	(0.909)			(0.000)	(0.397)	
income cubed			0.086				0.004				0.071	
			(0.134)				(0.886)				(0.323)	
Population density	0.155*	0.156*	0.156*	0.182*	0.016*	0.016*	0.016*	0.013*	0.120*	0.121*	0.121*	0.104*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)	(0.005)	(0.005)	(0.020)	(0.000)	(0.000)	(0.000)	(0.000)
% minorities	0.042**	0.043*	0.040**	0.052*	0.026*	0.026*	0.026*	0.034*	0.088*	0.084*	0.082*	0.079*
	(0.051)	(0.048)	(0.063)	(0.016)	(0.008)	(0.009)	(0.009)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)
% unemployed	-0.047	-0.047	-0.047	-0.040	-0.006	-0.004	-0.004	-0.025	-0.004	-0.016	-0.017	-0.026
	(0.218)	(0.218)	(0.221)	(0.305)	(0.777)	(0.841)	(0.837)	(0.268)	(0.933)	(0.765)	(0.751)	(0.614)
% labor in	0.055	0.068**	0.071**	0.142*	0.030**	0.035**	0.035**	0.029	0.109*	0.162*	0.172*	0.242*
manufacturing	(0.132)	(0.076)	(0.063)	(0.000)	(0.100)	(0.068)	(0.068)	(0.148)	(0.015)	(0.000)	(0.000)	(0.000)
% with high	-0.055	-0.044	-0.037	-0.054	-0.017	-0.010	-0.010	-0.003	0.008	0.057	0.057	-0.017
school	(0.211)	(0.331)	(0.418)	(0.333)	(0.476)	(0.686)	(0.696)	(0.918)	(0.89)	(0.323)	(0.326)	(0.803)
% of voters	0.226**	0.227**	0.225**	0.442*	-0.058	-0.059	-0.059	-0.043	-0.581*	-0.639*	-0.628*	-0.681*
registered	(0.059)	(0.058)	(0.059)	(0.000)	(0.222)	(0.215)	(0.217)	(0.356)	(0.000)	(0.000)	(0.000)	(0.000)
% of houses	0.054	0.067	0.068	0.063	-0.017	-0.011	-0.01	-0.021	0.037	0.072	0.075	0.050
rented	(0.213)	(0.135)	(0.130)	(0.179)	(0.378)	(0.586)	(0.597)	(0.340)	(0.470)	(0.155)	(0.138)	(0.314)
% female headed	-0.033	-0.029	-0.03	-0.087*	-0.008	-0.007	-0.007	-0.017	-0.122*	-0.120*	-0.121*	-0.125*
households	(0.230)	(0.291)	(0.279)	(0.007)	(0.646)	(0.697)	(0.689)	(0.390)	(0.001)	(0.001)	(0.001)	(0.001)
urban dummy	0.310*	0.303*	0.301*	0.262*	0.009	0.007	0.007	0.018	0.260*	0.256*	0.259*	0.320*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.704)	(0.770)	(0.769)	(0.456)	(0.000)	(0.000)	(0.000)	(0.000)
Adjusted R^2	0.47	0.471	0.472		0.314	0.314	0.313		0.769	0.779	0.779	
# of observations	509	509	509	509	820	820	820	820	305	305	305	305

Note: p-values are provided in parentheses. * indicates statistical significance at the 5% level or better; ** at the 10% level or better.

Table 2: Distribution of Income Elasticities of Pollutants

	CO		Ozone		NO_x	
Item	SP	Khanna	SP	Khanna	SP	Khanna
Mean	-0.069	-0.081	0.020	0.037	0.088	0.064
Standard dev.	0.211	0.066	0.160	0.014	0.300	0.225
Minimum	-0.916	-0.256	-0.583	0.033	-0.535	-0.720
Percentile Distribution						
10^{th}	-0.319	-0.176	-0.068	0.036	-0.350	-0.237
20^{th}	-0.286	-0.141	-0.049	0.039	-0.271	-0.132
30^{th}	-0.224	-0.112	-0.020	0.042	-0.152	-0.020
40^{th}	-0.139	-0.086	0.013	0.044	0.093	0.047
50^{th}	-0.062	-0.070	0.052	0.049	0.171	0.099
60^{th}	0.016	-0.055	0.103	0.052	0.242	0.131
70^{th}	0.087	-0.042	0.132	0.056	0.301	0.190
80^{th}	0.125	-0.025	0.142	0.061	0.343	0.244
90^{th}	0.175	-0.006	0.215	0.070	0.359	0.326
Maximum (100^{th})	1.649	0.089	0.853	0.106	1.319	0.644

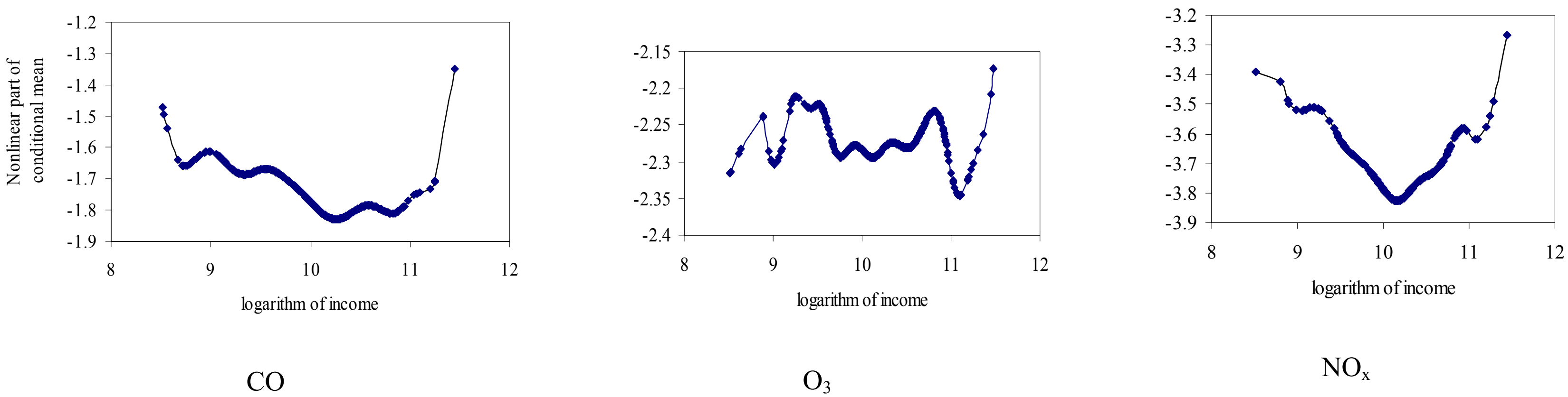

Figure 1: Concentration of Pollutant versus Income