The Journal of Environment & Development
Volume 15 Number 1
March 2006 22-41

10.1177/1070496505285466
http://jed.sagepub.com
hosted at
http://online.sagepub.com

Household Income and Pollution

Implications for the Debate About the Environmental Kuznets Curve Hypothesis

Florenz Plassmann
Neha Khanna
State University of New York at Binghamton

Country-level analyses of global Environmental Kuznets Curve (EKC) relationships that use multicountry panel data sets are likely to suffer from several types of aggregation bias that may explain why previous studies have yielded conflicting results. The authors analyze 1990 cross-sectional data for the United States for three pollutants and test the general EKC relationship as well as the pure income effect. Their results suggest that the income level at which households reduce their exposure to pollution depends on the nature of the pollutant. They find consistent evidence for such a relationship for coarse particulate matter but little evidence for nonmonotonic relationships for carbon monoxide and ground-level ozone.

Keywords: *correlated count data; MCMC; NAAQS; turning point*

I. Introduction

The March 2005 special issue of this journal included several articles on the Environmental Kuznets Curve (EKC) hypothesis. These articles focused on EKC drivers such as economic development (Marcotullio, Williams, & Marshall, 2005) and technological diffusion (e.g., Stern, 2005). We heed the call from the guest editors of this special issue (Leifman & Heil, 2005, pp. 13-14) and provide an "alternative perspective" by focusing on the role of consumer demand for environmental quality. Our goal is to determine the extent to which consumer preferences are an important determinant of an EKC relationship.[1] To that purpose, we use highly disaggregate data for the United States to estimate a relationship between household income and exposure to pollution, holding constant other factors that might affect the pollution-income relationship, and examine whether there is an income level at which consumers are willing to reduce their exposure to pollution. In other words, if structural factors such as the sectoral composition of the economy, education levels, unemployment rates, age dis-

Authors' Note: The authors thank Terence Fitz-Simmons for making the count data available to us, David Mintz for providing us with the ambient concentrations data, and Lisa Brooks for providing GIS support. Barry Jones, Aloke Bohara, Raymond Clémençon, and two anonymous referees provided helpful suggestions and comments. All errors are the authors'.

tribution of the population, and so on were to remain unchanged as an economy grew but household income were to increase monotonically, would the pollution-income relationship eventually turn downward as predicted by the EKC hypothesis?

Several recent articles, for example, List and Gallet (1999), Harbaugh, Levinson, and Wilson (2002), Millimet, List, and Stengos (2003), and Aldy (2005), have shown that estimated EKC relationships are highly sensitive to changes in the underlying data and model specification. We argue that earlier studies have used data that made it difficult to detect the relationship between economic growth and pollution. The main reason for this is that these studies have typically used multicountry data sets that combine environmental and economic data that reflect differences in consumer preferences over space and time. We consider it likely that the effect of changes in the demand for environmental quality is better captured by disaggregate data that link local pollution to local socioeconomic characteristics. Three earlier articles that use more disaggregated state-level data for the United States (Aldy, 2005; List & Gallet, 1999; Millimet et al., 2003) find that the estimated pollution-income relationships vary across states, lending support to our premise that differences in consumer preferences matter in empirical analyses of the EKC so that highly aggregated data are inappropriate for estimating such relationships.

We analyze the relationship between household income and air pollution caused by carbon monoxide (CO), ground level ozone (O_3), and coarse particulate matter (PM_{10}), using 1990 census tract data for the United States. All three pollutants have well-known adverse health effects (primarily respiratory and cardiovascular) and the U.S. Environmental Protection Agency (EPA) has established National Ambient Air Quality Standards (NAAQS) for these pollutants above which ambient concentrations are considered harmful to human health. Because we focus on the relationship between household income and (perceived) air quality, we measure pollution by the number of days during which the concentrations of these three pollutants exceeded their respective NAAQS at 704 locations. Exceedances of the NAAQS are a widely publicized measure of local air quality and are therefore more appropriate for our purposes than data on ambient pollutant concentrations that typically do not convey any meaningful information with regard to air quality to the lay consumer.

For PM_{10}, we find consistent evidence that pollution decreases at high levels of income. We find a concave but monotonically increasing relationship for O_3, and no evidence for correlation between CO exposure and income. Unlike earlier studies, our results are robust to changes in assumptions about the underlying distribution, whether we analyze the pollutants individually or jointly by incorporating the observed correlation between them, and to different functional forms and sets of covariates. We consider it likely that this robustness is due to the fact that we use "the right data for the right problem."

The remainder of the article is organized as follows: The next two sections describe the possible aggregation biases in panel data sets and the properties of our cross-sectional data set. Section IV describes our econometric model and Section V contains our results. Section VI concludes.

II. Choosing the Right Data

Most studies have tested the EKC hypothesis with multicountry panel data sets that contain annual information on either nationwide emissions or local concentrations of various pollutants (usually measured at several sites in a few major cities within each country).[2] These measurements of pollution are regressed on variables measured at the country level. The principal advantage of these data sets is their length and breadth of information.

However, use of pooled cross-section and time-series data from different countries to estimate a global pollution-income relationship may lead to at least three types of aggregation biases. First, ambient concentrations data (such as the widely used Global Environmental Monitoring System [GEMS] data) reflect site-specific pollution measured at various locations, whereas most of the variables on the right side of the regression equation are measured at the country level (most notably, the measure of income is usually per capita GDP). Changes in local pollution are more likely to be correlated with changes in the local economy than with changes in broad economic aggregates such as per capita GDP.

A second type of aggregation bias results from including heterogeneous countries in the same data set. Stern and Common (2001) and Harbaugh et al. (2002) suggest that estimates of an EKC relationship are sensitive to the range of countries included in the analysis. List and Gallet (1999) and Aldy (2005) indicate that the relationship may differ across regions within a country, and Millimet et al. (2003) show that even the pollution-income relationships estimated with state-level panel data for the United States are highly sensitive to modeling assumptions. This is not surprising if the relationship is influenced by local preferences for environmental quality (see Lieb, 2002; Plassmann & Khanna, in press), because there is no reason to expect that such preferences are identical across nations or regions. Furthermore, Chimeli and Braden (2005) suggest that the use of panel data ignores country-specific factors that might give rise to the EKC and leads to biased and inconsistent estimates when these factors are correlated with income.

A third type of aggregation bias stems from the fact that the relationship between income and pollution may shift over time, for example, either because of changes in the state of technology or because people become increasingly aware of the harmful effects of pollution. Thus, estimates of EKC relationships with time series data face the same problem as estimates of consumer demand curves: If the data represent observations from different curves rather than observations along a single curve, then estimates obtained from these data indicate the path along which the relationship has changed over time instead of the relationship itself. For example, List and Gallet (1999) and Millimet et al. (2003) analyze the EKC hypothesis using state-level panel data for the United States from 1929 to 1994 and Aldy (2005) uses similarly disaggregated data from 1960 to 1999. All three studies show that the estimated pollution-income relationships are region-specific and sensitive to modeling assumptions. This suggests that it is necessary to either explicitly account for changes in environmental awareness or use data that cover a time span during which preferences for environmental quality remained unchanged.

These aggregation biases may explain why international- and national-level panel data do not yield robust estimates of the EKC. A simple way to avoid these biases is to use cross-section data for a single country at the smallest feasible level of aggregation, which provide a snapshot of consumer preferences and eliminate the effect of spatial and temporal changes in people's awareness of the harmful effects of pollution. The main disadvantage of cross-sectional data is that they do not permit us to use location-specific dummies to capture the effect of omitted variables that are constant over time. However, including variables that capture the major differences between locations mitigates the effect of omitted variables. In addition, analyzing econometric models with different specifications indicates whether omitted variables have a qualitatively significant effect, because different specifications measure the correlation between included and omitted variables in different ways. If different analyses lead to identical conclusions, then the omitted variables are unlikely to lead to an incorrect interpretation of the results.

A final consideration is that consumers are unlikely to react to changes in pollution that they consider harmless. The EPA has established the NAAQS above which ambient concentrations are considered harmful to human health. Although the NAAQS does not necessarily represent an unambiguous scientific threshold above which pollution is dangerous (or below which pollution is harmless), most people assess their local air quality through exceedances of the NAAQS.[3] Observations below the threshold levels might generate sufficient noise to reduce the precision of the estimated relationship. Simply eliminating the observations for which the ambient concentrations do not exceed the NAAQS does not solve the problem, because the fact that the ambient concentration remained at or below the NAAQS is important information in itself. To reduce the effect of potentially noisy observations without eliminating information, it is important to left-censor the data and to record concentrations at relatively harmless levels as the lower bound of the data.

III. Our Data

To accommodate these issues, we analyze the relationship between household income and pollution using 1990 data for the United States, at the smallest geographic unit possible (census tracts).[4] The United States is at the upper end of the world's income scale, and these data are therefore likely to contain the threshold income levels beyond which pollution begins to decline (if such threshold levels exist). We measure pollution by the number of days in 1990 during which the concentrations of CO, O_3, and PM_{10} at 704 locations exceeded their respective NAAQS. We obtained data on the annual counts in 1990 from the EPA's USEPA-AQS database. This database also provides the geographic coordinates (latitude and longitude) for each monitor, which we used to identify the census tracts in which the monitors are located.

Empirical analyses of the EKC hypothesis seek to determine the relationship between economic growth and environmental degradation. In addition to leading to higher per capita GDP, economic growth has indirect effects like changes in the composition of the work force, improvements in medical services, changes in rural-urban settlement patterns, and so on. To be able to capture all these effects, most studies use only per capita GDP as a proxy for economic growth and other covariates that are likely to be uncorrelated with it.

Because we use household-level intranational data, we use median household income instead of per capita GDP. In addition, we also analyze the direct relationship between household income and changes in the demand for environmental quality (the "pure" income effect) and isolate the effect of income on local pollution by including variables that are likely to be correlated with income. The literature on the distribution of air pollution in the United States suggests that pollution in any given area is influenced by population density, racial composition, housing tenure, education, unemployment rate, and a population's propensity for collective action (see Brooks & Sethi, 1997, and the references cited there). We measure racial composition by the percentage of minorities, the level of education by the percentage of high school graduates, the unemployment rate by the percentage of the labor force that is unemployed, and housing tenure by the percentage of houses that are renter occupied. We obtained census-tract–level data on these variables from the 1990 census. We measure the propensity for collective action at the county level by the fraction of the voting age population that was registered to vote in the 1992 presidential elections.[5] We include three additional variables: the percentage of population below the poverty threshold, the percentage of female-headed households, and the percentage of population older than 65 years.

To account for differences in economic structure, we include the percentage of working age population in each census tract that is employed in manufacturing. We use the distance of the EPA monitors from the closest highway to measure the level of economic activity in the CO and O_3 analyses. This distance serves as a proxy for local economic activity because on-road vehicles are a primary source of emissions of CO and other precursors of ozone. The primary sources of PM_{10} are coal-burning facilities such as electric utilities and copper smelters, and for each monitor we include the number of electric utilities in that EPA region that are monitored under the EPA's Acid Rain Program. To incorporate regional differences, we include dummy variables for the 10 EPA regions.

Table 1 shows summary statistics of all nondummy variables. The range of household income extends from \$4,999 (or log-income = 8.52) to \$96,383 (or log-income = 11.48) with mean \$29,652 (or log-income = 10.30) and median \$28,397 (or log-income = 10.25). Eighty percent of the census tracts in our data set have a median household income between \$14,688 (or log-income = 9.60) and \$46,051 (or log-income = 10.74). We regard a turning point that falls well within this income range as an indicator of a nonmonotonic relationship between income and pollution. A turning point sufficiently far to the left or right of the median suggests a nonlinear but not necessarily nonmonotonic relationship between pollution and income. We use the

Table 1
Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Median household income	29,652	13,057	4,999	96,383
Population density (persons/square mile)	2,742	5,360	.57	102,938
Percentage minority	19.90	22.94	.00	100.00
Percentage of labor force unemployed[a]	7.73	6.24	.00	60.00
Percentage of labor force employed in manufacturing[a]	17.30	9.50	.00	83.33
Percentage high school graduates	71.91	15.84	.00	100.00
Percentage voting age population registered to vote[b]	72.07	10.48	44.72	113.26
Percentage of houses renter occupied	39.92	24.14	1.66	100.00
Percentage below poverty threshold	15.75	14.05	.00	100.00
Percentage female head of household	11.57	7.46	.00	52.72
Percentage population older than 65 years of age	12.15	8.23	.00	100.00
Distance from the nearest highway in meters	1,138	2,051	.59	27,773
Number of electric utilities in EPA region	30.48	22.13	0	106

a. We measure labor force as the population that is 16 years or older.
b. Refers to the 1992 presidential election. Values shown do not include the predicted data for Wisconsin and Alaska. The percentage exceeds 100 at several census tracts because of voter fraud (personal communication with a representative of Election Data Services).

95% confidence interval of our turning point estimate to assess whether the turning point is estimated with sufficient precision so as to be more than an artifact of the chosen functional form.

IV. A Model of Correlated Count Data

The number of days in a year during which the ambient concentration of a pollutant exceeds the NAAQS is a nonnegative integer, also known as "count data." Standard count data analyses assume that the data follow either a Poisson or a negative binomial distribution, and estimate the coefficients with maximum likelihood (e.g., Greene, 2002, pp. 740-747). Let c_{pi} denote the number of days during which the concentration of pollutant p, $p = 1, \ldots, P$, exceeds its NAAQS at location i, $i = 1, \ldots, I$. If we assume that the pollutants are independently Poisson distributed, then we obtain P univariate Poisson models,

$$c_{pi} \sim \text{Poisson}\,(\mu_{pi}), \tag{1}$$

with parameters $\mu_{pi} \in R^+$ that describe the means and variances of the distributions. Assuming that $\mu_{pi} = \exp(\chi_{pi}'\beta_p)$, where the χ_{pi} are covariate vectors and the β_p are the corresponding parameter vectors, turns equation 1 into a set of P unrelated Poisson regression models.

The Poisson regression model implies that the probability of pollutant p exceeding its NAAQS depends on the covariates χ_{pi}, whose effects are identical across census tract areas. But if this effect fluctuates across census tract areas, then the parameter of the Poisson distribution is not deterministic but a random variable itself. The standard way of accommodating such heterogeneity is to assume that $\mu_{pi} = \exp(\chi_{pi}'\beta_p)\varepsilon_{pi}$, where the ε_{pi} follow a univariate gamma distribution. Integrating the density functions over ε_{pi} turns equation 1 into a set of P unrelated Poisson-gamma, or negative binomial, regression models.

The asymmetry of the gamma distribution implies that an increase in c_{pi} by a factor of α is less likely than a decrease by the same factor. But in the absence of information that would warrant such a model, it is more appropriate to assume that an increase and a decrease by the same factor are equally likely. This can be achieved by assuming that $\mu_{pi} = \exp(\chi_{pi}'\beta_p)\exp(\varepsilon_{pi})$, where the ε_{pi} follow a univariate normal distribution. This distribution has the log-symmetry that the gamma distribution lacks and yields the Poisson-lognormal regression model. The ε_{pi} can be interpreted as pollutant-and-location specific latent, or random, effects that measure the effect of variables omitted from the covariate vectors. Because we have only a single observation for each location and cannot include location-specific fixed effects, these random effects are likely to improve the precision of our analysis.[6]

None of the models described above incorporate correlation between the pollutants because they assume independent μ_{pi}s. However, motor vehicles are a major source of emissions of CO and volatile organic compounds (VOCs, precursors of O_3), and the CO and O_3 counts are positively correlated. Electric utilities and copper smelters are a major source of PM_{10} and tend to be located in areas with relatively low population and highway densities. Counts of PM_{10} and the other two pollutants are therefore negatively correlated.

Such correlation can be accommodated by allowing the ε_{pi} to be correlated across pollutants. For example, if we assume that $\mu_{pi} = \exp(\chi_{pi}'\beta_p)\varepsilon_i$, where the ε_i follow a univariate gamma distribution, then integrating the density function of $c_i = (c_1i, \ldots, c_{Pi})$ over the common variable ε_i yields the P-variate Poisson-gamma regression model. The covariance between the concentrations of pollutants p and q at location i is given by $Cov(c_{pi}, c_{qi}) = \exp(\chi_{pi}'\beta_p)\exp(\chi_{qi}'\beta_q)\sigma$, where σ is the variance of the gamma distribution. Because all three terms are positive, the multivariate Poisson-gamma distribution accommodates only nonnegative correlation. An attractive alternative is the multivariate Poisson-lognormal regression model. This model can be obtained by assuming $\mu_{pi} = \exp(\chi_{pi}'\beta_p)\exp(\varepsilon_{pi})$, where the ε_{pi} follow a P-variate normal distribution with mean zero and covariance matrix Σ. Because $Cov(c_{\mathrm{p}i}, c_{qi}) = \exp(\chi_{pi}'\beta_p)\exp(\chi_{qi}'\beta_q)(\exp(\sigma_{pq})-1)$, where σ_{pq} is the element in row p and column q of Σ, the model accommodates positive and negative correlation between the elements of c_i.[7]

Table 2 shows estimates of the covariance matrix Σ and the corresponding correlation matrix of the latent effects of a joint Poisson-lognormal analysis of all three pol-

Table 2
Covariance and Correlation Matrices of the Latent Variables ε_p in the Joint Poisson-Lognormal Analysis in Table 3, Column 5

	Covariance Matrix Σ			Correlation Matrix		
	CO	O_3	PM_{10}	CO	O_3	PM_{10}
CO	8.8047***					
	(2.2235)			1		
O_3	–1.4831***	1.6109***		–.3977***	1	
	(.3649)	(.1219)		(.0844)		
PM_{10}	–1.4672*	–1.5087***	3.5391***	–.2670*	–.6408***	
	(.8725)	(.2482)	(1.0392)	(.1464)	(.0709)	1

Note: Standard errors are shown in parentheses. CO = carbon monoxide; O_3 = ground level ozone; PM_{10} = coarse particulate matter.
* indicates that the coefficient is different from zero at the 90% level. *** indicates that the coefficient is different from zero at the 99% level.

lutants. The latent effects from the three equations show statistically significantly negative correlation. This suggests that the multivariate Poisson-lognormal model is more appropriate for our data than three uncorrelated count models.[8]

The integral of the density function of C_i over $\varepsilon_i = (\varepsilon_{1i}, \ldots, \varepsilon_{Pi})$ does not have a closed form solution, which makes maximum likelihood analysis cumbersome. It is straightforward, however, to estimate the unknown parameters with simulation-based methods such as the Gibbs sampler, a Markov chain Monte Carlo (MCMC) method.[9] We closely follow the setup that Chib and Winkelmann (2001) suggest for this type of analysis, and we describe the Gibbs sampler for our multivariate Poisson-lognormal model in the technical supplement to this article. We obtained the results in the next section from 100,000 runs of the Gibbs sampler after a burn-in of 20,000 runs.

V. Results

Our benchmark specification corresponding to the EKC hypothesis assumes a quadratic relationship between income and pollution, omitting all nonincome covariates except "distance of the monitor from the nearest highway" (CO and O_3) and "number of electric utilities in EPA region" (PM_{10}), which are unlikely to be correlated with income. We report the coefficient estimates obtained under this specification in column 1 of Tables 3a–c. Because we are also interested in the pure income effect, we report the coefficient estimates for the pollution-income relationship obtained when all other covariates that are correlated with household income are included (column 2 of Tables 3a-c).

A commonly used estimator of the turning point of the quadratic income-pollution relationship is the ratio of the {estimator of the coefficient of income} and {minus 2 times the estimator of the coefficient of income squared}. However, the common

Table 3a
Analysis of Carbon Monoxide (CO)

	(1)	(2)	(3)	(4)
ln(income)[1]	1.2272	–3.3431	.3112	–1.5436
	(5.5665)	(6.4687)	(1.0682)	(11.2161)
ln(income)[2]	–.1080	.1931		–.1650
	(.2821)	(.3463)		(1.6210)
ln(income)[3]				.0180
				(.0751)
ln(population density)		1.5549***	1.4147***	1.4287***
		(.4916)	(.3252)	(.3504)
ln(percentage minority)		.6179	.5988	.6062*
		(.4890)	(.4173)	(.4179)
ln(percentage unemployment)		.1365	.1268	.1730
		(.7473)	(.6481)	(.6674)
ln(percentage employed in manufacturing)		.0603	.0522	.0710
		(.5312)	(.4610)	(.4751)
ln(percentage HS grad)		.0809	.2331	.2610
		(1.6686)	(1.5480)	(1.5099)
ln(percentage voters)		–.5756	–1.2673	–.8735
		(2.5253)	(2.2233)	(2.3865)
ln(percentage renters)		.7992	.7022	.9127
		(.9386)	(.7814)	(.8600)
ln(percentage below poverty)		.0060	.0155	.0508
		(.8118)	(.7454)	(.7551)
ln(percentage female head of household)		–.4031	–.4893	–.3853
		(.6269)	(.5403)	(.5768)
ln(percentage older than 65)		6.7004	6.1588	6.7028
		(5.6148)	(4.9465)	(5.3685)
Distance from highway	–.8170***	–.5490***	–.5279***	–.5278***
	(.2645)	(.2237)	(.1842)	(.1983)
Turning Point (TP)[a]	8.9616	9.0594		9.2136
	$7,797	$8,599		$10,032
Cov (β_1, β_2)	–1.5643	–2.2037		
Quantiles:				
2.5%	–13.2935	–9.7574		
5.0%	–3.2398	–.6466		
10.0%	3.3745	4.3006		
50.0%	8.9616	9.0594	N/A	N/A[b]
90.0%	14.9193	13.2683		
95.0%	20.6795	17.6554		
97.5%	32.7041	24.8578		

Note: Standard errors are shown in parentheses. Estimates of intercepts and regional dummies are not shown.
a. The turning point estimates are the modes of the empirical distributions of the turning point estimators. The turning point of the cubic relationship is the turning point of the peak.
b. For the cubic relationship in column 4, we were unable to obtain estimates of the distributions of the turning points because the majority of the runs of the Gibbs sampler implied complex roots for the cubic equation.
* indicates that the coefficient is different from zero at the 90% level. ** indicates that the coefficient is different from zero at the 95% level. *** indicates that the coefficient is different from zero at the 99% level.

Table 3b
Analysis of Ozone (O_3)

	(1)	(2)	(3)	(4)
ln(income)[1]	5.8727***	4.8198	.5132**	–3.837
	(2.9074)	(3.1234)	(.2774)	4 (9.2900)
ln(income)[2]	–.2590**	–.2137*		.7258
	(.1438)	(.1550)		(.9797)
ln(income)[3]				–.0335
				(.0352)
ln(population density)		.0020	–.0084	.0028
		(.0356)	(.0353)	(.0358)
ln(percentage minority)		.0159	.0082	.0222
		(.0677)	(.0681)	(.0698)
ln(percentage unemployment)		–.0396	–.0413	–.0601
		(.1371)	(.1350)	(.1376)
ln(percentage employed in manufacturing)		–.0186	.0159	–.0281
		(.1054)	(.1052)	(.1084)
ln(percentage HS grad)		.0385	.0487	–.0252
		(.2880)	(.2936)	(.2911)
ln(percentage voters)		–.9046***	–.8801***	–.8634**
		(.4523)	(.4551)	(.4488)
ln(percentage renters)		–.1650	–.1229	–.1815
		(.1390)	(.1355)	(.1423)
ln(percentage below poverty)		–.0881	–.1264	–.1009
		(.1415)	(.1401)	(.1395)
ln(percentage female head of household)		.3156***	.3856***	.3441***
		(.1440)	(.1407)	(.1427)
ln(percentage older than 65)		–1.3400	–1.2804	–1.3162
		(1.0808)	(1.0764)	(1.0859)
Distance from highway	.0670**	.0771**	.0822***	.0761**
	(.0377)	(.0410)	(.0409)	(.0408)
Turning Point (TP)[a]	11.2811	11.0416		10.8790
	$79,308	$62,417		$53,050
Cov (β_1, β_2)	–.4177	–.4868		
Quantiles:				
2.5%	6.6825	.8962		10.0909
5.0%	10.5652	5.5744		10.2304
10.0%	10.7046	9.9283		10.3613
50.0%	11.2811	11.0416	N/A	10.8790
90.0%	13.4576	13.9402		12.0522
95.0%	15.6850	16.8582		12.9040
97.5%	18.6154	23.5303		14.2135

Note: Standard errors are shown in parentheses. Estimates of intercepts and regional dummies are not shown.
a. The turning point estimates are the modes of the empirical distributions of the turning point estimators. The turning point of the cubic relationship is the turning point of the peak.
* indicates that the coefficient is different from zero at the 90% level. ** indicates that the coefficient is different from zero at the 95% level. *** indicates that the coefficient is different from zero at the 99% level.

Table 3c
Analysis of Particulate Matter (PM_{10})

	(1)	(2)	(3)	(4)
ln(income)[1]	7.3574*	8.8347**	.3592	–9.5380
	(5.0883)	(5.8701)	(.8195)	(11.0845)
ln(income)[2]	–.4586**	–.4536*		2.3760*
	(.2595)	(.3068)		(1.5804)
ln(income)[3]				–.1268**
				(.0725)
ln(population density)		.0368	.0497	.0741
		(.0986)	(.1001)	(.1017)
ln(percentage minority)		–.0001	–.0663	.0007
		(.2537)	(.2441)	(.2476)
ln(percentage unemployment)		.2869	.4603	.3723
		(.4629)	(.4402)	(.4448)
ln(percentage employed in manufacturing)		.0658	.0287	–.0087
		(.3156)	(.3301)	(.3314)
ln(percentage HS grad)		–.4489	–.2387	–.4338
		(.5244)	(.5340)	(.5304)
ln(percentage voters)		–4.6178***	–3.9503***	–4.3518***
		(1.6914)	(1.6376)	(1.6622)
ln(percentage renters)		–.6041	–.4627	–.5513
		(.4892)	(.4781)	(.4816)
ln(percentage below poverty)		.9486*	1.2972***	1.1741**
		(.5910)	(.5686)	(.6428)
ln(percentage female head of household)		.1883	.1272	–.1962
		(.4456)	(.4279)	(.4407)
ln(percentage older than 65)		4.9577	5.2233*	5.8717*
		(3.8155)	(3.6245)	(3.8050)
Distance from highway	.3048	.2942	.1043	.1741
	(.3370)	(.3558)	(.3604)	(.1741)
Turning Point (TP)[a]	8.1012	9.7439		9.9796
	$3,298	$17,049		$21,582
Cov (β_1, β_2)	–1.3166	–1.7567		
Quantiles:				
2.5%	–4.5745	4.5319		9.1095
5.0%	2.5437	7.1674		9.3167
10.0%	5.1199	8.2269		9.5066
50.0%	8.1012	9.7439	N/A	9.9756
90.0%	8.9479	11.4660		10.5803
95.0%	9.1752	13.0938		10.9419
97.5%	17.5628	16.0076		11.5808

Note: Standard errors are shown in parentheses. Estimates of intercepts and regional dummies are not shown.
a. The turning point estimates are the modes of the empirical distributions of the turning point estimators. The turning point of the cubic relationship is the turning point of the peak.
* indicates that the coefficient is different from zero at the 90% level. ** indicates that the coefficient is different from zero at the 95% level. *** indicates that the coefficient is different from zero at the 99% level.

assumption of (asymptotic) normality of the income coefficients implies that the distribution of this estimator is generally asymmetric and does not have a mean (see Plassmann & Khanna, 2002). We therefore estimate the turning point as the median of the empirical distribution of the turning point estimator, rather than as the ratio of the {coefficient of income} and {minus 2 times the coefficient of income squared}. At the bottom of Tables 3a–c, we show the quantiles of the empirical distributions of the turning point estimators. The quantiles corresponding to the 2.5th and the 97.5th percentile indicate the 95% confidence intervals.[10]

The EKC Relationship

Our benchmark quadratic specification for CO (column 1 of Table 3a) yields a turning point at the extreme lower end of the income data range. The estimated quantiles of the distribution of the turning point estimator show that the turning point estimate is highly imprecise, which indicates that the apparent nonmonotonicity is not an integral part of the estimated relationship. We obtained the same result with a cubic relationship that does not impose the strict symmetry of the quadratic functional form. Neither of the two income coefficients is significant, and we conclude that there is no evidence of correlation between CO exceedances and household income.

The estimated turning point from our benchmark specification for O_3 (column 1 of Table 3b) lies toward the right end of the income distribution. Although this turning point estimate is more precise than that obtained for CO, the confidence interval covers the entire range of the income data. Nevertheless, both income coefficients are strongly statistically significant. We conclude that there is some evidence of a concave relationship between income and O_3 exceedances, but we do not have sufficient data from census tracts with higher income levels to determine whether O_3 pollution will ultimately decrease with income.

In the case of PM_{10}, we also obtain a fairly imprecise estimate of the turning point, with the 95% confidence interval of the turning point estimator spanning more than the entire sample income range. However, because the estimated turning point ($3,298) lies at the lower end of our sample income range and the two income coefficients are statistically significant, we primarily estimate the right leg of the EKC relationship. We conclude that PM_{10} exceedances fall with household income.

The Pure Income Effect

To investigate the role of consumer preferences and the demand for environmental quality, we analyze the pollution-income relationship when income alone increases. In column 2 of Tables 3a-c, we report the results that we obtained when we included the other nonincome covariates. Adding these additional covariates does not affect our main conclusion with regard to the pollution-income relationship in the case of CO and O_3, but for PM_{10}, the results are noticeably different compared to the overall EKC relationship.

For CO, the precision of the turning point estimate of the pure pollution-income relationship is somewhat higher than that of the overall EKC relationship, but the 95% confidence interval still exceeds the sample income range by a wide margin. As before, the income coefficients are statistically insignificant and there is no evidence of correlation between income and CO exceedances. Motor vehicles are the main source of CO emissions, and the income elasticity of transportation (vehicle miles traveled) ranges from .5 to 1 (see Agras & Chapman, 1999). Although richer households consume more vehicle miles, they also tend to drive newer and more fuel-efficient vehicles. One interpretation of our result is that for higher income households, the increase in CO pollution caused by greater vehicle use tends to be offset by the fuel efficiency gains of newer cars.

Our analysis also suggests that CO pollution increases with population density and decreases with the distance of the EPA monitor to the nearest road. These results are intuitive because motor vehicles are the main source of CO emissions and densely populated areas tend to have high vehicle density as well. None of the other covariates is statistically significant.

The estimated turning point in the pure income relationship for O_3 (11.04 or \$62,317) is fairly close to the turning point obtained from the general EKC relationship (11.28 or \$79,221) and lies toward the right end of the sample income range. As before, we conclude that the estimated relationship is positive and likely to be concave but that our results do not necessarily imply a nonmonotonic relationship. A possible explanation lies in the fact that O_3 is formed by a chemical reaction of VOCs and nitrogen oxides (NO_x). The main sources of VOCs are chemical plants, refineries, and motor vehicles, and the main source of NO_x is fuel combustion (power plants, heating, and motor vehicles). Although more transportation tends to be associated with higher levels of these precursors of O_3, it is likely that households with higher incomes will reduce their exposure to O_3 by distancing themselves from the nonmobile sources of VOCs and NO_x. The relationship between O_3 and its precursors is highly nonlinear, and our results suggest that even greater reduction in these precursors is required for the O_3-income relationship to turn downward.

We also find that O_3 concentrations increase with the distance of the monitor from the closest highway. This suggests that most of the measured O_3 pollution is due to sources other than motor vehicles that tend to be located in areas with low population and highway density.[11]

PM_{10} is the only pollutant for which we obtain clear evidence for a nonmonotonic relationship with household income that is consistent with the EKC hypothesis. The quadratic specification (column 2 of Table 3c) yields a turning point of 9.74 (or \$17,049) that is fairly close to the median household income in our sample (10.25 or \$28,397). The quantiles at the bottom of Table 3c imply a relatively small confidence interval around the turning point estimate, suggesting that this turning point is more than just a functional artifact. Most of the anthropogenic emissions of PM_{10} are the result of nontransportation fuel combustion and industrial processing. Because it is possible to reduce one's exposure to PM_{10} pollution by moving away from polluting industries, it is reasonable that our results indicate an inverted U-shaped relationship for this pollutant. As the trade-off between finding a job at or near a polluting industry and being exposed to higher levels of pollution becomes less and less attractive as income increases, higher income households simply relocate. The statistically significant positive coefficient on the percentage of population below the poverty line supports this interpretation.

The percentage of registered voters has a strong negative effect on O_3 and PM_{10} pollution. This is consistent with the results of Brooks and Sethi (1997), who found that more politically active communities tend to be exposed to lower pollution.[12] Although one might expect a statistically significant negative coefficient in the case of CO, ambient CO concentrations are highly correlated with local emissions (EPA, 1997, 2002), so that reducing local CO pollution entails reducing emissions locally from its primary source (gasoline-powered vehicles). This has a high private cost because it necessitates lifestyle changes, such as a shift from private to public modes of transportation, or the adoption of alternative technologies (e.g., hybrid vehicles), which are fairly expensive. Political action is unlikely to be successful in reducing pollution in this situation. PM_{10} and O_3, on the other hand, have several point and nonpoint sources and it is possible to lower exposure to these pollutants through political action that keeps point sources out of a neighborhood.

VI. Tests of Robustness

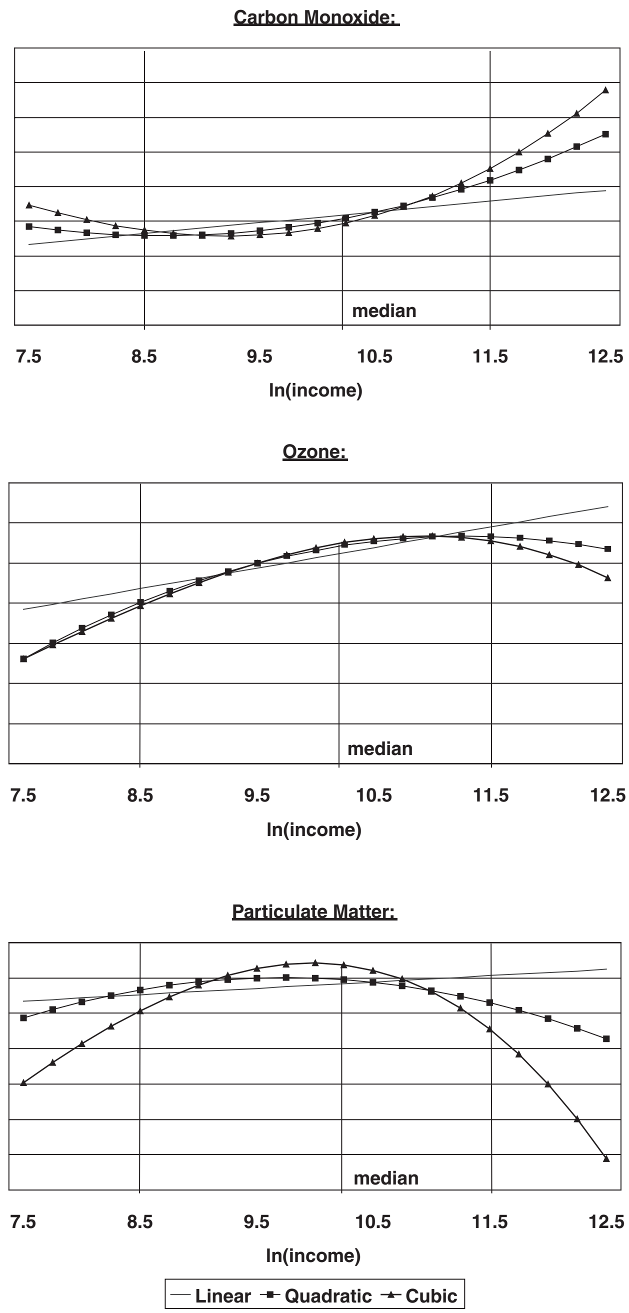
The EKC literature reports evidence to suggest that estimated pollution-income relationships are not very robust. We test the robustness of our estimates by comparing three different model specifications—linear, quadratic, and cubic—with respect to the pure income effect and report the results for these specifications in columns 2 through 4 of Tables 3a-c. Figures 1a-c contain the graphs of the estimated pollution-income relationships for the three functional forms for each pollutant.

All three analyses suggest that the incidence of CO pollution above its NAAQS is (mildly) increasing with income. The linear relationship is positive but not statistically significantly different from zero. The (second) turning point of the cubic equation is at a slightly higher income level than that of the quadratic equation, but both are to the left of more than 95% of the data. There is no indication of concavity at high levels of income.

In the case of O_3, the linear analysis suggests a statistically significantly positive relationship, and the quadratic and cubic models indicate that the growth rate of O_3 concentrations falls with income. The estimated turning points lie at the right end of the income distribution and the confidence intervals cover the whole range of the income data. All three analyses lead us to conclude that there is a monotonic and concave relationship between income and O_3.

The graph of the cubic analysis for PM_{10} shows that pollution increases with income as long as income is relatively low and that it falls more sharply beyond the turning point. The quadratic specification that imposes symmetry around the turning point yields a turning point estimate at a lower income level than the cubic specifica-

Figure 1
Comparison of Linear, Quadratic, and Cubic Pollution-Income Relationships



tion, which also suggests an asymmetry in slopes before and after the turning point. Based on the qualitative similarity of the results, we conclude that there is strong evidence of a nonmonotonic concave (inverted U-shaped) relationship between household income and PM_{10} pollution.

In the technical supplement to this article (available on the journal's Web site), we report two additional sets of specification and robustness tests. First, we assess the validity of our distributional assumption (normal, Poisson, negative binomial models) and our dependence assumption (individual versus joint analysis of the three pollutants). We find strong evidence that the joint Poisson-lognormal model is more appropriate for our data than any of the alternative models.

Second, we test the robustness of our results by repeating our analysis using the 1990 ambient concentrations of the three pollutants to determine the effect of left-censoring the data. This data set contains 509 sites with measurements of CO, 820 sites with measurements of O_3, and 1,331 sites with measurements of PM_{10}, but only 115 sites with measurements of all three pollutants. Rather than analyzing the pollutants jointly with fewer observations, we analyze them separately using all observations. The analysis of the additional data yields very similar relationships between household income and the three pollutants, and we conclude that our results are robust with respect to changes in the data set.

VII. Conclusion

The EKC hypothesis suggests a nonmonotonic relationship between economic growth and pollution. To the extent that this relationship is influenced by consumer preferences that can vary spatially and intertemporally, it is unlikely that there is a single global EKC so that analyses that use multicountry panel data sets suffer from aggregation bias. We conclude that it is necessary to first estimate local (regional or country-specific) relationships before attempting to construct a global EKC. Our analysis of cross-sectional census-tract-level data for the United States suggests that the income level at which households are willing to reduce their exposure to pollution depends on the nature of pollution. We find evidence for a nonmonotonic relationship between income and exposure to PM_{10}, but little evidence for such a relationship for CO and O_3.

From a policy perspective, it is important to know whether changes in income alone and without the compounding effects of associated changes, such as a change in the structural composition of the economy or the level of technology, can yield an inverted U-shaped pollution-income relationship. We find an inverted U-shaped pure income relationship between household income and PM_{10} with a peak at about \$20,000. PM_{10} is a point-source pollutant, and it is fairly straightforward and relatively inexpensive to reduce exposure to PM_{10} by relocating, without necessarily reducing global emissions. We do not find much evidence of an inverted U-shaped relationship between household income and the two non-point-source pollutants, O_3 and CO, that are generated primarily by the use of gasoline-powered transportation. The private abatement costs for O_3 and CO pollution are relatively high. Even in a country with one of the highest incomes in the world, household income has not yet reached the level beyond which the pollution-income relationship for such pollutants becomes negative.

Although economic growth and increases in household income are usually positively correlated, an analysis of the pure income relationship between household income and pollution (when only household income increases) has different policy implications than an analysis of the overall relationship between increases in income and pollution. The commonly estimated EKC relationship outlines the pollution-income trajectory as an economy passes through different stages of development and experiences a changing sectoral composition, unemployment rates, education levels, income and population distributions, and so on, along with a higher per capita income. The pure income effect emphasizes the effect of increasing household or per capita income without any change in the other variables typically associated with economic development. Our analysis suggests that economic growth by itself may be insufficient to generate a downturn in the pollution-income trajectory, especially for the pollutants with a high private abatement cost.

Most EKC studies using CO_2 data have not found any evidence of a nonmonotonic pollution-income relationship. For a global pollutant such as carbon dioxide (CO_2), whose abatement is likely to require a significant shift in lifestyle patterns and that has a high private abatement cost and a low private abatement benefit, it is unlikely that the pollution-income trajectory will become negative solely as the result of economic growth. It is much more likely that reducing global CO_2 emissions will require significant strides in economic development accompanied by a change in consumer attitudes toward pollution and a change in the state of technology as well.

Notes

1. Barbier (1997, p. 370) and Carson, Jeon, and McCubbin (1997, p. 434) have argued that most explanations for the EKC, such as changes in technology, civil and political liberties, trade policy, and environmental policy, are simply the vehicles that enable consumers to reveal their preferences for environmental quality.

2. Selden and Song (1994), Holtz-Eakin and Selden (1995), Stern and Common (2001), and Halkos and Tsionas (2001) use panel data on national emissions. Grossman and Krueger (1992, 1995), Torras and Boyce (1998), Barrett and Graddy (2000), and Harbaugh et al. (2002) use panel data on local concentrations.

3. For example, in many regions in the United States, it is common to hear the local O_3 rating based on the NAAQS and the EPA's Air Quality Index along with the daily weather forecast.

4. Other studies that use cross-section data for the United States include Carson et al. (1997), who analyze state-level air pollution data, Kahn (1998), who analyzes California vehicular emissions data, Berrens, Bohara, Gawande, and Wang (1997) and Gawande, Bohara, Berrens, and Wang (2000), who analyze county and metropolitan statistical area data for hazardous waste facilities, and Khanna (2002) and Khanna and Plassmann (2004), who analyze ambient air pollutant concentrations data for census tracts.

5. We obtained county-level election data from Election Data Services, Inc., which does not report voter turnout data for many jurisdictions. We therefore could not use the ratio of voter turnout to voter registration to capture collective action. North Dakota does not require voter registration, and we used the ratio of voter turnout to voting age population. The data set does not include information on Wisconsin and Alaska, and we predicted the log of the voter registration rate in these two states with an auxiliary regression, using county-level data for the entire United States.

6. The random effects in the Poisson-lognormal model require neither the assumption that the ε_{pi} are realizations from the same distribution nor that they are uncorrelated with the other covariates, the two main objections that are frequently raised against the least squares random-effects model. The Poisson-lognormal model easily accommodates observation-specific distributions as well as correlations between the ε_{pi} and the other covariates.

7. Aitchison and Ho (1989) show the derivation of the moments of the multivariate Poisson-lognormal distribution. Winkelmann (2000, pp. 182-184) describes the multivariate Poisson-lognormal regression model.

8. Recall that the covariance matrix of the latent effects measures the covariances of the concentrations as well as the covariances of variables that were omitted from the three regression equations. The estimates in Table 2 do not necessarily imply that all pollutants are negatively correlated, because the estimates might be driven by negative correlations of the effects of omitted variables.

9. MCMC methods are iterative techniques that use Markov chains to perform Monte Carlo integrations of integrals of interest. For details on MCMC methods and their applications to count data, see Casella and George (1992), Gilks, Richardson, and Spiegelhalter (1996), Chib and Greenberg (1996), Gamerman (1997), and Chen, Shao, and Ibrahim (2000).

10. A common method for assessing the precision of the turning point estimate is to use a normal approximation of the turning point estimator, the "delta method" (e.g., Greene, 2002, pp. 913-914). The asymmetry of the distributions of turning point estimators of polynomial regression functions makes the normal approximation inappropriate and implies that it is not possible to derive (asymptotic) 95% confidence intervals of the estimates by adding and subtracting 1.96 times the estimated standard error.

11. Alternatively, it is possible that O_3 concentrations are lower near highways because NO emissions from vehicles destroy it via the titration effect, whereas in areas further away from highways, the NO_x emissions may reproduce O_3 via photochemical reactions. We thank an anonymous referee for suggesting this alternative explanation.

12. Brooks and Sethi (1997) analyze the relationship between community characteristics and exposure to aggregate emissions of more than 300 chemicals reported in the EPA's Toxics Release Inventory.

References

Agras, J., & Chapman, D. (1999). The Kyoto Protocol, CAFE standards, and gasoline taxes. *Contemporary Economic Policy*, *17*(3), 296-308.

Aitchison, J., & Ho, C. H. (1989). The multivariate Poisson-lognormal distribution. *Biometrika*, *76*(4), 643-653.

Aldy, J. E. (2005). An Environmental Kuznets Curve analysis of U.S. state-level carbon dioxide emissions. *The Journal of Environment & Development*, *14*(1), 48-72.

Barbier, E. B. (1997). Introduction to the Environmental Kuznets Curve special issue. *Environment and Development Economics*, *2*, 369-381.

Barrett, S., & Graddy, K. (2000). Freedom, growth, and the environment. *Environment and Development Economics*, *5*(4), 433-456.

Berrens, R. P., Bohara, A. K., Gawande, K., & Wang, P. (1997). Testing the inverted-U hypothesis for US hazardous waste: An application of the generalized gamma model. *Economics Letters*, *55*, 435-440.

Brooks, N., & Sethi, R. (1997). The distribution of pollution: Community characteristics and exposure to air toxics. *Journal of Environmental Economics and Management*, *32*(2), 232-250.

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

Casella, G., & George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, *46*(3), 167-174.

Chen, M-H., Shao, Q-M., & Ibrahim, J. G. (2000). *Monte Carlo methods in Bayesian computing*. New York: Springer Verlag.

Chib, S., & Greenberg, E. (1996). Markov chain Monte Carlo simulation methods in econometrics. *Econometric Theory*, *12*, 409-431.

Chib, S., & Winkelmann, R. (2001). Markov chain Monte Carlo analysis of correlated count data. *Journal of Business and Economic Statistics*, *19*, 428-435.

Chimeli, A. B., & Braden, J. B. (2005). Total factor productivity and the Environmental Kuznets Curve. *Journal of Environmental Economics and Management*, *49*, 366-380.

Gamerman, D. (1997). *Markov chain Monte Carlo. Stochastic simulation for Bayesian inference*. London: Chapman and Hall.

Gawande, K., Bohara, A. K., Berrens, R. P., & Wang, P. (2000). Internal migration and the Environmental Kuznets Curve for US hazardous waste sites. *Ecological Economics*, *33*, 151-166.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice*. London: Chapman and Hall.

Greene, W. H. (2002). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Grossman, G. M., & Krueger, A. N. (1992, February). *Environmental impacts of a North American Free Trade Agreement* (Discussion Paper in Economics No. 158). Princeton, NJ: Woodrow Wilson School, Princeton University.

Grossman, G. M., & Krueger, A. N. (1995). Economic growth and the environment. *Quarterly Journal of Economics*, *112*(3), 353377.

Halkos, G. E., & Tsionas, E. G. (2001). Environmental Kuznets Curves: Bayesian evidence from switching regime models. *Energy Economics*, *23*, 191-210.

Harbaugh, W. T., Levinson, A., & Wilson, D. M. (2002). Reexamining the empirical evidence for an Environmental Kuznets Curve. *Review of Economics and Statistics*, *84*(3), 541-551.

Holtz-Eakin, D., & Selden, T. M. (1995). Stoking the fires? CO_2 emissions and economic growth. *Journal of Public Economics*, *57*, 85-101.

Kahn, M. E. (1998). A household level Environmental Kuznets Curve. *Economics Letters*, *59*, 269-273.

Khanna, N. (2002). The income elasticity of non-point source air pollutants: Revisiting the Environmental Kuznets Curve. *Economics Letters*, *77*(3), 387-392.

Khanna, N., & Plassmann, F. (2004). The demand for environmental quality and the Environmental Kuznets Curve hypothesis. *Ecological Economics*, *51*(3-4), 225-236.

Leifman, M., & Heil, M. (2005). Guest editors' note. *The Journal of Environment & Development*, *14*(1), 7-26.

Lieb, C. (2002). The Environmental Kuznets Curve and satiation: A simple static model. *Environment and Development Economics*, *7*, 429-448.

List, J. A., & Gallet, C. A. (1999). The Environmental Kuznets Curve: Does one size fit all? *Ecological Economics*, *31*(3), 409-423.

Marcotullio, P. J., Williams, E., & Marshall, J. D. (2005). Faster, sooner, and more simultaneously: How recent road and transportation CO_2 emission trends in developing countries differ from historic trends in the United States. *The Journal of Environment & Development*, *14*(1), 125-148.

Millimet, D. L., List, J. A., & Stengos, T. (2003). The Environmental Kuznets Curve: Real progress or misspecified models? *Review of Economics and Statistics*, *85*(4), 1038-1047.

Plassmann, F., & Khanna, N. (2002). *Assessing the precision of turning point estimates in polynomial regression functions* (Department of Economics Working Paper WP0213). Binghamton: State University of New York at Binghamton.

Plassmann, F., & Khanna, N. (in press). Preferences, technology and the environment: Understanding the Environmental Kuznets Curve hypothesis. *American Journal of Agricultural Economics*.

Selden, T. M., & Song, D. (1994). Environmental quality and development: Is there a Kuznets Curve for air pollution emissions? *Journal of Environmental Economics and Management*, *27*, 147162.

Stern, D. I. (2005). Beyond the Environmental Kuznets Curve: Diffusion of sulfur-emissions-abating technology. *The Journal of Environment & Development*, *14*(1), 101-124.

Stern, D. I., & Common, M. S. (2001). Is there an Environmental Kuznets Curve for sulfur? *Journal of Environmental Economics and Management*, *41*(2), 162-178.

Torras, M., & Boyce, J. K. (1998). Income, inequality, and pollution: A reassessment of the Environmental Kuznets Curve. *Ecological Economics*, *25*(2), 147-160.

U.S. Environmental Protection Agency (EPA). (1997, October). *The benefits and costs of the Clean Air Act, 1970-1990* (EPA-410-R-97-002). Washington, DC: Office of Air and Radiation, Office of Policy, Planning and Evaluation.

U.S. Environmental Protection Agency (EPA). (2002, September). *Latest findings on national air quality: 2001 status and trends* (EPA 454/K-02-001). Washington, DC: Author.

Winkelmann, R. (2000). *Econometric analysis of count data* (3rd ed.). New York: Springer Verlag.

Florenz Plassmann is associate professor of economics at Binghamton University (SUNY). His primary research interests are in public economics and applied microeconomics.

Neha Khanna is associate professor of Economics and Environmental Studies at Binghamton University (SUNY). Her research areas include the Environmental Kuznets Curve hypothesis, the world oil market and its implications for international security, the efficacy of voluntary pollution prevention programs, and climate change policy.