One Shape Does Not Fit All: A Nonparametric Instrumental Variable Approach to Estimating the Income-Pollution Relationship at the Global Level[1]

C.-Y. Cynthia Lin Lawell,[2] Krishna P. Paudel, Mahesh Pandit

Abstract

We examine the relationships among water pollution, income, and political institutions using country-level global water quality data over the period 1980 to 2012. In order to address concerns about the highly nonlinear relationship between pollution and income, the endogeneity of income, and the discrete nature of political variables, we use a nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables to identify these relationships. Results indicate an inverted-U shaped relationship between pollution and income consistent with an environmental Kuznets curve for one pollutant (lead), a cubic shape for three pollutants (nickel, mercury, and fecal coliform), and more highly nonlinear relationships for many of the other pollutants. For several stock pollutants (nickel, mercury, and arsenic), we find that pollution levels may continue to increase at higher levels of income, suggesting that stock pollutants may continue to accumulate in productive high-income countries even when marginal emissions have been reduced. We also find suggestive evidence that levels of pollutants resulting from industrial activity (e.g., chemical oxygen demand) may increase with income while those that are more driven by residential activity and population levels (e.g., fecal coliform) do not. By estimating a nonparametric relationship between pollution and political institutions and by accounting for the categorical nature of the political variables, we are able to detect a nonlinear relationship between pollution and political institutions as well, which for some pollutants is an inverted-U shaped curve.

Keywords: binary variable, environmental Kuznets curve, nonparametric instrumental variable regression, water pollution

***JEL* Codes:** Q53, C14

This draft: January 2018

---

[1] Lin Lawell is an associate professor and the Robert Dyson Sesquicentennial Chair in Environmental, Energy and Resource Economics in the Charles H. Dyson School of Applied Economics and Management at Cornell University, Ithaca, New York. Paudel and Pandit are Gilbert Durbin Endowed Professor and former doctoral student, respectively, in the Department of Agricultural Economics and Agribusiness at Louisiana State University (LSU) and LSU Agricultural Center, Baton Rouge, Louisiana. Senior authorship is shared equally by Lin Lawell and Paudel.

[2] Corresponding Author: Charles H. Dyson School of Applied Economics and Management, Cornell University, 407 Warren Hall, Ithaca, NY 14853-4203. Phone: (607) 255-4707. Fax: (607) 255-9984. Email: clinlawell@cornell.edu

1. Introduction

The environmental Kuznets curve (EKC) is a relationship between income and pollution which is hypothesized to have an inverted-U shape. The EKC hypothesis stems from Kuznets' famous work in income equality and poverty (Kuznets, 1955). The EKC hypothesis states that as income increases, pollution goes up initially but when income is high enough, pollution eventually declines. The income level at which pollution level is the highest is called a turning point.

There are numerous papers on the validity, application, and measurement of the EKC (Azomahou, Laisney and Nguyen Van, 2006; Carson, 2010). This is evident from the seminal work of Grossman and Krueger (1995), as well as from papers focusing specifically on air pollution (e.g., Hilton and Levinson, 1998; List and Gallet, 1999; Heerink, Mulatu and Bulte, 2001; Bruvoll and Medin, 2003; Deacon and Norman, 2006; Merlevede, Verbeke and Clercq, 2006; Liddle and Messinis, 2015; Liddle, 2015), water pollution (e.g., Jha and Murthy, 2003; Paudel and Schafer, 2009; Paudel, Zapata and Susanto, 2005; Lin and Liscow, 2013), deforestation (e.g., Heerink et al., 2001; Rodriguez-Meza, Southgate and Gonzalez-Vega, 2004; Barbier, 2004; Culas, 2007), hazardous waste and toxins (e.g., Gawande, Berrens and Bohara, 2001; Rupasingha et al., 2004), and carbon dioxide (e.g., Copeland and Taylor, 2004; Azomahou et al., 2006; Plassmann and Khanna, 2006; Poudel et al., 2009). However, critics have challenged both the findings and policy implications of these studies (Dasgupta et al., 2002; Stern, 2008). Some suggest that the pollutant-income relationship differs depending on the choice of the pollutant, study area, and time period (Harbaugh, Levinson and Wilson, 2002).

One strand in the EKC literature posits that there may be a political mechanism underlying the EKC relationship. These papers suggest that what cleaned up the environment was not rising income, but rather political institutions responding to public demand (Lomborg and Pope, 2003).

For example, Grossman and Krueger (1995) speculate that the strongest link between income and pollution in fact is via an induced policy response, and that these policies are in turn induced by popular demand. According to this line of reasoning, poor countries, at first, have so little development that they have high environmental quality. Then, countries' environments degrade as they develop and become richer. Finally, they reach a point at which environmental quality is poor enough and the people are rich enough that they desire to pay for improvements in environmental quality. At this point, they begin to demand changes from their government, and environmental degradation decreases. Similarly, Dasgupta and Mäler (1995) indicate that political rights and civil liberties are important components in protecting environmental rights. Congleton (1992) finds that political institutional arrangements, rather than resource endowments, largely determine policies concerning environmental regulation.

The importance of political institutions in the EKC relationship has also been examined empirically in papers that include political variables in addition to income in the EKC regression. Barrett and Graddy (2000) find that, for many pollution variables, "political reforms may be as important as economic reforms in improving environmental quality worldwide" (p. 433). However, they also find an absence of significant results for some pollution variables, which suggests that something other than an induced policy response may be affecting pollution levels. Lin and Liscow (2013) find that political institutions have a significant effect on environmental quality for five of the eleven water pollutants that they examine. Torras and Boyce (1998) hypothesize that changes in the distribution of power underlie the EKC relationship, and find that literacy, political rights, and civil liberties have particularly strong effects on environmental quality in low-income countries. Farzin and Bond (2006) develop and estimate an econometric model of the relationship between several local and global air pollutants and economic development while allowing for critical aspects of the sociopolitical-economic regime of a state. They find that democracy and freedom act as a conduit for people to demand for better environmental quality. Nguyen Van and Azomahou (2007) indicate that institutional failures (low political rights and civil liberties) worsen the deforestation process in developing countries. Galinato and Galinato (2012) examine the effects of political stability, corruption control, and economic growth on CO_2 emissions from deforestation, and find that political stability and corruption control do not significantly affect the income turning point, but both variables shift the forest-income curve up or down.[3]

Since different political mechanisms for the EKC may have different implications for people's marginal "willingness to pay" for environmental protection, Israel and Levinson (2004) develop and implement an empirical test for different political mechanisms and apply it to international survey data from the World Value Survey. Their results show strong relationships between marginal willingness to pay and individual characteristics, such as age, income, and education, but little evidence that marginal willingness to pay varies systematically with economic growth.

In the traditional EKC relationship, a parametric model is used to analyze the relationship between pollution levels and per capita income, where the pollution level is regressed on a quadratic or cubic function of per capita income. These types of ad hoc functional form specifications put an *a priori* restriction on how the relationship should look like in the empirical

[3] A related concept to political institutions that may need to be accounted for in the EKC relationship is social capital. Social capital is defined as shared norms, trust, and social networks that facilitate coordination and cooperation for mutually beneficial collective action. An example of social capital is membership in environmental groups. Paudel and Schafer (2009) and Paudel et al. (2011) include a social capital index in the EKC model.

estimation. However, pollutants are not created equally and different pollutants have different relationships with income (Dasgupta et al., 2002).

For example, stock pollutants like mercury keep accumulating in water bodies unless a substitute product is used in the production process or measures are taken to remove these stock pollutants from water bodies. It is therefore possible with stock pollutants that pollution levels may continue to increase at higher levels of income even as measures are taken to reduce emissions, since the stock pollutant may continue to accumulate in productive high-income countries even when marginal emissions have been reduced.

Moreover, the source and potential abaters of different pollutants may also affect their relationship with income. For example, industries may focus more on addressing chemical oxygen demand, while municipalities may focus more on addressing biological oxygen demand (Kemira, 2017). When incomes increase and the people are rich enough that they desire to pay for improvements in environmental quality, it is possible, for example, that these improvements come more from reductions in residential pollution rather than from industrial pollution. Likewise, agricultural pollutants such as phosphorus, dissolved oxygen,[4] and variants of nitrogen may have a different relationship with income than fecal coliform-related pollutants do.

Because different pollutants have different relationships with per capita income, it is therefore important to relax the assumption of a one size fit all functional form for pollutants. The

---

[4] Dissolved oxygen is an important parameter in assessing water quality because of its influence on the organisms living within a body of water. A dissolved oxygen level that is too high or too low can harm aquatic life and affect water quality (Fondriest Environmental, Inc., 2013). Agricultural practices may affect dissolved oxygen via channel alteration (e.g., straightening or deepening of streams); impoundments; the introduction of nutrients (e.g., fertilizers, animal wastes), chemical contaminants (e.g., heavy metals), or organic matter (e.g., sewage, animal wastes) to streams; or interactions with other stressors (EPA, 2016).

best approach may instead be a data driven nonparametric model to identify the relationship between income and various pollutants. Another advantage of nonparametric estimates is that they are more robust in detecting structures that sometimes remain undetected by traditional parametric estimation techniques.

Several authors (Millimet, List and Stengos, 2003; Paudel et al., 2005; Poudel, Paudel and Bhattarai, 2009; Zapata and Paudel, 2009) have refuted the parametric EKC model and suggest a need to include a nonparametric form of income in the regression. These semiparametric forms were found to perform better than the parametric form in specification tests. Although semiparametric and nonparametric methods are tedious, these methods are used by many researchers (Schmalensee, Stoker and Judson, 1998; List and Gallet, 1999; Millimet et al., 2003; Phu, 2003; Roy and Cornelis van Kooten, 2004; Paudel et al., 2005; Bertinelli and Strobl, 2005; Azomahou et al., 2006; Nguyen Van and Azomahou, 2007; Criado, 2008; Luzzati and Orsini, 2009; Poudel et al., 2009).

In addition to the concerns related to an ad hoc functional form, another concern raised by researchers is that the income variable in the EKC model could be endogenous. This endogeneity of income in the EKC model comes from simultaneity bias and omitted variable bias. The simultaneity bias is present because deteriorated water quality affects economic growth. Increases in pollution may harm people's health, for example, thereby reducing GDP. In addition, output and pollution may also be jointly produced in the production process, causing GDP and pollution to be simultaneously determined. Omitted variable bias in the EKC regression arises from such omitted variables as cultural or geographic factors that affect both environmental quality and income.

To address the potential endogeneity of income in EKC regressions, Lin and Liscow (2013) use a parametric instrumental variables regression approach with and without fixed effects. According to their results, evidence for an inverted-U relationship between income and environmental degradation are found for at least two out of the four IV specifications for seven out of eleven water pollutants: biological oxygen demand, chemical oxygen demand, arsenic, cadmium, lead, nickel, and fecal coliform. For these pollutants, there is both a peak and a trough. Their IV results therefore provide some support for an environmental Kuznets curve in global water quality. In contrast, the OLS results, which do not address the endogeneity of income, show no inverted-U relationship for any of the pollutants.

In addition to the concerns regarding parametric functional form and endogeneity of income, a third concern affecting EKC estimation is the discrete nature of the political variables. Although the literature on estimating the environmental Kuznets curve is growing fast and becoming very sophisticated in terms of empirical methodology used, hitherto published articles in the EKC literature have not properly addressed the properties of categorical, binary, and/or ordered explanatory variables in the model. One problem that arises in incorporating political rights and civil liberties variables, or any other categorical, ordered, or binary variable, in a semiparametric or nonparametric regression is that such variables cannot be treated as continuous variables. The conventional nonparametric approach to handle qualitative variables is a "frequency estimator" which involves splitting the samples into a number of cells (Racine and Li, 2004). In this paper, we use nonparametric functional forms for both categorical and continuous variables.

The highly nonlinear relationship between pollution and income, the endogeneity of income, and the discrete nature of political variables are therefore three concerns regarding EKC estimation. In order to address all three of these concerns, we use a nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables in order to better identify the relationships between pollution, income, and political institutions. We apply this method to analyze the relationships between water quality, per capita GDP, political rights, and civil liberties at the global level over the period 1980 to 2012.

Our results indicate an inverted-U shaped relationship between pollution and income consistent with an environmental Kuznets curve for one pollutant (lead), a cubic shape for three pollutants (nickel, mercury, and fecal coliform), and more highly nonlinear relationships for many of the other pollutants. For several stock pollutants (nickel, mercury, and arsenic), we find that pollution levels may continue to increase at higher levels of income, suggesting that stock pollutants may continue to accumulate in productive high-income countries even when marginal emissions have been reduced. We also find suggestive evidence that levels of pollutants resulting from industrial activity (e.g., chemical oxygen demand) may increase with income while those that are more driven by residential activity and population levels (e.g., fecal coliform) do not.

In general, we find that improved civil liberties and political rights are correlated with better water quality. By estimating a nonparametric relationship between pollution and political institutions and by accounting for the categorical nature of the political variables, we are able to detect a nonlinear relationship between pollution and political institutions as well, which for some pollutants is an inverted-U shaped curve.

2. Model

We first develop a theoretical model to illustrate why it is important to address the potential endogeneity of income when estimating the relationship between pollution and income; why it is important to control for political institutions; and why a nonparametric estimation may be better than a parametric estimation.

Let pollution *p* be given by the following function $g(\cdot)$ of income *y*, political variable *x*, and unobserved variable *z*:

$$p = g(y, x, z). \tag{1}$$

Let income *y* be given by the following function $f(\cdot)$ of pollution *p*, political variable *x*, and unobserved variable *z*:

$$y = f(p, x, z). \tag{2}$$

Substituting the income function $f(\cdot)$ into the pollution function $g(\cdot)$, one obtains:

$$p = g(f(p, x, z), x, z). \tag{3}$$

The EKC relationship between pollution and income is given by the partial derivative of the pollution function with respect to income, or $\frac{\partial g}{\partial y}$. The partial derivative of pollution with respect to income, $\frac{\partial g}{\partial y}$, is not necessarily the same as the total derivative of pollution with respect to income, $\frac{dp}{dy}$, as the following Proposition illustrates.

<u>Proposition 1.</u> The relationship between the partial derivative of pollution with respect to income and the total derivative of pollution with respect to income is given by:

$$\frac{\partial g}{\partial y} = \frac{dp}{dy} - \frac{\frac{\partial g}{\partial x}}{\frac{\partial f}{\partial p}\frac{\partial g}{\partial y}\frac{\partial f}{\partial x} + \frac{\partial f}{\partial p}\frac{\partial g}{\partial x} + \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{dz}{dx}} - \frac{\frac{\partial g}{\partial z}}{\frac{\partial f}{\partial p}\frac{\partial g}{\partial y}\frac{\partial f}{\partial z} + \frac{\partial f}{\partial p}\frac{\partial g}{\partial z} + \frac{\partial f}{\partial x}\frac{dx}{dz} + \frac{\partial f}{\partial z}}\ . \tag{4}$$

Proof. Taking the total derivative of (1), one obtains:

$$dp = \frac{\partial g}{\partial y} dy + \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial z} dz \,, \tag{5}$$

which, upon rearranging, yields:

$$\frac{\partial g}{\partial y} = \frac{dp}{dy} - \frac{\partial g}{\partial x}\frac{dx}{dy} - \frac{\partial g}{\partial z}\frac{dz}{dy} \,. \tag{6}$$

Taking the total derivative of (2), one obtains:

$$dy = \frac{\partial f}{\partial p} dp + \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial z} dz \,. \tag{7}$$

Substituting (7) into (6), one obtains:

$$\frac{\partial g}{\partial y} = \frac{dp}{dy} - \frac{\dfrac{\partial g}{\partial x}}{\dfrac{\partial f}{\partial p}\dfrac{dp}{dx} + \dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial z}\dfrac{dz}{dx}} - \frac{\dfrac{\partial g}{\partial z}}{\dfrac{\partial f}{\partial p}\dfrac{dp}{dz} + \dfrac{\partial f}{\partial x}\dfrac{dx}{dz} + \dfrac{\partial f}{\partial z}} \,. \tag{8}$$

The total derivative of pollution p with respect to political variable x is given by:

$$\frac{dp}{dx} = \frac{\partial g}{\partial y}\frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \,. \tag{9}$$

The total derivative of pollution p with respect to unobserved variable z is given by:

$$\frac{dp}{dz} = \frac{\partial g}{\partial y}\frac{\partial f}{\partial z} + \frac{\partial g}{\partial z} \,. \tag{10}$$

Substituting (9) and (10) into (8) yields the desired result. □

As seen in Proposition 1, if we ignore the partial derivatives in the second and third terms in the right-hand side of equation (4), we will have a biased estimate of the partial derivative of the pollution function with respect to income $\frac{\partial g}{\partial y}$, and therefore a biased estimate of the EKC relationship between pollution and income.

Corollary 1 describes the bias that arises if income is endogenous due to simultaneity bias. Simultaneity bias is present when pollution affects income so that $\frac{\partial f}{\partial p} \neq 0$. For example, one reason pollution may affect income is that deteriorated environmental quality may affect economic growth.

Corollary 1. (Endogeneity from simultaneity bias). If pollution affects income $\left(\frac{\partial f}{\partial p} \neq 0\right)$, then ignoring the effect $\frac{\partial f}{\partial p}$ of pollution on income will lead to a biased estimate of the partial derivative of the pollution function with respect to income $\frac{\partial g}{\partial y}$, and therefore a biased estimate of the EKC relationship between pollution and income, if either or both of the following are true:

(i) The political variable x affects pollution p: $\frac{\partial g}{\partial x} \neq 0$.

(ii) The unobserved variable z affects pollution p: $\frac{\partial g}{\partial z} \neq 0$.

Proof. If $\frac{\partial f}{\partial p} \neq 0$, then $\frac{\partial g}{\partial x} \neq 0$ is a sufficient condition for the second term on the right-hand side of equation (4) to be non-zero. If $\frac{\partial f}{\partial p} \neq 0$, then $\frac{\partial g}{\partial z} \neq 0$ is a sufficient condition for the third term on the right-hand side of equation (4) to be non-zero. □

In addition to simultaneity, a second reason income may be endogenous is that there may be omitted variables such as political institutions or cultural or geographic factors that affect both environmental quality and income. Corollary 2 describes the bias that arises if income is endogenous due to omitted variable bias. Omitted variable bias is present when the unobserved variable z affects both pollution and income (i.e., when $\frac{\partial g}{\partial z} \neq 0 \wedge \frac{\partial f}{\partial z} \neq 0$, where $\wedge$ is the logical operator denoting the logical conjunction "and").

Corollary 2. (Endogenity from omitted variable bias). If the unobserved variable z affects both pollution and income $\left(\frac{\partial g}{\partial z} \neq 0 \wedge \frac{\partial f}{\partial z} \neq 0\right)$, then ignoring the effects of the unobserved variable z on pollution and on income will lead to a biased estimate of the partial derivative of the pollution function with respect to income $\frac{\partial g}{\partial y}$, and therefore a biased estimate of the EKC relationship between pollution and income.

Proof. If $\frac{\partial g}{\partial z} \neq 0 \wedge \frac{\partial f}{\partial z} \neq 0$, then the third term on the right-hand side of equation (4) would be non-zero. □

Owing to simultaneity bias and omitted variable bias, it is important to address the endogeneity of income. In addition, owing to omitted variable bias, it is important to control for political institutions as well.

In addition to highlighting the importance of addressing endogeneity and of controlling for political institutions, our model also provides intuition why a non-parametric estimation may be better than a parametric estimation. This is formalized in Corollary 3 below.

Corollary 3. (Need for nonparametric model). If any of the partial derivatives in the second or third terms on the right-hand side of equation (4) are nonlinear; then the EKC relationship $\frac{\partial g}{\partial y}$, which is a nonlinear function of the partial derivatives, is highly nonlinear as well, and therefore best modeled with a nonparametric model.

Proof. Follows from Proposition 1. □

In order to address the concerns raised by our theory model regarding endogeneity and the highly nonlinear nature of the relationship between pollution and income, as well as a third concern regarding the discrete nature of political variables, we use a nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables in order to identify the relationship between income and pollution.

3. Methods

We are interested in identifying the relationships between water pollution, income, and political institutions. The highly nonlinear relationship between income and pollution, the endogeneity of income, and the discrete nature of political variables are three concerns regarding such EKC estimation. In order to address all three of these concerns, we use a nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables.

To address the potential endogeneity of income, Lin and Liscow (2013) use a parametric instrumental variables regression approach with debt service and age dependency ratio as instruments for per capita GDP. As their results suggest that the age dependency ratio is not strong instrument, in this paper we only use debt as an instrument for GDP.

For the instrumental variable for GDP, we try several variables such as share of GDP from manufacturing sector, age dependency ratio, and total debt service. In the end, we choose to use total debt service (% of GNI), as it has a very high correlation with our per capita GDP income variable. Debt is correlated with GDP, but does not have a direct effect on environmental quality, and is therefore a good instrument for GDP. Total debt service, which includes the principal repayments and interest actually paid on debt, is positively correlated with GDP because more debt is likely to be paid off when GDP is higher. Debt service may be correlated with types of degradation like deforestation, if countries liquidate natural assets to pay off debts (Kahn and McDonald, 1995), but there is less reason to believe that countries with high debts would pollute more, except indirectly through the effects of debt on GDP, since emitting more water pollution does not help them pay off debts the same way liquidating natural assets would. It is therefore GDP, not debt, that has a potential effect on water pollution. We test the validity of debt as an instrument for income using a Sargan test, and the results confirm that debt is an exogenous variable.

We also test for the endogeneity of political rights and civil liberties in all pollutant-income relationships using a Durbin-Wu-Hausman test. For the endogeneity test, we use age dependency and share of GDP from manufacturing sector as instruments in the model. According to our results, we find political rights and civil right liberties to be exogenous in the models for all pollutants except arsenic. Therefore, for the arsenic regression, we use instrument not only for GDP but also for political rights and civil liberties as well.

Recent papers in the econometrics literature, such as Darolles et al. (2011) and Horowitz (2011), have developed nonparametric instrumental variable estimation methods. In this paper, we use a method suggested by Horowitz (2011) to estimate a nonparametric instrumental variable EKC regression model that allows for the inclusion of continuous and discrete variables. Such a model addresses three concerns regarding EKC estimation: the highly nonlinear relationship between income and pollution, the endogeneity of income, and the discrete nature of political variables.

The model is given by:

$$p = g_y(y) + \sum_{j=1}^{J} g_j(x_j) + u \quad , \tag{11}$$

where p is pollution, $g_y(\cdot)$ is an unknown smooth function for income y, and $g_j(\cdot)$ is the unknown function for the $J = 2$ political factors x_j we use: civil liberties and political rights. Although the model is estimated using a nonparametric regression model, we separate the regressors y and x to emphasize the differences between the two types of variables: income y is continuous and endogenous, while the political factors x_j are discrete and exogenous.

Since the civil liberties and political rights variables are ordinal, we use the method suggested by Racine and Li (2004) to address continuous and categorical nature of variables. Further details on how to estimate a model with continuous and categorical variables in nonparametric estimation can be found in Ma and Racine (2013), Nie and Racine (2012), and Ma, Racine and Yang (2015). For simplicity, let us consider:

$$g(y, x) = g_y(y) + \sum_{j=1}^{J} g_j(x_j) \,. \tag{12}$$

Then, equation (11) can be written as:

$$p = g(y, x) + u; \qquad E[u \mid w, x] = 0 \tag{13}$$

for all instruments w and exogenous covariates x, which is equivalent to:

$$E[p - g(y, x) \mid w, x] = 0 \,. \tag{14}$$

In this model, y denotes per capita GDP, which is endogenous; x denotes the ordinal exogenous explanatory variables, political rights and civil liberties; and w denotes the instrument, debt. To address the ordinal and categorical nature of the political variables in a nonparametric model, we use a method suggested by Ma and Racine (2011), Nie and Racine (2012), and Ma, Racine and Yang (2015) to estimate the nonparametric instrumental variable model[5] given in equation (14). We use cross-validation as our bandwidth selection method to estimate the nonparametric model.

Details on estimating equation (14) can be found in Horowitz (2011). Here, we briefly describe the estimation procedure. We use a control function approach for nonparametric instrumental variable estimation. In the control function model,

$$p = g(y, x) + u \tag{15}$$

and

$$y = h(w) + v, \tag{16}$$

where $g(\cdot)$ and $h(\cdot)$ are unknown functions,

$$E[v \mid w] = 0 \quad \forall w \tag{17}$$

and

$$E[u \mid y, x, v] = E[u \mid v] \quad \forall y, x, v. \tag{18}$$

Assuming that the mean of y conditional on w exists, equations (16) and (17) can always be made to hold by setting $h(w) = E[y \mid w]$ (Horowitz, 2011). Identification in the control function approach comes from equation (18). It follows from equations (15) and (18) that:

$$E[p \mid y, x, v] = g(y, x) + k(v), \tag{19}$$

where $g(\cdot)$ and $k(\cdot)$ are unknown functions. In the above equation (19), v is unobservable and estimated consistently by the residuals from nonparametric estimation of $h(\cdot)$ in equation (16). We then estimate $g(\cdot)$ and $k(\cdot)$ using nonparametric additive models.

[5] We use the 'crs' R package to estimate nonparametric models which contain both categorical and continuous variables (Racine et al., 2014).

We use splines for the continuous variables and kernels for the categorical variables. We use a data-driven method to select bandwidths and to estimate all nonparametric functional components including $g(y,x)$ and $h(w)$. In particular, for bandwidth selection we use least squares cross-validation (Li and Racine, 2007). Our cross-validation criterion function is:

$$\min_{l} \sum_{i} \left(p_i - \hat{g}_{-i}(y,x) \right)^2 , \tag{20}$$

where $\hat{g}_{-i}(y,x)$ is the leave-one-out estimator of $g(y,x)$ omitting observation *i*, and where *l* is the vector of bandwidths used to construct $\hat{g}_{-i}(y,x)$.

4. Data

We use water pollution data from the Global Environment Monitoring System (GEMS) Water Dataset, which consists of triennial surveys of water quality statistics since 1979 from countries across the developed and developing world. We update the data set used by Grossman and Krueger (1995) and by Barrett and Graddy (2000) to include the years from 1991 to 2012.

The Global Environment Monitoring System (GEMS) for Water was established in 1978 to improve water quality monitoring in freshwater ecosystems (International Centre for Water Resources and Global Change, 2017) and to collect worldwide water quality data for assessments of status and trends in global inland water quality (United Nations Environment, 2017). Under the auspices of UNEP, the United Nations Environment Programme, GEMS involves the World Health Organisation (WHO), the World Meteorological Organisation (WMO) and UNESCO (International Centre for Water Resources and Global Change, 2017).

With the cooperation of participating countries, the GEMS water program is creating a unique global water quality monitoring network that provides water quality data to a central database (United Nations Environment, 2017). The data and information on water quality provided by participating countries provide a global overview of the condition of water bodies and statements on changes and trends at global, regional, and local levels (International Centre for Water Resources and Global Change, 2017). The GEMS water program also provides support and encouragement to developing countries that wish to establish monitoring programs and conduct assessments of water quality, through the provision of capacity development in the form of training, advice, and assessment tools (United Nations Environment, 2017).

The GEMS data set consist of over 70,000 observations of dozens of different types of water pollution, providing a substantive amount of data on varied measures of water quality. Each data point consists of the average over the three years of each triennial survey of one or more data point from one of the GEMS water program's hundreds of sites around the world. The year we assign to each data point is the middle of the three years.

The countries included in our data set are: Afghanistan, Algeria, Argentina, Australia, Austria, Bangladesh, Belgium, Bolivia, Brazil, Cambodia, Canada, Chile, China, Colombia, Congo, Cuba, Denmark, Ecuador, Egypt, Ecuador, Fiji, Finland, France, Germany, Ghana, Greece, Guatemala, Hong Kong, Hungary, India, Indonesia, Iran, Ireland, Israel, Italy, Japan, Jordan, Kenya, Korea, Laos, Lithuania, Luxembourg, Malaysia, Mali, Marshall Islands, Mexico, Morocco, Netherlands, New Zealand, Norway, New Zealand, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Russian Federation, Senegal, Singapore, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Tanzania, Thailand, Tunisia, Turkey, Uganda, United Kingdom, United States of America, Uruguay, Vietnam, and Zimbabwe.

We focus on four types of water pollutants: heavy metal (nickel, mercury, arsenic, cadmium, lead), pathogen-related (fecal coliform, total coliform), oxygen regime (dissolved oxygen (DO), chemical oxygen demand (COD), biological oxygen demand (BOD)), and nutrient-related (nitrate). All data are in the form of concentrations of mg/l except for the mercury data, which is in the form of μg/l; and the coliform data, which is in the form of measured count/100 ml.

Lead, mercury, arsenic, and cadmium are four important heavy metal pollutants responsible for adverse health effects in human beings (Jarup, 2003). These pollutants pollute water through deposition in soil and subsequent runoff to water bodies. Cadmium can pollute the soil and water through deposition of air pollution and through application of fertilizer and chemicals to crops. Mercury pollution in water bodies comes in the form of methyl mercury which can affect humans through consumption of fish. Lead pollution in water comes from air pollution settling in soil as well as from the application of organic and inorganic fertilizer in crop production. Although the phase-out of lead in gasoline has decreased the emission of lead (Hilton, 2006), lead pollution in soil remains a concern.

For our income measure, we use data on per capita gross domestic product (GDP) based on purchasing power parity (PPP) in constant 2005 international dollars from the World Development Indicators (WDI). In particular, GDP per capita is converted to constant 2005 international dollars using purchasing power parity rates. An international dollar has the same purchasing power over GDP as the U.S. dollar has in the United States. GDP at purchaser's prices is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products, and is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources.

For data on political mechanisms, we use indices on political rights (PR) and civil liberties (CL) from Freedom House. Each index varies from 1 to 7, with 1 meaning the most political rights or civil liberties. For example, the United States has a 1 in each category in all years, Indonesia has recently been in the middle of the range, and China has 7 in both categories for most years. Freedom House attempts to use a methodology not bound by culture, but instead uses standards drawn from the Universal Declaration of Human Rights (Freedom House, 2010). Political rights measure factors such as the fairness of the electoral process, the degree of political pluralism and participation, and the presence of a non-corrupt and transparent government (Freedom House, 2010). Civil liberties measure freedom of expression and beliefs, the ability to associate, the rule of law, and the degree of individual autonomy. The mean of the political rights variables is lower than that for civil liberties, which implies that political rights are more prevalent in many countries than civil liberties are. In previous studies (e.g., Torras and Boyce, 1998; Bhattarai and Hammig, 2001), political rights and civil liberties have been combined into one democracy measure that takes on values from 2 to 14.

For the instrumental variable for GDP, as explained above, we try several variables such as share of GDP from manufacturing sector, age dependency ratio (dependents—the population under age 15 and above age 65—as a proportion of the working age population), and total debt service (% of GNI), all from the World Development Indicators (WDI). In the end, we choose to use total debt service (% of GNI), as it has a very high correlation with our per capita GDP income variable.

Summary statistics for the water pollution variables are presented in Table 1; summary statistics for income, political variables, and debt are presented in Table 2.

As seen in Table 1, most pollutants exhibit a large range in values and a high standard deviation. According to exploratory plots of the data in Lin and Liscow (2013), the concentrations of the majority of the pollutants (chemical oxygen demand, total arsenic, dissolved oxygen, total lead, total nickel, and fecal coliform) are decreasing functions of per capita income,. The concentrations of only two pollutants (total cadmium and nitrate) exhibit increasing functions of per capita income. The concentrations of three pollutants (biological oxygen demand, total mercury, and total coliform) show no relationship with the income or political variables. Several of these trends are largely dependent upon the observations from only one or a few countries; for example, total cadmium's curve is dependent upon 1980s UK and 1990s France data. This suggests that water quality generally improves as countries develop.

Exploratory plots of the data also show that only a few of the pollutants (chemical oxygen demand, total arsenic, total mercury, and total cadmium) potentially have an inverted- U form for concentration with respect to income. Interestingly, a few of the pollutants (biological oxygen demand, chemical oxygen demand, total lead, fecal coliform) appear to have an inverted-U shape for the political variables as well. The high amounts of pollution and mid-range political variables for Mexico, India, and Colombia cause this phenomenon for both chemical and biological oxygen demand; this is also reflected in the OECD versus non-OECD plots, in which concentrations decrease for OECD countries with improving political institutions, while they increase for non-OECD countries with improving political institutions. These exploratory plots suggest that, to the extent that there is an EKC, it may be as much caused by political as income factors (Lin and Liscow, 2013).

5. Results

In Figure 1, we present the graphical results of the estimated relationships between water pollution and per capita GDP, political rights, and civil liberties resulting from the nonparametric instrumental variable estimation for each of the 11 water pollutants we examine. The left column of the figure represents the relationship between per capita GDP and pollution, the middle column represents the relationship between political liberties and pollution, and the last column represents the relationship between civil liberties and pollution.

We describe the results for each pollutant below. The pollutants are organized by type: heavy metal (nickel, mercury, arsenic, cadmium, lead), pathogen-related (fecal coliform, total coliform), oxygen regime (dissolved oxygen (DO), chemical oxygen demand (COD), biological oxygen demand (BOD)), and nutrient-related (nitrate).

5.1. Heavy metals

Nickel

We find a cubic relationship or N-shaped curve between nickel concentration and per capita GDP. We do not find that variations in either political rights or civil liberties impact the nickel concentration. When regressing nickel concentration on GDP, GDP squared, and GDP cubed, Grossman and Krueger (1995) find an environmental Kuznets curve shape for the nickel-income relationship. However, Bradford et al. (2005), who use a time series analysis, do not find the existence of an environmental Kuznets curve for this pollutant.

Mercury

Per capita GDP and mercury concentration seem to have a cubic relationship. We do not find that variations in either political rights or civil liberties impact the mercury concentration. Pandit and Paudel (2016) find an environmental Kuznets curve relationship for mercury pollution using 1985-2006 data from Louisiana watersheds.

Arsenic

Per capita GDP and arsenic pollution seem to have a V-shaped relationship. Arsenic pollution declines if there are no political rights (i.e., if the political rights index is high). Variations in civil liberties do not have any impact on the arsenic pollution.

Cadmium

Cadmium concentration seems to decline with increasing per capita GDP especially after per capita GDP hits the $6,000 level. However, we do not find any definitive relationships between cadmium concentration and per capita GDP.

Lead

We find that an inverted-U shaped relationship exists between per capita GDP and lead concentration. There is no distinct pattern on the relationship between civil liberties and lead pollution or political rights and lead pollution, although it looks like lead concentrations reach their peak when the political rights variable has the value of 5.

5.2. Pathogen-related

Fecal coliform

For fecal coliform, we find almost a cubic relationship between pollution and income. The lower turning points occur around \$4,000 whereas the upper turning point is around \$10,000. Variations in political rights or civil liberties do not seem to have any impact on the fecal coliform concentration in water bodies.

Total coliform

The relationship between total coliform and per capita GDP seem to follow almost a polynomial of 4th degree type of relationship. The pollution level seems to reduce substantially after the income level reaches the \$11,000 level. Bradford et al. (2005) do not find an environmental Kuznets curve relationship for total coliform. When the level of political rights is lower (and the political rights index is higher), total coliform concentration is lower.

5.3. Oxygen regime

Dissolved oxygen

At lower levels of GDP, the relationship between GDP and dissolved oxygen looks flat but once the GDP level is \$8000 or higher the dissolved oxygen level starts declining. We do see a clear quadratic relationship between political rights and GDP. There is no definitive relationship between variations in civil liberties and dissolved oxygen concentration.

Chemical oxygen demand

The relationship between GDP and chemical oxygen demand concentration looks like an N-shaped curve. We also see that higher civil liberties (and therefore a lower civil liberties index) are associated with higher levels of chemical oxygen demand and lower civil liberties are associated with lower levels of chemical oxygen demand. The relationship with political rights is flat.

Biological oxygen demand

The biological oxygen demand curve shows 5th degree of polynomial relationship with per capita GDP. We see a clear relationship between civil liberties and biological oxygen demand with higher civil liberties (and therefore a lower civil liberties index) associated with low biological oxygen demand levels and lower civil liberties associated with higher levels of pollution. We do not find that variations in political rights affect the level of biological oxygen demand.

5.4. Nutrient-related

Nitrate

Many studies (Paudel et al., 2011; Paudel and Schafer (2009) have found the existence of an environmental Kuznets curve for nitrate pollution. According to our results, however, there is no definitive relationship between nitrate pollution and per capita GDP. Higher civil liberties (and therefore a lower civil liberties index) lead to lower nitrate pollution but variations in political rights have no impact on the nitrate pollution. Using GEMS data, Bradford et al. (2005) also do not find support for an environmental Kuznets curve in nitrate pollution.

6. Conclusion

This study contributes to a better understanding of the relationships between water pollution and per capita GDP, civil liberties, and political rights at the global level. In order to address concerns about the highly nonlinear relationship between pollution and income, the endogeneity of income, and the discrete nature of political variables, we use a nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables in order to identify these relationships. The upshot of our paper is that one shape does not fit all the pollutants, and that the shapes vary quite a bit compared to what have been found so far in the literature.

According to our results, we find an inverted-U shaped relationship between pollution and income consistent with an environmental Kuznets curve for one pollutant (lead), a cubic shape for three pollutants (nickel, mercury, and fecal coliform), and more highly nonlinear relationships for many of the other pollutants.

For several stock pollutants (nickel, mercury, and arsenic), we find that pollution levels may continue to increase at higher levels of income, suggesting that stock pollutants may continue to accumulate in productive high-income countries even when marginal emissions have been reduced. We also find suggestive evidence that levels of pollutants resulting from industrial activity (e.g., chemical oxygen demand) may increase with income while those that are more driven by residential activity and population levels (e.g., fecal coliform) do not.

By using a nonparametric model that accounts for the discrete nature of the political variables, we find that fewer pollutants exhibit an environmental Kuznets curve than were previously found in studies using GEMS data, including those of Grossman and Krueger (1995), Ekins (1997), Barrett and Graddy (2000), Bradford et al. (2005), and Lin and Liscow (2013).

In terms of the political variables, we found that the arsenic and total coliform levels decline as the level of political rights declines (and as the political rights index increases), but lead and dissolved oxygen have an inverted-U shaped relationship political rights. For lead and dissolved oxygen, results suggest that as countries progress towards political rights, water pollution increases at first but then decreases after certain levels of political rights have been attained. Our results indicate that higher biological oxygen demand and nitrate pollution levels are associated with lower levels of civil liberties (higher civil liberties index) but that lower chemical oxygen demand levels are associated with lower levels of civil liberties (higher civil liberties index). Thus, factors affecting political rights such as the fairness of the electoral process, the degree of political pluralism and participation, and the presence of a non-corrupt and transparent government are beneficial for water quality to some extent. As seen in our theoretical model, even for those water pollutants with which political institutions do not appear to have any systematic relationship, it is still important to control for political institutions to address any endogeneity in income that may arise from omitted variable bias.

By estimating a nonparametric relationship between pollution and political institutions and by accounting for the categorical nature of the political variables, we are able to detect a nonlinear relationship between pollution and political institutions, which for some pollutants is an inverted-U shaped curve. We therefore improve upon the empirical analysis of Lin and Liscow (2013), whose model uses instrumental variables but, unlike the model in this paper, is neither nonparametric nor accounts for the discrete nature of the political variables. In contrast to our paper, Lin and Liscow (2013) are unable to tease out the nonlinear nature of some of the relationships, and instead find that the effect of political variables on pollution can be either positive or negative depending on pollutant and political variable.

Our results suggests several avenues for potential research. First, while the GEMS global water data set is arguably the most comprehensive and reliable global data set on country-level water quality to date, while the GEMS water program aims at improving water quality monitoring in freshwater ecosystems, and while the GEMS data set has been used by previous researchers in seminal papers on the topic (e.g., Grossman and Krueger, 1995; Barrett and Graddy, 2000), one may worry that water quality monitoring in developing countries may be weak. In future work, as water quality monitoring continues to improve, we hope to further analyze relationships between environmental degradation, income, and political institutions with improved data as such data becomes available.

A second avenue for future research is to extend our nonparametric instrumental variable approach to enable the exploration of possible correlations between pollutants. Pandit and Paudel (2016) use a seemingly unrelated partially linear model to address a potential correlation between pollutants (nitrogen, phosphorous, dissolved oxygen, and mercury) in an environmental Kuznets curve study. In future work, we hope to address correlations using a nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables.

A third avenue for future research is to build on our nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables in a way that addresses any nonstationarity in the variables.

A fourth avenue for future research is to build on our nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables in a way that addresses any transboundary pollutants. While the water pollutants in our data set are local water pollutants in inland water bodies for which transboundary issues across country borders may be second-order, in general it is possible for pollutants to cross country borders. Lin (2010) develops a spatial econometric approach to measuring pollution externalities; we hope in future work to combine spatial econometrics with our nonparametric instrumental variable approach that allows for the inclusion of continuous and discrete variables in a way that addresses any transboundary pollutants.

A fifth avenue for future research is to use new annual country-level data on institutions that the World Bank has begun collecting in 2008, including variables measuring political stability and absence of violence/terrorism; regulatory quality; rule of law; voice and accountability; government effectiveness; and control of corruption. Since these data are fairly recent, there is less of an overlap with the 1980-2012 time period of the data used in this study. However, as more years of this new World Bank institutions data become available, we hope in future work to use this data to further tease out the relationships between water pollution, income, and institutions.

A sixth avenue for future research is to further and more rigorously examine possible mechanisms for why pollution levels may increase with income for some pollutants. Doing so may require more and better data, especially at higher levels of income. In future work, as more and better data at higher levels of income become available, we hope to further analyze mechanisms by which pollution levels may increase at higher levels of income.

Our results suggest that there may be nonlinear relationships between environmental degradation, income, and political institutions. As a country's economy develops and its income rises, there may be nonlinear effects on water quality. Similarly, improvements in political institutions may have nonlinear effects on water quality as well. Our results also suggest that policies that improve income and/or political institutions may have nonlinear effects on water quality, and that these effects vary by pollutant. Policy-makers and academics alike should therefore be aware of the complex nature of the relationships between water pollution, income, and political institutions as they design strategies for environmental management and sustainable development.

References

Azomahou, T., F. Laisney, and P. Nguyen Van. 2006. "Economic Development and CO_2 Emissions: A Nonparametric Panel Approach." *Journal of Public Economics* 90 (6-7): 1347-1363.

Barbier, E.B. 2004. "Explaining Agricultural Land Expansion and Deforestation in Developing Countries." *American Journal of Agricultural Economics* 86 (5): 1347-1353.

Bhattarai, M. and M. Hammig. 2001. "Institutions and the Environmental Kuznets Curve for Deforestation: A Crosscountry Analysis for Latin America, Africa And Asia." *World Development* 29 (6): 995-1010.

Barrett, S., and K. Graddy. 2000. "Freedom, Growth, and the Environment." *Environment and Development Economics* 5 (4): 433-456.

Bertinelli, L., and E. Strobl. 2005. "The Environmental Kuznets Curve Semi-Parametrically Revisited." *Economics Letters* 88 (3): 350-357.

Bradford, D. F., R. A. Fender, S. H. Shore, and M. Wagner. 2005. "The environmental Kuznets curve: exploring a fresh specification." *B.E. Journal of Economic Analysis and Policy: Contributions* 4(1).

Bruvoll, A., and H. Medin. 2003. "Factors Behind the Environmental Kuznets Curve: A Decomposition of the Changes in Air Pollution." *Environmental and Resource Economics* 24 (1): 27-48.

Carson, R.T. 2010. "The Environmental Kuznets Curve: Seeking Empirical Regularity and Theoretical Structure." *Review of Environmental Economics and Policy* 4 (1): 3-23.

Congleton, R.D. 1992. "Policital Institutions and Pollution Control." *Review of Economics and Statistics* 74 (3): 412-421.

Copeland, B.R., and M.S. Taylor. 2004. "Trade, Growth, and the Environment." *Journal of Economic Literature* 42 (1): 7-71.

Criado, C.O. 2008. "Temporal and Spatial Homogeneity in Air Pollutants Panel EKC Estimations." *Environmental and Resource Economics* 40 (2): 265.

Culas, R.J. 2007. "Deforestation and the Environmental Kuznets Curve: An Institutional Perspective." *Ecological Economics* 61 (2-3): 429-437.

Darolles, S., Y. Fan, J.P. Florens, and E. Renault. 2011. "Nonparametric Instrumental Regression." *Econometrica* 79 (5): 1541-1565.

Dasgupta, P., and K.-G. Mäler. 1995. Poverty, Institutions, and the Environmental Resourcebase, ed. J. Behrman, and T.N. Srinivaan, vol. 3A. Amsterdam.

Dasgupta, S., B. Laplante, W. Hua, and D. Wheeler. 2002. "Confronting the Environmental Kuznets Curve." *Journal of Economic Perspectives* 16 (1): 147-168.

Deacon, R.T., and C.S. Norman. 2006. "Does the Environmental Kuznets Curve Describe How Individual Countries Behave?" *Land Economics* 82 (2): 291-315.

Ekins, P. 1997. "The Kuznets Curve For The Environment And Economic Growth: Examining The Evidence." *Environment and Planning A* 29 (5): 805-830.

Environmental Protection Agency [EPA]. 2016. "Dissolved Oxygen." URL: https://www3.epa.gov/caddis/ssr_do4s.html

Farzin, Y.H., and C.A. Bond. 2006. "Democracy and Environmental Quality." *Journal of Development Economics* 81 (1): 213-235.

Fondriest Environmental, Inc. 2013. "Dissolved Oxygen." Fundamentals of Environmental Measurements. URL: http://www.fondriest.com/environmental-measurements/parameters/water-quality/dissolved-oxygen/.

Freedom House. 2010. "Freedom in the World 2003: Survey Methodology", http://www.freedomhouse.org/research/freeworld/2003/methodology.htm. (accessed Dec 10, 2010)

Galinato, G.I., and S.P. Galinato. 2012. "The Effects of Corruption Control, Political Stability and Economic Growth on Deforestation-Induced Carbon Dioxide Emissions." *Environment and Development Economics* 17 (1): 67-90.

Gawande, K., R.P. Berrens, and A.K. Bohara. 2001. "A Consumption-Based Theory of the Environmental Kuznets Curve." *Ecological Economics* 37 (1): 101-112.

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

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

Heerink, N., A. Mulatu, and E. Bulte. 2001. "Income Inequality and the Environment: Aggregation Bias in Environmental Kuznets Curves." *Ecological Economics* 38 (3): 359-367.

Hilton, F.G. 2006. "Poverty and Pollution Abatement: Evidence from Lead Phase-Out." *Ecological Economics* 56 (1): 125-131.

Hilton, F., and A. Levinson. 1998. "Factoring the Environmental Kuznets Curve: Evidence from Automotive Lead Emissions." *Journal of Environmental Economics and Management* 35: 126-141.

Horowitz, J.L. 2011. "Applied Nonparametric Instrumental Variables Estimation." *Econometrica* 79 (2): 347-394.

Israel, D., and A. Levinson. 2004. "Willingness to Pay for Environmental Quality: Testable Empirical Implications of the Growth and Environment Literature." *B.E. Journal of Economic Analysis and Policy: Contributions* 3 (1): Article 2.

International Centre for Water Resources and Global Change. 2017. "Data: GEMS/Water-Programme". URL: http://www.waterandchange.org/en/#daten

Jarup, L. 2003. "Hazards of Heavy Mmetal Contamination." *British Medical Bulletin* 68: 167-182.

Jha, R., and K.V.B. Murthy. 2003. "An Inverse Global Environmental Kuznets Curve." *Journal of Comparative Economics* 31 (2): 352-368.

Kahn, J.R., and J.A. McDonald. 1995. "Third-World Debt and Tropical Deforestation." *Ecological Economics* 12: 107-123.

Kemira. 2017. "BOD and COD Removal." URL: http://www.kemira.com/en/industries-applications/pages/bodcod-removal.aspx

Kuznets, S. 1955. "Economic Growth and Income Equality." *American Economic Review* 45 (1): 1-28.

Li, Q., and J.S. Racine. 2007. *Nonparametric Econometrics: Theory and Practice*. New Jersey: Princeton University Press.

Liddle, B. 2015. "Urban Transport Pollution: Revisiting the Environmental Kuznets Curve." *International Journal for Sustainable Transportation* 9 (7): 502-508.

Liddle, B., and G. Messinis. 2015. "Revisiting Sulfur Kuznets Curves With Endogenous Breaks Modeling: Substantial Evidence of Inverted-Us/Vs For Individual OECD Countries." *Economic Modelling* 49: 278-285.

Lin, C.-Y.C. 2010. "A spatial econometric approach to measuring pollution externalities: An application to ozone smog." *Journal of Regional Analysis and Policy* 40 (1): 1-19.

Lin, C.-Y.C., and Z.D. Liscow. 2013. "Endogeneity in the Environmental Kuznets Curve: An Instrumental Variables Approach." *American Journal of Agricultural Economics* 95 (2): 268-274.

List, J.A., and C.A. Gallet. 1999. "The Environmental Kuznets Curve: Does One Size Fit All?" *Ecological Economics* 31 (3): 409-423.

Lomborg, B., and C. Pope 2003. "The Global Environment: Improving or Deteriorating?' John F. Kennedy, Jr. Forum at the Harvard Kennedy School of Government", http://www.iop.harvard.edu/programs/forum/transcripts/environment/03.13.03.pdf (accessed 12 Aug 2010)

Luzzati, T., and M. Orsini. 2009. "Investigating the Energy-Environmental Kuznets Curve." *Energy* 34 (3): 291-300.

Ma, S., and J.S. Racine. 2013. "Additive Regression Splines with Irrelevant Categorical and Continuous Regressors." *Statistica Sinica* 23 (2): 515-541.

Ma, S., J.S. Racine, and L. Yang. 2015. "Spline Regression in the Presence of Categorical Predictors." *Journal of Applied Econometrics* 30 (5): 705-717.

Merlevede, B., T. Verbeke, and M.D. Clercq. 2006. "The Ekc for So2: Does Firm Size Matter? ." *Ecological Economics* 59 (4): 451-461.

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

Nguyen Van, P., and T. Azomahou. 2007. "Nonlinearities and Heterogeneity in Environmental Quality: An Empirical Analysis of Deforestation." *Journal of Development Economics* 84 (1): 291-309.

Nie, Z., and J.S. Racine. 2012. "The Crs Package: Nonparametric Regression Splines for Continous and Categorical Predictors." *The R Journal* 4: 48-56.

Pandit, M., and K.P. Paudel. 2016. "Water Pollution and Income Relationships: A Seemingly Unrelated Partially Linear Analysis." *Water Resources Research* 52 (10): 7668–7689.

Paudel, K., B.N. Poudel, D. Bhandari, and T. Johnson. 2011. "Examining the Role of Social Capital in Environmental Kuznets Curve Estimation." *Global Journal of Environmental Science and Technology*: 1-11.

Paudel, K.P., and M.J. Schafer. 2009. "The Environmental Kuznets Curve under a New Framework: The Role of Social Capital in Water Pollution." *Environmental and Resource Economics* 42 (2): 265-278.

Paudel, K.P., H. Zapata, and D. Susanto. 2005. "An Empirical Test of Environmental Kuznets Curve for Water Pollution." *Environmental and Resource Economics* 31 (3): 325-348.

Phu, N.V. 2003. "A Semiparametric Analysis of Determinants of a Protected Area." *Appied Economics Letters* 10 (10): 661-665.

Plassmann, F., and N. Khanna. 2006. "Household Income and Pollution: Implications for the Debate About the Environmental Kuznets Curve Hypothesis." *Journal of Environment and Development* 15 (1): 22-41.

Poudel, B.N., K.P. Paudel, and K. Bhattarai. 2009. "Searching for an Environmental Kuznets Curve in Carbon Dioxide Pollutant in Latin American Countries." *Journal of Agricultural and Applied Economics* 41 (1): 13-27.

Racine, J., and Q. Li. 2004. "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data." *Journal of Econometrics* 119 (1): 99-130.

Racine, J.S., Z. Nie, and B.D. Ripley. 2014. crs: Categorical Regression Splines. R package version 0.15-23. http://CRAN.R-project.org/package=crs

Rodriguez-Meza, J., D. Southgate, and C. Gonzalez-Vega. 2004. "Rural Poverty, Household Responses to Shocks, and Agricultural Land Use: Panel Results for El Salvador." *Environment and Development Economics* 9 (2): 225-239.

Roy, N., and G. Cornelis van Kooten. 2004. "Another Look at the Income Elasticity of Non-Point Source Air Pollutants: A Semiparametric Approach." *Economics Letters* 85 (1): 17-22.

Rupasingha, A., S.J. Goetz, D.L. Debertin, and A. Pagoulatos. 2004. "The Environmental Kuznets Curve for U.S. Counties: A Spatial Econometric Analysis with Extensions." *Papers in Regional Science* 83 (2): 407-424.

Schmalensee, R., T.M. Stoker, and R.A. Judson. 1998. "World Carbon Dioxide Emissions: 1950–2050." *Review of Economics and Statistics* 80 (1): 15-27.

Stern, D.I. 2008. The Rise and Fall of the Environmental Kuznets Curve, ed. J. Martinez-Alier, and I. Ropke, Elgar Reference Collection. Cheltenham, U.K. and Northampton, Mass.: Elgar, pp. 519-539.

Torras, M., and J.K. Boyce. 1998. "Income, Inequality, and Pollution: A Reassessment of the Environmental Kuznets Curve." *Ecological Economics* 25 (2): 147-160.

United Nations Environment. 2017. "GEMS Water History and Mandate." URL: https://www.unep.org/gemswater/who-we-are/history-and-mandate

Zapata, H.O., and K.P. Paudel. 2009. "Functional Form of the Environmental Kuznets Curve," ed. T.B. Fomby, and R.C.l. Hill, vol. 25, Emerald Group Publishing Limited, pp. 471-493.

Table 1: Summary statistics for water pollutants

Variable Name	M. type	Mean	Std. Dev.	Min	Max	Observations	
Nickel	overall	0.014	0.030	0.000	0.326	N	246
	between		0.015	0.000	0.067	n	30
	within		0.024	-0.041	0.325	T-bar	8.20
Mercury	overall	0.336	0.713	0.000	7.900	N	447
	between		0.459	0.000	2.468	n	44
	within		0.617	-2.133	6.937	T-bar	10.16
Arsenic	overall	0.017	0.068	0.000	0.785	N	309
	between		0.084	0.000	0.518	n	38
	within		0.051	-0.250	0.516	T-bar	8.13
Cadmium	overall	0.023	0.097	0.000	1.000	N	475
	between		0.061	0.000	0.257	n	45
	within		0.081	-0.222	0.857	T-bar	10.56
Lead	overall	0.030	0.106	0.000	1.067	N	500
	between		0.127	0.000	0.500	n	50
	within		0.079	-0.440	0.860	T-bar	10.00
Fecal coliform	overall	47,982.37	229,659.500	0.000	3681414.000	N	467
	between		96,137.040	0.000	515869.100	n	42
	within		201,667.600	-411961.200	3383963.000	T-bar	11.12
Total coliform	overall	134,726.9	660,444.300	0.000	10400000.000	N	431
	between		421,087.000	0.000	2593846.000	n	47
	within		541,430.600	-2456222.000	7985642.000	T-bar	9.17
Dissolved oxygen	overall	8.389	2.497	0.000	42.500	N	914
	between		1.943	3.556	11.586	n	70
	within		1.584	3.983	41.169	T-bar	13.06
Chemical oxygen demand	overall	24.740	31.787	0.873	393.400	N	531
	between		23.420	2.011	96.650	n	52
	within		22.680	-52.046	331.192	T-bar	10.21
Biological oxygen demand	overall	4.189	9.658	0.348	192.400	N	688

	between		6.456	0.831	33.465	n	56
	within		8.134	-27.394	163.506	T-bar	12.29
Nitrate	overall	1.281	2.302	0.010	18.565	N	294
	between		2.957	0.067	15.283	n	41
	within		1.300	-4.069	11.908	T-bar	7.17

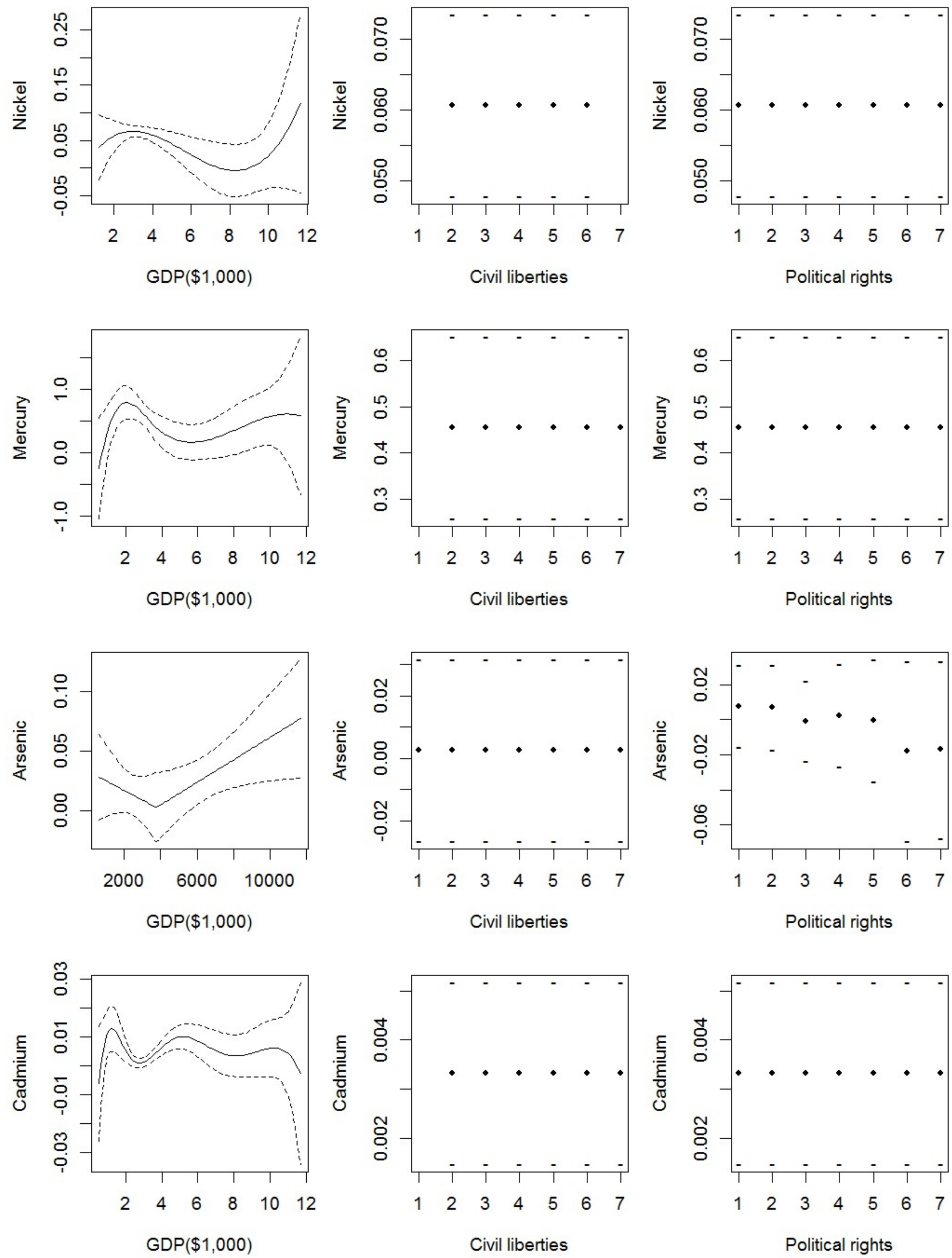
Notes: All data are in the form of concentrations of mg/l except for the mercury data, which is in the form of μg/l; and the coliform data, which is in the form of measured count/100 ml. The second column under the M.type in the table indicates whether the values shown are "overall", "between" or "within". "Overall" statistics are based on the total number of observations, "between" statistics describe properties between individual countries, and "within" statistics describe the properties over time. The "between" min and max refer to the minimum and maximum, respectively, of the average value for each country. The "within" min is the maximum negative deviation from each country's average. The "within" max is the maximum positive deviation from each country's average plus the global average. In the observation column, N is the number of observations, n is the number of countries, and T-bar is the average number of years per country. The data set is an unbalanced panel.

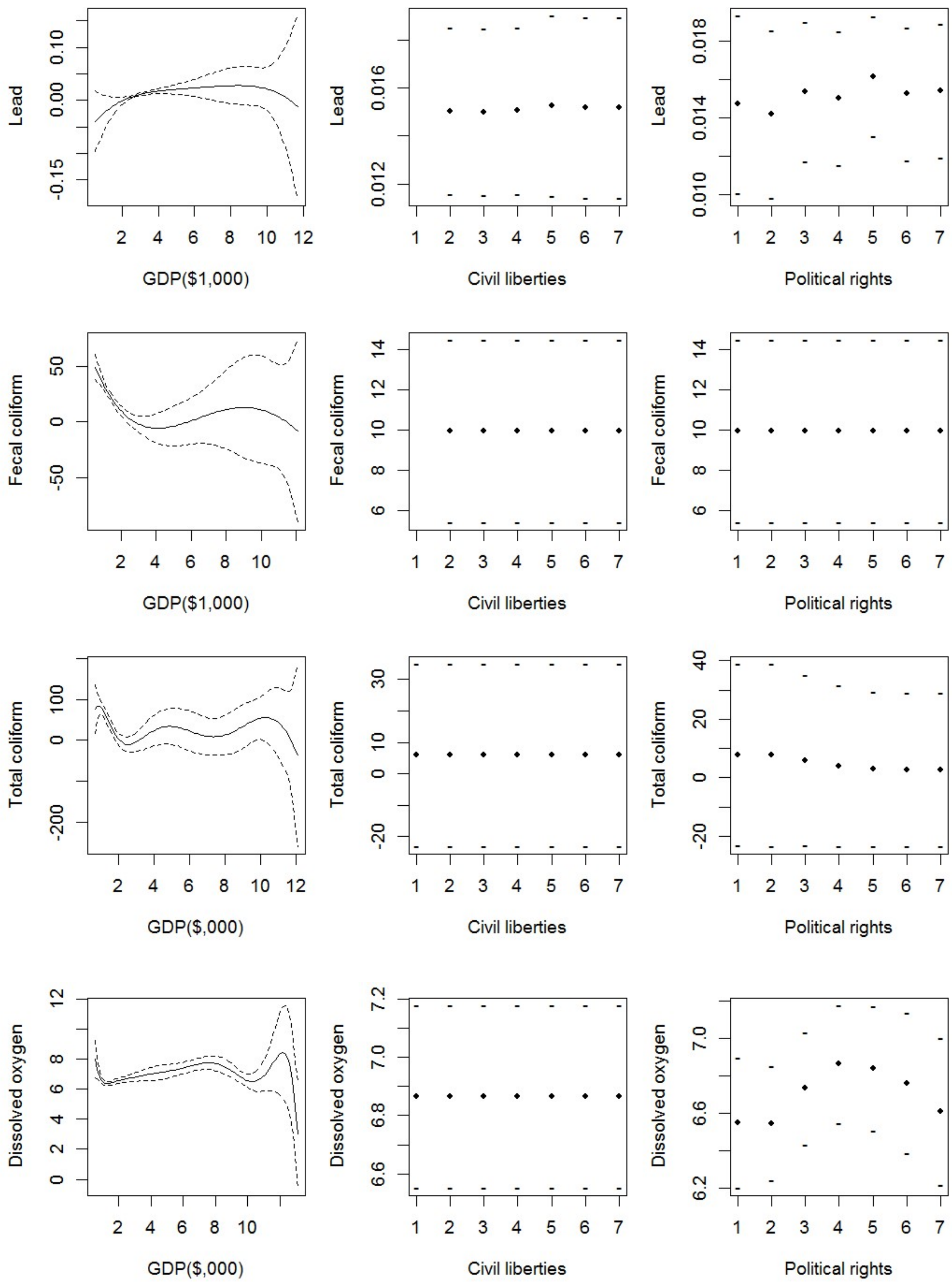
Table 2: Summary statistics for income, political variables, and debt

Variable Name.	M. type	Mean	Std. Dev.	Min	Max	Observations	
Per Capita GDP	overall	10.362	12.506	0.102	123.433	N	5447
	between		12.265	0.482	70.805	n	182
	within		3.600	-23.051	62.990	T-bar	29.93
Political Rights	overall	3.660	2.228	1.000	7.000	N	6017
	between		2.006	1.000	7.000	n	203
	within		1.006	-0.109	8.751	T-bar	29.64
Civil Liberties	overall	3.665	1.935	1.000	7.000	N	6017
	between		1.756	1.000	7.000	n	203
	within		0.839	0.588	7.635	T-bar	29.64
Debt	overall	5.184	5.942	0.000	135.376	N	3478
	between		3.317	0.084	17.891	n	127
	within		4.946	-11.697	128.428	T	27.39

Notes: Per capita GDP is in constant 2005 international dollars. The political rights and civil liberties variables each vary from 1 to 7, with 1 meaning the most political rights or civil liberties. Total debt service is in percent of GNI. The second column under the M.type in the table indicates whether the values shown are "overall", "between" or "within". "Overall" statistics are based on the total number of observations, "between" statistics describe properties between individual countries, and "within" statistics describe the properties over time. The "between" min and max refer to the minimum and maximum, respectively, of the average value for each country. The "within" min is the maximum negative deviation from each country's average. The "within" max is the maximum positive deviation from each country's average plus the global average. In the observation column, N is the number of observations, n is the number of countries, and T-bar is the average number of years per country. The data set is an unbalanced panel.

Figure 1


Nickel
-0.05 0.05 0.15 0.25
2 4 6 8 10 12
GDP($1,000)
0.050 0.060 0.070
1 2 3 4 5 6 7
Civil liberties
Political rights
Mercury
-1.0 0.0 1.0
0.3 0.4 0.5 0.6
Arsenic
0.00 0.05 0.10
2000 6000 10000
-0.02 0.00 0.02
-0.06 -0.02 0.02
Cadmium
-0.03 -0.01 0.01 0.03
0.002 0.004


Lead
GDP($1,000)
Civil liberties
Political rights
Fecal coliform
Total coliform
GDP($,000)
Dissolved oxygen

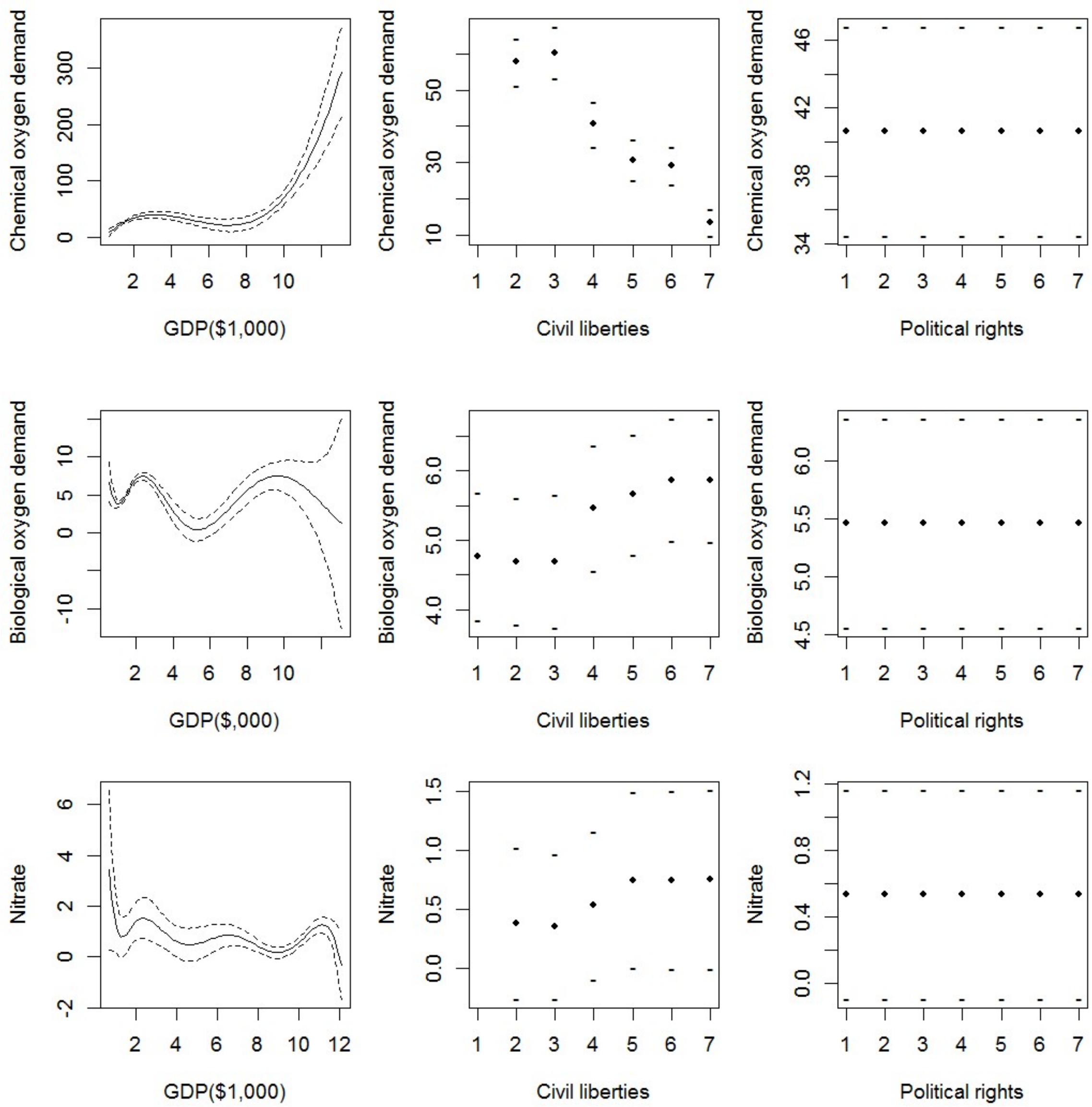


Figure 1. Relationships between pollution and per capita GDP; pollution and civil liberties; and pollution and political rights obtained from using a nonparametric instrumental variable estimation.

Notes: Dotted lines in the GDP-pollutant relationship are 95% confidence interval bands. The lightly dotted points in civil liberties and political rights are confidence interval bands. All water pollutant data are in the form of concentrations of mg/l except for the mercury data, which is in the form of μg/l; and the coliform data, which is in the form of measured count/100 ml. Per capita GDP is in thousand constant 2005 international dollars. The political rights and civil liberties variables each vary from 1 to 7, with 1 meaning the most political rights or civil liberties. In plotting the GDP-pollutant relationship, the ordinal categorical variables are fixed at their middle value of 4.