# Endogeneity in the Environmental Kuznets Curve: An Instrumental Variables Approach

## C.-Y. Cynthia Lin and Zachary D. Liscow

The effects of increasing income on environmental quality is an issue that has long puzzled economists. For over a decade, economists have theorized that a graph of environmental degradation versus income often looks something approximating an inverted-U shape, dubbed the environmental Kuznets curve (EKC) after Simon Kuznets' work in the 1950s and 1960s on income equality (Kuznets 1955, 1965). Among the reasons why economists have found the effects of increasing income on environmental quality so intriguing is that the answers to this question would help resolve fundamental issues concerning humanity's ability to develop economically, while still preserving the environment.

Some economists hypothesize that there is a causal relationship between income and environmental degradation, and that the relation is in the shape of an inverted U: as countries "get rich, … first [environmental] problems increase, and then they decrease" (Lomborg and Pope 2003, p.9). According to this theory, the solution to environmental problems is to alleviate poverty.

Other economists agree with the shape of the relationship between income and environmental degradation but disagree with the claim of causality and, more importantly, with the conclusion that by mitigating poverty, one would also improve environmental quality. Instead, they suggest that there are very important omitted variables, and that what cleaned up the environment was not rising income, but rather political institutions responding to public demand (Lomborg and Pope 2003). As Dasgupta and Maler (1995, p. 2412) state: "The connection between environmental protection and civil and political rights is a close one. As a general rule, political and civil liberties are instrumentally powerful in protecting the environmental resource base, at least when compared with the absence of such liberties in countries run by authoritarian regimes."

To examine the relationship between environmental quality, income, and political institutions, we use a multivariate regression analysis of water pollution on income and institutional variables. According to the results, evidence for an inverted-U relationship between income and environmental degradation were found for seven out of eleven water pollutants. Political institutions have a significant effect on environmental quality for five of the eleven pollutants.

The key innovation of this study is to use instrumental variables to mitigate the problems caused by simultaneity bias and omitted variable bias. For example, it is entirely plausible that water pollution harms economic development. Likewise, an omitted variable such as a cultural or geographic factor may affect both environmental quality and income. Our study is important not only because of its methodology, but also because its results will expand the state of knowledge regarding the consequences—especially on environmental quality—of income, civil liberties and political institutions.

## Data

For our measure of environmental quality, we choose to focus on water. We choose water because it seems the environmental variable most likely to show the importance of political institutions. Since water is probably the most important and fought-over public good, it is also the most politicized. As the ever-expanding human population continues to place increasing demands on the global water supply, the issue of water quality is becoming even more crucial.

We use data from the Global Environmental Monitory System GEMS/Water dataset, which consists of triennial surveys of water quality statistics from 1979 to 1999 from across the developed and developing world. This study updates the data set used by Grossman and Krueger (1995) and by Barrett and Graddy (2000) to include the years from 1991 to 2000.

The data set consists 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 one or more data point from one of GEMS/water's hundreds of sites around the world.

Following Grossman and Krueger (1995), we use the data on biological oxygen demand, chemical oxygen demand, dissolved oxygen, nitrate, arsenic, cadmium, lead, mercury, nickel, total coliforms, and fecal coliform. All data is 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 data set also includes water temperature (in degrees Celsius), which we use as control. The year assigned to each data point is the middle of the three years.

To this data, we add data on gross domestic product (GDP) per capita at purchaser's prices in constant 2000 international dollars from the World Development Indicators (WDI). We also add GDP squared and GDP cubed.

For data on political mechanisms, we use the indices on political rights and civil liberties from Freedom House. Political rights measures factors like 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 2004). Civil liberties measures the freedom of expression and belief, the ability to associate, the rule of law, and the degree of

individual autonomy. 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 a 7 in both categories for most years. Freedom House attempts to use a methodology not bound by culture, but rather using standards drawn from the Universal Declaration of Human Rights (Freedom House 2004).

Also from the World Development Indicators (WDI), we add data on the percentage of GDP that comes from manufacturing as a control. Also from the WDI dataset, we add an 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) as instruments.

Tables 1a and 1b in the online Appendix A present summary statistics for our data set.[1]

## Graphical analysis

In this section we examine the relationships among environmental quality, income and political institutions by graphical analysis. To assess these relationships at an aggregate level, we first pool the data over all years and all countries to plot each of the eleven water quality variables versus income, political rights, and civil liberties. There are thus 33 relationships in total. The plots are available in the online Appendix B.[2] In order to see the nuances of the data, we also subdivided the plots by country, by year, and by development level (OECD versus non-OECD) for each of these 33 relationships. For example, the plots by country allow easy understanding of which countries lead to an inverted-U shape.

As seen from a summary of the patterns gleaned from the exploratory plots presented in table 2 in the online Appendix A, there seems to be no connection between the type of pollutant and the relationship between the pollution's concentration and the state of political or economic development. However, there are still several interesting trends.

The concentrations of the majority of the pollutants (chemical oxygen demand, total arsenic, dissolved oxygen, total lead, total

---

[1] The online appendix A is available at: http://www.des.ucdavis.edu/faculty/Lin/water_EKC_AppA.pdf.
[2] The online appendix B is available at: http://www.des.ucdavis.edu/faculty/Lin/water_EKC_AppB.pdf.

nickel, and fecal coliform) are decreasing functions of per capita income, political rights, and civil liberties. The concentrations of only two pollutants (total cadmium and nitrate) exhibit increasing functions of per capita income, political rights, and civil liberties. 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.

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 results suggest that, to the extent that there is an EKC, it may be as much caused by political as income factors.

In order to see the nuances of the data, we also subdivided the plots in the online appendix by country. As evidenced by these plots, the shape of the relationships of each pollutant with GDP, political rights and civil liberties in the pooled plots is governed primarily by the cross-sectional variation between countries. They thus reflect cross-sectional variation in pollution levels between countries at different points in their development paths, rather than time series variation in pollution levels within countries developing over time either politically or economically. With few exceptions (i.e. Ireland), the countries in this dataset do not vary substantially in their GDP or institutional indicators over the 20-year time horizon of this dataset. Given that this study uses a dataset over a longer period of time than those used in many studies in the EKC literature, these results suggest that the EKC may not truly reflect individual countries' trajectory over time, but instead perhaps other factors not traditionally captured in the EKC literature.

## Econometric Methodology

Our regression model is the following:

$$pollution_{it} = \alpha_0 + \alpha_1 y_{it} + \alpha_2 y_{it}^2 + \alpha_3 y_{it}^3 \\ + \alpha_4 pr_{it} + \alpha_5 cl_{it} + x_{it}'\beta \\ + \varepsilon_{it},$$

where $pollution_{it}$ is the water pollutant concentration for country $i$ in year $t$, $y_{it}$ is country $i$'s per capita GDP in year $t$, $pr_{it}$ the political rights index in country $i$ in year $t$, $cl_{it}$ is the civil liberties index in country $i$ in year $t$, and $x_{it}$ is a vector of controls including population density, water temperature, year, and manufacturing value added. The cubic trend in income is consistent with previous studies (see e.g., List and Gallet 1999), while the addition of political variables is less common. Allowing for the possibility that a country's political institutions may have a lagged effect on pollution concentrations, we also run a model lagging the political variables:

$$pollution_{it} = \alpha_0 + \alpha_1 y_{it} + \alpha_2 y_{it}^2 + \alpha_3 y_{it}^3 \\ + \alpha_4 pr_{i,t-1} + \alpha_5 cl_{i,t-1} \\ + x_{it}'\beta + \varepsilon_{it}.$$

Finally, to control for any time-invariant unobservables that vary by country, we also run the following fixed effects model:

$$pollution_{it} = \alpha_0 + \alpha_1 y_{it} + \alpha_2 y_{it}^2 + \alpha_3 y_{it}^3 \\ + \alpha_4 pr_{it} + \alpha_5 cl_{it} + x_{it}'\beta \\ + \mu_i + \varepsilon_{it},$$

where $\mu_i$ is a country-specific fixed effect. A fixed effects model is more appropriate than a random effects model because the unobservables captured by the fixed effect are likely to be correlated with the regressors.

In all the models above, a negative second derivative of pollution with respect to income for some range of income would be consistent with an inverted-U shape:

$$\frac{\partial^2 pollution_{it}}{\partial y_{it}^2} = 2\alpha_2 + 6\alpha_3 y_{it} < 0.$$

If political institutions facilitate environmental improvement, we would expect a positive coefficient on the political rights and civil liberties indices, where lower values of the indices indicate a stronger political institution.

There are two types of endogeneity problems that plague regressions of environmental quality on institutional and income variables and that have been largely ignored by previous literature on the subject. One type is the simultaneity bias introduced by the reverse causality of GDP and environmental degradation. While the increases in economic activity that come along with increases in GDP may increase pollution, increases in pollution may, at the same time, harm people's health, for example, thereby reducing GDP. Output and pollution may also be jointly produced in the production process, causing GDP and pollution to be simultaneously determined.

A second type of endogeneity problem arises from omitted variable bias. While including policy variables helps reduce the problem of the endogeneity of GDP, it is still quite plausible that a third variable jointly causes both economic growth and environmental degradation—perhaps cultural or geographic factors not now in the regression formula.

In order to mitigate the problems of endogeneity, we innovate upon the previous literature by employing an instrumental variables approach for the regressions both with and without the fixed effects in order to identify the coefficient on income. The instruments are debt service and age dependency ratio. These instruments are reasonably credible instruments for GDP; while they are correlated with GDP, they do not have an effect on environmental quality, except through their effect on 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, but there is little reason to believe that countries with high debts would pollute more. Countries with a higher age dependency ratio will have lower rates of growth and GDPs, both because countries with large populations of young are likely to be less productive on average and because poorer countries tend to have this demographic profile.

For each of the 11 water pollutants, we run five regressions. The first regression is OLS where standard errors are clustered by country. The second and third regressions are instrumental variables (IV) generalized method of moments (GMM) regressions; the second regression uses the contemporaneous political variables and the third regression uses the lagged political variables. In both IV GMM regressions, per capita GDP, per capita GDP squared and per capita GDP cubed are instrumented with age dependency ratio, total debt service, age dependency ratio squared, total debt service squared, age dependency ratio cubed and total debt service cubed. For the IV GMM specifications, a robust weighting matrix that is optimal when the error term is heteroskedastic is used.

To address any potential weak instruments problem, the fourth regression is a limited information maximum likelihood (LIML) regression using age dependency ratio and total debt service as instruments for per capita GDP. We report the LIML estimate and a coverage-corrected standard error for the coefficient on per capita GDP based on the conditional likelihood ratio (CLR) approach developed by Moreira (2003). Computation of the conditional p-value for the CLR test uses the algorithm of Andrews, Moreira, and Stock (forthcoming). Andrews, Moreira, and Stock (2004) showed that the CLR test is approximately optimal. In particular, it dominates the Anderson and Rubin (1949) test and the Lagrange multiplier (score) test proposed independently by Kleibergen (2002) and Moreira (2001).

To capture any country-specific unobservables that are invariant over time, the fifth regression is a fixed effects regression where per capita GDP, per capita GDP squared and per capita GDP cubed are instrumented with age dependency ratio, total debt service, age dependency ratio squared, total debt service squared, age dependency ratio cubed and total debt service cubed.

## Regression Results

Before adjusting the standard errors, we test for heteroskedasticity using a Breusch-Pagan / Cook-Weisberg test. For all pollutants, as reported in table 4 in the online Appendix A, we reject the null of constant variance. We therefore need to adjust the standard errors for heteroskedasticity. Thus, with OLS, standard errors are clustered by country. With IV GMM, we use a robust weighting

matrix that is optimal when the error term is heteroskedastic.

To compare the instrumental variables result with OLS, we first run an OLS model for each of the pollutants. The results are reported as specification (1) in tables 4a-4k in the online Appendix A. The OLS results show no evidence for an inverted-U shaped relationship. However, the OLS results may be biased since income owing to the endogeneity of income.

For each pollutant, we conduct a Durbin-Wu-Hausman test to test for the endogeneity of income. This is a test of whether the residual from a regression of income on all the exogenous variables has a significant coefficient when added to the original model. The null hypothesis is that income is exogenous. According to the results, income is endogenous for the regressions of chemical oxygen demand, dissolved oxygen, arsenic, and fecal coliform. Thus, at least for these pollutants, instrumental variables are needed to overcome endogeneity.

Table 3 in the online Appendix A presents the results of the first stage regression of per capita GDP on the instruments and on the other exogenous variables. The F-statistic for the joint test of the instruments is 39.11 when the exogenous variables include the contemporaneous political variables and 41.12 when the exogenous variables include the lagged political variables. The instruments are thus correlated with the endogenous variable.

We use a Hansen overidentification test to test whether the instruments are uncorrelated with the error term. As reported in table 4,

for all the pollutants for which we reject exogeneity of per capita GDP (chemical oxygen demand, dissolved oxygen, arsenic, and fecal coliform) and therefore for which instruments are needed, we cannot reject the null hypothesis that the instruments are uncorrelated with the error term, so the instruments are admissible.

The results from the IV GMM specifications in tables 4a-4k are robust to whether the political variables are lagged (specification 3) or not (specification 2). To address any potential weak instruments problem, we run a conditional IV LIML regression using age dependency ratio and total debt service as instruments for per capita GDP, and report the LIML estimate and a coverage-corrected standard error for the coefficient on per capita GDP. In contrast with the OLS results, some pollutants show an inverted-U relationship under the IV GMM and conditional IV specifications. Pollutants exhibiting an inverted-U relationship have a cubic relationship with income, which leads to both a peak and a trough.

When country fixed effects are included with the IV estimation, some of the environmental Kuznets relationships go away. As with the other IV results, pollutants exhibiting an inverted-U relationship have a cubic relationship with income, which leads to both a peak and a trough. Table I lists, for each of the four IV specifications, the peak of the EKC for pollutants that exhibit an inverted-U relationship under that specification.

Thus, according to the results, evidence for an inverted-U relationship between income and environmental degradation were found for

**Table 1. Peaks for Pollutants Exhibiting an Inverted-U Relationship**

| | IV GMM | IV GMM – lagged political variables | COND IV | IV FE |
|---|---|---|---|---|
| Biological oxygen demand | $8362 | $8434 | $8194 | |
| Chemical oxygen demand | $7228 | $7234 | $8170 | Significant cubic term |
| Dissolved oxygen | | | | |
| Nitrate | | | | Significant linear term |
| Arsenic | $9055 | $10,883 | No peak | $12,468 |
| Cadmium | $10,000 | $8937 | $7927 | No peak |
| Lead | $7932 | $8104 | $8242 | |
| Mercury | | | | |
| Nickel | | | $6667 | $9000 |
| Total coliforms | | | | $12861 |
| Fecal coliform | $4298 | $4253 | $4446 | Significant linear term |

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. The 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.

The results also provide some evidence for the importance of political institutions. The political indicators have a statistically significant effect on water pollution in at least two IV specifications for five out of eleven pollutants: nitrate, arsenic, cadmium, lead, and mercury. The direction of the effect varies by pollutant and political variable. For some pollutants, political institutions have a positive effect on water pollution; for others, political institutions have a negative effect. For some pollutants, the two political variables have opposite effects on water pollution.

## Conclusions

This study is suggestive of a likely shape of the water quality-income relationship and also of the importance of political institutions. The key innovation is the use of instrumental variables and country-specific fixed effects in our econometric methodology. Evidence for an inverted-U relationship between income and environmental degradation were found for seven out of eleven water pollutants. Political institutions have a significant effect on environmental degradation for five out of eleven water pollutants.

For international donors trying to improve the lives of the world's poor, while improving the environment they inhabit, this study offers a few implications. First, at least with some of the water pollutants, political institutions matter. Second, this study suggests that water quality may affect income to a significant extent, as evidenced by the Durbin-Wu-Hausman tests and the instrumental variables regressions. If accurate, then this study provides more evidence for why water supplies should be protected and cleaned—it may help countries gain more income. Finally, the relationships between environmental degradation, income and political institutions found in this study suggest that those in the field

and academia should be open to relationships between these key components of sustainable development.

## References

Anderson, T.W., and H. Rubin. 1949. "Estimators of the parameters of a single equation in a complete set of stochastic equations", *Annals of Mathematical Statistics* **20**: 46–63.

Andrews, D.W.K., M. Moreira, and J. Stock. 2004, "Optimal invariant tests for instrumental variables regression", Unpublished manuscript.

Andrews, D.W.K., M. Moreira, and J. Stock. Forthcoming. "Performance of conditional Wald tests in IV regression with weak instruments", *Journal of Econometrics*.

Barrett, S. and K. Graddy. 2000, "Freedom, growth, and the environment", *Environment and Development Economics* **5**: 433–456.

Dasgupta, P. and K.-G. Maler. 1995. "Poverty, institutions, and the environmental resource-base", In: Behrman, J. and T.N. Srinivaan (Eds.), Handbook of Development Economics, vol. 3A. Elsevier Science, Amsterdam. (Chapter 39).

Freedom House. 2004. "Freedom in the world 2003: Survey methodology", http://www.freedomhouse.org/research/freeworld/2003/methodology.htm. Cited 10 Aug 2004.

GEMS Water. 1979-1999. "Publications and multimedia". http://www.gemswater.org/publications/index-e.html.

Grossman, G. and A. Kreuger. 1995. "Economic growth and the environment", *Quarterly Journal of Economics* **110**(2):353–377.

Kleibergen, F. 2002. "Pivotal statistics for testing structural parameters in instrumental variables regression", *Econometrica* **70**: 1781–1803.

Kuznets, S. 1955. "Economic growth and income equality", *American Economic Review* **45**(1):1–28.

Kuznets, S. 1965. *Economic growth and structural change*, New York: Norton.

List, J.A. and C.A. Gallet.1999. "The environmental Kurnets curve: does one size fit all?", *Ecological Economics* **31**: 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. Cited 12 Aug 2004.

Moreira, M. 2001. "Tests with correct size when instruments can be arbitrarily weak", Center for Labor Economics Working Paper 37, UC Berkeley.

Moreira, M. 2003. "A conditional likelihood ratio test for structural models", *Econometrica* **71**: 1027–1048.

World Development Indicators On-line. 2004. Accessed through Harvard University libraries. http://devdata.worldbank.org.ezp2.harvard.edu/dataonline/. Cited 13 Jul 2004.