Unveiling the Income-Environment Relationship: 
An Exploration into the Determinants of Environmental Quality

Nazrul Islam  
Dept. of Economics, Emory University

Jeffrey Vincent  
Harvard Institute for International Development (HIID)

Theodore Panayotou  
Harvard Institute for International Development (HIID)

First Draft: September 1996  
Current Draft: June 1998

----------------------------------------------------------------------------------------------------------------------

This paper was presented earlier at seminars organized by HIID, HIID/ADB, and Department of Economics, Emory University. We would like to thank the participants of these seminars for their comments. However, we are alone responsible for remaining errors. Research assistance by Xiang Yu is gratefully acknowledged. We would like to thank Gene Grossman and Alan Krueger for making the data sets and programs of their 1995 *QJE* paper available on the internet. Special thanks are due to Alan Krueger for his prompt response to e-mail queries. The views expressed in this paper are authors’ own and are not to be ascribed to the organizations they represent.
**Unveiling the Income-Environment Relationship:**  
An Exploration into the Determinants of Environmental Quality

1. Introduction

In recent years, the relationship between economic growth and environmental quality has drawn considerable attention from researchers. Since the most prevalent indicator of economic growth is *income level*, the question has been further concretized into one of relationship between environmental quality and level of income. A whole body of literature has now emerged addressing this narrowed down question. Shafik and Bandyopadhay (1992), Hettige, Lucas and Wheeler (1992), Selden and Song (1994), Grossman and Krueger (1995), Kauffmann, Davidsdotter, and Garnham (1995), Panayotou (1995), and Vincent (1997) are a few examples. In these works, various measures of environmental quality have been, in many different ways, regressed on income in order to ascertain the nature of this relationship. Broadly, we refer to these works as studies of the *Income-Environment Relationship (IER)*. Both cross-sectional and panel data have been used in this research. For some environmental indicators, the obtained relationship displayed an inverted-\(U\) shaped pattern, and this has given rise to the hypothesis of *Environmental Kuznets Curve (EKC)*.

So far, the IER has been generally considered in its reduced form. There are two main reasons for this. *First*, specifying a structural form of the IER is not easy. At any point of time, environmental quality is the resultant of two sets of forces: forces that generate pollution, and forces that work towards abatement of pollution. Efforts to model these different forces and their interaction are still in a preliminary stage. There are models that reproduce the inverted-\(U\) shaped relationship between income level and pollution.¹ These are useful for reduced form IER regressions. However, these do not help much with structural specification of this relationship. In some cases, pollution and abatement processes have been modeled separately but with such variables on which data are hard to obtain. This makes integration of these models and their empirical estimation difficult. These reasons explain why the reduced form approach has appealed to so many researchers. As Grossman and Krueger (1995) note, it gives the ‘net effect’ of income

---

¹ See for example, Selden and Song (1993).
on environment, and documentation of this relationship is ‘an important first step’ in the study of income-environment relationship.

However, an important limitation of the reduced form approach is that it does not reveal why the observed relationship exists. In the reduced form IER, income serves as an omnibus representative of many factors. The estimated relationship, therefore, can not throw light on the separate effects of these factors. Hence it leaves us without any handle on the relationship and without any understanding how policies can influence it. The present study takes Grossman and Krueger’s (1995) study of the reduced form IER as the point of departure and attempts to uncover the structural relationships that lie beneath it. One way to get at the structural determinants of environmental quality is to ask the question what are the forces for which income acts as the proxy in the reduced form IER. We identify three such forces and name their respective effects on the environment as: (a) the Level effect, (b) the Composition effect, and (c) the Abatement effect. We first show that the relationship among these effects is multiplicative and then specify each of these effects. Possibilities of specification of these different effects are rich. However, a priori reasoning suggests that the level effect is likely to be monotonically increasing in output per unit of area, and the composition effect is to be quadratic in output per capita. The abatement effect is likely to yield a downward sloping curve between pollution and per capita income. We embed these specifications into the multiplicative relationship to obtain an equation that makes it possible to recover the structural relationships from the estimated results.

This framework is implemented using a panel data set on suspended particulate matter (SPM) in urban air around the world. We focus on SPM for three reasons. First, it has been included in almost all the previous studies of the IER. Second, more data are available for it than any other pollutant. Third, among air pollutants, SPM is increasingly recognized as the one that has the greatest impact on human health (see Wilson and Spengler, 1996). The estimation results generally confirm the hypotheses regarding the structural relationships mentioned above. The level effect is found to be monotonically increasing with output level per unit of area. The composition effect displays an inverted-U shape with respect to the share of industry in the GDP. The abatement effect reveals itself in a declining relationship between pollution and per capita income. The pattern of decline conforms more to a backward-J than to an inverted-J.

The paper is organized as follows. Decomposition of the ambient pollution level into its structural sources is discussed in section-2. Section-3 presents the specification-framework for the
exercise. Section-4 discusses the data and econometric issues of implementation of this framework using a global SPM data set. The results are presented and discussed in section-5. Concluding observations are offered in section-6.

2. Identification of the Underlying Determinants of Environmental Quality

2.1 Decomposition of Ambient Pollution Level

The ambient pollution level in an area is influenced both by the pollution that is emitted in that area and by other factors such as location, topography, weather patterns etc. We term these later factors as ancillary determinants of environmental quality, and we shall deal with their role shortly. Abstracting from the role of these ancillary determinants, the ambient pollution level in an area can be decomposed as follows:\(^2\)

\[
\text{Ambient Pollution Level} = \frac{\text{Actual Pollution}}{\text{Area}} = \frac{\text{GDP}}{\text{Area}} \times \frac{\text{Actual Pollution}}{\text{GDP}}
\]

\[
= (\text{GDP per unit of Area}) \times (\text{Pollution Intensity of GDP})
\]

Pollution intensity of GDP, in turn, depends on two sets of forces. On the one hand are forces that work for generation of pollution. On the other hand are forces that work for abatement.\(^3\) Actual emission, and hence pollution intensity of GDP, is the resultant of these two opposing sets of forces. Algebraically, we can, therefore, write:

\[
\text{Pollution Intensity of GDP} = \frac{\text{Actual Pollution}}{\text{GDP}} = \frac{\text{Potential Pollution}}{\text{GDP}} \times \frac{\text{Actual Pollution}}{\text{Potential Pollution}}.
\]

\(^2\) Previous studies have offered similar decomposition of pollution into its various sources. For one such example, see Griffith (1994) or Cropper (1996). However, to our knowledge, no previous study has estimated a structural model based on such decomposition.

\(^3\) By ‘abatement,’ we refer to both within-process reduction of pollution and clean up at the end of the production process. The latter may also be referred to as post-process reduction of pollution.
The potential pollution-to-GDP ratio depends on the composition of GDP, while the actual-to-potential pollution ratio depends on abatement efforts. These two terms can, therefore, be referred to as the Composition effect (C) and the Abatement effect (A). It is clear that GDP per unit area represents the Level effect (L). Combining, therefore, we get the decomposition:

\[ \text{Ambient Pollution Level} = \left( \frac{\text{GDP}}{\text{Area}} \right) \times \left( \frac{\text{Composition of GDP}}{\text{Population}} \right) \times \left( \frac{\text{Abatement Efforts}}{\text{Area}} \right) = L \times C \times A. \] (3)

We base our analysis on equation (3).\(^4\) The relationship of GDP with the first term (L) is apparent. This is not so for the other two terms. Before proceeding to specification of L, C, and A, it is therefore worthwhile to consider the different ways in which income affects all these terms.

**Level Effect**

The level effect refers to income as an indicator of level of economic activity. Economic activity entails interaction with environment and is generally pollution-generating and resource-depleting. Hence, higher the level of economic activity per capita or per unit of area, the higher is likely to be the level of pollution, measured in commensurate terms. This implies a monotonically increasing relationship between income and pollution, as shown in Figure-1a.

**Composition Effect**

The composition effect depends on two other assumed relationships. Of these, the first is between income level and structure of the economy. Starting with Kuznets’ pioneering work, Chenery and others have shown that the structure of the economy, as captured by the sectoral

---

\(^4\)Note that this decomposition can be carried further to bring population density explicitly into the picture. Since GDP per unit area is the product of GDP per capita and density of population, equation (3) can also be written as

\[ \text{Ambient Pollution Level} = \left( \frac{\text{GDP}}{\text{Population}} \right) \times \left( \frac{\text{Population}}{\text{Area}} \right) \times \left( \frac{\text{Composition of GDP}}{\text{Area}} \right) \times \left( \frac{\text{Abatement Efforts}}{\text{Population}} \right). \] (4)

We use equation (3) primarily for parsimony. Also, the impact of population density can be gauged from the impact of GDP per unit of area. Another subtlety regarding this decomposition is the following. Ambient pollution level may be the result not only of current emission but also of past emission. Introduction of this distinction will require further modification of the equations above. However, current emission is likely to be the most important determinant of current ambient level. Also, current emission level is a good indicator of pollution accumulated from past. Hence, for practical purposes, equation (3) or (4) may suffice.
composition of output or employment, evolves in a predictable manner with rise of income.\textsuperscript{5} This transformation, in turn, reflects the underlying process of industrialization. This is generally captured by the evolution of industry’s share in output of the country. This share first increases and then declines as a country graduates from pre-industrial to industrial and then to post-industrial phases of development. The second assumption required for composition effect is that, as a producing sector, industry is more polluting and resource depleting than either agriculture or services. When combined, these two assumptions yield the inverted-\textit{U} shaped relationship between pollution and income level shown in Figure-1b.

\textit{Abatement Effect}

There are both demand- and supply-side aspects of abatement effect. The demand side derives from an analog of Engel’s Law regarding relationship between demand for food and level of income. At low levels of income, people are more concerned with meeting their urgent material needs, and they worry less about the quality of environment. However, as the level of income rises, people become freed up from pressing material needs and can better appreciate the value of environmental quality and, hence, start demanding it. This is the demand side of abatement.

The supply side of abatement effect refers to the following. At low levels of income, societies are less able to devote resources for abatement and pollution control measures even if there is a perceived need and demand for better environmental quality. Higher income levels make this possible. The economy can then spend more either on importing the necessary technologies for pollution abatement and clean-up or on developing these on their own by financing appropriate R&D activities. This supply side of abatement, together with the demand side discussed above, is likely to yield an inverse relationship between pollution and income. One particular hypothesis regarding this relationship is that of an inverted-\textit{J}, as shown in Figure-1c.\textsuperscript{6}

Obviously, the precise nature of the curves depicted in Figure-1 is a matter of debate. The description and the graphs presented here are merely illustrative and depend on many assumptions that may not always hold. Thus, for example, the two assumptions needed for an inverted-\textit{U} shaped composition effect may both go wrong. With inordinate use of machinery, chemical

\textsuperscript{5} This relationship between income level and the structure of the economy, in turn, has its own underlying determinants, with Engel's law regarding falling share of expenditure on food with rise of income being one.

\textsuperscript{6}Song (1994) presents a model with such an inverted-\textit{J} relationship.
fertilizers, and pesticides, modern agriculture may prove to be more detrimental to the environment than some types of manufacturing. Similarly, many lines of service industry, with their tendency to produce large amounts of solid and, sometimes, hazardous waste, may do more damage to the environment than many industrial sectors. With regard to abatement effect, it may not be the case that both demand for better environmental quality and supply of pollution control and abatement measures are delayed till the attainment of high level of income.\(^7\) Hence, the abatement effect may result in a steeper decrease of pollution even at lower levels of income. This will produce an abatement effect curve that has the shape of a backward-J instead of an inverted-J.

In view of the disputed nature of the relationships, it is necessary to allow as much flexibility in the specification of the effects as possible. This is precisely what we shall do. However, before coming to the issue of specification, it is worthwhile to check what kinds of variables correspond to the various effects that we have identified above.

### 2.2 Correspondence Between Various Effects and Possible Explanatory Variables

Table-1 gives a correspondence between the effects and the possible variables. As already mentioned, theoretical models formalizing the processes of pollution and abatement are still in a preliminary stage. Hence, the correspondence shown in the table is heuristic and not a formal one. However, note that, for level effect, the corresponding variable is unambiguously defined and is unique. This is, however, not the case with composition effect and abatement effect. For each of these, there is a multiplicity of corresponding variables. This is particularly true of abatement effect. In specifying these effects, one, therefore, faces the question which variable to include as well as the question in what form to include.

\(^7\) The analogy with the demographic transition can hardly be missed here. For long, it was postulated that the demographic transition can be achieved only at higher levels of income. However, recent experience has shown that this is not necessarily the case. Provided certain other conditions are fulfilled, fewer children and smaller families can become the norm even at relatively low levels of income. The same may be true with respect to environmental quality.
Table-1

Correspondence Between Effects and Candidate Variables

<table>
<thead>
<tr>
<th>Type of Effect</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level Effect</strong></td>
<td>• GDP per unit of area</td>
</tr>
<tr>
<td><strong>Composition Effect</strong></td>
<td>• Sectoral shares in GDP</td>
</tr>
<tr>
<td></td>
<td>• Technological features of the sectors, etc.</td>
</tr>
<tr>
<td><strong>Abatement Effect (demand side)</strong></td>
<td>• Per capita income</td>
</tr>
<tr>
<td></td>
<td>• Level of education and awareness</td>
</tr>
<tr>
<td></td>
<td>• Political rights and leadership</td>
</tr>
<tr>
<td></td>
<td>• Civil liberties</td>
</tr>
<tr>
<td></td>
<td>• Inequality of income distribution, etc.</td>
</tr>
<tr>
<td><strong>Abatement Effect (supply side)</strong></td>
<td>• Level of R&amp;D expenditure</td>
</tr>
<tr>
<td></td>
<td>• Degree of openness of the economy</td>
</tr>
<tr>
<td></td>
<td>• Direct expenditure on pollution control and abatement, etc.</td>
</tr>
</tbody>
</table>

3. Specification

3.1 Specification of the Effects

Specifications proposed for \( L, C, \) and \( A \) are given below. This is followed by discussion of their rationale.

\[
L = \alpha_0 + \alpha_1 Y + \alpha_2 Y^2, \\
C = \beta_1 Q + \beta_2 Q^2 + \beta_3 Q^3, \text{ and} \\
A = \gamma_0 + \gamma_1 I + \gamma_2 I^2 + \gamma_3 I^3,
\]

where,

\( Y \) = Output (GDP) per unit of area,  \\
\( Q \) = Indicator of composition of output,  \\
\( I \) = lagged per capita income.
Specification of the Level Effect

Specification of the level effect is somewhat made easy by the fact that the corresponding variable is relatively unambiguous and unique. In section 2 and in Figure-1a, we suggested that the level effect is likely to be monotonically increasing. A linear specification could capture this property. However, a priori, we do not know whether the increase will be linear or not. In order to allow the data to express themselves freely in this regard, we allow equation (5) to be a quadratic function of $Y$.

Specification of the Composition Effect

In case of the composition effect, for reasons mentioned above, the precise representation of $Q$ is an issue. The first and simplest approach is to take $Q$ to be equal to the share of industry in the total output. However, it is possible to contend that the share of industry does not contain the entire information regarding composition of output that is relevant for the relationship being investigated. A second approach, therefore, may be to include information regarding shares of some other sectors, like agriculture, services, etc. Also, information regarding shares may be complemented by information on some other attributes of the sectors. However, a parallel inclusion of such additional variables will make the overall equation extremely large. As a third approach, therefore, $Q$ may be taken as a composite index of the sectoral shares and attributes and be constructed and analyzed in a way similar to that of factor analysis. To keep the analysis tractable, in this exercise we take $Q$ to be the share of industry in total output. With this settled, the next question is of appropriate form of inclusion of $Q$. In section 2, the composition effect was hypothesized to be quadratic in nature. To ensure more flexibility, we allow a cubic form.\(^8\)

Specification of the Abatement Effect

In view of the multiplicity of corresponding variables, all three approaches to inclusion of variables that we discussed above in the context of composition effect are relevant for specification of the abatement effect. Obviously, per capita income level is the most important variable on the demand side of abatement. However, as we have seen, it is important on the supply

---

\(^8\) Note that the composition effect can be operative only through levels of output. Hence its specification may not require an intercept once such a term has already been included in the specification of the level effect.
side as well, working as a proxy for R&D and other pollution control expenditures. Theoretically, 
$I$ in equation (4) can also be a composite index of all or a subset of the variables corresponding to 
the demand and supply of abatement. It is also possible to think of other variables, besides income 
per capita, entering into this equation in a parallel manner. However, such entry would make the 
system unwieldy. Moreover, reliable data across countries on environmental policies and 
environmental expenditure simply do not exist. Hence, in this exercise, we set $I$ equal to per capita 
income. We use lagged income in order to allow some time lag between the rise in income and its 
effect to be transmitted to the pollution level. In section-2, we hypothesized the curve for 
abatement effect to be generally decreasing and having either an inverted or backward-$J$ shape. In 
specifying abatement effect, we allow the cubic form so that any of these different shapes can 
surface.

Finally, to keep the estimation problem tractable, we refrain from adding disturbance term 
in the specification of each of the effects separately. Instead, we allow a stochastic component to 
the overall equation for $E$, as we shall see shortly.

### 3.2 The Overall Equation and Recovery of the Structural Relationships

We now use equation (5)-(7) to substitute for $L$, $C$, and $A$ in equation (3) to get

\[(8)\]

\[E = L \ast C \ast A,\]

\[= (\alpha_0 + \alpha_1 Y + \alpha_2 Y^2) \ast (\beta_1 Q + \beta_2 Q^2 + \beta_3 Q^3) \ast (\gamma_0 + \gamma_1 I + \gamma_2 I^2 + \gamma_3 I^3).\]

There are two estimation strategies that can be adopted regarding equation (8). The first is 
to view it as a problem of non-linear estimation. The second is to try to keep the estimation linear. 
There are several problems with the non-linear estimation strategy. First, note that the equation is 
non-linear not only in variables but also in parameters. Second, the number of parameters is very 
large for a problem of non-linear estimation. Third, given the nature of non-linearity and number 
of parameters involved, convergence of the iterative procedure required for non-linear estimation 
is likely to be a problem. Fourth, even if there is convergence, it will be difficult to know whether 
the convergence is local or global. This is particularly true in view of the fact that initial values for 
these parameters required for starting the iteration are not readily available. Given this array of
problems, it is desirable and safe to keep the estimation linear. This can be done in the following way.

Multiplying out, we get the following expanded version of equation (8):

\[
E = \alpha_0 \beta_1 Y_0 Q + \alpha_0 \beta_2 Y_0 Q^2 + \alpha_0 \beta_3 Y_0 Q + \alpha_0 \beta_2 Y_0 Q^2 + \alpha_0 \beta_3 Y_0 Q^3 \\
+ \alpha_2 \beta_1 Y_0 Q^2 + \alpha_2 \beta_2 Y_0 Q^3 + \alpha_2 \beta_3 Y_0 Q^4 + \alpha_2 \beta_2 Y_0 Q^3 I + \alpha_2 \beta_3 Y_0 Q^4 I \\
+ \alpha_0 \beta_1 Y_0 Q^3 I + \alpha_1 \beta_1 Y_0 Q^4 I + \alpha_1 \beta_2 Y_0 Q^5 I + \alpha_1 \beta_3 Y_0 Q^6 I + \alpha_1 \beta_2 Y_0 Q^5 I \\
+ \alpha_2 \beta_2 Y_1 Y_0 Q^6 I + \alpha_2 \beta_3 Y_1 Y_0 Q^7 I + \alpha_2 \beta_3 Y_1 Y_0 Q^8 I \\
+ \alpha_1 \beta_1 Y_1 Y_0 Q^9 I + \alpha_1 \beta_2 Y_1 Y_0 Q^{10} I + \alpha_1 \beta_3 Y_1 Y_0 Q^{11} I + \alpha_1 \beta_1 Y_1 Y_0 Q^{12} I \\
+ \alpha_2 \beta_2 Y_2 Y_0 Q^{13} I + \alpha_1 \beta_2 Y_2 Y_0 Q^{14} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{15} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{16} I \\
+ \alpha_1 \beta_3 Y_2 Y_0 Q^{17} I + \alpha_1 \beta_3 Y_2 Y_0 Q^{18} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{19} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{20} I \\
+ \alpha_1 \beta_3 Y_2 Y_0 Q^{21} I + \alpha_1 \beta_3 Y_2 Y_0 Q^{22} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{23} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{24} I \\
+ \alpha_1 \beta_1 Y_2 Y_0 Q^{25} I + \alpha_1 \beta_2 Y_2 Y_0 Q^{26} I + \alpha_1 \beta_3 Y_2 Y_0 Q^{27} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{28} I \\
+ \alpha_1 \beta_3 Y_2 Y_0 Q^{29} I + \alpha_1 \beta_3 Y_2 Y_0 Q^{30} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{31} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{32} I \\
+ \alpha_1 \beta_3 Y_2 Y_0 Q^{33} I + \alpha_1 \beta_3 Y_2 Y_0 Q^{34} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{35} I + \alpha_2 \beta_3 Y_2 Y_0 Q^{36} I.
\]

Using notations for reduced form coefficients, we can write equation (9) as follows:

\[
E = \pi_0 Q + \pi_2 Q^2 + \pi_3 Q^3 + \pi_4 YQ + \pi_6 YQ^2 + \pi_8 YQ^3 + \pi_4 Y^2 Q + \pi_8 Y^2 Q^2 + \pi_9 Y^2 Q^3 \\
+ \pi_0 QI + \pi_1 QI^2 + \pi_3 QI^3 + \pi_4 YQI + \pi_6 YQ^2 I + \pi_8 YQ^3 I + \pi_4 Y^2 QI + \pi_8 Y^2 Q^2 I + \pi_9 Y^2 Q^3 I \\
+ \pi_1 Q^2 I + \pi_2 Q^2 I^2 + \pi_4 Q^2 I^3 + \pi_6 Q^2 I^4 + \pi_8 Q^2 I^5 + \pi_9 Q^2 I^6 \\
+ \pi_1 Q^3 I + \pi_2 Q^3 I^2 + \pi_4 Q^3 I^3 + \pi_6 Q^3 I^4 + \pi_8 Q^3 I^5 + \pi_9 Q^3 I^6 \\
+ \pi_1 Q^4 I + \pi_2 Q^4 I^2 + \pi_4 Q^4 I^3 + \pi_6 Q^4 I^4 + \pi_8 Q^4 I^5 + \pi_9 Q^4 I^6 \\
+ \pi_1 Q^5 I + \pi_2 Q^5 I^2 + \pi_4 Q^5 I^3 + \pi_6 Q^5 I^4 + \pi_8 Q^5 I^5 + \pi_9 Q^5 I^6 \\
+ \pi_1 Q^6 I + \pi_2 Q^6 I^2 + \pi_4 Q^6 I^3 + \pi_6 Q^6 I^4 + \pi_8 Q^6 I^5 + \pi_9 Q^6 I^6.
\]

The correspondence between the \( \pi \)'s and the underlying structural coefficients is obvious from the formulations above. Note that even the reduced form equation (10) is non-linear in variables. However, this non-linearity can be handled by corresponding transformation of the variables on the right hand side. The other issue is whether the structural parameters can be uniquely recovered from the reduced form coefficients. There are thirty-six reduced form coefficients to correspond to ten structural parameters. Clearly, there are numerous restrictions on the reduced form coefficients.\(^9\) Numerical recovery of the structural parameters will require

\[^9\text{One way of formulating these restrictions is as follows:}\]

\[
(11) \quad \frac{\pi_1}{\pi_{10}} = \frac{\pi_2}{\pi_{11}} = \frac{\pi_3}{\pi_{12}} = \frac{\pi_4}{\pi_{13}} = \frac{\pi_5}{\pi_{14}} = \frac{\pi_6}{\pi_{15}} = \frac{\pi_7}{\pi_{16}} = \frac{\pi_8}{\pi_{17}} = \frac{\pi_9}{\pi_{18}} = \frac{\gamma_0}{\gamma_1},
\]

\[
(12) \quad \frac{\pi_{10}}{\pi_{19}} = \frac{\pi_{11}}{\pi_{20}} = \frac{\pi_{12}}{\pi_{21}} = \frac{\pi_{13}}{\pi_{22}} = \frac{\pi_{14}}{\pi_{23}} = \frac{\pi_{15}}{\pi_{24}} = \frac{\pi_{16}}{\pi_{25}} = \frac{\pi_{17}}{\pi_{26}} = \frac{\pi_{18}}{\pi_{27}} = \frac{\gamma_1}{\gamma_2}, \text{and}
\]
imposition of these constraints and checking whether we can find a *unique* correspondence between the reduced form coefficients and the structural parameters. However, imposition of the restrictions will bring us back to non-linear estimation. The important thing to realize here is that, for our purpose, it is *not* necessary to recover all the structural parameters *separately.* This is because recovery of the structural *relations* is not the same as recovery of structural *parameters.* To recover structural relations, estimates of some of the *combinations* of the structural parameters may be sufficient. This is particularly true when the relationships considered are of polynomial form, as it is in our case. For such relationships, numerical values of individual parameters are of limited value because the effect of an explanatory variable on the dependent variable is given by a function of several parameters and specified values of the remaining explanatory variables. Thus, we can recover the structural relationships without having to impose the restrictions, and this saves us from the pitfalls of non-linear estimation. In the following, we illustrate how the specified structural relationships can be recovered from the linear estimation results.

**Recovery of the Level Effect**

Based on the estimation results of the overall equation, the level effect can be traced out using the following equation:

\[
\hat{E}_L = (\pi_1 \bar{Q} + \pi_2 \bar{O}^2 + \pi_3 \bar{O}^3 + \pi_{10} \bar{Q} \bar{I} + \pi_{11} \bar{O}^2 \bar{I} + \pi_{12} \bar{O}^3 \bar{I}) \\
+ (\pi_{19} \bar{Q}^2 \bar{I}^2 + \pi_{20} \bar{O}^2 \bar{I}^2 + \pi_{21} \bar{O}^3 \bar{I}^2 + \pi_{28} \bar{O}^2 \bar{I}^3 + \pi_{29} \bar{O}^3 \bar{I}^3 + \pi_{30} \bar{O}^3 \bar{I}^3) \\
+ (\pi_{3} \bar{Q} + \pi_{5} \bar{O}^2 + \pi_{6} \bar{O}^3 + \pi_{11} \bar{Q} \bar{I} + \pi_{14} \bar{O}^2 \bar{I} + \pi_{15} \bar{O}^3 \bar{I}) \\
+ (\pi_{22} \bar{Q} \bar{I}^2 + \pi_{23} \bar{O}^2 \bar{I}^2 + \pi_{24} \bar{O}^3 \bar{I}^2 + \pi_{31} \bar{Q} \bar{I}^3 + \pi_{32} \bar{O}^2 \bar{I}^3 + \pi_{33} \bar{O}^3 \bar{I}^3) \gamma Y \\
+ (\pi_{9} \bar{Q} + \pi_{5} \bar{O}^2 + \pi_{16} \bar{Q} \bar{I} + \pi_{17} \bar{O}^2 \bar{I} + \pi_{18} \bar{O}^3 \bar{I} \\
+ \pi_{25} \bar{Q} \bar{I}^2 + \pi_{26} \bar{O}^2 \bar{I}^2 + \pi_{27} \bar{O}^3 \bar{I}^2 + \pi_{34} \bar{O} \bar{I}^3 + \pi_{35} \bar{O}^2 \bar{I}^3 + \pi_{36} \bar{O}^3 \bar{I}^3) \gamma^2 Y^2.
\]

(14)

As the equation shows, the level effect is traced out by fixing the values of \(Q\) and \(I\) at their respective sample means\(^\text{10}\) and computing the predicted values of \(E\) for different values of \(Y\). The

\[
\pi_{19} = \pi_{20} = \pi_{21} = \pi_{22} = \pi_{23} = \pi_{24} = \pi_{25} = \pi_{26} = \pi_{27} = \gamma_2 \\
\pi_{28} = \pi_{29} = \pi_{30} = \pi_{31} = \pi_{32} = \pi_{33} = \pi_{34} = \pi_{35} = \pi_{36} = \gamma_3
\]

(13)

\(^{10}\) A symbol with bar above it represents the mean of the corresponding variable.
expressions in the three sets of parentheses respectively give the intercept and the coefficients of the linear and quadratic terms of this relationship.

**Recovery of the Composition Effect**

Similarly, recovery of the composition effect can proceed using the following equation:

\[
\hat{E}_c = (\hat{\pi}_1 + \hat{\pi}_4 \bar{Y} + \hat{\pi}_7 \bar{Y}^2 + \hat{\pi}_{10} \bar{I} + \hat{\pi}_{13} \bar{YI} + \hat{\pi}_{16} Y^2 \bar{I} \\
+ \hat{\pi}_{19} \bar{I}^2 + \hat{\pi}_{22} \bar{YI}^2 + \hat{\pi}_{25} \bar{Y}^2 \bar{I}^2 + \hat{\pi}_{28} \bar{I}^3 + \hat{\pi}_{31} \bar{YI}^3 + \hat{\pi}_{34} \bar{Y}^2 \bar{I}^3) Q \\
(\hat{\pi}_2 + \hat{\pi}_5 \bar{Y} + \hat{\pi}_8 \bar{Y}^2 + \hat{\pi}_{11} \bar{I} + \hat{\pi}_{14} \bar{YI} + \hat{\pi}_{17} Y^2 \bar{I} \\
+ \hat{\pi}_{20} \bar{I}^2 + \hat{\pi}_{23} \bar{YI}^2 + \hat{\pi}_{26} \bar{Y}^2 \bar{I}^2 + \hat{\pi}_{29} \bar{I}^3 + \hat{\pi}_{32} \bar{YI}^3 + \hat{\pi}_{35} \bar{Y}^2 \bar{I}^3) Q^2 \\
(\hat{\pi}_3 + \hat{\pi}_6 \bar{Y} + \hat{\pi}_9 \bar{Y}^2 + \hat{\pi}_{12} \bar{I} + \hat{\pi}_{15} \bar{YI} + \hat{\pi}_{18} Y^2 \bar{I} \\
+ \hat{\pi}_{21} \bar{I}^2 + \hat{\pi}_{24} \bar{YI}^2 + \hat{\pi}_{27} \bar{Y}^2 \bar{I}^2 + \hat{\pi}_{30} \bar{I}^3 + \hat{\pi}_{33} \bar{YI}^3 + \hat{\pi}_{36} \bar{Y}^2 \bar{I}^3) Q^3.
\]

We see that the composition effect is traced out by fixing the values of \(Y\) and \(I\) at their respective sample means and computing the values of \(E\) for different values of \(Q\). The expressions in the parentheses correspond to the coefficients of the linear, quadratic, and cubic terms of this relationship.

**Recovery of the Abatement Effect**

Finally, the abatement effect can be recovered using the following equation:

\[
\hat{E}_a = (\pi_1 \bar{Q} + \pi_2 \bar{Q}^2 + \pi_3 \bar{Q}^3 + \pi_4 \bar{Y} \bar{Q} + \pi_5 \bar{Y} \bar{Q}^2 + \pi_6 \bar{Y} \bar{Q}^3 + \pi_7 \bar{Y} \bar{Q}^2 + \pi_8 \bar{Y} \bar{Q}^2 + \pi_9 \bar{Y} \bar{Q}^3) \\
+ (\pi_{10} \bar{Q} + \pi_{11} \bar{Q}^2 + \pi_{12} \bar{Q}^3 + \pi_{13} \bar{Y} \bar{Q} + \pi_{14} \bar{Y} \bar{Q}^2 + \pi_{15} \bar{Y} \bar{Q}^3 + \pi_{16} \bar{Y} \bar{Q}^2 + \pi_{17} \bar{Y} \bar{Q}^2 + \pi_{18} \bar{Y} \bar{Q}^3) \bar{I} \\
+ (\pi_{19} \bar{Q} + \pi_{20} \bar{Q}^2 + \pi_{21} \bar{Q}^3 + \pi_{22} \bar{Y} \bar{Q} + \pi_{23} \bar{Y} \bar{Q}^2 + \pi_{24} \bar{Y} \bar{Q}^3 + \pi_{25} \bar{Y} \bar{Q}^2 + \pi_{26} \bar{Y} \bar{Q}^2 + \pi_{27} \bar{Y} \bar{Q}^3) \bar{I}^2 \\
+ (\pi_{28} \bar{Q} + \pi_{29} \bar{Q}^2 + \pi_{30} \bar{Q}^3 + \pi_{31} \bar{Y} \bar{Q} + \pi_{32} \bar{Y} \bar{Q}^2 + \pi_{33} \bar{Y} \bar{Q}^3 + \pi_{34} \bar{Y} \bar{Q}^2 + \pi_{35} \bar{Y} \bar{Q}^2 + \pi_{36} \bar{Y} \bar{Q}^3) \bar{I}^3.
\]

In tracing out the abatement effect, the values of \(Y\) and \(Q\) are fixed at their respective sample means, and the values of \(E\) are computed for the corresponding values of \(I\). Expressions in the parentheses give the intercept and the coefficients of the linear, quadratic, and cubic terms, respectively.

The above completes the exposition of the framework for identification of the various determinants of environmental quality. We now proceed to an implementation of this framework.
on the basis of a global data set on suspended particulate matter (SPM) in urban air. We begin by considering certain general issues of implementation.

4. Data and Estimation Issues

4.1 General Issues regarding Data

We use the Global Environmental Monitoring System (GEMS) data on suspended particulate matter in urban air, as compiled by Grossman and Krueger (1995), for the exercise in this paper. This provide information regarding our variable $E$. For the right hand side variables we rely on, in addition to the Grossman-Krueger data, the World Tables and the Summers-Heston data set. As Grossman and Krueger explained, for many of these variables, site-specific data are not available, and we are constrained to use the economy-wide values for all the sites in a particular country. GDP per unit area (square mile) of a city is obtained by multiplying the GDP per capita for the corresponding country by the density of population of the city concerned. This gives the variable $Y$. It is not perfect because ideally we would like to multiply the GDP per capita of the city by its density of population. However, it is difficult to get data on per capita GDP for cities. Similarly, for $Q$, we would ideally like to have the share of industry in the GDP of the individual cities concerned. However, it is difficult to obtain data of that kind. With regard to the abatement effect, we use lagged per capita GDP as the variable $I$.\footnote{It is again a case of substitution of the economy-wide value of the variable for all cities in a country.} The total number of observations in Grossman-Krueger’s data set was 1021. However, for 120 of these observations, data on $Q$ were not available, and therefore we had to drop these. Our sample, therefore, differs slightly from that of Grossman and Krueger.

4.2 Some Descriptive Aspects of the Data

The data come from 23 countries for the period of 1977-88. Canada, China, and the USA are, by far, the most important sources of data on SPM. The total number of observations tends to be higher for the early eighties than for more recent period. Based on the country means, the value
of \( E \) (measured as microgram per cubic meter of air) ranges from a low of 19 for New Zealand to a high of 530 for Pakistan. For GDP (measured in 1985 PPP adjusted thousand dollars) per square mile, \( Y \), we find an enormous range of values, with a minimum of 0.0125 for Ghana and a maximum of 35.55 for Venezuela. For \( Q \), the share of industry in GDP, we find that, based on country averages, its value ranges from a minimum of 0.1947 for Ghana to 0.4588 for Brazil. \( I \), per capita lagged income, is the variable whose country wise distribution we are most familiar with. The country means of \( I \) (measured in 1985 PPP adjusted thousand dollars) ranges from a low of 0.645 for India to a high of 14.881 for the US.

The SPM data for each country generally come from more than one city and monitoring station. In total, the number of cities represented in the data set is 56, and the number of sites is 145. The number of sites is generally the highest for the USA, averaging to about 20. China comes in second in terms of number of sites, Canada third. The rest of the countries generally have less than five sites, with Japan exceeding this mark in some recent years. There is also considerable flux or, if you would like, turnover in the sites reporting SPM data. The highest number of sites for which we have data in any particular year\(^{12} \) is 82. This may be compared with 145, the overall number of sites represented in the data set. Additional information about the distribution properties of the variables concerned is presented in the form of histograms in Figures 2 to 5.

### 4.3 Econometric Issues

The income-environment relationship is basically a dynamic relationship with respect to a site or a country. However, to the extent that long time series on the relevant variables are difficult to come by for individual sites or countries, researchers have used cross-country data for estimating the relationship. Use of panel data is obviously an improvement in this regard because it brings into play both the time series and cross-sectional dimensions of the data. In using either cross-sectional or panel data, one needs to deal with the role of, what we earlier termed as, ancillary determinants of ambient pollution level. The panel data setup is helpful in this regard. One property of the ancillary determinants is that they are specific to monitoring site, and they

---

\(^{11}\) It is obtained by averaging the per capita GDP of the preceding three years, and is the same as the lagged income variable in Grossman and Krueger (1995). It would have been better to use GNP, instead of GDP, in constructing this variable. Unfortunately, the Summers-Heston data set does not provide information on GNP.

\(^{12}\) Which happens to be the year of 1982.
either do not change over time or change very slowly. Also, often these ancillary determinants are
either unobservable or unmeasurable. Even if they are observable and measurable, in most cases,
data on them are not available. In panel estimation, it is possible to control for these ancillary
determinants through inclusion of time-invariant, site-specific individual effects, $\mu_i$.

Most often, $\mu_i$’s are modeled either as fixed effects or as random effects. Grossman and
Krueger (1995) used the random-effects specification. Under this specification, the error term $\varepsilon_{it}$
is assumed to be a composite of the time invariant term $\mu_i$ and a transitory term $\nu_{it}$ that varies
both across time and across site or station. That is, we have

$$ (17) \quad \varepsilon_{it} = \mu_i + \nu_{it}, $$

There are quite a few arguments for random effects specification. First, the Bruesch and
Pagan Lagrangian Multiplier test applied to the results from estimating equation (9) (discussed
more fully below) overwhelmingly rejects the null $H_0 : \text{Var}(\mu_i) = 0$. The sample value of the $\chi^2$
statistic is 701.44, and the associated $p$-value is less than 0.01 percent. This strongly supports the
composite error specification.

Second and more helpful in this regard are the results obtained from analysis of variance of
the concerned variables. These results are presented in Table-2. As we can see from this table, the
overwhelming part of variation in the data lies along the ‘between’ dimension of data, rather than
the ‘within’ dimension. This is true not only for the dependent variable, $E$, but also for all the
explanatory variables. Going by the sum of squares (SS), between variation accounted for 93
percent of total variation in $E$. This proportion for $Y$, $Q$, and $I$ is 97, 95, and 99 percent
respectively. Between sum of squares (SSB) exceeds the within sum of squares (WSS) by factors
of 13, 35, 21, and 111 for $E$, $Y$, $Q$, and $I$, respectively. Similarly, the between mean sum of squares
(MSB) exceeds within mean sum of squares (MSW) by 69, 183, 106, and 709 times, respectively,
for the same variables as above. Given this structure of the data, the fixed effects estimator, which
throws out all the between-variation in the data, cannot be a good choice. All these considerations
favor the random effects specification of $\mu_i$.

Before proceeding with estimation of equation (9), it is worthwhile to consider the reduced
form IER for SPM that was estimated by Grossman and Krueger.
5. Results for Suspended Particulate Data

5.1 The Reduced Form Income-Environmental Relationship

The equation estimated by Grossman and Krueger had the following specification:

\[ E_{it} = \delta_0 + \delta_1 N_{it} + \delta_2 N_{it}^2 + \delta_3 N_{it}^3 + \lambda_1 LN_{it} + \lambda_2 LN_{it}^2 + \lambda_3 LN_{it}^3 + \epsilon_{it}. \]

where \( N_{it} \) and \( LN_{it} \) are current and lagged per capita income, respectively.\(^{13}\) Note that \( LN \) of this equation is the same as \( I \) of equation (5) to (16). We use different notations to avoid confusion between current and lagged values of income. To the extent that we had to drop a few observations, we re-estimate this equation on the basis of our sample to see whether the slight change in the sample makes any important difference. The numerical results can be seen in Table-A1. The graph of the predicted IER is presented in Figure-6. It is clear that the modification of the sample does not matter much. The graph in Figure-6 is almost identical to the SPM graph of Grossman and Krueger (1995, p. 363). It has a mild cubic nature with a high and a low peak at $-7.745 and $22,530, respectively. Both these peaks are outside the range of the income data and are results of extrapolation. Within the range, the curve slopes downward everywhere, although the slope becomes less steep as income attains higher values. The analysis of the same data using equation (8) will show how widely different forces are working underneath this seemingly placid relationship!

5.2 Regression Results

The overall statistics of the regression of equation (9) are shown in Table-3. As is known, the \( R^2 \) from random effects GLS estimation do not have the classical properties. The reported values shown in the table are \( R^2 \)'s from simple bivariate regressions of the fitted values for within, between, and overall equations on the respective set of original values. Going by these measures, we find that the regression does remarkably well. The overall fit, as measured by \( R^2 \)-overall, is 63 percent, which is quite high given the nature of the data. This is also confirmed by

---

\(^{13}\) The equation also included other variables representing some of the ancillary determinants of environmental quality.
the \( p \)-value (less than 0.01 percent) of the \( \chi^2 \) statistic for the test of joint significance of the explanatory variables. As expected, the fit is much better along the \textit{between} dimension of the data. \( R^2 \)-\textit{between} equals 0.60. However, even along the \textit{within} dimension, the fit is quite respectable, with a \( R^2 \)-\textit{within} equaling 0.21. Looking at the estimated values of \( \sigma_\mu \), \( \sigma_v \), and \( \sigma_\varepsilon \), we see that \( \mu_i \), the time-invariant individual component of the error, prove to be dominant. Its standard deviation, \( \sigma_\mu \), is about 2.2 times larger than \( \sigma_v \), the standard deviation of the pure transitory component, \( \varepsilon_i \).

The basic regression results are provided in Table-4, which gives the parameter estimates, standard errors, and levels of significance. Recall that the total number of interactive regressors is 36. As many as 29 of these prove to be significant at 10 percent level. Ten of these are significant at 5 percent level, and another 5 prove to be significant even at one percent level. As all these regressors are basically various multiplicative combinations of only three underlying variables, the danger of multicollinearity loomed large. In view of this, the results on the significance of the individual right hand side terms of the regression are encouraging. The regression also includes a time trend variable (\textit{year}) in order to capture the secular global advance in technology and awareness. This variable proves to be significant at 10 percent level. The overall intercept term also proves to be highly significant.

Although we may be curious about the sign and numerical magnitudes of the coefficients of the individual terms, examining these may not be that productive because the relationships of interest have been specified in the form of polynomials. Individual coefficients are of limited value in revealing these relationships. This brings us to the task of tracing out the polynomials.

\textbf{5.3 Estimated Structural Relationships}

\textit{Estimated Level Effect}

We first turn to the recovery of the level effect. This is done using equation (14). The numerical results for this equation turn out to be the following:

\begin{equation}
\hat{E}_L = -15.298 + 6.425 Y - 0.0343 Y^2
\end{equation}
As equation (14) shows, these results are on the basis of evaluation of the equation at the sample mean levels of $Q$ and $I$. The year was set equal to 1982 which is the closest integer to the value of the sample mean of the variable “year.” The graph corresponding to this level effect equation is presented in Figure-7. The methodology of producing these graphs is similar to that used in Grossman and Krueger (1995). Thus, the circles, or, if you like, the bubbles in the graph give information regarding fit of the curve. These are obtained on the basis of the residuals. The distance of these circles from the line of the graph indicate the size of the residuals, and the diameters of the circles are proportional to the number of observations in the corresponding interval of the variable on the horizontal axis, in the present case, $Y$.

We find that the curve for level effect is, indeed, monotonically increasing (within the range of $Y$ considered) albeit with decreasing slope. The coefficient on the quadratic term is negative but small relative to the coefficient on the linear term. Table-5 presents the SMP levels predicted by this equation for some specified levels of GDP per square mile, $Y$. It also shows the numerical magnitudes of the slopes of the curve at those levels of $Y$. These give the marginal increases in the SPM level for a $1,000 increase in GDP per square mile, as predicted by the level effect alone. The aspects of increasing level and decreasing slope of the level effect curve are both apparent from the numbers of column (2) and (3) of Table-5. Finally, the numbers in column (4) of this table show the change in SMP between the levels of $Y$ shown in the first column.

Overall, therefore, we find that, other things remaining constant, the ambient SPM level increases with increase in GDP per unit of area. For the relevant range, this relationship appears to be almost linear. However, the slope actually decreases with increasing $Y$ albeit at a very slow pace. Extrapolation based on equation (19) shows that the level effect tapers off (slope being zero) at $Y$ equaling to $93,796!$ This indicates that the level effect is certain to bedevil environment for a long time to come.

**Estimated Composition Effect**

---

14 Only the last two digits were used to denote the years. Hence, the actual number entering for the year variable in evaluating this equation is 82. The same is done to derive the numerical evaluation of the equations for composition and abatement effects below.

15 In producing this graph and computing the level figures presented in Table-5, the constant term of equation (12) was augmented by a value of 16 to avoid negative ranges. This modification does not affect its slope and the changes in the SPM levels over intervals of $Y$ computed on its basis.
The composition effect is recovered using equation (15). This equation, when evaluated at sample means of \( Y \) and \( I \), gives the following results:

\[
\hat{E}_c = 673.207 - 818.703 \ Q + 30186.179 \ Q^2 - 33595.984 \ Q^3
\]

As can be seen, the equation for the composition effect has large numerical coefficients for all the terms of the polynomial. This is because the numerical values of \( Q \) are much smaller than that of either \( Y \) or \( I \). It has a low peak at \( Q \) equal to 0.21, a high peak at \( Q \) equal to 0.39, and a point of inflection at a \( Q \) equal to 0.30. The graph corresponding to the equation is presented in Figure-8.\(^{16}\) Clearly, for most part of the range of \( Q \), the relationship is quadratic, with an inverted-U shape. It is only for very low values of \( Q \) (under 0.2) that the curve has a declining segment. However, as the bubble for this interval shows, the fit of the curve for this range is not that great. The location of the bubble far below the curve reveals that the actual values of \( E \) for this range of \( Q \) were much lower than those predicted by the curve. Also, the size of the bubble shows that the number of observations belonging to this range is small. This may entice one to do away with the cubic specification and constrain the composition effect to be quadratic. We, however, refrain from doing so in order to preserve the generality of the framework.

Table-6 shows some computed values of \( E \) on the basis of the equation for composition effect. The slope figures are presented in column (3).\(^{17}\) Note that the range of values of \( Q \) is theoretically limited to \([0,1]\). In actual data, it is limited to an even shorter interval. That is why, even though the slope figures are rather high, the relevant range over which \( E \) changes with changes in \( Q \) is not that large. This can be seen from column (4) of Table-6. The numbers in this column give the changes in the SPM level, as predicted by equation (20), across intervals of \( Q \) shown in column (1). We can also see that, barring the first entry, these bear out the largely quadratic nature of the curve. An initial phase of increase is followed by a phase of decrease.

Thus, the estimated composition effect curve demonstrates, by and large, a hump-shaped relationship of the SPM level with the share of industry in GDP, \( Q \). The peak is not reached until a long range of values of \( Q \) is passed, during which the SPM level increases almost linearly. Beyond

\(^{16}\) In order to avoid the negative range, the equation was normalized upward through augmentation of the constant term by 115. Obviously, this does not affect the slopes and the changes in the SPM levels over intervals of \( Q \) computed on the basis of the equation.

\(^{17}\) The level values are conditional on the normalization.
that level, the ambient level of SPM seems to be negatively related with $Q$, when the levels of $Y$ and $I$ are held constant. Note that this inverted-$U$ curve should not be confused with the humped curve that we presented in Figure-1b. There the pollutant level was graphed against income level. In contrast, what we have here on the horizontal axis is $Q$, and not income level. Nevertheless, we find a generally inverted-$U$ shaped relationship between $E$ and $Q$.

**Estimated Abatement Effect**

Finally, we come to the abatement effect. This is recovered using equation (16). Evaluation of this equation at sample means of $Y$ and $Q$ yields the following result:

\[ \hat{E}_A = 272.35 - 54.10 I + 4.233 I^2 - 0.1007 I^3 \]

This equation has a low peak at $I$ equaling $9,856$, and a high peak at $I$ equaling $18,168$. It also has an inflection point at $I$ equaling $14,012$. However, clearly, the linear and quadratic terms dominate this equation. The graph corresponding to this equation is presented in Figure-9.

In our discussion of section-2, we noted that the hypothesis most often put forward regarding abatement effect is one of an inverted-$J$ shape. However, we also observed that the abatement effect curve may also have a backward-$J$ shape. The curve in Figure-9 resembles more a backward-$J$ than an inverted-$J$. After displaying a steep decline over lower intervals of $I$, the curve flattens out. In fact, after an income level of $9,856$, the curve displays an upward slope, albeit mild. However, since the cubic term in the equation enters negatively, this upward trend is soon arrested, and beyond an income level of $18,168$, the slope of the curve again turns negative.

The figures of Table-7 give a numerical illustration of this pattern of movement. Recall that the inverted-$J$ hypothesis predicts that one has to wait until a high level of income is reached before the beneficial impact of income on pollution reduction becomes effective. In contrast, a backward-$J$ pattern suggests that increase in income has large pollution-reducing effect even at low levels of income. Given that most of the countries of the world are characterized by low level of income (see Figure-3), the finding of a backward-$J$ shaped abatement effect curve may prove encouraging.

Note that although this is a graph with pollution on the vertical axis and level of per capita income on the horizontal axis, it is not the same as the graphs that depict reduced form relationship between these two variables. The curve in Figure-9 shows the effect on pollution of
increase in income when other important variables, like $Y$ and $Q$, are held constant. The advantage is that we can now see the effect of per capita income on environmental quality when this effect is not confounded by the effects of rising GDP per unit of area and changes in the composition of output.

6. Concluding Remarks

The exercise above has shown that it is possible to unveil the income-environment relationship, which has until now remained shrouded in mystery. We have argued that in the reduced form IER, income actually proxies for a number of different forces that influence environmental quality in different ways. In particular, we identify Level effect, Composition effect, and Abatement effect. We start by deriving a multiplicative relationship among these effects and allow sufficient flexibility in their individual specifications. This gives an equation with structural moorings. We estimate this equation on the basis of global data on suspended particulate matter in urban air. The structural relationships are then recovered from the estimation results.

The estimated curves of the Level, Composition, and Abatement effects are generally found to conform to the a priori theoretical predictions. The level effect is found to be monotonically increasing over the relevant range of values of GDP per unit of area, $Y$. The composition effect displays mostly an inverted-$U$ shape. The abatement effect curve proves to be generally declining.

This does not mean that all the results are as expected. The composition effect shows a declining segment for the lower range of $Q$, which is something difficult to explain. The fit for that part of the curve is, however, poor. Similarly, the abatement effect curve appeared to have the shape of a backward-$J$ rather than of an inverted-$J$.

The results obtained in this paper should be considered as preliminary. There are several reasons for this. First, we do not claim that the specifications used in this paper are final. Second, and more important, there were numerous gaps between specification and implementation. These gaps arose mainly because of data limitations. For example, theory requires data on the explanatory variables to be site-specific. In reality, in many cases, we have to be satisfied with country-level data. Also, theory points to a host of variables that correspond to the Composition
and Abatement effects. In actual implementation, we use just one variable for each of these
effects. Even with respect to the variables that we used, there were limitations.

Therefore, there is considerable scope to improve upon the results presented in this paper. However, what the exercise shows is that income-environment relationship needs not remain
shrouded in mystery. Despite data limitations, it is possible to go beyond the reduced form and uncover the underlying structural relationships. In this paper, we have provided a framework for unveiling the income-environment relationship and have illustrated on the basis of the global SMP
data that it actually works. With more research, this framework can certainly be improved and extended.
References


Summers, Robert and Alan Heston (1988), A New Set of International Comparisons of Real


Abstract

This paper decomposes the reduced form income-environmental relationship into its structural sources. It identifies a *level effect*, a *composition effect*, and an *abatement effect*. This decomposition is implemented using global data on solid particulate matter in air. The level effect and composition effects are found to conform to their *a-priori* expected linear and quadratic forms, respectively. The abatement effect is found to be generally downward sloping and of backward-J shape. (JEL O0, Q0, Z0)
Nor is there much satisfaction in contemplating the world with nothing left to the spontaneous activity of nature; with every rood of land brought into cultivation, which is capable of growing food for human beings; every flowery waste or natural pasture ploughed up, all quadrupeds or birds which are not domesticated for man's use exterminated as his rivals for food, every hedgerow or superfluous tree rooted out, and scarcely a place left where a wild shrub or flower could grow without being eradicated as a weed in the name of improved agriculture. If the earth must lose that great portion of its pleasantness which it owes to things that the unlimited increase of wealth and population would extirpate from it, for the mere purpose of enabling it to support a larger, but not a better or happier population, I sincerely hope, for the sake of posterity, that they will be content to be stationary, long before necessity compels them to it.

(John Stuart Mill, *Principles of Political Economy*, Book IV, Chapter IV)
Different Effects of Income on Environment

a) Level Effect

b) Composition Effect

c) Abatement Effect
Figure 2: Frequency Distribution of SPM Level (E)

Median Concentration of SPM in Air (microgram per cubic meter)
Figure 3: Frequency Distribution of GDP per square mile (Y)

Figure 4: Frequency Distribution of Share of Industry in GDP (Q)
Reduced Form Relationship between Income and SPM Level Suspended Particulate Matter (SPM) in Urban Air

Figure-6

Median SPM Level (microgram per cubic meter of air)

GDP Per Capita ($000 in 1985 PPP prices)

0 2 4 6 8 10 12 14 16 18

0 50 100 150 200 250 300

Figure-6
LEVEL EFFECT
Suspended Particulate Matter, SPM

Median SPM Level (microgram per cubic meter of air)

GDP Per Square Mile ($'000, 1985 PPP)

Figure-7
COMPOSITION EFFECT
Suspended Particulate Matter (SPM) in Air

Median SPM Level (microgram per cubic meter)

Figure-8

Share of Industry in GDP

Figure-8
ABATEMENT EFFECT
Suspended Particulate Matter (SPM) in Air

Median SPM Level (microgram per cubic meter)

Lagged Per Capita GDP ($'000 in 1985 PPP prices)

Figure-9
Table-2

Results from Analysis of Variance of the Variables

<table>
<thead>
<tr>
<th>Item</th>
<th>$SMP$ Level (E)</th>
<th>GDP per sq. mile (Y)</th>
<th>Share of industry in GDP (Q)</th>
<th>Lagged GDP per capita (I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of squares, Total (SST)</td>
<td>15,239,065</td>
<td>73.216</td>
<td>4.1732</td>
<td>32,097</td>
</tr>
<tr>
<td>Degrees of freedom for SST</td>
<td>900</td>
<td>900</td>
<td>900</td>
<td>900</td>
</tr>
<tr>
<td>Sum of squares, Between</td>
<td>14,144,495</td>
<td>71.176</td>
<td>3.979</td>
<td>31,861</td>
</tr>
<tr>
<td>Degrees of freedom for SSB</td>
<td>144</td>
<td>144</td>
<td>144</td>
<td>144</td>
</tr>
<tr>
<td>Sum of squares, Within (SSW)</td>
<td>1,094,570</td>
<td>2.040</td>
<td>0.194</td>
<td>286</td>
</tr>
<tr>
<td>Degrees of freedom for SSW</td>
<td>756</td>
<td>756</td>
<td>756</td>
<td>756</td>
</tr>
<tr>
<td>Mean sum of squares, Total (MST)</td>
<td>16,932</td>
<td>81</td>
<td>0.0046</td>
<td>36</td>
</tr>
<tr>
<td>Mean sum of squares, Between (MSB)</td>
<td>98,226</td>
<td>494</td>
<td>0.0276</td>
<td>221</td>
</tr>
<tr>
<td>Mean sum of squares, Within (MSW)</td>
<td>1,448</td>
<td>3</td>
<td>0.0003</td>
<td>0.312</td>
</tr>
<tr>
<td>SSB as ratio of SST</td>
<td>0.93</td>
<td>0.97</td>
<td>0.95</td>
<td>0.99</td>
</tr>
<tr>
<td>SSW as ratio of SST</td>
<td>0.07</td>
<td>0.03</td>
<td>0.05</td>
<td>0.01</td>
</tr>
<tr>
<td>SSB as ratio of SSW</td>
<td>12.92</td>
<td>34.89</td>
<td>20.51</td>
<td>111.40</td>
</tr>
<tr>
<td>MSB as ratio of MST</td>
<td>5.80</td>
<td>6.06</td>
<td>6</td>
<td>6.20</td>
</tr>
<tr>
<td>MSW as ratio of MST</td>
<td>0.09</td>
<td>0.03</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>MSB as a ratio of MSW</td>
<td>67.84</td>
<td>183.13</td>
<td>106.15</td>
<td>709.16</td>
</tr>
</tbody>
</table>

Test for equality of Within and Between Variances

<table>
<thead>
<tr>
<th>Value of F-statistic</th>
<th>67.84</th>
<th>183.15</th>
<th>107.72</th>
<th>708.73</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value of F</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Bartlett’s Test for equality of within group variances

<table>
<thead>
<tr>
<th>Value of Chi-square statistic</th>
<th>1150.63</th>
<th>1810.64</th>
<th>522.08</th>
<th>956.19</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value of the Chi-square</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
### Table-3

Panel Regression Using Particulate Data  
(Overall Statistics)

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations</td>
<td>901</td>
</tr>
<tr>
<td>Number of countries</td>
<td>23</td>
</tr>
<tr>
<td>Number of cities</td>
<td>56</td>
</tr>
<tr>
<td>Number of sites</td>
<td>145</td>
</tr>
<tr>
<td>R-square, Within</td>
<td>0.2088</td>
</tr>
<tr>
<td>R-square, Between</td>
<td>0.6003</td>
</tr>
<tr>
<td>R-square, Overall</td>
<td>0.6341</td>
</tr>
<tr>
<td>Sample value of $\chi^2$ for overall fit</td>
<td>439.36</td>
</tr>
<tr>
<td>p-value of the $\chi^2$</td>
<td>0.0000</td>
</tr>
<tr>
<td>Estimated value of $\sigma_{p,i}$</td>
<td>71.2670</td>
</tr>
<tr>
<td>Estimated value of $\sigma_{v,u}$</td>
<td>33.7140</td>
</tr>
<tr>
<td>Estimated value of $\sigma_{e}$</td>
<td>78.8392</td>
</tr>
</tbody>
</table>
### Detailed Results of the Panel Regression

*Dependent Variable: Median SPM level in air*

<table>
<thead>
<tr>
<th>RHS Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Z-value</th>
<th>p-value of z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>797.94</td>
<td>290.03</td>
<td>2.751</td>
<td>0.006</td>
</tr>
<tr>
<td>Year</td>
<td>-1.52</td>
<td>0.89</td>
<td>-1.856</td>
<td>0.063</td>
</tr>
<tr>
<td>$Q$</td>
<td>-5859.07</td>
<td>4821.78</td>
<td>1.215</td>
<td>0.224</td>
</tr>
<tr>
<td>$Q^2$</td>
<td>30952.03</td>
<td>22071.12</td>
<td>1.402</td>
<td>0.161</td>
</tr>
<tr>
<td>$Q^3$</td>
<td>-45756.06</td>
<td>28913.42</td>
<td>1.583</td>
<td>0.114</td>
</tr>
<tr>
<td>$YQ$</td>
<td>753.47</td>
<td>441.76</td>
<td>1.706</td>
<td>0.088</td>
</tr>
<tr>
<td>$YQ^2$</td>
<td>-4760.31</td>
<td>2600.88</td>
<td>-1.830</td>
<td>0.067</td>
</tr>
<tr>
<td>$YQ^3$</td>
<td>6880.05</td>
<td>3749.30</td>
<td>1.835</td>
<td>0.067</td>
</tr>
<tr>
<td>$Y^2Q$</td>
<td>-47.64</td>
<td>16.86</td>
<td>-2.826</td>
<td>0.005</td>
</tr>
<tr>
<td>$Y^2Q^2$</td>
<td>262.49</td>
<td>98.12</td>
<td>2.675</td>
<td>0.007</td>
</tr>
<tr>
<td>$Y^2Q^3$</td>
<td>-342.41</td>
<td>140.42</td>
<td>-2.439</td>
<td>0.015</td>
</tr>
<tr>
<td>$QI$</td>
<td>1349.17</td>
<td>2235.84</td>
<td>0.603</td>
<td>0.546</td>
</tr>
<tr>
<td>$Q^2I$</td>
<td>-16692.65</td>
<td>12685.82</td>
<td>-1.316</td>
<td>0.188</td>
</tr>
<tr>
<td>$Q^3I$</td>
<td>32914.63</td>
<td>17849.66</td>
<td>1.844</td>
<td>0.065</td>
</tr>
<tr>
<td>$YQI$</td>
<td>-605.21</td>
<td>309.47</td>
<td>-1.956</td>
<td>0.051</td>
</tr>
<tr>
<td>$YQ^2I$</td>
<td>3944.03</td>
<td>1786.98</td>
<td>2.207</td>
<td>0.027</td>
</tr>
<tr>
<td>$YQ^3I$</td>
<td>-6020.66</td>
<td>2556.56</td>
<td>-2.355</td>
<td>0.019</td>
</tr>
<tr>
<td>$Y^2QI$</td>
<td>31.13</td>
<td>10.65</td>
<td>2.922</td>
<td>0.003</td>
</tr>
<tr>
<td>$Y^2Q^2I$</td>
<td>-176.65</td>
<td>61.92</td>
<td>-2.853</td>
<td>0.004</td>
</tr>
<tr>
<td>$Y^2Q^3I$</td>
<td>243.02</td>
<td>89.49</td>
<td>2.716</td>
<td>0.007</td>
</tr>
<tr>
<td>$QI^2$</td>
<td>-488.31</td>
<td>396.91</td>
<td>-1.230</td>
<td>0.219</td>
</tr>
<tr>
<td>$Q^2I^2$</td>
<td>3907.22</td>
<td>2362.68</td>
<td>1.654</td>
<td>0.098</td>
</tr>
<tr>
<td>$Q^3I^2$</td>
<td>-6759.88</td>
<td>3515.11</td>
<td>-1.923</td>
<td>0.054</td>
</tr>
<tr>
<td>$YQI^2$</td>
<td>104.46</td>
<td>53.17</td>
<td>1.965</td>
<td>0.049</td>
</tr>
<tr>
<td>$YQ^2I^2$</td>
<td>-659.35</td>
<td>313.74</td>
<td>-2.102</td>
<td>0.036</td>
</tr>
<tr>
<td>$YQ^3I^2$</td>
<td>992.52</td>
<td>461.34</td>
<td>2.151</td>
<td>0.031</td>
</tr>
<tr>
<td>$Y^2QI^2$</td>
<td>-4.43</td>
<td>1.81</td>
<td>-2.443</td>
<td>0.015</td>
</tr>
<tr>
<td>$Y^2Q^2I^2$</td>
<td>25.22</td>
<td>10.63</td>
<td>2.373</td>
<td>0.018</td>
</tr>
<tr>
<td>$Y^2Q^3I^2$</td>
<td>-35.04</td>
<td>15.58</td>
<td>-2.249</td>
<td>0.025</td>
</tr>
</tbody>
</table>
Table-4 (continued)

Detailed Results of the Panel Regression
(Dependent Variable: Median SPM level in air)

<table>
<thead>
<tr>
<th>RHS Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>Z-value</th>
<th>p-value of z</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Qi^3$</td>
<td>28.46</td>
<td>19.74</td>
<td>1.442</td>
<td>0.149</td>
</tr>
<tr>
<td>$Qi^2I^3$</td>
<td>-202.36</td>
<td>121.73</td>
<td>1.662</td>
<td>0.096</td>
</tr>
<tr>
<td>$Qi^3I^3$</td>
<td>332.58</td>
<td>188.19</td>
<td>1.767</td>
<td>0.077</td>
</tr>
<tr>
<td>$YQi^3$</td>
<td>-4.71</td>
<td>2.58</td>
<td>1.826</td>
<td>0.068</td>
</tr>
<tr>
<td>$YQi^2I^3$</td>
<td>29.19</td>
<td>15.60</td>
<td>1.871</td>
<td>0.061</td>
</tr>
<tr>
<td>$YQi^3I^3$</td>
<td>-43.61</td>
<td>23.60</td>
<td>1.848</td>
<td>0.065</td>
</tr>
<tr>
<td>$Y^2Qi^3$</td>
<td>0.18</td>
<td>0.09</td>
<td>2.033</td>
<td>0.042</td>
</tr>
<tr>
<td>$Y^2Qi^2I^3$</td>
<td>-1.01</td>
<td>0.52</td>
<td>1.952</td>
<td>0.051</td>
</tr>
<tr>
<td>$Y^2Qi^3I^3$</td>
<td>1.41</td>
<td>0.77</td>
<td>1.830</td>
<td>0.067</td>
</tr>
</tbody>
</table>
### Table 5

#### Level Effect

(Partial effect of GDP per square mile on the level of SPM in air)

<table>
<thead>
<tr>
<th>GDP per square mile (Y)</th>
<th>SPM Level in the air (E)</th>
<th>Slope of the Abatement Curve ( \frac{\partial E}{\partial Y} )</th>
<th>Change in E over the interval ( \Delta E )</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>1.344</td>
<td>6.418</td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>3.906</td>
<td>6.391</td>
<td>2.562</td>
</tr>
<tr>
<td>1,000</td>
<td>7.093</td>
<td>6.357</td>
<td>3.187</td>
</tr>
<tr>
<td>5,000</td>
<td>31.971</td>
<td>6.083</td>
<td>24.878</td>
</tr>
<tr>
<td>10,000</td>
<td>61.527</td>
<td>5.740</td>
<td>29.556</td>
</tr>
<tr>
<td>15,000</td>
<td>89.370</td>
<td>5.398</td>
<td>27.844</td>
</tr>
<tr>
<td>20,000</td>
<td>115.501</td>
<td>5.055</td>
<td>26.131</td>
</tr>
<tr>
<td>25,000</td>
<td>139.920</td>
<td>4.713</td>
<td>24.418</td>
</tr>
<tr>
<td>30,000</td>
<td>162.625</td>
<td>4.370</td>
<td>22.706</td>
</tr>
<tr>
<td>35,000</td>
<td>183.618</td>
<td>4.028</td>
<td>20.993</td>
</tr>
<tr>
<td>40,000</td>
<td>202.899</td>
<td>3.685</td>
<td>19.281</td>
</tr>
</tbody>
</table>

**Notes:**

1. The SPM level figures in the table are conditional on the normalization chosen.
2. \( \Delta E \) represent change in E associated with the change in Y over the previous figure on this variable in the table. Thus, for example, SPM level change by 2.562 as a result of a partial effect of increase in Y from $100 to $500. These change figures do not depend on the normalization.
3. The slope figures are also invariant to normalization.
### Table-6

Composition Effect

(Partial effect of the share of industry in GDP on the SPM level in air)

<table>
<thead>
<tr>
<th>Share of industry in the GDP ($Q$)</th>
<th>Level of SPM in air ($E$)</th>
<th>Slope of the Composition Effect curve</th>
<th>Change in $E$ over the interval ($\Delta E$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>126.754</td>
<td>-1393.58</td>
<td></td>
</tr>
<tr>
<td>0.20</td>
<td>90.546</td>
<td>-138.75</td>
<td>-36.208</td>
</tr>
<tr>
<td>0.25</td>
<td>104.480</td>
<td>612.14</td>
<td>13.934</td>
</tr>
<tr>
<td>0.30</td>
<td>143.361</td>
<td>859.09</td>
<td>38.880</td>
</tr>
<tr>
<td>0.35</td>
<td>181.990</td>
<td>602.10</td>
<td>38.629</td>
</tr>
<tr>
<td>0.40</td>
<td>195.172</td>
<td>-158.83</td>
<td>13.181</td>
</tr>
<tr>
<td>0.45</td>
<td>157.708</td>
<td>-1423.70</td>
<td>-37.464</td>
</tr>
<tr>
<td>0.50</td>
<td>44.403</td>
<td>-3192.51</td>
<td>-113.306</td>
</tr>
</tbody>
</table>

**Notes:**

1. The SPM level figures in the table are conditional on the normalization chosen.
2. $\Delta E$ represent change in $E$ associated with the change in $S$ over the previous figure on this variable in the table. Thus, for example, SPM level change by 13.934 as a result of a partial effect of increase in $Q$ from 0.20 to 0.25. These change figures do not depend on the normalization.
3. The slope figures have not been presented because the value of $Q$ is numerically bounded between (0,1) and hardly reaches 1.
### Table 7

**Abatement Effect**

*(Partial Effect of Lagged Per Capita GDP on SPM Level in Air)*

<table>
<thead>
<tr>
<th>Lagged Per Capita GDP (I)</th>
<th>Level of SMP in air (E)</th>
<th>Slope of the Abatement Curve ((\partial E/\partial I))</th>
<th>Change in the SMP level over the interval ((\Delta E))</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>246.353</td>
<td>-49.938</td>
<td>-23.961</td>
</tr>
<tr>
<td>1,000</td>
<td>222.392</td>
<td>-45.931</td>
<td>-42.101</td>
</tr>
<tr>
<td>2,000</td>
<td>180.291</td>
<td>-38.371</td>
<td>-34.843</td>
</tr>
<tr>
<td>3,000</td>
<td>145.448</td>
<td>-31.416</td>
<td>-28.190</td>
</tr>
<tr>
<td>4,000</td>
<td>117.258</td>
<td>-25.065</td>
<td>-22.141</td>
</tr>
<tr>
<td>5,000</td>
<td>95.118</td>
<td>-19.318</td>
<td>-16.696</td>
</tr>
<tr>
<td>6,000</td>
<td>78.422</td>
<td>-14.175</td>
<td>-11.855</td>
</tr>
<tr>
<td>7,000</td>
<td>66.567</td>
<td>-9.636</td>
<td>-7.618</td>
</tr>
<tr>
<td>8,000</td>
<td>58.949</td>
<td>-5.701</td>
<td>-3.986</td>
</tr>
<tr>
<td>9,000</td>
<td>54.963</td>
<td>-2.371</td>
<td>-0.958</td>
</tr>
<tr>
<td>10,000</td>
<td>54.005</td>
<td>0.355</td>
<td>1.466</td>
</tr>
<tr>
<td>11,000</td>
<td>55.471</td>
<td>2.477</td>
<td>3.286</td>
</tr>
<tr>
<td>12,000</td>
<td>58.757</td>
<td>3.995</td>
<td>4.502</td>
</tr>
<tr>
<td>13,000</td>
<td>63.259</td>
<td>4.908</td>
<td>5.113</td>
</tr>
<tr>
<td>14,000</td>
<td>68.372</td>
<td>5.217</td>
<td>5.120</td>
</tr>
<tr>
<td>15,000</td>
<td>73.493</td>
<td>4.923</td>
<td>4.523</td>
</tr>
<tr>
<td>16,000</td>
<td>78.016</td>
<td>4.023</td>
<td>3.322</td>
</tr>
<tr>
<td>17,000</td>
<td>81.338</td>
<td>2.520</td>
<td>1.517</td>
</tr>
<tr>
<td>18,000</td>
<td>82.855</td>
<td>0.413</td>
<td>-0.893</td>
</tr>
<tr>
<td>19,000</td>
<td>81.962</td>
<td>-2.299</td>
<td>-3.907</td>
</tr>
<tr>
<td>20,000</td>
<td>78.055</td>
<td>-5.615</td>
<td>-3.907</td>
</tr>
</tbody>
</table>

**Notes:**

1. The SPM level figures in the table are conditional on the normalization chosen.
2. \(\Delta E\) represent change in E associated with the change in I over the previous figure on this variable in the table. Thus, for example, SPM level change by -23.961 as a result of a partial effect of increase in I from $500 to $1,000. These change figures do not depend on the normalization.
3. The slope figures are also invariant to normalization.
Regression Results for the Reduced Form Income-Environment Relationship

**Overall Statistics of the Regression:**

- Number of observations = 901
- R-square, Within = 0.0004
- R-square, Between = 0.4208
- R-square, Overall = 0.4857
- Chi-square = 96.40
- p-value of Chi-square = 0.000

Dependent variable = Median SPM level (microgram per cubic meter)

<table>
<thead>
<tr>
<th>RHS Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>z-value</th>
<th>p-value of z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>235.11</td>
<td>18.27</td>
<td>12.869</td>
<td>0.000</td>
</tr>
<tr>
<td>N</td>
<td>37.27</td>
<td>22.38</td>
<td>1.664</td>
<td>0.096</td>
</tr>
<tr>
<td>N$^2$</td>
<td>-5.06</td>
<td>2.96</td>
<td>-1.711</td>
<td>0.087</td>
</tr>
<tr>
<td>N$^3$</td>
<td>0.16</td>
<td>0.10</td>
<td>1.616</td>
<td>0.106</td>
</tr>
<tr>
<td>LN</td>
<td>-47.08</td>
<td>24.56</td>
<td>-1.917</td>
<td>-0.055</td>
</tr>
<tr>
<td>LN$^2$</td>
<td>4.65</td>
<td>3.46</td>
<td>1.345</td>
<td>0.179</td>
</tr>
<tr>
<td>LN$^3$</td>
<td>-0.15</td>
<td>0.13</td>
<td>-1.152</td>
<td>0.249</td>
</tr>
</tbody>
</table>

Notes:

1. N denotes per capita GDP
2. LN denotes lagged per capital GDP defined as average per capita GDP over the previous three years.