Household Income and Pollution: Implications for the Debate about the Environmental Kuznets Curve Hypothesis

Florenz Plassmann  
Department of Economics  
State University of New York at Binghamton  
Binghamton, NY 13902-6000, USA  
Email: fplass@binghamton.edu; Phone: 607 (777)-4304  

and  

Neha Khanna∗, ∗∗  
Department of Economics & Environmental Studies Program  
State University of New York at Binghamton  
Binghamton, NY 13902-6000, USA  
Email: nkhanna@binghamton.edu; Phone 607 (777)-2689

Abstract
The results of many studies that examine the relationship between income and pollution with multi-country, macro-level panel data sets are sensitive to the number and type of countries included in the analysis as well as to model specification. We argue that this sensitivity is the result of several aggregation biases in multi-country, macro-level panel data sets. Using 1990 count data at the census tract level for the United States, we find relationships between household income and carbon monoxide, ground level ozone, and coarse particulate matter pollution that are robust with respect to changes in model assumptions and data measurement. Our results suggest that the income level at which households are willing to reduce their exposure to pollution depends on the nature of pollution.

Journal of Economic Literature Classification Codes: Q25, Q01, C35, C15.

Keywords: correlated count data, MCMC, NAAQS, turning point

∗ Corresponding author.

** We thank Terence Fitz-Simmons for making the count data available to us, David Mintz for providing us with the ambient concentrations data, and Lisa Brooks for providing GIS support. Barry Jones and Alok Bohara provided helpful suggestions and comments on an earlier draft.
I. INTRODUCTION

The March 2005 special issue of this Journal included several papers on the Environmental Kuznets Curve (EKC) hypothesis. Economists have suggested two primary explanations for this ‘inverted U-shaped’ relationship between pollution and income. The first relies on the idea that pollution first rises and then falls as an economy passes through different stages of development. The second explanation, based on consumer preferences, implies that accepting higher pollution as the price of higher personal income becomes increasingly unpalatable as household income increases.1 Barbier (1997, p. 370) and Carson et al. (1997, p.434) have argued that other explanations like changes in technology, civil and political liberties, trade policy, and environmental policy, are simply the vehicles that enable consumers to reveal their preferences for environmental quality.2

The special issue included papers focusing on EKC drivers such as economic development (see Marcotullio et al., 2005) and technological diffusion (see, for example, Stern, 2005). This paper heeds the call from the guest editors of this special issue (Leifman and Heil, 2005, pp. 13-14) and provides an ‘alternative perspective’ by focusing on the role of consumer demand for environmental quality. We estimate a relationship between household income and exposure to pollution using highly disaggregated data for the United States, with a view to finding the income level at which consumers are willing to reduce their exposure to pollution.

---

1 In the absence of technological progress, more production also leads to more pollution (the “scale effect”), but this effect by itself does not yield a non-monotonic relationship between growth and pollution.

2 For example, it is unlikely that technological development will by itself lead to less pollution intensive techniques. It is more intuitive that the development of less pollution intensive techniques occurs because there is a demand for improved environmental quality that ensures that resources are channeled into this type of research.
We subject our estimated pollution-income relationship to various tests of robustness and find that, unlike earlier studies, our results are remarkably robust to changes in assumptions about the underlying distribution, whether we analyze the pollutants individually or jointly by incorporating the observed correlation between them, and to different functional forms and sets of covariates. We believe that this robustness is due to the fact that we use ‘the right data for the right problem’.

Several recent papers, for example, List and Gallet (1999), Harbaugh et al. (2002), Millimet et al. (2003), and Aldy (2005, this Journal) have shown that estimated EKC relationships are highly sensitive to changes in the underlying data and model specification. We find it somewhat surprising that there is no clear evidence in favor of an inverted-U relationship between pollution and income because many governments have passed legislation to regulate and reduce pollution as their citizens have become aware of its harmful effects, and we would expect the data to reflect this. Datasets that have been used to analyze the EKC hypothesis also include countries/regions that have and have not yet experienced the structural shift to a more service-oriented economy.

We argue that earlier studies used data that made it difficult to detect the relationship between economic growth and pollution. First, given the relatively short time span that available data sets cover, empirical analyses largely rely on cross-sectional data (Chimeli and Braden, 2005), and as such, have a better chance of capturing the effects of preferences for environmental quality than of changes in economic structure. Second, earlier studies have typically used multi-country data sets that are likely to suffer from aggregation biases caused by combining local environmental data with national economic data, as well as intertemporal and spatial differences in preferences. We argue that the effect of changes in the demand for environmental quality is
better captured by disaggregate data that link local pollution to local socio-economic characteristics. Three earlier papers that utilize more disaggregated state level data for the United States, namely List and Gallet (1999), Millimet et al. (2003) and Aldy (2005), find that the estimated pollution-income relationships do, indeed, vary across states, lending support to our premise that consumer preferences matter, and that highly aggregated data may be inappropriate for estimating such relationships.

Chimeli and Braden (2005) provide additional theoretical support for utilizing cross-sectional data to estimate the EKC. Using an infinite horizon dynamic model, they argue that the cross-sectional pollution-income relationship is likely to be different from a pollution-income relationship estimated over time because the sources of income differences across countries at a point in time differ from those within a country over time. They show that an inverted U-shaped pollution-income relationship is possible in a cross-section of countries in the steady state, even if the trajectory of environmental quality to the steady state is monotonic.

We support our claim by analyzing the relationship between household income and air pollution using data on three pollutants: carbon monoxide (CO), ground level ozone (O₃), and coarse particulate matter (PM₁₀), using 1990 census tract data for the United States. All three pollutants have well known adverse health effects (primarily respiratory and cardio-vascular) and the United States Environmental Protection Agency (EPA) has established National Ambient Air Quality Standards (NAAQS) for these gases above which ambient concentrations are considered harmful to human health. Since we focus on the relationship between household income and (perceived) air quality, we measure pollution by the number of days during which the concentrations of these three pollutants exceeded their respective NAAQS at 704 locations. Exceedences of the NAAQS are a widely publicized measure of local air quality and we prefer
them to data on ambient pollutant concentrations which typically do not convey any meaningful information regarding air quality to the lay consumer.

The number of days during which the concentration of a pollutant exceeded its NAAQS is a non-negative integer. To accommodate the count data nature of our dependent variables we analyze the data using Poisson-lognormal models, while accounting for the correlation in pollution levels at each site. For PM$_{10}$, we find consistent evidence that the relationship is concave and non-monotonic. We find a concave but monotonically increasing relationship for O$_3$, and a moderately convex and monotonically increasing relationship for CO. Our results are robust with respect to changes in model specification and data measurement.

A caveat regarding the interpretation of our results is worth noting. Our critique of the data used in earlier analyses is motivated by the claim that the estimable relationship between household income and pollution is driven by consumer preferences, but our results say little about the claim that economic growth ultimately lowers pollution. Although the demand for improved environmental quality increases with income, households can satisfy this demand either by implementing cleaner technologies or, if possible, by spatially separating their consumption from production. In the latter case, households simply relocate to less polluted regions, or “export” the pollution to other regions (see Rothman, 1997), and global or even national pollution may remain unchanged while local pollution decreases. Thus, we estimate a specific pollution-income relationship for the United States. To the extent that the EKC relationship is rooted in consumer preferences that may vary spatially and intertemporally, we believe that it is necessary to first estimate local (regional or country specific) relationships before attempting to construct a global EKC (see also, Chimeli and Braden, 2005).
II. CHOOSING THE RIGHT DATA

(a) *Three types of aggregation biases*

Most studies have tested the EKC hypothesis with multi-country panel data sets that contain annual information on either nation-wide emissions or local concentrations of various pollutants (usually measured at several sites in a few major cities within each country).\(^3\) These measurements of pollution are regressed on variables measured at the country-level. The principal advantage of these data sets is the length and breadth of information that they contain.

However, using pooled cross-section and time-series data from different countries to estimate a global pollution-income relationship may lead to at least three types of aggregation biases. First, ambient concentrations data (such as the widely used GEMS data) reflect site-specific pollution measured at various locations, while most of the variables on the right-hand side of the regression equation are measured at the country level (most notably, the measure of income is usually per capita GDP). Changes in local pollution are more likely to be correlated with changes in the local economy than with changes in broad economic aggregates such as per capita GDP.\(^4\)

A second type of aggregation bias results from including heterogeneous countries in the same data set. Stern and Common (2001) and Harbaugh *et al.* (2002) suggest that estimates of an EKC relationship are sensitive to the range of countries included in the analysis. List and

---


\(^4\) In addition, the lack of variation of country-level variables across sites makes it inappropriate to treat site-specific measurements as uncorrelated observations (see Moulton, 1990).
Gallet (1999) and Aldy (2005) indicate that the relationship may differ across regions within a country, and Millimet et al. (2003) have shown that even the pollution-income relationships estimated with state level panel data for the United States are highly sensitive to modeling assumptions. This is not surprising if the relationship is driven primarily by local preferences for environmental quality (see Lieb, 2002, and Plassmann and Khanna, 2003), because there is no reason to expect that such preferences are identical across nations or regions. Furthermore, Chimeli and Braden (2005) argue that the use of panel data ignores country-specific factors that might give rise to the EKC, and creates problems of bias and inconsistency when these factors are correlated with income.

A third type of aggregation bias stems from the fact that the relationship between income and pollution may be driven by people’s awareness of the harmful effects of pollution, and that the relationship shifts over time as this awareness increases. Estimates of EKC relationships with time series data face the same problem as estimates of consumer demand curves: if the data represent observations from different curves rather than observations along a single curve, then estimates obtained from these data indicate the path along which the relationship has changed over time instead of the relationship itself.\(^5\) For example, List and Gallet (1999) and Millimet et al. (2003) analyze the EKC hypothesis using state level panel data for the United States from

\(^5\) If the change in preferences were identical in all countries in every year, then it would be easy to estimate its effect with year dummies. Because it is likely that the change in preferences is not homogeneous across countries, we would need country-specific year dummies to estimate the effect for each country. This would require a data set with more than one observation per country-year. If the entries in such a data set came from different regions within the country, we would need to assume that preferences change in the same way in all regions. If the data were collected in a single region at different points in time during the year, we would need to assume that preferences do not change during the year but only between years. Both assumptions are unlikely to be valid.
1929-1994 and Aldy (2005) uses similarly disaggregated data from 1960-1999. All three studies show that the estimated pollution-income relationships are region specific and sensitive to modeling assumptions. This suggests that it is necessary to either explicitly account for changes in environmental awareness, or to use data that cover a time span during which preferences for environmental quality remained unchanged.

(b) Cross sectional data for a single country instead of multi-country panel data sets

We believe that these aggregation biases explain why national and regional level panel data do not yield robust estimates of the EKC. A simple way to avoid these biases is to use cross-section data for a single country at the smallest feasible level of aggregation. Such data are likely to provide information about the part of the pollution-income relationship that stems from consumer preferences. By using cross section data we capture a snap-shot of consumer preferences and eliminate the impact of changes in people’s awareness of the harmful effects of pollution. The main conceptual disadvantage of cross-section data is that they may not reflect the change in pollution that occurs as the economy goes through different stages of development. However, available multi-country panel data sets contain environmental information that rarely spans more than 20 or 30 years, and even less for many developing countries. Changes in economic structure occur relatively slowly and it is likely that the multi-country panel data sets that are typically used in the existing EKC literature do not reflect these changes very well either.

---

6 Three notable exceptions are List and Gallet (1999), Millimet et al. (2003), and Aldy (2005) that estimate EKC relationships using state level emissions data, albeit only for the United States. The first two studies utilize data from 1929-1994, and the third study utilizes data from 1960-1999.
The main econometric disadvantage of cross sectional data is that they do not permit us to use location specific dummies to capture the effect of omitted variables that are constant over time. Including variables that capture the major differences between locations mitigates the impact of omitted variables. In addition, analyzing econometric models with different specifications indicates whether omitted variables have a qualitatively significant effect, because different specifications measure the correlation between included and omitted variables in different ways. If different analyses lead to identical conclusions, then the omitted variables are unlikely to lead to an incorrect interpretation of the results.

\(c\) Pollution in general versus pollution that people consider harmful

In order to capture the pollution-income relationship rooted in consumer preferences, it is necessary to account for people’s awareness of the harmful effects of pollution. Most people are unaware of the dose-response relationships that outline the potential consequences of exposure to pollution. For this reason, the EPA has established the NAAQS above which ambient concentrations are considered harmful to human health. While the NAAQS do not necessarily represent an unambiguous scientific threshold above which pollution is dangerous (or below which pollution is harmless) most people assess their local air quality through exceedences of the NAAQS.\(^7\) Therefore, we expect to find a preference based pollution-income relationship primarily with respect to pollution levels that exceed these thresholds. Observations below the threshold levels might generate sufficient noise to reduce the precision of the estimated relationship. Simply eliminating the observations for which the ambient concentrations do not

\(^7\) For example, in many areas it is common to hear the local O\(_3\) rating based on the NAAQS and the EPA’s Air Quality Index along with the daily weather forecast
exceed the NAAQS does not solve the problem, because the fact that the ambient concentration
remained at or below the NAAQS is important information in itself. To reduce the impact of
potentially noisy observations without eliminating information, it is important to left-censor the
data and to record concentrations at relatively harmless levels as the lower bound of the data.
We are not aware of any existing study that focuses pollution that people consider harmful.

III. Our Data

We analyze the relationship between household income and pollution using 1990 data for
the United States, at the smallest geographic unit possible (census tracts).8 The United States is
at the upper end of the world’s income scale, and our data have a good chance of containing the
threshold income levels beyond which pollution begins to decline. We measure pollution by the
number of days in 1990 during which the concentrations of CO, O3, and PM10 at 704 locations
exceeded their respective NAAQS.9 This measure of pollution has five advantages over the

8 Other studies that use cross-section data for the United States include Berrens et al. (1997) and Gawande et al.
who analyzes state level air pollution data, Khanna (2002) and Khanna and Plassmann (2004) who analyze ambient
air pollutant concentrations data for census tracts, and Kahn (1998) who analyzes California vehicular emissions
data.

9 We obtained data on the annual counts in 1990 from the EPA’s USEPA-AQS database. This database also
provides the geographic coordinates (latitude and longitude) for each monitor, which we used to identify the census
tracts in which the monitors are located. In section VI.4 we also use the ambient concentrations data for our
pollutants to test the robustness of our results.
measures typically used in other studies. First, using data based on ambient concentrations rather than emissions ensures that we measure the pollution that people are exposed to rather than pollution that is generated locally but might mostly affect other regions. Second, all pollution measurements were taken by a single agency, which reduces the likelihood of measurement errors compared to data that were collected by different agencies in different countries, and possibly with different techniques. Third, by using local cross sectional data, we eliminate the possibility of aggregation biases due to spatial and intertemporal changes in preferences. Fourth, exceedences of the NAAQS explicitly account for short-term exposure at high concentrations that may be just as harmful as sustained exposure at low concentrations. Last but not least, using data on the number of days during which the level of pollution was considered harmful to human health ensures that we concentrate on those data that are most likely to indicate a preference-based relationship between household income and pollution.

The disadvantage of our cross-sectional data is that they rule out a location specific fixed effects model and we need to measure all relevant differences between census tracts. However, we use regional dummies to capture major differences between the 10 EPA regions, and we include all covariates that other studies have identified as relevant (see below).

---

10 We distinguish between ambient concentrations and pollution, and we consider as pollution only those concentrations that are harmful to human health. It is therefore possible that a census tract has higher ambient concentrations of a particular pollutant than another census tract, but we consider it no more polluted if the concentrations remain below the NAAQS.

11 It is possible that the spatial variation in preferences is so substantial that each census tract has a unique EKC. But it is impossible to identify such local EKCs with the currently available data. We therefore assume that preferences for environmental quality are uniformly distributed within in each census tract.
Empirical analyses of the EKC hypothesis seek to determine the relationship between economic growth and environmental degradation. In addition to leading to higher per capita GDP, economic growth has indirect effects like changes in the composition of the work force, improvements in medical services, changes in rural-urban settlement patterns, etc. To be able to capture all these effects, most studies use only per capita GDP as a proxy for economic growth, and other covariates that are likely to be uncorrelated with it.

In contrast to other studies, we analyze the relationship between household income and changes in the demand for environmental quality. To do this, we isolate the impact of income on local pollution by including variables that are likely to be correlated with income. The literature on the distribution of air pollution in the United States suggests that pollution in any given area is influenced by population density, racial composition, housing tenure, education, unemployment rate, and a population’s propensity for collective action.\textsuperscript{12} We include three additional variables:

\textsuperscript{12} See Brooks and Sethi (1997) and the references cited there. We measure racial composition by the percentage of minorities, the level of education by the percentage of high school graduates, the unemployment rate by the percentage of the labor force that is unemployed, and housing tenure by the percentage of houses that are renter occupied. We obtained census tract level data on these variables from the 1990 Census.

We measure the propensity for collective action at the county level by the fraction of the voting age population that was registered to vote in the 1992 Presidential elections. We obtained county level election data from Election Data Services, Inc, which does not report voter turnout data for many jurisdictions. We therefore could not use the ratio of voter turnout to voter registration to capture collective action. North Dakota does not require voter registration, and we used the ratio of voter turnout to voting age population. The data set does not include information on Wisconsin and Alaska, and we predicted the log of the voter registration rate in these two states with an auxiliary regression, using county level data for the entire United States.
the percentage of population below the poverty threshold, the percentage of female-headed households, and the percentage of population older than 65 years.

To account for differences in economic structure, we include the percentage of working age population in each census tract that is employed in manufacturing.\textsuperscript{13} We use the distance of the EPA monitors from the closest highway to measure the level of economic activity in the CO and O\textsubscript{3} analyses. This distance serves as a proxy for local economic activity because on-road vehicles are a primary source of emissions for these gases. The primary sources of PM\textsubscript{10} are coal-burning facilities such as electric utilities and copper smelters, and for each monitor we include the number of electric utilities in that EPA region that are monitored under the EPA’s Acid Rain Program. To incorporate regional differences, we include dummy variables for the 10 EPA regions. Table I shows summary statistics of all non-dummy variables.

**IV. A MODEL OF CORRELATED COUNT DATA**

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

\textsuperscript{13} Manufacturing is an aggregate category and pollution intensity may vary across the different industries within the manufacturing sector. Unfortunately, detailed data on employment in different manufacturing industries are not available at the census tract level.
\[ c_{pi} \sim \text{Poisson} \left( \mu_{pi} \right) \] (1)

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

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

The asymmetry of the gamma distribution implies that an increase in \( c_{pi} \) by a factor of \( \alpha \) is less likely than a decrease by the same factor. But in the absence of information that would warrant such a model, it is more appropriate to assume that an increase and a decrease by the same factor are equally likely. This can be achieved by assuming that \( \mu_{pi} = \exp(x_{pi}' \beta_p) \exp(\varepsilon_{pi}) \), where the \( \varepsilon_{pi} \) follow a univariate normal distribution. This distribution has the log-symmetry that the gamma distribution lacks, and yields the Poisson-lognormal regression model.\(^{14}\) The \( \varepsilon_{pi} \) can be interpreted as pollutant-and-location specific latent, or random, effects that measure the

\(^{14}\) Both, the negative binomial and the Poisson-lognormal regression model, accommodate overdispersed data, while the Poisson model describes data whose variance equals the mean. As long as the regression equation is specified correctly, maximum likelihood estimators of Poisson regression coefficients are consistent even if the data are overdispersed (Gourieroux et al., 1984).
impact of variables omitted from the covariate vectors. Because we have only a single observation for each location and cannot include location-specific fixed effects, these random effects are likely to improve the analysis.¹⁵

None of the models described above incorporate correlation between the pollutants because they assume independent $\mu_{pi}$’s. However, motor vehicles are a major source of emissions of CO and volatile organic compounds, VOCs (precursors of O₃), and the CO and O₃ counts are positively correlated. Electric utilities and copper smelters are a major source of PM₁₀, and tend to be located in areas with relatively low population and highway densities. Counts of PM₁₀ and the other two pollutants are therefore negatively correlated.

Such correlation can be accommodated by allowing the $\epsilon_{pi}$ to be correlated across pollutants. For example, if we assume that $\mu_{pi} = \exp(x_{pi}' \beta_p) \epsilon_i$, where the $\epsilon_i$ follow a univariate gamma distribution, then integrating the density function of $c_i = (c_{1i}, \ldots, c_{Pi})$ over the common variable $\epsilon_i$ yields the $P$-variate Poisson-gamma regression model. The covariance between the concentrations of pollutants $p$ and $q$ at location $i$ is given by $\text{Cov}(c_{pi}, c_{qi}) = \exp(x_{pi}' \beta_p) \exp(x_{qi}' \beta_q) \sigma$, where $\sigma$ is the variance of the gamma distribution. Because all three terms are positive, the multivariate Poisson-gamma distribution accommodates only non-negative correlation. An attractive alternative is the multivariate Poisson-lognormal regression model. This model can be obtained by assuming $\mu_{pi} = \exp(x_{pi}' \beta_p) \exp(\epsilon_{pi})$, where the $\epsilon_{pi}$ follow a $P$-variate normal distribution with mean zero and covariance matrix $\Sigma$. Because $\text{Cov}(c_{pi}, c_{qi}) =

¹⁵ The random effects in the Poisson-lognormal model require neither the assumption that the $\epsilon_{pi}$ are realizations from the same distribution nor that they are uncorrelated with the other covariates, the two main objections that are frequently raised against the least squares random-effects model. The Poisson-lognormal model easily accommodates observation-specific distributions as well as correlations between the $\epsilon_{pi}$ and the other covariates.
\[ \exp(x_{pl}'\beta_p)\exp(x_{qj}'\beta_q)(\exp(\sigma_{pq})-1), \]

where \(\sigma_{pq}\) is the element in row \(p\) and column \(q\) of \(\Sigma\), the model accommodates positive and negative correlation between the elements of \(c_i\).\(^{16}\)

The integral of the density function of \(C_i\) over \(\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{ip})\) does not have a closed form solution, which makes maximum likelihood analysis cumbersome. It is straightforward, however, to estimate the unknown parameters with simulation-based methods such as the Gibbs sampler, a Markov chain Monte Carlo (MCMC) method.\(^{17}\) We closely follow the setup that Chib and Winkelmann (2001) suggest for this type of analysis, and we describe the Gibbs sampler for our multivariate Poisson-lognormal model in the appendix.

V. RESULTS

Table II shows estimates of the covariance matrix \(\Sigma\) and the corresponding correlation matrix of the latent effects of a joint Poisson-lognormal analysis of all three pollutants. The latent effects from the three equations show statistically significantly negative correlation.\(^{18}\) This suggests that the multivariate Poisson-lognormal distribution is more appropriate for these data

\(^{16}\) Aitchison and Ho (1989) show the derivation of the moments of the multivariate Poisson-lognormal distribution. Winkelmann (2000, pp. 182-184) describes the multivariate Poisson-lognormal regression model.

\(^{17}\) MCMC methods are iterative techniques that use Markov chains to perform Monte Carlo integrations of integrals of interest. For details on MCMC methods, see Casella and George (1992), Gilks et al. (1996), Chib and Greenberg (1996), Gamerman (1997), Carlin and Louis (2000), and Chen et al. (2000).

\(^{18}\) The covariance matrix of the latent effects measures the covariances of the concentrations as well as the covariances of variables that were omitted from the three regression equations. The estimates in Table II do not necessarily imply that all pollutants are negatively correlated, because the estimates might be driven by negative correlations of the effects of omitted variables.
than either the multivariate Poisson or the multivariate negative binomial distributions that accommodate only positive correlation, and we report the corresponding coefficient estimates in column 5 of Tables IIIa–c. This is our benchmark specification that assumes a quadratic relationship between income and pollution and includes all covariates. In this section, we discuss the estimated pollution-income relationship for each pollutant obtained using this specification and defer the discussion of the estimated coefficients on the other covariates as well as the other specifications reported in Tables IIIa-c until the next section where we assess the robustness of our results.

Before we discuss our results we need to address an important issue regarding the location and precision of the estimated turning points. A common method for assessing the precision of the turning point estimate is to use a normal approximation of the turning point estimator, the “delta method” (see Greene, 2002, pp. 913-914). However, the distributions of turning point estimators of polynomial regression functions are often asymmetric, making the normal approximation inappropriate, and it is not possible to derive (asymptotic) 95 percent confidence intervals of the estimates by adding and subtracting 1.96 times the estimated standard error (see Plassmann and Khanna, 2002). At the bottom of Tables IIIa – c, we show estimates of the quantiles of the distributions of the turning point estimators. The quantiles corresponding to the 2.5th and the 97.5th percentile indicate the 95 percent confidence intervals, and the modes of the distributions (the 50th percentile) yield estimates of the turning points.19

Put simply, we regard a turning point that falls well within the sample income range as an indicator of a non-monotonic relationship between income and pollution. A turning point

---

19 The ratio of (the coefficient of income) and (minus 2 times the coefficient of income squared) is an imprecise estimator of the turning point because the distribution of the turning point estimator has no mean.
sufficiently far to the left or right of the median suggests a non-linear but not necessarily non-monotonic relationship between pollution and income. The range of log-household income extends from 8.52 ($4,999) to 11.48 ($96,383) with mean 10.30 ($29,652) and median 10.25 ($28,397). 80 percent of the census tracts in our data set have a log of median household income between 9.60 ($14,688) and 10.74 ($46,051).

*Carbon monoxide*

Our benchmark quadratic specification (Table IIIa, column 5) yields a turning point at the extreme lower end of the income data range, and the apparent non-monotonicity is likely to be an artifact of the polynomial functional form. The estimated quantiles of the distribution of the turning point estimator show that the turning point estimate is fairly imprecise, which also indicates that it is not an integral part of the estimated relationship. Overall we conclude that there is no evidence that either the growth rate of CO concentrations or the concentrations themselves decrease with income.

Motor vehicles are the main source of CO emissions, and the income elasticity of transportation (vehicle miles traveled) ranges from 0.5 to 1 (see Agras and Chapman, 1999). While richer households consume more vehicle miles, they also tend to drive newer and more fuel-efficient vehicles. Our result implies that for higher income households the increase in CO pollution due to greater vehicle use tends to be offset by the fuel efficiency gains of newer cars.

*Ozone*

Our benchmark specification for O₃ is reported in column 5 of Table IIIb. The estimated turning point lies towards the right end of the income distribution, and while the turning point
estimate is more precise than that obtained for CO, the confidence interval covers the entire range of the income data. We conclude that there is evidence of a concave relationship between income and O₃ exceedences, but we do not have sufficient data from census tracts with higher income levels to determine whether O₃ pollution will ultimately decrease with income.

O₃ is formed by a chemical reaction of VOCs and nitrogen oxides (NOₓ). The main sources of VOCs are chemical plants, refineries, and motor vehicles, and the main source of NOₓ is fuel combustion (power plants, heating, and motor vehicles). While more transportation tends to be associated with higher levels of these precursors of O₃, we expect that higher income households will reduce exposure to O₃ by distancing themselves from the non-mobile sources of VOCs and NOₓ. The relationship between O₃ and its precursors is highly non-linear, and our results suggest that even greater reduction in these precursors is required for the O₃-income relationship to turn downwards.

**Particulate Matter**

PM₁₀ is the only pollutant for which our analysis suggests a non-monotonic relationship with household income. The quadratic specification (Table IIIc, column 5) yields a turning point of 9.74, which is reasonably close to the median of our income data (10.25 = ln($28,397)). The quantiles at the bottom of Table IIIc imply a relatively small confidence interval around the turning point estimate, suggesting that this turning point is more than just a functional artifact.

Most of the anthropogenic emissions of PM₁₀ are the result of non-transportation fuel combustion and industrial processing. Because it is possible to reduce exposure to PM₁₀ pollution by moving away from polluting industries, it is reasonable that our results indicate an inverted U-shaped relationship for this pollutant. As the trade-off between finding a job at or
near a polluting industry and being exposed to higher levels of pollution becomes less and less attractive with increases in income, higher income households simply relocate.

VI. TESTS OF ROBUSTNESS

As mentioned earlier, the literature reports evidence to suggest that estimated EKCs are not very robust. To test the robustness of our estimates, we assess the distributional assumption (normal, Poisson, negative binomial, Poisson-lognormal) in columns 1 – 4, the dependence assumption (individual versus joint analysis of the three pollutants) in columns 4 and 5, and the heterogeneity assumption (different functional forms and different sets of covariates) in columns 5 – 8 of Tables IIIa-c.20, 21

1. Assessing distributional and dependence assumptions

Table IV shows the empirical frequency distributions of the three pollutants. For CO and PM\textsubscript{10}, the vast majority of the observations are zeros, while the distribution of O\textsubscript{3} is much more dispersed. The negative binomial and the Poisson-lognormal distributions can describe any of the three distributions, while the Poisson distribution is likely to be adequate only for PM\textsubscript{10}, whose variance is closest to its mean. The normal distribution is ill suited for the analysis of any

---

20 In the first three columns, we used $p$-values to determine whether the coefficients are statistically significantly different from zero. The posterior distributions of most coefficients in Columns 4 – 8 are asymmetric and the standard error is not a reliable guide for the assessment of statistical significance. We therefore established statistical significance from the posterior distributions of the parameters.

21 We applied the specifications in columns 6 to 8 to all four count models, but show only the results of the joint Poisson-lognormal model, which we consider to be most appropriate. The results of the three other models are similar, given their respective assumptions about distribution and dependence.
of the three pollutants, but we expect it to do least badly for the analysis of O₃, whose distribution is closest to the right leg of the normal distribution.

Comparing columns 1 – 4 of Tables IIIa–c confirms this intuition. For many coefficients in the analyses of CO and PM₁₀, the estimates of the normal model differ from the coefficient estimates of the three count models by more than two standard errors. In the analysis of O₃, all 95 percent confidence intervals of the normal model include the point estimates of the count models, while most 95 percent confidence intervals of the count models do not include the normal point estimates. We conclude that it is inappropriate to analyze our data with a normal model. For the overdispersed pollutants CO and O₃, the estimates of the negative binomial and the Poisson-lognormal model are more similar to each other than to the Poisson estimates, which is consistent with our intuition that the Poisson-model is inappropriate for our data. The three count models yield fairly similar coefficient estimates for PM₁₀, whose distribution shows the least degree of overdispersion.

We argue that the multivariate Poisson-lognormal model is the most appropriate for our data based on the estimate of the covariance matrix in Table II. We report the results from the individual Poisson-lognormal models for each of the three pollutants in column 4 of Tables IIIa–c: comparing the estimates in columns 4 and 5 shows the influence of this dependence assumption. Overall, the results in columns 4 and 5 are qualitatively similar.

2. Different functional forms with respect to income

Next we compare three different model specifications -- linear, quadratic, and cubic -- with respect to the income covariate. We report the results for these specifications in columns 5-
7 of Tables IIIa-c, while Figures 1a-c contain graphical depictions of the estimated pollution-income relationships for each pollutant, respectively.\textsuperscript{22}

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

In the case of O\textsubscript{3}, the linear analysis suggests a statistically significantly positive relationship, and the quadratic and cubic models indicate that the growth rate of O\textsubscript{3} concentrations falls with income. The estimated turning points lie at the right end of the income distribution and the confidence intervals cover the whole range of the income data. All three analyses lead us to conclude a monotonic, albeit concave, relationship between income and O\textsubscript{3}.

The graph of the cubic analysis for PM\textsubscript{10} shows that pollution increases with income as long as income is relatively low, and that it falls more sharply beyond the turning point. The quadratic specification, which imposes symmetry around the turning point, yields a turning point estimate at a lower income level than the cubic specification, which also suggests an asymmetry in slopes before and after the turning point. Based on the qualitative similarity of the results, we

\textsuperscript{22} Because we are interested in a comparison of shapes, we shifted the quadratic and cubic relationships vertically so that their average impact over the log-income range \([8.52, 11.48]\) is the same as the linear average, and ignored the presence of non-income covariates, regional fixed effects, and observations-specific random effects in the equations. We therefore do not show units on the vertical axes. An alternative is to evaluate the equations at the medians and means of the non-income covariates. Because this practice ignores the impact of the observation-specific random effects, it overstates the differences between the curves. We think that, in the presence of such random effects, our method is more appropriate to assess the differences between functional forms.
conclude that there is strong evidence of a non-monotonic concave (inverted U-shaped) relationship between household income and PM$_{10}$ pollution.

3. *The other covariates*

We obtain consistent estimates of the coefficients of the non-income covariates across analyses. All analyses in Table IIIa suggest that CO pollution increases with population density and decreases with the distance of the EPA monitor to the nearest road. These are intuitive results because motor vehicles are the main source of CO emissions and densely populated areas tend to have high vehicle density as well. None of the other covariates is statistically significant.

The percentage of registered voters has a strong negative impact on O$_3$ and PM$_{10}$ pollution. This is consistent with the results of Brooks and Sethi (1997), who found that more politically active communities tend to be exposed to lower pollution. Although one might expect a statistically significant negative coefficient in the case of CO, ambient CO concentrations are highly correlated with local emissions (EPA, 1997), so that reducing local CO pollution entails reducing emissions locally from its primary source: gasoline powered vehicles. This has a high private cost since it necessitates lifestyle changes, such as a shift from private to public modes of transportation, or the adoption of alternative technologies (for example, hybrid vehicles), which are fairly expensive. Political action is unlikely to be successful in reducing pollution in this situation. PM$_{10}$ and O$_3$, on the other hand, have several point and non-point sources and it is possible to lower exposure to these pollutants through political action that keeps point sources out of a neighborhood.

---

$^{23}$ Brooks and Sethi (1997) analyze the relationship between community characteristics and exposure to aggregate emissions of over 300 chemicals reported in the EPA’s Toxics Release Inventory.
All Poisson-lognormal analyses indicate that O₃ concentrations increase with the distance of the monitor from the closest highway. This suggests that most of the measured O₃ pollution is due to sources other than motor vehicles that tend to be located in areas with low population and highway density. O₃ pollution also increases with the percentage of female-headed households. PM₁₀ pollution increases with the percentage of population below the poverty line and, possibly, with the unemployment rate. All three measures are negatively correlated with income, and our estimates indicate that low-income households are willing to accept proximity to a polluting industry as the price of finding employment.

Omitting all non-income covariates except ‘distance of the monitor to the nearest highway’ (which is least likely to be correlated with income) affects the results only in the case of CO where it yields a decreasing rather than increasing relationship over the whole income range (column 8 of Table IIIa). This change is not surprising because most of the omitted covariates are highly correlated with income. For example, our analysis in column 5 shows that household income and population density are positively correlated with CO pollution, and that the effect of density exceeds the effect of income. Because we use data on the census tract level, median household income and population density are negatively correlated ($\rho = -0.15$). If we omit population density, then the income coefficients measure the joint effect of both variables, and suggest a negative relationship between household income and CO pollution.

4. Comparing count data with ambient concentrations data

Finally, we attempt to repeat our analysis using the “precursor” of our data set with the ambient concentrations of the three pollutants to determine the effect of left-censoring the data. Unfortunately, the EPA was unable to make these data available to us, but they supplied us with
data on the 1990 concentrations of the three pollutants measured at different sites. Because many sites in this concentrations data set differ from the sites in the count data set, we could use the new data to test the robustness of our results. The analysis of the additional data yields very similar relationships between household income and the three pollutants, and we conclude that our results are robust with respect to changes in the data set. We report the results of the additional analysis in a supplementary paper that is available at http://bingweb.binghamton.edu/~fplass/papers/IncomeEKC_Supplement.pdf.

VII. CONCLUSION

The EKC hypothesis suggests a non-monotonic relationship between economic growth and pollution. To the extent that this relationship is driven by changes in the demand for environmental quality, it is unlikely that there is a single global EKC. Country-level analyses of global EKC relationships that use multi-country panel data sets therefore suffer from several types of aggregation bias. We suspect that these biases are the main reason why previous studies have yielded conflicting results. Analyzing cross sectional data for a single country avoids these biases, and yields estimates of local relationship between household income and pollution. For three pollutants in the United States in 1990, we find that such relationships are not sensitive to either changes in model specification or differences in data.

Our results suggest that the income level at which households are willing to reduce their exposure to pollution depends on the nature of pollution. We find an inverted U-shaped relationship between household income and PM$_{10}$ with a peak at about $20,000. PM$_{10}$ is a point-source pollutant, and it is fairly straightforward and relatively inexpensive to reduce exposure to PM$_{10}$ by relocating, without necessarily reducing global emissions. We do not find unambiguous
evidence of an inverted U-shaped relationship between household income and the two non point-source pollutants, \( O_3 \) and \( CO \), that are generated primarily by the use of gasoline powered transportation. The private abatement costs for \( O_3 \) and \( CO \) pollution are relatively high. Even in a country with one of the highest incomes in the world, household income has not yet reached the level beyond which the pollution-income relationship for such pollutants becomes negative. Despite the Clean Air Act Amendments of 1990, \( O_3 \) pollution in the United States has remained fairly constant over the last decade. \( CO \) pollution has declined somewhat (EPA, 2002). A likely reason for this decline is a change in consumer preferences due to increased awareness of the harmful effects of \( CO \), coupled with the relatively easy switch to oxygenated gasoline.

Although economic growth and increases in household income are usually positively correlated, analyzing the relationship between household income and pollution has somewhat different implications than analyzing the relationship between economic growth and pollution. An inverted U-shaped relationship between economic growth and pollution suggests that promoting economic growth is an attractive possibility for improving local environmental quality. An inverted U-shaped relationship between household income and pollution shows that households accept higher pollution as the price of higher income only up to a certain level of income, and that richer households stay away from polluted areas. A negative relationship between household income and pollution does not imply that increasing local income will automatically be associated with an improvement in local environmental quality, because it may be cheaper to relocate to a less polluted area than to reduce local pollution. Alternatively, households may lobby to relocate polluting industries to other areas, and local environmental quality improves at the expense of environmental quality elsewhere.
Table I. Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median household income</td>
<td>29,652</td>
<td>13,057</td>
<td>4,999</td>
<td>96,383</td>
</tr>
<tr>
<td>Population density (persons / square mile)</td>
<td>2,742</td>
<td>5,360</td>
<td>0.57</td>
<td>102,938</td>
</tr>
<tr>
<td>Percent minority</td>
<td>19.90</td>
<td>22.94</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Percent of labor force unemployed^a</td>
<td>7.73</td>
<td>6.24</td>
<td>0.00</td>
<td>60.00</td>
</tr>
<tr>
<td>Percent of labor force employed in manufacturing^a</td>
<td>17.30</td>
<td>9.50</td>
<td>0.00</td>
<td>83.33</td>
</tr>
<tr>
<td>Percent high school graduates</td>
<td>71.91</td>
<td>15.84</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Percent voting age population registered to vote^b</td>
<td>72.07</td>
<td>10.48</td>
<td>44.72</td>
<td>113.26</td>
</tr>
<tr>
<td>Percent of houses renter occupied</td>
<td>39.92</td>
<td>24.14</td>
<td>1.66</td>
<td>100.00</td>
</tr>
<tr>
<td>Percent below poverty threshold</td>
<td>15.75</td>
<td>14.05</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Percent female head of household</td>
<td>11.57</td>
<td>7.46</td>
<td>0.00</td>
<td>52.72</td>
</tr>
<tr>
<td>Percent population above 65 years of age</td>
<td>12.15</td>
<td>8.23</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Distance from the nearest highway in meters</td>
<td>1,138</td>
<td>2,051</td>
<td>0.59</td>
<td>27,773</td>
</tr>
<tr>
<td>Number of electric utilities in EPA region</td>
<td>30.48</td>
<td>22.13</td>
<td>0</td>
<td>106</td>
</tr>
</tbody>
</table>

^a We measure labor force as the population that is 16 years or older.

^b Refers to the 1992 presidential election. Values shown do not include the predicted data for Wisconsin and Alaska. The percentage exceeds 100 at several census tracts because of voter fraud (personal communication with a representative of Election Data Services).
Table II. Covariance and correlation matrices of the latent variables $\varepsilon_p$

in the joint Poisson-lognormal analysis in Table III, Column 5

<table>
<thead>
<tr>
<th></th>
<th>CO</th>
<th>O$_3$</th>
<th>PM$_{10}$</th>
<th>CO</th>
<th>O$_3$</th>
<th>PM$_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariance matrix $\Sigma$</td>
<td>8.8047***</td>
<td>1.6109***</td>
<td>3.5391***</td>
<td>CO</td>
<td>O$_3$</td>
<td>PM$_{10}$</td>
</tr>
<tr>
<td>CO</td>
<td>(2.2235)</td>
<td>(0.3649)</td>
<td>(0.8725)</td>
<td>1</td>
<td>-0.3977***</td>
<td>-0.6408***</td>
</tr>
<tr>
<td>O$_3$</td>
<td>-1.4831***</td>
<td>(0.1219)</td>
<td>(0.2482)</td>
<td>-0.2670*</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PM$_{10}$</td>
<td>-1.4672*</td>
<td>-1.5087***</td>
<td>(1.0392)</td>
<td>-0.2670*</td>
<td>-0.6408***</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes: Standard errors are shown in parentheses.
A ‘***’ indicates that the coefficient is different from zero at the 99% level, and a ‘*’ indicates that the coefficient is different from zero at the 90% level.
## Table IIIa. Analysis of Carbon Monoxide (CO)

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Normal</th>
<th>Poisson</th>
<th>Negative Binomial</th>
<th>Poisson-lognormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis</td>
<td>Individual (1)</td>
<td>Individual (2)</td>
<td>Individual (3)</td>
<td>Individual (4)</td>
</tr>
<tr>
<td>ln(income)</td>
<td>-6.5642**</td>
<td>-3.1012***</td>
<td>-24.1871***</td>
<td>-2.3580**</td>
</tr>
<tr>
<td>ln(income)$^2$</td>
<td>0.3103</td>
<td>0.1835</td>
<td>1.2050**</td>
<td>0.1310</td>
</tr>
<tr>
<td>ln(income)$^3$</td>
<td>0.4611</td>
<td>0.3709</td>
<td>0.5441</td>
<td>0.3590</td>
</tr>
<tr>
<td>ln(population density)</td>
<td>0.1263**</td>
<td>0.8565***</td>
<td>1.0838***</td>
<td>1.3850***</td>
</tr>
<tr>
<td>ln(percent minority)</td>
<td>0.0690</td>
<td>0.2475</td>
<td>0.1497</td>
<td>0.3924</td>
</tr>
<tr>
<td>ln(percent unemployment)</td>
<td>0.0902</td>
<td>0.1624</td>
<td>0.3289</td>
<td>0.3035</td>
</tr>
<tr>
<td>ln(percent in manufact.)</td>
<td>0.0568</td>
<td>0.2135</td>
<td>0.2286</td>
<td>0.4514</td>
</tr>
<tr>
<td>ln(percent HS grad)</td>
<td>0.0111</td>
<td>0.2608</td>
<td>0.8691**</td>
<td>0.1420</td>
</tr>
<tr>
<td>ln(percent voters)</td>
<td>0.1909</td>
<td>-0.0690</td>
<td>-0.3546</td>
<td>-0.2065</td>
</tr>
<tr>
<td>ln(percent renters)</td>
<td>0.1514</td>
<td>0.5470</td>
<td>0.5034</td>
<td>0.8718</td>
</tr>
<tr>
<td>ln(percent &lt; poverty)</td>
<td>-0.0929</td>
<td>-0.0181</td>
<td>0.2224</td>
<td>0.1933</td>
</tr>
<tr>
<td>ln(percent female head of household)</td>
<td>0.0518</td>
<td>0.4332</td>
<td>-0.3460</td>
<td>-0.4740</td>
</tr>
<tr>
<td>ln(percent above 65)</td>
<td>0.2376</td>
<td>0.5349</td>
<td>0.4002</td>
<td>0.6052</td>
</tr>
<tr>
<td>ln(Pseudo $R^2$)</td>
<td>0.0837</td>
<td>0.4247</td>
<td>0.1899</td>
<td>0.0060</td>
</tr>
</tbody>
</table>

Turning Point (TP)

<table>
<thead>
<tr>
<th>Quantiles:</th>
<th>2.5%</th>
<th>5.0%</th>
<th>10.0%</th>
<th>50.0%</th>
<th>90.0%</th>
<th>95.0%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cov ($\beta_1$, $\beta_2$)</td>
<td>-4.3824</td>
<td>-2.6901</td>
<td>-5.9525</td>
<td>-2.3494</td>
<td>-2.2037</td>
<td>-1.5643</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Standard errors are shown in parentheses. Estimates of intercepts and regional dummies are not shown.

A "***" indicates that the coefficient is different from zero at the 99% level, a "**" indicates that the coefficient is different from zero at the 95% level, and a "*" indicates that the coefficient is different from zero at the 90% level.

(a) The turning point estimates are the modes of the empirical distributions of the turning point estimators.
(b) For the cubic relationship in Column 7, we were unable to obtain estimates of the distributions of the turning points, because the majority of the runs of the Gibbs sampler implied complex roots for the cubic equation.

Household Income and Pollution —Tables and Figures 29
### Table IIIb. Analysis of Ozone (O₃)

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Analysis</th>
<th>Normal</th>
<th>Poisson</th>
<th>Negative Binomial</th>
<th>Individual</th>
<th>Joint</th>
<th>Joint</th>
<th>Joint</th>
<th>Joint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>ln(income) **</td>
<td></td>
<td>0.3782</td>
<td>11.3383***</td>
<td>3.8431</td>
<td>4.1218</td>
<td>4.8198</td>
<td>0.5132**</td>
<td>-3.8374</td>
<td>5.8727***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(29.6153)</td>
<td>(3.4942)</td>
<td>(3.2016)</td>
<td>(3.2496)</td>
<td>(3.1234)</td>
<td>(0.2774)</td>
<td>(9.2900)</td>
<td>(2.9074)</td>
</tr>
<tr>
<td>ln(income)³</td>
<td></td>
<td>0.2096</td>
<td>-0.5284***</td>
<td>-0.1724</td>
<td>-0.1805</td>
<td>-0.2137*</td>
<td></td>
<td>0.7258</td>
<td>-0.2590**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.4919)</td>
<td>(0.2096)</td>
<td>(0.1563)</td>
<td>(0.1611)</td>
<td>(0.1550)</td>
<td></td>
<td>(0.9797)</td>
<td>(0.1438)</td>
</tr>
<tr>
<td>ln(income)³²</td>
<td></td>
<td>0.1911</td>
<td>0.0167</td>
<td>0.0040</td>
<td>0.0030</td>
<td>0.0020</td>
<td>-0.0084</td>
<td></td>
<td>0.0028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.4041)</td>
<td>(0.0354)</td>
<td>(0.0302)</td>
<td>(0.0359)</td>
<td>(0.0356)</td>
<td>(0.0353)</td>
<td></td>
<td>(0.0358)</td>
</tr>
<tr>
<td>ln(population density)</td>
<td></td>
<td>0.2029</td>
<td>-0.0081</td>
<td>-0.0113</td>
<td>0.0158</td>
<td>0.0159</td>
<td>0.0082</td>
<td></td>
<td>0.0222</td>
</tr>
<tr>
<td>ln(percentage minority)</td>
<td></td>
<td>0.5738</td>
<td>(0.0752)</td>
<td>(0.0533)</td>
<td>(0.0684)</td>
<td>(0.0677)</td>
<td>(0.0681)</td>
<td></td>
<td>(0.0698)</td>
</tr>
<tr>
<td>ln(percentage unemployement)</td>
<td></td>
<td>0.4995</td>
<td>0.0768</td>
<td>0.1187</td>
<td>-0.0463</td>
<td>-0.0396</td>
<td>-0.0413</td>
<td></td>
<td>-0.0601</td>
</tr>
<tr>
<td>ln(percentage empl. in manufact.)</td>
<td></td>
<td>1.1558</td>
<td>(0.1489)</td>
<td>(0.1156)</td>
<td>(0.1374)</td>
<td>(0.1371)</td>
<td>(0.1350)</td>
<td></td>
<td>(0.1376)</td>
</tr>
<tr>
<td>ln(percentage HS grad.)</td>
<td></td>
<td>-0.6672</td>
<td>-0.0908</td>
<td>-0.0640</td>
<td>-0.0155</td>
<td>-0.0186</td>
<td>0.0159</td>
<td></td>
<td>-0.0281</td>
</tr>
<tr>
<td>ln(percentage voters)</td>
<td></td>
<td>-19.3147***</td>
<td>-2.4140***</td>
<td>-1.4665***</td>
<td>-0.8274**</td>
<td>-0.9046***</td>
<td>-0.8801***</td>
<td></td>
<td>-0.8634**</td>
</tr>
<tr>
<td>ln(percentage renters)</td>
<td></td>
<td>-2.6765</td>
<td>-0.3664***</td>
<td>-0.1984*</td>
<td>-0.1619</td>
<td>-0.1650</td>
<td>-0.1229</td>
<td></td>
<td>-0.1815</td>
</tr>
<tr>
<td>ln(percentage poverty)</td>
<td></td>
<td>1.1817</td>
<td>0.1668</td>
<td>-0.0824</td>
<td>-0.0647</td>
<td>-0.0881</td>
<td>-0.1264</td>
<td></td>
<td>-0.1009</td>
</tr>
<tr>
<td>ln(percentage female head of household)</td>
<td></td>
<td>1.5468</td>
<td>0.3217**</td>
<td>0.2407*</td>
<td>0.3060***</td>
<td>0.3156***</td>
<td>0.3856***</td>
<td></td>
<td>0.3441***</td>
</tr>
<tr>
<td>ln(percentage above 65)</td>
<td></td>
<td>1.1074</td>
<td>(0.1585)</td>
<td>(0.1282)</td>
<td>(0.1411)</td>
<td>(0.1440)</td>
<td>(0.1407)</td>
<td></td>
<td>(0.1427)</td>
</tr>
<tr>
<td>ln(Distance from highway)</td>
<td></td>
<td>-0.0406</td>
<td>-2.0072</td>
<td>-2.2614**</td>
<td>-1.3860</td>
<td>-1.3400</td>
<td>-1.2804</td>
<td></td>
<td>-1.3162</td>
</tr>
<tr>
<td>ln(Pseudo R²)²</td>
<td></td>
<td>(9.0215)</td>
<td>(1.4292)</td>
<td>(0.8848)</td>
<td>(1.0822)</td>
<td>(1.0808)</td>
<td>(1.0764)</td>
<td></td>
<td>(1.0859)</td>
</tr>
<tr>
<td>Turning Point (TP)</td>
<td></td>
<td>0.2180</td>
<td>0.0179</td>
<td>0.0431</td>
<td>0.0737**</td>
<td>0.0771**</td>
<td>0.0822***</td>
<td>0.0761**</td>
<td>0.0670**</td>
</tr>
<tr>
<td>Cov (β₁, β₂ )</td>
<td></td>
<td>0.4581</td>
<td>(0.0488)</td>
<td>(0.0344)</td>
<td>(0.0414)</td>
<td>(0.0410)</td>
<td>(0.0409)</td>
<td>(0.0408)</td>
<td>(0.0377)</td>
</tr>
</tbody>
</table>

| Quantiles:          |                   | 0.1724 | 0.2819 | 0.0577           |       |       |       |       |       |
| Turning Point (TP)  |                   | 9.2778 | 10.7252 | 11.8617          | 11.0236   | 11.0416 | 11.8790 | 11.2811 |
| Cov (β₁, β₂ )       |                   | -44.0839 | -0.9155  | -0.4990          | -0.5214   | -0.4868 | -0.4177 |       |       |

Notes: Standard errors are shown in parentheses. Estimates of intercepts and regional dummies are not shown. A ‘***’ indicates that the coefficient is different from zero at the 99% level, a ‘**’ indicates that the coefficient is different from zero at the 95% level, and a ‘*’ indicates that the coefficient is different from zero at the 90% level.

(a) The turning point estimates are the modes of the empirical distributions of the turning point estimators.
Table IIIc. Analysis of Particulate Matter (PM$_{10}$)

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Normal</th>
<th>Poisson</th>
<th>Negative Binomial</th>
<th>Poisson-lognormal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis</td>
<td>Individual (1)</td>
<td>Individual (2)</td>
<td>Individual (3)</td>
<td>Joint (5)</td>
</tr>
<tr>
<td>ln(income)</td>
<td>2.9611* (1.7802)</td>
<td>22.8391*** (8.9942)</td>
<td>23.3054*** (8.8741)</td>
<td>8.8347** (4.3871)</td>
</tr>
<tr>
<td>ln(income)$^2$</td>
<td>-0.1498* (0.0868)</td>
<td>-1.1388*** (0.4496)</td>
<td>-1.1903*** (0.4907)</td>
<td>-0.0604*** (0.2239)</td>
</tr>
<tr>
<td>ln(income)$^3$</td>
<td>0.0260 (0.0164)</td>
<td>0.0487 (0.0829)</td>
<td>0.0654* (0.0802)</td>
<td>0.0368 (0.0518)</td>
</tr>
<tr>
<td>ln(population density)</td>
<td>0.0260 (0.0164)</td>
<td>0.0487 (0.0829)</td>
<td>0.0654* (0.0802)</td>
<td>0.0368 (0.0518)</td>
</tr>
<tr>
<td>ln(percent minority)</td>
<td>-0.0381 (0.0313)</td>
<td>-0.3321* (0.1944)</td>
<td>-0.1778 (0.1881)</td>
<td>-0.3486* (0.1292)</td>
</tr>
<tr>
<td>ln(percent unemployment)</td>
<td>0.1225** (0.0510)</td>
<td>0.5374* (0.2921)</td>
<td>0.5556* (0.3215)</td>
<td>0.5352** (0.2403)</td>
</tr>
<tr>
<td>ln(percent in manufact.)</td>
<td>-0.0704 (0.0518)</td>
<td>-0.2320 (0.2065)</td>
<td>-0.1362 (0.2037)</td>
<td>-0.1706 (0.1611)</td>
</tr>
<tr>
<td>ln(percent HS grad)</td>
<td>-0.3523** (0.1515)</td>
<td>-0.7750** (0.3515)</td>
<td>-0.5121 (0.3371)</td>
<td>-0.7034** (0.2493)</td>
</tr>
<tr>
<td>ln(percent voters)</td>
<td>-0.4688** (0.0197)</td>
<td>-3.2733*** (1.1637)</td>
<td>-3.5410*** (1.0369)</td>
<td>-3.6767*** (0.8789)</td>
</tr>
<tr>
<td>ln(percent percent poverty)</td>
<td>-0.0854 (0.0771)</td>
<td>-0.4349 (0.4834)</td>
<td>-0.5233 (0.4300)</td>
<td>-0.3158 (0.2597)</td>
</tr>
<tr>
<td>ln(percent female head of household)</td>
<td>0.0811 (0.0497)</td>
<td>0.8853* (0.4885)</td>
<td>0.7361* (0.4211)</td>
<td>0.8156*** (0.2493)</td>
</tr>
<tr>
<td>ln(percent above 65)</td>
<td>-0.5154 (0.4779)</td>
<td>0.0263 (2.4119)</td>
<td>2.7633 (2.5010)</td>
<td>-0.6743 (1.8267)</td>
</tr>
<tr>
<td>Number of elect. utilities</td>
<td>0.0259 (0.0599)</td>
<td>-0.1593* (0.0824)</td>
<td>-0.1337 (0.0934)</td>
<td>0.1218 (0.1796)</td>
</tr>
<tr>
<td>Pseudo $R^2$</td>
<td>0.1091 (0.1023)</td>
<td>0.2537 (0.9793)</td>
<td>0.1502 (10.0073)</td>
<td>0.7439 (9.7937)</td>
</tr>
</tbody>
</table>

Turning Point (TP) Cov ($\beta_1, \beta_2$)

<table>
<thead>
<tr>
<th>Quantiles:</th>
<th>2.5%</th>
<th>5.0%</th>
<th>10.0%</th>
<th>50.0%</th>
<th>90.0%</th>
<th>95.0%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-4.5745</td>
<td>2.5437</td>
<td>5.1199</td>
<td>8.1012</td>
<td>8.9479</td>
<td>9.1752</td>
<td>17.5628</td>
</tr>
</tbody>
</table>

Notes: Standard errors are shown in parentheses. Estimates of intercepts and regional dummies are not shown. A ‘***’ indicates that the coefficient is different from zero at the 99% level, a ‘**’ indicates that the coefficient is different from zero at the 95% level, and a ‘*’ indicates that the coefficient is different from zero at the 90% level.

(a) The turning point estimates are the modes of the empirical distributions of the turning point estimators.
Table IV. Empirical Distributions of CO, O₃, and PM₁₀

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>&gt;9</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>647</td>
<td>14</td>
<td>14</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>0.44</td>
<td>6.15</td>
</tr>
<tr>
<td>O₃</td>
<td>111</td>
<td>106</td>
<td>72</td>
<td>64</td>
<td>42</td>
<td>29</td>
<td>39</td>
<td>29</td>
<td>23</td>
<td>22</td>
<td>167</td>
<td>8.69</td>
<td>261.65</td>
</tr>
<tr>
<td>PM₁₀</td>
<td>632</td>
<td>41</td>
<td>13</td>
<td>7</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.20</td>
<td>0.60</td>
</tr>
</tbody>
</table>
Figure 1a Joint Poisson-lognormal analyses of Carbon Monoxide. The vertical lines show range of our income data.

Figure 1b Joint Poisson-lognormal analyses of Ozone. The vertical lines show the range of our income data.
Figure 1c Joint Poisson-lognormal analyses of Particulate Matter. The vertical lines show the range of our income data.
Appendix

In this appendix, we describe the setup of the Gibbs sampler, a Markov chain Monte Carlo (MCMC) method, that we use to analyze the multivariate Poisson-lognormal model. Implementation of the Gibbs sampler requires knowledge of the full-conditional distributions of all parameters of interest. Such full conditional distributions are derived from the joint distribution of the data and the model parameters of interest. In our model, the parameters of interest are the three parameter vectors that describe the impact of the covariates and the regional dummies on each pollutant, $\beta_p$, $p = 1, 2, 3$, as well as the covariance matrix $\Sigma$.

Let $N$ be the number of monitors at which the gases are measured, and let $c_{pi}$ be the number of days at which the concentration of pollutant $p$ at monitor $i$, $i = 1, \ldots, N$, exceeds its NAAQS. Denote the collection of the $c_{pi}$ over all locations by the vector $C_p = (c_{p1}, \ldots, c_{pN})$. Denote the covariates in the analysis of pollutant $p$ by the $N \times M_p$ matrix $X_p$, where $M_p$ is the number of covariates used in the analysis of pollutant $p$, and let $x_{pi}$ be the $i^{th}$ row of $X_p$.

We assume that all $c_{pi}$ follow univariate Poisson distributions with density functions $f_c(c_{pi} | \mu(c_{pi}))$, with $\mu(c_{pi}) = \exp(x_{pi}^T \beta_p^* + \varepsilon(c_{pi}))$. We introduce correlation between the $c_{pi}$ by assuming that the three latent effects $\varepsilon(c_{pi})$ at location $i$ follow a tri-variate normal distribution with density function $f_{\varepsilon} (0, \Sigma)$. With respect to the specification of the priors of the parameters of interest, we follow the standard assumptions that the $\beta_p$ follow multivariate normal distributions with density functions $f_{\beta} (\beta_p | b_p, B_p^{-1})$, and that the inverse of the covariance matrix $\Sigma$ follows a Wishart distribution with density function $f_{\Sigma} (\Sigma^{-1} | \delta, R)$.$^{24}$

---

$^{24}$ We closely follow the setup suggested by Chib and Winkelmann (2001).
These assumptions yield the posterior density function $f(C_1, C_2, C_3, \beta_1, \beta_2, \beta_3, \Sigma | X_1, X_2, X_3)$, which is given by

$$f(C_1, C_2, C_3, \beta_1, \beta_2, \beta_3, \Sigma | X_1, X_2, X_3) = \prod_{p=1}^{3} \prod_{i=1}^{N} f_c(c_{pi} | \mu(c_{pi})) \cdot \prod_{p=1}^{3} f_\beta(\beta_p | b_p, B_p^{-1}) \cdot \prod_{i=1}^{N} f_\epsilon(0, \Sigma) \cdot f_\Sigma(\Sigma^{-1} | \delta, R),$$

where $f_c(c_{pi} | \mu(c_{pi})) = f(c_{pi} | \mu(c_{pi}), \Sigma)$.

Convergence of the Gibbs sampler is slow if the posterior correlations of the parameters are high, and reparameterization can help to reduce the posterior correlations. We follow Ibrahim et al. (2000) and center the latent effects on their means, so that $\epsilon(c_{pi}) = \mu(c_{pi}) - x_{pi}\beta'_p$.

This permits us to write the posterior density as

$$f(C_1, C_2, C_3, \beta_1, \beta_2, \beta_3, \Sigma | X_1, X_2, X_3) = \prod_{p=1}^{3} \prod_{i=1}^{N} f_c(c_{pi} | \mu(c_{pi})) \cdot \prod_{p=1}^{3} f_\beta(\beta_p | b_p, B_p^{-1}) \cdot \prod_{i=1}^{N} f_\epsilon(\mu_i - x_i\beta', \Sigma) \cdot f_\Sigma(\Sigma^{-1} | \delta, R),$$

where $\mu_i = (\mu(c_{i1}), \mu(c_{i2}), \mu(c_{i3}))'$, $x_i = (x_{i1}, x_{i2}, x_{i3})'$, and $\beta = (\beta_1, \beta_2, \beta_3)'$. Besides reducing the posterior correlation and thereby improving convergence, centering the latent effects removes the coefficients of the covariates from $f_c(c_{pi} | \mu(c_{pi}))$. The centered model in equation A.2 yields full conditional distributions of $\beta$ that are multivariate normal, while the full conditional distributions of $\beta$ from the non-centered model in equation A.1 can be determined only up to the normalizing constant. Sampling from non-standard distributions is time consuming, while sampling from multivariate normal distributions is very fast. This makes it possible to update all
components of $\beta$ for each equation simultaneously, even if the number of covariates is large. Simultaneous updating of the model components greatly improves mixing of the Gibbs sampler.

It is straightforward to show that the full conditional distributions of $\beta_p$ that follow from the model assumptions in equation 3 are multivariate normal, or

$$
\beta_p \mid \sim N\left[V_p^{-1}\left(B_p^{-1}b_p + \sum_{i=1}^{N} x_{pi}\Sigma^{-1}\eta\right) ; V_p\right],
$$

(A.3)

with

$$
V_p = \left(B_p^{-1} + \sum_{i=1}^{N} x_{pi}\Sigma^{-1}x_{pi}'\right).
$$

Similarly, the full conditional distribution of $\Sigma^{-1}$ is Wishart, or

$$
\Sigma^{-1} \mid \sim W\left[R + \sum_{i=1}^{N} (\mu_i - x_i\beta')(\mu_i - x_i\beta')', \delta + N\right].
$$

(A.4)

For both distributions, very efficient algorithms are available to generate pseudo random numbers. The densities of the full conditional distributions of $\mu(c_{pi})$ can be derived only up to the normalizing constant, so that

$$
f(\mu(c_{pi})) \propto f_x(c_{pi} \mid \mu(c_{pi})) \cdot f_x(\mu_i - x_i\beta', \Sigma)
$$

(A.5)

25 The dot indicates the full set of conditioning variables.

26 See, for example, Gelfand et al. (1990, p.979).

27 To generate samples from the multivariate normal distribution, we used the algorithm based on the Cholesky decomposition that is described in Ripley (1987, pp. 98-99). To generate samples from the Wishart distribution, we used the algorithm described in Odell and Feiveson (1966).
Chib et al. (1998) show how to use the Metropolis-Hastings algorithm to sample from equation 6. Because these distributions are log-concave, one can also use the adaptive rejection method of Gilks and Wild (1992) to draw samples from these distributions.28

We obtained the results shown in the next section from 100,000 runs of the Gibbs sampler after a burn-in of 20,000.29 We used the program Bayesian Output Analysis (BOA) for convergence diagnostics, and checked the robustness of our results by running several alternate chains with different starting values.

---

28 See, for example, Chen et al. (2000, p.42).

29 We reduced the autocorrelation that we observed in some variables (although generally not in the income variable) by running long chains that we thinned by using only every 50th sample.
REFERENCES


