Population Growth and Local Air Pollution: Methods, Models, and Results
Author(s): James C. Cramer
Published by: Population Council
Stable URL: http://www.jstor.org/stable/3115267
Accessed: 20/09/2013 20:00

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Population Council is collaborating with JSTOR to digitize, preserve and extend access to Population and Development Review.

http://www.jstor.org
Population Growth and Local Air Pollution: Methods, Models, and Results

James C. Cramer

Air pollution is a serious threat to human and environmental health in virtually all of the large cities of the world (Brennan 1996; Schneider 1998; WHO/UNEP 1992; World Resources Institute 1996). There are many pollutants; the most common include oxides of sulphur and nitrogen (SO_2 and NO_x respectively), carbon monoxide (CO), ozone (O_3), suspended particulate matter (SPM), and lead (Pb). A recent survey of 20 of the world’s largest megacities (projected population over 10 million in the year 2000) found that World Health Organization guidelines were exceeded by at least one of these pollutants in every megacity (WHO/UNEP 1992). High levels of pollution have been shown to have adverse impacts on respiratory, cardiovascular, and neurological systems in humans, as well as adverse effects on vegetation, agricultural crops, and forests and damage to building surfaces.

Most research on urban air pollution has focused on the sources rather than the causes of pollution. The most important sources include fuel combustion for domestic cooking and heating, power generation, motor vehicles, industrial processes, and disposal of solid wastes by incineration (WHO/UNEP 1992). The most important sources of pollution and types of pollutants vary from city to city. Domestic fuel combustion tends to be more important in cities in lower-income countries, while motor vehicles tend to be more important in cities in higher-income countries. Research on the sources of pollution tends to be quite specific and quantitative; for example, motor vehicle emissions account for 44 percent of SPM and 73 percent of NO_x in Jakarta (World Resources Institute 1996: 6).

Discussion of the causes of air pollution, in contrast, tends to be quite general and vague. It is widely assumed that population growth, industrialization, and socioeconomic development all contribute to air pollution, but their relative contributions are rarely quantified. These relationships,
furthermore, are complex. With socioeconomic development, for example, pollution from domestic fuel combustion often diminishes while pollution from motor vehicles increases. Pollution tends to rise with income because of increased production and consumption, but at high levels of income pollution may decline because of investments in pollution-reducing technologies (Dietz and Rosa 1997). In addition, it is widely acknowledged that government policies and programs aimed at reducing air pollution can be quite effective; thus another fundamental cause of air pollution is the absence or insufficiency of pollution controls and regulations.

Population growth and socioeconomic development are general, ultimate causes of air pollution. More proximate and specific causes link these ultimate causes to the sources of pollution. Proximate determinants include the nature of the transportation system (e.g., whether public transportation is available), the age of the vehicle fleet, types of fuels available (e.g., dependence on coal, use of unleaded gasoline), industrial composition (some industries pollute more than others), and urban spatial structure (degree of sprawl and congestion). Finally, meteorology and topography are important determinants of the impact of atmospheric emissions on ambient air quality. Causal models that incorporate ultimate causes, proximate determinants, sources, and emissions or ambient air quality are rare indeed.

Even rarer are models that explicitly recognize multiple causal directions. As I have noted, population growth is widely regarded as an important cause of air pollution, and air pollution has been shown to have serious adverse impacts on human health. This suggests a simple model of reciprocal causality with negative feedback: population growth causes air pollution, and air pollution reduces the rate of population growth. In this instance the feedback probably is rather weak, since several studies suggest that urban air pollution may account for about 2 percent of all local deaths (WHO/UNEP 1992; World Resources Institute 1996). Other types of feedback may be more important. In the United States, for example, much migration is guided by the search for improved quality of life, not just a better job; urban air pollution may have a much larger impact on net migration than on mortality (Hunter 1998). In addition, increasing levels of pollution may stimulate or provoke government pollution control programs (at least, one would hope so); this also would provide negative feedback, damping the growth of air pollution.

This chapter has two aims. The first aim is to survey the methodological approaches that have been used to study the relationship between population growth and local air pollution. The second aim is to describe a statistical modeling approach and to illustrate it with an empirical case study. The case study utilizes data from California and incorporates multiple causes of pollution as well as feedback loops. I conclude with a discussion of the applicability of the proposed approach to other settings.
Research approaches in the study of population and pollution

Two fundamentally different approaches have been used in research on population growth and air pollution: simulation models and statistical models. In both approaches most research has been at the national or global level, not local; researchers have considered only the impact of population growth on pollution and have ignored the possibility of reciprocal causality and feedback. Simulation is the more common approach, so I will review these models first.

Simulation models

The simplest simulation models use population projections and data on per capita emissions of pollutants. Prospective projections compare the emissions that would occur with no population growth to the emissions that are expected to occur with population growth (Birdsall 1992; Bongaarts 1992; Heilig 1994; Meyerson 1998). Results often are expressed in terms of the percentage of future emissions attributable to population growth. Such simulations sometimes use constant per capita emissions and sometimes incorporate projections of the trend in per capita emissions. Often these simulations are disaggregated into at least two population categories, for example, the more developed and the less developed countries. An especially sophisticated prospective projection (O’Neill, MacKellar, and Lutz 2001) examines greenhouse gas emissions in terms of detailed projections of population and of per capita emissions. The model is disaggregated and takes into account indirect effects of population growth, whereby reduced rates of population growth stimulate more rapid increases in affluence, partially offsetting the beneficial direct effects of slowed population growth.

Essentially identical methods have been used with historical data and retrospective projections (e.g., Harrison 1993; Moomaw and Tullis 1999). Usually these efforts have been based on the slightly more sophisticated IPAT model, in which the trend in emissions (I for “impact”) is a function of population growth (P), the trend in consumption per capita (A for “affluence”), and the trend in emissions per unit of consumption (T for “technology”). Again, results usually are expressed in terms of the percentage of emissions growth that is attributable to population growth. The IPAT model has been widely criticized (Demeny 1991; Dietz and Rosa 1994; O’Neill, MacKellar, and Lutz 2001), both for being a tautological identity and for its assumption that the components (population, affluence, and technology) are independent. These criticisms apply equally to the prospective projections based on population and emissions per capita.

More sophisticated projection models disaggregate the population into a number of categories and utilize complex methods for projecting per capita
emissions. Ridker (1972), for example, used a variety of age groups and household composition categories in the population projection. He used econometric models to estimate household consumption functions for a variety of goods, an elaborate input–output model to convert end-use consumption into necessary resource inputs, and additional econometric models to predict the emissions produced in providing these inputs. Ridker then compared the consumption streams for two different population projections, representing slow and moderate population growth, and estimated the emissions of criteria pollutants (e.g., CO, SPM, NOX) implied by each consumption stream. Comparison of the relative rates of emissions to the relative rates of population growth yielded estimates of the pollution elasticity with respect to population: for example, 0.46 for SPM, 0.26 for CO, and 0.29 for NOX.

The most elaborate projection models utilize computer simulation models based on numerous simultaneous equations depicting the interrelations among many pertinent factors. These models are widely used in urban planning to predict the emissions implied by specific projects or changes in land use; until recently they have not been used often by demographers to study the impacts of population growth. De la Barra (1989) presents the theoretical underpinnings of integrated land use and transportation models, and Wegener (1994) compares the leading models currently in use. These models simulate spatial distribution, land use changes, and patterns of transportation that accompany population growth. They can be used to predict emissions from transportation, industry, and residential activities. These models, however, are most useful in studying the impacts of specific projects or plans, such as construction of a new highway or imposition of new zoning ordinances. In using the models to simulate different scenarios of population growth, numerous assumptions about transportation infrastructure are required, and results may be sensitive to the assumptions.

Statistical models

The second approach to research on population and pollution involves statistical analysis of data on pollution and the factors that cause pollution. This requires identifying all relevant variables and the functional forms by which they are linked to pollution. Very sophisticated models have been developed to link pollution to its sources and to climate and meteorology, but models linking pollution to its anthropogenic causes are relatively primitive. Most such models either examine a single causal factor or make highly restrictive assumptions about functional form. One of the best studies is by Dietz and Rosa (1997), who transform the IPAT model into a stochastic statistical model. They utilize data from 111 countries in 1989 to regress CO₂ emissions on population and affluence (per capita gross domestic prod-
uct). They find that both population and affluence have non-proportional (i.e., nonlinear) effects on CO₂ emissions. A serious deficiency in their specification of the IPAT model is their lack of data on technology.

A more ambitious model has been developed recently and applied to local-level data on pollution in California (Cramer 1998; Cramer and Cheney 2000). This model also is a stochastic transformation of the IPAT model, but it includes measures of technology as well as population and affluence; it has been used to examine a variety of types of pollutants. It warrants detailed description.

Air pollution is a serious problem in California, and stringent air quality standards have been imposed by state and federal authorities. An elaborate regulatory system has been developed to implement mitigation measures and to monitor progress toward attainment of the air quality standards. Because of these regulations, air quality in California actually is improving despite rapid population and economic growth. The regulatory system generates two types of data on air pollution: actual ambient air quality is measured at hundreds of monitoring stations throughout the state to test for attainment of the air quality standards; and emissions of specific pollutants from specific sources are estimated in each county of the state, to aid in the development of new regulations and to monitor progress toward attainment of the standards.

Both types of data—observed ambient air quality and estimated emissions—were examined in connection with population growth, using similar models. In both cases, air quality was measured as a trend, defined as the ratio of pollution levels in two successive time periods. Population also was measured as a trend (i.e., ratio) over the same time period. Trends in air quality and population were compared across different geographic units, usually counties but also cities and places. Trends in regulatory effort¹ and trends in per capita income also were included in the model. The former was used as an indicator of technology, since most technological changes that improve air quality in California have been mandated by government regulations; the latter was used as an indicator of trends in overall consumption and production per capita. Other possible causal factors, such as cultural beliefs and social institutions, were assumed not to vary across areas within California.

Research on population growth and observed trends in ambient air quality is complicated by the pervasive wind currents in California. Pollution measured at a specific site may have been produced by the population at that site,² or it may have blown in from upwind. Carbon monoxide is a primary pollutant and is readily measured as soon as it is produced, so wind currents are largely irrelevant in the model of atmospheric CO. The research showed a strong relationship between local population growth and trends in atmospheric CO at a site. Ozone, in contrast, is the product of a
chemical reaction involving two precursor primary pollutants (reactive organic gases and oxides of nitrogen) and sunlight. This chemical reaction may take place long after the primary pollutants were emitted, so levels of ozone at a site are strongly affected by wind currents. The research failed to demonstrate any relationship between population growth and trends in ozone, despite efforts to identify major wind currents and to incorporate trends in upwind population and upwind regulatory effort into the models (Cramer and Cheney 2000).

Research on population growth and trends in emissions is, in many ways, more straightforward. Wind currents are irrelevant, so upwind causal factors need not be identified. Minor errors in estimating emissions should be similar in successive years and cancel out in the calculation of trends. Major, systematic errors in estimating emissions should be duplicated in estimates of regulatory effort, since the methods for estimating emissions and regulatory effort are similar; hence these errors should be controlled by including regulatory effort in the model.

Research on population growth and trends in emissions in California counties from 1980 to 1990 did, in fact, demonstrate large effects of population on pollution, controlling for trends in regulatory effort and per capita income (Cramer 1998). The complexity that arose here is that population growth had a greater effect on some sources of emissions than on others. Emissions were aggregated into 13 broad source categories in order to obtain reliable estimates of emissions for counties. For source categories associated mainly with consumption or lifestyle activities, the effects of population growth were large (elasticities near 1.0, with per capita income and regulatory effort controlled); examples of these source categories are residential and commercial sources, solvent use, passenger vehicles, and off-road vehicles. For source categories associated mainly with production for a national or global market, the effects of local population growth were small; examples of these source categories are industrial and agricultural processes and off-road mobile equipment. The overall impact of population growth depended on the relative balance of types of source categories.

The statistical analyses of air pollution and population growth in California are deficient in several respects. Wind currents introduce enormous complexity into models of ambient air quality, so at this early stage of demographic research it is premature to utilize these data; at present, research should focus on emissions. In the models reviewed above, emissions were aggregated into 13 broad categories of sources, but little justification was provided for these particular groupings. More specific sources of emissions should be examined in detail before sources are combined in categories. Finally, and most importantly, the possibility of feedback effects and reciprocal causality was noted (Cramer 1998) but could not be examined because of limitations of the data. The empirical case study below, also using
data from California, is intended to replicate the earlier models and rectify these deficiencies.

Population and pollution in California

The aim of this empirical case study is to describe the statistical approach to research on population growth and air pollution in detail, illustrating its implementation, strengths, and weaknesses with real data. I investigate the relationship between population growth and pollution, taking into account multiple causes of pollution and possibilities of feedback or reciprocal causality. I look specifically at population growth and atmospheric emissions in California for the period 1975–95. Available data do not permit the construction of a complete causal model. Instead, the research reported here should be considered a modest first step in this direction. I use longitudinal data and lagged variables to represent alternative causal directions in separate models. In one model I estimate the effect of population growth on air quality, with statistical controls for other plausible causal factors. In a second set of models I explore the effects of air quality on population growth and the other causal factors. The results confirm the need for complex causal models and provide a starting point for further work.

The setting

California is a highly urbanized and industrialized state with a large capitalistic agricultural sector. The state is ecologically diverse, with a long coastline, large inland desert, rich farmland, extensive forests, and imposing mountains. In part because of its environmental diversity and beauty and the abundance of resources, California’s population has grown rapidly. The population quadrupled in size from 1940 to 1990, from under 7 million to over 30 million persons; it is projected to double again in the next 50 years (California Department of Finance 1998b). Much of this growth is due to migration, both from other states and from other countries, but natural increase also contributes significantly (California Department of Finance 1998a). The economy has grown rapidly as well, leading to a very high average standard of living. Real per capita income, however, has grown little since the mid-1970s.

Population and economic growth have put severe stresses on the environment, and this in turn has spawned a very active environmental movement. Among the more serious environmental problems in the state are water scarcity and pollution; diminished biodiversity; loss of wetlands; soil and forest degradation; toxic, radioactive, and solid waste disposal; and urban and suburban sprawl (Palmer 1993). But the most notorious and possibly most serious environmental problem is air pollution. Smog in Los An-
geles is legendary, but other areas suffer as well. Of the 20 most polluted metropolitan areas in the United States in 1995, eight were in California (United States Environmental Protection Agency 1996: Tables A-18, A-19).

Despite rapid population and economic growth, air quality actually has improved in most areas of California since the late 1970s. This is because the vigorous environmental movement has succeeded in enacting legislation creating an elaborate regulatory structure. There is a statewide Air Resources Board and 14 regional (i.e., air basin) air quality management districts; many counties also have air quality management offices. Responsibility for controlling and reducing air pollution is shared among the different regulatory layers. In general, point sources are regulated by counties, area sources by regional districts, and mobile sources by the state agency. Most emissions are caused by mobile sources, especially automobiles, and enormous reductions in emissions have been achieved by statewide legislation mandating cleaner cars and cleaner fuels. Regional and county regulations increasingly are attacking other sources of emissions. Regulations generally are directed at specific pollutants from specific sources.

Success in reducing air pollution is due as much to stringent air quality standards as to the elaborate regulatory structure. Ambient air quality standards are set by both the California and federal governments, although California’s own standards often are more restrictive. The state standard for ozone, for example, limits the maximum concentration of ozone to 0.09 parts per million for any one-hour average; noncompliance is measured as the number of days in which the peak one-hour concentration exceeds this standard. Standards also exist for carbon monoxide, oxides of sulphur, and several other pollutants. Attainment of these standards is determined by measurements taken at hundreds of monitoring stations scattered throughout the state.

If an area is not in compliance with the ambient air quality standards, it (i.e., a county or an air basin) must develop a plan for attaining compliance and for monitoring progress toward attainment. These plans are based on estimates of emissions of specific pollutants from specific sources. Emissions are estimated by local, regional, or state regulatory agencies depending on the type of source; for example, emissions from mobile and many area sources are estimated by the state, and emissions from some area sources and from point sources are estimated by the regions and counties. In general, plans and regulations are directed at the most important sources of emissions first, although concepts of efficiency also play a role in decisionmaking.

The improvement in air quality since 1975 is impressive but uneven. Improvement is seen both for actual concentrations of ozone, CO, and other pollutants (California Air Resources Board 1995b; United States Environmental Protection Agency 1996) and for estimated emissions of reactive organic gases (ROG), CO, and SO$_x$ (but not NO$_x$ or SPM) (California Air
Resources Board 1993). Much of the improvement is due to dramatic reductions in emissions from automobiles and from point-source industrial processes. Even among other (i.e., area) sources of emissions, however, improvements have been made in ROG and SO$_x$ emissions. For these two pollutants, statewide emissions from area sources were about one-third lower in 1995 than in 1975.7

Trends in emissions vary, however. For the other three pollutants considered here—CO, NO$_x$, and PM10 (small particulate matter)—emissions from area sources increased by about half from 1975 to 1995. Trends also vary between counties. While ROG emissions (from the sources examined here) declined by a third statewide from 1975 to 1995, they declined by over half in a number of counties, increased somewhat in many other counties, and more than doubled in two counties. Similar variation occurs with other pollutants.

The challenge confronting us is to explain these trends. Population growth often is cited as the cause of environmental degradation, but this explanation is problematic. Population growth, of course, cannot explain why some emissions increased while others declined. Furthermore, it turns out that declines in emissions often occurred in the larger, more urbanized counties where population growth was greatest, while large relative increases in emissions occurred in small, rural counties. In other contexts, observations of similar discrepancies between trends in population and trends in environmental degradation have led to the conclusion either that population is irrelevant or that the effects of population are conditional (i.e., that the effect of population growth depends upon the institutional and social structure, level of development, type of economy, cultural values, and the like). Such conclusions, however, may be premature.

Statistical models

The approach taken here is that environmental trends and the seemingly anomalous effects of population can only be understood clearly within the context of a properly specified causal model. One key characteristic of such a model is that it is multivariate. Other causal factors may account for divergent trends in degradation; after these other factors are taken into account, the effects of population may appear consistent and systematic. A second characteristic of such a model is that all variables should be measured as relative trends. An absolute increase in population may have very different absolute impacts on the environment in different economic and institutional settings (e.g., at low and high standards of living), whereas a specific percentage increase in population may have the same relative impact in quite diverse settings. These two methodological ploys—specification of relative trends within a multivariate model—are essential for cor-
rect estimation of the effects of population growth on environmental degradation. Full understanding of these effects requires that population growth and environmental degradation be examined within a complete causal model that allows for indirect effects and feedback effects.

The research proceeds in two steps: analysis of the effects of causal factors on trends in emissions, and analysis of feedback effects of emissions on the causal factors. For the first part of the research, I use the same model as in past research (described above and in Cramer 1998). The standard IPAT model \((I=PA T)\) suggests that emissions are produced by three factors: population, affluence, and technology. I modify this model in four ways: I use unconventional definitions of the three causal factors (particularly technology); I measure each causal factor as a trend from one time period to the next; I take the logarithm of each side of the equation, making it additive rather than multiplicative; and I add a residual term, making the model stochastic. The residual term incorporates both random measurement errors in the estimates of emissions and the effects of numerous unobserved variables such as local values, cultures, institutions, and technological changes not mandated by the regulatory system. The model, then, is a type of production function:

\[
\ln(I') = b_0 + b_1 \ln(P') + b_2 \ln(A') + b_3 \ln(R') + e
\]

where the trend in county emissions \((I')\) is regressed on county trends in population \((P')\), per capita income \((A')\), and regulatory effort \((R', \text{ representing mandated technology})\). Because all variables are measured in logarithmic transformation, the coefficients represent elasticities.

The measurement of trend in regulatory effort is crucial. Air quality has improved because of regulations, so without proper controls for this factor the detrimental effects of population growth cannot be detected. Because regulations are costly, they are imposed only after careful analysis of expected benefits. Estimates of emissions reductions due to any regulation are based on laboratory and field experiments, simulations, and experiences elsewhere. For planning purposes, these estimates are used to produce projections of future emissions relative to a baseline level, all else remaining constant. Estimates of emission reductions are called control factors. For example, a control factor of 0.8 for CO emissions in 1995 indicates that the expected cumulative effect of all CO regulations is to reduce CO emissions by 20 percent from the baseline year (e.g., 1975) to 1995. A control factor of 1.0 indicates that no regulations have been imposed. Control factors are computed by state and regional technical staff for each criterion pollutant separately by source (for over 100 specific sources) for each county and each year. In the research on population growth and trends in air quality, control factors are used to construct measures of trends in regulatory effort from one time period to the next.
I estimate the modified IPAT model repeatedly for various criterion pollutants and sources of emissions. Data for all variables are available for the years 1975, 1980, 1985, 1987, 1988, 1990, and 1995 (though for only eight counties in 1988). Trends are estimated from one year to the next, so intervals range from one to five years in duration. With 56 counties and six time intervals, the data comprise a cross-section of time series. Because there are many more counties than time periods and the entire population of relevant counties is included, the correct statistical model for these data is a fixed-effects panel model (Baltagi 1995; Greene 1993; Judge et al. 1985). In effect, this is an ordinary least squares regression with a dummy variable for each county. A log transform is used for all trend variables, so that regression coefficients are interpreted as elasticities.\(^\text{10}\)

The analysis of trends in emissions serves to replicate past research on the effects of population growth on air quality. Perhaps the main novelty of the current research, however, is the effort to transcend such a narrow conceptualization of population–environment dynamics. In the second half of the research, each of the causal factors—trends in population, regulatory effort, and per capita income—is itself specified as an outcome variable. The analysis is restricted to the same set of four variables; no new variables are introduced, as would be required in a full-scale causal analysis of any of the outcomes (e.g., population growth surely is determined by other factors besides pollution, environmental regulations, and per capita income!).

While it is highly likely that the trend in emissions is caused by trends in the causal factors, the reverse is unlikely. Short-term trends in population, regulations, and per capita income most likely are caused by initial levels of the other causal factors, not their trends. For example, population growth over a one- to five-year period probably is influenced by the initial level of pollution (if at all), not by the trend in pollution. By distinguishing levels from trends, the analyses of emissions and its causal factors can be disaggregated into separate models; nonrecursive models for reciprocal causality are not needed. Using longitudinal data to represent a dynamic process by recursive equations is fairly common (Cramer 1980). The specific models used in the analyses of the causal factors are described in a later section. Each model includes initial level of emissions as a predictor, so separate regressions are estimated for distinct combinations of pollutants and sources of emissions. Fixed-effects models are used here as well.\(^\text{11}\)

Data quality and sample restrictions

Observations have been included in the sample only if they were deemed reliable and relevant. For example, data from two of California’s 58 counties have been excluded because these two counties (Alpine and Sierra) have
very small populations sparsely distributed across vast mountainous terrain. Trivial sources of emissions also have been excluded; these are sources that produce, on average, less than 100 tons of emissions of a specific pollutant per county per year. Emissions from natural causes and from unplanned wildfires have been excluded on the assumption that they are unrelated to local population growth. Emissions from several additional sources (e.g., utility equipment, solvents used in dry cleaning) have been excluded because these emissions were estimated from data on county population; only emissions estimates made without reference to the causal factors are used here (California Air Resources Board 1995a).

Estimates of control factors (i.e., regulatory effort) are not available for certain sources of emissions, so these sources also have been excluded from the analysis. Most industrial sources of emissions are excluded, because control factors are not available for specific point sources. Emissions from automobiles and trucks also have been excluded, because control factors are calculated for specific models of vehicles and are not available for county-wide fleets. These exclusions are not as serious as they may first appear. Most industrial sources, while theoretically interesting, would have been excluded anyway, because most industries in California today emit trivial amounts of pollution (less than 100 tons per county per year on average, according to the official emissions inventory). Automobiles and trucks are very important sources of emissions, but the earlier research already showed clearly that these emissions are strongly associated with local population growth; replication is not essential here.

A small number of observations have been omitted from the sample because of apparent errors of measurement that caused undue influence or extreme outliers in data analysis. Separate regressions were estimated for each pollutant and each source included in the analysis. Each regression was carefully analyzed for influential cases or outliers. For example, from the regression of emissions of reactive organic gases from solvents used in commercial and industrial degreasing operations, the plot of standardized residuals against predicted values is shown in Figure 1. Two observations, from county 24 (Merced) in 1988 and 1990, appear as extremely large residuals (one positive, one negative). The same two extreme residuals are seen in the partial regression (i.e., added variable) plot of trend in emissions on trend in population, shown in Figure 2. These extreme residuals are caused by an abnormally small volume of emissions recorded in 1988; emissions of ROG from degreasing in Merced County were greater than 70 tons per year in 1985 and 1987 and greater than 60 tons per year in 1990 and 1995, but were recorded as only 21 tons in 1988. This appears to be an error in the data file; emissions of 71 tons in 1988 would be consistent with the general trend in Merced County. The very low figure for 1988 creates an extremely steep trend downward in emissions from 1987...
to 1988, and an extremely steep trend upward in emissions from 1988 to 1990 (hence the large negative and positive residuals). Rather than assume that the correct figure for 1988 is 71 tons, the two extreme observations (trends 1987–88 and 1988–90) have been excluded from analysis. A total of 126 such outlying or influential observations were excluded.

Another, related problem can be detected in Figures 1 and 2: 1980 data points (i.e., trends from 1975 to 1980) tend to cluster in the upper-right quadrant of the residuals-by-predicted plot and the partial regression plot. These data points have high predicted values because population growth was particularly rapid during the 1975–80 period. The systematically large positive residuals in 1980, however, are troubling. These large positive residuals seem to be associated with a change in methods used to estimate emissions. The change was implemented between 1975 and 1979 (California Air Resources Board 1996), and the new method seems to produce larger estimates of emissions than the old method. With sources for which estimates of emissions were produced by the state agency, the bias

FIGURE 1 Plot of residuals by predicted values, from a regression of trend in ROG emissions from degreasing solvents on trends in population, regulatory effort, and per capita income: Fixed-effects panel model for California counties, 1975–95
FIGURE 2 Partial regression plot of trend in ROG emissions from degreasing solvents on trend in population, controlling for trends in regulatory effort and per capita income: Fixed-effects panel model for California counties, 1975–95

\[
\text{coef} = 0.79525114, \ se = 0.08010512, \ t = 9.93
\]

due to the change in methods should be similar across counties, and the bias can be controlled by a dummy variable for 1980\(^{13}\) (not included in the regression shown in Figures 1 and 2). With sources for which estimates of emissions were produced by regional or county agencies, the impact of the centrally mandated change might vary across counties; in these cases, the 1980 data (i.e., the distorted 1975–80 trend) have been excluded from analysis.

The specific regressions of emissions trends that are estimated here are shown in Table 1, along with the number of observations available for each regression after all the various exclusions. Five criteria pollutants are examined: carbon monoxide (CO), reactive organic gases (ROG), oxides of nitrogen and of sulphur (NO\(_x\) and SO\(_x\)), and small particulate matter (PM10). Twenty-nine regressions are estimated: 7 sources of CO, 3 sources of NO\(_x\), 12 sources of ROG, 2 sources of SO\(_x\) and 5 sources of PM10. In the majority of regressions there are over 200 observations (up to 56 counties and up to six time intervals each), while in two regressions there are fewer
TABLE 1  Data source and number of observations used to estimate fixed-effects regression models of emissions, by type and source of emission

<table>
<thead>
<tr>
<th>Emissions source</th>
<th>Data sourcea</th>
<th>Number of observations</th>
<th>CO</th>
<th>NOx</th>
<th>ROG</th>
<th>SO2</th>
<th>PM10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel combustion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural</td>
<td>D</td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential</td>
<td>A</td>
<td>288b</td>
<td>288</td>
<td>288</td>
<td>287c</td>
<td>288</td>
<td></td>
</tr>
<tr>
<td>Waste burning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural debris</td>
<td>D</td>
<td>195b</td>
<td>180c</td>
<td></td>
<td></td>
<td>183b</td>
<td></td>
</tr>
<tr>
<td>Forest management</td>
<td>D</td>
<td>137b</td>
<td>116b</td>
<td></td>
<td></td>
<td>113b</td>
<td></td>
</tr>
<tr>
<td>Range management</td>
<td>D</td>
<td>174b,c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solvent use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asphalt paving</td>
<td>D</td>
<td>230c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degreasing</td>
<td>A</td>
<td>286c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printing</td>
<td>D</td>
<td>118c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misc. procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petroleum marketing</td>
<td>A</td>
<td>286c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction, demolition</td>
<td>A</td>
<td>288</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farming operations</td>
<td>D</td>
<td>64c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>218b,c</td>
</tr>
<tr>
<td>Pesticide application</td>
<td>A</td>
<td>284c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Waste disposal</td>
<td>D</td>
<td>125c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other mobile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobile equipment</td>
<td>A</td>
<td>288</td>
<td>230c</td>
<td>268c</td>
<td>221c</td>
<td>286c</td>
<td></td>
</tr>
<tr>
<td>Off-road vehicles</td>
<td>A</td>
<td>286c</td>
<td>286c</td>
<td>286c</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Emissions data provided by air management districts are indicated by D; emissions estimated by the California Air Resources Board are indicated by A.

bBecause of lack of regulations, control factors for this type and source of emission do not vary (all or virtually all are equal to unity), so control factor is not included in the emissions model.

cOne or more outliers have been omitted.

than 100 observations (because certain sources are not present in every county, not all counties have data for all six time intervals, and the sample restrictions affect some counties more than others). Table 1 also indicates the pollutants and sources for which there were influential cases or extreme outliers (i.e., cases excluded because of data errors), and the sources for which 1975–80 trends were excluded (data source D).

Models of trends in emissions

The purpose of this analysis is to replicate past research, in which population growth had a large impact on emissions from sources associated with consumption and lifestyle, but only a small impact on emissions from sources associated with production for a global economy. Several of the specific sources of emissions examined here clearly belong to the “consumption and lifestyle” category; these include residential fuel combustion,
petroleum marketing (i.e., gas stations), asphalt paving, and off-road vehicles. With these sources of emissions, we expect to find a large regression coefficient for population growth. Other sources clearly belong to the “global production” category, for which we expect to find a relatively small regression coefficient for population growth; these include agricultural fuel combustion, burning of agricultural debris, waste burning associated with forest and range management, miscellaneous farming operations, and off-road mobile equipment.

Several of the specific sources of emissions listed in Table 1 do not clearly belong to one or the other of these two broad categories; or, more accurately, they belong to both categories. These include activities that take place in both the local commercial sector and the larger industrial sector, with neither sector clearly predominant a priori: solvent use in printing and degreasing, construction and demolition, and unspecified waste disposal. Pesticide use also is included here because pesticides are used extensively in residential and commercial settings as well as in large-scale agriculture. For these sources, we cannot predict in advance whether the regression coefficient for population growth will be large or small; most likely, it will be intermediate.

For each pollutant and source shown in Table 1, the trend in emissions is regressed on trends in population, per capita income, and regulatory effort (with all variables in log transform). The regression coefficients for trend in population from these equations are shown in Table 2. These coefficients are elasticities. For example, the coefficient from the regression of CO emissions from residential fuel combustion of 0.611 indicates that a 10 percent increase in population is associated with about a 6 percent increase in emissions.

Nearly all of the specific hypotheses are weakly supported by the results in Table 2. As expected, relatively large coefficients (larger than 0.6) are found for emissions from residential fuel combustion, off-road vehicles, and asphalt paving; and relatively small coefficients (smaller than 0.4) are found for agricultural fuel combustion, debris burning, mobile equipment, and farming operations (with PM10 but not ROG) and from forest management. The major exceptions to the hypotheses are the small coefficient found for petroleum marketing and the somewhat large coefficient found for range management; one or two exceptions out of 20 coefficients are within the normal range of error and require no special explanation.

These results, while largely consistent with the hypotheses, are considered weak support because the “large” coefficients are not as large as expected, and the “small” coefficients are not as small as expected. In past research with broader source categories, coefficients for “consumption and lifestyle” sources were greater than 0.8 and often near unity, while coefficients for production sources were less than 0.2 and often near zero. I have no explanation for the current, more moderate results.
### Table 2: Partial regression coefficients for population trend in fixed-effects panel models of trends in emissions, by type and source of emission

<table>
<thead>
<tr>
<th>Emissions source</th>
<th>Type of model</th>
<th>Regression coefficients</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CO</td>
<td>NOx</td>
<td>ROG</td>
<td>SOx</td>
<td>PM10</td>
</tr>
<tr>
<td>Fuel combustion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural</td>
<td>D</td>
<td>0.306</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential</td>
<td>A</td>
<td>0.611*</td>
<td>0.592*</td>
<td>0.616*</td>
<td>0.573*</td>
<td>0.610*</td>
</tr>
<tr>
<td>Waste burning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural debris</td>
<td>D</td>
<td>0.316*</td>
<td></td>
<td>0.389*</td>
<td></td>
<td>0.342*</td>
</tr>
<tr>
<td>Forest management</td>
<td>D</td>
<td>0.087</td>
<td>0.178</td>
<td></td>
<td></td>
<td>0.065</td>
</tr>
<tr>
<td>Range management</td>
<td>D</td>
<td>0.466*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solvent use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asphalt paving</td>
<td>D</td>
<td>0.601*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degreasing</td>
<td>A</td>
<td>0.456*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printing</td>
<td>D</td>
<td>1.080*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misc. procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petroleum marketing</td>
<td>A</td>
<td>0.248*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction, demolition</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td>0.663*</td>
<td></td>
</tr>
<tr>
<td>Farming operations</td>
<td>D</td>
<td>0.756*</td>
<td></td>
<td></td>
<td>0.334*</td>
<td></td>
</tr>
<tr>
<td>Pesticide application</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td>-0.449*</td>
<td></td>
</tr>
<tr>
<td>Waste disposal</td>
<td>D</td>
<td>0.678*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other mobile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobile equipment</td>
<td>A</td>
<td>0.377*</td>
<td>-0.008</td>
<td>0.343*</td>
<td>0.400*</td>
<td></td>
</tr>
<tr>
<td>Off-road vehicles</td>
<td>A</td>
<td>0.720*</td>
<td>0.676*</td>
<td></td>
<td>0.762*</td>
<td></td>
</tr>
</tbody>
</table>

*Model A includes a dummy variable identifying 1980 data; model D simply excludes 1980 data. Both models include controls for trends in per capita income and control factor (when the latter is not constant—see Table 1).

*p < .05.

With the sources for which hypotheses are ambiguous, results are mixed. The coefficient for degreasing is intermediate, neither large nor small; but the coefficients for printing, construction and demolition, and waste disposal clearly are large. Evidently the latter activities are associated primarily with consumption, not with production for national or global markets.

While population growth is our main concern here, other results from these equations should be mentioned briefly. For those pollutants and sources for which regulations exist, so that regulatory effort varies and could be included in the model, the coefficient for regulatory effort in most cases is near unity. This is the result that should occur if the "regulatory effort" variable was properly constructed, so it adds credibility to the data and approach. In most equations, the coefficient for per capita income is small and not significant; in those cases where it is significant, the coefficient is as often negative as positive. In some countries, for example China in the 1980s, rapidly increasing prosperity may be a major factor in environmental degradation. In the California context of very high standards of living, a
strong environmental movement, and high environmental consciousness, improvements in income can as easily be used for environmentally friendly consumption (e.g., cleaner cars, more efficient home appliances) as for environmentally harmful consumption (Dietz and Rosa 1997). The unimportance of trend in per capita income in these equations may be due to the offsetting effects of increasing but sometimes environmentally friendly consumption; thus, it was unexpected but is not surprising.

Models of feedback effects

As noted earlier, the model analyzed in the preceding section is causally incomplete and deficient. A more complete causal model depicting the dynamic linkages among the four variables considered here is displayed in Figure 3. The right-hand portion of this model shows the linkages examined in the preceding section: effects of trends in population, regulatory effort, and per capita income on emissions. What is of interest now is the left-hand portion of the model, which shows that the causal factors may be linked to each other (with appropriate time lags) and that each in turn may be affected by the level of pollution.

**FIGURE 3  Dynamic recursive causal model of emissions**

<table>
<thead>
<tr>
<th>Level</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_0$</td>
<td>$T_0 - T_1$</td>
</tr>
<tr>
<td>$P$</td>
<td>$dP$</td>
</tr>
<tr>
<td>$CF_i$</td>
<td>$dCF_i$</td>
</tr>
<tr>
<td>$E$</td>
<td>$dE_i$</td>
</tr>
<tr>
<td>$E_i$</td>
<td>$dI$</td>
</tr>
<tr>
<td>$I$</td>
<td></td>
</tr>
</tbody>
</table>

Where:
- $P$ = population
- $CF_i$ = control factor for emissions from source $i$
- $E$ = emissions from all sources
- $E_i$ = emissions from source $i$
- $I$ = per capita income
Each of the causal factors is the outcome of a complex process involving numerous determinants. Local population growth, for example, is the sum of natural increase and net migration; in the California context, these depend upon such things as relative wages and the demand for labor, the cost of living, racial and ethnic composition of the population and the existence of immigrant communities to serve as network magnets, and such quality of life factors as crime rate, quality of schools, recreation opportunities, and climate (Cadwallader 1992; Hunter 1998). Introduction and implementation of environmental regulations and changes in county per capita income are at least equally complex phenomena. I have made no effort to specify comprehensive models of the causal factors. Instead, I restrict my attention to the four variables in the emissions model and empirically explore relationships among them, as shown in Figure 3. Using these variables, I construct and estimate a plausible model of each causal factor in turn. As a further simplification I examine the models of the causal factors independently, under the implausible assumption that the residuals in the three models are uncorrelated. The results that follow, therefore, should be considered exploratory, incomplete, and suggestive only.

Population growth

The model in Figure 3 indicates that the short-term trend in population may be affected by all three of the variables considered here: regulatory effort, emissions, and per capita income. Regulations in place in a county at the beginning of a time interval could affect population growth during that interval in several ways, both of which would seem to impede growth: regulations could increase the cost of living; and regulations could symbolize an alienating bureaucratic political system that stifies individual behavior (possibly relevant in the California “frontier” context). The level of emissions can affect subsequent population growth in two ways: by reducing natural increase (presumably mainly by raising mortality), or by discouraging in-migration or encouraging out-migration (Hunter 1998). The health effects of air pollution are very real but demographically relatively small (Anderson et al. 1996; Morgan et al. 1998; Schwartz 1993), so migration probably is the important factor here. Air pollution mainly affects in-migration, not out-migration (Cadwallader 1992). Finally, the initial level of per capita income could affect subsequent population growth in several ways, with diverse consequences: for example, wealth may serve as a magnet, attracting migrants; high-income areas tend to have high housing prices, discouraging in-migration; and higher-income areas are more likely to support growth-control measures. Because of these diverse consequences, one cannot predict a priori the direction of effect of level of per capita income on population growth.

Estimating the effects of initial levels of regulatory effort, emissions, and per capita income on subsequent population growth is complicated by
the problem that regulatory effort and emissions are measured separately for each of the 29 combinations of pollutants and sources of emissions shown in Table 1, while per capita income and population growth are the same for all 29 categories. Estimating 29 separate equations for population growth is redundant and inefficient, but aggregating emissions and regulatory effort into single measures is difficult. The measure of emissions that is empirically feasible yet relevant here is emissions of a particular pollutant from all sources;\(^1\) this yields five measures and requires a separate regression equation for each.

An analogous measure of regulatory effort (i.e., overall regulation of a particular pollutant, regardless of source) would be relevant but is not feasible. Initial level of regulatory effort is defined as the cumulative control factor for a specific pollutant and source from 1975 to the initial year in question; this is a decimal number less than or equal to unity, and these numbers cannot meaningfully be summed (as is done with emissions), averaged, multiplied, or otherwise combined with the data available. The solution adopted here is to pool all observations for a particular pollutant into one sample, and to estimate a separate regression for each of the five pooled samples. Within each pooled sample, most counties in a given year appear in the sample multiple times (once for each source of emissions, except in cases of missing or deleted data). The values of population growth and initial levels of emissions from all sources and per capita income are identical for each of these multiple appearances of a county in a given year, while the value of regulatory effort differs and is specific to the particular source of emissions.\(^1\) In this specification the coefficient for "control factor" represents the average impact of regulations of particular sources, not the impact of overall regulatory effort. For unregulated sources, all control factors are unity; a dummy variable is used to identify these sources.

Estimates of coefficients from the five regressions of trend in population are presented in Table 3. The coefficient for initial level of emissions from all sources is consistently negative and significant in the five regressions, although in several cases it is relatively small in magnitude. Nevertheless it is clear that population growth is slower in those counties with higher levels of air pollution. This is strong evidence for an equilibrating feedback effect: population growth contributes to air pollution, and this pollution in turn retards further growth.

The coefficient for per capita income also is consistently negative and significant in the five equations, indicating that population growth is slower in higher-income counties. Most population growth in California during this time period was due either to immigration or to the high fertility of earlier immigrants. Apparently immigrants are not drawn to high-income counties, perhaps because of the cost of housing or lack of demand for their labor skills. Finally, the coefficient for initial regulatory effort generally is small and not significant. This may be an artifact of inappropriate
### TABLE 3  Regressions of trend in population on initial levels of emissions from all sources, source-specific control factor, and per capita income*: Fixed-effects panel models, by type of pollutant

<table>
<thead>
<tr>
<th>Predictors, initial levels</th>
<th>CO</th>
<th>NO\textsubscript{a}</th>
<th>ROG</th>
<th>SO\textsubscript{a}</th>
<th>PM10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emissions, all sources</td>
<td>-0.203**</td>
<td>-0.020*</td>
<td>-0.042**</td>
<td>-0.035**</td>
<td>-0.184**</td>
</tr>
<tr>
<td>If initial year is 1975</td>
<td>-0.003</td>
<td>0.051**</td>
<td>0.061**</td>
<td>0.048**</td>
<td>-0.007</td>
</tr>
<tr>
<td>Control factor</td>
<td>0.301*</td>
<td>0.028</td>
<td>0.002</td>
<td>-0.003</td>
<td>-0.267</td>
</tr>
<tr>
<td>If source is not regulated</td>
<td>-0.001</td>
<td>\textsuperscript{b}</td>
<td>-0.003</td>
<td>\textsuperscript{b}</td>
<td>0.001</td>
</tr>
<tr>
<td>Per capita income</td>
<td>-0.078**</td>
<td>-0.164**</td>
<td>-0.185**</td>
<td>-0.224**</td>
<td>-0.032</td>
</tr>
<tr>
<td>Constant</td>
<td>2.748**</td>
<td>1.774**</td>
<td>2.170**</td>
<td>2.377**</td>
<td>1.885**</td>
</tr>
<tr>
<td>(R^2) within</td>
<td>0.377**</td>
<td>0.352**</td>
<td>0.329**</td>
<td>0.385**</td>
<td>0.345**</td>
</tr>
<tr>
<td>Pooled N</td>
<td>1451</td>
<td>804</td>
<td>2537</td>
<td>508</td>
<td>1090</td>
</tr>
</tbody>
</table>

*All variables are in log transform, so coefficients represent elasticities.
\textsuperscript{a}Dummy variable omitted; all sources were regulated at least somewhat.
\textsuperscript{b}**p < .01
\textsuperscript{p} < .05

measurement, or it may indicate that air pollution regulations have no effect on population growth.

### Regulatory effort

Regulatory effort is measured for specific pollutants and sources of emissions, so it is examined in separate regressions for each of the 29 combinations of pollutants and sources shown in Table 1. The causal model in Figure 3 shows regulatory effort responding to four factors: initial level of population, per capita income, emissions from all sources, and emissions from the specific source being regulated. The importance of emissions is self-evident: regulations would not be imposed unless pollution were a problem. Demonstrating this linkage is important, however, because there is always the possibility that regulations are not imposed even when pollution is a problem. There also is the question of whether regulations are imposed in response to the overall level of pollution, regardless of which sources contribute most to the problem; or whether specific sources are regulated whenever they produce emissions, without regard to the aggregate problem of level of emissions from all sources combined. In each equation of regulatory effort during a time interval, where regulatory effort is for a particular pollutant and source, two measures of initial level of emissions are specified: emissions from that particular source, and emissions from all other sources combined (net of that particular source).

Population size could affect regulatory effort in several ways. In counties with large populations, problems of interdependence and externalities may be more apparent and public tolerance of regulations may be greater.
Counties with large populations may have larger and better-trained planning departments that can correctly assess air quality and the need for regulations. Finally, public support for environmental protection is strong in California, but so is opposition by well-organized interest groups; large local populations may be more difficult to mobilize in favor of regulation. The overall impact of population size is difficult to predict. The effects of per capita income are more clearcut, at least judging from the broad historical and cross-national record: higher-income people are more concerned with quality of life issues (like clean air), and are better able to escape onerous aspects of regulations. Thus, I expect that regulatory effort is greater in higher-income counties.

Empirical problems again must be confronted. The majority of pollutants and sources are unregulated; with 19 of the 29 combinations of pollutant and source, there is too little variation in regulatory effort between counties for reliable analysis. In the remaining ten cases, initial levels of population and emissions from all sources are very highly correlated (r > 0.85), so coefficients estimated for these variables are unreliable. Other correlations among the predictors are modest in size, so collinearity should not affect other coefficients.17

Because of these empirical problems, I will describe the regression results but not present them in detail. In all ten regressions of regulatory effort during a time interval, the coefficients for initial level of emissions from the particular source under consideration are large in magnitude, significant at the .01 level, and negative. Control factors are measured as expected reductions in emissions, so a smaller control factor (e.g., 0.6 compared to 0.8) represents a greater regulatory effort. The negative coefficients, thus, indicate that counties with more emissions from a particular source made greater regulatory efforts to reduce those emissions. This is highly reassuring; planning and regulation are not always so rational. Coefficients for net emissions from all other sources, while unreliable because of collinearity, generally are quite small as well as not significant, and they are as likely to be positive as negative. This suggests that regulations are imposed in response to specific emissions, not in response to overall levels of emissions (i.e., regulations target the actual offenders).

Coefficients for initial level of population are large, significant, and positive in six of the ten regressions of regulatory effort; they are small or negative in the remaining four regressions. Given the inconsistency in results and the problem of severe collinearity, these results at best are suggestive; what they suggest is that regulatory efforts generally are weaker in counties with larger populations.

Coefficients for initial level of per capita income generally are small, not significant, and inconsistent in direction. Per capita income does not seem to be an important determinant of county regulatory effort.
Trend in per capita income

The causal model in Figure 3 shows the trend in county per capita income responding to three factors: initial levels of population, emissions from all sources, and regulatory effort. The first two of these factors are included largely for completeness and symmetry, not because of strong hypotheses or theory. In agricultural or forest settings, air pollution can reduce yields and diminish incomes. This may be a problem in certain local settings in California, but it is unlikely to be an important consideration in statewide comparisons of counties. In any case, initial levels of population and emissions from all sources are highly correlated, precluding reliable empirical estimates of relationships.

There is, however, considerable interest in regulatory effort. The general issue is whether environmental regulation obstructs or diminishes economic prosperity (Repetto et al. 1996); here the specific question is whether greater initial regulatory effort is associated with slower growth in per capita income. As was the case with population growth, the trend in per capita income is the same in a county regardless of source of emissions, but regulatory effort is measured with respect to each source of emissions. As before, I pooled observations into five samples, one for each pollutant. In all five regressions of trend in per capita income, the coefficient for regulatory effort is small and not significant. Clearly the determinants of trend in per capita income are complex and many relevant factors are not specified in these equations, so these results are at best suggestive. Nevertheless, the results are not consistent with a hypothesis that regulatory effort seriously diminishes prosperity.

Summary

This case study of population growth and air pollution has proceeded in two steps. The first step was an analysis of the direct effects of three causal factors—population, per capita income, and regulatory effort—on emissions. In the second step I explored how the three causal factors are interrelated and how each, in turn, is affected by emissions. The purpose of this second step was to identify possible indirect causal effects on emissions and feedback effects from emissions. The two steps can be conducted separately by taking advantage of time lags in longitudinal data.

Two types of feedback are significant. One is seemingly obvious: initial levels of emissions influence subsequent regulatory effort. This shows that California’s regulatory structure works. State and federal laws require that if ambient air quality standards are violated, mitigation plans must be developed and implemented. The results presented here show that areas with high levels of emissions do indeed enact regulations to reduce those
emissions; the regulations are directed specifically at the offending sources of emissions. The second type of feedback is less obvious and more intriguing: high initial levels of emissions deter subsequent population growth. More detailed research is needed to determine whether pollution impedes in-migration or encourages out-migration (Hunter 1998), and whether pollution matters because of market prices (e.g., for housing) or perceptions of quality of life. In any case, both types of feedback are negative: high initial levels of emissions produce responses that dampen subsequent growth in pollution. Because of the difficulty of comparing levels and trends of variables, it is not clear how strong the dampening feedback is; statistically and theoretically, it is significant.

Two types of indirect effect on emissions also are significant. First, size of population is associated with regulatory effort. Other things being equal (such as initial level of pollution), areas with larger populations are less likely to implement additional regulations. Indirectly, then, a large population is associated with a greater increase in emissions (or smaller decline in emissions); this indirect effect is consistent with and amplifies the direct effect of population growth on emissions. Secondly, per capita income is associated with population growth. Areas with high levels of per capita income experience slower population growth; as a consequence, indirectly these areas experience less pollution. Additional research is needed to illuminate the mechanisms and motivations involved here.

Several non-significant (non)findings are noteworthy. In the data examined here, initial regulatory effort appears to be unrelated both to subsequent population growth and to the trend in per capita income (other things being equal). Other types of government regulation may be offensively intrusive and burdensome, but efforts to maintain clean air seem to be inconsequential in these two respects. The converse is not necessarily true. While initial size of population is associated with subsequent regulatory effort, as described above, initial level of per capita income is not associated with subsequent regulatory effort. Surprisingly, while affluent areas prevent or avoid population growth, they neither promote nor oppose clean air regulations (other things being equal).

The indirect and feedback effects described above come from very incompletely specified causal models in which many relevant causal factors are omitted. They must, therefore, be considered suggestive hypotheses, not conclusive results.

Applicability of the approach elsewhere

The statistical approach to studying population growth and air pollution produces useful and insightful results in California. The obvious question,
then, is whether the approach can be applied elsewhere. The answer, of course, hinges upon the institutional context and data availability "elsewhere." In different institutional settings the specific model described here may require minor modifications or, alternatively, the same model may be used but it will produce different results requiring different interpretations.

The issue of interpretation arises because of the problem of units of analysis. In the analysis of emissions (the first stage of the model), the differential effects of population growth occur because of boundaries. Population and emissions are measured within specific political jurisdictions, for example counties in California. Local population has large effects on emissions from sources that are associated with activities that take place within these local boundaries, and has small effects on emissions from sources that are heavily involved in transportation or trade across the boundaries. Different results might be obtained for different jurisdictions or sets of boundaries and in areas with different patterns of internally versus externally generated activities. This is a standard problem in statistical geography; it necessitates care in interpreting results, not a different model.

An important boundary problem that was overlooked in the case study presented here (because of lack of data) involves emissions from mobile sources such as automobiles, trucks, buses, ships, and airplanes. Emissions from such sources will be measured for specific jurisdictions or sites, but the people using these mobile sources may be transients, not residents of that jurisdiction or site. This problem will be more acute the smaller the jurisdiction, so larger areal units should be used. It may be necessary, nevertheless, to modify the basic model to take account of transient populations.

Regulatory effort is specified as one of the three causal factors in the model for California because the environmental regulatory system is highly developed and most changes in technology that have important impacts on emissions are mandated by this system. In other settings the regulatory system may be less effective or absent entirely, and modifications of the model may be warranted. For example, regulatory effort may be omitted as a causal factor, or a measure of market-driven technological change might be added to the model. Such changes would require modification of both the model of emissions and the models of feedback.

The model of emissions is estimated separately for different sources of emissions. The specific sources used for analysis undoubtedly would vary across settings, and the results obtained may depend on the specific sources used. Relatively detailed sources were used here, compared to the broad categories used earlier (Cramer 1998); even here, however, the sources are aggregated, and this aggregation can pose problems. In California the most important instance of aggregation is the source category of automobiles. Automobile emissions vary dramatically depending on the model and age of the car, so changes in fleet composition can have large impacts on the trend in automobile emissions. Some changes in fleet composition may
be associated with per capita income and thus would be controlled in the model (had data been available to permit analysis of automobile emissions); other changes in fleet composition, however, may be correlated with population growth, and their effects may be confounded. Other examples of aggregated source categories include residential fuel combustion, which lumps together different end uses (e.g., cooking, space heating, water heating) and different fuels (wood, natural gas, electricity); and solvent use from degreasing, which reflects the composition of industrial and commercial activities. In other settings, different levels of aggregation of sources may be desirable in order to minimize the problem of changes in source composition and to focus attention on policy-relevant sources (e.g., changes in residential fuels or modes of transportation).

In California emissions are produced by many sources, although automobiles are the most important. In some settings, especially in lower-income countries, most emissions may be produced by only one or two sources, such as domestic fuel combustion, burning of forests or agricultural wastes, a local industry, or automobiles and trucks. In such cases research should focus on just the important sources and the model should be adapted to the characteristics of each specific source.

The model of population growth and air pollution described here is feasible only because time series data on four key variables are available for a large sample of local jurisdictions. If a long time series of data is available but for only one or a few jurisdictions, the same conceptual model can be used but with a different estimator (e.g., an autoregressive time series model instead of a fixed-effects panel model; see Greene 1993; Kmenta 1986). If data are available for a large sample of jurisdictions but for only one or two time periods, then a simultaneous-equations model with reciprocal causality (endogeneity) will be necessary; this will require additional variables for proper identification. In any case, data on population and some measure of income or consumption probably are readily available, but in many settings estimates of emissions and control factors (the basis for measuring regulatory effort) may not be available. If monitoring-station data on actual ambient air quality are used in place of estimated emissions, then several changes in the model may be necessary. Measures of ambient air quality, of course, combine pollution from all sources, so separate models for different sources could not be estimated and the richness of the model of emissions would be lost. Seasonal data on ambient air quality should be used, selecting seasons when wind currents are minimal. In addition, attention should be focused on primary pollutants that are detected immediately after emission, such as CO and suspended particulates, rather than secondary pollutants that accrue over time such as ozone. Alternatively, upwind population could be added to the model as another causal factor, although experience with such a model is not promising (Cramer and Cheney 2000).
Future directions

Aside from replicating this research in settings other than California, what needs to be done? Certainly one necessary step is to improve the model by including additional causal factors. In the model of emissions, if source categories cannot be disaggregated satisfactorily because of data limitations, then measures of composition should be added to the model. For example, Cramer (1998) combined all industrial sources into one source category because the emissions from any one industry are so small and unreliably estimated; in this case it would have been desirable to include a measure of industrial composition in the model to control for the effects of changing composition (e.g., decline in petrochemicals, growth in electronics and textiles) on the trend in emissions. Another desirable addition to the model would be a measure of culture or lifestyle. At present the only indicator of what people do (as distinct from how many people are doing it) is per capita income. Income surely is important, but it is not the only determinant of automobile use, size of house (a key factor in residential fuel combustion), disposal of waste, or demand for pollution-intensive commodities (Stern, Young, and Druckman 1992).

The models of feedback presented here are even more deficient, as explicitly noted above. Local population growth depends upon relative wages, demand for labor, cost of living, proximity to sources of migration and immigration, racial and ethnic composition of the population, quality of schools, local amenities (entertainment and recreation facilities, local university, and the like), and climate as well as the factors specified here—initial levels of income, pollution, and regulatory effort (Cadwallader 1992; Gober 1993). The imposition and implementation of environmental regulations depend upon the local industrial structure, strength of local and regional environmental organizations, existence of governmental coalitions and credibility of the regulatory agency, and the educational level and environmental attitudes of the public as well as the factors specified here (Marzotto, Burnor, and Bonham 2000; Sabatier and Jenkins-Smith 1993). These additional variables should be included in the models of feedback in order to get good estimates of the effects of the variables of interest.

If most emissions are produced by only a few sources, then the general model should be modified to suit the specific characteristics of those sources. Here many sources were examined, and the same general model was applied to each for the sake of comparability; this ignored the unique features of each source. Residential fuel combustion, for example, depends upon household size and composition as well as number of households; it is not a function merely of total number of persons in a jurisdiction. In California most residential fuel combustion is for space heating and cooling; emissions from residential fuel combustion depend upon the local climate and vegetation (e.g., tree shade), daytime occupancy of the home...
(which depends upon age structure of the household and employment and school enrollment), and thermal comfort thresholds of the inhabitants (Cramer et al. 1984, 1985). Clearly, different models would be appropriate in different settings.

In many settings the majority of emissions are produced by automobiles. Often more than one-half of the emissions from automobiles are produced during the first minute after a cold start, so a model of automobile emissions would have to account for number of trips as well as duration of each trip (Cameron 1991). Number of trips is related to the density of automobiles (e.g., number of cars and drivers per household), which in turn is related to affluence. Number of trips also is related to the activity schedules of household members and to the spatial locations of residence, work, schools, commercial shops, and recreation. Duration of trips depends on urban spatial structure and on congestion (or average speed of trips). Models of emissions from automobiles (and other modes of transportation), therefore, may shift the focus away from local population size toward the spatial distribution of population and land uses. This may be a desirable direction for research, as local jurisdictions may have more policy leverage over land use and population distribution than over population size.

Notes

This research was partially supported by a grant from the Committee on Research, University of California, Davis. Data were kindly provided by the California Air Resources Board; the advice and assistance of CARB staff, especially Martin Johnson and Andy Delao, are greatly appreciated. Robin Cheney helped construct the data file with great care, dedication, and humor.

1 The quantitative measure of regulatory effort is described below in the case study.

2 Part of the "population at that site" may be transient and not measured well in official statistics. In California, motor vehicles are a major source of atmospheric emissions, and much motor vehicle usage is for commuting and long-distance travel or hauling. Thus many drivers may not be residents of "the site" and not counted as part of the local population. This problem should affect the model of CO as well as the model of ozone; since the former "works" well, the problem does not seem to be serious.

3 These also are called volatile organic compounds (VOC).

4 Five of the six most polluted areas are in California: Riverside-San Bernardino, Los Angeles, Bakersfield, Ventura County, and Fresno. (The sixth site is Houston.) Other California sites in the "worst 20" are Oakland, Sacramento, and San Diego.

5 Point sources of emissions are large manufacturing plants. Area sources are smaller, dispersed stationary sources such as dry cleaning establishments, farms, and residential homes. Mobile sources include automobiles, trucks, other forms of transportation, equipment (from large tractors to forklifts and lawn mowers), and recreational vehicles (e.g., boats, snowmobiles).

6 That is, the main consideration is to achieve the largest possible reduction in emissions. The effort or cost required to achieve this reduction is a secondary consideration, mainly because historical experience indicates that costs or effort tend to be greatly exaggerated a priori and because the health and economic benefits of reducing pollution exceed almost any plausible cost, but are proportional to the amount of reduced emissions.
7 This and subsequent statements about trends in emissions are based on estimates from my data set and thus refer only to the specific sources of emissions included in this research, most of which are area sources.

8 For example, Modoc County is a small, rural county with nearly pristine air quality. Population growth in Modoc County causes far less pollution (in tons of emissions per year) than comparable growth in Los Angeles County, but the percentage increase in pollution may be the same in the two counties.

9 A control factor of 1.0 could indicate that regulations are expected to be completely ineffective, but this is unlikely. In principle, control factors can exceed 1.0. This would occur if a regulation were withdrawn, causing an increase in emissions; or if a regulation designed to reduce emissions of one criterion pollutant inadvertently caused increased emissions of another pollutant.

10 The coefficient for population growth, for example, indicates the percentage change in emissions due to a one percent increase in population (other things being equal).

11 The 29 equations of trends in emissions (for different combinations of pollutants and sources of emissions) are not independent, so a "seemingly unrelated equations" (SUE) approach might seem warranted (Greene 1993; Kmenta 1986). The same could be said for the equations for trends in the causal factors. According to Greene (1993: 488–489), however, SUE has no advantage if the explanatory variables are identical in the equations, and little advantage if the explanatory variables are highly correlated. This is the case here, so SUE is not warranted.

12 Outliers and influential cases were examined carefully and conservatively, and were omitted from the sample only if they appeared to be the result of errors in the data. Observations were not excluded simply because they were incongruent with the model (i.e., they showed up as large residuals) or were unduly influential (i.e., associated with large values of Cook's Distance). For guidelines on assessing outliers, see Bollen and Jackman (1990), Cook and Weisberg (1999), and Fox (1991).

13 In the regressions of emissions trend, the dummy variable controls for systematic bias in the 1975–80 trend relative to other trend intervals. In the regressions of trends in population, regulatory effort, and per capita income, the dummy variable controls for systematic bias in the 1975 level of emissions.

14 Consider, for example, that in some counties in recent time periods (but not in other counties or earlier time periods) outdoor barbecues, lawn mowers, household appliances, and commuting patterns (e.g., ride-sharing, parking availability) have been regulated.

15 This, of course, is a massive simplification necessitated by the data. The relevant concept for migration is air pollution that is visible, either to the eye or to the nose (or suspected or statistically demonstrated and widely publicized, even if not directly detectable by individuals). We have data only for specific emissions such as CO and ROG (which cannot readily be aggregated, e.g., total tons of “junk”), not for pollutants like ozone. Finally, our measure of “total” emissions is summed over available source categories; it does not include emissions from automobiles or trucks, which contribute the majority of emissions of most pollutants.

16 Because of the substantial redundancy of multiple observations that are identical in most respects, the sample sizes in effect are inflated; to compensate, a more stringent level of significance is utilized here.

17 Collinearity was not a problem in regressions of trend in emissions or trend in population.

18 In the latter case, “external” population is more important than local population.

References


