Volume Title: Economic Analysis of Environmental Problems

Volume Author/Editor: Edwin S. Mills, ed.

Volume Publisher: NBER

Volume ISBN: 0-87014-267-4

Volume URL: http://www.nber.org/books/mill75-1

Publication Date: 1975

Chapter Title: Acute Relationships Among Daily Mortality, Air Pollution, and Climate

Chapter Author: Lester Lave, Eugene Seskin

Chapter URL: http://www.nber.org/chapters/c2839

Chapter pages in book: (p. 325 - 350)

Acute Relationships Among Daily Mortality, Air Pollution, and Climate

Lester B. Lave, Carnegie-Mellon University
and Eugene P. Seskin, Carnegie-Mellon University

Introduction

"Killer fogs" and other acute air pollution episodes have occurred in the past fifty years.[1] The highly industrialized Meuse Valley of Belgium experienced climatic conditions permitting the buildup of abnormally high levels of air pollutants (particularly sulfur dioxide) during December 1930. Over a five-day period, approximately 6,000 people became ill and approximately 60 died (most of whom were elderly persons or those with previous heart and lung conditions). This was more than 10 times the number of deaths which would normally be expected. See [9].

In October 1948, a similar situation took place in Donora, Pennsylvania. Within three days, almost 6,000 people (over 40 per cent of the population) became ill and about 20 deaths were reported. This, again, was approximately 10 times the expected number of deaths and again the aged were most susceptible (the average age of the dead was 65). See [32].

London was enveloped by a dense fog in December 1952 and, in a two-week period, 4,000 excess deaths were attributed to the abnormally high concentrations of sulfur dioxide and smoke. Unlike the previous episodes, all age groups were affected. See [29].

Severe air pollution episodes, such as the three mentioned, do adversely

NOTE: We thank Martin S. Geisel for helpful comments. This research was supported by a grant from Resources for the Future, Inc. and by Fellowship AP48992-03, Air Pollution Control Office, Environmental Protection Agency.

1. For a detailed review of the following three episodes as well as several other air pollution incidents, see [1] and [26].

affect health. Little is known of the consequences of acute episodes of lesser severity. Long-term exposure to low levels of pollution have been investigated statistically by Lave and Seskin [22–24] across 117 U. S. Standard Metropolitan Statistical Areas (SMSAs) in 1960. We have also examined annual observations on a number of cities to explore further the association between air pollution and mortality. A consistent and significant association between air pollution and mortality was exhibited in each of these studies. We shall extend this investigation by looking at the former question of the association between daily air pollution levels and daily mortality. Before presenting the analysis, we review briefly studies which were similar in scope.

A Brief Review and Critique

Greenburg, et al. [16] investigated a November 1953 New York City episode where sulfur dioxide and smokeshade reached unusually high levels. Using analysis of variance to compare the period with six control years and assuming a three-day lag in the effect of air pollution on mortality, they concluded that there was a statistically significant increase in the number of deaths and that the increase was generally distributed over all age groups.[2] In a subsequent study, Greenburg, et al. [14] looked at pediatric and adult clinic visits during this period. Using the same control periods and again assuming a three-day lag, they found an increase in upper respiratory illnesses and cardiac visits at the four hospitals under observation. Other air pollution episodes in New York City were also scrutinized by Greenburg and his associates [15] with less consistent results.

McCarroll and Bradley [27] examined daily mortality, air pollution (sulfur dioxide and smokeshade), and weather data for New York City during a three-year period (1962–1964). The interval of time included five air pollution episodes. Using only graphical techniques, McCarroll and Bradley concluded that the periodic peaks in mortality were associated with periods of high air pollution.[3] They further noted that such peaks were not followed by drops of a sufficient degree to compensate for the excess deaths, hence they inferred that what they were detecting was

2. This was based on investigations of the Donora episode discussed above.

3. In an attempt to control for variations in climatic and other conditions, they included a 15-day moving average as a baseline from which to judge mortality peaks. In addition, graphs of temperature and wind speed were included for comparison.

more than a hastening effect. Finally, they remarked that the influence on death rates occurred in the 45–64 year group as well as the elderly over 65.

During Thanksgiving of 1966 still another episode occurred in New York City. Glasser, et al. [12] investigated the relationship between mortality and morbidity and indices of SO_2 and smokeshade. For a seven-day period, they found a total of 168 excess deaths, using a number of different control periods, some of which took into account possible lag effects. They also found the number of clinic visits for bronchitis and asthma at seven New York hospitals rose. It was concluded that the rise in temperature at the time of the episode could not account for the observed increase in mortality, although it might have been a contributing factor [12; p. 694]. In another study covering this episode, Becker, et al. [2] utilized questionnaires in a study of 2,052 executive and clerical personnel and found that ". . . as the air pollution levels increased, a greater response to symptoms of dyspnea, cough, sputum, wheeze, eye irritation, and general discomfort was elicited from the study subjects" [2; p. 419]. The direct effects of weather conditions were not accounted for, although they did consider the smoking histories of individuals.

These studies are difficult to evaluate for a number of reasons. Informal analyses such as the use of graphical techniques, are inadequate. Ad hoc procedures for analyzing small data sets require extreme care in interpretation; there is not justification for using different ad hoc assumptions (such as different lag structures) in analyzing other data sets.

In addition there are a number of more general problems that should be noted. The first surrounds the use of questionnaires and other voluntary responses to measure the effect of air pollution. Awareness of high pollution levels is increased by news coverage and other publicity. Thus, a major part of the measured response is likely to be due to the type and extent of news coverage.

Another problem surrounds the estimation of the number of "excess" deaths during a pollution episode. Expected mortality is estimated by computing averages for adjoining periods or for a previous period. No adjustments are made for population changes or variations in other factors known to affect the mortality rate. As an example of the magnitude of this problem, extrapolating the experience in Donora to London would have led to a prediction that over one million people would have been made ill by the fog, whereas the actual number was much smaller.

Finally, there are many factors that affect the daily mortality rate (or morbidity rate) which have not been controlled. McCarroll and Bradley found that during one episode there was a sharp increase in the number

of deaths although the pollution measures they considered were not unusually high (an atmospheric inversion did take place). They pointed out the need for studying other relevant factors. Regarding the Thanksgiving study by Glasser, et al., Eckhardt remarked in a letter to the editor, "Lung cancer deaths can hardly be related to air pollution unless the following factors have been investigated: Is the case postoperative? How many days postoperative? How long did the patient have lung cancer" [7; p. 837]? In addition to these factors, the researchers were aware of the difficulty in assessing the effects of the holiday itself. Eckhardt again commented, "Increased food intake and relaxation of salt restrictions for cardiacs on festive days like Thanksgiving, I am sure exact their toll. . . . Psychiatrists tell us that suicides increase over holidays, and the National Safety Council tells us that highway deaths go up with four-day weekends" [7; p. 837]. Ideally, all these factors should be accounted for; however, data enabling such detailed analysis is seldom available.

Multivariate statistical analysis offers one approach in coping with the possible effects of a number of "relevant" variables on mortality. It is known that climatological characteristics affect mortality [3, 4, 31], air pollution affects climate [11], and air pollutants and climate may have synergistic effects [25]. A multivariate statistical model can simultaneously consider these interrelationships.

Such a study on short-term effects was undertaken by Hodgson [21]. Using multiple regression, he analyzed mortality, air pollution, and certain meteorological factors between 1962 and 1965 in New York City. Hodgson addressed the question: "How is air pollution during [time] t influencing short-run future mortality" [21; p. 590]? He found that mortality from respiratory and heart diseases for all ages was significantly related to the level of air pollution. Mortality from other causes was not so related. Furthermore, the effects of the environmental factors on mortality occurred on the day of increased pollution and extreme temperature with lesser effects on the day following. Finally, he concluded from his results that the increase in mortality observed was not merely the bunching together of deaths of persons already ill.

Some questions arise with regard to Hodgson's analysis. Although the author professes concern with day-to-day variations in the health effects of pollution, most of his analysis is based on monthly averages. Detailed analysis of the acute reaction is lost in this aggregation. In addition, aggregating the daily data into months enhances the multicollinearity problems among the explanatory variables (especially when lagged periods are included) and so Hodgson utilized moving averages of the

monthly means. As pointed out by Hodgson, this sacrifices much of the short-term effects. In defense, he states that the index will then reflect the cumulative effect of air pollution. It is true that the moving averages will represent the mean pollution levels for the number of months included, but they will neither reflect longer-term effects nor day-to-day effects.

Hodgson does analyze daily observations at the end of his paper, but again he uses moving averages and the interesting lag effects are masked. The daily results are not found to be significantly different from the monthly regressions. A natural hypothesis with such a study is that air pollution merely hastens death of those who would have died shortly. Hodgson remarks, ". . . if the sole effect of air pollution was to redistribute the deaths within a short time interval, for example, a month or less, then the average daily mortality for a given month would be independent of the concentration of air pollution during the month and no statistically significant relation would be observed between monthly mortality and air pollution" [21; p. 593]. The assertion is incorrect since monthly averages would only mask such an effect, not eliminate it. At both the beginning and end of the month there would be days where one or two-day shifts in deaths would still be present. To examine the redistribution question a more complicated lag structure is needed as well as a comparison with cross-section data.

Another recent study employing regression analysis was conducted by Hexter and Goldsmith [18]. Daily mortality in Los Angeles County was related to carbon monoxide and oxidant concentrations and maximum temperature for a four-year period (1962–1965) via cross-spectral analysis. A significant association between mortality and carbon monoxide pollution was found, while no association was demonstrated between mortality and oxidant.[4]

Although Hexter and Goldsmith included maximum temperature as a factor likely to be important in explaining mortality, they did not adequately take account of other possibly relevant factors. This was emphasized in a subsequent letter to the editor. Mosher, et al. [30] found that for the period corresponding to the Hexter and Goldsmith study the

4. They made no attempt to resolve the conflict with the Hechter and Goldsmith study [17] which found that daily Los Angeles deaths from 1956 to 1958 were not correlated with daily oxidant and carbon monoxide concentrations. Many investigators have detected an association with daily data. Mills [28] found a significant association between Los Angeles smog and day-by-day respiratory and cardiac deaths, as did Brant and Hill [6] and Brant [5] when they examined the effects of oxidants on weekly hospital admissions in Los Angeles for respiratory and cardiovascular dysfunction.

correlations between carbon monoxide and nitric oxide or nitrogen dioxide were significant at a level of more than 99.9 per cent. Thus, one cannot isolate which of the three pollutants is the "true cause" of illness. Ellsaesser [8] questioned the ad hoc procedure employed by Hexter and Goldsmith. He argued that their method combined with their failure to take account of other important factors made the results suspect. For instance, the population (and therefore the population at risk) increased over time. As another example, Ellsaesser discussed the interrelationships among carbon monoxide, traffic fatalities and mortality.

Hexter and Goldsmith [19] replied that further analysis showed that changes in population structure were not affecting their previous results. They also found a negative association between CO and traffic fatalities when a death-specific regression was analyzed.

Glasser and Greenburg [13] examined daily deaths in New York City between the years 1960–1964 (excluding April to September[5]). They attempted to explain deviations in daily deaths from a five-year "normal" by air pollution measures (SO_2 and smokeshade), and climatological variables (temperature deviation from normal, wind speed, sky cover, and rainfall).[6] Using descriptive statistics, cross tabulations, and regression analyses, they found a relationship between daily mortality and air pollution (primarily as measured by SO_2).[7]

One must be curious as to the omission of six months from each year. It is true that seasonal factors and air pollution might be confounded during the summer months, however, regression analyses, which include both the relevant weather variables and pollution variables should sort out the individual effects. If the two semi-annual periods differ in some respects, two models can be developed and these can be tested in order to determine whether the underlying structures are similar. Analysis of this time of year is of particular interest in view of the number of inversions and high pollution episodes occurring during the summertime.

Glasser and Greenburg measured mortality in terms of deviations from a 15-day moving average. The measure is suggestive of cycles in the data

5. These months were omitted ". . . because of the generally low levels of air pollution and generally high temperatures during these months. It was believed that including such data would complicate the analysis since seasonal factors would be confounded with air pollution factors" [13; p. 336].

6. The "normal" measure was defined as the mean for corresponding days in each year of a 15-day moving average centered on the 8th day.

7. After an examination of the individual effects of SO_2 and smokeshade, it was concluded that the exclusion of smokeshade as a measure of air pollution would not substantially alter the results.

and of lags in the effects of pollution, however, no explanations are presented. Lagged variables (particularly pollution) are not utilized, hence the investigation never questions the timing of the effect.

Data

As with previous work in the field, the present study is limited by the availability of pollution data and to a lesser extent mortality data. Daily observations on pollution levels (24-hour averages) were obtained from the Continuous Air Monitoring Program (CAMP). There were seven cities and eleven pollutants for which data was collected; however, the series are far from complete. In fact, a complete series for a given city during a single calendar year for one pollutant was nonexistent. Thus, we restricted ourselves to the most complete set of series available.[8] The cities remaining under consideration were Chicago, Denver, Philadelphia, St. Louis, and Washington, D.C. The pollutants remaining were carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and hydrocarbons (HC).

Corresponding to the daily pollution readings, we obtained daily death counts for the five cities from 1962 to 1966 from a special study by the Environmental Protection Agency. These deaths were classified into 47 different causes ranging from tuberculosis to motor vehicle accidents; however, in this paper we limit ourselves to analyzing total deaths.

Finally, we secured climatological data on daily weather factors for the cities from the Department of Commerce. Air pollution, mortality, and climate information constituted our data base. We selected periods when the air pollution data were complete for analysis. (See table 1 for a list of the individual data sets.)

The Model

According to Hodgson, "Nobody knows how air pollution causes death and, unfortunately, controlled scientific enquiry into the nature and mechanism of the deaths has still to be designed" [20; p. 15]. Our own literature review concurs with Hodgson that there is little theory available to specify the functional form of the relationship. There is no alter-

8. Those daily observations which were missing were estimated by simple interpolation.

TABLE 1

Data Sets With Means and Standard Deviations of Pollutants and Deaths

Dates		Pollutants[a] Mean	Pollutants[a] Standard Deviation	Deaths Mean	Deaths Standard Deviation
		Chicago			
8/63–8/64	NO	0.090	0.038	114.62	13.99
8/63–8/64	NO_2	0.045	0.016	114.62	13.99
9/62–6/63	SO_2	0.167	0.131	121.05	16.94
9/63–6/64	SO_2	0.185	0.146	116.37	13.10
8/63–5/64	HC	3.005	0.819	114.80	13.59
9/63–5/64[b]	NO	0.099	0.040	116.39	12.96
	NO_2	0.041	0.012	116.39	12.96
	SO_2	0.198	0.147	116.39	12.96
	HC	2.992	0.779	116.39	12.96
		Denver			
3/65–7/66	NO_2	0.034	0.010	25.17	5.58
3/65–7/66	SO_2	0.017	0.009	25.17	5.58
4/65–7/66	HC	2.377	0.704	25.18	5.67
		Philadelphia			
9/65–10/66	CO	7.256	2.397	25.21	5.38
9/64–4/65	NO	0.056	0.043	24.85	5.06
9/65–9/66	NO_2	0.037	0.014	25.27	5.41
		St. Louis			
4/64–2/66	CO	6.501	2.526	14.18	4.04
9/64–4/65	NO	0.039	0.024	14.68	3.99
9/64–6/66	NO_2	0.028	0.011	14.32	4.07
1/65–6/66	SO_2	0.048	0.028	14.38	4.13
		Washington, D.C.			
4/65–10/66	CO	3.397	1.532	27.93	9.88
11/64–7/65	HC	2.397	0.758	27.15	5.78

[a] Figures for pollution in parts per million (ppm).
[b] Nine month Chicago data set (see text).

native to investigating a number of possible specifications and to performing a sensitivity analysis.

A function explaining mortality can be written as in equation 1:

$$MR = MR(G, H, SE, P, C, e), \tag{1}$$

where G represents genetic characteristics, H represents personal habits

(e.g., exercise, eating), *SE* represents socioeconomic characteristics (e.g., income, age, race, occupation, medical care, etc.), *P* represents environmental pollution, *C* represents climatological characteristics, and *e* is all other factors. Some of these factors are difficult to measure (e.g., genetic factors and nutritional history), while data on other factors have not been collected (e.g., smoking habits).

From day to day, within a single city, we further hypothesize that *G*, *H*, and *SE* remain essentially constant. Thus, for examining day-to-day changes in mortality, the relevant function is:

$$MR = MR'(P, C, e), \tag{2}$$

where these variables are defined above.

We expect mortality at time t to be associated with current pollution (at time t) and pollution levels on immediately preceding days as well as with a number of meteorological variables. The coefficients of the lagged pollution variables should shed light on the short-term effects of pollution.

More precisely, we will be considering models with finite lags which are linear, multiplicative, or quadratic (including interaction terms). One such model is shown as equation 3:

$$\begin{aligned} MR_t = {} & 30.056 + \underset{(1.16)}{0.046}\, SO_{2_t} - \underset{(-1.19)}{0.051}\, SO_{2_{t-1}} - \underset{(-.88)}{0.038}\, SO_{2_{t-2}} + \underset{(.80)}{0.035}\, SO_{2_{t-3}} \\ & - \underset{(-.48)}{0.021}\, SO_{2_{t-4}} - \underset{(-.14)}{0.006}\, SO_{2_{t-5}} - \underset{(-.14)}{0.001}\, \text{Wind} - \underset{(-.68)}{0.009}\, \text{Rain} \\ & - \underset{(-3.81)}{0.057}\, \text{Mean } T - \underset{(-1.72)}{1.585}\, \text{Sun.} - \underset{(-1.18)}{1.088}\, \text{Mon.} - \underset{(-1.81)}{1.676}\, \text{Tues.} \\ & - \underset{(-1.61)}{1.504}\, \text{Thurs.} - \underset{(-1.60)}{1.481}\, \text{Fri.} - \underset{(-.78)}{0.725}\, \text{Sat.} \end{aligned} \tag{3}$$

This model and an explanation of the variables are discussed below.

Method and Results

Having little a priori knowledge of the underlying relationships among mortality, air pollution, and weather factors, we initially fit models such as equation 3 to Chicago, Denver, Philadelphia, St. Louis, and Washington, D. C. using ordinary least squares. In each case we regressed the daily deaths on the level of pollution at time t, at time t − 1, and so on up to time t minus five days. We also included a number of weather factors which were similar to those found in the recent study by Glasser

and Greenburg [13] cited above. These consisted of average wind speed (Wind), rainfall (Rain), and mean temperature (Mean T). In addition, we included a set of 1–0 dummy variables for the days of the week, since it was hypothesized that both pollution and mortality are cyclic over the week.[9]

Results are shown in tables 2 through 4 and the first regression in table 2 is reproduced in equation 3 above. The daily number of people dying in Denver from March 1965 through July 1966, is regressed on air pollution (SO_2), climate variables, and day-of-the-week variables and 5.6 per cent of the variance is explained. The t statistics are shown in parentheses below the estimated coefficients (see table 2, footnote 1). The coefficients of the SO_2 variables suggest that there is a contemporaneous effect of high levels on the number of deaths, but the association is not statistically significant. When high levels of SO_2 occur, one, two, four, and five days prior to the day of observed deaths, there is a negative correlation with deaths, although none of these effects is significant. Of the climate variables, only mean temperature is significant (indicating that as the mean temperature increases, the number of deaths decreases). The greatest numbers of deaths seem to occur on Wednesdays and the smallest numbers on Tuesdays.

Results are similar for four of the five cities and the five pollutants in that no consistent pattern of statistically significant coefficients can be discerned, and hence there is no evidence of an effect of air pollution on daily mortality. The only exception is Chicago, especially for the pollutant SO_2.* In St. Louis one coefficient is positive and significant, although the sum of pollution coefficients indicates that the effect was negligible (see the second regression of table 2). For Chicago, three coefficients are significant (and all positive), indicating a close association between daily deaths and SO_2, and the sum of pollution coefficients indicates a strong effect (see the third regression of table 2).[10] Some climate variables are significant in each city. Average wind speed is significant only for St. Louis, indicating a negative relationship with daily mortality. This is plausible since wind cleanses the air.[11] Rainfall never attains

9. It was hypothesized that pollution would be lower on weekends due to decreased industrial activity. Silverman [33] has detected a definite day-of-the-week influence on emergency hospital admissions. We felt this effect might also carry over to daily deaths.

10. A useful way in which to view the magnitude of this association for Chicago is to consider a 50 per cent reduction in the mean level of SO_2 pollution. A decrease of .092 ppm in the mean values of the pollution variables is estimated to lead to a decrease of 6.3 deaths per day (approximately 5.4 per cent).

11. The simple correlation between Wind and SO_{2t} is −.30.

* We have already discussed the results of using the sulfur dioxide measure in Denver.

TABLE 2

Regression Analysis Comparing Daily Deaths in Denver, St Louis, and Chicago

Variable	Denver, 3/65–7/66		St. Louis, 1/65–6/66		Chicago, 9/63–6/64	
	Coefficients[a]	Means[b]	Coefficients[a]	Means[b]	Coefficients[a]	Means[b]
Dependent variable		25.17		14.38		116.37
Independent variables:						
Constant	30.056		16.849		93.268	
(ppm × 1,000)						
SO_{2_t}	0.046		−0.005		0.026	
	(1.16)	17.45	(−0.70)	47.54	(2.90)	184.79
$SO_{2_{t-1}}$	−0.051		0.003		0.004	
	(−1.19)	17.44	(0.44)	47.56	(0.39)	184.24
$SO_{2_{t-2}}$	−0.038		0.017		0.028	
	(−0.88)	17.44	(2.33)	47.51	(2.87)	184.00
$SO_{2_{t-3}}$	0.035		−0.008		−0.007	
	(0.80)	17.43	(−1.18)	47.49	(−0.71)	183.95
$SO_{2_{t-4}}$	−0.021		−0.004		0.028	
	(−0.48)	17.43	(−0.51)	47.42	(2.88)	184.02
$SO_{2_{t-5}}$	−0.006		−0.002		−0.011	
	(−0.14)	17.40	(−0.30)	47.41	(−1.38)	183.93
Wind (mph × 10)	−0.001		−0.011		−0.004	
	(−0.14)	78.97	(−1.90)	98.49	(−0.22)	113.03
Rain (in. × 100)	−0.009		0.003		0.054	
	(−0.68)	4.96	(0.38)	8.53	(1.25)	6.03
Mean T	−0.057		−0.036		0.182	
	(−3.81)	51.99	(−3.92)	53.63	(3.38)	47.59
Sunday	−1.585		1.173		2.868	
	(−1.72)	0.14	(1.74)	0.14	(1.11)	0.14
Monday	−1.088		0.486		5.107	
	(−1.18)	0.14	(0.73)	0.14	(1.98)	0.14
Tuesday	−1.676		0.291		4.329	
	(−1.81)	0.14	(0.44)	0.14	(1.70)	0.14
Thursday	−1.504		0.262		0.388	
	(−1.61)	0.14	(0.39)	0.14	(0.15)	0.14
Friday	−1.481		0.724		−1.451	
	(−1.60)	0.14	(1.08)	0.14	(−0.56)	0.14
Saturday	−0.725		0.767		3.628	
	(−0.78)	0.14	(1.14)	0.14	(1.39)	0.14
R^2	.056		.050		.266	

[a] Figures in parentheses are t statistics. A value of 1.65 indicates significance at the .05 level using a one-tailed test on a sample of infinite size.

[b] For computational ease, the means of some variables were multiplied by 10, 100, or 1,000 as indicated next to the relevant variables. This applies to all subsequent tables.

statistical significance. Mean temperature is strongly negative for Denver and St. Louis, whereas it is positive for Chicago (a result which needs further investigation). While many of the day-of-the-week variables are statistically significant, it is difficult to explain the different patterns in the three cities.[12] Finally, the coefficient of determination (R^2) is quite small for Denver and St. Louis (.056 and .050, respectively), while for Chicago almost 27 per cent of the variation in daily deaths is explained by the variables.

A natural question arises as to why the Chicago results differ from the results for the other cities. One possibility concerns the unusually high levels of pollution prevalent in Chicago. For each pollutant under consideration, the mean level in Chicago exceeds the mean value for any other city (see table 1). This is especially noteworthy for SO_2 where Chicago's mean levels are almost ten times the levels of Denver and almost four times the levels of St. Louis. Thus, it may be that at levels of pollution substantially below those found in Chicago, acute effects are not important. A closely related issue involves the relative size of the cities. Because of its larger population, the mean number of deaths per day in Chicago is more than four times as large as any other city in our sample (see table 1). Since deaths occur in discrete units (one at a time), effects of pollution in a smaller city may be lost when scattered over a five-day period.[13] We shall examine these conjectures in more detail below.

For whatever reason, only Chicago regressions showed a significant relationship between daily mortality and air pollution. We have attempted to investigate why this occurred and to examine the Chicago relationship more intensively. Much of the remaining estimation is carried out on a nine-month data set (denoted by superscript b in table 1) which contains measures of four pollutants.

Results for Chicago

In addition to the explanations given above that Chicago has much greater levels of pollution and is a large enough city to be able to observe

12. The day-of-the-week variables were not found to be correlated with the pollution levels.

13. Close associations between daily mortality and air pollution have been found for New York City and Los Angeles, both of which are large cities with high levels of air pollution [13, 17].

the effects of pollution more easily, there is the possibility that we have misspecified the relation. We have examined this possibility in a number of ways. First we attempted to see if the structure was the same for months with low pollution as it was for months with high pollution. We divided the two Chicago data sets with SO_2 as the pollutant on the basis of the high and low pollution levels during the year. F ratios were computed to test the hypothesis that the structures dicered between high and low pollution periods. The F ratios were 1.25 and 1.32 for the two data sets. Since $F_{.05}(30\ \infty) = 1.46$, one cannot reject the null hypothesis that the coefficients are identical. While this test is supportive of the notion that the relationship was correctly specified, one can object that it is not relevant to the comparison between Chicago and the other cities, since even the low Chicago months had pollution levels far greater than those of the other cities.

Second, we attempted to fit a relation which would be more sensitive to different effects of pollution according to the level of pollution. We divided the pollution variable into five new variables, each of which represented a different range of pollution levels.[14] These "piecewise" linear variables can approximate much more complicated functional forms. For both data sets, the coefficients for the smallest two variables were insignificant, both statistically and numerically. This is evidence that daily effects of relatively low levels of air pollution are not important.

Third, we fit other functional forms, quadratic and log-linear and included interaction terms (e.g., HC $\times$ NO). The coefficients of the quadratic variables were not statistically significant and added little to the explanatory power of the regression. The log-linear specification exhibited results similar to the linear specification; the elasticities at the mean were quite close for the pollution variables. Of the interaction variables, only HC $\times$ SO_2 attained statistical significance.

Fourth, we included all four pollutants simultaneously, as shown in the first regression of table 3. Only NO and SO_2 possessed statistically significant coefficients (both were positive). The coefficients for the weather variables were quite similar to those in the previous Chicago regression in table 2.[15] The day-of-the-week dummy variables were almost identical in magnitude and significance to those reported in table 2. It is noteworthy that almost 36 per cent of the variation in daily deaths was explained by the variables.

14. In this formulation, piecewise lags were excluded.
15. The sign of "Wind" changed, but was quite insignificant.

TABLE 3

Three Alternative Specifications Analyzing Daily Deaths in Chicago

	Original		*Deaths as Deviations from a Moving Average*		*Episodic Pollution*	
Variable	*Coefficients*[a]	*Means*	*Coefficients*[a]	*Means*	*Coefficients*[a]	*Means*
Dependent variable		116.39		0.14		116.39
Independent variables:						
Constant	96.591		−19.181		100.038	
HC_t	0.012 (0.72)	299.22				
HC_{t-1}	−0.001 (−0.05)	298.91				
HC_{t-2}	0.003 (0.14)	299.16				
HC_{t-3}	−0.015 (−0.74)	299.10				
HC_{t-4}	0.009 (0.45)	298.49				
HC_{t-5}	−0.006 (−0.36)	297.90				
NO_t	−0.027 (−0.86)	99.24	−0.032 (−1.16)	100.25	−0.025 (−0.82)	99.24
NO_{t-1}	0.007 (0.23)	99.33	0.007 (0.27)	100.38	−0.013 (−0.47)	99.33
NO_{t-2}	0.029 (0.90)	99.41	0.024 (0.96)	100.61	0.027 (0.96)	99.41
NO_{t-3}	0.058 (1.82)	99.45	0.027 (1.07)	100.62	0.036 (1.49)	99.45
NO_{t-4}	−0.007 (−0.23)	99.24	−0.016 (−0.63)	100.48		
NO_{t-5}	0.046 (1.68)	99.14	0.015 (0.66)	100.56		
NO_{2t}	0.053 (0.58)	40.91				
NO_{2t-1}	−1.28 (−1.22)	40.94				
NO_{2t-2}	0.010 (0.10)	40.96				
NO_{2t-3}	−0.155 (−1.49)	40.94				
NO_{2t-4}	0.009 (0.08)	40.94				

(*continued*)

TABLE 3 (*concluded*)

Variable	Original		Deaths as Deviations from a Moving Average		Episodic Pollution	
	Coefficients[a]	*Means*	*Coefficients*[a]	*Means*	*Coefficients*[a]	*Means*
$NO_{2_{t-5}}$	−0.079					
	(−0.94)	40.92				
SO_{2_t}	0.022		0.019	205.07	0.023	198.26
	(2.31)	198.26	(2.15)		(2.45)	
$SO_{2_{t-1}}$	0.007		−0.002	204.85	0.003	198.12
	(0.66)	198.12	(−0.25)		(0.29)	
$SO_{2_{t-2}}$	0.020		0.019	204.60	0.017	198.04
	(1.92)	198.04	(1.95)		(1.62)	
$SO_{2_{t-3}}$	−0.006		−0.014	203.88	0.001	198.02
	(−0.61)	198.02	(−1.44)		(0.10)	
$SO_{2_{t-4}}$	0.023		0.026	203.81		
	(2.16)	198.02	(2.72)			
$SO_{2_{t-5}}$	−0.014		−0.018	203.84		
	(−1.52)	198.03	(−2.16)			
$\left(\prod_{k=0}^{3} NO_{t-k}\right)$[b]					0.002	143.54
					(0.31)	
$\left(\prod_{k=0}^{3} SO_{2_{t-k}}\right)$[b]					0.0001	6904.47
					(1.01)	
Wind	0.002		−0.011	116.71	−0.008	115.03
	(0.07)	115.03	(−0.46)		(−0.34)	
Rain	0.067		0.049	6.25	0.061	5.98
	(1.48)	5.98	(1.16)		(1.36)	
Mean T	0.171		0.233	43.76	0.073	44.74
	(2.42)	44.74	(3.86)		(1.10)	
Sunday	3.125		1.189	0.14	3.551	0.14
	(1.14)	0.14	(0.46)		(1.32)	
Monday	5.112		4.366	0.14	2.358	0.14
	(1.87)	0.14	(1.67)		(0.87)	
Tuesday	3.659		3.577	0.14	4.135	0.14
	(1.38)	0.14	(1.40)		(1.52)	
Thursday	0.743		0.125	0.14	3.598	0.14
	(0.28)	0.14	(0.05)		(1.35)	
Friday	−1.252		−1.215	0.14	−0.284	0.14
	(−0.46)	0.14	(−0.46)		(−0.11)	
Saturday	3.282		2.951	0.14	−1.467	0.14
	(1.22)	0.14	(1.14)		(−0.54)	
R^2	.357		.164		.297	

[a] Figures in parentheses are t statistics.
[b] Multiplied by 10^{-6}.

Fifth, we estimated the specification of Glasser and Greenburg [13]. We altered the dependent variable, so that instead of examining deaths at time t in terms of (current and lagged) pollution and current weather factors, we examined deviations from a 15-day moving average centered on the eighth day.[16] Inasmuch as only NO and SO_2 were significant in explaining daily deaths in Chicago (results just cited), we limited ourselves to current and lagged values of these two pollutants in rerunning the previous regression. We again included the climatological and day-of-the-week variables. (This is presented as the second regression in table 3.)

As can be seen by comparing this regression with the previous one, there was little difference in the results. The NO variables were no longer statistically significant, although the magnitude and significance of the SO_2 variables remained essentially unchanged.[17] The significant weather variables and dummy variables exhibited little change when the "new" dependent variable was included. The explanatory power of the regression R^2 dropped from .357 to .164. Since the formulation had no theoretical justification and since it was less satisfactory empirically, we rejected it.

It is difficult to compare these results with those of Glasser and Greenburg since they included neither lagged pollution nor day-of-the-week variables.[18] Of their meteorological variables, wind speed displayed mixed signs when it was statistically significant (as did ours, although it was not significant) and rainfall exhibited a positive coefficient when significant (as did ours). Because their temperature variable was expressed as a deviation from normal, we cannot easily compare it with our mean temperature measure.

Sixth, we investigated whether a consecutive period of several days of high pollution was more important than isolated days of high pollution. Many of the episodic studies have found significant effects on mortality during such occurrences [12, 16, 27]. We defined a new SO_2 variable as the product of the SO_2 levels for the current day and the three preceding days.[19] We also defined a similar NO variable.

16. In notational terms, the left-hand side of equation 3 becomes

$$MR_t - \sum_{k=-7}^{7} \frac{MR_{t+k}}{15}.$$

17. The coefficient of SO_{2t-5} was negative and quite significant. This will be discussed below.

18. They did disaggregate their data by day of the week and run separate regressions.

19. In notational terms, $P_t = \prod_{k=0}^{3} P_{t-k}$.

The results of adding these new pollution variables are presented as the third regression in table 3. Comparing the first and third regressions, the magnitude and significance of the separate pollution variables exhibited little change. The new pollution variables were statistically unimportant.[20] In examining the weather variables one finds that mean temperature lost significance while precipitation gained significance. "Wind" remained unimportant. The explanatory power of this regression was .297.

Seventh, we tried other specifications, including moving averages and variables measuring the change in pollution levels on consecutive days. None proved as satisfactory as the simple linear form in terms either of explanatory power or the significance of the coefficients.

Eighth, we tried a final test to judge the linearity of the specification. Both cross-section and time-series data should reflect the same underlying structure. Using measures of particulates and sulfates, we had previously determined that a linear specification fit cross-section data as well as other alternatives. We used these cross-section data to estimate a piecewise linear specification similar to that estimated above for the time-series data. To make the comparison with the daily results, we reran the regression including mean sulfates as the only measure of air pollution since it is related to measures of SO_2. The results were quite similar to the piecewise linear form of the daily data.

Ninth, having settled on the linear form with five days of lags, we attempted to refine the estimates by using the Almon technique, as programmed and discussed by Gaver [10]. This procedure imposes structure on the lag coefficients by constraining them to fit a polynomial curve of a specified degree. The process often results in the reduction of large standard errors in the distributed lag coefficients which may arise from multicollinearity in the lagged values of the pollution variables.[21] The technique allows considerable flexibility; we began by fitting second and third-degree polynomials, using lags of five to ten days. The third-degree polynomial added little to the analysis so we confined our interest to the second-degree polynomial. The only other qualification we imposed was that the current pollution coefficient be positive inasmuch as a negative coefficient was deemed unreasonable.[22] In this case the effect of such a

20. We experimented with other variables, but were unable to detect an episodic effect.

21. For example, the simple correlations between NO_t and NO_{t-1} and between SO_{2t} and SO_{2t-1} were .59 and .78, respectively.

22. In practice, this qualification was unnecessary for SO_2; however initially the zero lag of NO had a negative (although statistically insignificant) coefficient. Thus, only the latter variable had such a restriction.

front-end restriction was simply to shift the polynomial so that the t + 1 lag (one future day) had a zero weight.

In table 4 the results of a regression are shown which employed the

TABLE 4

***Regression of Refined Lag Structures of* NO *and* SO_2**

Independent Unlagged Variables	*Coefficients*[a]	*Means*[b]	*Independent Lagged Variables*	*Weight*[a]	*Independent Lagged Variables*	*Weight*[a]
Constant	85.58		NO_t	0.007	SO_{2_t}	0.017
				(2.53)		(2.94)
Wind	0.013		NO_{t-1}	0.013	$SO_{2_{t-1}}$	0.015
	(0.65)	114.55		(2.61)		(5.39)
Rain	0.053		NO_{t-2}	0.016	$SO_{2_{t-2}}$	0.012
	(1.22)	6.34		(2.71)		(4.33)
Mean T	.154		NO_{t-3}	0.018	$SO_{2_{t-3}}$	0.009
	(2.66)	45.17		(2.85)		(2.72)
Sunday	2.409		NO_{t-4}	0.017	$SO_{2_{t-4}}$	0.005
	(.91)	.14		(3.00)		(1.84)
Monday	4.544		NO_{t-5}	0.015	$SO_{2_{t-5}}$	0.001
	(1.73)	.14		(2.94)		(.32)
Tuesday	3.391		NO_{t-6}	0.010	$SO_{2_{t-6}}$	−0.004
	(1.31)	.14		(1.98)		(−.73)
Thursday	0.177		NO_{t-7}	0.004		
	(.07)	.14		(.51)		
Friday	−1.537		NO_{t-8}	−0.005		
	(−.58)	.14		(−.39)		
Saturday	3.378					
	(1.27)	.14				
$R^2 = .291$						

[a] The figures in parentheses are t statistics.
[b] Mean of dependent variable is 116.18.

Almon technique using second-degree polynomials in estimating the lag structure of the two pollutants.[23] The climate variables and the day-of-the-week variables are essentially unchanged from the previous results. The magnitude and significance of the pollution variables display inter-

23. This particular regression was chosen because it had the minimum estimated standard error of disturbances adjusted for the degrees of freedom.

esting characteristics. The maximum effect of NO on daily deaths occurs at a lag of three days. This could have implications as to the physiological mechanism involved. The significant SO_2 weights (t to $t - 4$)[24] imply a different physical response. The maximum effect on deaths takes place simultaneously with peaks in SO_2 pollution levels and then drops off. (Figures 1 and 2 illustrate the two underlying structures.)

Figure 1

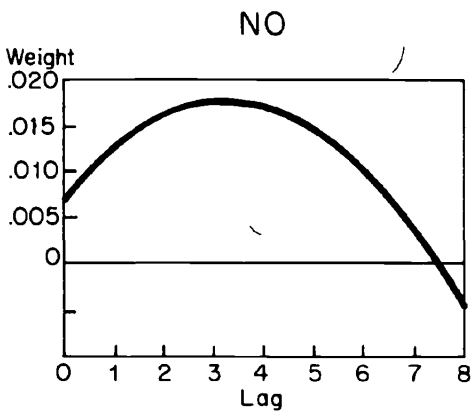
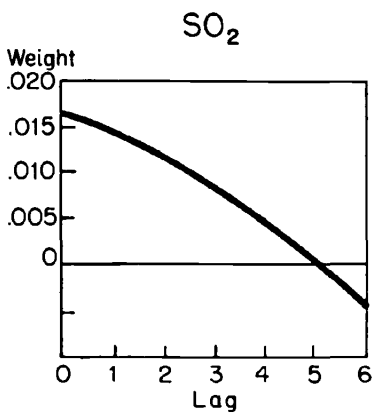


Figure 2



Are Deaths Merely Shifted by a Few Days?

One of the persistent questions is whether the estimated effect of air pollution on health has any policy significance. Surely an individual would have to be extraordinarily ill (or marginally viable) if an increase in air pollution results in death. Surely such an individual would die within a short time anyway, and so the air pollution episode can have no effect on the annual death rate, but only act to reallocate deaths within a short interval. If so, abating the air pollution episodes would do no more than add a few days of life to the few individuals dying and would have no policy significance.

Note that the statistically significant association between daily pollution and daily mortality is not proof that such an assertion is incorrect.

24. The fact that SO_{2t-5} was no longer important probably indicates that its previous significance was an artifact of multicollinearity among the lagged variables.

The pollution might only be reallocating deaths within a few days. One test of the assertion would be to investigate a long enough lag structure to be able to identify decreases in mortality that might follow a few days after an air pollution episode. Another test would be to dampen short-term effects by considering longer intervals (such as the monthly periods used by Hodgson). A more powerful test would be a comparison of time-series results within a city with cross-section results among cities.

A number of models have been estimated above using lags of five to ten days. None of these results give any indication of a decrease in deaths following a period of high pollution. While this test is not conclusive (since the lag may not have been sufficiently long), it does tend to indicate that the reallocation (if it exists) is not over a period as short as ten days.

Cross-section data have been analyzed using annual data for 117 cities in 1960 [24]. Since the studies relate annual mortality rates to annual air pollution levels (holding other factors constant), the effects are much longer than day-to-day. That analysis estimated that a 50 per cent abatement in air pollution (sulfates and particulates) would lead to a 4.5 per cent decrease in the mortality rate.[25] It is interesting to compare this estimate with the results of the daily analysis; for daily regressions, a 50 per cent abatement in air pollution is estimated to lead to a 5.4 per cent reduction in the total mortality rate. The closeness of these two estimates indicates that the estimated effects from the daily time-series are similar to those from the cross-section. Thus we conclude that the two effects are similar and that the increase in deaths when air pollution levels rise is not merely a shifting of these deaths by a few days.

Summary and Conclusions

In this paper we have taken up the question of whether daily mortality is affected by daily air pollution levels. Our method of investigation has consisted of examining some of the relevant literature on the subject as well as performing our own statistical analysis. Taken together, we believe that the evidence supports the conclusion that in large cities there are significant effects of air pollution on the daily death rate.

Our statistical analysis examined the effect of five pollutants in five cities. In only one of the five cities investigated, Chicago, was a significant relation displayed between pollutants and daily mortality when climate

25. See footnote 10.

and weekly effects where controlled. We conjectured that the effect was due to the relatively large number of daily deaths and to the high pollution level. We also investigated a number of other formulations, but found the simple linear one to be at least as good as any alternative we tried.

Only two pollutants, SO_2 and NO, were statistically significant, with SO_2 much more important. Careful investigation of the lag structure revealed that high SO_2 levels had a large immediate effect that gradually diminished over time. For NO, the effect was delayed by several days, reached its peak, and then diminished.

In examining the question of whether the deaths associated with pollution are merely displaced by a few days, we found our estimate corresponded closely to those from cross-section data. Thus we concluded that displacing deaths by a few days is not an important consideration in our estimate.

These results have a number of policy implications if one accepts the association we estimated as a causal one. The daily results strengthen this conclusion and indicate that the short-term effects are important. Authorities should continue to take steps to abate pollution during episodes. But the linear form also indicates that they should worry about air pollution during "clean" periods as well.

References

1. Ashe, W. F., "Exposure to High Concentrations of Air Pollution (1)—Health Effects of Acute Episodes," Proceedings of the National Conference on Air Pollution, Washington, D. C., *Public Health Service Publication No. 654*, Nov. 18–20, 1958, p. 188 (1959).
2. Becker, W. H., et al., "The Effect on Health of the 1966 Eastern Seaboard Air Pollution Episode," *Archives of Environmental Health* 16 (1968): 414.
3. Bokonjić, R., and N. Zec, "Strokes and the Weather: A Quantitative Statistical Study," *Journal of Neurological Sciences* 6 (1968): 483.
4. Boyd, J. T., "Climate, Air Pollution, and Mortality," *British Journal of Preventative and Social Medicine* 14 (1960): 123.
5. Brant, J. W. A., "Human Cardiovascular Diseases and Atmospheric Air Pollution in Los Angeles, California," *International Journal of Air Pollution* 9 (1965): 219.
6. ——— and S. R. G. Hill, "Human Respiratory Diseases and Atmospheric Air Pollution in Los Angeles, California," *International Journal of Air Pollution* 8 (1964): 259.

7. Eckhardt, R. B., "Variations Affecting Death Rate," *Archives of Environmental Health* (Letters) 17 (1968): 837.
8. Ellsaesser, H. W., "Air Pollution in Los Angeles," *Science* (Letters) 173 (1971): 576.
9. Firket, M., "Sur les causes des accidents servenus dans la valleé de la Meuse, lors des brouillards de décembre 1930," *Bulletin de Académie Royale Medicale de Belgique* 11 (1931): 683.
10. Gaver, K. M., "A Note on the Polynomial Technique for Estimating Distributed Lags," unpublished (1971).
11. Georgii, H., "The Effects of Air Pollution on Urban Climates," *World Meteorological Organization Symposium on Urban Climates and Building Climatology, Brussels* (Oct. 1968).
12. Glasser, M., et al., "Mortality and Morbidity during a Period of High Levels of Air Pollution," *Archives of Environmental Health* 15 (1967): 684.
13. Glasser, M., and Greenburg, L., "Air Pollution, Mortality, and Weather," *Archives of Environmental Health* 22 (1971): 334.
14. Greenburg, L., et al., "Air Pollution and Morbidity in New York City," *Journal of the American Medical Association* 182 (1962): 161.
15. ———, "Intermittent Air Pollution Episode in New York City, 1962," *Public Health Reports 78* (1963): 1061.
16. ———, "Report of an Air Pollution Incident in New York City, November 1953," *Public Health Reports 77* (1962) p. 7.
17. Hechter, H. H. and J. R. Goldsmith, "Air Pollution and Daily Mortality" *American Journal of Medical Sciences* 241 (1961): 581.
18. Hexter, A. C. and J. R. Goldsmith, "Carbon Monoxide: Association of Community Air Pollution with Mortality," *Science* 172 (1971): 265.
19. ———, "Air Pollution in Los Angeles." *Science* (Letters) 173 (1971): 576.
20. Hodgson, T. A., Jr., "The Effect of Air Pollution on Mortality in New York City, A Statistical Study," unpublished.
21. Hodgson, T. A., Jr., "Short-Term Effects of Air Pollution on Mortality in New York City," *Environmental Science and Technology* 4 (1970): 589.
22. Lave, L. B. and E. P. Seskin, "Air Pollution and Human Health," *Science* 169 (1970): 723.
23. ———, "Health and Air Pollution: The Effects of Occupation Mix," *Swedish Journal of Economics* 73 (1971): 76.
24. ———, "Air Pollution, Climate, and Home Heating: Their Effects on U. S. Mortality," *American Journal of Public Health* 62 (1972): 909.
25. Lewis, H. R., *With Every Breath You Take* (New York: Crown, 1965), p. 66–67, 181–182.
26. McCarroll, J., "Measurements of Morbidity and Mortality Related to Air Pollution," *Journal of Air Pollution Control* 17 (1967): 203.
27. McCarroll, J. and W. Bradley, "Excess Mortality as an Indicator of Health Effects of Air Pollution," *American Journal of Public Health* 56 (1966): 1933.
28. Mills, C. S., "Respiratory and Cardiac Deaths in Los Angeles Smogs" *American Journal of Medical Sciences* 233 (1957): 379.

29. *Mortality and Morbidity During the London Fog of December 1952,* Reports on Public Health and Related Subjects 95 (Her Majesty's Stationary Office: London Ministry of Health, 1954).
30. Mosher, J. C., et al., "Air Pollution in Los Angeles," *Science* (Letters) 173 (1971): 576.
31. Rogot, E. and W. C. Blackwelder, "Associations of Cardiovascular Mortality with Weather in Memphis, Tennessee," *Public Health Reports 85* (1970), p. 25.
32. Schrenk, H. H., Heimann, H., Clayton, G. D., et al., "Air Pollution in Donora, Pa.," *Public Health Service, Bulletin 306,* Federal Security Agency, Division of Industrial Hygiene (1949).
33. Silverman, L., "The Determinants of Total Emergency Admissions to Pittsburgh Hospitals," unpublished (1972).

COMMENT

Thomas A. Hodgson, Jr., Social Security Administration

Although I do not accept the Lave and Seskin critique of the literature without objections, I shall not raise them here, as they do not bear upon the authors' analysis. Lave and Seskin have made an important contribution in their paper. For the first time, to my knowledge, a lagged structure is identified for air pollutants that does not exhibit a hit and miss pattern of significant coefficients with reasonable signs; a pattern of many earlier studies which did not inspire confidence that the estimated relations revealed anything at all about the time-dependent structure relating mortality to air pollution. The smooth functional forms taken by the weights of nitric oxide and sulfur dioxide in the final Lave and Seskin model are encouraging evidence that they have uncovered at least a glimpse, if not more, of the underlying structure. The different paths taken by the two sets of weights may, in fact, result from different physiological responses to nitric oxide and sulfur dioxide, or, possibly, to other guilty agents for which nitric oxide and sulfur dioxide are indices. These initial results indicate it may be possible to describe the time-dependent physiological responses induced by individual air pollutants even when the biological mechanisms that are responsible remain a mystery. Further study on the dynamic aspects of the relation between exposure and response, both experimentally in the laboratory and epidemiologically, is called for. Increased knowledge of the time paths of responses may provide clues helpful in discovering the nature of the biological mechanism, which, in turn, will improve the specification of subsequent epidemiological and statistical models.

My reservations on the Lave and Seskin study concern specification of variables. They do not, however, raise serious questions regarding the principal results.

First, rainfall can reduce concentrations by washing pollutants from the atmosphere. I am not aware, however, of evidence of an independent effect of rainfall upon those categories of mortality generally considered responsive to air pollution which warrants including rainfall as an explanatory variable.

Second, there is substantial evidence that mortality is positively related to thermal stress. Thermal stress is relieved by loss of body heat through evaporative cooling and convection from the body.[1] The efficacy of these cooling processes to dispense body heat is proportional to the ventilation rate, which, in turn, depends on wind speed.[2] Moisture content of the air also affects the amount of body heat lost through evaporation, while solar radiation is a primary source of radiant heat gained by the body.[3] Since wind speed, moisture, and solar radiation all influence the effect of heat upon the body,[4] they are candidates for inclusion in the model. In practice, however, they often show up as insignificant explanatory variables. This may be because wind speed, humidity, and solar radiation exert their deleterious effects in conjunction with temperature, with the variation in mortality due to heat reflected in the temperature variable. A complicating factor in specifying climatic variables is the large amount of time spent indoors, in artificial environmental conditions that may not be reflected from day to day and place to place by ambient air temperature, moisture content, solar radiation, and wind speed.

The use of mean temperature, however, does not adequately take into account the detrimental effects of both extremely low and extremely high temperatures. While more research, at least epidemiologically, has examined the effect of extreme heat on mortality, it is known that coldness affects bodily functions.[5]

Thirdly, the day of the week in itself certainly does not affect mortality. If the correlation is not spurious, but a proxy for some other explanatory variable or variables, the question is, what are these variables? My inclination is not to include days of the week without a greater a priori justification than a previously observed statistical relation between emergency hospital admissions and day of the week unless the coefficients of other variables with a stronger a priori justification are severely and adversely affected by excluding days of the week. Emergency rooms in many hospitals are routinely used as a substitute for a family physician. These visits may be related to days of the week in a way that deaths are not.

Fourthly, lack of significant coefficients for those variables formed by

the product of four consecutive daily levels of a pollutant does not preclude the possibility of an episodic effect that is generated by only two, or three, consecutive days of high pollution. Including additional days can obscure the effect induced by a shorter period, depending on the serial correlation between consecutive levels of a pollutant. This is probably not the case in the Lave and Seskin analysis given the rather high serial correlations of nitric oxide and sulfur dioxide and the small t statistics of the episodic variables. On the other hand, attainment of statistically significant coefficients for these variables in the face of serial correlation of consecutive values would raise the question of which of those periods included are actually responsible for the observed effect. The episodic effect should be more systematically investigated, and perhaps the authors' footnote on having experimented with other variables without observing an episodic effect bears witness to their having done so.

Although available evidence supports the hypothesis that deaths are not merely shifted by a few days as a result of increased air pollution, I do not believe the Lave and Seskin argument is conclusive in itself. It is not clear to me whether the estimated reduction in mortality following a fifty per cent reduction in pollution according to the daily regressions is based upon all of the five cities considered by the authors, or just Chicago. In any case, we do not know if this small sample is representative of the larger set of 117 cities. Lave and Seskin have shown that the results obtained for Chicago cannot be replicated for Denver and St. Louis. Although every city contains individuals at risk of death when exposed to air pollution, the populations at risk differ among cities with respect to racial and ethnic composition, medical care, diet, income, and other factors affecting health status, including the nature of pollution which threatens them. Accordingly, caution is urged when extrapolating results of daily regressions in one or several cities to a large number of cities. Some hasty calculations (which can be redone to obtain more precise values) from my work on daily mortality and air pollution in New York City indicate a fifty per cent reduction in air pollution would have lead to approximately a nine per cent decline in total deaths and fifteen per cent fewer deaths from heart and respiratory diseases. The data in my analysis are from the period November 1962 through May 1965 and coincide with much of the data utilized by Lave and Seskin. Thus, if the differences between time-series and cross-section results are to be taken as a measure of deaths that were merely shifted a few days, the percentage is five times larger for New York City. The Lave and Seskin daily regressions do not closely reflect the New York City experience.

We should not be surprised, however, at differing figures among in-

dividual cities. Nevertheless, Lave and Seskin have raised an interesting point. Certainly, some deaths are merely shifted a few days, and more detailed comparison between time-series analyses within cities and a cross-section analysis among cities may help us understand this aspect of the nature and magnitude of the short-term response of mortality to air pollution.

I believe Lave and Seskin rightfully conclude "there are significant effects of air pollution on the daily death rate," and they have contributed new support for this contention. Numerous investigators have sought to examine the relationship between air pollution and human health. Diverse populations, time periods, indices of air pollution and other explanatory variables, and methods of analysis have been employed. It is difficult to make comparisons among some studies, and lack of a consistent approach has prevented the close replication of results, which has long been an essential ingredient in epidemiological proof of causation. On the other hand, the frequent indictment of air pollution indicates contaminated air is deleterious to human health. Our knowledge of this phenomenon is meager on many counts; for example, the biological mechanism through which community air pollution exerts its effect on morbidity and mortality and those components of the ambient air that are in fact guilty. Investigators have precious little data or a priori knowledge as a basis for seeking answers; but I believe epidemiologic study such as presented here is important, provides insight, and must be continued.

References

1. Clarke, John F., "Some Effects of the Urban Structure on Heat Mortality," *Environmental Research* 5 (1972): 97.
2. Clarke, "Effects of the Urban Structure," pp. 97, 99.
3. Clarke, "Effects of the Urban Structure," p. 98.
4. Clarke, John F., "Some Climatological Aspects of Heat Waves in the Contigous United States," *Environmental Research* 5 (1972): 84.
5. Hodgson, Thomas A., "Short-Term Effects of Air Pollution on Mortality in New York City," *Environmental Science and Technology* 4, no. 7 (July 1970): 589–597.