Estimating the Independent Effects of Multiple Pollutants in the Presence of Measurement Error: An application of a measurement error resistant technique

Ariana Zeka and Joel Schwartz

Estimating the Independent Effects of Multiple Pollutants in the Presence of Measurement Error: An application of a measurement error resistant technique

Ariana Zeka[1*], Joel Schwartz[1]

[1] Exposure, Epidemiology, and Risk Program, Environmental Health Department
Harvard School of Public Health, Boston, MA 02215, USA

* Correspondence to: Ariana Zeka
Exposure, Epidemiology, and Risk Program
Environmental Health Department
Harvard School of Public Health
401 Park Drive, Suite 415 W
P.O. Box 15698, MA 02215, USA
Tel: 1 617 998 1001
Fax: 1 617 384 8745
E-mail: azeka@hsph.harvard.edu

Running Title: Measurement Error Correction

Key Words: air pollution, daily-mortality, measurement error, particulate matter, carbon monoxide.

Abbreviations:
CO – Carbon monoxide
IHAPSS – Internet-based Health and Air Pollution Surveillance System
NMMAPS – National Mortality and Morbidity Air Pollution Study
NO_2 – Nitrogen dioxide
O_3 - Ozone
PM_{10} – Particulate matter of aerodynamic diameter less or equal to 10 $\mu m/m^3$
SO_2 – Sulfur dioxide

Text Word Count: (3,800)

Abstract Word Count: (200)

Abstract

Introduction

Materials and Methods

Results

Discussion

References

Tables

ABSTRACT

Misclassification of exposure usually leads to biased estimates of exposure-response associations. This is particularly an issue in cases with multiple correlated exposures, where the direction of bias is uncertain. It is necessary to address this problem when considering associations with important public health implications such as the one between mortality and air pollution, since biased exposure effects can result in biased risk assessments. The National Morbidity and Mortality Air Pollution Study (NMMAPS) recently reported results from an assessment of multiple pollutants and daily-mortality in 90 US cities. The independent associations of the selected pollutants with daily-mortality were assessed in two-pollutant models. Excess mortality was found to be associated with PM_{10}, but not with other pollutants, in these two pollutant models. The extent of bias due to measurement error in these reported results is unclear. Schwartz and Coull recently proposed a method that deals with multiple exposures and, under certain conditions, is resistant to measurement error. We applied this method to re-analyze the data from NMMAPS. For PM_{10}, we found similar results to those reported previously from NMMAPS (0.24 % increase in deaths per $10\mu g/m^3$ increase in PM_{10}). In addition we report an important effect of carbon monoxide which had not been observed previously.

INTRODUCTION

Growing evidence from published studies has shown increased all-cause and specific-cause mortality from short-term exposures to air pollution (Fairley 1999; Katsouyanni et al. 1997; Pope at al. 1995; Schwartz 1993; Schwartz and Dockery 1992). An important piece of that evidence comes from the National Mortality and Morbidity Air Pollution Study (NMMAPS) conducted over 90 US cities (Dominici et al. 2000a; Samet et al. 2000a; Samet et al. 2000b; Samet et al. 2000c). Recent updates of this study reported excess mortality in association with exposures to particulate matter of aerodynamic diameter less then or equal to 10 μm (PM_{10}), while no independent associations with gaseous pollutants were observed (Dominici et al. 2003; Dominici et al. 2002). In these previous studies, effects of pollutants were examined using two-, and multiple pollutant models.

From the public health perspective, when considering the evidence of a positive association between air pollution and mortality, it is important to confirm that such an effect is not biased due to exposure misclassification, and if so to correct for that bias. The magnitude and direction of uncertainty in the observed effects of air pollution due to exposure measurement error, has been argued by several investigators as a limitation in making causal inference for the link between air pollution and health outcomes (Lipfert and Wyzga 1997, 1999). In a single pollutant model, exposure measurement error, due to the non-differential misclassification, will underestimate the *"true"* effects of exposure-response associations (bias towards the null). Because of this, risk assessments based on the findings of observational epidemiologic studies may underestimate the benefits of

5

reducing exposures. This is particularly true for air pollution studies, which, unlike cancer risk assessment, rely on maximum likelihood estimates of risk coefficients, and not upper confidence estimates.

The situation is more complex in the case of multiple correlated pollutants. Here, the measurement error in one pollutant will tend to bias the risk coefficient of that pollutant toward the null. However, measurement error in the second pollutant will contribute some bias to the coefficient of the first pollutant. The direction of the bias will depend on the sign of the correlation between the pollutants. In rare cases, when the correlation is high between the two pollutants, and the measurement error in the second is large, this can result in an upward bias in the risk coefficient of the first pollutant. This may lead to an overestimation of exposure effects of the better measured pollutant (bias away from the null) (Schwartz 2000; Zeger et al. 2000). In the context of studies of air pollution, the upward bias has been cited as one reason for the positive associations between air pollutants and health outcomes (Lipfert and Wyuzga 1997, 1999). In recent analysis of this issue, Zeger et al. (2000) demonstrated that in the case of two pollutants measured with error, the correlation between the two pollutants, the variances of measurement errors of these two pollutants, and the correlation between the two errors would predict the magnitude and direction of bias. The study showed that even with hypothetically large differences in the four parameters, upward bias was unlikely (Zeger et al. 2000). Schwartz and Coull (2003) reported a similar finding in a simulation study, where under certain assumptions such as high correlation between the pollutants and their errors (>0.95), and/or large difference in the error variances, upward bias was not likely to occur.

Why is there measurement error in air pollution studies? A recent development in air pollution epidemiology, called time-series design, is based on series of air pollution concentrations and health outcomes (events) over a certain period of observation (which may be months or years). Through this, one can estimate the average number of events that occur with changes in air pollution concentrations. The unit of analysis is the day, and the outcome data are the counts of events (mortality, or other health outcomes). Exposure data are usually ambient concentrations of different air pollutants measured continuously (hourly or daily) from fixed-site monitoring stations. However, monitored ambient concentrations of air pollutants are not representative of personal exposures, which are important when evaluating the relation of exposure-health outcome at the individual level. Unfortunately, in time-series studies of air pollution there are no available calibration data (routinely measured series of personal exposure data) on which to base a measurement error correction. Dominici et al. (2000b) described a method to estimate a correction factor, using information on ambient and personal exposures in several cities in the US, addressing only one pollutant (PM_{10}). However, the applicability of that approach to the population-based time-series is difficult, due to the fact that a more complex scenario of multiple air pollutants is generally present, and, only information on ambient exposures is usually available.

In recent papers, Schwartz and Coull (Schwartz 2000; Schwartz and Coull 2003) developed an approach that uses hierarchical modeling to assess exposure-health outcome associations, which is resistant to exposure measurement error. The method is useful in

studies with multiple exposures, as in the case of air pollution time-series studies, and can provide bias-corrected estimates for the multiple exposures in the presence of measurement error. The present work aimed in two ways. First, to validate recent findings in air pollution epidemiology, we applied this approach to examine the independent effects of PM_{10} and several gaseous air pollutants on daily deaths, using recent data and results from NMMAPS. Second, we demonstrated the method with the intention to make it compelling to other researchers as a useful tool in assessing causal relationships.

MATERIALS AND METHODS

The NMMAPS study was an analysis of the association between air pollution and daily deaths in 90 US cities. The cities included essentially the entire urban US population living in counties with regular air pollution monitoring. Data on daily deaths in this study were obtained from records of the National Center for Health Statistics, and air monitoring data were obtained from the US Environmental Protection Agency, for the period from 1987 to 1994. This study conducted Poisson regressions, relating the daily death counts as a function of each day to air pollution concentrations on the same day, the previous day, and two days before the event, controlling for weather and season. Under such model, a Poisson process is assumed to describe the number of deaths per day, with events following a binomial process with low probability of occurrence. The model had the following form: $\log[E(Y)] = \alpha + \Sigma\beta_j X_j + \Sigma\beta_k Z_k$, where $E(Y)$ was the expected daily death count, β_j were the coefficients measuring the effects of j pollutants, and β_k were the coefficients measuring the effects of k predictors (weather, season). The study looked at the effects on daily mortality from ambient concentrations of PM_{10}, sulfur dioxide (SO_2),

8

nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3). Further details have

been published elsewhere (Dominici et al. 2003; Samet et al. 2000b; Samet et al. 2000c).

The hierarchical model

The approach from Schwartz and Coull (2003) applied to this study is described as

follows. If an outcome were linearly associated with two exposures, in our case two

pollutants (X_1 and X_2), than we would have a model such as:

$$E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \qquad (1)$$

where $E(Y)$ is the expected daily mortality, and β_1 and β_2 are the unbiased effects of

respectively X_1 and X_2. If X_1 and X_2 were correlated with each other than we could also

fit a model like the following:

$$X_2 = \gamma_0 + \gamma_1 X_1 + e \qquad (2)$$

If we now substitute X_2 in equation (1) with equation (2), then we would obtain:

$$E(Y) = (\beta_0 + \beta_2 \gamma_0) + (\beta_1 + \gamma_1 \beta_2) X_1 \qquad (3)$$

If we were to regress Y against X_1 alone, then:

$$E(Y) = \delta_0 + \delta_1 X_1 \qquad (4)$$

And by comparing (3) and (4) we would obtain:

$$E(\delta_1) = \beta_1 + \beta_2 \gamma_1 \qquad (5)$$

Hence, as equation (5) shows, by regressing the coefficient relating X_1 to mortality (δ_1)

against the coefficient relating X_2 to X_1 (γ_1), we can recover β_2, the coefficient relating X_2

to mortality. If instead of substituting for X_2 we had substituted for X_1, then we would

similarly have obtained an estimate of β_1.

The advantage of this method is seen when we consider the impact of measurement error. If X_1 and X_2 are both measured with error, then the coefficients γ_1 and δ_1 in equations (2) and (4) are both biased; however that bias depends only on the variance of X_1 and its measurement error, which are the same in both equations, and cancel out in equation (5). This results in the estimate of β_2 in equation (5) being unbiased. The extension of the approach to models with additional predictors, and Poisson regressions, is provided in Schwartz and Coull (2003).

Now, let us consider the application of this hierarchical analysis to an epidemiologic study. We applied the method to the NMMAPS study to estimate the unbiased independent effects of each of the five pollutants (PM_{10}, SO_2, NO_2, CO, and O_3) on daily mortality. One can think of the application of the hierarchical approach of Schwartz and Coull (2003) as a three step analysis. As an example, we are presenting each step using two pollutants (for example SO_2 and PM_{10}). If we were to estimate the "true" effects of each of the two pollutants on daily mortality, then β_1 (in equation (1) above) would represent the unbiased effect of SO_2, and β_2 the unbiased effect of PM_{10}. However, we are unable to estimate β_1 and β_2 directly, due to measurement error in each pollutant. The three stage method then comes into play.

In the *first stage*, daily values of the various air pollutants, in each city, were regressed against each other to obtain regression slopes for each pollutant-pollutant combination (pollutant pairs), using least squares linear regression. In this step we applied equation (2) above, which in our example would take the form: $PM_{10} = \gamma_0 + \gamma_1 SO_2$

+ e. To enhance comparability with the original NMMAPS results, we obtained the daily

air pollution data used by the NMMAPS researchers, so that their cleaning and averaging

procedures would be reflected in our analysis. This data has been made publicly available

as part of the Internet-based Health and Air Pollution Surveillance System (IHAPSS

2003).

In the *second stage*, single pollutant models were fit to the mortality data, in each

city, for each of the pollutants being examined (equation (4) above). In such case, if

estimating the effects of SO_2 on daily mortality then: $E(Y) = \delta_0 + \delta_1 SO_2$. The NMMAPS

study already provided results for the single pollutant models in each city, relating daily

death counts to daily concentration of each pollutant (PM_{10}, SO_2, NO_2, CO, and O_3) on

the day before (lag 1) the event, using Poisson regression modeling (Samet et al. 2000b;

Dominici et al. 2003). Therefore, the slopes relating mortality and each of the air

pollutants the day before death, by city, as provided by the NMMAPS, were the

coefficients (δ_1) of the *second stage* of the hierarchical approach.

Finally, the slopes (δ_l) obtained from the second stage were regressed against

those (γ_l) obtained from the first stage, across the 90 US cities (equation 5), using least

squares linear regression. The slope of this final regression is, ideally, an unbiased

estimate of β_2, the effect of PM_{10} on daily mortality, controlling for the effect of SO_2 and

of measurement error.

The method allows one to recover the independent effect of each pollutant, controlling

for any other pollutant, using an approach that is in principle, unbiased by measurement

error and which, in simulations, appears to be relatively unbiased under moderate

11

violations of the model assumptions (Schwartz and Coull 2003). Obviously, for this approach to work, the independent variable (γ_1) in equation (5) must vary across the cities.

In the application of the three stage analyses, the five pollutants were paired with each other. That is, for each pollutant, four unbiased independent estimates relating that pollutant to daily mortality were obtained, each estimate controlling for the effects of the other four pollutants. One limitation of the hierarchical modeling is loss in precision in these estimates, since the last regression (equation 5) has only 90 observations. This is reflected in the confidence limits of the estimates (Table 2). We could improve power in the effect of each pollutant, by averaging its four estimates, using the following formula:

$$\overline{\beta}_j = \frac{\sum_{i=1}^{k} w_{ji}\beta_{ji}}{\sum_{i=1}^{k} w_{ji}}$$

Where $\overline{\beta}_j$ is the weighted average slope (*meta-slope*) for pollutant j, i.e. PM_{10}, defined over k number slopes (obtained from third stage regressions), and β_{ji} is the unbiased slope of pollutant j (i.e. PM_{10}) controlling for pollutant i (i.e. SO_2), where i = 1 to k. Weights for the slopes were defined as follows:

$$w_{ji} = \frac{1}{v_{ji}}$$

Where v_{ji} is the variance of β_{ji}. The variance of $\overline{\beta}_j$ is calculated as: $\overline{v}_j = \sum_{i=1}^{k} v_{ji}$

RESULTS

The relationship between pollutant pairs differed across the 90 cities (Table 1). The heterogeneity in the pollutant-pollutant regression slopes (γ_1) among all the cities, as shown in Table 1, assured sufficient variability in the independent variable of the third stage regression to proceed with the analysis. Power in a linear regression is increased by increasing the variability in the independent variable (γ_1 in this case) and by reducing variability in the residuals. Table 1 indicates that at least the first condition is met. The range of variability in slopes relating PM_{10} to other pollutants was lower for traffic–related pollutants (CO and NO_2) than for SO_2 or O_3. This likely reflects traffic particles always being a substantial component of PM_{10}, whereas the correlation with SO_2 is more varied since it depends on the sulfur content of fuel. SO_2 is poorly correlated with O_3, resulting in a very small mean slope, and large range of variation. The range of variation in the slope relating NO_2 to SO_2 was smaller than for the other pollutants. It is possible this reflects the importance of diesel emissions for NO_2 concentrations in urban areas. Diesel fuel has much higher sulfur content than gasoline, and this may contribute to a tighter spread of the association between the two pollutants across cities.

The bias-corrected estimates from the third stage analysis are presented in Table 2. The results presented are percent increase in daily deaths for a $10\mu g/m^3$ increase in PM_{10}, or a 10 ppb increase in each of the gaseous air pollutants, except CO (100 ppb). For PM_{10}, the percent increase in daily mortality ranged from 0.14 to 0.35 (controlling for other pollutants), with an overall estimate of 0.24 percent (95 percent confidence interval: 0.05 percent, 0.42 percent). In contrast, we found small and non-significant associations of daily deaths with SO_2, NO_2, and O_3.

Increased daily-mortality in association with CO was found in the present analyses. An overall relative excess daily-mortality of 0.06 percent was observed per increments of 100 ppb of CO, estimated with fair precision (95 percent confidence interval: 0.02 percent, 0.10 percent).

DISCUSSION

The validity of exposure-response associations in epidemiologic studies depends on the precision of exposure measurements (Dominici et al. 2000b; Schwartz and Coull 2003; Zeger et al. 2000). Environmental studies of air pollution often lack precisely measured exposures, which can lead to exposure misclassification and biased estimates of exposure-response associations (Dominici et al. 2000b; Schwartz and Coull 2003; Zeger et al. 2000). Usually, an exposure measured with error will bias the association towards the null. A more complicated problem occurs when two or more exposures are measured with error, and are correlated with each other. Zeger et al. (2000) reported that in such a case, the better measured exposure may, rarely, capture some of the effect of the other exposure, which could bias away from the null. In any event, the extent of downward bias in each pollutant effect is still dependent on the measurement error in the other exposure. Hence, environmental studies of multiple air pollutants may inherit bias in both directions and with varying degrees. Underestimating the public health consequences of air pollution exposure can result in suboptimal measures to reduce these health consequences, particularly when cost-benefit or cost–effectiveness analysis is used as part of the decision process. Uncertainties about upward bias in the effect estimates can

undermine the credibility of observed associations, raising questions about the appropriateness of proposed air quality standards. Hence reducing both potential errors can improve environmental health.

Schwartz and Coull (2003) recently described a method that uses hierarchical modeling to deal with confounding and measurement error bias in epidemiological studies. Their approach yielded an exposure-response estimate which was unbiased under certain assumptions, and showed small downward biases when those assumptions were not met -- for example when the measurement errors among the multiple exposures were correlated. Although not entirely eliminated, the bias effect in this case was much less than the one produced by the two-pollutant model under the same circumstances (Schwartz and Coull 2003).

The present study applied the method of Schwartz and Coull (2003) to data and results from the National Morbidity and Mortality Study (NMMAPS) to estimate the independent effects of air pollutants on daily mortality. The application of this approach to the NMMAPS results was used in the case of two concurrent pollutants, both assumed to be measured with error. The method provided minimally biased independent effect estimates for each pollutant-daily mortality association. The price of this reduced bias was a reduction in precision. However, when we averaged over the results for each different pollutant, important associations appeared for PM_{10} and CO.

Recent results of NMMAPS had found positive association between PM_{10} and daily-mortality for 90 US cities (Dominici et al. 2003; Samet et al. 2000b; Samet et al. 2000c). The relative increase in daily mortality was 0.21 percent (SE = 0.06 percent) per 10 $\mu g/m^3$ increase of PM_{10} concentration one day prior the event. The presence of other pollutants in the model did not change this effect (Dominici et al. 2003). Our estimate for this pollutant was slightly greater after reducing measurement error bias (0.24 percent increase in mortality; SE = 0.09 percent).

The NMMAPS had reported no independent associations between daily mortality and concentration of other air pollutants one day prior the event, including SO_2, NO_2, CO, and O_3 (Dominici et al. 2003). The findings for gaseous pollutants from that study were based on two and multiple pollutant models. For comparison with our results we report the estimates from this previous study for increments in concentration of 10 ppb. The estimates for SO_2 from the two and multiple pollutant models from the NMMAPS ranged between 0.4 and 0.5 percent increase in daily mortality. NO_2 showed relative increases in daily mortality from 0.3 to about 0.4 percent. Relative increase in daily mortality for O_3 varied between 0.08 to 0.2 percent. Percent increases from 0.02 to 0.06 in daily mortality were associated with increments of 100 ppb in CO concentrations. None of these reported effects were estimated precisely, which resulted to the summary of the evidence from this previous study of no association between any of the gaseous pollutants and daily mortality (Dominici et al. 2003).

We found the effects for SO_2, controlling for other pollutants, to vary greatly, with an overall estimate of 0.1 percent increase in daily mortality. For NO_2, we found essentially no effect on daily mortality (estimate = -0.004 percent). O_3 effects in our study were found between two and three orders of magnitude smaller than the observed effects for the same pollutant from the NMMAPS. However, none of these estimates were precise, which made our summary finding for these pollutants qualitatively similar to the one reported from the NMMAPS study.

Unlike in the NMMAPS finding, we observed an association between CO and daily mortality. The estimates controlling for other pollutants, ranged from -0.02 to 0.09 percent, with the greater effects being estimated fairly precise. The pooled effect of CO showed a 0.06 percent increase in mortality with a tight confidence interval (95% CI: 0.02 – 0.10).

One explanation for the different finding for CO in our study, compared to that of NMMAPS, could be related to the high degree of measurement error in this pollutant. There is possibility that the spatial heterogeneity in ambient concentrations of CO is greater than that of any other air pollutant. This would result in a greater amount of measurement error when monitoring ambient concentrations of CO from a central monitoring site. Whether the greater effect seen using the hierarchical modeling approach reflects a true association with CO *per se*, or CO is a surrogate for traffic particles or some other component of vehicular exhaust (Sarnat et al. 2001), is not clear. Nevertheless, the results for CO indicate the potential usage of the approach and suggest

that attention should be focused on CO or on traffic pollution. The similar results of this study with those of NMMAPS for PM_{10} were reassuring.

There are several limitations in the application of the Schwartz and Coull approach (2003) that must be acknowledged. First, because the third stage regression had only 90 observations, the power of the method was reduced. The power depends in part on the R^2 of the third stage models. In our case, these were low (Table 2), forcing us to apply meta-analysis as an *ad hoc* approach to improve power. In other applications, the R^2 may be larger, and this approach may be unnecessary. We are currently developing a multivariate version of the approach that is less *ad hoc* in improving power, by using multiple predictors at the last stage regression.

Second, we did not have seasonal specific regression coefficients relating air pollution to mortality in the NMMAPS. The use of this approach with season specific relationships offers increased power because of the increased number of risk coefficients by pollutant, and also because the relationship between many pollutants varies seasonally. We expect these to be available from NMMAPS in the future.

Third, in the present analyses we assumed that the concentration-response relation between air pollution and daily deaths is linear. This question has been subject to a number of investigations using splines and smoothing in single city studies and combining splines or smooth curves in multi-city studies (Schwartz et al. 2002, Schwartz and Zanobetti 2000). At least for PM_{10}, linear associations were seen in concentration ranges similar to those of the present study.

Finally, our model implies that each pollutant, as measured at a central site in each city, is a surrogate for exposure to the same pollutant. However, Sarnat and coworkers (2001) have proposed recently that ambient concentrations of gaseous air pollutants may be serving as surrogates not for exposure to those gases, but for exposure to particles or particles from particular sources. For example, in Baltimore, CO was a good surrogate for exposure to particles from traffic (Sarnat et al. 2001). This suggests that there must be caution in interpreting the present results for CO. We think that personal exposure studies of intermediate outcomes may be necessary to resolve the question.

Despite these limitations, the approach of Schwartz and Coull (2003) is useful to studies of air pollution and mortality. In our study, it provided evidence of a slightly greater effect of PM_{10} than that reported previously by NMMAPS, when controlling for measurement error in other pollutants, and suggested that greater attention be paid to CO and possibly traffic pollution.

References

Dominici F, McDermott A, Daniels M, Zeger SL, Samet JM. 2003. Mortality among residents of 90 cities. Health Effects Institute. Revised analyses of time-series studies of air pollution and health. Special Report. Capital City Press, Montpelier VT.

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On the use of generalized additive models in time-series studies of air pollution and health. Am J Epidemiol 156:193-203.

Dominici F, Samet J, Zeger S. 2000a. Combining evidence on air pollution and daily Mortality from the twenty largest US cities: a hierarchical modeling strategy (with discussion). J R Stat Soc Ser A 163:263-302.

Dominici F, Zeger SL, Samet JM. 2000b. A measurement error model for time-series studies of air pollution and mortality. Biostatistics 1:157-175.

Fairley D. 1999. Daily mortality and air pollution in Santa Clara County, California: 1989-1996. Environ Health Perspect 107:637-641.

Interned-based Health and Air Pollution Surveillance Systems (IHAPSS) 2003. Available: http://www.ihapss.jhsph.edu/data/data_download.htm [accessed 15 September 2003].

Katsouyanni K, Touloumi G, Spix C, Schwartz J, Balducci F, Medina S, et al. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. Air Pollution and Health: a European Approach. BMJ 314:1658-1663.

Lipfert FW, Wyzga RE. 1997. Air pollution and mortality: the implications of uncertainties in regression modeling and exposure measurement. J Air Waste Manag Assoc 47:517-523.

------. 1999. Statistical considerations in determining the health significance of constituents of airborne particulate matter. J Air Waste Manag Assoc 49:182-91.

National Morbidity and Mortality Air Pollution Study (NMMAPS) 2003. Available: http://www.biostat.jhsph.edu/biostat/research/update.main.htm [accessed 15 September 2003].

Pope CA, 3rd, Bates DV, Raizenne ME. 1995. Health effects of particulate air pollution: time for reassessment? Environ Health Perspect 103:472-480.

Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. 2000a. Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. N Engl J Med 343:1742-1749.

Samet JM, Dominici F, Zeger SL, Schwartz J, Dockery DW. 2000b. The National Morbidity, Mortality, and Air Pollution Study. Part I: Methods and methodologic issues. Res Rep Health Eff Inst 5-14; discussion 75-84.

Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery DW, et al.. 2000c. The National Morbidity, Mortality, and Air Pollution Study. Part II: Morbidity and mortality from air pollution in the United States. Res Rep Health Eff Inst 94:5-70; discussion 71-9.

Sarnat JA, Schwartz J, Catalano PJ, Suh HH. 2001 Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? Environ Health Perspect 109:1053-1061.

Schwartz J. 1993. Air pollution and daily mortality in Birmingham, Alabama. Am J Epidemiol 137:1136-1147.

------. 2000. Daily deaths are associated with combustion particles rather than SO_2 in Philadelphia. Occup Environ Med 57:692-697.

Schwartz J, Coull B. 2003. Control for Confounding in the presence of measurement error in hierarchical models. Biostatistics 4:569-582.

Schwartz J, Dockery DW. 1992. Increased mortality in Philadelphia associated with daily air pollution concentrations. Am Rev Respir Dis 145:600-604.

Schwartz J, Laden F, Zanobetti A. 2002. The concentration-response relation between PM(2.5) and daily deaths. Environ Health Perspect 110:1025-1029.

Schwartz J, Zanobetti A. 2000. Using meta-smoothing to estimate dose-response trends across multiple studies, with application to air pollution and daily death. Epidemiology 11:666-672.

Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, et al.. 2000 Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ Health Perspect 108:419-426.

Table 1. Mean and standard deviation (SD) of the distribution of pollutant-pollutant regression slopes (γ_1) [a], and their coefficients of variation (CV – in percent), across 90 US cities.

Pollutant-Pollutant Regression Variables		Pollutant-Pollutant Regression Slopes		
Dependent Pollutant	Independent Pollutant	Mean [b]	SD	CV (%)
PM_{10}	SO_2	1.31	1.45	110.3
	NO_2	0.83	0.64	77.2
	CO	0.01	0.01	62.7
	O_3	0.27	0.38	143.6
SO_2	PM_{10}	0.09	0.08	87.9
	NO_2	0.22	0.20	91.6
	CO	0.003	0.003	92.0
	O_3	-0.03	0.10	377.6
NO_2	SO_2	2.64	5.09	192.8
	PM_{10}	-0.22	1.83	831.1
	CO	-0.01	0.08	1030.7
	O_3	0.19	2.03	1064.6
CO	SO_2	61.58	74.64	121.2
	NO_2	34.19	19.39	56.7
	PM_{10}	13.73	9.26	67.4
	O_3	-9.58	9.47	98.8
O_3	SO_2	136.66	535.82	392.1
	NO_2	24.76	95.82	387.0
	CO	-0.59	3.26	550.6
	PM_{10}	-63.99	292.61	457.2

[a] The slopes were obtained from *first stage* regressions (equation (2)), paring each of the five pollutants with the other four (pollutant pairs), for each US city.
[b] Mean and SD represent the mean and standard deviation of the distribution of first stage regression slopes across the 90 cities, for each pollutant pair.

Table 2. Percent increase in daily deaths associated with each pollutant, controlling for measurement error in other pollutants, based on data and results from the NMMAPS study, including the period between 1987 and 1994.

	Independent Effects of Pollutants on Daily Mortality [a]														
Paired with pollutant [b]	**PM_{10}**			**SO_2**			**NO_2**			**CO**			**O_3**		
	R^{2} [c]	**Slope** [d]	**T** [e]	**R^2**	**Slope**	**T**	**R^2**	**Slope**	**T**	**R^2**	**Slope**	**T**	**R^2**	**Slope**	**T**
PM_{10}				0.048	1.14	1.77	0.008	0.033	0.69	0.029	0.078	1.56	0.008	0.0003	0.77
SO_2	0.047	0.28	1.73				<0.001	-0.004	0.06	0.154	0.086	3.25	0.012	-0.0004	0.83
NO_2	0.012	0.16	0.83	0.007	-0.29	0.54				0.007	0.032	0.64	0.024	-0.0019	1.15
CO	0.006	0.14	0.69	0.030	0.65	1.33	0.001	-0.004	0.27				0.045	0.0011	1.87
O_3	0.036	0.35	1.70	0.039	-0.76	1.52	0.007	-0.025	0.63	0.002	-0.018	0.34			
Meta-slope [f]		0.24	2.80		0.10	0.34		-0.004	0.27		0.062	3.16		0.0002	0.74

[a] Corrected slopes (effects) removing the effect of each pollutant shown in the first column.
[b] The independent pollutant of *first stage* regression.
[c] Adjusted R^2 from regressions of *third stage* (see text for more detail).
[d] Slope for PM_{10} presented as percent increase per $10\mu g/m^3$. Slopes for other pollutants are presented as percent increase per 10 ppb (100 ppb for CO).
[e] t-statistic from regressions of the *third stage*.
[f] Meta-slope of the four alternative estimates.