Harvard University

Harvard University Biostatistics Working Paper Series

Year 2006 *Paper* 46

An Informative Bayesian Structural Equation Model to Assess Source-Specific Health Effects of Air Pollution

Margaret C. Nikolov* Brent A. Coull†

Paul J. Catalano‡ John J. Godleski**

*Harvard School of Public Health, meg.nikolov@gmail.com

†Harvard University, bcoull@hsph.harvard.edu

‡Harvard School of Public Health and Dana Farber Cancer Institute, catalano@hsph.harvard.edu

**Harvard School of Public Health, jgodlesk@hsph.harvard.edu

An Informative Bayesian Structural Equation Model to Assess Source-Specific Health Effects of Air Pollution

Margaret C Nikolov, Brent A Coull, Paul J Catalano, John J Godleski

SUMMARY

A primary objective of current air pollution research is the assessment of health effects related to specific sources of air particles, or particulate matter (PM). Quantifying source-specific risk is a challenge, because most PM health studies do not directly observe the contributions of the pollution sources themselves. Instead, given knowledge of the chemical characteristics of known sources, investigators infer pollution source contributions via a source apportionment or multivariate receptor analysis applied to a large number of observed elemental concentrations. Although source apportionment methods are well-established for exposure assessment, little work has been done to evaluate the appropriateness of characterizing unobservable sources thus in health effects analyses. In this article, we propose a structural equation framework to assess source-specific health effects using speciated elemental data. This approach corresponds to fitting a receptor model and the health outcome model jointly, such that inferences on the health effects account for the fact that uncertainty is associated with the source contributions. Since the structural equation model (SEM) typically involves a large number of parameters, for small sample settings we propose a fully Bayesian estimation approach that leverages historical exposure data from previous related exposure studies. We compare via simulation the performance of our approach in estimating source-specific health effects to that of two existing approaches, a tracer approach and a two-stage approach. Simulation results suggest that the proposed informative Bayesian SEM is effective in eliminating the bias incurred by the two existing approaches, even when the number of exposures is limited. We employ the proposed methods in the analysis of a concentrator study investigating the association between ST-segment, a cardiovascular outcome, and major sources of Boston PM, and discuss the implications of our findings with respect to the design of future PM concentrator studies.

1 Introduction

Epidemiological studies have consistently demonstrated increased morbidity and mortality outcomes associated with elevated levels of air pollution (Dockery *et al.* 1993; Dominici *et al.* 2002). Although the health risks associated with high concentrations of air pollution tend to be small, the exposed population is large such that the potential burden of morbidity and mortality attributable to air pollution is considerable (Dominici, Sheppard, and Clyde 2003). In response to these findings, the United States Congress in 1998 mandated extensive research into health effects associated with ambient air particulate matter (PM; Lippmann *et al.* 2003).

One of the major current objectives of PM research is to assess the health effects related to specific sources of air pollution, such as power plants and motor vehicles. Estimation of source-specific health effects is of primary importance from a regulatory standpoint. Recent research suggests that emissions from different sources exhibit differing levels of toxicity (Clarke *et al.* 2000; Laden *et al.* 2000; Godleski *et al.* 2002; Wellenius *et al.* 2003). The most direct way to reduce health effects of ambient air particles is to regulate the sources of pollution having adverse effects. In order to set protective standards, researchers must first establish the risk associated with the different pollution sources. Quantifying source-specific risk is a challenge because current studies investigating the health effects of air pollution do not observe the PM contributions of the sources directly. Rather, exposures consist of samples of ambient air, or concentrated versions of ambient air, which reflect dynamic mixtures of source contributions. However, by taking into account the chemical fingerprints of known sources, an assessment of the chemical composition of exposure provides indirect information on the source contributions.

The exposure assessment literature contains an ample amount of research that focuses on estimation of source-specific contributions from a complex mixture of air pollution (i.e., Koutrakis and Spengler 1987; Kavouras *et al.* 2001; and for review see Seigneur *et al.* 1999; Hopke 2003; Kim *et al.* 2004). Methods such as source apportionment and multivariate receptor modeling use factor analytic techniques to estimate the contributions of a small number of pollution sources from the measured mixture components, elements and other compounds. Alternatively, a distinct set of tracer elements

may be selected to represent the known sources, and the elemental concentrations of the tracers taken as surrogates for the source contributions.

Although receptor modeling is well-developed for exposure assessment, little work has been done to evaluate the appropriateness of characterizing unobservable sources in this way to estimate source-specific health effects. The problem can be thought of as an exposure measurement error problem (Carroll, Ruppert, and Stefanski 1995), whereby the PM exposure generated from a particular source is estimated rather than known or measured directly (Dominici, Sheppard, and Clyde 2003). At present, existing source-specific health effects analyses rely on approaches that do not take into account the uncertainty associated with estimated source contributions. A "two-stage" strategy uses estimated source contributions from a factor analysis to assess the impacts of specific pollution sources on health effects (Laden *et al.* 2000; Clarke *et al.* 2000). This two-step approach has several advantages. First, many authors have noted that the use of factor scores in regression settings is appropriate when the independent variables are highly collinear or when the underlying factors have a natural intuitive meaning (Mardia, Kent, and Bibby 1979). Both of these considerations apply in the elemental composition framework considered in PM research. Concentrations of elements that are markers of the same pollution source are typically highly correlated, and the aggregated factors represent the pollution sources themselves (Koutrakis and Spengler 1987). A variation on this two-stage strategy is the tracer approach, whereby the estimated source contributions are replaced by the elemental concentrations of a distinct set of tracers in the health effects analysis (Wellenius *et al.* 2003). In either case, reducing the dimensionality into a small number of "source" factors or tracers will typically provide more stable estimates than using all measured elemental concentrations, since the former approach is much less likely to suffer from multicollinearity.

Despite the practical advantages of implementing the two-stage or tracer approaches discussed above, the statistical properties of the resulting estimates of the health effects of PM are not well-understood. Previous statistical research has shown that, in simpler models, measurement error associated with estimated latent variables can lead to bias in the subsequent regression coefficient estimates (Tsiatis, De Gruttola, and Wulfsohn 1995; Roberts, Ryan, and Wright 2003). It is unclear whether this bias will occur in PM research as the mixtures typically observed in PM exposures are quite different from

those in other latent variable settings. As noted by Dominici, Sheppard, and Clyde (2003), there remains much work to be done in order to understand these estimates from a statistical standpoint and to assess the reliability of these estimates of association between pollution sources and health outcomes.

In this paper, we propose a structural equation framework for assessing source-specific health effects using speciated data in the form of elemental concentrations. This approach corresponds to jointly fitting a multivariate receptor model to the exposure data and a model for the health outcome given source contributions. Because the source contributions and health effects are modeled jointly, resulting inferences on the health effects account for the fact that uncertainty is associated with the exposures of interest.

This work is motivated by animal toxicology studies evaluating the mechanisms of morbidity and mortality associated with inhalation of concentrated air particles (CAPs) conducted at the Harvard School of Public Health (HSPH) (Godleski *et al.* 2000; Coull, Catalano, and Godleski 2000). Harvard researchers have implemented multiple animal toxicology studies to investigate the adverse effects of PM on cardiopulmonary and respiratory activity in canines and rats. Samples of ambient Boston aerosol are collected and are concentrated approximately 30 times by the Harvard Ambient Particle Concentrator (HAPC) (Sioutas, Koutrakis, and Burton 1995; Sioutas *et al.* 1995; Godleski *et al.* 2000) without altering the physical and chemical composition of the mixture. Animals are then exposed to the concentrated complex mixture for a given period of time, and cardiac and respiratory outcomes are monitored on each exposed animal. Because exposure is generated from ambient pollution, exposures are essentially random across days and, hence, a complete exposure assessment is made for each concentrated exposure. Data from these studies consist of the measured elemental concentrations of the concentrated air pollution mixture and the recorded health outcomes on the exposed animals. Because of the complexity of these studies, in any one study investigators typically expose animals on approximately 20 unique exposure days.

The structural equation model (SEM), as well as the factor analysis model used in the two-stage approach, typically involves a large number of parameters. Given the high dimensionality of the

model, the typical exposure study has an insufficient number of exposures to obtain reliable parameter estimates using maximum likelihood (ML). An approach for handling this problem is to consider a reduced number of elemental species for a health effects analysis (Clarke *et al.* 2000). To overcome the small sample problem, we propose a fully Bayesian estimation approach that leverages historical exposure data from previous concentrator studies in defining informative priors on the parameters relating the measured exposures to the source contributions. This serves to pool exposure information from studies which are consistent in their collection and analysis of CAPs data.

The remainder of this paper is arranged as follows. Section 2 describes in detail the design and data from a study evaluating the effects of CAPs on myocardial ischemia in dogs (Wellenius *et al.* 2003). Section 3 presents the SEM and Section 4 discusses the informative Bayesian approach to estimation. Section 5 presents a simulation study to examine the statistical properties of health effect estimates obtained with the tracer, two-stage, and structural equation methodologies. Section 6 demonstrates an application of the informative Bayesian SEM to analyze the Wellenius *et al.* (2003) study. Finally, in Section 7 we discuss our findings along with implications for the design of future PM concentrator studies.

2 Data

Wellenius *et al.* (2003) described results from a concentrator study examining the effects of inhaled CAPs on myocardial ischemia in dogs. The study subjects were six retired mongrel breeder dogs, each initially fitted with a balloon occluder around the left anterior descending coronary artery. The study design consisted of pairs of dogs undergoing three or four consecutive days of exposure and evaluation. On each day of the cycle, the dogs were put in side-by-side chambers and simultaneously underwent a continuous 6-hour exposure to either CAPs or filtered air (Sham). The dogs were randomly assigned to CAPs exposure; one dog was exposed to CAPs on the second day of the cycle, while the other dog received CAPs exposure on the third day. Immediately following each exposure period, the dogs underwent a 5 minute coronary artery occlusion, and were monitored via continuous electrocardiogram (ECG). The primary outcome of interest was peak ST-segment elevation, a marker for myocardial ischemia. The study protocol was repeated multiple times for a total of 18

Table 1: Order and timing of exposures (Wellenius *et al.* 2003)

Sequence	Dog	Start Date	Day 1	Day 2	Day 3	Day 4
1	1	9/13/2000	Sham	Sham	CAPs	-
2	2	9/13/2000	Sham	CAPs	Sham	-
3	3	12/13/2000	Sham	CAPs	Sham	-
4	2	12/13/2000	Sham	Sham	CAPs	-
5	2	1/9/2001	Sham	Sham	CAPs	-
6	3	1/24/2001	Sham	Sham	CAPs	-
7	2	1/24/2001	Sham	CAPs	Sham	-
8	3	2/6/2001	Sham	CAPs	Sham	Sham
9	2	2/6/2001	Sham	Sham	CAPs	Sham
10	4	2/13/2001	Sham	CAPs	Sham	-
11	3	2/20/2001	Sham	Sham	CAPs	Sham
12	2	2/20/2001	Sham	CAPs	Sham	Sham
13	5	2/27/2001	Sham	CAPs	Sham	Sham
14	6	2/27/2001	Sham	Sham	CAPs	Sham
15	3	3/7/2001	Sham	CAPs	Sham	-
16	5	3/12/2001	Sham	Sham	CAPs	Sham
17	6	3/12/2001	Sham	CAPs	Sham	Sham
18	6	3/27/2001	Sham	CAPs	Sham	Sham

Note that four pairings were not complete due to failed sequences.

complete and successful exposure cycles. Table 1 summarizes the study design. This experiment was conducted according to the principles and regulations of the National Institutes of Health under protocols approved by the Harvard Medical Area Standing Committee on Animals.

Samples of the CAPs exposures were collected and analyzed. Each CAPs exposure was measured for sulfate (SULF) via ion chromatography, black carbon (BC) using an aethalometer, elemental carbon (EC) and organic carbon (OC) determined with a thermal and optical reflectance method, and elemental concentrations (in ug/m^3) collected via X-ray fluorescence (XRF), specifically: aluminum (Al), arsenic (As), barium (Ba), bromine (Br), calcium (Ca), cadmium (Cd), chlorine (Cl), chromium (Cr), copper (Cu), iron (Fe), potassium (K), manganese (Mn), nickel (Ni), sodium (Na), lead (Pb), sulfur (S), selenium (Se), silicon (Si), titanium (Ti), vanadium (V), and zinc (Zn). The Sham exposures were assumed to have zero concentration of all elements, as this had been confirmed in earlier test runs of the concentrator (Sioutas, Koutrakis, and Burton 1995; Sioutas *et al.* 1995; Lawrence *et al.* 2004).

Table 2: Major Sources of Boston Air Pollution

Source	Elements
Road Dust	silicon and aluminum
Power Plants	sulfur and sulfate
Oil Combustion	nickel and vanadium
Motor Vehicles	black carbon, organic carbon, elemental carbon

Wellenius *et al.* (2003) considered linear mixed models for log transformed peak ST-segment, using individual elemental concentrations as tracer representatives of known sources of air pollution in Boston. Table 2 summarizes four major sources of PM pollution based on existing knowledge of the composition of Boston aerosol (Koutrakis and Spengler 1987; Oh *et al.* 1997; Oh 2000; Clarke *et al.* 2000; Batalha *et al.* 2002). The authors chose silicon, sulfur, nickel, and black carbon to represent resuspended road dust, coal-fired power plants, oil combustion (primarily for home heating), and motor vehicle exhaust, respectively, and found a strong positive association between log peak ST-segment and resuspended road dust. A question that naturally arises is whether measurement error associated with tracer representatives of the source contributions obscured the relationship between log peak ST-segment and the other pollution sources. In this article, we analyze the data using methods that account for the uncertainty in estimated source contributions.

3 Model and Notation

3.1 Modeling Framework

We propose a full-likelihood approach that estimates the health effects by fitting the receptor and health outcome models jointly. A general framework for our joint model is

$$\mathbf{X}_t = \mathbf{\Lambda}\boldsymbol{\eta}_t + \boldsymbol{\epsilon}_t^{\mathbf{X}} \tag{1}$$

$$Y_t = \alpha + \boldsymbol{\beta}^T\boldsymbol{\eta}_t + \epsilon_t^Y \tag{2}$$

where for a given time t, $\mathbf{X}_t$ is the vector of P elemental concentrations, $\boldsymbol{\eta}_t$ is the vector of the K unobserved source contributions, and Y_t is the health outcome. We assume that Y_t represents a single

continuous variable and that K is known. The model for $\mathbf{X}_t$ is the factor analysis model for the exposure analysis (Park, Guttorp, and Henry 2001), where $\mathbf{\Lambda}$ is the $(P \times K)$ matrix of factor loadings, also known as the factor pattern, and $\boldsymbol{\epsilon}_t^{\mathbf{X}} \overset{iid}{\sim} MVN_P(\mathbf{0}, \mathbf{\Psi})$ for diagonal $\mathbf{\Psi}$. The vector $(\lambda_{1k}, \lambda_{2k}, ..., \lambda_{Pk})$ may be viewed as the profile of pollution source k. The parameters $\boldsymbol{\beta}$ quantify the K source-specific health effects and $\epsilon_t^Y \overset{iid}{\sim} N(0, \sigma_Y^2)$. Standard factor analysis assumes $\boldsymbol{\eta}_t \overset{iid}{\sim} MVN_K(\boldsymbol{\mu}, \mathbf{\Sigma})$.

This model falls within the structural equation framework (Bollen 1989; Budtz-Jorgensen *et al.* 2003). The *measurement model* describing the relationship between the latent variables and the observed measures may be expressed as

$$\mathbf{X}_t^* = \begin{pmatrix} \mathbf{X}_t \\ Y_t \end{pmatrix} = \boldsymbol{\alpha} + \mathbf{\Lambda}^* \boldsymbol{\eta}_t + \boldsymbol{\epsilon}_t^{\mathbf{X}^*} \tag{3}$$

$$\begin{pmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{Pt} \\ Y_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \alpha \end{pmatrix} + \begin{pmatrix} \lambda_{11} & \lambda_{12} & ... & \lambda_{1K} \\ \lambda_{21} & \lambda_{22} & ... & \lambda_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{P1} & \lambda_{P2} & ... & \lambda_{PK} \\ \beta_1 & \beta_2 & ... & \beta_K \end{pmatrix} \begin{pmatrix} \eta_{1t} \\ \eta_{2t} \\ \vdots \\ \eta_{Kt} \end{pmatrix} + \begin{pmatrix} \epsilon_t^{X_1} \\ \epsilon_t^{X_2} \\ \vdots \\ \epsilon_t^{X_P} \\ \epsilon_t^{Y} \end{pmatrix}$$

where $\boldsymbol{\epsilon}_t^{\mathbf{X}^*} \overset{iid}{\sim} MVN_{P+1}(\mathbf{0}, \mathbf{\Psi}^*)$, $\mathbf{\Psi}^* = \begin{pmatrix} \mathbf{\Psi} & \mathbf{0} \\ \mathbf{0}^T & \sigma_Y^2 \end{pmatrix}$, and the *structural model* demonstrating the relationship *amongst* the latent variables is simply the model on $\boldsymbol{\eta}$, which we specify in the next section.

3.2 Distributional Assumptions

We extend the standard SEM (3) in two ways. First, we truncate a normal distribution for $\boldsymbol{\eta}$ to ensure the physical non-negativity of source contributions,

$$\boldsymbol{\eta} \overset{iid}{\sim} MVN_K(\boldsymbol{\mu}, \mathbf{\Sigma}) I(\boldsymbol{\eta} \geq \mathbf{0}).$$

Alternatively, we could specify a lognormal distribution. In our application, we assess the sensitivity of our conclusions to distributional assumptions on $\boldsymbol{\eta}$ by fitting the SEM both ways. The specification

of non-negative source contributions extends the standard factor analysis, which does not restrict the domain of $\boldsymbol{\eta}$ in the model and allows negative source contributions. Positive Matrix Factorization (Paatero and Tapper 1994; Kim *et al.* 2004) is a alternative method that uses constrained weighted least squares to ensure non-negative source contributions.

Second, to accommodate the repeated measures design of these studies, we build random effects into the model for the health outcome,

$$Y_{st} = \alpha + \boldsymbol{\beta}^T \boldsymbol{\eta}_t + \mathbf{Z}_{st}^T \mathbf{b}_s + \epsilon_{st}^Y$$

where Y_{st} is the health outcome and $\mathbf{Z}_{st}$ is the vector of covariates for unit s at time t, $\mathbf{b}_s$ is the vector of random effects for unit s, $\mathbf{b}_s \overset{iid}{\sim} MVN(\mathbf{0}, \boldsymbol{\Sigma}_{\mathbf{b}})$, $\epsilon_{st}^Y \overset{iid}{\sim} N(0, \sigma_Y^2)$, and $\mathbf{b} \perp \epsilon^Y$ (Diggle *et al.* 2002).

3.3 Model Identifiability

The SEM specified in (3) is not identifiable without further assumptions. Because the source profiles are unknown *and* the source contributions are unobserved, the structural equation model does not have a unique solution. However, the model may be made identifiable by constraining parameters in $\boldsymbol{\Lambda}$. We consider the following two sets of identifiability conditions, which result in a confirmatory, rather than exploratory, factor analysis (Park, Spiegelman, and Henry 2002).

C1: There are at least $K-1$ zero elements in each column of $\boldsymbol{\Lambda}$

C2: The rank of $\boldsymbol{\Lambda}^{(k)}$ is $K-1$, where $\boldsymbol{\Lambda}^{(k)}$ is the matrix composed of the rows containing the assigned 0s in the kth column with those assigned 0s deleted.

C3: $\lambda_{pk} = 1$ for some p $(p = 1, 2, .., P)$ for each $k = 1, 2, .., K$

D1: There are at least K rows in $\boldsymbol{\Lambda}$ with each of the K rows containing only one nonzero element.

D2: Same as C2

D3: Same as C3

The C1-C3 conditions assume that there are at least $K-1$ elements per source that are not associated with that source, and that this group of elements is not the same for all sources. The D1-D3 conditions assume that there is at least one "tracer" element for each source in the sense that the tracer does not load on other sources. Further, as noted by Park, Guttorp, and Henry (2001), these conditions identify the loadings up to normalization. Thus, for each source, we specify one loading to be equal to 1, effectively placing the source contribution on the scale of the element having the constrained loading of 1 for that source.

The C1-C3 conditions and the D1-D3 conditions are each sufficient but not necessary to establish identifiability. While there exist alternative conditions, other commonly used proposals are also sufficient but not necessary. For instance, Park, Spiegelman, and Henry (2002) proposed sufficient conditions which, instead of placing constraints on the factor loadings, assume that some sources are absent on some days. These authors argued that in some settings, this alternative set of constraints may be plausible if one knows that a particular source, such as a power plant in the region, has been shut down for some period of time. In the same vein, Bandeen-Roche (1994) considered situations in which a subset of the source contributions is known. In our setting, however, we do not have information on the presence or absence of a particular source on a particular day. Thus, given the existing literature on the pollution mixture in the Boston area (Oh *et al.* 1997; Oh 2000), it seems safer to assume that certain elements are not markers for certain sources. Bollen (1989) also gave some rules that help determine whether a model is identifiable, but, as noted by this author, these rules are also either necessary, or sufficient, but not both.

4 Estimation

Standard SEMs may be fit via ML using existing latent variable software, such as Mplus (Muthen and Muthen 1998) or the `sem` package in R (R Development Core Team 2003). Due to the large number of parameters involved, these methods require a large sample size for asymptotic optimality of the resulting ML estimators. In studies motivating this research, the number of exposures rarely exceeds 20 and maximum likelihood methods break down. The problem is due to the large number

parameters in the factor analysis model (1). Given P elements and K factors, $\mathbf{\Lambda}$ is of dimension $(P \times K)$. Even with K^2 (tracer) identifiability constaints, we have $K(P-K)$ free parameters to estimate in $\mathbf{\Lambda}$ alone. Add to this estimation of the P specific variances in $\mathbf{\Psi}$ and, in the case of uncorrelated factors, the K variances in $\mathbf{\Sigma}$. For even moderate P and K, we have a large number of parameters to estimate. Twenty days of exposure are simply too few to obtain a reliable receptor model fit in most cases.

To overcome the small number of unique exposure days, we propose an informative Bayesian approach to model fitting. This approach is especially appealing considering that HSPH researchers have conducted multiple concentrator studies, all of which are consistent in their collection and analysis of exposure data. It is reasonable to pool the exposure data from prior studies to estimate the profiles of PM sources in Boston. An informative Bayesian approach leverages historical exposure data to obtain more reliable estimates of the source profiles, thus improving our ability to estimate the health effects investigated in an individual study.

The Bayesian approach incorporates information from previous studies through specification of the priors. In this case, a preliminary factor analysis of the historical exposure data provides prior information on the unknown factor pattern $\mathbf{\Lambda}$. Let $\boldsymbol{\lambda}$ be the $K \times (P-K)$ vector of all unconstrained factor loadings, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1^T, \boldsymbol{\lambda}_2^T, ..., \boldsymbol{\lambda}_K^T)^T$, where $\boldsymbol{\lambda}_k$ is the vector of unconstrained loadings for source k $(k = 1, ..., K)$. Let $\widehat{\boldsymbol{\lambda}}^{(hist)}$and $\widehat{Var}(\widehat{\boldsymbol{\lambda}}^{(hist)})$ represent the posterior mean and covariance of the factor loadings obtained from a Bayesian factor analysis of the historical data. For the informative Bayesian SEM, the prior distribution on the free parameters in $\mathbf{\Lambda}$ may be defined as:

$$\boldsymbol{\lambda} \sim MVN_{K(P-K)}(\widehat{\boldsymbol{\lambda}}^{(hist)}, \widehat{Var}(\widehat{\boldsymbol{\lambda}}^{(hist)})), \tag{4}$$

while the constrained loadings are treated as fixed constants in the likelihood. The information in the prior on $\boldsymbol{\lambda}$ supplements the 20 or so days of exposure data and provides for better estimation of the SEM. In addition to leveraging historical exposure data to aid estimation, the Bayesian approach is flexible in handling the physical constraints of air pollution data, such as non-negative source contributions. This strategy is demonstrated in the application in Section 6.

5 Simulation Study

We conducted a simulation study to examine the statistical properties of the health effect estimates obtained via the tracer, two-stage, and structural equation approaches. In the interest of direct comparison between the various approaches, we assume a normal mean zero distribution on the source contributions to ensure that our assessments are not confounded by distributional assumptions made by different implementations of SEMs.

In order to make our findings most relevant to the HSPH concentrator studies, we based our simulations on the known sources of Boston PM pollution described in Table 2. We obtained realistic parameter settings for $\mathbf{\Lambda}$, $\mathbf{\Sigma}$, and $\mathbf{\Psi}$ from a confirmatory factor analysis on the complete aggregated exposure data ($N = 178$). As noted by Park, Guttorp, and Henry (2001), it is important to first select a subset of species that are contributed by major pollution sources. Thus, we conducted our analysis on a subset of $P = 13$ elements deemed to be major components the four known sources of Boston PM; silicon (Si), sulfur (S), nickel (Ni), organic carbon (OC), aluminum (Al), titanium (Ti), calcium (Ca), sulfate (SULF), selenium (Se), vanadium (V), bromine (Br), black carbon (BC), and elemental carbon (EC). Since convergence problems are common when elemental concentrations are on widely different scales, each element was scaled by its sample standard deviation, which is equivalent to conducting a factor analysis on the sample correlation matrix, as opposed to the sample covariance matrix.

We constrained one "tracer" element for each of the $K = 4$ sources according to the D1-D3 identifiability conditions. We chose silicon, sulfur, nickel, and organic carbon to identify road dust, power plants, oil combustion, and motor vehicles, respectively. A preliminary exploratory factor analysis justified the "tracer" identifiability conditions, since the estimated factor loadings of silicon, sulfur, nickel, and organic carbon were low (< 0.2) on all but a single source.

The parameters were defined as follows:

$$\mathbf{\Lambda} = \begin{pmatrix} & RoadDust & PowerPlants & OilCombustion & MotorVehicles \\ Si & 1 & 0 & 0 & 0 \\ S & 0 & 1 & 0 & 0 \\ Ni & 0 & 0 & 1 & 0 \\ OC & 0 & 0 & 0 & 1 \\ Al & 0.88 & 0.00 & 0.01 & 0.00 \\ Ti & 0.83 & 0.08 & 0.34 & 0.09 \\ Ca & 0.91 & 0.02 & 0.31 & 0.17 \\ SULF & 0.00 & 0.95 & 0.00 & 0.01 \\ Se & 0.02 & 0.65 & 0.05 & 0.26 \\ V & 0.16 & 0.04 & 1.02 & 0.03 \\ Br & 0.18 & 0.58 & 0.26 & 0.43 \\ BC & 0.17 & 0.41 & 0.44 & 0.65 \\ EC & 0.13 & 0.27 & 0.51 & 0.81 \end{pmatrix}$$

$$diag(\mathbf{\Psi}) = \begin{pmatrix} \psi_{Si} \\ \psi_{S} \\ \psi_{Ni} \\ \psi_{OC} \\ \psi_{Al} \\ \psi_{Ti} \\ \psi_{Ca} \\ \psi_{SULF} \\ \psi_{Se} \\ \psi_{V} \\ \psi_{Br} \\ \psi_{BC} \\ \psi_{EC} \end{pmatrix} = \begin{pmatrix} 0.08 \\ 0.05 \\ 0.22 \\ 0.45 \\ 0.05 \\ 0.28 \\ 0.35 \\ 0.05 \\ 0.31 \\ 0.05 \\ 0.31 \\ 0.11 \\ 0.10 \end{pmatrix}$$

$$\mathbf{\Sigma} = \begin{pmatrix} & RoadDust & PowerPlants & OilCombustion & MotorVehicles \\ RoadDust & 2.36 & 0 & 0 & 0 \\ PowerPl & 0 & 1.60 & 0 & 0 \\ OilComb & 0 & 0 & 1.49 & 0 \\ Vehicles & 0 & 0 & 0 & 1.62 \end{pmatrix}$$

Settings for health outcome parameters were motivated by the Wellenius *et al.* (2003) investigation of PM effects on heart rate; $\alpha = 86$, $\beta_{HE} = 2$, and $\sigma_Y = 8$, yielding an effect size of $\delta = 0.25$.

To simulate exposure, we generated source contributions from $\boldsymbol{\eta} \sim MVN_4(\mathbf{0}, \boldsymbol{\Sigma})$. Then, given each set of source contributions, we simulated the elemental concentrations from $\mathbf{X}|\boldsymbol{\eta} \sim MVN_{13}(\boldsymbol{\Lambda}\boldsymbol{\eta}, \boldsymbol{\Psi})$. We generated health outcomes assuming a health effect from a single source. Specifically, for a given source k, the health outcome was simulated from the simple linear regression model,

$$Y_{kt} \sim N(\alpha + \beta_{HE} \times \eta_{kt}, \sigma_Y^2)$$

where η_{kt} is the contribution of source k at time t. For example, $\mathbf{y}_1$ is a vector of simulated health outcomes where the health effect is associated with the first factor, road dust. In the simulation, we generate the health effect on each of the four factors individually, such that for each set of exposures, we generate four sets of health outcomes, $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, and $\mathbf{y}_4$, where the health effect corresponds to the different pollution sources, road dust, power plants, oil combustion, and motor vehicles, respectively. We analyzed these simulated health outcomes separately.

Exposures, source contributions and elemental concentrations, were generated for $N \in \{20, 100\}$ days. Although the HSPH concentrated particle experiments typically do not run with 100 exposure days, we included this hypothetical scenario to confirm that any deficiencies of the ML SEM are due to a small number of exposure days. The health outcomes were generated for two animals per exposure day, for a total of $2N \in \{40, 200\}$ outcomes.

We obtained health effect estimates using five different strategies:

1. ***Known source contributions***: Although source contributions are not directly measured in the studies motivating this research, here we simulate them so that they are effectively known. We estimate the health effects based on the known source contributions,

$$Y_t = \alpha + \boldsymbol{\beta}^T \boldsymbol{\eta}_t + \epsilon_t^Y$$

   where $\epsilon_t^Y \overset{iid}{\sim} N(0, \sigma_Y^2)$.

2. ***Tracer approach***: We estimated the health effects based on the elemental concentrations of

the distinct set of $K = 4$ tracers,

$$Y_t = \alpha + \boldsymbol{\beta}^T \mathbf{x}_t^{(\mathsf{T})} + \epsilon_t^Y$$

where $\mathbf{x}_t^{(\mathsf{T})}$ is the (4×1) vector of simulated concentrations for the tracer elements silicon, sulfur, nickel, and organic carbon at time t.

3. ***Two-stage approach***: We first conducted a confirmatory factor analysis on all simulated elements, constraining the factor loadings for the tracer elements, silicon, sulfur, nickel, and organic carbon, according to the D1-D3 identifiability conditions. We then fit the health effects model on the estimated source contributions.

$$Y_t = \alpha + \boldsymbol{\beta}^T \widehat{\boldsymbol{\eta}_t} + \epsilon_t^Y$$

4. ***ML SEM***: We estimated the receptor and health outcome models (1) and (2) jointly using ML in Mplus (Muthen and Muthen 1998). This approach imposed the D1-D3 identifiability conditions, but did not use any historical exposure information.

5. ***Bayesian SEM***: We estimated the receptor and health outcome models jointly using an informative Bayesian approach. To obtain informative priors on the source profiles, we conducted a confirmatory factor analysis on a simulated historical dataset of $N = 200$ exposures, based on the D1-D3 identifiability constraints and tracers, silicon, sulfur, nickel, and organic carbon. We defined informative priors on the source profiles using (4), and set vague priors on the remaining parameters; $IG(0.01, 0.01)$ on $\{\boldsymbol{\Sigma}_{kk}\}$, $\{\boldsymbol{\Psi}_{pp}\}$, and σ_Y^2, and $N(0, 1000)$ on α and $\{\beta_k\}$. The Bayesian SEM was fit using the Markov chain Monte Carlo (MCMC) method in WinBUGS (Spiegelhalter, Thomas, and Best 2000); for each fit, we ran 25,000 iterations, discarding 20,000 as burn-in and thinning by five, for a total of 1,000 posterior samples for estimation and inference. We randomly checked convergence on multiple simulated data sets and saw evidence of good mixing and convergence in every case.

We ran 500 simulations to assess the statistical properties of the health effect estimates of a typical concentrator study with $N = 20$ days of exposure and an additional 500 simulations to evaluate

Table 3: Simulation Study: Health Effect Estimates (SEs) for $N = 20$

Method	Road Dust	Power Plants	Oil Combustion	Motor Vehicles
Known $\boldsymbol{\eta}$	1.94 (0.04)	2.04 (0.05)	1.99 (0.05)	1.99 (0.05)
Tracer	1.85 (0.04)	1.99 (0.05)	1.74 (0.05)	1.58 (0.05)
Two-Stage	1.91 (0.04)	1.93 (0.05)	1.66 (0.05)	1.83 (0.06)
ML SEM	1.92 (0.04)	1.97 (0.05)	1.77 (0.06)	1.89 (0.06)
Bayes SEM	1.90 (0.04)	2.02 (0.05)	1.91 (0.06)	1.95 (0.05)

the estimates of a hypothetical study with $N = 100$ exposure days. Tables 3 and 5 summarize the simulation results based on $N = 20$ days of exposure, and Tables 4 and 6 summarize the results for the study with $N = 100$ exposure days.

Table 3 and Table 4 display the health effect estimates, and corresponding simulation standard errors, obtained with the five different methodologies. Although the health effects were estimated with a model that included terms to represent all four sources, the tables present only the estimate for the source on which the health effect was simulated. For example, the first column in Tables 3 and 4 contains the health effect estimates corresponding to road dust, since this column reflects the analysis of the $\mathbf{y}_1$ outcome, where the health effect was simulated on the road dust source. The estimated coefficients for the other three sources, power plants, oil combustion, and motor vehicles, were always all approximately zero and, hence, are not included in the tables. In all cases, our estimates of the null coefficients were unbiased, and therefore, we display only the estimates for which the truth is $\beta = 2$.

The health effects estimates based on known source contributions represent the "gold standard." However, although source contributions are available in a simulation study, they are not observable in the studies motivating this research. Therefore, health effect estimates based on known $\boldsymbol{\eta}$ are unobtainable in practice and are provided for reference only.

The health effects estimates obtained with the tracer approach demonstrate the typical attenuation of effect associated with measurement error in this simple setting. In fact, we can calculate the at-

Table 4: Simulation Study: Health Effect Estimates (SEs) for $N = 100$

Method	Road Dust	Power Plants	Oil Combustion	Motor Vehicles
Known $\boldsymbol{\eta}$	2.06 (0.02)	2.00 (0.02)	2.00 (0.02)	1.97 (0.02)
Tracer	2.00 (0.02)	1.94 (0.02)	1.73 (0.02)	1.56 (0.02)
Two-Stage	2.05 (0.02)	1.99 (0.02)	1.98 (0.02)	1.94 (0.02)
ML SEM	2.05 (0.02)	1.99 (0.02)	1.98 (0.02)	1.95 (0.02)
Bayes SEM	2.09 (0.02)	1.98 (0.02)	1.96 (0.02)	1.90 (0.02)

tenuation factor γ associated with each tracer estimate, since we know the amount of measurement error associated with each of the tracer elements, quantified by $\psi_{Si}, \psi_S, \psi_{Ni}$, and ψ_{OC}. Because our simulations are based on a factor pattern with a unique tracer for each source, uncorrelated factors, and normality, for a given source k,

$$\gamma_k = \frac{\boldsymbol{\Sigma}_{kk}}{\boldsymbol{\Sigma}_{kk} + \boldsymbol{\Psi}_{kk}}.$$

Here, the attenuation factors are 0.97, 0.97, 0.87, and 0.78 for the road dust, power plants, oil combustion, and motor vehicles effects, respectively. In our simulation study, we are able to correct for the bias induced by measurement error and obtain reliable health effects estimates using the tracer approach. However, in practical settings, the variance parameters $\{\boldsymbol{\Sigma}_{kk}\}$ and $\{\boldsymbol{\Psi}_{kk}\}$ are typically unknown, and the non-negativity of source contributions violates the assumption of normality; given these limitations, correcting for measurement error induced bias is no longer straightforward.

Alternatively, the two-stage approach amounts to estimating the correction terms and adjusting the health effect estimates accordingly by using $\widehat{\boldsymbol{\eta}} = \widehat{E}(\boldsymbol{\eta}|data)$ in the health outcome model. In this way, the two-stage approach may be viewed as a form of regression calibration (Carroll, Ruppert, and Stefanski 1995). The simulation study demonstrates attenuation in the two-stage health effect estimates based on $N = 20$ exposure days; however, we attribute this bias to the small number of exposures. In the study based on $N = 20$ exposures, the receptor model failed to converge in 19 out of 500 (3.8%) simulations. However, in the study based on $N = 100$ days of exposure, all receptor models converged. Furthermore, the two-stage estimates based on $N = 100$ exposure days are all very similar to the estimates obtained with the known source contributions, and all estimates are within twice the simulation standard error of the truth.

Table 5: Simulation Study: Power (Size) Analysis for $N = 20$

Method	Road Dust	Power Plants	Oil Combustion	Motor Vehicles
Known $\boldsymbol{\eta}$	54% (0.056)	44% (0.047)	40% (0.049)	40% (0.039)
Tracer	49% (0.050)	43% (0.045)	34% (0.053)	35% (0.039)
Two-Stage	53% (0.095)	42% (0.081)	36% (0.097)	34% (0.087)
ML SEM	62% (0.133)	53% (0.129)	47% (0.133)	46% (0.122)
Bayes SEM	50% (0.052)	44% (0.051)	35% (0.050)	37% (0.040)

Table 6: Simulation Study: Power (Size) Analysis for $N = 100$

Method	Road Dust	Power Plants	Oil Combustion	Motor Vehicles
Known $\boldsymbol{\eta}$	100% (0.047)	99% (0.057)	98% (0.043)	99% (0.049)
Tracer	100% (0.049)	99% (0.055)	96% (0.058)	97% (0.055)
Two-Stage	100% (0.048)	99% (0.057)	98% (0.054)	98% (0.050)
ML SEM	100% (0.049)	99% (0.059)	98% (0.057)	98% (0.053)
Bayes SEM	100% (0.055)	98% (0.048)	97% (0.052)	98% (0.049)

The SEM approach appears to offer a clear advantage to the tracer and two-stage approaches, particularly in the case of a small number of exposures. However, in this small sample context, the SEM estimates are distinguished by the method of estimation. The health effects estimates obtained with the informative Bayesian SEM are most similar to those obtained with known source contributions, and are within twice the simulation standard error of the truth in almost all cases for $N = 20$. In contrast, the estimates obtained via ML appear to be biased downward, and the ML SEM estimate for oil combustion is well beyond twice the simulation standard error from the truth. As in the case of the two-stage approach, we attribute these deficiencies in the ML SEM to the small number of exposures. In the study based on $N = 20$ exposure days, the ML method failed to converge in approximately 4% of the simulations. (The number of failures are 22, 22, 20, and 21 for the analysis of $\mathbf{y}_1$, $\mathbf{y}_2$, $\mathbf{y}_3$, and $\mathbf{y}_4$, respectively.) However, when we increase the number of exposures to $N = 100$, all 500 simulations converged and the ML SEM performs almost exactly the same as the "gold standard" that uses known $\boldsymbol{\eta}$.

Tables 5 and 6 provide the estimated power, defined as the proportion of 95% confidence (credible)

Table 7: Robustness Study: Health Effect Estimates (SEs) for $N = 20$

Method	Road Dust	Power Plants	Oil Combustion	Motor Vehicles
Known $\boldsymbol{\eta}$	2.02 (0.07)	1.94 (0.05)	2.04 (0.05)	2.10 (0.05)
Tracer	1.96 (0.07)	1.90 (0.04)	1.81 (0.05)	1.65 (0.05)
Bayes SEM (D1-D3)	2.09 (0.07)	1.98 (0.05)	1.96 (0.06)	2.03 (0.05)
Bayes SEM (C1-C3)	2.11 (0.07)	1.96 (0.05)	2.00 (0.06)	2.05 (0.05)

Table 8: Robustness Study: Power (Size) Analysis for $N = 20$

Method	Road Dust	Power Plants	Oil Combustion	Motor Vehicles
Known $\boldsymbol{\eta}$	54% (0.047)	41% (0.052)	40% (0.051)	45% (0.059)
Tracer	51% (0.053)	38% (0.061)	34% (0.054)	35% (0.062)
Bayes SEM (D1-D3)	49% (0.045)	41% (0.066)	36% (0.057)	42% (0.058)
Bayes SEM (C1-C3)	50% (0.047)	40% (0.063)	38% (0.057)	42% (0.060)

intervals that do not contain the null value of zero when $\beta_{HE} = 2$, as well as the estimated size, the proportion of 95% confidence (credible) intervals that do not contain a true value of $\beta^*_{HE} = 0$. In the study based on $N = 20$ exposure days, the two-stage approach and the ML SEM exceed the expected size of 0.05 in all cases, indicating that these approaches are too liberal when the number of exposures is limited. For methods of approximately the same size (excluding the two-stage and ML SEM), the Bayesian SEM is comparable to the "gold standard" based on known source contributions and has virtually the same sensitivity for detecting a true effect as the tracer approach. In the study based on $n = 100$ exposure days, the two-stage approach and ML SEM are of the appropriate size (≈ 0.05), and all methods are very powerful ($> 95\%$) at detecting a true effect in this setting.

Finally, we conducted an additional simulation study designed to investigate the impact of choosing incorrect identifiability constraints. One could argue that setting a single factor loading to zero when it is really greater than zero should have little impact on the resulting health effect estimates, whereas incorrectly setting $K \times (K-1)$ loadings equal to zero may collectively have a larger effect. To check the impact of this misspecification, we conducted a simulation study assuming all loadings were nonzero, replacing the zero loadings in the previous simulation study to randomly generated values from a uniform(0,0.2) distribution. For each simulated data set, we estimated the health

effects using known source contributions, the tracer approach, the informative Bayesian SEM fit with the D1-D3 conditions, and the informative Bayesian SEM fit with the less restrictive C1-C3 identifiability conditions. Table 7 presents the source-specific health effect estimates and standard errors. Table 8 provides the corresponding power and size. These results are consistent with the results from the previous simulations for $N = 20$ exposure days, in that the Bayesian SEM approaches yield estimates similar to those obtained if the true source contributions are known and tests of the appropriate size. Thus, the second study suggests that the proposed identifiability constraints do not have a large impact on inference as long as the loadings set to zero are not much larger than 0.2.

6 Data Analysis

6.1 Joint Model

In this section, we implement our informative Bayesian SEM to analyze the source-specific PM health effects on myocardial ischemia in dogs (Wellenius *et al.* 2003). Analyses of the Wellenius data did not detect any pairing or period (day) effects, but did suggest there may be a carryover effect from the CAPs exposure. As a result, the authors excluded Sham exposures following the CAPs exposure in their analyses. The final analysis was based on a total of 43 measured health outcomes, corresponding to 18 CAPs and 25 Sham exposures.

Wellenius *et al.* (2003) found a large amount of variability within each dog by cycle combination and therefore included a random effect for sequence (see Table 1). The authors estimated the source-specific health effects of PM with the following model,

$$Y_{td} = \alpha + \boldsymbol{\beta}^T \mathbf{x}_{td}^{(\mathrm{T})} + b_t + \epsilon_{td}^Y$$

where Y_{td} is log(peak ST-segment) and $\mathbf{x}_{td}^{(\mathrm{T})}$ is the vector of elemental concentrations for silicon, sulfur, nickel, and black carbon for sequence t and day d, b_t is the random sequence effect, $b_t \overset{iid}{\sim} N(0, \sigma_b^2)$, $\epsilon_{td}^Y \overset{iid}{\sim} N(0, \sigma_Y^2)$, and $b \perp \epsilon^Y$.

Accordingly, we fit the following informative SEM model,

$$\mathbf{X}_t = \boldsymbol{\Lambda}\boldsymbol{\eta}_t + \boldsymbol{\epsilon}_t^{\mathbf{X}}$$

$$Y_{td} = \alpha + \boldsymbol{\beta}^T(I_{CAPS_{td}} \times \boldsymbol{\eta}_t) + b_t + \epsilon_{td}^Y$$

where $\boldsymbol{\epsilon}_t^{\mathbf{X}} \overset{iid}{\sim} MVN_P(\mathbf{0}, \boldsymbol{\Psi})$ for diagonal $\boldsymbol{\Psi}$, $b_t \overset{iid}{\sim} N(0, \sigma_b^2)$, $\epsilon_{td}^Y \overset{iid}{\sim} N(0, \sigma_Y^2)$, and $b \perp \epsilon^Y$. $I_{CAPS_{td}}$ is an indicator of CAPs exposure on day d in sequence t. This indicator provides for the Sham exposures and operates on the assumption that all source contributions are null in filtered air; i.e. if the dog in sequence t is exposed to Sham on day d, $I_{CAPS_{td}} = 0$, and

$$Y_{td} = \alpha + b_t + \epsilon_{td}^Y.$$

Based on this specification, the receptor model applies to the CAPs exposures only, while the health effects regression is fit on all outcomes. Finally, to respect the non-negativity of source contributions, we truncate the normal distribution on the latent variables,

$$\boldsymbol{\eta}_t \overset{iid}{\sim} MVN_K(\boldsymbol{\mu}, \boldsymbol{\Sigma})I(\boldsymbol{\eta}_t \geq \mathbf{0}).$$

As noted in Section 3.2, we also fit the model assuming lognormal source contributions to assess the sensitivity of our conclusions to distributional assumption on $\boldsymbol{\eta}$.

6.2 Prior Elicitation

To obtain prior information on the source profiles, we fit a Bayesian confirmatory factor analysis on the scaled historical data ($N = 160$). We assumed the $K = 4$ major sources of PM as described in Table 2, and we assumed that these sources are independent. To be consistent with the Wellenius *et al.* (2003) analysis, here we chose silicon to identify road dust, sulfur for power plants, nickel for oil combustion, and black carbon for motor vehicles.

It is thought that BC is a better marker of motor vehicles than OC. However, exploratory factor analyses consistently estimated moderate loadings for BC on several factors. Therefore, in our

analysis, we apply the more flexible C1-C3 identifiability conditions described in Section 3.3. According to these conditions, we need to constrain $K - 1 = 3$ loadings to zero on each profile, while ensuring a distinct set of constraints for each source. In order to set meaningful constraints, we consulted our exploratory results and identified distinct sets of three near zero (< 0.2) loadings per factor. We constrain to zero the following factor loadings: sulfur, nickel, and elemental carbon on road dust; silicon, vanadium, and organic carbon on power plants; aluminum, sulfate, and organic carbon on oil combustion; and aluminum, sulfate, and nickel on motor vehicles.

We fit the confirmatory factor analysis to the historical data using MCMC in WinBUGS (Spiegelhalter, Thomas, and Best 2000). We set vague priors on all parameters, specifying $IG(0.01, 0.01)$ on $\{\mathbf{\Sigma}_{kk}\}$ and $\{\mathbf{\Psi}_{pp}\}$, $N(0, 1000)$ on $\{\mu_k\}$, and, following Park, Guttorp, and Henry (2001), we truncate normal distributions for the unconstrained $\{\lambda_{pk}\}$,

$$\lambda_{pk} \sim N(0, 10000)I(\lambda_{pk} \geq 0),$$

since negative components of source profiles are not interpretable. We ran 25,000 iterations, discarding 20,000 as burn-in and thinning by five, for a total of 1,000 posterior samples for estimation. Evaluation of autocorrelation and trace plots supported convergence. Table 9 displays the posterior source profiles.

Finally, since meteorology is likely to impact the dynamics of PM, one might expect source profiles to differ depending on weather conditions, for example warm weather versus cold weather. As shown in Table 1, all of the Wellenius *et al.* (2003) exposures were collected in months of typically cold weather. The historical data, on the other hand, consists of exposures collected year-round. To ensure that our prior information on the source profiles is relevant to the Wellenius *et al.* (2003) study, we conducted the identical confirmatory factor analysis on the subset of historical data consisting of exposures collected in the months of September through March ($N = 85$). The posterior profiles estimated on the restricted set of exposures were very similar to those estimated on the complete set of historical data; we therefore used all historical data to define informative priors on the source profiles.

Table 9: Posterior Source Profiles

Element	$\widehat{\boldsymbol{\lambda}}_{RD}$	$\widehat{\boldsymbol{\lambda}}_{PP}$	$\widehat{\boldsymbol{\lambda}}_{OC}$	$\widehat{\boldsymbol{\lambda}}_{MV}$
Si	1*	0*	0.05	0.26
S	0*	1*	0.02	0.04
Ni	0*	0.03	1*	0*
OC	0.04	0*	0*	1.48
Al	1.00	0.01	0*	0*
Ti	0.96	0.02	0.18	0.51
Ca	0.95	0.02	0.21	0.32
SULF	0.01	0.99	0*	0*
Se	0.02	0.66	0.04	0.38
V	0.09	0*	1.04	0.04
Br	0.10	0.54	0.18	0.81
BC	0.04	0.44	0.42	1*
EC	0*	0.29	0.42	1.28

* denotes constrained parameters

6.3 Health Effects Analysis

We fit the informative Bayesian SEM to the Wellenius *et al.* (2003) data using MCMC in WinBUGS (Spiegelhalter, Thomas, and Best 2000). We defined informative priors on the source profiles based on the posterior means in Table 9 and posterior covariances from the Bayesian factor analysis of the historical data. Here again, we truncated the normal distribution for the factor loadings,

$$\boldsymbol{\lambda} \sim MVN_{K(P-K)}(\widehat{\boldsymbol{\lambda}}^{(hist)}, \widehat{Var}(\widehat{\boldsymbol{\lambda}}^{(hist)}))I(\boldsymbol{\lambda} \geq \mathbf{0}).$$

We set vague priors on all other parameters, $IG(0.01, 0.01)$ on $\{\boldsymbol{\Sigma}_{kk}\}$, $\{\boldsymbol{\Psi}_{pp}\}$, σ_Y^2, and σ_b^2, and $N(0, 1000)$ on $\{\mu_k\}$, α, and $\{\beta_k\}$. We ran 25,000 iterations, discarding 20,000 as burn-in and thinning by five, for a total of 1,000 posterior samples. We examined diagostic trace and autocorrelaion plots and found satisfactory convergence.

Table 10 displays the posterior means, standard errors, and 95% credible intervals for the source-specific health effects estimated with the informative Bayesian SEM. For each pollution source, the health effect estimate is on the scale of the element whose factor loading is constrained to 1. For example, we interpret the health effect estimate of road dust as the change in log peak ST-segment

Table 10: Bayesian Structural Equation Results

Source	$\hat{\beta}$	$\hat{se}(\hat{\beta})$	95% Credible Interval	90% Credible Interval
Road Dust	0.154*	0.063	(0.030 , 0.276)	(0.048 , 0.252)
Power Plants	-0.071	0.072	(-0.217 , 0.072)	(-0.188 , 0.047)
Oil Combustion	-0.034	0.071	(-0.167 , 0.115)	(-0.144 , 0.084)
Motor Vehicles	0.062	0.141	(-0.217 , 0.341)	(-0.160 , 0.293)

Table 11: Source Tracer Results

Tracer	$\hat{\beta}$	$\hat{se}(\hat{\beta})$	95% Confidence Interval	90% Confidence Interval
Silicon	0.137*	0.048	(0.037 , 0.237)	(0.055 , 0.220)
Sulfur	-0.063	0.079	(-0.227 , 0.100)	(-0.199 , 0.072)
Nickel	-0.036	0.069	(-0.178 , 0.107)	(-0.154 , 0.082)
Black Carbon	0.014	0.088	(-0.170 , 0.198)	(-0.138 , 0.166)

associated with an increase in the *contribution of road dust* on the scale of one standard deviation increase in the concentration of silicon. Table 11 presents the corresponding source-specific health effect estimates obtained from the tracer analysis of the scaled data. We interpret the health effect estimates obtained with the tracer approach as the change in log peak ST-segment associated with one standard deviation increase in the concentration of corresponding tracer. For instance, here we interpret the health effect estimate of road dust as the change in log peak ST-segment associated with one standard deviation increase in the *concentration of silicon*. The estimates provided in Table 11 do not precisely correspond to those reported in Wellenius *et al.* (2003), which were based on ECG recordings from two precordial leads. Since readings from the two leads were highly correlated ($r > 0.8$), we restrict our analysis to log peak ST-segment recorded on a single lead (V5).

In accordance with Wellenius *et al.* (2003), we found a significant effect of resuspended road dust on myocardial ischemia in dogs. Implementing our informative Bayesian SEM, we estimated an increase of 0.154 in log peak ST-segment associated with an increase in road dust contribution on the scale of one standard deviation of silicon, with 95% credible interval (0.030,0.276). We did not find a significant change in log peak-ST segment associated with pollution from coal-fired power plants, oil combustion for home heating, or motor vehicle exhaust.

Thus, the informative Bayesian structural equation results suggest that the conclusions in the original analysis were not driven by unequal amounts of measurement error associated with the tracer representations of the four pollution sources. Given that the SEM analysis yields an estimate of the road dust coefficient similar to the silicon coefficient in the tracer analysis, one might wonder whether the more complex analysis was worth the effort. Then again, the SEM analysis does more than reinforce the significance of the road dust effect; it also reaffirms the lack of evidence of an association between this cardiac outcome and pollution from the other sources, in particular motor vehicles. Not surprisingly given the results of our simulation study, the estimated health effect estimate for motor vehicles, $\hat{\beta}_{MV} = 0.062$, is more than four times greater than the corresponding attenuated tracer estimate for black carbon, $\hat{\beta}_{BC} = 0.014$; however, this estimated association remains insignificant. Confirmation of a non-significant effect of motor vehicles is equally important in this setting.

To ensure that our conclusions are not sensitive to distributional assumptions, we re-ran the informative Bayesian SEM under the specification of lognormal source contributions. We estimated $\hat{\beta}_{RD} = 0.164$, $\hat{\beta}_{PP} = -0.066$, $\hat{\beta}_{OC} = -0.030$, and $\hat{\beta}_{MV} = 0.026$, and detected a significant effect of road dust only. Thus, both sets of distributional assumptions yield findings that agree with those in Wellenius *et al.* (2003).

7 Discussion

In this paper, we considered methods to assess source-specific health effects of complex mixtures of PM pollution. One objective was to evaluate the statistical properties of estimates obtained with methods currently used in practice, the tracer approach and the two-stage approach, for multivariate pollution patterns typical of Boston aerosol. In a simulation study, we showed that the health effect estimates obtained using the tracer approach are attenuated, both in small and large sample cases, which was expected having framed the problem from a measurement error perspective. Our results suggest that the ability of the tracer approach to detect source-specific effects will vary by source, due to differing amounts of measurement error associated with the tracers of the different sources. In particular, the common marker for traffic particles, black carbon, has a relatively large degree of

error associated with it, which may reflect a regional component of black carbon in addition to local traffic in the Boston area; as such, the tracer approach may underestimate the true effect associated with motor vehicles. The two-stage approach is similarly susceptible to bias, although only in the case of small samples. For large samples, the two-stage estimates appear unbiased as we would expect of a regression calibration.

As an alternative to the tracer and two-stage approaches, we proposed a structural equation model to account for the uncertainty associated with latent source contributions, along with a Bayesian approach to model fitting. This approach leverages exposure information from previous related concentrator studies. Simulations suggest that the proposed informative Bayesian SEM is effective in eliminating bias in estimated source-specific health effect estimates, even when the number of exposures is limited. We demonstrated the flexibility of the Bayesian approach to accommodate complex study designs and non-normality in our analysis of the Wellenius *et al.* (2003) study. As an added advantage, the informative Bayesian SEM may be implemented in freely available software, such as the WinBUGS package (Spiegelhalter, Thomas, and Best 2000).

Our findings in this paper have implications for the design of future PM concentrator studies. The results demonstrate the benefits of using exposure data from existing, relevant exposure studies where possible. However, not all studies have the benefit of such prior knowledge. When historical data are not available, investigators should maximize the number of unique exposure days, subject to cost constraints. A large number of exposure days will allow one to use two-stage models that employ source apportionment techniques to address errors in the tracer characterization of the pollution sources. Unfortunately, because even one run of a concentrator exposure can be costly, such "many-exposure" designs may be prohibitive. In cases where both historic and current exposure information is limited, the tracer approach is preferable to the two-stage analysis since the latter may be unstable and experience convergence problems; however, one should take care to choose "good" tracers that minimize the error associated with these surrogates. In these settings, the tracer approach is still useful in screening for source-specific health effects, but is likely to yield attenuated estimates of effects.

The purpose of this article was to assess the performance of methods for source characterization in concentrator studies. We therefore focused on a specific form of a factor analysis model in the structural equations framework, and have not addressed all of the interesting modeling issues that arise in the development of a "good" receptor model. For instance, although the estimation of the number of sources can often be challenging (Park, Speigelman, and Henry 2002), we assumed that we have good prior knowledge on the number of major pollution sources in the Boston area. This assumption is probably reasonable in our setting, as the pollution mixture in this area has been extensively studied for almost two decades (Koutrakis and Spengler 1987; Oh *et al.* 1997; Oh 2000). We note that existing exposure studies suggest that there exists an additonal pollution source in the Boston aerosol comprised of sodium and chlorine, often referred to as sea salt. However, from a regulatory perspective this exposure is not of primary importance, and hence of less interest in particulate matter health studies. Secondly, the majority of the source apportionment literature for exposure assessment of particulate matter (Park, Guttorp, and Henry 2001) assumes that the sources of exposure are independent. To maintain consistency with existing methodology, we specified independent priors on the latent source contributions in the implementation of our Bayesian SEM. However, while source-specific exposures are assumed to be independent a priori, the Bayesian approach uses data to update the priors, thus allowing for correlation in the posterior distributions of the source contributions. Given that we might expect source contributions to be correlated, at least in part due to meterologic conditions, this is an appealing feature of our approach. Furthermore, the methods proposed in this paper extend naturally to account for systematic factors, like meteorology, that are likely to affect pollution levels from the different sources similarly (see for instance, Gryparis *et al.* 2005), by allowing the means of the unobserved sources to depend on covariates.

8 Acknowledgements

This research was supported by NIEHS grant ES07142 (MCN), American Chemistry Council grant 2843 and NIEHS grant ES012044 (BAC), and NIH grant ES012972 (JJG).

9 References

1. Bandeen-Roche, K. (1994). Resolution of additive mixtures into source components and contributions: A compositional approach. *Journal of the American Statistical Association* **89**, 1450–1458.

2. Batalha, J. R. F., Saldiva, P. H. N., Clarke, R. W., Coull, B. A., Stearns, R. C., Lawrence, J., Murthy, G. G. K., Koutrakis, P., and Godleski, J. J. (2002). Concentrated ambient air particles induce vasoconstriction of small pulmonary arteries in rats. *Environmental Health Perspectives* **110**, 1191–1197.

3. Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

4. Budtz-Jorgensen, E., Keiding, N., Grandjean, P., Weihe, P., and White, R. F. (2003). Statistical methods for the evaluation of health effects of prenatal mercury exposure. *Environmentrics* **14**, 105–120.

5. Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*. London: Chapman & Hall.

6. Clarke, R. W., Coull, B. A., Reinisch, U., Catalano, P., Killingsworth, C. R., Koutrakis, P., Kavouras, I., Murthy, G. G. K., Lawrence, J., Lovett, E., Wolfson, J. M., Verrier, R. L., and Godleski, J. J. (2000). Inhaled concentrated ambient particles are associated with hematologic and bronchoalveolar lavage changes in canines. *Environmental Health Perspectives* **12**, 1179–1187.

7. Coull, B. A., Catalano, P. J., and Godleski, J. J. (2000). Semiparametric analyses of cross-over data with repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics* **5**, 417–429.

8. Diggle, P. J., Heagerty, P., Liang, K. Y., and Zeger, S. L. (2002). *The Analysis of Longitudinal Data: 2nd Edition*. Oxford, England: Oxford University Press.

9. Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E. (1993). An association between air pollution and mortality in six US cities. *New England Journal of Medicine* **329**, 1753–1759.

10. Dominici, F., Daniels, M., Zeger, S. L., and Samet, J. M. (2002). Air pollution and mortality: estimating regional and national dose-response relationships. *Journal of the American Statistical Association* **97**, 100–111.

11. Dominici, F., Sheppard, L., and Clyde, M. (2003). Health effects of air pollution: A statistical review. *International Statistical Review* **71**, 243–276.

12. Godleski, J. J., Clarke, R. W., Coull, B. A., Saldiva, P. H. N., Jiang, N.-F., Lawrence, J., and Koutrakis, P. (2002). Composition of inhaled urban air particles determines acute pulmonary responses. *Annals of Occupational Hygiene* **46**, 419–424.

13. Godleski, J. J., Verrier, R. L., Koutrakis, P., Catalano, P., Coull, B. A., Reinisch, U., Lovett, E. G., Lawrence, J., Murthy, G. G. K., Wolfson, J. M., Clarke, R. W., and Nearing, B. D. (2000). Mechanisms of morbidity and mortality from exposure to ambient air particulate. *Health Effects Institute Reserach Report* **91**, 1–103.

14. Gryparis, A., Coull, B. A., Schwartz, J., and Suh, H. H. (2005). Semiparametric latent variable regression models for spatial-temporal modeling of mobile source particles in the Greater Boston area. Submitted for publication. Technical report. `http://www.bepress.com/harvardbiostat/`.

15. Hopke, P. K. (2003). Recent developments in receptor modeling. *Journal of Chemometrics* **17**, 255–265.

16. Kavouras, I. G., Koutrakis, P., Cereceda-Balic, F., and Oyola, P. (2001). Source apportionment of PM10 and PM2.5 in five Chilean cities using factor analysis. *Journal of the Air & Waste Management Association* **51**, 451–464.

17. Kim, E., Hopke, P. K., Larson, T. V., and Covert, D. S. (2004). Analysis of ambient particle size distributions using unmix and positive matrix factorization. *Environmental Science & Technology* **38**, 202–209.

18. Koutrakis, P. and Spengler, J. D. (1987). Source apportionment of ambient particles in Steubenville, Ohio using specific rotation factor analysis. *Atmospheric Environment* **21**, 1511–1519.

19. Laden, F., Neas, L. M., Dockery, D. W., and Schwartz, J. (2000). Association of fine particulate matter from different sources with daily mortality in six U. S. cities. *Environmental Health Perspectives* **108**, 941–947.

20. Lawrence, J., Wolfson, J. M., Ferguson, S., Koutrakis, P., and Godleski, J. J. (2004). Performance stability of the Harvard ambient particle concentrator. *Aerosol Science and Technology* **38**, 219–227.

21. Lippmann, M., Frampton, M., Schwartz, J., Dockery, D., Schlesinger, R., Koutrakis, P., Froines, J., Nel, A., Finkelstein, J., Godleski, J., Kaufman, J., Koenig, J., Larson, T., Luchtel, D., Liu, L. J. S., Oberdorster, G., Peters, A., Sarnat, J., Sioutas, C., Suh, H., Sullivan, J., Utell, M., Wichmann, E., and Zelikoff, J. (2003). The US Environmental Protection Agency particulate matter health effects research centers program: A midcourse report of status, progress, and plans. *Environmental Health Perspectives* **111**, 1074–1092.

22. Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press.

23. Muthen, L. K. and Muthen, B. O. (1998). *Mplus User's Guide (3rd Edition)*. Los Angeles, CA: Muthen & Muthen.

24. Oh, J. A. (2000). Characterization and source apportionment of air pollution in Nashville, TN and Boston, MA. Doctoral Thesis. Department of Environmental Health, Harvard School of Public Health.

25. Oh, J. A., Suh, H. H., Lawrence, J. E., Allen, G. A., and Koutrakis, P. (1997). Characterization of particulate mass concentrations in South Boston, MA. Proceedings of AWMA/EPA Symposium on "Measurement of Toxic and Related Air Pollutants", April 29-May 1, 1997, Research Triangle Park, NC. AWMA publication number VIP-74 (Pittsburgh, PA), 397–407.

26. Paatero, P., and Tapper, U. (1994). Positive Matrix Factorization - A nonnegative factor model with optimal utilization of error-estimates of data values. *Environmetrics* **5**, 111–126.

27. Park, E. S., Guttorp, P., and Henry, R. C. (2001). Multivariate receptor modeling for temporally correlated data by using MCMC. *Journal of the American Statistical Assocation* **96**, 1171–1183.

28. Park, E. S., Spiegelman, C. H., and Henry, R. C. (2002). Bilinear estimation of pollution source profiles and amounts by using multivariate receptor models. *Environmetrics* **13**, 775–798.

29. R Development Core Team. (2003). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. `http://www.R-project.org`.

30. Roberts, K., Ryan, L. M., and Wright, R. J. (2003). On the use of Rasch models for handling high-dimensional covariates in epidemiological studies. Technical Report. Department of Biostatistics, Harvard School of Public Health. Under revision for *Biometrics*.

31. Seigneur, C., Pai, P., Hopke, P. K., and Grosjean, D. (1999). Modeling atmospheric: Particulate matter. *Environmental Science & Technology* **33**, 80A–86A.

32. Sioutas, C., Koutrakis, P., and Burton, R. M. (1995). A technique to expose animals to concentrated fine ambient aerosols. *Environmental Health Perspectives* **103**, 172–177.

33. Sioutas, C., Koutrakis, P., Ferguson, S. T., and Burton, R. M. (1995). Development and evaluation of a prototype ambient particle concentrator for inhalation exposure studies. *Inhalation Toxicology* **7**, 633–644.

34. Spiegelhalter, D., Thomas, A., and Best, N. (2000). *WinBUGS Version 1.3. User's Manual*. MRC Biostatistics Unit. Institute of Public Health, Cambridge. `http://www.mrc-bsu.cam.ac.uk/bugs`.

35. Tsiatis, A. A., De Gruttola, V., and Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error; applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Assocation* **90**, 27–37.

36. Wellenius, G. A., Coull, B. A., Godleski, J. J., Koutrakis, P., Okabe, K., Savage, S. T., Lawrence, J. E., Murthy, G. G. K., and Verrier, R. L. (2003). Inhalation of concentrated ambient air particles exacerbates myocardial ischemia in conscious dogs. *Environmental Health Perspectives* **111**, 402–408.