Biometrics 70, 164–174
March 2014

DOI: 10.1111/biom.12125

Estimating Acute Air Pollution Health Effects from Cohort Study Data

Adam A. Szpiro,[1,*] Lianne Sheppard,[2] Sara D. Adar,[3] and Joel D. Kaufman[4]

[1]Department of Biostatistics, University of Washington, Seattle, Washington, U.S.A.
[2]Departments of Biostatistics and Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, U.S.A.
[3]Department of Epidemiology, University of Michigan, Ann Arbor, Michigan, U.S.A.
[4]Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, U.S.A.
*email: aszpiro@u.washington.edu

Summary. Traditional studies of short-term air pollution health effects use time series data, while cohort studies generally focus on long-term effects. There is increasing interest in exploiting individual level cohort data to assess short-term health effects in order to understand the mechanisms and time scales of action. We extend semiparametric regression methods used to adjust for unmeasured confounding in time series studies to the cohort setting. Time series methods are not directly applicable since cohort data are typically collected over a prespecified time period and include exposure measurements on days without health observations. Therefore, long-time asymptotics are not appropriate, and it is possible to improve efficiency by exploiting the additional exposure data. We show that flexibility of the semiparametric adjustment model should match the complexity of the trend in the health outcome, in contrast to the time series setting where it suffices to match temporal structure in the exposure. We also demonstrate that pre-adjusting exposures concurrent with the health endpoints using trends in the complete exposure time series results in unbiased health effect estimation and can improve efficiency without additional confounding adjustment. A recently published article found evidence of an association between short-term exposure to ambient fine particulate matter ($PM_{2.5}$) and retinal arteriolar diameter as measured by retinal photography in the Multi-Ethnic Study of Atherosclerosis (MESA). We reanalyze the data from this article in order to compare the methods described here, and we evaluate our methods in a simulation study based on the MESA data.

Key words: Air pollution; Environmental epidemiology; Generalized least squares; Mixed models; Semiparametric regression; Time series; Unmeasured confounding.

1. Introduction

Epidemiologic evidence demonstrates an association between exposure to fine particulate matter ($PM_{2.5}$) air pollution and adverse health effects. Since air pollution is a modifiable risk factor, it is important to accurately estimate the magnitude of health effects and understand their mechanisms and time scales (Brook et al., 2010). The Environmental Protection Agency (EPA) has a legislative mandate to set standards for short-term and long-term air pollution levels to protect human health. Epidemiologic evidence plays a central role in establishing the scientific basis for these regulations (Environmental Protection Agency, 2006).

Short-term air pollution exposure on a time scale of hours or days is most likely associated with acute or transient health outcomes. A traditional approach to assessing the acute impact of short-term exposure uses population outcomes such as hospitalization or mortality rates in time series studies (Schwartz, 1994; Sheppard et al., 1999; Samet et al., 2000; Dominici, McDermott, and Hastie, 2004). Other designs used for this purpose include case-crossover studies (Janes, Sheppard, and Lumley, 2005), which are closely related to time series methods, and panel studies in which a small cohort of individuals are followed longitudinally (Dominici, Sheppard, and Clyde, 2003; Janes, Sheppard, and Shepherd, 2008).

Air pollution cohort studies have focused primarily on the cross-sectional effect of long-term air pollution exposure on chronic health outcomes (Dockery et al., 1993; Pope et al., 2002; Miller et al., 2007). Long-term exposure could refer to a subject's entire lifetime or to a period on the order of a year or more. There is growing interest in exploiting cohort data to estimate associations between short-term exposure and acute health effects in order to better understand the biological mechanisms by which air pollution causes disease.

We consider the recently published analysis by Adar et al. (2010) of the association between $PM_{2.5}$ exposure and retinal microvasculature as a marker of subclinical cardiovascular disease in the Multi-Ethnic Study of Atherosclerosis (MESA). While Adar et al. (2010) evaluate both chronic and acute effects on retinal arteriolar and vascular outcomes, we focus on the acute association with daily air pollution exposure. The dominant $PM_{2.5}$ variability is temporal rather than spatial, so we follow the approach in Adar et al. (2010) and treat exposure as a spatially homogeneous time series within metropolitan areas.

The exposure and the outcome can have seasonal and meteorological trends, so we need to control for shared sources of temporal variability to estimate unconfounded associations with air pollution. A methodology developed for time series studies is to include semiparametric spline terms in a regression model to adjust for temporal confounding. For this approach to be effective, we need to ensure that the spline terms contain sufficient degrees of freedom (df) to fully adjust for the temporal structure. A number of methods have been proposed for selecting df in time series studies (Dominici et al., 2004; Peng, Dominici, and Louis, 2006), but the existing literature does not address the implications for cohort study data.

The first objective of this article is to adapt the semiparametric regression methodology to cross-sectional cohort studies. The theory does not carry over directly for a number of reasons: (i) the relevant asymptotics are different, as in a cohort study we are concerned with large n asymptotics corresponding to a large number of subjects, whereas in time series studies the interest is in large T asymptotics corresponding to long study time periods; (ii) there can be multiple or no health observations on a given day, in contrast to a time series study where a single population-level health outcome is available on each day in a given geographic region; (iii) different assumptions about sources of randomness in the exposure may be appropriate for the two study designs; (iv) inter-subject variability makes it more difficult to accurately identify the seasonal and meteorological trends in cohort health data than in time series data; and (v) we need to be concerned with subject-specific covariates in cohort data such as blood pressure that could have their own temporal trends.

Our second objective is to propose a more efficient alternative to semiparametric regression. Since cohort study data often include air pollution measurements on days without health outcomes, semiparametric regression does not utilize all of the available exposure data. An alternative is to pre-adjust the exposure for temporal variability due to seasonality or meteorology and then use this modified exposure to estimate an unconfounded effect by ordinary least squares (OLS) or generalized least squares (GLS), without further adjustment in the disease model. Similar ideas have been considered for time series studies, but it is not clear that there is an advantage in that setting since the conventional approach already utilizes all of the available exposure data (Fung et al., 2003).

We summarize the data and findings from Adar et al. (2010) in Section 2, and we introduce notation and describe our statistical framework in Section 3. In Section 4, we formalize the semiparametric regression methodology for cohort studies and discuss the required number of df to obtain unbiased effect estimates and valid standard errors. In Section 5 we describe the pre-adjustment methodology. We illustrate our findings with a simulation study in Section 6 and reanalyze the retinal arteriolar data from MESA in Section 7. We conclude in Section 8 with a discussion, including guidance on when pre-adjustment followed by OLS or GLS is preferable to semiparametric regression.

2. Retinal Arteriolar Diameter and Air Pollution

A recently published analysis of the MESA cohort found evidence of an association between decreased retinal arteriolar diameter and elevated exposure to $PM_{2.5}$ air pollution on the previous day (Adar et al., 2010). As discussed by Adar et al. (2010), previous studies have found that changes in the microvasculature, including retinal arteriolar diameter, are associated with increased risk of myocardial infarction, stroke, and cardiovascular mortality, independent of other traditional risk factors. Therefore, these findings provide support for the hypothesis that reported associations between air pollution and the development and exacerbation of clinical cardiovascular disease are related to microvascular phenomena.

MESA is a prospective cohort study designed to examine the progression of subclinical cardiovascular disease (CVD). It enrolled 6814 men and women 45–84 years of age who were free of clinical CVD at entry from six U.S. communities in Baltimore, Chicago, Los Angeles, New York, Minneapolis-St. Paul, and Winston-Salem. Details of the sampling, recruitment, and data collection are described by Bild et al. (2002). The MESA cohort provides an excellent infrastructure for assessing the relationship between air pollution and various indicators of cardiovascular disease, particularly within the framework of MESA Air, an ancillary study to MESA funded by the EPA that includes collection of additional air quality monitoring and health endpoint data (Kaufman et al., 2012).

Retinal arteriolar diameter, a marker of microvasculature phenomena, was measured in MESA participants by retinal photography. Retinal photography was performed during the second MESA examination between August 2002 and January 2004. A total of 6176 individuals had retinal photographs taken, and 4607 subjects had complete data for inclusion in this analysis. Retinal arteriolar diameters within an area equal to 0.5–1 disc diameters from the optic disc margin are summarized as central retinal arteriolar equivalents (CRAE).

Air pollution exposures on the day prior to retinal photography were assigned based on the area-wide average concentrations from EPA Air Quality System (AQS) monitoring stations with complete time series during the period of interest. In light of the complex topography in the Los Angeles basin, the analysis incorporated four sub-regions: coastal Los Angeles, downtown Los Angeles, Riverside, and the area between Los Angeles and Riverside, giving a total of nine separate regions in our analysis. The data in Figure 2 show clear evidence of temporal trends in meteorology and $PM_{2.5}$ concentrations. Inter-subject variability makes it difficult to determine if there are temporal trends in CRAE measurements, but as noted by Adar et al. (2010) there is scientific reason to believe such trends are present. In a multivariate linear model with a full suite of subject-specific covariates and semiparametric adjustment for season with 12 df per year in each region and for meteorology with 6 df in each region, Adar et al. (2010) found a -0.4 μm (95% CI -0.8 to 0.1) decrease in CRAE per 10 $\mu g/m^3$ increase in the previous day's $PM_{2.5}$ concentration.

The analytic approach used by Adar et al. (2010) was chosen to be consistent with standard practice in air pollution time series studies. The present work was motivated by a desire to (i) improve precision by more fully utilizing exposure data on days when no health outcome measurements were available and (ii) determine if alternative criteria for selecting df were either necessary to ensure unbiasedness or preferable to increase precision. In Section 7, we will reanalyze this dataset using the methods proposed here.

3. Statistical Framework

3.1. *Overview of Model*

Consider a cohort study with a continuous health outcomes y_i, exposures x_i, and subject-specific covariates $\mathbf{z}_i$ for subjects $i = 1, \ldots, n$, measured at follow-up times t_i. We assume the t_i can take values in $\{1, \ldots, T\}$ and note that the exposure is defined and observable at every time in $\{1, \ldots, T\}$, while the outcome and subject-specific covariates are only observed on days that study subjects have clinical follow-up. We refer to the units of time as days, although other timescales can also be considered. We focus on the asymptotic properties of estimators for large n, keeping T fixed, since the number of subjects is the natural asymptotic scaling for a cohort study.

In Section 3.2, we describe a model for the random follow-up times t_i. In Sections 3.3 and 3.4, respectively, we describe models for the x_i and $\mathbf{z}_i$ conditional on the t_i. Finally, in Section 3.5 we describe a model for the y_i conditional on the x_i, $\mathbf{z}_i$, and t_i.

3.2. *Follow-Up Time*

In many observational studies, including MESA, follow-up times are determined by clinic visit dates. We assume clinics make an effort to schedule multiple appointments on a subset of the available dates, resulting in clusters of subjects with the same follow-up times, as we see in the MESA data. It is impossible to know the exact procedure by which this occurs, so we adopt the following model. Assume the study participants are divided into clusters of varying sizes such that the cluster visit days are chosen independently of each other, and each subject is pre-determined to be in a particular cluster. Notice that under this model, there is no way of knowing from the data whether individuals with clinical follow-up on the same day are part of a shared cluster. Our analyses assume that they are, which can lead to slightly conservative inference since the number of independent clusters is underestimated.

We have also evaluated the performance of our estimation methodology in simulations with alternative clustering mechanisms and with no clustering (i.e., independent follow-up times). The results are similar for different clustering mechanisms and the differences between methodologies are less pronounced when there is no clustering since there are fewer days without health outcomes (not shown).

3.3. *Exposure Model*

We assume there is a shared time series of exposures $x(\cdot)$ defined on $t \in \{1, \ldots, T\}$ such that conditional on t_i we can write $x_i = x(t_i)$ and

$$x(t_i) = g(t_i) + \eta(t_i), \tag{1}$$

where $g(\cdot)$ is a smooth function of time and $\eta(\cdot)$ is residual temporal variation.

An important question is whether to regard the function $\eta(\cdot)$ as deterministic or random variation around the temporal trend $g(\cdot)$. While it is conventional to regard $\eta(\cdot)$ as stochastic, say with the $\eta(t)$ for $t \in \{1, \ldots, T\}$ i.i.d. normal with mean zero and variance σ_η^2 (Dominici et al., 2004), it is not clear what stochastic data-generating mechanism could underly such a construction in a cohort study. Furthermore, even if it is appropriate to regard $\eta(\cdot)$ as stochastic, the assumption that there is no autocorrelation may be problematic.

We believe it is most natural to regard $\eta(\cdot)$ as deterministic and assume the sources of randomness in hypothetical repeated experiments are the choice of subjects in the cohort, their disease states, and the follow-up days t_i on which their disease states are measured. In what follows, we consider the implications of treating $\eta(\cdot)$ as either deterministic or stochastic (i.i.d. normal as described above).

3.4. *Subject-Specific Covariates*

Some subject-specific covariates will have temporal structure (e.g., blood pressure) while others will be independent of time (e.g., height). To accommodate both types of covariate, we decompose the subject-specific covariates as $\mathbf{z}_i = \mathbf{w}(t_i) + \boldsymbol{\zeta}_i$, where $\mathbf{w}(\cdot)$ is the temporal trend component and the $\boldsymbol{\zeta}_i$ are independent of time and have mean zero. Notice that unlike our model for air pollution exposure, subject-specific covariates are not purely a function of time and we always model the residual term $\boldsymbol{\zeta}_i$ as i.i.d. normal, with the stochasticity derived from random selection of subjects from the superpopulation.

3.5. *Disease Model*

Finally, we assume a linear disease model

$$y_i = x_i\beta_x + \mathbf{z}_i\boldsymbol{\beta}_z + f(t_i) + \varepsilon_i, \tag{2}$$

where β_x is the parameter of interest, $f(\cdot)$ is a smooth function of time, and the ε_i are i.i.d. normal random variables with mean zero and variance σ_ε^2. Our objective is to derive efficient and unbiased estimates of β_x. We omit dependence on meteorology to simplify notation, but no substantive changes are required to include this in the analysis.

We observe each of the y_i and $\mathbf{z}_i$ and the corresponding follow-up times t_i. We also assume we are able to measure the shared exposure time series $x(\cdot)$ without error, so that we know $x_i = x(t_i)$. The challenge in estimating β_x is to control for temporal confounding that manifests itself as a correlation between $f(\cdot)$ and the exposure time series $x(\cdot)$. If we observed $f(t_i)$ for each t_i we could easily adjust for temporal confounding by including it in the regression model. Since we do not observe the $f(t_i)$, we need to assume a flexible structure for $f(\cdot)$ and exploit this structure to adjust for temporal confounding.

3.6. *Regression Splines*

We extend the framework in Dominici et al. (2004) and assume that $f(\cdot)$, $g(\cdot)$, and $\mathbf{w}(\cdot)$ can be represented by regression splines with m_1, m_2, and m_3 df, respectively. There is always some error in assuming that a smooth function can be fully represented by a particular regression spline basis, but if we allow sufficient df this error is relatively small. Let $h_1(\cdot), h_2(\cdot), \ldots$ be a possibly infinite sequence of orthogonal regression spline basis functions, and for any positive integer m let $\mathbf{H}_m(\cdot) = (h_1(\cdot), \ldots, h_m(\cdot))$ be the vector-valued function comprised of the first m basis functions. We can write $f(\cdot) = \mathbf{H}_{m_1}(\cdot)\boldsymbol{\gamma}_{m_1}$ and $g(\cdot) = \mathbf{H}_{m_2}(\cdot)\boldsymbol{\alpha}_{m_2}$ for some $m_1 \times 1$ and $m_2 \times 1$ vectors of coefficients $\boldsymbol{\gamma}_{m_1}$ and $\boldsymbol{\alpha}_{m_2}$, respectively, and $\mathbf{w}(\cdot) = \mathbf{H}_{m_3}(\cdot)\boldsymbol{\delta}_{m_3}$ for some $m_3 \times r$ matrix of coefficients $\boldsymbol{\delta}_{m_3}$.

In practice we do not know how many df are needed to adequately describe $f(\cdot)$, $g(\cdot)$, or $\mathbf{w}(\cdot)$. It is tempting to estimate these quantities from the data using a method such as generalized cross-validation or Akaike Information Criterion (AIC). These methods have been applied for estimating the degree of smoothness in time series datasets where the residual variability is relatively small (Peng et al., 2006). However, such methods favor parsimony and may underestimate the required number of df if the smooth trends are difficult to identify in the data. In addition, a data-driven approach such as this requires using some or all of the data twice, making it difficult to estimate valid standard errors.

We prefer to determine m_1, m_2, or m_3 based on scientific judgment about the degree of variability in the seasonal and meteorological trends in the outcome, exposure, and subject-specific covariates (Schwartz, 2006) and to assess the sensitivity of our findings by fitting the model with additional df (Peng et al., 2006). We assume we can estimate m_1, m_2, or m_3 well based on scientific considerations, or at least that we have valid lower bounds for these quantities. Finally, to simplify the exposition we assume $m_3 \leq \min(m_1, m_2)$. The arguments that follow can be adapted easily to situations where $m_3 > \min(m_1, m_2)$.

4. Semiparametric Regression Model

The first approach to adjusting for temporal confounding is semiparametric regression. As adapted by Dominici et al. (2004) for time-series studies, the semiparametric regression methodology is to estimate β_x by OLS from the model

$$y_i = x_i\beta_x + \mathbf{z}_i\boldsymbol{\beta}_z + \mathbf{H}_m(t_i)\boldsymbol{\gamma}_m + \tilde{\varepsilon}_i \tag{3}$$

for some value of m. If $m < m_1$ then $\tilde{\varepsilon}_i$ may not be identical to ε_i. Assuming the degrees of smoothness of $f(t)$ and $g(t)$ are known, two natural choices are to take $m = m_1$ or $m = m_2$, which correspond to including sufficient df in the disease model to account for the trend in the health outcome or the exposure, respectively. Dominici et al. (2004) demonstrate for time series studies that it is sufficient to take $m = \min(m_1, m_2)$. We generalize their development to cohort studies and explain why in this setting it is preferable to choose $m \geq m_1$.

4.1. *Sufficient Degrees of Freedom to Account for the Trend in the Outcome ($m = m_1$)*

The analysis is straightforward if we set $m = m_1$, since the model in (3) fully adjusts for $f(\cdot) = \mathbf{H}_{m_1}(\cdot)\boldsymbol{\gamma}_{m_1}$. We can rely on fixed covariate regression results, conditioning first on the t_i and on $\eta(\cdot)$ if it is random, to conclude that there is no bias in estimating β_x and that classical standard error estimates are valid.

4.2. *Sufficient Degrees of Freedom to Account for the Trend in the Exposure ($m = m_2$)*

Suppose we set $m = m_2$ in a scenario where m_1 is greater than m_2 (the results from Section 4.1 are applicable if m_2 is greater than or equal to m_1). Define $\boldsymbol{\gamma}_{m_1/m_2} = (\boldsymbol{\gamma}_{m_2+1}, \ldots, \boldsymbol{\gamma}_{m_1})$ and the vector-valued function $\mathbf{H}_{m_1/m_2}(\cdot) = (h_{m_2+1}(\cdot), \ldots, h_{m_1}(\cdot))$. For fixed T and conditional on the t_i, we define $\mathbb{H}_{m_1/m_2} = (\mathbf{H}_{m_1/m_2}(t_1)^\top, \ldots, \mathbf{H}_{m_1/m_2}(t_n)^\top)^\top$, and denote $\bar{\mathbb{H}}_{m_1/m_2}$ in the special case of $n = T$ and $t_i = i$ corresponding to exactly one observation per day. We similarly define $\mathbb{H}_{m_2}$ and $\bar{\mathbb{H}}_{m_2}$. Considering a fixed sequence of $\eta(t_i)$ (conditionally, if $\eta(\cdot)$ is random), we define $\boldsymbol{\eta} = (\eta_{t_1}, \ldots, \eta_{t_n})^\top$ and $\bar{\boldsymbol{\eta}} = (\eta(1), \ldots, \eta(T))^\top$. We can now adapt an argument from Dominici et al. (2004) to show that as $n \to \infty$,

$$E(\hat{\beta}_x - \beta_x) \overset{\text{a.s.}}{\to} \frac{\bar{\boldsymbol{\eta}}^\top \bar{\mathbb{H}}_{m_1/m_2}\boldsymbol{\gamma}_{m_1/m_2}}{\bar{\boldsymbol{\eta}}^\top \left(I - \bar{\mathbb{H}}_{m_2}\left(\bar{\mathbb{H}}_{m_2}^\top \bar{\mathbb{H}}_{m_2}\right)^{-1}\bar{\mathbb{H}}_{m_2}^\top\right)\bar{\boldsymbol{\eta}}}. \tag{4}$$

See Web Appendix A for details.

The right-hand side of (4) is non-zero for a general deterministic function $\eta(\cdot)$, so $\hat{\beta}_x$ is asymptotically biased. Symmetry implies the right-hand side of (4) has zero expectation if the $\eta(t)$ for $t \in \{1, \ldots, T\}$ are i.i.d. normal with mean zero, in which case we conclude $\hat{\beta}_x$ is asymptotically unbiased for large n. In an asymptotic analysis appropriate for time series studies but not cohort studies, Dominici et al. (2004) show the right-hand side of (4) converges to zero as the study duration, T, converges to infinity, even for fixed $\eta(\cdot)$.

Standard error estimation is also an open problem for $m = m_2$. Classical fixed covariate regression results do not apply since the bias is only eliminated by averaging over realizations of $\eta(\cdot)$, and random covariate regression methods with robust "sandwich" standard errors (White, 1980) do not apply since the shared random function $\eta(\cdot)$ induces correlation across all study subjects. We recommend selecting m based on m_1, as in Section 4.1.

5. Pre-Adjusting the Exposure

We consider an alternative to semiparametric regression that can be more efficient if fitting (3) with sufficiently large m requires too many df relative to the available health data. The idea is to remove the temporal trend from the exposure time series and then estimate β_x without further concern for confounding. This is particularly appealing when there are many days with exposure data on which there is no health data. We assume in this section that clinic visits are equally likely on each day in $\{1, \ldots, T\}$, with obvious modifications for other follow-up day probability distributions.

We estimate $g(\cdot)$ by $\hat{g}(\cdot) = \mathbf{H}_m(\cdot)\hat{\boldsymbol{\alpha}}_m$ where $\hat{\boldsymbol{\alpha}}_m$ is the OLS estimate from fitting

$$x(\cdot) = \mathbf{H}_m(\cdot)\boldsymbol{\alpha}_m + \tilde{\eta}(\cdot)$$

based on the data $\{x(1), \ldots, x(T)\}$ and $\{\mathbf{H}_m(1), \ldots, \mathbf{H}_m(T)\}$. If $m < m_2$ then $\tilde{\eta}(\cdot)$ may not be identical to $\eta(\cdot)$. We define $\hat{\eta}(\cdot) = x(\cdot) - \hat{g}(\cdot)$ to be the pre-adjusted exposure from which the estimated trend is removed, and we estimate β_x by OLS from

$$y_i = \hat{\eta}(t_i)\beta_x + \hat{g}(t_i)\breve{\beta}_x + \mathbf{z}_i\boldsymbol{\beta}_z + (f(t_i) + \varepsilon_i), \tag{5}$$

regarding $f(t_i) + \varepsilon_i$ as the unobserved random noise. equation (2) implies that (5) holds with $\breve{\beta}_x = \beta_x$, but we estimate these quantities separately and only interpret $\hat{\beta}_x$ since we will show that for sufficiently large m it is not confounded by $f(\cdot)$.

For any two random functions of time $\phi(\cdot)$ and $\psi(\cdot)$, we define stochastic orthogonality on $\{1,\ldots,T\}$ by $E\sum_{t=1}^{T}\phi(t)\psi(t)=0$. A straightforward extension of Lemma 1 in White (1980) shows that $\hat{\beta}_x$ estimated from (5) is strongly consistent for β_x if $\hat{\eta}(\cdot)$ is stochastically orthogonal to $\hat{g}(\cdot)$, $f(\cdot)$, and each element $w_k(\cdot)$ of $\mathbf{w}(\cdot)$ for $k=1,\ldots,r$.

It is always true that $\hat{\eta}(\cdot)$ is stochastically orthogonal to $\hat{g}(\cdot)$ since $\sum_{t=1}^{T}\hat{\eta}(\cdot)\hat{g}(\cdot)$ holds by construction. In the next two subsections, we give conditions on m to guarantee stochastic orthogonality with $f(\cdot)$ and the $w_k(\cdot)$, and we discuss calculation of standard errors.

5.1. *Sufficient Degrees of Freedom to Account for the Trend in the Outcome ($m=m_1$)*

If we set $m=m_1$, then by construction $\sum_{t=1}^{T}\hat{\eta}(t)h_k(t)=0$ for $k=1,\ldots,m_1$, from which it follows that $\hat{\eta}(\cdot)$ is stochastically orthogonal to $f(\cdot)$, regardless of whether $\eta(\cdot)$ is fixed or random. Unlike the semiparametric regression model in the Section 4.1, however, this is not sufficient to guarantee strong consistency of $\hat{\beta}_x$ due to inclusion of time-varying subject-specific covariates z_i in the model. We have assumed $m_1\geq m_3$ so that similar logic guarantees that $\hat{\eta}(\cdot)$ is stochastically orthogonal to the $w_k(\cdot)$. If there is reason to believe that $m_3>m_1$ then m should be chosen to be at least as large as m_3. The required orthogonality conditions hold for fixed $\eta(\cdot)$, or conditionally if $\eta(\cdot)$ is random, so GEE standard errors (Liang and Zeger, 1986) can account for clusters of subjects with follow-ups on the same day.

5.2. *Sufficient Degrees of Freedom to Account for the Trend in the Exposure ($m=m_2$)*

Suppose now that we set $m=m_2$ in a scenario where $m_1>m_2$. We cannot rely on the arguments from Section 5.1 for a fixed $\eta(\cdot)$ to conclude that $\hat{\beta}_x$ is strongly consistent for β_x. However, if we assume the $\eta(t)$ for $t=1,\ldots,T$ are i.i.d. random variables with mean zero and are independent of the t_i, ε_i, and $\boldsymbol{\zeta}_i$, then it follows immediately that $E\hat{\eta}(t)=0$ for each $t=1,\ldots,T$ and that $\hat{\eta}(\cdot)$ is stochastically orthogonal to $f(\cdot)$ and the $w_k(\cdot)$, and there is no asymptotic bias. However, similar to Section 4.2, it is not clear how to calculate standard errors in this setting because the estimated subject exposures $\hat{\eta}(t_i)$ in (5) are all correlated with each other due to their shared dependence on $\eta(\cdot)$. Therefore, we recommend choosing m based on the smoothness of temporal trends in the outcome as in Section 5.1.

5.3. *GLS to Improve Efficiency*

Comparing (3) and (5), we see a tradeoff in efficiency between semiparametric regression and pre-adjusting the exposure. If we pre-adjust the exposure, we can estimate $\hat{\beta}_x$ from a model with m fewer degrees of freedom. However, this comes at the cost of adding $f(\cdot)$ to the unmeasured residual in the disease model, suggesting that it is better to use semiparametric regression when $f(\cdot)$ is large enough to be a precision variable. We now describe a strategy for improving efficiency of using a pre-adjusted exposure in such situations.

When we estimate β_x from (5) by OLS there is a loss of efficiency from weighting the contrasts between all subjects equally, especially if the shared component of the residual $f(\cdot)$ is relatively large. It is preferable to assign larger weights to contrasts for pairs of subjects i and j such that $f(t_i)$ and $f(t_j)$ are similar, which is analogous to assigning larger weights to within-subject contrasts in a crossover experiment (Diggle et al., 2002, p. 63). One way of achieving this in our setting is by GLS estimation with a suitably chosen reweighting matrix.

Since we do not explicitly estimate the spline coefficients for $f(\cdot)$, we cannot immediately calculate optimal weights. However, we can construct approximate weights based on an estimate of the average magnitude of the γ_k, which is obtained by regarding the γ_k as random effects in a linear mixed effects model. We emphasize that the γ_k are fixed in a given geographic region, so we do not posit a random data-generating mechanism, but we can still formally derive a mixed effects model by regarding the γ_k as exchangeable (Gelman, 2005; Hoff, 2009; Hodges and Reich, 2010). It may be that some of the γ_k represent seasonality and others represent meteorology, in which case we allow different random effect variances for the distinct groups of exchangeable coefficients.

Take $m=m_1$ as in Section 5.1 and consider the formal mixed model

$$y_i=\hat{\eta}(t_i)\beta_x+\hat{g}(t_i)\breve{\beta}_x+\mathbf{z}_i\boldsymbol{\beta}_z+\mathbf{H}_m(t_i)\boldsymbol{\gamma}_m+\varepsilon_i, \tag{6}$$

where unlike our treatment of (3) we regard $\boldsymbol{\gamma}_m$ as a random effect. If all components of $\boldsymbol{\gamma}_m$ are exchangeable (e.g., if they are all coefficients for temporal spline functions), the random effect model has a diagonal homoscedastic covariance matrix. If there are multiple groups of exchangeable coefficients in $\boldsymbol{\gamma}_m$ we allow separate homoscedastic diagonal covariance matrices for each group, with a separate variance estimate for each group.

Let W^{-1} be the marginal covariance matrix for the y_i based on estimated variances of $\boldsymbol{\gamma}_m$ and the ε_i. This can be obtained by restricted maximum likelihood (REML) using standard software such as the NLME package (Pinheiro et al., 2010) in R (R Development Core Team, 2010). We obtain $\hat{\beta}_x$ by fitting (5) using GLS with weight matrix W. GLS is more efficient than OLS because it takes advantage of the shared residual structure between observations.

Point estimates from GLS are identical to those from directly fitting the mixed model in (6), so little additional programming is required. However, standard errors based on treating the fixed γ_k as if they were stochastic cannot be assumed to be valid, and in some of our simulations they underestimate the variability of $\hat{\beta}_x$ (not shown). Therefore, we emphasize the GLS interpretation and further discuss standard error estimation in Web Appendix B.

6. Simulations

We simulate data according to (1) and (2) with $\beta_x=-0.5$ and no subject-specific covariates. The time period is $T=546$ days from September 2002 through February 2004 in six regions denoted by $R=0,\ldots,5$. We consider cohorts with 1527 clusters (as in MESA) or 300 clusters (smaller cohort) of average size 3 (range 1–10). The cluster sizes are based on the distinct groups of MESA subjects with follow-ups on the same date.

The temporal trend for the exposure is

$$g(t)=0.87\sin\left(\frac{2\pi}{365}(t+60R)\right),$$

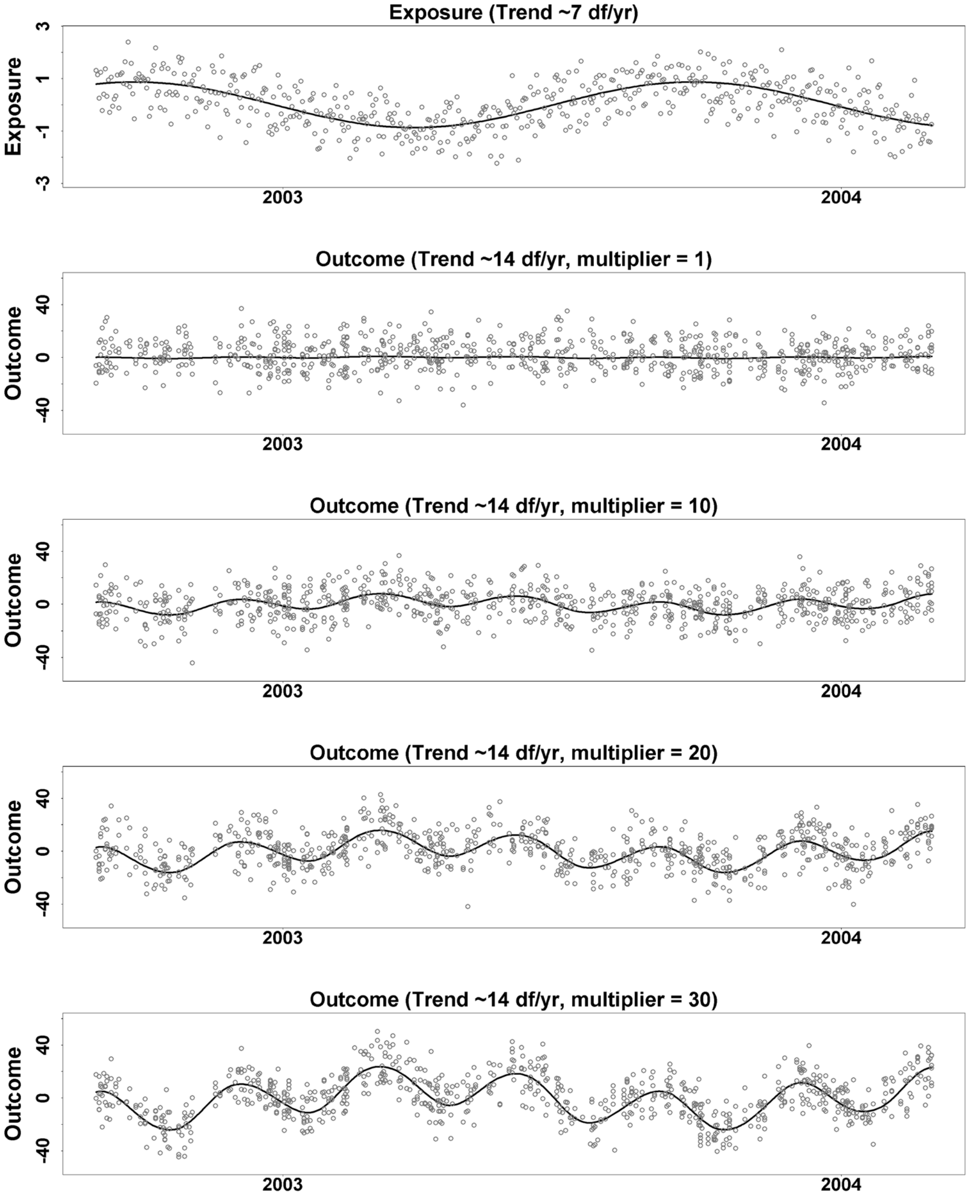

Figure 1. Simulated exposure and outcome data for a single region from a cohort with 1527 clusters across six regions (total of 781 subjects in this region). The gray dots are simulated observations, and the black curves are the assumed underlying trends.

which corresponds to a seasonal annual pattern with a different phase in each region. The outcome trend has the inverse seasonal structure plus additional finer scale structure

$$f(t) = \alpha\left[-0.35\sin\left(\frac{2\pi}{365}(t+60R)\right) - 0.48\sin\left(\frac{8\pi}{365}(t+60R)\right)\right].$$

We set the outcome trend multiplier $\alpha = 1, 10, 20, 30$, with $\alpha = 1$ corresponding to the magnitude of seasonal variation observed in the MESA data. Using B-splines, visual inspection shows that 7 df per year in each region are adequate to accurately model $g(\cdot)$ ($m_2 = 63$) while 14 df per year in each region are required for $f(\cdot)$ ($m_1 = 126$) (not shown).

The residuals in the health model are independent Gaussian random variables with $\sigma_\varepsilon^2 = 124$. The non-smooth part of the exposure $\eta(\cdot)$ is generated as independent Gaussian random variables with $\sigma_\eta^2 = 0.48$. The values of σ_η^2 and σ_ε^2 are based on corresponding residual variances in our analysis of the MESA data. We primarily report results based on a single realization of $\eta(\cdot)$ across Monte Carlo simulations, which we have argued in Section 3.3 is more scientifically plausible than the alternative scenario of independent realizations in each Monte Carlo simulation. The results with random $\eta(\cdot)$ are similar to what we report below, except that there is no bias even if we only use 7 df per year for seasonal adjustment.

Example realizations of simulated data are shown in Figure 1. Simulation results with 5000 Monte Carlo simulations

Table 1

Simulation results with 7 degrees of freedom per year, based on 5000 Monte Carlo realizations for each scenario

	Small study population (300 clusters)				Large study population (1527 clusters)			
	Rel. bias	SD	E(SE)	95% CI	Rel bias	SD	E(SE)	95% CI
Outcome trend multiplier $\alpha = 1$								
NO-ADJ	0.37 (0.012)	0.41	0.40	92%	0.36 (0.005)	0.17	0.18	83%
SEMIPAR (7 df/year)	0.01 (0.018)	0.65	0.63	95%	0.00 (0.007)	0.25	0.25	95%
SEMIPAR + GEE (7 df/year)	0.01 (0.018)	0.65	0.55	91%	0.00 (0.007)	0.25	0.24	94%
PRE-OLS (7 df/year)	0.00 (0.016)	0.56	0.54	94%	0.00 (0.007)	0.24	0.24	95%
PRE-GLS (7 df/year)	0.00 (0.016)	0.56	0.54	94%	0.00 (0.007)	0.24	0.24	95%
Outcome trend multiplier $\alpha = 10$								
NO-ADJ	3.58 (0.014)	0.49	0.42	2%	3.58 (0.006)	0.21	0.19	0%
SEMIPAR (7 df/year)	−0.04 (0.021)	0.73	0.65	92%	−0.05 (0.008)	0.28	0.26	92%
SEMIPAR + GEE (7 df/year)	−0.04 (0.021)	0.73	0.62	91%	−0.05 (0.008)	0.28	0.28	94%
PRE-OLS (7 df/year)	−0.07 (0.019)	0.66	0.64	94%	−0.06 (0.008)	0.28	0.29	96%
PRE-GLS (7 df/year)	−0.07 (0.019)	0.66	0.63	94%	−0.06 (0.008)	0.28	0.29	96%
Outcome trend multiplier $\alpha= 20$								
NO-ADJ	7.15 (0.019)	0.67	0.48	0%	7.15 (0.008)	0.29	0.21	0%
SEMIPAR (7 df/year)	−0.09 (0.026)	0.93	0.70	86%	−0.11 (0.010)	0.36	0.28	87%
SEMIPAR + GEE (7 df/year)	−0.09 (0.026)	0.93	6.50	91%	−0.11 (0.010)	0.36	0.38	96%
PRE-OLS (7 df/year)	−0.15 (0.025)	0.88	0.87	94%	−0.11 (0.011)	0.37	0.42	97%
PRE-GLS (7 df/year)	−0.14 (0.025)	0.88	0.85	94%	−0.11 (0.010)	0.36	0.39	96%
Outcome trend multiplier $\alpha= 30$								
NO-ADJ	10.71 (0.025)	0.90	0.58	0%	10.73 (0.011)	0.39	0.25	0%
SEMIPAR (7 df/year)	−0.14 (0.034)	1.20	0.79	80%	−0.16 (0.013)	0.47	0.32	81%
SEMIPAR + GEE (7 df/year)	−0.14 (0.034)	1.20	1.02	91%	−0.16 (0.013)	0.47	0.51	96%
PRE-OLS (7 df/year)	−0.23 (0.033)	1.16	1.16	94%	−0.17 (0.014)	0.49	0.57	97%
PRE-GLS (7 df/year)	−0.21 (0.033)	1.17	1.10	93%	−0.17 (0.013)	0.47	0.51	96%

The exposure deviations η_t are fixed across all Monte Carlo realizations. For each simulation scenario and estimation method, we report the mean relative bias in estimating $\beta_x = -0.5$ and Monte Carlo standard error in parentheses, the empirical standard deviation of $\hat{\beta}_x$, the mean estimated standard error, and coverage of 95% Wald confidence intervals. No adjustment is denoted by NO-ADJ, semiparametric adjustment is denoted by SEMIPAR, and pre-adjustment followed by OLS and GLS are denoted by PRE-OLS and PRE-GLS, respectively. SEMIPAR + GEE refers to the variant of using GEE standard error estimates with SEMIPAR estimation.

in each scenario are reported in Tables 1 (7 df per year) and 2 (14 df per year). We report relative bias, the observed standard deviation (SD) of $\hat{\beta}_x$, the mean estimated standard error (SE) of $\hat{\beta}_x$, and coverage of 95% confidence intervals (CIs). No adjustment is denoted by NO-ADJ, semiparametric adjustment is denoted by SEMIPAR, and pre-adjustment followed by OLS and GLS are denoted by PRE-OLS and PRE-GLS, respectively. For NO-ADJ and SEMIPAR we use classical standard errors estimates, and for PRE-OLS and PRE-GLS we use GEE standard error estimates. We also consider, as a variant, using GEE standard error estimates with SEMIPAR.

6.1. *Bias and Confidence Interval Coverage*

There is noticeable bias in $\hat{\beta}_x$ when no seasonal adjustment is made. The bias is completely eliminated by SEMIPAR, PRE-OLS, and PRE-GLS when we use 14 df per year in each region, the number of df required to account for seasonality in the outcome. There is some residual bias if we only use 7 df per year, especially for larger values of α.

All three adjustment approaches have good inferential properties when we use 14 df/year. The mean estimated SEs are close to the observed SDs, and we see nearly nominal coverage for 95% CIs. SE estimates are slightly conservative for PRE-OLS with 1527 clusters. This may be attributed to our data-based determination of which subjects to include in a cluster.

The SE estimates from SEMIPAR with 7 df/year are too small for larger values of α, resulting in undercoverage of 95% CIs. Results are improved by GEE SEs, but bias remains and numerical instability is a concern due to the small number of independent clusters compared to df in the SEMIPAR model. In particular, standard software fails to calculate GEE standard errors in a small number of our 5000 realizations (7 df/year: 1 with 300 clusters, 0 with 1527 clusters; 14 df/year: 81 with 300 clusters, 1 with 1527 clusters). We report these results only for 7 df/year and exclude the one problematic realization. The SE estimates from PRE-OLS and PRE-GLS remain accurate when we use 7 df/year and, despite the residual bias, 95% CI coverage is close to nominal.

6.2. *Relative Efficiency*

We focus on simulations with 300 clusters where we use 14 df per year for adjustment (the first four columns in Table 2). Similar patterns in relative efficiency are evident when we simulate a larger cohort with 1527 clusters, although the differences are considerably smaller. Given the residual bias with 7 df per year, relative efficiency may be of less interest, but the general patterns are similar, although somewhat less pronounced.

Table 2

Simulation results with 14 degrees of freedom per year, based on 5000 Monte Carlo realizations for each scenario

	Small study population (300 clusters)				Large study population (1527 clusters)			
	Rel. bias	SD	E (SE)	95% CI	Rel. bias	SD	E (SE)	95% CI
Outcome trend multiplier $\alpha = 1$								
NO-ADJ	0.37 (0.012)	0.41	0.40	92%	0.36 (0.005)	0.17	0.18	83%
SEMIPAR (14 df/year)	0.01 (0.022)	0.78	0.77	94%	0.00 (0.007)	0.26	0.26	95%
PRE-OLS (14 df/year)	0.01 (0.016)	0.56	0.54	94%	0.00 (0.007)	0.24	0.24	95%
PRE-GLS (14 df/year)	0.01 (0.016)	0.56	0.54	94%	0.00 (0.007)	0.24	0.24	95%
Outcome trend multiplier$\alpha = 10$								
NO-ADJ	3.58 (0.014)	0.49	0.42	2%	3.58 (0.006)	0.21	0.19	0%
SEMIPAR (14 df/year)	0.01 (0.022)	0.78	0.77	94%	0.01 (0.007)	0.26	0.26	95%
PRE-OLS (14 df/year)	−0.01 (0.019)	0.66	0.64	94%	0.00 (0.008)	0.28	0.30	96%
PRE-GLS (14 df/year)	0.00 (0.019)	0.66	0.63	94%	0.00 (0.007)	0.25	0.25	95%
Outcome trend multiplier α= 20								
NO-ADJ	7.15 (0.019)	0.67	0.48	0%	7.15 (0.008)	0.29	0.21	0%
SEMIPAR (14 df/year)	0.01 (0.022)	0.78	0.77	94%	0.01 (0.007)	0.26	0.26	95%
PRE-OLS (14 df/year)	−0.04 (0.025)	0.90	0.88	94%	0.00 (0.011)	0.38	0.43	97%
PRE-GLS (14 df/year)	0.00 (0.021)	0.74	0.69	94%	0.01 (0.007)	0.26	0.26	95%
Outcome trend multiplier $\alpha = 30$								
NO-ADJ	10.71 (0.025)	0.90	0.58	0%	10.73 (0.011)	0.39	0.25	0%
SEMIPAR (14 df/year)	0.02 (0.022)	0.78	0.77	94%	0.01 (0.007)	0.26	0.26	95%
PRE-OLS (14 df/year)	−0.06 (0.034)	1.18	1.18	94%	0.00 (0.014)	0.50	0.58	98%
PRE-GLS (14 df/year)	0.01 (0.021)	0.75	0.71	93%	0.01 (0.007)	0.26	0.26	95%

The exposure deviations η_t are fixed across all Monte Carlo realizations. For each simulation scenario and estimation method, we report the mean relative bias in estimating $\beta_x = -0.5$ and Monte Carlo standard error in parentheses, the empirical standard deviation of $\hat{\beta}_x$, the mean estimated standard error, and coverage of 95% Wald confidence intervals. No adjustment is denoted by NO-ADJ, semiparametric adjustment is denoted by SEMIPAR, and pre-adjustment followed by OLS and GLS are denoted by PRE-OLS and PRE-GLS, respectively.

We first consider scenarios with relatively small magnitudes of trend in the outcome. With $\alpha = 1$, the SD of $\hat{\beta}_x$ is 0.78 using SEMIPAR and 0.56 using PRE-OLS and PRE-GLS, which implies the relative efficiency of SEMIPAR compared to either PRE-OLS or PRE-GLS is 0.52 (ratio of variances). Similarly, with $\alpha = 10$, the relative efficiency of SEMIPAR compared to either PRE-OLS or PRE-GLS is 0.72. Consistent with our expectations, we gain efficiency by pre-adjusting the exposure, and there is no benefit from using GLS rather than OLS since the trend is not an important precision variable.

Turning now to scenarios with larger magnitudes of trend in the outcome, PRE-GLS is consistently the most efficient analysis. The relative efficiency of PRE-OLS is 0.68 for $\alpha = 20$ and 0.40 for $\alpha = 30$, and the relative efficiency of SEMIPAR is 0.95 for $\alpha = 20$ and 0.92 for $\alpha = 30$. Thus, consistent with our expectations, we see that ignoring the trend as a precision variable in PRE-OLS results in less efficiency compared to either SEMIPAR or PRE-GLS. Furthermore, it turns out that PRE-GLS is slightly more efficient that SEMIPAR when it is important to take advantage of the structure in the trend.

7. Application to Retinal Arteriolar Data

We reanalyze the data from Adar et al. (2010) to compare the impact of different temporal adjustment methods. We adjust for the full set of subject-specific covariates and use SEMIPAR, PRE-OLS, and PRE-GLS for calendar date, temperature, and relative humidity with separate B-splines with interactions by region. We also include a day-of-week term with regional interaction. We vary the df in each region between 0 and 20 per year for seasonality and 0 and 9 for the meteorology variables. For PRE-GLS, calendar date, temperature, and relative humidity are independent random effects in the mixed model formulation.

Figure 3a suggests that results are minimally sensitive to the method of adjustment for temporal confounding. This is consistent with Figure 2, since the trend in the outcome appears small compared to the overall variability, similar to $\alpha = 1$ in our simulations. Closer examination of the results in Figure 3a, however, reveals efficiency gains. If we follow Adar et al. (2010) and use 12 df per year for calendar time and 6 df for meteorology in each region, the relative efficiency of SEMIPAR compared to PRE-OLS or PRE-GLS is 0.76. The 95% confidence interval for SEMIPAR crosses zero, while the confidence intervals for PRE-OLS and PRE-GLS do not, meaning our proposed methodology results in a statistically significant association, whereas SEMIPAR does not. Of course, great care is needed in interpreting such marginally significant findings. To further illustrate the efficiency gains, in Figure 3b we show results for a randomly selected subset of 1000 MESA subjects. With the same df as above, the relative efficiency of SEMIPAR compared to PRE-OLS or PRE-GLS is 0.69.

8. Discussion

The need to adjust for temporal confounding in estimating acute air pollution effects is well known, especially in time series studies. Extension of time series methods to air pollution cohort data requires some care due to differences in data availability and plausible assumptions about randomness. A noteworthy difference is that air pollution cohort studies will often include data for exposure on days where there is no

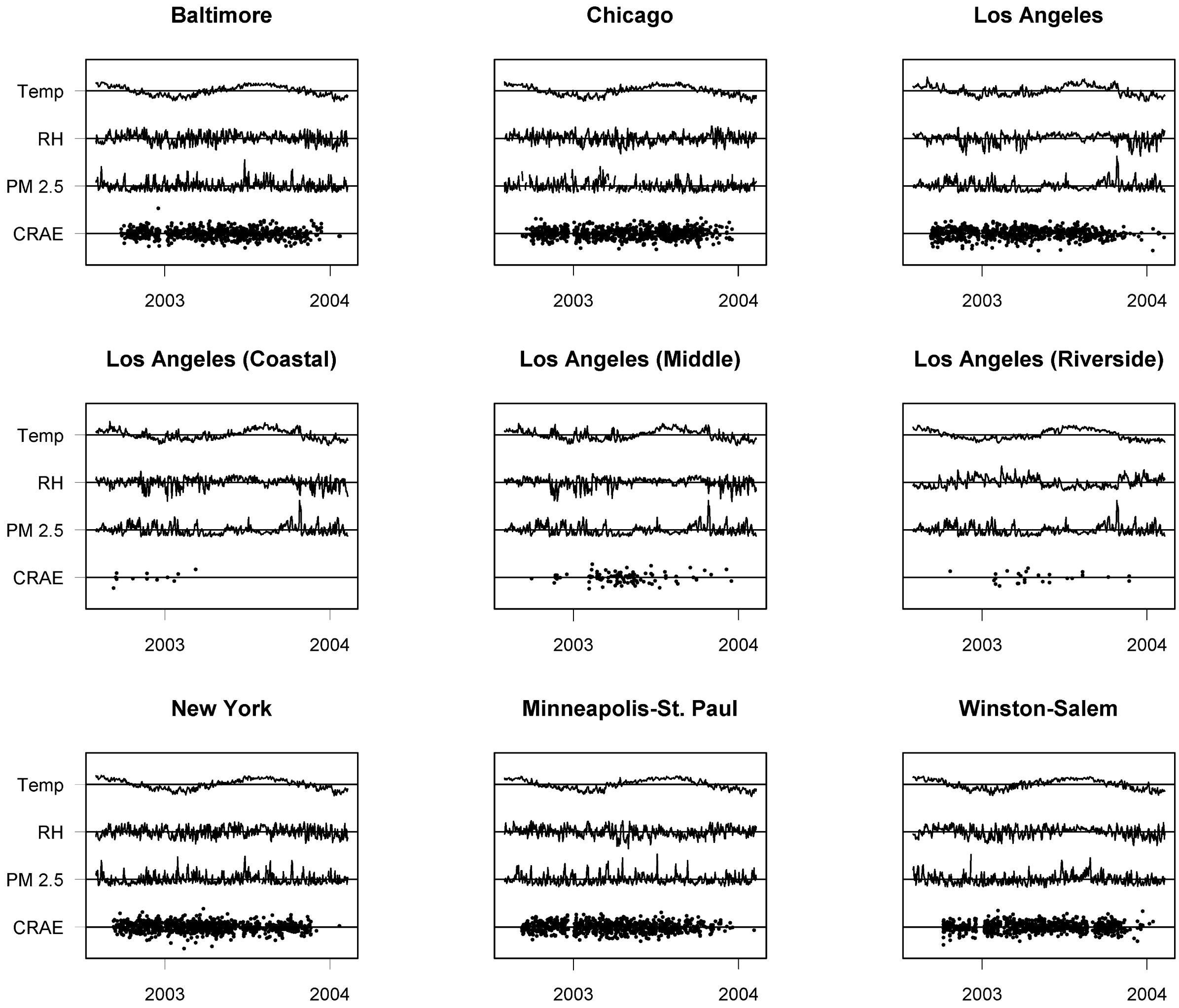

Figure 2. Data from the study of the association between short-term air pollution exposure and retinal arteriolar diameter in six MESA cities (plus four meteorology zones in Los Angeles). For each plot, time series are shown of temperature, relative humidity, and $PM_{2.5}$ concentration in the first three rows. The fourth row shows all available measurements of central retinal arteriolar equivalents (CRAE) on each day.

outcome data. We demonstrate that pre-adjusting the exposure rather than fitting a semiparametric regression model can result in increased efficiency by utilizing the additional exposure data. This approach can be improved further by estimating the health effect parameter of interest using GLS with a weight matrix determined by a formal random effects model.

Our simulation studies suggest that the advantage of pre-adjustment is most pronounced when two conditions are met, namely (i) the cohort is relatively small (in particular, smaller than MESA) such that there is no health outcome data on most days in the study period and (ii) the magnitude of temporal trend in the outcome data is small compared to the overall variability. When the trend in the outcome data is larger, the temporal adjustment terms are precision variables in the semiparametric adjustment, and pre-adjustment followed by OLS is less efficient. However, even in that situation pre-adjustment with GLS is at least as efficient as the standard semiparametric approach. Therefore, we recommend pre-adjustment followed by GLS in smaller cohort studies where there is reason to be concerned about the number of degrees of freedom required to robustly adjust for temporal confounding.

Our development emphasizes the importance of adjusting for temporal confounding with a sufficiently rich model to account for temporal trends in the outcome, and ideally any subject-specific covariates with temporal structure. Inter-subject variability presents a challenge for estimating the required model richness from cohort data, so it is even more important than in time series studies to rely on prior scientific knowledge and to conduct sensitivity analyses with different levels of temporal adjustment. While we have focused on cross-sectional cohort studies, similar issues can arise if longitudinal cohort data or certain types of panel study data are used to study acute air pollution health effects (Dominici et al., 2003).

Hodges and Reich (2010) point out the danger of introducing bias by using a formal mixed effects model as a device

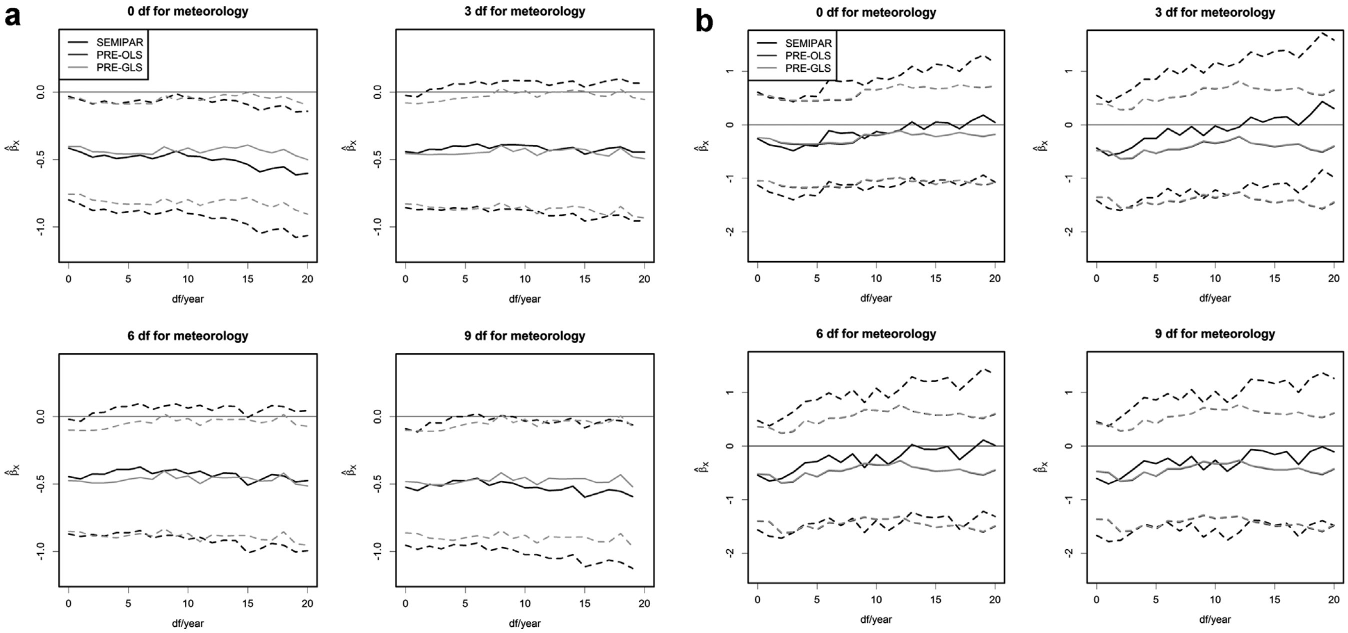

Figure 3. Estimated increase in CRAE (μm) corresponding to a 10 μg/m^3 increase in daily PM$_{2.5}$ concentration. Semiparametric regression and detrended exposure with and without shrinkage are employed with varying degrees of freedom to control for meteorology (separate splines for relative humidity and temperature in each zone) and calendar time (separate splines in each city). Solid lines are point estimates and dashed lines are 95% confidence intervals. Semiparametric adjustment is denoted by SEMIPAR and pre-adjustment followed by OLS and GLS are denoted by PRE-OLS and PRE-GLS, respectively. Point estimates for PRE-OLS and PRE-GLS are indistinguishable. (a) Full MESA cohort (4,607) subjects. (b) Randomly selected subset of MESA cohort (1,000 subjects).

for smoothing when there is no random effect in the data-generating mechanism. It would appear that our GLS approach has the potential to introduce the bias they describe, but it does not because by construction the pre-adjusted exposure is orthogonal to the random effect basis functions and to other covariates in the model in (6). If we were to use shrinkage or penalization directly in the semiparametric regression model in (3), there would be a possibility of bias as described by Hodges and Reich (2010).

A recent article proposed Bayesian adjustment for confounding (BAC), a form of model averaging, to parsimoniously adjust for confounding in time series studies (Wang, Parmigiani, and Dominici, 2012). A salient feature of BAC is joint estimation of the exposure and disease models, in contrast to our two-stage approach. The confounding adjustment in BAC is approximate and is most appropriate when there are not sufficient data to support a complete confounder adjustment, whereas our methods efficiently use all of the available data to fully adjust for confounding, assuming the data are sufficient.

In our example from MESA, exposure is defined to be the concentration on the day prior to measurement of the retinal arteriolar diameter. Other lags have been considered, including the day of exposure and the average of several days prior to exposure (Dominici et al., 2006). It is straightforward to adapt our discussion to any such pre-specified exposure lag or averaging period. The unconstrained distributed lag model (Schwartz, 2000), which provides a more flexible framework for combining exposures from multiple days, however, presents additional complications because the exposure is multivariate. Further research is needed to determine how the pre-adjustment methodology can be adapted to this setting.

We have treated short-term air pollution as spatially fixed within regions. This is reasonable given that there is much more temporal than spatial variability, but there is some exposure misclassification from ignoring the spatial variability The primary result of this is loss of statistical power to detect small effects (Zeger et al., 2000). In principle, it is possible to adapt a spatio-temporal prediction model such as the one described by Szpiro et al. (2010) to produce daily exposure predictions at subject locations. However, this has the potential to introduce additional exposure misclassification (Szpiro and Paciorek, 2013), so inference about short-term air pollution effects may not be improved.

9. Supplementary Material

Web Appendices referenced in Sections 4.2 and 5.3 and example code are available with this paper at the *Biometrics* website on Wiley Online Library.

Acknowledgements

This research was funded by NIH/NIEHS through 5P50ES015915 and by the United States EPA through R831697 and RD-83479601. It has not been subjected to the EPA's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

References

Adar, S. D., Klein, R., Klein, B. E. K., Szpiro, A. A., Cotch, M. F., Wong, T. Y., O'Neill, M. S., Shrager, S., Barr, R. G., Siscovick, D., Daviglus, M. L., Sampson, P. D., and Kaufman, J. D. (2010). Air pollution and the human microvasculature in vivo assessed via retinal imaging: The Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *PLoS Medicine* **7**, 1–11.

Bild, D. E., Bluemke, D. A., Burke, G. L., Detrano, R., Diez Roux, A. V., Folsom, A. R., Greenland, P., Jacobs, D. R., Kronmal, R., Liu, K., Nelson, J. C., O'Leary, D., Saad, M. F., Shea, S., Szklo, M., and Tracy, R. P. (2002) Multi-ethnic study of atherosclerosis: Objectives and design. *American Journal of Epidemiology* **156**, 871.

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Sicovick, D., Smith, S. C., Whitsel, L., Kaufman, J. D., and on behalf of the American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation* **121**, 2331–2378.

Diggle, P., Heagerty, P., Liang, K.-Y., and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. Oxford: Oxford University Press.

Dockery, D. W., Pope, C. A., Xu, X., Spangler, 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 cities. *New England Journal of Medicine* **329**, 1753–1759.

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

Dominici, F., McDermott, A., and Hastie, T. J. (2004). Improved semiparametric time series models of air pollution and mortality. *Journal of the American Statistical Association* **99**, 938–948.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., and Samet., J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Journal of the American Medical Association* **295**, 1127–1134.

Environmental Protection Agency. (2006). *National Ambient Air Quality Standards.*

Fung, K. Y., Krewski, D., Chen, Y., Burnett, R., and Cakmak, S. (2003). Comparison of time series and case-crossover analyses of air pollution and hospital admission data. *International Journal of Epidemiology* **32**, 1064–1070.

Gelman, A. (2005). Analysis of variance: Why it is more important than ever. *The Annals of Statistics* **33**, 1–31.

Hodges, J. and Reich, B. (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician* **64**, 325–334.

Hoff, P. (2009). *A First Course in Bayesian Statistical Methods.* Dordrecht: Springer.

Janes, H., Sheppard, L., and Lumley, T. (2005). Overlap bias in the case-crossover design with applications to air pollution exposures. *Statistics in Medicine* **24**, 285–300.

Janes, H., Sheppard, L., and Shepherd, K. (2008). Statistical analysis of air pollution panel studies: an illustration. *Annals of Epidemiology* **18**, 792–802.

Kaufman, J. D., Adar, S. D., Allen, R., Barr, R. G., Budoff, M., Burke, G., Casillas, A., Cohen, M., Curl, C., Daviglus, M., Diez-Roux, A., Jacobs, D., Kronmal, R., Larson, R., Liu, l.-J., Lumley, T., Navas-Acien, A., O'Leary, D., Rotter, J., Sampson, P. D., Sheppard, L., Siscovick, D., Stein, J., and Szpiro, A. A. (2012). Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease. the multi-ethnic study of atherosclerosis and air pollution (MESA Air). *American Journal of Epidemiology* **176**, 825–837.

Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13–22.

Miller, K. A., Sicovick, D. S., Sheppard, L., Shepherd, K., Sullivan, J. H., Anderson, G. L., and Kaufman, J. D. (2007). Long-term exposure to air pollution and incidence of cardiovascular events in women. *New England Journal of Medicine* **356**, 447–458.

Peng, R. D., Dominici, F., and Louis, T. A. (2006). Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society, Series A* **169**, 179–198.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and the R Development Core Team (2010). *nlme: Linear and Nonlinear Mixed Effects Models.* R package version 3. 1–97

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Ito, K. Krewski, D., and Thurston, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* **9**, 1132–1141.

R Development Core Team. (2010). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Samet, J. M., Dominici, F., Curriero, F., Coursac, I., and Zeger, S. L. (2000). Particulate air pollution and mortality: Findings from 20 US cities. *New England Journal of Medicine* **343**, 1742–1749.

Schwartz, J. (1994). Nonparametric smoothing in the analysis of air pollution and respiratory illness. *Canadian Journal of Statistics* **22**, 471–487.

Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology* **11**, 320.

Schwartz, J. (2006). Comment on: Model choice in time series studies of air pollution and mortality. *Journal of the Royal Statistical Society, Series A* **169**, 198–200.

Sheppard, L., Levy, D., Norris, G., Larson, T. V., and Koenig, J. Q. (1999). Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987–1994. *Epidemiology* **10**, 23–30.

Szpiro, A., Sampson, P. D., Sheppard, L., Lumley, T., Adar, S. D., and Kaufman, J. D. (2010). Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics* **21**, 606–631.

Szpiro, A. A. and Paciorek, C. J. (in press). Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics.*

Wang, C., Parmigiani, G., and Dominici, F. (2012). Bayesian effect estimation accounting for adjustment uncertainty. *Biometrics* **68**, 661–686.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* **48**, 817–838.

Zeger, S. L., Thomas, D., Dominici, F., Samet, J. M., Schwartz, J., Dockery, D., and Cohen, A. (2000). Exposure measurement error in time-series studies of air pollution: Concepts and consequences. *Environmental Health Perspectives* **108**, 419–423.

Received August 2012. *Revised October* 2013.
Accepted October 2013.