# Regression calibration with instrumental variables for longitudinal models with interaction terms, and application to air pollution studies

## M. Strand[a,b]*, S. Sillau[c], G. K. Grunwald[b] and N. Rabinovitch[d]

In this paper, we derive forms of estimators and associated variances for regression calibration with instrumental variables in longitudinal models that include interaction terms between two unobservable predictors and interactions between these predictors and covariates not measured with error; the inclusion of the latter interactions generalize results we previously reported. The methods are applied to air pollution and health data collected on children with asthma. The new methods allow for the examination of how the relationship between health outcome leukotriene E4 ($LTE_4$, a biomarker of inflammation) and two unobservable pollutant exposures and their interaction are modified by the presence or absence of upper respiratory infections. The pollutant variables include secondhand smoke and ambient (outdoor) fine particulate matter. Simulations verify the accuracy of the proposed methods under various conditions. © 2015 The Authors. *Environmetrics* published by John Wiley & Sons Ltd.

Keywords: measurement error; errors in variables; ambient $PM_{2.5}$; cigarette smoke; $LTE_4$; cotinine

## 1. INTRODUCTION

Exposures to ambient air pollution and secondhand smoke (SHS) have been implicated as health risks, particularly for some conditions such as asthma (Chilmonczyk *et al.,* 1993; Roemer *et al.,* 1993; Peters *et al*., 1997; Vedal *et al.,* 1998; Ostro *et al*., 2001; Wallace *et al*., 2003; Rabinovitch *et al*., 2006, 2011). In terms of subject exposure, measurement error is inherent to air pollution studies because most have used fixed (central) monitors at a distance from study subjects. Even for studies that have used personal monitors, determining fractions of pollution due to SHS and outdoor sources involves estimation. Regression calibration with instrumental variables (RCIV; Carroll *et al*., Chapter 6, 2006; Buonaccorsi, Chapter 5, 2010; Buzas and Stefanski, 1996) can be used to minimize bias in health-effect estimates due to measurement error by employing data from both fixed and personal monitors as well as other variables related to exposure.

Regression calibration methods are well established and have been extended for longitudinal or clustered data within recent decades, most commonly by employing linear mixed models (Bartlett *et al*., 2009); nonlinear mixed models, Ko and Davidian, 2000; or generalized linear mixed models (GLMM; Carroll *et al*., 1997; Wang *et al*., 1998, 1999). Specific extensions for RCIV using GLMMs have also been made (Li and Wang, 2012). These articles provide framework for estimation but do not give explicit forms for specific models with interaction terms. Other recent measurement error methods articles involve interaction terms, but not longitudinal data or instrumental variables (e.g., Wong *et al.,* 2004; Huang *et al*., 2005). We previously derived explicit forms of RCIV estimators and variances for longitudinal models with interaction terms between the unobservable predictors, that is, the true exposure measurements of interest, while adjusting for covariates not involved in interactions (Strand *et al*., 2014). Here, we extend these forms to allow interaction between the unobservable predictors and the covariates. This extension allows us to account for situations where the health–exposure relationship may be modified by subject characteristics or conditions.

Regression calibration with instrumental variables algorithms employ unbiased but measured-with-error variables (unbiased surrogate variables) and variables that are linearly related to the true predictors of interest (instrumental variables) to obtain consistent and nearly

*    *Correspondence to: M. Strand, Division of Biostatistics & Bioinformatics, National Jewish Health, 1400 Jackson St., M222a, Denver, CO 80206, U.S.A. E-mail: strandm@njhealth.org*

a    *Division of Biostatistics & Bioinformatics, National Jewish Health, Denver, CO, U.S.A.*

b    *Department of Biostatistics & Informatics, Colorado School of Public Health, University of Colorado Denver, Denver, CO U.S.A.*

c    *Department of Neurology, Colorado School of Medicine, University of Colorado Denver, Denver, CO U.S.A.*

d    *Department of Pediatrics, National Jewish Health, Denver, CO U.S.A.*

unbiased estimators of the coefficients of the true predictors. Two particular algorithms are often used to carry out RCIV, one that obtains the calibrated estimates as functions of estimates from the two separate model fits (RCIV1) and another that obtains predicted values from one model fit, then plugs these predicted values into a second model in place of the unknown predictors (RCIV2). Carroll *et al.* (2006) and Hardin and Carroll (2003) describe methods to obtain asymptotic variances for estimated coefficients of interest in generalized linear models using "stacked estimating equations" for RCIV1 and/or RCIV2 estimators, but recommend the use of empirical (sandwich) variances due to model approximations with this approach. Thurston *et al.* (2003) proved that RCIV1 and RCIV2 estimators have the same asymptotic variances when the outcome follows a generalized linear model of exponential family form and when there may be multiple covariates measured with error. In this article, we determine both model-based and empirical asymptotic variances of RCIV1 estimators in linear mixed models with specific interaction terms by applying the delta method. This approach allows for true covariance structures to be retained (based on underlying models and assumptions), although in practice, we have found that approximations to covariance structures have worked adequately.

Our motivating application involves a longitudinal air pollution and health study conducted on children with asthma at the National Jewish Health in Denver and reported in Rabinovitch *et al.* (2011) and Strand *et al.* (2014). In the first article, we examined the relationship between a biomarker of inflammation, leukotriene E4 (LTE$_4$), and two instrumental variables related to pollutant exposures: urinary cotinine (related to SHS exposure) and outdoor fine particulate matter (related to personal ambient fine particulate matter exposure), plus their interaction. Using these instrumental variables plus a small number of days of measured-with-error but unbiased exposure concentrations from personal monitors (the unbiased surrogate variables), we then used RCIV in the second article to estimate the slopes of the two unobserved pollutant exposure variables, plus their interaction. Both reports demonstrated that while increases in exposures to ambient fine particulate matter and SHS each had adverse effects on health, greater dose–response relationship for one pollutant occurred when the other pollutant was at lower levels. Some preliminary analyses suggested that the health–pollutant relationships for subjects might depend on whether or not children had upper respiratory infections (URIs or "colds"). The extended models presented in this paper allowed us to examine this, and could be used to study other potential effect modifiers.

In this paper, we first derive the form of regression calibration estimators for the extended RCIV1 model, which includes interactions between unobservable predictors and observable covariates, and derive asymptotic standard errors for all estimators in this model using the delta method. We then revisit the air pollution and health data to better understand the role of URIs in the health-exposure relationship. Finally, the precision and accuracy of the estimation method are assessed via simulation.

## 2. METHODS

### 2.1. The models

We consider linear mixed models, including random intercepts to account for subject heterogeneity in the responses, and a serial correlation structure for within-subject repeated measures. Although fitting models does involve approximation to the covariance structures as will be discussed later, including these two features in the approximated structures has allowed us to obtain reasonably accurate inferential results for parameters of interest, as verified through Monte Carlo simulation.

The continuous outcome variable $Y_{ij}$ for subject $i$ at time $j$, $i = 1, \cdots, n$ and $j = 1, \cdots, r_i$ is expressed as a function of the unobservable predictors ($X_{1ij}$, $X_{2ij}$ and their interaction), covariates measured without error that are involved in interactions with at least one of the unobservable predictors (denoted as $\mathbf{Z}_{ij}^*$), covariates without error that are not involved in interactions with the unobservable predictors (denoted as $\mathbf{Z}_{ij}$), and random terms to incorporate the correlation due to longitudinal measurement of subjects. If each $Z^*$ term is allowed to interact with all terms in $\mathbf{X}_{ij} = (1, X_{1ij}, X_{2ij}, X_{1ij} \cdot X_{2ij})$ and we let $\mathbf{Z}_{ij} = (Z_{1ij}, Z_{2ij}, \cdots, Z_{kij})$, $\mathbf{Z}_{ij}^* = (Z_{1ij}^*, Z_{2ij}^*, \cdots, Z_{mij}^*)$, then the model of interest is

$$Y_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} + \left(\mathbf{Z}_{ij}^* \otimes \mathbf{X}_{ij}\right)\boldsymbol{\beta}^{\mathbf{Z}^*} + b_i^Y + \epsilon_{ij} \tag{1}$$

where $b_i^Y \sim N(0, \sigma_{bY}^2)$ is a random intercept term for subjects and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \epsilon_{i2}, \cdots, \epsilon_{ir_i})^t \sim N(\mathbf{0}, \mathbf{R}_{\epsilon_i})$. We assume that the random intercept and error terms are independent of each other and that the subjects are independent. We use the spatial power structure for $\mathbf{R}_{\epsilon_i}$, which has the form $Cov(\epsilon_{ij}, \epsilon_{ij'}) = \mathbf{R}_{\epsilon_i}(j, j') = \phi_\epsilon^{|t_{ij} - t_{ij'}|}\sigma_\epsilon^2$, for days $t_{ij}$ and $t_{ij'}$ in the study period associated with observations $j$ and $j'$, respectively, for subject $i$. This is a generalization of the AR(1) structure and is useful when data are collected intermittently, such as in our data application. (In our application, we do not consider days without data to be missing values because data collection was designed to be intermittent.) The fixed-effect parameter vectors are defined as $\boldsymbol{\beta}^{\mathbf{X}} = (\beta_0^X, \beta_1^X, \beta_2^X, \beta_3^X)^t$, $\boldsymbol{\beta}^{\mathbf{Z}} = (\beta^{Z_1}, \beta^{Z_2}, \dots, \beta^{Z_k})^t$, $\boldsymbol{\beta}^{\mathbf{Z}^*} = ([\boldsymbol{\beta}^{Z_1^*}]^t, [\boldsymbol{\beta}^{Z_2^*}]^t, \cdots, [\boldsymbol{\beta}^{Z_m^*}]^t)^t$, where $\boldsymbol{\beta}^{Z_p^*} = (\beta_0^{Z_p^*}, \beta_1^{Z_p^*}, \beta_2^{Z_p^*}, \beta_3^{Z_p^*})^t$, for $p = 1, ..., m$. The $\otimes$ symbol indicates the Kronecker product; because $\mathbf{X}_{ij}$ and $\mathbf{Z}_{ij}^*$ are both row vectors, the resulting product will also be a row vector with $4m$ elements: $(Z_{1ij}^*\mathbf{X}_{ij}, \dots, Z_{mij}^*\mathbf{X}_{ij})$. Let $\boldsymbol{\beta} = ([\boldsymbol{\beta}^{\mathbf{X}}]^t, [\boldsymbol{\beta}^{\mathbf{Z}}]^t, [\boldsymbol{\beta}^{\mathbf{Z}^*}]^t)^t$ denote the entire set of $4(m+1) + k$ fixed-effect parameters in model (1).

In our data application, $Y$ is natural log leukotriene E4, or ln(LTE$_4$), and $X_1$ and $X_2$ are SHS and fine ambient particulate matter exposure concentrations, respectively. Both of the pollutant variables are considered in terms of fine particulate matter (technically, particulate matter less than 2.5 microns in diameter, or PM$_{2.5}$), and measured in $\mu$g per $m^3$. Our primary models included an indicator for URI, or "cold" status ($Z_1^* = 1$ for presence, 0 for absence) as the covariate allowed to interact with $X$ variables, and study year ($Z_1, \dots, Z_4$; the last year was the reference year) as covariates not involved in interactions.

If not all interactions between $Z^*$ and $X$ variables are required, then the model can be reformulated as

$$Y_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} + \sum_{p=1}^{m} Z_{pij}^* \mathbf{X}_{ij}^{(p)} \boldsymbol{\beta}_{(p)}^{Z_p^*} + b_i^Y + \epsilon_{ij} \tag{2}$$

where $\mathbf{X}_{ij}^{(p)}$ is the (nontrivial) subset of $\mathbf{X}_{ij}$ to be included in interactions with $Z_{pij}^*$, for $p = 1, \ldots, m$ and $\boldsymbol{\beta}_{(p)}^{Z_p^*}$ is the corresponding subset of $\boldsymbol{\beta}^{Z_p^*}$.

Instead of observing the predictors of interest, $X_1$ and $X_2$, we observe instrumental variables $M_1$ and $M_2$ that are linearly related to them:

$$X_{1ij} = \theta_0 + \theta_1 M_{1ij} + b_i^{X_1} + \omega_{1ij}$$
$$X_{2ij} = \pi_0 + \pi_1 M_{2ij} + b_i^{X_2} + \omega_{2ij}$$
(3)

where $b_i^{X_p} \sim N(0, \sigma_{b^{X_p}}^2)$ are random intercepts for subjects, for $p = 1, 2$, and $\omega_{1ij}$ and $\omega_{2ij}$ are error terms with $\boldsymbol{\omega}_{pi} = (\omega_{pi1}, \omega_{pi2}, \ldots, \omega_{pir_i})^t \sim N(\mathbf{0}, \mathbf{R}_{\boldsymbol{\omega}_{pi}})$. As with (1) [or (2)], we assume that the random intercept and error terms are independent (for each equation), subjects are independent, and assume the spatial power structure for $\mathbf{R}_{\boldsymbol{\omega}_{pi}}$. Instrumental variables in our application are cotinine ($M_1$, measured in ng per mg of creatinine), a metabolite of nicotine that can be measured in the urine and used as a biomarker of exposure to SHS, and ambient fine particulate matter concentrations measured by a fixed outdoor monitor ($M_2$, measured in $\mu$g per $m^3$).

We can generalize (3) so that $M_1$ and $M_2$ are included in both equations. Such an expanded model was considered in Strand *et al.* (2014). One of the difficulties of this model is that matrices associated with estimators have linear dependencies that need to be dealt with. It does not produce problems with estimators, but just complicates derivations. In this article, we will only consider (3) and not the expanded version primarily for the sake of simplifying notation, although results could be generalized to that case. In addition, the use of (3) seemed to be adequate for the data application, although the expanded version was used in the previous article to demonstrate the methods.

The primary difficulty in measuring $X_1$ and $X_2$ is because they are mixed together when measured, along with pollution from other sources. Although we do not observe $X_1$ and $X_2$, in our application, there is a small amount of data for "unbiased but measured-with-error" versions of these variables, denoted as $W_{1ij} = X_{1ij} + U_{1ij}$ and $W_{2ij} = X_{2ij} + U_{2ij}$, respectively, where $U_{pij}$ has mean 0 and is assumed to be independent of $b_i^{X_p}$, $p = 1, 2$. Incorporating $W_1$ and $W_2$ into (3) yields

$$W_{1ij} = \theta_0 + \theta_1 M_{1ij} + b_i^{X_1} + \omega'_{1ij}$$
$$W_{2ij} = \pi_0 + \pi_1 M_{2ij} + b_i^{X_2} + \omega'_{2ij}$$
(4)

where $\omega'_{pij} = \omega_{pij} + U_{pij}$, for $p = 1, 2$. Portions of total measured fine particulate matter attributable to SHS ($W_1$) and ambient sources ($W_2$) were estimated using personal monitors, chemical signatures, and methods as described in Strand *et al.* (2014). This approach is expensive, cumbersome, and still results in measurements with error. But as long as we can assume that the $W$ variables are unbiased for the $X$ variables, we can apply RCIV. In (4), let $\boldsymbol{\omega}'_{pi} = (\omega'_{pi1}, \omega'_{pi1}, \ldots, \omega'_{pir_i})^t$ denote the error vector for subject $i$ and $p = 1, 2$. We use the spatial power structure for $\mathbf{R}_{\boldsymbol{\omega}'_{pi}}$, although it is likely to just be an approximation.

Letting $\mathbf{M}_{ij} = (1, M_{1ij}, M_{2ij}, M_{1ij} \cdot M_{2ij})$, the health outcome model fit for RCIV1 is

$$Y_{ij} = \mathbf{M}_{ij} \boldsymbol{\gamma}^{\mathbf{M}} + \mathbf{Z}_{ij} \boldsymbol{\gamma}^{\mathbf{Z}} + \left( \mathbf{Z}_{ij}^* \otimes \mathbf{M}_{ij} \right) \boldsymbol{\gamma}^{\mathbf{Z}^*} + b_i^{Y*} + \epsilon_{ij}^*$$
(5)

This is essentially model (1) with $\mathbf{M}_{ij}$ used in place of $\mathbf{X}_{ij}$, that is, unobservable predictors are replaced with their instrumental counterparts. Collectively, we define $\boldsymbol{\gamma} = ([\boldsymbol{\gamma}^{\mathbf{M}}]^t, [\boldsymbol{\gamma}^{\mathbf{Z}}]^t, [\boldsymbol{\gamma}^{\mathbf{Z}^*}]^t)^t$ as the $4(m + 1) + k$ fixed-effect parameters in model (5) that correspond with the parameters in $\boldsymbol{\beta}$ for model (1). Asterisks are included on the random terms in (5) to distinguish them from those in (1). A version of (5) that involves partial interactions can also be written [analogous to how (2) relates to (1)].

Models (4) and (5) are fitted to obtain estimates that are then combined via RCIV1 to achieve the calibrated estimates, which are derived in the next section. The true covariance structure for response variables in models (4) and (5) as well as between responses in (4) and (5) based on underlying models (1) and (3) are displayed and/or discussed in Appendix A. These covariance structures are quite complicated, but our work has shown that using the "spatial power plus random intercept" structures as approximations in models (4) and (5) works reasonably well. This is true for models here as well as those studied in Strand *et al.* (2014). To recognize these forms as only approximations, asterisks are included on the covariance parameters to distinguish them from parameters in true underlying structures. While using these simpler covariance structures has appeared to be adequate, we also found that ignoring the repeated measures completely leads to very over-inflated variances.

## 2.2. Point estimation

Let $\mathbf{W}_{ij} = (1, W_{1ij}, W_{2ij}, W_{1ij} \cdot W_{2ij})$. Based on models (1) and (4), and assuming $(b_i^{X_1} + \omega'_{1ij})$ and $(b_i^{X_2} + \omega'_{2ij})$ are uncorrelated, we can write $E(\mathbf{W}_{ij} | \mathbf{M}_{ij}) = \mathbf{M}_{ij} \mathbf{C}$, where

$$\mathbf{C} = \begin{pmatrix} 1 & \theta_0 & \pi_0 & \theta_0 \pi_0 \\ 0 & \theta_1 & 0 & \theta_1 \pi_0 \\ 0 & 0 & \pi_1 & \theta_0 \pi_1 \\ 0 & 0 & 0 & \theta_1 \pi_1 \end{pmatrix}$$
(6)

For our models and assumptions,

$$
\begin{aligned}
E\left(Y_{ij}|\mathbf{M}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*\right) &= E_{\mathbf{x}}\left(E\left(Y_{ij}|\mathbf{X}_{ij},\mathbf{M}_{ij},\mathbf{W}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*\right)|\mathbf{M}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*\right) \\
&= E_{\mathbf{x}}\left(E\left(Y_{ij}|\mathbf{X}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*\right)|\mathbf{M}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*\right) \qquad \dagger \\
&= E_{\mathbf{x}}\left(\mathbf{X}_{ij}\boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} + \left(\mathbf{Z}_{ij}^*\otimes\mathbf{X}_{ij}\right)\boldsymbol{\beta}^{\mathbf{Z}^*}|\mathbf{M}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*\right) \\
&= E_{\mathbf{x}}(\mathbf{X}_{ij}|\mathbf{M}_{ij})\boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} + \left[\mathbf{Z}_{ij}^*\otimes E_{\mathbf{x}}(\mathbf{X}_{ij}|\mathbf{M}_{ij})\right]\boldsymbol{\beta}^{\mathbf{Z}^*} \qquad \ddagger \\
&= E_{\mathbf{x}}(\mathbf{W}_{ij}|\mathbf{M}_{ij})\boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} + \left[\mathbf{Z}_{ij}^*\otimes E_{\mathbf{x}}(\mathbf{W}_{ij}|\mathbf{M}_{ij})\right]\boldsymbol{\beta}^{\mathbf{Z}^*} \qquad \# \\
&= \mathbf{M}_{ij}\mathbf{C}\boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} + \left[\mathbf{Z}_{ij}^*\otimes(\mathbf{M}_{ij}\mathbf{C})\right]\boldsymbol{\beta}^{\mathbf{Z}^*}
\end{aligned}
$$

In the aforementioned equations, we employ the following assumptions that are standard in deriving regression calibration estimators: $\dagger$ $\mathbf{W}_{ij}$ and $\mathbf{M}_{ij}$ are surrogates for $\mathbf{X}_{ij}$ in predicting $Y_{ij}$, such that $E(Y_{ij}|\mathbf{X}_{ij},\mathbf{M}_{ij},\mathbf{W}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*) = E(Y_{ij}|\mathbf{X}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*)$ (see Carroll *et al.*, 2006 for a definition of surrogacy involving distributions); $\ddagger$ $E(\mathbf{X}_{ij}|\mathbf{M}_{ij},\mathbf{Z}_{ij},\mathbf{Z}_{ij}^*) = E(\mathbf{X}_{ij}|\mathbf{M}_{ij})$; $\#$ random terms between models in (4) are uncorrelated and $\mathbf{M}_{ij}$ is uncorrelated with $U_{pij}$, $p = 1,2$, such that $E(\mathbf{X}_{ij}|\mathbf{M}_{ij}) = E(\mathbf{W}_{ij}|\mathbf{M}_{ij})$. For a given application, whether these assumptions are met should be considered carefully. Violations to these assumptions are a matter of degree; small violations may be inconsequential, but larger violations could be very problematic. We return to this again in the Discussion.

In order to determine forms of estimators of $\boldsymbol{\beta}^{\mathbf{X}}$, $\boldsymbol{\beta}^{\mathbf{Z}}$ and $\boldsymbol{\beta}^{\mathbf{Z}^*}$, we first equate terms in the preceding mean derivation with those in the mean of (5), given $\mathbf{M}_{ij}$ and covariates:

$$
\mathbf{M}_{ij}\mathbf{C}\boldsymbol{\beta}^{\mathbf{X}} = \mathbf{M}_{ij}\boldsymbol{\gamma}^{\mathbf{M}}
$$

$$
\mathbf{Z}_{ij}\boldsymbol{\beta}^{\mathbf{Z}} = \mathbf{Z}_{ij}\boldsymbol{\gamma}^{\mathbf{Z}}
$$

$$
\left[\mathbf{Z}_{ij}^*\otimes(\mathbf{M}_{ij}\mathbf{C})\right]\boldsymbol{\beta}^{\mathbf{Z}^*} = \left(\mathbf{Z}_{ij}^*\otimes\mathbf{M}_{ij}\right)\boldsymbol{\gamma}^{\mathbf{Z}^*}
$$

By properties of the Kronecker product, we can re-express $\mathbf{Z}_{ij}^*\otimes(\mathbf{M}_{ij}\mathbf{C}) = (\mathbf{Z}_{ij}^*\times\mathbf{I}_m)\otimes(\mathbf{M}_{ij}\mathbf{C}) = (\mathbf{Z}_{ij}^*\otimes\mathbf{M}_{ij})\times(\mathbf{I}_m\otimes\mathbf{C})$, and hence, the last aforementioned equation becomes

$$
\left(\mathbf{Z}_{ij}^*\otimes\mathbf{M}_{ij}\right)(\mathbf{I}_m\otimes\mathbf{C})\boldsymbol{\beta}^{\mathbf{Z}^*} = \left(\mathbf{Z}_{ij}^*\otimes\mathbf{M}_{ij}\right)\boldsymbol{\gamma}^{\mathbf{Z}^*}
$$

where $\mathbf{I}_m$ is an $m\times m$ identity matrix. We also note that $(\mathbf{I}_m\otimes\mathbf{C})^{-1} = (\mathbf{I}_m\otimes\mathbf{C}^{-1})$. By utilizing all of these equations and replacing parameters with their maximum likelihood estimators in the mixed model fits, we obtain

$$
\hat{\boldsymbol{\beta}}^{\mathbf{X}} = \hat{\mathbf{C}}^{-1}\hat{\boldsymbol{\gamma}}^{\mathbf{M}}
$$

$$
\hat{\boldsymbol{\beta}}^{\mathbf{Z}} = \hat{\boldsymbol{\gamma}}^{\mathbf{Z}}
$$

$$
\hat{\boldsymbol{\beta}}^{\mathbf{Z}^*} = (\mathbf{I}_m\otimes\hat{\mathbf{C}}^{-1})\hat{\boldsymbol{\gamma}}^{\mathbf{Z}^*}
$$

where $\hat{\mathbf{C}}$ indicates that all parameters in $\mathbf{C}$ are replaced with their estimators; this matrix is invertible as long as estimates yield a nonsingular matrix, which is expected. (e.g., $\hat{\theta}_1$ and $\hat{\pi}_1$ must be nonzero.) The matrices $\mathbf{C}^{-1}$ and $\hat{\mathbf{C}}^{-1}$ can also easily be written in algebraic form.

Often, not all interactions may be of interest. A formulation of the model and estimators for partial interactions is given in Appendix B.

## 2.3. Variance of $\hat{\boldsymbol{\beta}}$

As in Strand *et al.* (2014), we employed the delta method to obtain the approximate asymptotic variance of $\hat{\boldsymbol{\beta}}$. We use the term "approximate" because of the approximations of covariance structures in the model; asymptotic methods are inherently approximate. Based on the delta method, $Var(\hat{\boldsymbol{\beta}}) = \boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}^t$ where

$$
\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{\theta} & \boldsymbol{\Sigma}_{\theta\pi} & \boldsymbol{\Sigma}_{\theta\gamma} \\ \boldsymbol{\Sigma}_{\pi\theta} & \boldsymbol{\Sigma}_{\pi} & \boldsymbol{\Sigma}_{\pi\gamma} \\ \boldsymbol{\Sigma}_{\gamma\theta} & \boldsymbol{\Sigma}_{\gamma\pi} & \boldsymbol{\Sigma}_{\gamma} \end{pmatrix} \tag{7}
$$

is the covariance matrix for estimators of fixed-effect parameters in (4) and (5) and $\boldsymbol{\Delta}$ is a matrix of partial derivatives with $(i,j)^{th}$ element $\frac{\partial\boldsymbol{\beta}[i]}{\partial\boldsymbol{\alpha}[j]}$, where $\boldsymbol{\beta}[i]$ denotes the $i^{th}$ element of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}[j]$ denotes the $j^{th}$ element of $\boldsymbol{\alpha} = (\theta_0,\theta_1,\pi_0,\pi_1,\boldsymbol{\gamma}^t)^t$. The diagonal elements in (7) are the covariance matrices for fixed-effect estimators in model (4) (first two diagonal matrices) and model (5) (third diagonal matrix), and the off-diagonal matrices contain covariances between fixed-effect parameter estimators between models. Specific forms of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$ based on models (1), (4), and (5) and the defined estimators are given in Appendix C. In (7), note that $\boldsymbol{\Sigma}_{\theta\pi} = \boldsymbol{\Sigma}_{\pi\theta}^t$, $\boldsymbol{\Sigma}_{\theta\gamma} = \boldsymbol{\Sigma}_{\gamma\theta}^t$, $\boldsymbol{\Sigma}_{\pi\gamma} = \boldsymbol{\Sigma}_{\gamma\pi}^t$.

**Table 1.** Descriptive statistics of variables involved in the analysis

| Statistic | $Y$ | $W_1$ | $W_2$ | $M_1$ | $M_2$ | $Z_1^*$ |
|---|---|---|---|---|---|---|
| Number of subjects with data available | 85 | 64 | 50 | 85 | 90 | 86 |
| Average number of responses per subject | 24.5 | 11.5 | 11.3 | 24.6 | 392.7 | 59.1 |
| Minimum of subject means | 3.85 | 0.80 | 1.41 | −0.75* | 2.48 | 0 |
| Average of subject means | 4.49 | 1.51 | 1.92 | 1.67 | 2.67 | 0.18 |
| Maximum of subject means | 5.93 | 3.22 | 2.34 | 5.32 | 2.80 | 0.73 |
| Average of subject SD's | 0.35 | 0.42 | 0.41 | 0.64 | 0.62 | 0.31 |
| SD of subject means | 0.36 | 0.49 | 0.19 | 1.35 | 0.07 | 0.17 |

$Y$=ln(LTE$_4$); $W_1$=ln(SHS exposure+1); $W_2$=ln(ambient fine particulate matter exposure + 1); $M_1$=ln(cotinine); $M_2$=ln(ambient fine particulate matter from fixed monitor); $Z_1^*$=upper respiratory infection (1=present, 0=absent), that is, 'cold'. Units for pollutant exposure variables and ambient fine particulate matter measured at the fixed monitor were $\mu$g/m$^3$; units for LTE$_4$ were pg per mg of creatinine; units for cotinine were ng per mg of creatinine (all before transformation). Results are combined across study years, $Z_1, \ldots, Z_4$.

\* Negative values occurred due to natural log transformation.

Both model-based and empirical forms of $Var(\hat{\boldsymbol{\beta}})$ were calculated (Appendix C). When random terms between models in (4) as well as between those in (4) and (5) are uncorrelated, $\boldsymbol{\Sigma}_{\boldsymbol{\theta\pi}} = \mathbf{0}$, although $\boldsymbol{\Sigma}_{\boldsymbol{\theta\gamma}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\pi\gamma}}$ are not necessarily $\mathbf{0}$. For model-based variances, we set all of these submatrices to $\mathbf{0}$, primarily to avoid non-convergence or non-positive definite matrix issues. For the empirical-based estimates of $Var(\hat{\boldsymbol{\beta}})$, we did not set off-diagonal elements to $\mathbf{0}$. One disadvantage of doing this is the same subject-day records must be used to fit different models in (4) and (5) so that off-diagonal covariances can be estimated (with our given approach).

In the following are our steps to calculate model-based and empirical forms of $Var(\hat{\boldsymbol{\beta}})$ used in the data application and simulations, followed by some important notes.

Steps to calculate the model-based form of $Var(\hat{\boldsymbol{\beta}})$:

1. Set off-diagonal submatrices in (7) to $\mathbf{0}$.
2. Obtain diagonal submatrices in (7) from mixed model fits. Numerical forms of the diagonal covariance matrices are obtained easily with common statistical packages (for example, using the COVB option in ODS OUTPUT in SAS PROC MIXED or in the y.fit\$varFix object in R, when using the *lme* function from the *nlme* package).
3. Compute $Var(\hat{\boldsymbol{\beta}}) = \boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}^t$ using forms of $\boldsymbol{\Delta}$ as shown in Appendix C, replacing parameters with their estimates.

Steps to calculate the empirical form of $Var(\hat{\boldsymbol{\beta}})$:

1. Calculate empirical versions of submatrices in (7) using forms shown in Appendix C, where the middle "Var" and "Cov" quantities are calculated based on squared residuals or residual cross products.
2. Compute $Var(\hat{\boldsymbol{\beta}}) = \boldsymbol{\Delta}\boldsymbol{\Sigma}\boldsymbol{\Delta}^t$ using forms of $\boldsymbol{\Delta}$ as shown in Appendix C, replacing parameters with their estimates.

## 3. APPLICATION

In our analyses, we consider the same data as described and analyzed in Strand *et al.* (2014), with the addition of the time-varying "cold" indicator variable, obtained through surveys with the children. The variables are defined in Section 2.1, with summary statistics presented in Table 1. Within a study year, data collection occurred approximately from October through April. Data from the first 2 study years (2002–2004) were available to fit (4), while data from all 4 years (2002–2006) were available to fit (5). Consequently, model-based variances but not empirical variances were calculated for $\hat{\boldsymbol{\beta}}$. For ease of interpretation of parameter estimates, we mean-corrected the $W$ variables in the fit of (4). Instrumental data were much more abundant and easier to collect than data from personal monitors, making RCIV a natural choice for analysis. Our assumed model for (1) was

$$
\begin{aligned}
Y_{ij} = {} & \beta_0^X + \beta_1^X X_{1ij} + \beta_2^X X_{2ij} + \beta_3^X X_{1ij} X_{2ij} + \beta^{Z_1} Z_{1ij} + \beta^{Z_2} Z_{2ij} + \beta^{Z_3} Z_{3ij} + \beta^{Z_4} Z_{4ij} \\
& + \beta_0^{Z_1^*} Z_{1ij}^* + \beta_1^{Z_1^*} Z_{1ij}^* X_{1ij} + \beta_2^{Z_1^*} Z_{1ij}^* X_{2ij} + \beta_3^{Z_1^*} Z_{1ij}^* X_{1ij} X_{2ij} + b_i^Y + \epsilon_{ij}
\end{aligned}
\tag{8}
$$

Table 2 shows results of model (4) and (5) fits, while Table 3 shows estimates of the key parameters in (8) via RCIV1. (Note that year 4 was set as the reference, so $\beta^{Z_4} = \gamma^{Z_4} = 0$.) Our data suggested that the instrumental ($M$) variables were only weak to moderately associated with the personal exposure data. This was one of the reasons why more data were sought to fit (4), even though unbiased surrogate data did

**Table 2.** Estimates of parameters in (4) and (5) for the data application (SE in parentheses)

| Predictor or covariance quantity | Regression of $W_1^a$ (personal SHS exposure) on $M_1$ | Regression of $W_2^a$ (personal ambient fine particulate matter) on $M_2$ | Regression of $Y$ (LTE$_4$) on $M_1, M_2, M_1 \cdot M_2$ |
|---|---|---|---|
| Intercept | $\hat{\theta}_0 = -0.26\ (0.07)$ | $\hat{\pi}_0 = -0.52\ (0.09)$ | $\hat{\gamma}_0^M = 4.12\ (0.08)$ |
| Cotinine ($M_1$) | $\hat{\theta}_1 = 0.14\ (0.03)$ | | $\hat{\gamma}_1^M = 0.09\ (0.03)$ |
| Ambient PM$_{2.5}$ ($M_2$) | | $\hat{\pi}_1 = 0.17\ (0.03)$ | $\hat{\gamma}_2^M = 0.05\ (0.02)$ |
| $M_1 \cdot M_2$ | | | $\hat{\gamma}_3^M = -0.02\ (0.01)$ |
| Cold ($Z_1^*$) | | | $\hat{\gamma}_0^{Z_1^*} = 0.20\ (0.15)$ |
| $Z_1^* \cdot M_1$ | | | $\hat{\gamma}_1^{Z_1^*} = -0.07\ (0.08)$ |
| $Z_1^* \cdot M_2$ | | | $\hat{\gamma}_2^{Z_1^*} = -0.05\ (0.05)$ |
| $Z_1^* \cdot M_1 \cdot M_2$ | | | $\hat{\gamma}_3^{Z_1^*} = 0.03\ (0.03)$ |
| Year 1 ($Z_1$)[b] | | | $\hat{\gamma}^{Z_1} = 0.11\ (0.04)$ |
| Year 2 ($Z_2$)[b] | | | $\hat{\gamma}^{Z_2} = 0.24\ (0.04)$ |
| Year 3 ($Z_3$)[b] | | | $\hat{\gamma}^{Z_3} = 0.30\ (0.04)$ |
| Random subject variance | $\hat{\sigma}_{b^{X_1}}^{*2} = 0.15$ | $\hat{\sigma}_{b^{X_2}}^{*2} = 0.02$ | $\hat{\sigma}_{b^Y}^{*2} = 0.11$ |
| Correlation | $\hat{\phi}_{\omega_1'}^{*} = 0.47$ | $\hat{\phi}_{\omega_2'}^{*} = 0.08$ | $\hat{\phi}_{\epsilon}^{*} = 0.42$ |
| Residual variance | $\hat{\sigma}_{\omega_1'}^{*2} = 0.22$ | $\hat{\sigma}_{\omega_2'}^{*2} = 0.17$ | $\hat{\sigma}_{\epsilon}^{*2} = 0.12$ |

$W$ variables were mean-corrected before the fit of (4). There were 458, 560 and 1438 records for analysis in the regression models for $W_1$, $W_2$ and $Y$, respectively; see the text for more detail about the variables and models. Correlations are meaningful for responses on consecutive days. Ambient fine particulate matter is denoted as ambient PM$_{2.5}$.

[a] Mean corrected before model fit.

[b] Relative to Year 4.

**Table 3.** Estimates of parameters (or linear combinations of parameters) in model (8) for ln(LTE$_4$) based on the RCIV1 estimation method

| Model term | No cold | | Cold | | Difference |
|---|---|---|---|---|---|
| | Estimate (SE) | *P*-value | Estimate (SE) | *P*-value | P-value |
| Intercept | $\hat{\beta}_0^X = 4.33\ (0.049)$ | <0.001 | $\hat{\beta}_0^X + \hat{\beta}_0^{Z^*} = 4.41\ (0.06)$ | <0.001 | 0.02 |
| SHS slope (at mean ambient PM$_{2.5}$) | $\hat{\beta}_1^X = 0.22\ (0.11)$ | 0.04 | $\hat{\beta}_1^X + \hat{\beta}_1^{Z^*} = 0.27\ (0.20)$ | 0.18 | 0.75 |
| Ambient PM$_{2.5}$ slope (at mean SHS) | $\hat{\beta}_2^X = 0.07\ (0.11)$ | 0.52 | $\hat{\beta}_2^X + \hat{\beta}_2^{Z^*} = 0.09\ (0.30)$ | 0.76 | 0.92 |
| Interaction | $\hat{\beta}_3^X = -0.86\ (0.48)$ | 0.07 | $\hat{\beta}_3^X + \hat{\beta}_3^{Z^*} = 0.23\ (1.45)$ | 0.88 | 0.36 |
| Ambient PM$_{2.5}$ slope (in absence of SHS) | 1.37 (0.73) | 0.06 | −0.25 (2.04) | 0.90 | 0.34 |

Model-based asymptotic standard errors were derived using the delta method as described in Section 2.3, and *p*-values are based on Wald Z-tests assuming asymptotic normality. Because $W$ variables were mean corrected before analysis, interpretations for SHS and ambient PM$_{2.5}$ slopes (rows 2 and 3) are relevant at the mean of the other pollutant, and the intercept is the estimate of mean ln(LTE4) for the reference year (Year 4) at the mean of both pollutants. The larger SEs for "Cold" are due to fewer subject-days for these conditions, relative to "No cold". Ambient fine particulate matter is denoted as ambient PM$_{2.5}$.

not exist in the last 2 years of the study. Figure 1 shows predicted values based on calibrated estimates; the figure demonstrates that while predicted values for LTE$_4$ generally increase as pollutants increase, there was noticeable negative interaction between pollutants for those without colds ($p = 0.07$); when one pollutant was lower, the dose–response relationship between LTE$_4$ and the other pollutant was stronger. Such a pattern was not observed for those with colds ($p = 0.88$ for interaction). The quartiles used in Figure 1 were determined based on distributions of predicted values of $W_1$ and $W_2$ from fits of (4). This differs from the approach taken for the figure in Strand *et al.* (2014), for
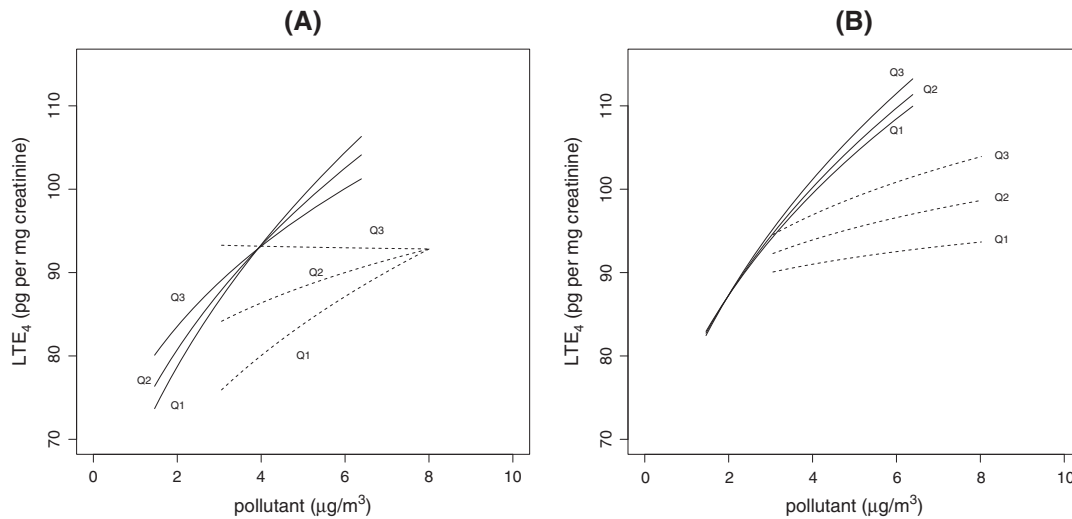
**Figure 1.** Relationships between LTE$_4$ and exposure to one pollutant, modified by exposure to a second pollutant, for children on (A) days without colds and (B) days with colds. Solid lines show estimated mean LTE$_4$ as a function of secondhand smoke (SHS), by quartiles of ambient fine particulate matter. Dashed lines show estimated mean LTE$_4$ as a function of ambient fine particulate matter, by quartiles of SHS. Quartiles were determined from distributions of predicted values from model fits of (4). Quartile symbols are Q1=25th percentile, Q2=50th percentile, Q3=75th percentile. Predicted mean values were obtained using RCIV1 methods, using estimates, as shown in Table 3, and scaled for the average across years. Because the response variable was analyzed on the natural log scale, the predicted values have the form $c \cdot exp(\mathbf{x}\hat{\boldsymbol{\beta}}^{\mathbf{X}} + z_1^* \mathbf{x}\hat{\boldsymbol{\beta}}^{Z_1^*})$ on the original scale, where $c = 1.18$ for the average across years. The graph illustrates that for no-cold days, increases in LTE$_4$ per unit increase in one pollutant are steepest when the second pollutant is lower. However, when children have colds, LTE$_4$ is higher and there is less interaction between pollutants on health. Both outcome and predictors were analyzed on the natural log scale but inverted back for presentation, resulting in curves

which quartiles from distributions of observed values of $W_1$ and $W_2$ were used. The distributions of predicted values are less variable, and so using quartiles from these distributions leads to estimates more on the interior of the distributions of exposures rather than on the extremes.

The estimated slope of the SHS exposure variable in predicting log LTE4—fixing ambient fine particulate matter exposure at its mean level—was approximately three times greater than for ambient fine particulate matter (fixing SHS exposure at its mean level), both for those without and with colds (Table 3). Although this difference was not significant for either those with or without colds ($p > 0.2$), the consistency of estimates between sick and well children is notable. For those without colds, the SHS slope at a fixed level $c$ for ambient fine particulate matter was greater than the slope of ambient fine particulate matter slope at level $c$ of SHS with marginal significance ($p = 0.047$), while it was very insignificant for those with colds. (This is a different test than when comparing at the mean of the "other" pollutant because the means of the two pollutants differed.) The weaker significance of tests involving those with colds is at least in part due to fewer records under those conditions (approximately 20%).

In the absence of SHS, well children had a marginally significant relationship between ambient fine particulate matter and log(LTE4), while there was no apparent relationship when children were sick (last row, Table 3). The comparison of SHS slopes in absence of ambient fine particulate matter is less relevant and not included because children are generally exposed to ambient fine particulate matter, but not necessarily cigarette smoke.

The increase in LTE$_4$ per IQR increase in one pollutant at the mean of the other pollutant ranged from approximately 3 to 15% both for those with and without colds. Fixing both SHS and ambient fine particulate matter at mean values, those with colds had LTE$_4$ values that were approximately 8% higher than those without ($p = 0.02$; obtained as $[exp(4.41) - exp(4.33)]/exp(4.33)$, based on intercepts shown in Table 3). This suggests that having a cold has an acute effect on LTE$_4$ that is approximately the same as an IQR increase in a pollutant (fixing non-involved pollutants at mean values).

The larger random subject variances for $W_1$ and $Y$ models are expected; the first occurring due to differences in smoking behaviors of those living with the study subjects, and the second due to natural subject heterogeneity in LTE$_4$ levels. Both the random subject variance and within-subject correlation were relatively low for $W_2$ models, but for consistency, we kept the terms in that model. Appendix D has a discussion of the predicted values for models without covariate interaction terms.

## 4. SIMULATION

Simulations were conducted in order to determine accuracy of the regression calibration estimators and to assess the coverage rates of confidence intervals. The models used were similar to those in the data application, without year indicators (i.e., no $Z$ variables in the model). One indicator covariate ($Z_1^*$) was included and was allowed to interact with both $X$ variables as well as their interaction (i.e., the "full" interaction model (1) was considered). $Z_1^*$ was included to mimic the "cold" variable, with the same approximate rate of event occurrence ($z_1^* = 1$ for about 20% of the data). The simulations generally used parameter settings that lead to parameter estimates similar to those observed in the data application.

The simulation conditions were as follows: sample sizes ranged from $n = 50$ to $n = 500$ subjects, with $r = 10$ to $r = 20$ repeated measures per subject on consecutive days. Complete records were simulated so that both model-based and empirical variance estimators could be determined. Errors and random subject intercepts were generated from normal distributions. Additional conditions (e.g., different correlation parameter values or error distribution types) were considered in Strand *et al.* (2014) for models without covariates.

Overall, the results here are similar to those presented in Strand *et al.* (2014) for models without covariate interaction terms, but more complex pollutant models. The results of some of the simulations are presented in Tables 4 and 5. Generally, the results indicate that the methods described in Section 2 yield confidence intervals for parameters of interest that have appropriate (nominal) coverage rate, with some slight overcoverage when using model-based variances, for smaller sample sizes.

**Table 4.** MSE and Bias simulation results for RCIV1

| Para-meter | MSE | | | | | Bias | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 50$ $r = 10$ | $n = 100$ $r = 10$ | $n = 200$ $r = 10$ | $n = 100$ $r = 20$ | $n = 500$ $r = 10$ | $n = 50$ $r = 10$ | $n = 100$ $r = 10$ | $n = 200$ $r = 10$ | $n = 100$ $r = 20$ | $n = 500$ $r = 10$ |
| $\beta_0^X$ | 3.12 | 1.83 | 1.20 | 1.04 | 0.87 | 0.04 | 0.07 | −0.003 | 0.003 | −0.005 |
| $\beta_1^X$ | 2.08 | 1.26 | 0.80 | 0.69 | 0.58 | −0.04 | −0.05 | 0.01 | 0.001 | 0.005 |
| $\beta_2^X$ | 1.64 | 0.97 | 0.64 | 0.57 | 0.46 | −0.02 | −0.04 | 0.0003 | −0.003 | 0.003 |
| $\beta_3^X$ | 1.09 | 0.67 | 0.43 | 0.37 | 0.31 | 0.02 | 0.02 | −0.005 | 0.000 | −0.003 |
| $\beta_0^{Z^*}$ | 6.22 | 3.70 | 2.79 | 2.66 | 2.23 | −0.05 | −0.17 | 0.06 | 0.01 | −0.08 |
| $\beta_1^{Z^*}$ | 5.65 | 4.48 | 4.25 | 4.18 | 4.03 | 0.02 | 0.11 | −0.05 | −0.005 | 0.05 |
| $\beta_2^{Z^*}$ | 4.35 | 3.42 | 3.25 | 3.20 | 3.07 | 0.02 | 0.09 | −0.03 | −0.007 | 0.04 |
| $\beta_3^{Z^*}$ | 2.93 | 2.30 | 2.17 | 2.13 | 2.05 | −0.01 | −0.06 | 0.02 | 0.004 | −0.03 |

One thousand replicates per condition, using normal errors; $r$ = number of consecutive-day repeated measures per subject; $n$ = subject sample size. Covariance and fixed-effect parameter settings were chosen so that fitted models had similar parameter estimates as those observed for the data application; for fixed-effect parameters, the following values were used: $\beta_0^X = 1.4$, $\beta_1^X = 1.9$, $\beta_2^X = 1.4$, $\beta_3^X = -0.9$, $\beta_0^{Z_1^*} = 3.1$, $\beta_1^{Z_1^*} = -2.05$, $\beta_2^{Z_1^*} = -1.6$, $\beta_3^{Z_1^*} = 1.1$ (related to $X$ variables that were not mean corrected). For more details, see Section 4.

**Table 5.** Simulation confidence interval coverage rates for RCIV1

| Conditions | Parameter set | Model-based variance | | Empirical variance | |
|---|---|---|---|---|---|
| | | Min | Max | Min | Max |
| $n = 50, r = 10$ | $\beta^X$ | **96.6** | **96.9** | 95.7 | 96.2 |
| | $\beta^{Z^*}$ | **96.9** | **98.6** | **92.4** | 95.1 |
| $n = 100, r = 10$ | $\beta^X$ | 94.7 | 95.1 | 94.7 | 95.1 |
| | $\beta^{Z^*}$ | 95.2 | **96.9** | 94.5 | 95.3 |
| $n = 200, r = 10$ | $\beta^X$ | 95.5 | 95.9 | 95.0 | 95.7 |
| | $\beta^{Z^*}$ | **97.5** | **98.1** | 95.4 | **96.5** |
| $n = 100, r = 20$ | $\beta^X$ | 94.2 | 94.8 | 94.3 | 94.9 |
| | $\beta^{Z^*}$ | 96.1 | **96.7** | 94.0 | 95.3 |
| $n = 300, r = 10$ | $\beta^X$ | 94.5 | 95.0 | 94.6 | 95.0 |
| | $\beta^{Z^*}$ | 95.9 | **96.6** | 94.8 | 95.1 |
| $n = 500, r = 10$ | $\beta^X$ | 94.5 | 95.3 | 95.1 | 95.4 |
| | $\beta^{Z^*}$ | 95.4 | 95.9 | 94.2 | 95.5 |

One thousand replicates per condition; $r$ = number of consecutive-day repeated measures per subject; $n$ = subject sample size. Values are minimum (or maximum) confidence interval (CI) coverage rate among four parameters in set. For 1000 replicates, a correct method will have coverage within 1.4% of 95%, with 95% probability; CI coverages outside of this range are in bold. Other conditions are the same as mentioned in Table 4.

The properties of estimators and confidence intervals seemed to differ slightly between coefficients related to "main effects" ($\beta_0^X, \ldots, \beta_3^X$) and interactions ($\beta_0^{Z^*}, \ldots, \beta_3^{Z^*}$). In particular, more precision was apparent for estimators of "main effect" coefficients, likely because of more data available to estimate them (subject-days with no colds); the interaction coefficients ($\beta_0^{Z^*}, \ldots, \beta_3^{Z^*}$) involve comparisons between "cold" versus "no-cold" conditions.

Mean bias was not significantly nonzero (for $\alpha = 0.05$ with $t$-tests) across simulation sets. There were slight trends of bias in some simulations, although it did not appear to adversely affect confidence interval coverage rates, and the bias tended to average out between sets of parameter estimates. For example, when $\beta_0^X$ and $\beta_3^X$ had negative bias, $\beta_1^X$ and $\beta_2^X$ were positive, and vice versa. Similar results were observed for $\beta_0^{Z^*}, \ldots, \beta_3^{Z^*}$ although even more pronounced (likely because of greater standard errors for those estimates). These patterns were not consistent across simulations, and together with the nonsignificant observed bias, there is no indication of systematic problems with the estimation method.

The results suggest that using empirical variances may be more reliable (when possible), although model-based variances did not perform too poorly. Because of the fact that the variances are asymptotic in nature, it is not recommended that sample sizes much less than 100 subjects be used. The simulations discussed earlier and presented in the tables generated data from random terms that were independent between as well as within models. A set of simulations was also conducted using correlated random terms between models in (4) in order to determine how violating the assumption affected results, as discussed in Section 5.

# 5. DISCUSSION

In our analyses, we found that while pollutants appeared to have additive effects when children were sick, there was a marginally significant interaction between pollutants when they were not sick, such that the health–pollutant relationship for one pollutant was stronger when the other pollutant was lower. This difference in interaction terms (i.e., the three-way interaction) suggests a complex relationship between exposure to pollutants and cold status of children with asthma, with respect to $LTE_4$ levels. One possible explanation for the observed results is that when they are not sick, children are more able to maintain homeostatis by limiting $LTE_4$ responses during co-exposures to pollutants, relative to when they are sick. However, more research is certainly warranted in order to not only test this hypothesis but also to verify the observed patterns here.

The conditions and restrictions of the study should be kept in mind here, which include but are not limited to the following: fine particulate matter is only one type of air pollution that does not include toxic gases such as carbon monoxide; the study group here involves children with asthma and not the general population; relatively low levels of fine particulate matter and SHS were observed, and results cannot necessarily be extrapolated to higher, unobserved levels—ambient fine particulate matter concentrations in parts of China are often well over 10 times that observed in Denver (e.g., see Venners *et al.*, 2003); and the toxicity of ambient fine particulate matter in Denver may differ from that in other cities because of different compositions.

Specific forms of regression calibration estimators depend on exactly which terms are included in the models. Special cases for models with one predictor measured with error could be obtained easily from the results here. Methods could be extended to models with more than two predictors using the same approach given in this paper; however, notation would become burdensome.

Results here and in Strand *et al.* (2014) suggest that regression calibration for more complicated longitudinal models with interaction terms can yield fairly accurate estimators despite approximations made to covariance structures. In both of these studies, some over-coverage was apparent for relatively small sample sizes (no larger than 100) for model-based variances estimators. Buonaccorsi *et al.* (2000) also found overestimated variances of regression calibration estimators associated with random effect models, for estimation without instrumental variable estimators; he also suggested modified approaches to reduce this overestimation. Whether or not these results are due to the same issue, the slight overestimation of variances of $\hat{\boldsymbol{\beta}}$ here did not appear to be large enough to consider adjustments, as long as adequate sample sizes are used.

Certain assumptions were made that allowed us to derive the point estimators, and we discuss a few here with respect to our application. The non-differential error (or surragocy) assumption essentially states that the unbiased surrogate and instrumental variables do not provide any more information in predicting the health outcome beyond what the actual exposure variables and covariates do. Because cotinine is a metabolite specific to nicotine, we believe it is reasonable to assume that it will not provide more information in predicting $LTE_4$ than actual SHS exposure. An exception to this would be the fact that cotinine may be affected by the gas portion of secondhand smoke, which could potentially influence $LTE_4$ and is not included in estimated or actual SHS particulate exposure. (In this article, "SHS" only includes particulate.) Similarly, we believe it is reasonable to assume that outdoor fine particulate matter measured at a fixed location will not provide more information in predicting $LTE_4$ than the direct exposures to this pollutant. The assumption that $W$ variables are surrogates for $X$ variables seems even more logical, although the unbiased assumption is not easily verifiable.

Further examination of our data indicated the assumption that random terms between pollutant models in (4) are uncorrelated was violated to a small degree. However, additional simulations that included correlation of approximately the same magnitude as was observed had a negligible effect on the point estimates. Another assumption in our models is that the instrumental variables are not measured with error. There is some potential for error in cotinine and creatinine measurements, although it is not expected to be greater than that of the $W$ variables.

One may wonder whether regressing the health outcome directly on unbiased surrogate variables is a reasonable alternative to using regression calibration as described in this article. While easier, this approach not only has the disadvantage that the measurement error in the predictors will attenuate the estimates by introducing "classical measurement error", but also much less data are available to fit this model. This approach yielded effects that were not close to significant (results not presented).

In our application, we used a binary covariate that interacted with the pollutant variables to illustrate the methods. Because it is just a binary variable, one could also perform stratified analyses for those with colds and those without, eliminating $Z^*$ variables from the models. However, doing such would not allow for statistical comparisons to be made between those with and those without colds. Using our presented methods, we could also add continuous covariates into the model such as subject height. In some analyses with other data, we did find height to be an effect modifier in the pollution–health relationship. However, for these particular data, height was insignificant and did not contribute to the model and thus was not included.

## APPENDIX A. TRUE COVARIANCE STRUCTURES

In order to derive the covariance between measures $Y_{ij}$ and $Y_{ij'}$ in (5), we first define terms. Let $V_{ijj'}^{X_p} = \sigma_{b^{X_p}}^2 + \mathbf{R}_{\boldsymbol{\omega}_{pi}}(j, j')$, where $\mathbf{R}_{\boldsymbol{\omega}_{pi}}(j, j')$ is the $j^{th}$ row and $(j')^{th}$ column of $\mathbf{R}_{\boldsymbol{\omega}_{pi}}$, and define $\mu_{ij}^{X_p} = E(X_{pij})$ for $p = 1, 2$. Specifically, $\mu_{ij}^{X_1} = \theta_0 + \theta_1 M_{1ij}$ and $\mu_{ij}^{X_2} = \pi_0 + \pi_1 M_{2ij}$ based on (3) [although the next forms will still be applicable for more general two-pollutant models as considered in Strand *et al.* (2014)]. In our derivations, a basic assumption is that random terms between models [including between models in (4) as well as between models in (4) and (5)] are uncorrelated. These same assumptions were used in deriving the "true" covariance structure in Strand *et al.* (2014). To simplify the following notation, let $1\mathbf{Z}_{ij}^* = (1, \mathbf{Z}_{ij}^*)$ and $\boldsymbol{\beta}^{\mathbf{X}, \mathbf{Z}^*} = ([\boldsymbol{\beta}^{\mathbf{X}}]^t, [\boldsymbol{\beta}^{\mathbf{Z}^*}]^t)^t$. We can write

$$Cov(\mathbf{X}_{ij}^t | \mathbf{M}_{ij}, \mathbf{X}_{ij'}^t | \mathbf{M}_{ij'}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & V_{ijj'}^{X_1} & 0 & \mu_{ij'}^{X_2} V_{ijj'}^{X_1} \\ 0 & 0 & V_{ijj'}^{X_2} & \mu_{ij'}^{X_1} V_{ijj'}^{X_2} \\ 0 & \mu_{ij}^{X_2} V_{ijj'}^{X_1} & \mu_{ij}^{X_1} V_{ijj'}^{X_2} & V_{4,4} \end{pmatrix}$$

where $V_{4,4} = \mu_{ij}^{X_2} \mu_{ij'}^{X_2} V_{ijj'}^{X_1} + \mu_{ij}^{X_1} \mu_{ij'}^{X_1} V_{ijj'}^{X_2} + V_{ijj'}^{X_1} V_{ijj'}^{X_2}$. Now based on our models,

$$Cov\left(Y_{ij} | \mathbf{M}_{ij}, \mathbf{Z}_{ij}, \mathbf{Z}_{ij}^*; Y_{ij'} | \mathbf{M}_{ij'}, \mathbf{Z}_{ij'}, \mathbf{Z}_{ij'}^*\right) = f + \sigma_{b^Y}^2 + \mathbf{R}_{\boldsymbol{\epsilon}_i}(j, j') \tag{A.1}$$

where $f = (\boldsymbol{\beta}^{\mathbf{X}, \mathbf{Z}^*})^t \left([(1\mathbf{Z}_{ij}^*)^t (1\mathbf{Z}_{ij'}^*)] \otimes Cov[\mathbf{X}_{ij}^t | \mathbf{M}_{ij}, \mathbf{X}_{ij'}^t | \mathbf{M}_{ij'}]\right) \boldsymbol{\beta}^{\mathbf{X}, \mathbf{Z}^*}$ is a function of fixed-effect and covariance parameters in (1) and (3) as well as instrumental variables and covariates. Equation (A.1) demonstrates that the covariance between responses within a subject based on (5) involves the additional term $f$, relative to the covariance between such responses in (1). Via simulation, we found that using an approximated covariance structure in the fit of (5) that essentially dropped $f$ from the calculation worked well, with an additional advantage that unknown parameters in $f$ from models (1) and (3) were removed from the calculation. A substantially nonzero $f$ may not indicate whether the approximation will work because of the fact that the parameters will take on different values and meanings in the approximated structure.

The true covariance structures for responses in (4) differ from those in (3) because of the added error $U_p$. If $U_p$ is uncorrelated over time and uncorrelated with $\omega_{pij}$, then the covariance structure simply adds $Var(U_p)$ down the diagonal relative to that in (3), and this can easily be modeled using standard software. However, in our application, the best results appeared to be obtained when using the same structure in (4) as was assumed in (3), that is, the "random intercept plus spatial power" structure (although it is still likely just an approximation).

The covariance between $Y_{ij}$ and $W_{pij'}$, given instrumental variables and covariates for $p = 1, 2$ [i.e., covariances between responses in (4) and (5)] are functions of $\mathbf{Z}_{ij}^*$, $\boldsymbol{\beta}^{\mathbf{X}}$, $\boldsymbol{\beta}^{\mathbf{Z}^*}$, $V_{ijj'}^{X_p}$, and $\mu_{ij}^{X_p}$ (the full form is available from the authors upon request). However, model-based forms set between-model covariances to 0, as discussed in Section 2.3, and empirical covariance forms employ residuals.

## APPENDIX B. PARTIAL INTERACTION MODELS

Often, an analyst may be interested in only including a subset of covariate interactions rather than "all possible" interactions in the model. In this case, the mean values become

$$E(Y_{ij} | \mathbf{M}_{ij}, \mathbf{Z}_{ij}, \mathbf{Z}_{ij}^*) = \mathbf{M}_{ij} \mathbf{C} \boldsymbol{\beta}^{\mathbf{X}} + \mathbf{Z}_{ij} \boldsymbol{\beta}^{\mathbf{Z}} + \sum_{p=1}^{m} Z_{pij}^* \mathbf{M}_{ij}^{(p)} \mathbf{C}_{(p)} \boldsymbol{\beta}_{(p)}^{Z_p^*}$$

where $\mathbf{C}_{(p)}$ and $\mathbf{M}_{ij}^{(p)}$ are the subsets of $\mathbf{C}$ [in (6)] and $\mathbf{M}_{ij}$, respectively, that correspond to the subset $\mathbf{X}_{ij}^{(p)}$ of $\mathbf{X}_{ij}$ in (2). To illustrate, consider the case where there are three covariates; the first ($Z_1^*$) is allowed to interact with all terms in $\mathbf{X}_{ij}$; the second ($Z_2^*$) with all terms except $X_1 \cdot X_2$; and the third ($Z_1$) not with any $X$ terms. We note that $\mathbf{X}_{ij}^{(1)} = (1, X_1, X_2, X_1 \cdot X_2)$ and $\mathbf{X}_{ij}^{(2)} = (1, X_1, X_2)$. Model (2) in this case becomes

$$Y_{ij} = \beta_0^X + \beta_1^X X_{1ij} + \beta_2^X X_{2ij} + \beta_3^X X_{1ij} X_{2ij} + \beta^{Z_1} Z_{1ij}$$

$$+ \beta_0^{Z_1^*} Z_{1ij}^* + \beta_1^{Z_1^*} Z_{1ij}^* X_{1ij} + \beta_2^{Z_1^*} Z_{1ij}^* X_{2ij} + \beta_3^{Z_1^*} Z_{1ij}^* X_{1ij} X_{2ij}$$

$$+ \beta_0^{Z_2^*} Z_{2ij}^* + \beta_1^{Z_2^*} Z_{2ij}^* X_{1ij} + \beta_2^{Z_2^*} Z_{2ij}^* X_{2ij}$$

$$+ b_i^Y + \epsilon_{ij}$$

and

$$\mathbf{C}_1 = \begin{pmatrix} 1 & \theta_0 & \pi_0 & \theta_0 \pi_0 \\ 0 & \theta_1 & 0 & \theta_1 \pi_0 \\ 0 & 0 & \pi_1 & \theta_0 \pi_1 \\ 0 & 0 & 0 & \theta_1 \pi_1 \end{pmatrix} \quad , \qquad \mathbf{C}_2 = \begin{pmatrix} 1 & \theta_0 & \pi_0 \\ 0 & \theta_1 & 0 \\ 0 & 0 & \pi_1 \end{pmatrix}$$

The estimators for the partial interaction model can be expressed as

$$\hat{\boldsymbol{\beta}}^{\mathbf{X}} = \hat{\mathbf{C}}^{-1} \hat{\boldsymbol{\gamma}}^{\mathbf{M}}$$

$$\hat{\boldsymbol{\beta}}^{\mathbf{Z}} = \hat{\boldsymbol{\gamma}}^{\mathbf{Z}}$$

$$\hat{\boldsymbol{\beta}}^{Z_p^*} = \hat{\mathbf{C}}_{(p)}^{-1} \hat{\boldsymbol{\gamma}}_{(p)}^{Z_p^*}, \qquad p = 1, \ldots, m$$

where $\hat{\boldsymbol{\gamma}}_{(p)}^{Z_p^*}$ are estimated coefficients of $Z_p^* \mathbf{M}_{ij}^{(p)}$ in the model analogous to (5) but for partial interactions. Note: $m = 2$ for the illustration.

If we apply the constraint that lower-order terms must always be included in models when higher-order terms are, then there are four possible nontrivial subsets of $\mathbf{X}_{ij}$ that could be used for $\mathbf{X}_{ij}^{(p)}$ for each $p$: $(1, X_1, X_2, X_1 \cdot X_2)$, $(1, X_1, X_2)$, $(1, X_1)$, and $(1, X_2)$. Because there are only four possible forms of $\mathbf{C}$ with this reasonable constraint, constructing a program to handle different partial models for real data would not be too difficult.

## APPENDIX C. CALCULATING VARIANCE: FORMS OF $\boldsymbol{\Delta}$ AND $\boldsymbol{\Sigma}$

Let the predictors in (5) associated with $\boldsymbol{\gamma}^t$ be denoted as $(\mathbf{MZ})_{ij} = (\mathbf{M}_{ij}, \mathbf{Z}_{ij}, \mathbf{Z}_{ij}^* \otimes \mathbf{M}_{ij})$. Then let $(\mathbf{MZ})_i$ denote the subject version, obtained by stacking $(\mathbf{MZ})_{ij}$ rows vertically, for $j = 1, \cdots, r_i$. Define $\mathbf{M}_{pij} = (1, M_{pij})$ for $p = 1, 2$, and let $\mathbf{Y}_i$, $\mathbf{W}_{pi}$, and $\mathbf{M}_{pi}$ denote subject versions of $Y_{ij}$, $W_{pij}$, and $M_{pij}$, respectively. ($\mathbf{Y}_i$, $\mathbf{W}_{1i}$ and $\mathbf{W}_{2i}$ are each $r_i \times 1$ vectors; $\mathbf{M}_{pi}$ is an $r_i \times 2$ matrix; $(\mathbf{MZ})_i$ is an $r_i \times [4(m+1)+k]$ matrix).

Elements of (7) in the text are calculated as follows.

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} \mathbf{M}_{1i} \right)^{-1} \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} Var(\mathbf{W}_{1i}|\mathbf{M}_{1i}) \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} \mathbf{M}_{1i} \right) \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} \mathbf{M}_{1i} \right)^{-1}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\pi}} = \left( \sum_{i=1}^n \mathbf{M}_{2i}^t \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right)^{-1} \left( \sum_{i=1}^n \mathbf{M}_{2i}^t \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} Var(\mathbf{W}_{2i}|\mathbf{M}_{2i}) \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right) \left( \sum_{i=1}^n \mathbf{M}_{2i}^t \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right)^{-1}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\gamma}} = \left( \sum_{i=1}^n (\mathbf{MZ})_i^t \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right)^{-1} \left( \sum_{i=1}^n (\mathbf{MZ})_i^t \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} Var(\mathbf{Y}_i|(\mathbf{MZ})_i) \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right) \left( \sum_{i=1}^n (\mathbf{MZ})_i^t \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right)^{-1}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta\pi}} = \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} \mathbf{M}_{1i} \right)^{-1} \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} Cov(\mathbf{W}_{1i}|\mathbf{M}_{1i}, \mathbf{W}_{2i}|\mathbf{M}_{2i}) \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right) \left( \sum_{i=1}^n \mathbf{M}_{2i}^t \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right)^{-1}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta\gamma}} = \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} \mathbf{M}_{1i} \right)^{-1} \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} Cov(\mathbf{W}_{1i}|\mathbf{M}_{1i}, \mathbf{Y}_i|(\mathbf{MZ})_i) \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right) \left( \sum_{i=1}^n (\mathbf{MZ})_i^t \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right)^{-1}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\pi\gamma}} = \left( \sum_{i=1}^n \mathbf{M}_{2i}^t \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right)^{-1} \left( \sum_{i=1}^n \mathbf{M}_{2i}^t \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} Cov(\mathbf{W}_{2i}|\mathbf{M}_{2i}, \mathbf{Y}_i|(\mathbf{MZ})_i) \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right) \left( \sum_{i=1}^n (\mathbf{MZ})_i^t \mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i}^{-1} (\mathbf{MZ})_i \right)^{-1}$$

where $\mathbf{V}_{\mathbf{Y}_i|(\mathbf{MZ})_i} = Var(\mathbf{Y}_i|(\mathbf{MZ})_i)$ and $\mathbf{V}_{\mathbf{W}_{pi}|\mathbf{M}_{pi}} = Var(\mathbf{W}_{pi}|\mathbf{M}_{pi})$, for $p = 1, 2$, based on models (4) and (5) and other terms are as previously defined. Empirical forms can be obtained by employing residuals to estimate the middle "$Var$" and "$Cov$" quantities in the aforementioned equations. Model-based forms of block diagonal elements reduce to

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \left( \sum_{i=1}^n \mathbf{M}_{1i}^t \mathbf{V}_{\mathbf{W}_{1i}|\mathbf{M}_{1i}}^{-1} \mathbf{M}_{1i} \right)^{-1}$$

$$\Sigma_{\pi} = \left( \sum_{i=1}^{n} \mathbf{M}_{2i}^{t} \mathbf{V}_{\mathbf{W}_{2i}|\mathbf{M}_{2i}}^{-1} \mathbf{M}_{2i} \right)^{-1}$$

$$\Sigma_{\gamma} = \left( \sum_{i=1}^{n} (\mathbf{MZ})_{i}^{t} \mathbf{V}_{\mathbf{Y}_{i}|(\mathbf{MZ})_{i}}^{-1} (\mathbf{MZ})_{i} \right)^{-1}$$

Model-based block off-diagonal elements were set to **0** for analysis.

The $(i,j)^{th}$ element of $\boldsymbol{\Delta}$ is $\frac{\partial \boldsymbol{\beta}[i]}{\partial \boldsymbol{\alpha}[j]}$, where $\boldsymbol{\beta}[i]$ denotes the $i^{th}$ element of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}[j]$ denotes the $j^{th}$ element of $\boldsymbol{\alpha}$. The matrix can be expressed easily by considering submatrices with similar forms. Let

$$\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{X\theta} & \boldsymbol{\Delta}_{X\pi} & \boldsymbol{\Delta}_{\gamma} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{k} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{\Delta}_{Z_{1}^{*}\theta} & \boldsymbol{\Delta}_{Z_{1}^{*}\pi} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Delta}_{\gamma} & \cdots & \mathbf{0} \\ \boldsymbol{\Delta}_{Z_{2}^{*}\theta} & \boldsymbol{\Delta}_{Z_{2}^{*}\pi} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Delta}_{Z_{m}^{*}\theta} & \boldsymbol{\Delta}_{Z_{m}^{*}\pi} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Delta}_{\gamma} \end{pmatrix}$$

where

$$\boldsymbol{\Delta}_{\gamma} = \mathbf{C}^{-1}$$

$$\boldsymbol{\Delta}_{X\theta}^{t} = \frac{1}{\pi_{1}\theta_{1}^{2}} \begin{pmatrix} \theta_{1}c_{1}^{M} & 0 & -\theta_{1}\gamma_{3}^{M} & 0 \\ -\theta_{0}c_{1}^{M} & c_{1}^{M} & \theta_{0}\gamma_{3}^{M} & -\gamma_{3}^{M} \end{pmatrix}$$

$$\boldsymbol{\Delta}_{X\pi}^{t} = \frac{1}{\pi_{1}^{2}\theta_{1}} \begin{pmatrix} \pi_{1}c_{2}^{M} & -\pi_{1}\gamma_{3}^{M} & 0 & 0 \\ -\pi_{0}c_{2}^{M} & \pi_{0}\gamma_{3}^{M} & c_{2}^{M} & -\gamma_{3}^{M} \end{pmatrix}$$

$$c_{1}^{M} = \pi_{0}\gamma_{3}^{M} - \pi_{1}\gamma_{1}^{M}$$

$$c_{2}^{M} = \theta_{0}\gamma_{3}^{M} - \theta_{1}\gamma_{2}^{M}$$

Matrices $\boldsymbol{\Delta}_{Z_{p}^{*}\theta}^{t}$ and $\boldsymbol{\Delta}_{Z_{p}^{*}\pi}^{t}$ are the same as $\boldsymbol{\Delta}_{X\theta}^{t}$ and $\boldsymbol{\Delta}_{X\pi}^{t}$, respectively, except all "$X$" are replaced with "$Z_{p}^{*}$", for $p = 1, ..., m$; note that $\mathbf{I}_{k}$ is a $k \times k$ matrix and that dimensions of $\mathbf{0}$ matrices can be determined from surrounding matrices.

The formulas for $\boldsymbol{\Delta}$ and $\boldsymbol{\Sigma}$ given here are similar to those presented in Strand *et al.* (2014), with some key differences as follows: (i) those in Strand *et al.* (2014) are derived for models that have $M_{1}$ and $M_{2}$ in both equations for (3) and (4); (ii) formulas given in the Appendix of Strand *et al.* (2014) do not include covariates, although non-interactive covariates were generally considered in that article. (Including non-interactive covariates would essentially involve replacing $\mathbf{M}_{i}^{red}$ with $(\mathbf{M}_{i}^{red}, \mathbf{Z}_{i})$ for every occurrence in $\boldsymbol{\Sigma}_{\gamma}$, $\boldsymbol{\Sigma}_{\theta\gamma}$, and $\boldsymbol{\Sigma}_{\pi\gamma}$, based on notation used in that article; note also that conditioning on $\mathbf{Z}_{i}$ in formulas was suppressed part-way through that article for convenience.)

Similar steps as those provided here could be taken to generalize $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$ to account for the more general forms of (3) and (4).

## APPENDIX D. PREDICTED VALUES FOR MODELS WITHOUT COVARIATE INTERACTION ($Z^{*}$) TERMS

When there are no covariate interaction terms in (1) (i.e., no $Z^{*}$ variables), predicted values as a function of pollutants, $\mathbf{x} = (1, x_{1}, x_{2}, x_{1} \cdot x_{2})$, have the form $exp(\mathbf{x}\hat{\boldsymbol{\beta}}^{X} + \mathbf{z}\hat{\boldsymbol{\beta}}^{Z}) = c \cdot exp(\mathbf{x}\hat{\boldsymbol{\beta}}^{X})$, where $c$ is a constant that depends on which values of non-interactive covariates are considered. As an example, consider the analysis performed in Strand *et al.* (2014) that was the same as the analysis in this article, but without the cold indicator variable (and hence no covariate interaction terms), and performed only for days when children were not sick. The value of $c$ depends on the combination of study years being considered. For the reference year (Year 4), $c = exp(0) = 1$; for the average across years, $c = exp((0.048 + 0.204 + 0.292 + 0)/4) = 1.15$, based on estimates from that article. Although not specifically stated, actual plotted values in Figure 1 of Strand *et al.* (2014) are predicted values averaged over the 4 years.

# REFERENCES

Bartlett JW, De Stavola BL, Frost C. 2009. Linear mixed models for replication data to efficiently allow for covariate measurement error. *Statistics in Medicine* **28**:3158–3178.

Buonaccorsi JP. 2000. Estimation in longitudinal random effects models with measurement error. *Statistica Sinica* **10**:885–903.

Buonaccorsi JP. 2010. *Measurement Error: Models, Methods and Applications*. Chapman and Hall/CRC: Boca Raton.

Buzas JS, Stefanski LA. 1996. Instrumental variable estimation in generalized measurement error models. *Journal of the American Statistician* **91**:999–1006.

Carroll RJ, Ruppert D, Stephanski LA. 2006. *Measurement Error in Nonlinear Models: A Modern Perspective,* 2nd edition. Chapman and Hall/CRC: Boca Raton.

Carroll RJ, Lin X, Wang N. 1997. *Generalized linear mixed measurement error models.* In Modeling Longitudinal and Spatially Correlated Data, Gregoire T (ed). Springer Verlag, New York.

Chilmonczyk BA, Salmun LM, Megathlin KN, Neveaux LM, Palomaki GE, Knight GJ, Pulkkinnen AJ, Haddow JE. 1993. Association between exposure to environmental tobacco smoke and exacerbations of asthma in children. *New England Journal of Medicine* **328**:1665–1669.

Hardin JW, Carroll RJ. 2003. Variance estimation for the instrumental variables approach to measurement error in generalized linear models. *The Stata Journal* **3**(4):342–350.

Huang L-S, Wang H, Cox C. 2005. Assessing interaction effects in linear measurement error models. *Applied Statistics* **54**(1):21–30.

Ko H, Davidian M. 2000. Correcting for measurement error in individual-level covariates in nonlinear mixed effects models. *Biometrics* **56**:368–375.

Li H, Wang L. 2012. Consistent estimation in generalized linear mixed models with measurement error. *Journal of Biometrics and Biostatistics* S7–007. doi:10.4172/2155-6180.S7-007.

Ostro B, Lipsett M, Mann J, Braxton-Owens H, White M. 2001. Air pollution and exacerbation of asthma in African-American children in Los Angeles. *Epidemiology* **12**:200–208.

Peters A, Dockery DW, Heinrich J, Wichmann HE. 1997. Short-term effects of particulate air pollution on respiratory morbidity in asthmatic children. *European Respiratory Journal* **10**:872–879.

Rabinovitch N, Silveira L, Gelfand EW, Strand M. 2011. The response of children with asthma to ambient particulate is modified by tobacco smoke exposure. *American Journal or Respiratory and Critical Care Medicine* **184**:1350–1357.

Rabinovitch N, Strand M, Gelfand EW. 2006. Particulate levels are associated with early asthma worsening in children with persistent disease. *American Journal of Respiratory and Critical Care Medicine* **173**:1098–1105.

Roemer W, Hoek G, Brunekreef B. 1993. Effect of ambient winter air pollution on respiratory health of children with chronic respiratory symptoms. *American Review of Respiratory Disease* **147**:118–124.

Strand M, Sillau S, Grunwald GK, Rabinovitch N. 2014. Regression calibration for models with two predictor variables measured with error and their interaction, using instrumental variables and longitudinal data. *Statistics in Medicine* **33**:470–87.

Strand M, Hopke PK, Zhao W, Vedal S, Gelfand EW, Rabinovitch N. 2007. A study of health effect estimates using competing methods to model personal exposures to ambient PM$_{2.5}$. *Journal of Exposure Science and Environmental Epidemiology* **17**:549–558.

Strand M, Vedal S, Rodes C, Dutton SJ, Gelfand EW, Rabinovitch N. 2006. Estimating effects of ambient PM2.5 exposure on health using PM$_{2.5}$ component measurements and regression calibration. *Journal of Exposure Science and Environmental Epidemiology* **16**:30–38. [Erratum: 2007; **17**: 122.]

Thurston SW, Spiegelman D, Ruppert D. 2003. Equivalence of regression calibration methods in main study/external validation study designs. *Journal of Statistical Planning and Inference* **113**:527–539.

Vedal S, Petkau J, White R, Blair J. 1998. Acute effects of ambient air particles in asthmatic and nonasthmatic children. *American Journal or Respiratory and Critical Care Medicine* **157**:1034–1043.

Venners SA, Wang B, Peng Z, Xu Y, Wang L, Xu X. 2003. Particulate matter, sulfur dioxide, and daily mortality in Chongqing, China. *Environmental Health Perspectives* **111**(4):562–567.

Wallace LA, Mitchell H, OŠConnor GT, Neas L, Lippmann M, Kattan M, Koenig J, Stout JW, Vaughn BJ, Wallace D. 2003. Inner-city asthma study. Particle concentrations in inner-city homes of children with asthma: the effect of smoking, cooking, and outdoor pollution. *Environmental Health Perspectives* **111**:1265–1272.

Wang N, Lin X, Gutierrez RG, Carroll RJ. 1998. Generalized linear mixed measurement error models. *Journal of the American Statistical Association* **93**:249–261.

Wang N, Lin X, Gutierrez RG. 1999. A bias correction regression calibration approach in generalized linear mixed measurement error models. *Communications in Statistics – Theory and Methods* **28**(1):217–232.

Wong MY, Day NE, Luan JA, Wareham NJ. 2004. Estimation of magnitude in gene–environment interactions in the presence of measurement error. *Statistics in Medicine* **23**:987–998.