1-3-2007

A BAYESIAN HIERARCHICAL MODEL FOR CONSTRAINED DISTRIBUTED LAG FUNCTIONS: ESTIMATING THE TIME COURSE OF HOSPITALIZATION ASSOCIATED WITH AIR POLLUTION EXPOSURE

Roger Peng
Johns Hopkins Bloombreg School of Public Health, Department of Biostatistics, rpeng@jhsph.edu

Francesca Dominici
Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics

Leah J. Welty
Northwestern University, Department of Preventive Medicine

A Bayesian hierarchical model for constrained distributed lag functions: Estimating the time course of hospitalization associated with air pollution exposure

Roger D. Peng Francesca Dominici Leah J. Welty

January 3, 2007

Abstract

Numerous time series studies have provided strong evidence of an association between increased levels of ambient air pollution and increased levels of hospital admissions, typically at 0, 1, or 2 days after an air pollution episode. An important research aim is to extend existing statistical models so that a more detailed understanding of the time course of hospitalization after exposure to air pollution can be obtained. Information about this time course, combined with prior knowledge about biological mechanisms, could provide the basis for hypotheses concerning the mechanism by which air pollution causes disease. Previous studies have identified two important methodological questions: (1) How can we estimate the shape of the distributed lag between increased air pollution exposure and increased mortality or morbidity? and (2) How should we estimate the cumulative population health risk from short-term exposure to air pollution? Distributed lag models are appropriate tools for estimating air pollution health effects that may be spread over several days. However, estimation for distributed lag models in air pollution and health applications is hampered by the substantial noise in the data and the inherently weak signal that is the target of investigation. We introduce an hierarchical Bayesian distributed lag model that incorporates prior information about the time course of pollution effects and combines information across multiple locations. The model has a connection to penalized spline smoothing using a special type of penalty matrix. We apply the model to estimating the distributed lag between exposure to particulate matter air pollution and hospitalization for cardiovascular and respiratory disease using data from a large United States air pollution and hospitalization database of Medicare enrollees in 94 counties covering the years 1999–2002.

KEY WORDS: Distributed lag model; air pollution; hierarchical model; time series, reproducible research

1 Introduction

Time series studies of air pollution and health in the United States and around the world have provided consistent strong evidence of an adverse effect of ambient air pollution levels on mortality and morbidity (Health Effects Institute, 2003; Pope and Dockery, 2006, and references therein). In particular, multi-site studies, which combine information from many locations, have produced robust and consistent results demonstrating an adverse health effect of particulate matter (PM) and ozone. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS) in the U.S. and the Air Pollution and Health: A European Approach (APHEA) study in Europe are prominent examples of such multi-site studies (Bell et al., 2004; Peng et al., 2005; Katsouyanni et al., 2001; Samoli et al., 2001, 2003). More recently, the Medicare Air Pollution Study (MCAPS) showed a strong association between fine particulate matter (PM less than 2.5 μm in diameter) and hospitalization for cardiovascular and respiratory diseases in 204 U.S. counties (Dominici et al., 2006). These large studies all make clear the advantages of the multi-site approach: Combining information across locations improves the precision of relative risk estimates and allows for the examination of variation in estimates between cities.

Previous time series studies of the health effects of PM have generally focused on estimating the effect of increased levels of pollution over a fixed time lag. For example, hospitalization rates will often be compared with PM levels on the same day or one day previous. However, this model, referred to as a *single lag* model, assumes that all of the effect of pollution (if any) plays out at a fixed number of days in the future. While such an assumption might be plausible for modeling a given individual's response, it is less realistic for describing population level associations.

A *distributed lag* model, which allows the effect of an increase in PM levels to play out over multiple days in the future, is a more informative tool for characterizing the time course of the effect of increased pollution levels on hospitalization. For example, it might be reasonable to assume that at the population level, the effect of PM on mortality and morbidity outcomes is distributed *smoothly* over multiple days into the future. Distributed lag models have been used for decades in economics (Almon, 1965; Leamer, 1972; Shiller, 1973) and have been applied more recently in the area of environmental epidemiology to estimate the short-term effects of air pollution on mortality and morbidity (e.g. Zanobetti et al., 2000; Schwartz, 2000; Bell et al., 2004).

Distributed lag models offer the advantage of being able to estimate the *distributed lag function*, which describes the change over time in the relative risk associated with a given day's increase in air pollution levels. Information about the shape of the distributed lag function, combined with knowledge about biological mechanisms, could provide useful evidence concerning the time course of disease progression and perhaps shed light on the mechanism by which air pollution causes disease.

Given the distributed lag function we can also estimate the cumulative effect of a single day's increase in pollution levels. The cumulative effect can be obtained by integrating (summing) the distributed lag function over the time period where the effect of PM is deemed relevant. It has been hypothesized that if the effect of air pollution is in fact spread out over multiple days, then single lag models will tend to underestimate the cumulative effect of an air pollution episode (Zanobetti et al., 2000; Schwartz, 2000; Zanobetti et al., 2002;

Goodman et al., 2004; Roberts, 2005; Schwartz, 2006). An alternative hypothesis, sometimes referred to as the "harvesting" or "mortality displacement" hypothesis, claims that air pollution episodes deplete a frail pool of individuals and decrease the number of susceptible people on future days (Schimmel and Murawski, 1976). Such a phenomenon would lead to a distributed lag function that is negative for certain periods and, when summed over the relevant time period, may result in a cumulative effect that is less then estimates obtained from single lag models (Zeger et al., 1999; Zanobetti et al., 2000; Dominici et al., 2002b).

The MCAPS study (Dominici et al., 2006) linked data from the U.S. Environmental Protection Agency's (EPA) $PM_{2.5}$ monitoring network with health data from the United States Medicare system and found strong associations between $PM_{2.5}$ and hospitalization for cardiovascular and respiratory diseases. Their approach used single lag models and estimated independent effects at fixed lags of 0, 1, and 2 day. One issue highlighted by that study is the variation in risk estimates across different lags. Without an estimate of the distributed lag function, it is difficult to interpret the different estimates across lags since one would typically expect them to be correlated.

Until recently, the monitoring pattern of particulate matter data made the application of distributed lag models challenging for more than a handful of counties. Before 1999, the US EPA collected data on particulate matter less than 10 μm in diameter (PM_{10}), generally on a 1-in-6 day basis (i.e. for every six days, one measurement of PM_{10} is made). Beginning in 1999, the EPA began collecting data for particulate matter less than 2.5 μm in diameter ($PM_{2.5}$), sometimes referred to as fine particulate matter. Monitoring for $PM_{2.5}$ is generally conducted on a 1-in-3 day basis, but there are many more counties where measurements are taken everyday. With the emergence of the new $PM_{2.5}$ monitoring network, we are able to fit distributed lag models to data from more counties than previously possible and capture more precisely the time-course of particulate matter health effects.

Previous approaches to applying distributed lag models to air pollution and health data have generally been implemented in two stages. In the first stage data for a specific location are used to estimate the "raw" distributed lag function for that location. This function is then smoothed using a polynomial or other smoother (e.g. Almon, 1965; Corradi, 1977; Zanobetti et al., 2000). If data are available from multiple locations, then the smooth curves can be combined in a second stage to form an overall estimate. The combining is typically accomplished via an hierarchical model over the locations. Welty et al. (2005) proposed a Bayesian model for estimating the distributed lag function in a time series study of a single location. They used an informative prior that constrains the shape of the function by allowing effects corresponding to early lags to take any value while effects at more distant lags are constrained to be near zero and relatively smooth. Their approach can also be formulated as an application of penalized splines with a special type of penalty term. The use of prior information in the estimation of distributed lag functions is an important step forward, however there is a need to extend the model to take advantage of the availability of data from multiple locations.

We introduce a Bayesian hierarchical model for constrained distributed lag functions and use the model to estimate the distributed lag between ambient air pollution exposure and hospitalizations for cardiovascular and respiratory diseases. The model uses prior knowledge about the shape of the distributed lag function and extends the work of Welty et al. (2005) by introducing an hierarchical structure for combining information from multiple sites. The

incorporation of both prior information and information from other locations provides much improved estimates of the distributed lag function than have been obtained previously and gives us the ability to explore heterogeneity in the shapes of the location-specific functions. We apply this model to a large multi-site $PM_{2.5}$ and hospitalization database from the United States covering the years 1999–2002.

2 A Bayesian Hierarchical Model for Constrained Distributed Lag Functions

Conceptually, our approach to modeling can be thought of in multiple stages. First, county-specific distributed lag functions are estimated via log-linear Poisson regressions. These county-specific functions are constrained based on prior information about the time course of hospitalization. The constrained county-specific distributed lag functions are then combined across counties to obtain a smooth overall estimate. The products of the model include posterior distributions for the constrained county-specific distributed lag functions and the overall combined distributed lag function.

2.1 Distributed lag functions

Given time series data $y_1, y_2, \ldots$ on an outcome such as daily hospitalization counts, and corresponding time series data $x_1, x_2, \ldots$ on an exposure such as ambient air pollution levels, a log-linear Poisson distributed lag model of order L specifies the relationship between X and Y to be of the form

$$\begin{aligned} Y_t &\sim \text{Poisson}(\mu_t) \\ \log \mu_t &= \sum_{\ell=0}^{L-1} \theta_\ell \, x_{t-\ell} \end{aligned} \tag{1}$$

for $t \geq L-1$. The vector of coefficients $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_{L-1})$, as a function of the lag number ($\ell = 0, \ldots, L-1$), is what we call the *distributed lag function*. This function is sometimes referred to as the impulse-response function because it describes the effect on the outcome series of a single impulse in the exposure series (Chatfield, 1996). For example, if we have an exposure series of the form $x_0 = 1, x_1 = 0, x_2 = 0, \ldots$, i.e. a spike at $t = 0$, then the log-relative risk (over L days) associated with that spike is

$$\xi = \sum_{\ell=0}^{L-1} \theta_\ell.$$

We can interpret $1000 \times (\exp(\xi) - 1)$ as the cumulative percent increase in hospitalizations for a 10 μg/m^3 increase in pollution (a standard increment for reporting particle air pollution relative risks).

2.2 County-specific model

Our approach begins with a model for the county-specific air pollution and hospitalization data. This model relates day-to-day changes in air pollution levels to day-to-day changes in

hospitalization counts for given county, controlling for other time-varying factors that might confound the relationship of interest.

Let the vector $\mathbf{y}_c = (y_{c1}, y_{c2}, \ldots, y_{cT})$ represent the daily time series of hospitalization counts for county c and let the vector $\mathbf{d}_c$ be the time series of the numbers of people at risk. The matrix $\mathbf{X}_c$ represents the exposure of interest and includes the corresponding time series of air pollution levels and lagged versions of that series. $\mathbf{X}_c$ is of dimension $T \times L$, where L is the order of the distributed lag model specified in (1). In our setup, the first column of $\mathbf{X}_c$ is the original air pollution time series (lag 0), the second column is the original series lagged by one day (lag 1), etc. For each county c we also observe p other time-varying covariates which are combined in a $T \times p$ matrix $\mathbf{Z}_c$. Then for county c, our county-specific log-linear Poisson model is of the form

$$\begin{array}{rcl} \mathbf{y}_c \mid \mathbf{X}_c, \mathbf{Z}_c & \sim & \text{Poisson}(\mu_c(\boldsymbol{\theta}_c, \boldsymbol{\beta}_c)) \\ \log \mu_c(\boldsymbol{\theta}_c, \boldsymbol{\beta}_c) & = & \mathbf{X}_c\boldsymbol{\theta}_c + \mathbf{Z}_c\boldsymbol{\beta}_c + \log \mathbf{d}_c \end{array} \tag{2}$$

where $c = 1, \ldots, n$ and each of the n counties are assumed to be mutually independent. The length L vector of parameters $\boldsymbol{\theta}_c$ is the distributed lag function and parameters in $\boldsymbol{\beta}_c$ are nuisance parameters.

In the county-specific model (2), we incorporate into $\mathbf{Z}_c$ certain time-varying factors that might confound the relationship between air pollution and hospitalization (Kelsall et al., 1997; Samet et al., 1998; Dominici et al., 2002a). In particular, we include average daily temperature, dew point temperature, and indicators for the day of the week. We also include a smooth function of time to adjust for seasonal variation that is common to both the air pollution and hospitalization time series. This smooth function of time is modeled using natural splines and the natural spline basis is included in $\mathbf{Z}_c$.

2.3 Constraining the distributed lag function and combining information

Our model takes advantage of two types of information to constrain the distributed lag function at each of the counties. The first type is prior information about the shape of the distributed lag function over time, which we incorporate by using an informative prior distribution on the county-specific distributed lag function $\boldsymbol{\theta}_c$. The second type of information is information from other counties, which we incorporate by placing an hierarchical structure on the county-specific distributed lag functions.

The prior distribution on $\boldsymbol{\theta}_c$ places few constraints on the values of the distributed lag function coefficients at early lags. At more distant lags, the variance of the coefficients is tapered towards zero and neighboring coefficients are constrained to be more correlated. The rationale behind this prior distribution is that the effects of air pollution at early lags are not well understood because of our lack of knowledge about biological mechanisms and the time course of the disease process. In addition, competing hypotheses mentioned previously about the shape of the distributed lag function suggest that fewer constraints should be placed at early lags. At longer lags it is reasonable to believe that the effects of air pollution on the outcome should approach zero smoothly, particularly since we are primarily interested in short-term effects.

Specifically, for $\boldsymbol{\theta}_c$, we let

$$\boldsymbol{\theta}_c \mid \boldsymbol{\mu}, \boldsymbol{\eta}, \sigma_\eta^2 \sim \mathcal{N}(\boldsymbol{\mu}, \sigma_\eta^2 \Omega(\boldsymbol{\eta})) \tag{3}$$

where $\boldsymbol{\mu} = (\mu_0, \mu_1, \ldots, \mu_{L-1})$ is a national average distributed lag function which describes, on average across all counties, the response spread over L days associated with a unit increase in pollution on a single day. The covariance matrix Ω is parametrized by the vector $\boldsymbol{\eta} = (\eta_1, \eta_2)$ where η_1 controls the rate at which the variance of the distributed lag function coefficients taper to zero and η_2 controls the rate at which neighboring coefficients become more correlated. The parameter σ_η^2 is the prior variance of $\theta_{c,0}$, the first distributed lag coefficient, which we take as fixed and known a priori.

In our model, we assume that the variance of θ_ℓ^c tapers to zero exponentially as a function of ℓ, so that

$$\text{Var}(\theta_{c,\ell}) = \sigma_\eta^2 \, \exp(-\eta_1 \, \ell)$$

for $\ell = 0, 1, 2, \ldots, L-1$. We further assume that the covariance of neighboring coefficients at lags ℓ_1 and ℓ_2 is proportional to

$$\text{Cov}(\theta_{c,\ell_1}, \theta_{c,\ell_2}) \propto [1 - \exp(-\eta_2 \, \ell_1)][1 - \exp(-\eta_2 \, \ell_2)],$$

so that neighboring coefficients at large lags have a correlation close to 1.

The matrix $\Omega(\boldsymbol{\eta})$ describes the heterogeneity across counties of the distributed lag functions $\boldsymbol{\theta}_c$. With this model we assume that there will be more heterogeneity in the coefficients at early lags and less heterogeneity in the coefficients at longer lags. Note that direct extension of the Welty, et al. model would assume that $\boldsymbol{\theta}_c$ is the same for all counties. Here, we extend the model to the multi-site context by allowing for some variation in the shapes of the county-specific distributed lag functions.

We assume that the national average distributed lag function $\boldsymbol{\mu}$ has a prior distribution

$$\boldsymbol{\mu} \mid \boldsymbol{\gamma}, \sigma_\gamma^2 \sim \mathcal{N}(\mathbf{0}, \sigma_\gamma^2 \, \Omega(\boldsymbol{\gamma})). \tag{4}$$

where Ω has the same form as in 3 but is parametrized by the vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2)$ and σ_γ^2, the prior variance of μ_0.

To complete the model specification the parameters $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ are each assumed to have uniform hyperprior distributions over a fixed range (details of the ranges are given in the Appendix) and σ_η^2 and σ_γ^2 are fixed at specific values. Following Welty et al. (2005), we set σ_η and σ_γ to be approximately 10 times the square root of the estimated variance of the maximum likelihood estimate of $\theta_{c,0}$. Sensitivity analysis indicates that the resulting posterior distributions are not affected as long as σ_η and σ_γ are not too small.

We implement a Gibbs sampler to obtain samples from the posterior distributions of the unknown parameters $\boldsymbol{\mu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\eta}$, and $\boldsymbol{\theta}_c$ for $c = 1, \ldots, n$. Briefly, the full conditionals for $\boldsymbol{\eta}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\theta}_c$ for $c = 1, \ldots, n$ are sampled using a Metropolis-Hastings rejection step and the full conditional for $\boldsymbol{\mu}$ is sampled in closed form. Rather than put a prior distribution on the many nuisance parameters in $\boldsymbol{\beta}_c$, we make use of a profile likelihood $L_p(\boldsymbol{\theta}_c) = L(\boldsymbol{\theta}_c, \widehat{\boldsymbol{\beta}}_c(\boldsymbol{\theta}_c))$, where for each given value of $\boldsymbol{\theta}_c$, we maximize the full likelihood L with respect to $\boldsymbol{\beta}_c$. In the Metropolis-Hastings step taken to sample from the full conditional for $\boldsymbol{\theta}_c$, we use the profile likelihood for $\boldsymbol{\theta}_c$ to calculate the acceptance ratio for the proposal. Full details of the sampling procedures can be found in Appendix B.

2.4 Other constrained distributed lag models

Many other types of distributed lag models have been applied in the literature and we do not attempt to describe them all. Most of these models take a two-stage approach, where

at the first stage, the distributed lag function is estimated independently for each location using log-linear Poisson regressions. As part of the first stage, each distributed lag function may be smoothed with a polynomial or other parametric smoother. At the second stage, the individual distributed lag functions are typically combined using a Normal hierarchical model to obtain an overall average estimate. This pooling is achieved by replacing the Poisson likelihood in (2) with a Normal approximation,

$$L_{\mathcal{N}}(\boldsymbol{\theta}_c) \propto \exp\left\{(\widehat{\boldsymbol{\theta}}_c - \boldsymbol{\theta}_c)'\widehat{\Sigma}_c^{-1}(\widehat{\boldsymbol{\theta}}_c - \boldsymbol{\theta}_c)\right\} \tag{5}$$

where $\widehat{\boldsymbol{\theta}}_c$ and $\widehat{\Sigma}_c$ are estimated in the first stage.

We implement for comparison an alternate distributed lag model where the distributed lag function is constrained to be a step function. This model restricts the effect at lags 0–2, lags 3–6, and lags 7–13 to be constant, respectively. Hence, for each county there are only 3 parameters of interest in this model and we pool them together using the Normal approximation in (5).

2.5 Connection to penalized spline smoothing

It is possible to make a connection between our Bayesian hierarchical model and smoothing via penalized splines. One can use the prior distributions described in the previous section and reformulate them as a penalized spline problem where one essentially does two levels of smoothing, once to constrain the county-specific distributed lag functions and once to combine information across counties. To make the computations more transparent, we will demonstrate this connection using the Normal approximation to the likelihood for $\boldsymbol{\theta}_c$ as in (5).

We can model the estimated distributed lag function $\widehat{\boldsymbol{\theta}}_c$ as a linear combination of basis functions (e.g. B-splines, natural splines), $\widehat{\boldsymbol{\theta}}_c = U\boldsymbol{\alpha}_c + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \widehat{\Sigma}_c)$ and $\widehat{\boldsymbol{\theta}}_c$ and $\widehat{\Sigma}_c$ are estimated using data for county c. U is a $L \times k$ basis matrix and $\boldsymbol{\alpha}_c$ is a k-vector of coefficients. The penalized spline solution solves the following optimization problem

$$\min_{\boldsymbol{\alpha}_c} (\widehat{\boldsymbol{\theta}}_c - U\boldsymbol{\alpha}_c)'\widehat{\Sigma}_c^{-1}(\widehat{\boldsymbol{\theta}}_c - U\boldsymbol{\alpha}_c) + \boldsymbol{\alpha}_c' D^{-1}\boldsymbol{\alpha}_c.$$

where D^{-1} is a penalty matrix which for now we assume incorporates a scalar penalty parameter. Since the penalty term $\boldsymbol{\alpha}_c' D^{-1}\boldsymbol{\alpha}_c$ is proportional to the minus log-density of the Normal distribution, we can rewrite the problem as

$$\begin{aligned} \widehat{\boldsymbol{\theta}}_c \mid U\boldsymbol{\alpha}_c &\sim \mathcal{N}(U\boldsymbol{\alpha}_c, \widehat{\Sigma}_c) \\ \boldsymbol{\alpha}_c &\sim \mathcal{N}(\mathbf{0}, D) \end{aligned} \tag{6}$$

where the solution is the posterior mode of $\boldsymbol{\alpha}_c$ under the Normal prior in (6).

Note that using our previous notation, we have that $\boldsymbol{\theta}_c = U\boldsymbol{\alpha}_c$. Given (3) and (4), we can write the marginal distribution of $\boldsymbol{\theta}_c$ as

$$\boldsymbol{\theta}_c \sim \mathcal{N}(\mathbf{0},\, \Omega(\boldsymbol{\eta}) + \Omega(\boldsymbol{\gamma})) \tag{7}$$

where we have absorbed σ_η^2 and σ_γ^2 into $\Omega(\boldsymbol{\eta})$ and $\Omega(\boldsymbol{\gamma})$, respectively, to reduce the clutter. The prior in (6) implies that $U\boldsymbol{\alpha}_c \sim \mathcal{N}(\mathbf{0}, UDU')$. Using this fact and the marginal distribution in (7), the penalty matrix D must satisfy $UDU' = \Omega(\boldsymbol{\eta}) + \Omega(\boldsymbol{\gamma})$, which has the solution

$$D_{\eta,\gamma} = (U'U)^{-1}U'(\Omega(\boldsymbol{\eta}) + \Omega(\boldsymbol{\gamma}))U(U'U)^{-1}. \tag{8}$$

Now we have shown that our prior distribution on $\boldsymbol{\theta}_c$ can be translated into a penalty matrix for spline coefficients in a penalized spline problem. Given values of $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ and using the penalty matrix in (8), we can calculate the penalized spline coefficient estimates as

$$\widehat{\boldsymbol{\alpha}}_c = D_{\eta,\gamma} U' (U D_{\eta,\gamma} U' + \widehat{\Sigma}_c)^{-1} \widehat{\boldsymbol{\theta}}_c$$

and the smoothed county-specific distributed lag function is $U\widehat{\boldsymbol{\alpha}}_c$.

Similar calculations can be made to show the second level of smoothing that averages information across counties. Analogous to (6), we can write the the second level of our hierarchical model as

$$\begin{aligned} \boldsymbol{\theta}_c \mid W\boldsymbol{\delta} &\sim \mathcal{N}(W\boldsymbol{\delta}, \Omega(\boldsymbol{\eta})) \\ \boldsymbol{\delta} &\sim \mathcal{N}(\mathbf{0}, H) \end{aligned} \tag{9}$$

where W is a basis matrix, $\boldsymbol{\delta}$ is a vector of coefficients, and H is a penalty matrix in the penalized spline problem. The distribution in (4) and the prior for $\boldsymbol{\delta}$ in (9) imply that we need to find a matrix H such that $WHW' = \Omega(\boldsymbol{\gamma})$. The solution for H has the form $H_\gamma = (W'W)^{-1} W' \Omega(\boldsymbol{\gamma}) W (W'W)^{-1}$ and we can subsequently solve for $\widehat{\boldsymbol{\delta}} = H_\gamma W' (W H_\gamma W' + \Omega(\boldsymbol{\eta}))^{-1} \bar{\boldsymbol{\theta}}$, where $\bar{\boldsymbol{\theta}} = (1/n) \sum_{c=1}^{n} \boldsymbol{\theta}_c$.

One can see that if we replace the basis matrices U and W with the $L \times L$ identity matrix, then we revert back to our original formulation and obtain the same answers as our original Bayesian hierarchical model. Our model places a prior directly on the distributed lag function, while the penalized spline approach is one step removed via the linear maps U and W, with a prior placed on the spline coefficients. In this application it seems more natural to place a prior distribution on the distributed lag function directly. However, it can be shown that either approach can result in similar estimates of the distributed lag function (Welty et al., 2005).

3 Data

The hospitalization data consist of daily counts of hospital admissions for the years 1999–2002 constructed from the National Claims History Files (NCHF) of Medicare, which contain the billing claims of all Medicare enrollees. Medicare enrollees make up almost the entire U.S. population over 65 years of age, or approximately 48 million people. Each billing claim obtained from the NCHF contains the date of service, treatment, disease classification (via *International Classification of Diseases, Ninth Revision [ICD-9]* codes), age, gender, self-reported race, and place of residence (five digit ZIP code and county). The daily counts for a given county were computed by summing the total number of admissions with a primary diagnosis for a specific disease. For computing hospitalization rates, a time series of the numbers of individuals at risk in each county for each day was constructed.

The $PM_{2.5}$ data were obtained from the EPA's Air Quality System database (formerly the AIRS database) which makes available data from a network of monitors stationed around the country. Temperature and dew point temperature data were assembled from the National Climatic Data Center on the Earth-Info CD database and linked by county with the pollution and hospitalization data.

The hospitalization, air pollution, and weather data were originally used for the MCAPS study and further details about the construction of the dataset can be found in Dominici et al. (2006). This analysis was restricted to 94 counties for which daily data on $PM_{2.5}$ were available. The included counties were further constrained to have a population over 200,000 and at least one full year of $PM_{2.5}$ data. The resulting study population for the 94 counties consisted of 6.3 million Medicare enrollees living on average 6 miles from a $PM_{2.5}$ monitor. The locations of the 94 counties are shown in Figure 1.

Figure 2 shows as an example some hospitalization and $PM_{2.5}$ time series data for Cook County, Illinois, which contains many features that are common to the other counties used in the analysis. The hospitalization time series shows rates of admissions (per 100,000) for a primary diagnosis of chronic obstructive pulmonary disease with acute exacerbation. The winter season tends to to see many more hospitalizations and results in a seasonal peak around the first of the year. $PM_{2.5}$ has a similar pattern with a winter peak and a summer trough, a pattern that is common in western and midwestern counties. Eastern counties tend to have the opposite pattern with summer peaks in $PM_{2.5}$ levels and winter troughs. Another common feature shown in Figure 2(b) is that $PM_{2.5}$ is measured somewhat sparsely in the year 1999 and is monitored more regularly starting in 2000.

For all of the counties used in this analysis, there were occasional missing $PM_{2.5}$ values. With the exception of the year 1999, the missingness tended to be sporadic and seemingly at random. Rather than treat the missing $PM_{2.5}$ values specially or implement an imputation scheme, we chose to analyze only the days for which observations were available. There were few, if any, missing values in the hospitalization and meteorological data.

4 Application to Air Pollution and Hospitalization Data

We applied the Bayesian hierarchical distributed lag model (BDLM) to the $n = 94$ counties with Medicare, air pollution, and weather data described in Section 3. The data for each county spanned $T = 1,461$ days (the 4 years from 1999 to 2002). We chose to examine two specific causes of hospitalization: chronic obstructive pulmonary disease with acute exacerbation (COPDAE) and ischemic heart disease. These outcomes were chosen because they represent both respiratory and cardiovascular diseases and have been shown in previous studies to be strongly associated with $PM_{2.5}$ exposure. For the distributed lag function, we chose to fit a model with a maximum lag of two weeks, so that $L = 14$ in (1).

Figure 3 shows the national average distributed lag functions for COPDAE and ischemic heart disease. Each of these plots shows the posterior mean for $\boldsymbol{\mu}$ for each outcome and pointwise 95% posterior intervals for each lag coefficient. At each lag the plotted coefficient can be interpreted as the percent increase in hospitalization for a 10 $\mu g/m^3$ increase in $PM_{2.5}$.

For COPDAE, a 10 $\mu g/m^3$ increase in $PM_{2.5}$ appears to be associated with two "waves" of admissions, with the first arriving 1 day after the increase and the second arriving a few days later. For ischemic heart disease, there is about a 0.2% increase in admissions on the same day, followed by an approximately 0.4% decrease in admissions on the following day. Then at lag 2, the relative risk jumps to a 0.60% increase in admissions, beyond which the distributed lag function for ischemic heart disease is essentially zero. The impact of $PM_{2.5}$ on ischemic heart disease appears to be more immediate, being felt in the first 0–2 days, while the impact on COPDAE appears to be spread out over a week or more. In addition, the

shape of the distributed lag function for ischemic heart disease suggests a possible mortality displacement scenario described previously.

Figure 3 also illustrates the effect of the prior on $\boldsymbol{\mu}$, which is to smooth and taper the lag coefficients as the lag increases. The distributed lag function appears to taper and smooth out much more rapidly for ischemic heart disease than for COPDAE, a feature that is corroborated by the joint posterior distributions for γ_1 and γ_2 shown in Figure 4. The data for both outcomes prefer a large value for γ_1, but for ischemic heart disease the marginal distribution for γ_2 is shifted somewhat higher than that of COPDAE, indicating a preference for more smoothness.

The county-specific Bayesian distributed lag functions for COPDAE and ischemic heart disease are are shown in Figures 5 and 6, respectively. Each figure shows the posterior mean and pointwise 95% posterior intervals of $\boldsymbol{\theta}_c$ for the largest 25 counties in the study. For COPDAE, the estimated county-specific distributed lag functions are a mix of shapes including large immediate effects (Sacramento, Broward), somewhat smaller delayed effects (Los Angeles CA, Franklin OH, Pinellas FL), and more moderate effects spread out over a longer period of time (Bronx NY, Palm Beach FL, Salt Lake UT). Ischemic heart disease appears to exhibit somewhat less heterogeneity in the shapes of the distributed lag functions with most of the effects occurring at lags 0–2.

Figure 7 shows the estimated cumulative effect of $PM_{2.5}$ on both outcomes from a few different models. For each outcome we plot the estimate originally reported in the MCAPS study for a single lag model applied to 204 U.S. counties (Dominici et al., 2006), the estimate obtained from a single lag model applied to the 94 counties used in this study (for the exposure lag we chose lag 0 for COPDAE and lag 2 for ischemic heart disease), the estimate obtained from using a 14 day distributed lag model with a step function (as described in Section 2.4), and the estimate obtained from our Bayesian hierarchical distributed lag model (BDLM). It is useful to think of the single lag model here as a heavily constrained distributed lag model where the coefficient for lag 0 is free and the coefficients for lags 1–13 are forced to be zero.

The estimates of the cumulative effects for COPDAE and ischemic heart disease are remarkably similar across models. The MCAPS point estimate was reported as 0.91 and posterior mean from the BDLM model was 0.84. One can see from the difference in posterior intervals from the "MCAPS" and the "Single lag" estimates that the loss of 110 counties in this study only results in a small loss of efficiency in the estimate of the cumulative effect. However, the increased number of parameters introduced by the distributed lag model (even the 3-parameter "Step DLM" model) results in a substantial increase in the variance of the estimate. For the ischemic heart disease outcome, the estimate from the BDLM is 0.66 compared to the MCAPS estimate of 0.44. This higher effect was also captured by the step-function distributed lag model but the BDLM appears to exhibit less variance in its estimate.

5 Discussion

We have proposed a Bayesian hierarchical distributed lag model (BDLM) for combining constrained distributed lag functions and have applied the model to the problem of estimating the distributed lag between day-to-day changes in ambient air pollution levels and day-to-day changes in hospitalization rates. The model uses an informative prior regarding the

time course of the short-term health effects of air pollution and combines information from multiple locations. We have applied the model to a large air pollution and hospitalization database for United States residents enrolled in Medicare, examining the distributed lag between $PM_{2.5}$ exposure and hospitalization for ischemic heart disease and COPD with acute exacerbation.

The model that we have proposed allows us to see how the risk of hospital admission due to air pollution changes over short periods of time. The prior distribution on $\boldsymbol{\mu}$, the national average distributed lag function, gives more freedom to coefficients at early lags, reflecting our lack of knowledge in that time period, and becomes increasingly constrained as the lag increases. In this case we have chosen a period of two weeks for the exposure period, after which we assume that the effect of air pollution is zero. In addition, the model allows for estimating county-specific distributed lag functions and examining the heterogeneity in the shapes of the distributed lag functions across counties.

The national average distributed lag functions for COPDAE and ischemic heart disease present two different shapes, indicating different time courses for the relative risks for these disease categories. The effect of $PM_{2.5}$ on COPDAE admissions appears to be spread over a longer time period than the effect on ischemic heart disease admissions, which is more immediate. The shape of the distributed lag function for COPDAE suggests that there are two "waves" of admissions associated with an increase in $PM_{2.5}$ at a 1 day lag and at a 3 to 5 day lag, respectively.

The nature and characteristics of acute exacerbations of COPD are known to be heterogeneous across people (Sapey and Stockley, 2006) and exacerbations are often a cause of hospital admission after initial treatment outside the hospital has failed (Seemungal et al., 2000). Given those factors, it is perhaps not surprising that on average, the effect of $PM_{2.5}$ on COPDAE is spread out over a longer period of time. In addition, the severity of exacerbations specifically due to bacterial or viral infection would likely depend on the nature of the bacteria or viruses involved, thereby introducing more heterogeneity into the time at which a person is admitted to the hospital for exacerbation (Wilkinson et al., 2006).

We found little evidence that the effect of increases in $PM_{2.5}$ levels on admission for ischemic heart disease extends much beyond 2 days after a given episode. In addition, the shape of the distributed lag function for ischemic heart disease suggests some weak evidence of mortality displacement. From Figure 3(b) it appears that there may be an initial susceptible wave of admissions at lag 0, followed by fewer than expected admissions at lag 1. Then at lag 2 a larger second wave of admissions occurs. Cardiovascular effects of PM are thought to be generally related to neurogenic and inflammatory processes (Pope et al., 2003). The results from our analysis suggest that for ischemic heart disease in particular, the biological mechanism involved has a relatively short time course, with the bulk of people admitted to the hospital within two days of an increase in $PM_{2.5}$ levels.

Given that the effect of COPDAE appears to be spread out over a week or more, one might expect that the cumulative effect would be greater than the estimate obtained from a single lag model that only captures the effect spread over one day. Previous studies have shown increased relative risk estimates when a distributed lag model has been applied. For example, Bell et al. (2004), applying a distributed lag model to 94 US cities, found the estimate of the cumulative effect of ozone on daily mortality to be approximately twice as large as estimates obtained via single lag models. The APHEA multi-city European study found that when using

the average of PM_{10} over multiple days as the exposure, the effect was larger (Katsouyanni et al., 1997, 2001). In addition, various single city studies have shown larger effects from using distributed lag models (Health Effects Institute, 2003).

In this analysis the cumulative effect of $PM_{2.5}$ on COPDAE admissions was similar to the effect estimated via the single lag model, indicating that in this case the single lag model captures the extent of the cumulative percent increase in hospital admissions. With ischemic heart disease, the single lag model underestimates the cumulative effect, which the BDLM estimates as approximately a 0.6% increase in admissions for a 10 $\mu g/m^3$ increase in $PM_{2.5}$. However, given the variability in the BDLM estimate, there does not appear to be a significant difference between the single lag estimate and the BDLM estimate.

It is interesting to note the bias-variance trade off involved in choosing between using a single lag or distributed lag model. Even with a relatively large dataset, estimation of the distributed lag function resulted in an substantial increase in the variance of the cumulative effect compared to single lag models. While one might consider the single lag model's restriction to 1-day effects a limitation (and perhaps a source of bias), one must also consider the dramatic increase in precision that the model provides. If one is solely interested in the cumulative short-term effect of an increase in air pollution levels, the benefits of the distributed lag model's greater flexibility may not outweigh the cost of incurring much greater variability in the resulting estimate.

The principal benefit of the distributed lag model is its ability to estimate the shape of the population-level response to increased air pollution levels. Our model provides a useful parametrization that can easily incorporate prior knowledge and be applied to large multi-site databases. However, we should be careful not to overinterpret the findings of our analysis. Even with the constraints imposed by the prior, the uncertainty of the estimates in Figure 3 is still large, particularly for estimates at early lags. In addition, Medicare data are collected for administrative purposes and disease diagnoses are known to be subject to some missclassification. However, such missclassification would only bias our results if the daily pattern of diagnosis and coding varied in a a way that was correlated with $PM_{2.5}$ levels. Finally, while our results are interesting and suggest some possible hypotheses, more focused studies (perhaps involving susceptible sub-populations) will have to be conducted to obtain more precise information about the biological mechanisms involved.

References

Almon, S. (1965), "The distributed lag between capital appropriations and expenditures," *Econometrica*, 33, 178–196.

Bell, M. L., McDermott, A., Zeger, S. L., Samet, J. M., and Dominici, F. (2004), "Ozone and Short-term Mortality in 95 US Urban Communities, 1987-2000," *Journal of the American Medical Association*, 292, 2372–2378.

Chatfield, C. (1996), *The Analysis of Times Series: An Introduction*, Chapman & Hall/CRC, 5th ed.

Corradi, C. (1977), "Smooth distributed lag estimators and smoothing spline functions in Hilbert spaces," *Journal of Econometrics*, 5, 211–220.

Dominici, F., Daniels, M., Zeger, S. L., and Samet, J. M. (2002a), "Air Pollution and Mortality: Estimating Regional and National Dose-Response Relationships," *Journal of the American Statistical Association*, 97, 100–111.

Dominici, F., McDermott, A., Zeger, S. L., and Samet, J. M. (2002b), "Airborne particulate matter and mortality: Time-scale effects in four US Cities," *American Journal of Epidemiology*, 157, 1053–1063.

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.

Goodman, P. G., Dockery, D. W., and Clancy, L. (2004), "Cause-specific mortality and the extended effects of particulate pollution and temperature exposure," *Environmental Health Perspectives*, 112, 179–185.

Health Effects Institute (2003), *Revised Analyses of Time-Series Studies of Air Pollution and Health. Special Report.*, Health Effects Institute, Boston MA.

Katsouyanni, K., Toulomi, G., Samoli, E., Gryparis, A., LeTertre, A., Monopolis, Y., Rossi, G., Zmirou, D., Ballester, F., Boumghar, A., and Anderson, H. R. (2001), "Confounding and Effect Modification in the Short-term Effects of Ambient Particles on Total Mortality: Results from 29 European Cities within the APHEA2 Project," *Epidemiology*, 12, 521–531.

Katsouyanni, K., Toulomi, G., Spix, C., Balducci, F., Medina, S., Rossi, G., Wojtyniak, B., Sunyer, J., Bacharova, L., Schouten, J., Ponka, A., and Anderson, H. R. (1997), "Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project," *British Medical Journal*, 314, 1658–1663.

Kelsall, J. E., Samet, J. M., Zeger, S. L., and Xu, J. (1997), "Air Pollution and Mortality in Philadelphia, 1974–1988," *American Journal of Epidemiology*, 146, 750–762.

Leamer, E. E. (1972), "A class of informative priors and distributed lag analysis," *Econometrica*, 40, 1059–1081.

Peng, R. D., Dominici, F., Pastor-Barriuso, R., Zeger, S. L., and Samet, J. M. (2005), "Seasonal Analyses of Air Pollution and Mortality in 100 US Cities," *American Journal of Epidemiology*, 161, 585–594.

Pope, C. A., Burnett, R. T., Thruston, G. D., Calle, E., Thun, M. J., Krewski, D., and Goldeski, J. (2003), "Cardiovascular Mortality and Long-term Exposure to Particulate Air Pollution: Epidemiological Evidence of General Pathophysiological Pathways of Disease," *Circulation*, 6, 71–77.

Pope, C. A. and Dockery, D. W. (2006), "Health effects of fine particulate air pollution: lines that connect," *Journal of the Air and Waste Management Association*, 56, 709–742.

Roberts, S. (2005), "An investigation of distributed lag models in the context of air pollution and mortality time series analysis," *Journal of the Air and Waste Management Association*, 55, 273–282.

Samet, J., Zeger, S., Kelsall, J., Xu, J., and Kalkstein, L. (1998), "Does Weather Confound or Modify the Association of Particulate Air Pollution with Mortality?" *Environmental Research, Section A*, 77, 9–19.

Samoli, E., Schwartz, J., Wojtyniak, B., Touloumi, G., Spix, C., Balducci, F., Medina, S., Rossi, G., Sunyer, J., Bacharova, L., Anderson, H. R., and Katsouyanni, K. (2001), "Investigating Regional Differences in Short-Term Effects of Air Pollution on Daily Mortality in the APHEA Project: A Sensitivity Analysis for Controlling Long-Term Trends and Seasonality," *Environmental Health Perspectives*, 109, 349–353.

Samoli, E., Touloumi, G., Zanobetti, A., Le Tertre, A., Schindler, C., Atkinson, R., Vonk, J., Rossi, G., Saez, M., Rabczenko, D., Schwartz, J., and Katsouyanni, K. (2003), "Investigating the dose-response relation between air pollution and total mortality in the APHEA-2 multicity project," *Occupational and Environmental Medicine*, 60, 977–982.

Sapey, E. and Stockley, R. A. (2006), "COPD exacerbations 2: Aetiology," *Thorax*, 61, 250–258.

Schimmel, H. and Murawski, T. J. (1976), "The relation of air pollution to mortality," *Journal of Occupational Medicine*, 18, 316–333.

Schwartz, J. (2000), "The Distributed Lag between Air Pollution and Daily Deaths," *Epidemiology*, 11, 320–326.

— (2006), "Comments on 'Model choice in time series studies of air pollution and mortality'," *Journal of the Royal Statistical Society, Series A*, 169, 198–200.

Seemungal, T. A. R., Donaldson, G. C., Bhowmik, A., Jeffries, D. J., and Wedzicha, J. A. (2000), "Time Course and Recovery of Exacerbations in Patients with Chronic Obstructive Pulmonary Disease," *American Journal of Respiratory and Critical Care Medicine*, 161, 1608–1613.

Shiller, R. J. (1973), "A distributed lag estimator derived from smoothness priors," *Econometrica*, 41, 775–788.

Welty, L. J., Zeger, S. L., and Dominici, F. (2005), "Bayesian Distributed Lag Models: Estimating the Effects of Particulate Matter Air Pollution on Daily Mortality," Tech. Rep. 96, Johns Hopkins University Department of Biostatistics, http://www.bepress.com/jhubiostat/paper96.

Wilkinson, T. M. A., Hurst, J. R., Perera, W. R., Wilks, M., Donaldson, G. C., and Wedzicha, J. A. (2006), "Effect of Interactions Between Lower Airway Bacterial and Rhinoviral Infection in Exacerbations of COPD," *Chest*, 129, 317–324.

Zanobetti, A., Schwartz, J., Samoli, E., Gryparis, A., Touloumi, G., Atkinson, R., Le Tertre, A., Bobros, J., Celko, M., Goren, A., Forsberg, B., Michelozzi, P., Rabczenko, D., Aranguez, R. E., and Katsouyanni, K. (2002), "The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement," *Epidemiology*, 13, 87–93.

Zanobetti, A., Wand, M., Schwartz, J., and Ryan, L. (2000), "Generalized additive distributed lag models: quantifying mortality displacement," *Biostatistics*, 1, 279–292.

Zeger, S. L., Dominici, F., and Samet, J. M. (1999), "Harvesting-resistant estimates of pollution effects on mortality," *Epidemiology*, 89, 171–175.

A Figures

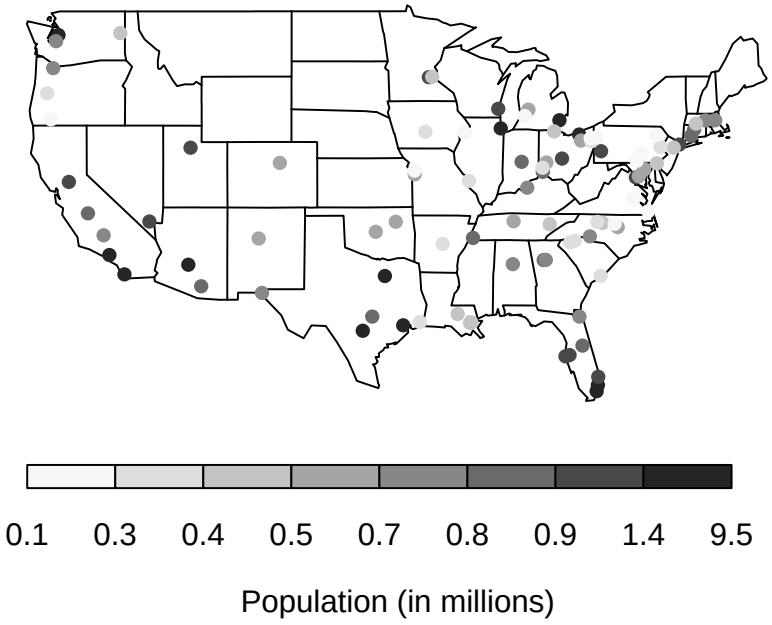

Figure 1: Locations of 94 U.S. counties which have daily data for particulate matter < 2.5 μm in diameter for 1999–2002.

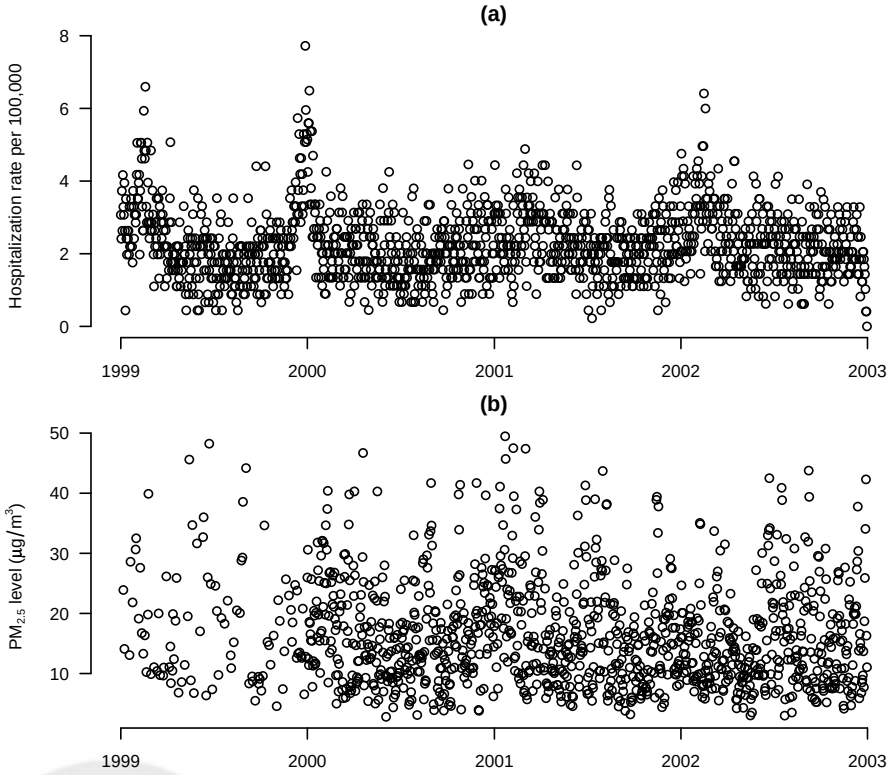

Figure 2: (a) Daily hospitalization rates for chronic obstructive pulmonary disease with acute exacerbation for people aged ≥ 65 living in Cook County, Illinois, 1999–2002; (b) daily levels of particulate matter < 2.5 μm in diameter ($PM_{2.5}$) for Cook County, Illinois, 1999–2002.

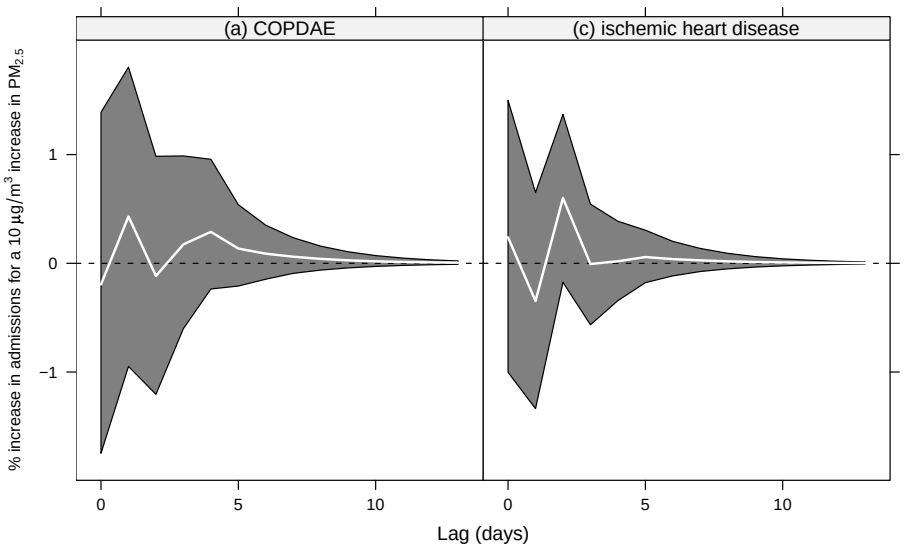

Figure 3: National average distributed lag functions for (a) COPD with acute exacerbation and (b) ischemic heart disease. Results are from a Bayesian hierarchical distributed lag model applied to 94 U.S. counties, 1999-2002. Each plot shows the posterior mean (white line) and pointwise 95% posterior intervals (shaded gray region) for each lag coefficient.

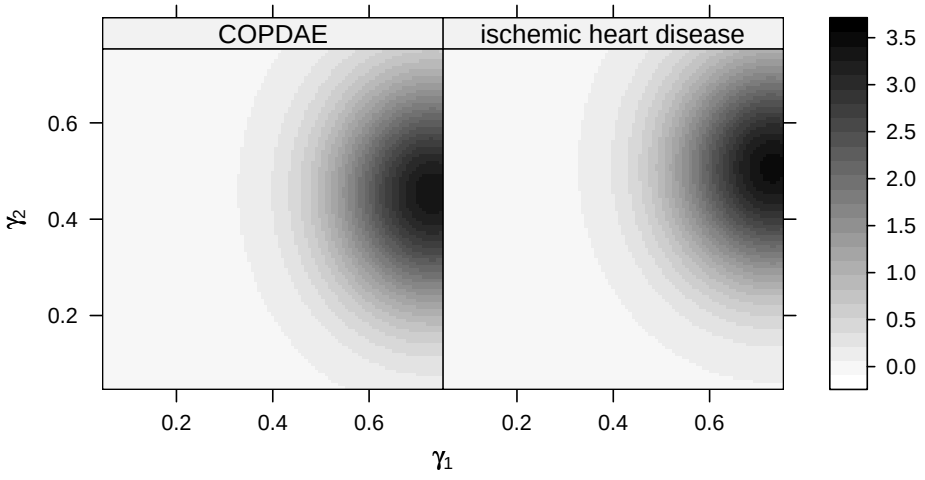

Figure 4: Joint posterior distributions for γ_1 and γ_2 for both COPDAE and ischemic heart disease.

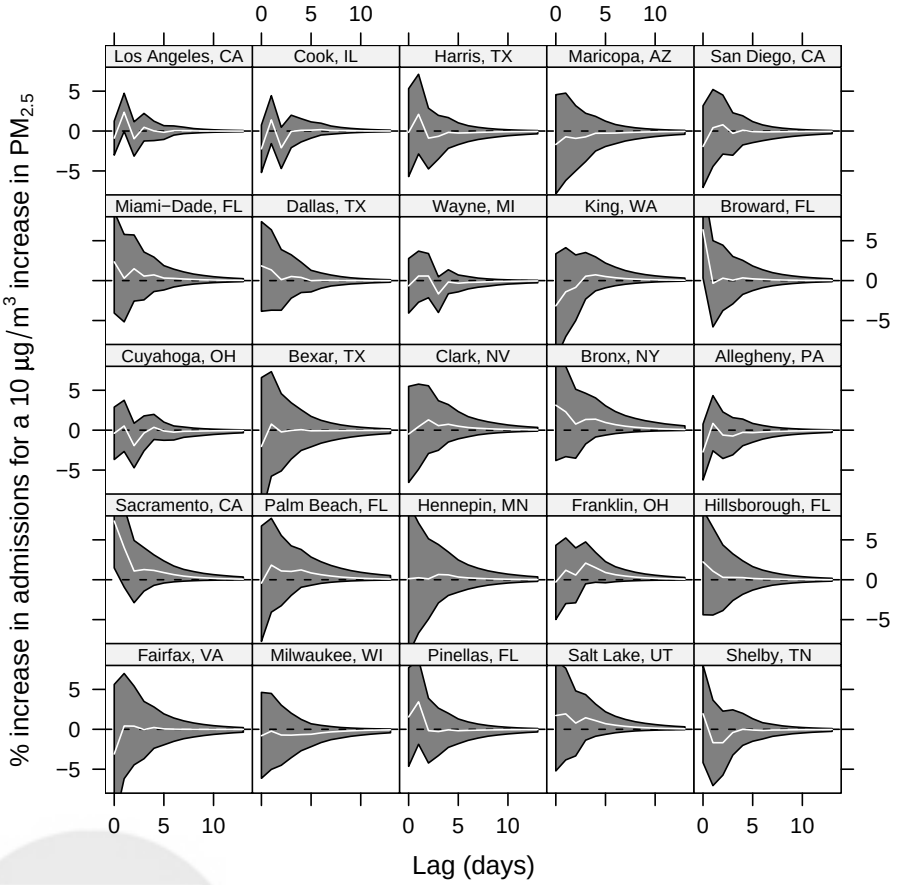

Figure 5: County-specific Bayesian distributed lag functions (with 95% posterior regions) showing the effect of $PM_{2.5}$ on hospitalization for COPD with acute exacerbation. Only the largest 25 counties (by population) are shown here, with the largest county (Los Angeles, CA) in the top left corner.

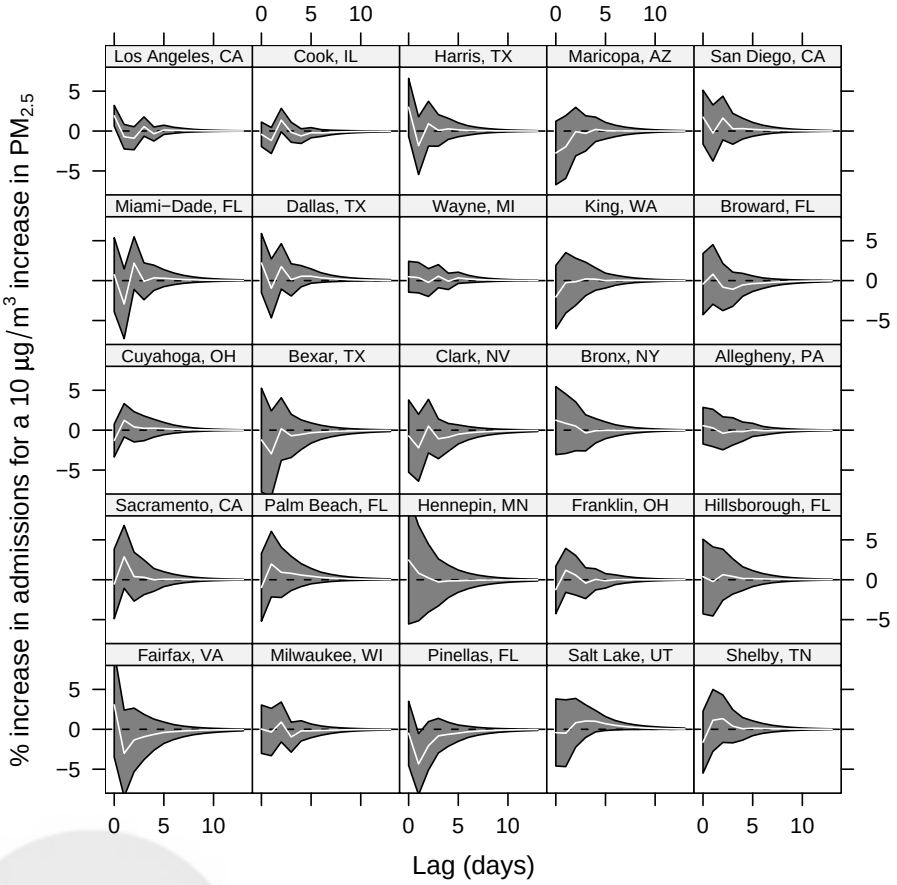

Figure 6: County-specific Bayesian distributed lag functions (with 95% posterior regions) showing the effect of $PM_{2.5}$ on hospitalization for ischemic heart disease. Only the largest 25 counties (by population) are shown here, with the largest county (Los Angeles, CA) in the top left corner.

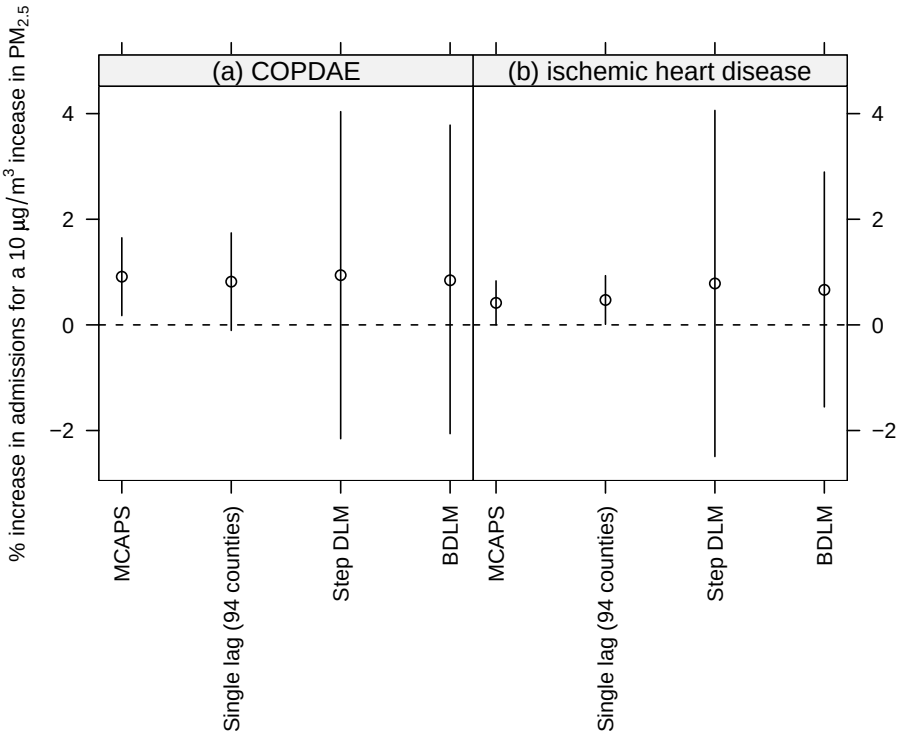

Figure 7: Estimates and 95% posterior intervals for the cumulative effect of $PM_{2.5}$ (a) COPD with acute exacerbation and (b) heart failure. Estimates for "MCAPS" and "Single lag (94 counties)" come from single lag models applied to the original MCAPS study and to the 94 counties used in this study, respectively; the "Step DLM" estimates come from a 14-day distributed lag model using a step function (county-specific estimates are pooled using the Normal approximation); the "BDLM" estimates come from applying the Bayesian hierarchical distributed lag model using a 14-day distributed lag.

B Details of Gibbs Sampler

In order to draw samples for the posterior distribution of the parameters in our Bayesian hierarchical distributed lag model, we implement an hybrid Gibbs sampler.

1. *Sampling* $\boldsymbol{\theta}_c$. In order to sample from the full conditional for $\boldsymbol{\theta}_c$ we implement a Metropolis-Hastings rejection scheme. The proposal distribution for sampling from the full conditional of $\boldsymbol{\theta}_c$ is constructed by first estimating $\boldsymbol{\theta}_c$ in a county-specific log-linear Poisson regression model. Once the estimate $\widehat{\boldsymbol{\theta}}_c$ and its estimated covariance matrix $\widehat{\Sigma}_c$ is obtained, we can compute the posterior distribution of $\boldsymbol{\theta}_c$ given $\widehat{\boldsymbol{\theta}}_c$ and the current value of the national average $\boldsymbol{\mu}$, i.e.

$$\boldsymbol{\theta}_c \mid \widehat{\boldsymbol{\theta}}_c, \boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\mu} + B_1(\widehat{\boldsymbol{\theta}}_c - \boldsymbol{\mu}), \sigma_\gamma^2(I - B_1)\Omega(\boldsymbol{\gamma})) \tag{10}$$

   where $B_1 = \sigma_\gamma^2\Omega(\boldsymbol{\gamma})\,[\widehat{\Sigma}_c + \sigma_\gamma^2\Omega(\boldsymbol{\gamma})]^{-1}$. Given the proposal distribution in (10), the full conditional for $\boldsymbol{\theta}_c$ from which we sample is then proportional to

$$p(\boldsymbol{\theta}_c \mid \cdot) \propto L_p(\boldsymbol{\theta}_c)\varphi(\boldsymbol{\theta}_c \mid \boldsymbol{\mu}, \sigma_\eta^2\Omega(\boldsymbol{\eta}))$$

   where $\varphi(\boldsymbol{\theta}_c \mid \boldsymbol{\mu}, \sigma_\eta^2\Omega(\boldsymbol{\eta}))$ is the multivariate normal density with mean $\boldsymbol{\mu}$ and covariance matrix $\sigma_\eta^2\Omega(\boldsymbol{\eta})$ and $L_p(\boldsymbol{\theta}_c)$ is the profile likelihood for $\boldsymbol{\theta}_c$.

2. *Sampling* $\boldsymbol{\mu}$. The full conditional for $\boldsymbol{\mu}$ is proportional to

$$\begin{aligned} p(\boldsymbol{\mu} \mid \cdot) &\propto \left\{\prod_{c=1}^{n} \varphi(\boldsymbol{\theta}_c \mid \boldsymbol{\mu},\, \sigma_\eta^2\Omega(\boldsymbol{\eta}))\right\} \varphi(\boldsymbol{\mu} \mid \mathbf{0},\, \sigma_\gamma^2\Omega(\boldsymbol{\gamma})) \\ &= \mathcal{N}(B_2\,\bar{\boldsymbol{\theta}},\, (I - B_2)\,\sigma_\gamma^2\Omega(\boldsymbol{\gamma})) \end{aligned}$$

   where $B_2 = \sigma_\gamma^2\Omega(\boldsymbol{\gamma})\,[\sigma_\gamma^2\Omega(\boldsymbol{\gamma}) + \sigma_\eta^2\Omega(\boldsymbol{\eta})/n]^{-1}$ and $\bar{\boldsymbol{\theta}} = \frac{1}{n}\sum \boldsymbol{\theta}_c$.

3. *Sampling* $\boldsymbol{\eta}$ *and* $\boldsymbol{\gamma}$. The full conditionals for $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ are proportional to $\prod \varphi(\boldsymbol{\theta}_c \mid \boldsymbol{\mu}, \sigma_\eta^2\Omega(\boldsymbol{\eta}))$ and $\varphi(\boldsymbol{\mu} \mid \mathbf{0}, \sigma_\gamma^2\Omega(\boldsymbol{\gamma}))$, respectively. We put uniform priors on both (η_1, η_2) and (γ_1, γ_2). In order to preserve numerical stability, we placed upper and lower bounds on each parameter so that both η_1 and η_2 were restricted to be in the range $[0.2, 0.8]$ while γ_1 and γ_2 were restricted to be in the range $[0.05, 0.75]$. These bounds were chosen based on previous work and some experimentation. Upper bounds that were much larger than these values often produced covariance matrices that were not invertible. We subsequently used uniform proposal distributions and a Metropolis-Hastings rejection step to sample from the full conditionals of $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$.