UNDERESTIMATION OF STANDARD ERRORS IN MULTI SITE TIME SERIES STUDIES

Abstract

Multi-site time series studies of air pollution and mortality and morbidity have figured prominently in the literature as comprehensive approaches for estimating short-term effects of air pollution on health. Hierarchical models are generally used to combine site-specific information and estimate pooled air pollution effects taking into account both within-site statistical uncertainty, and across-site heterogeneity.

Within a site, characteristics of time series data of air pollution and health (small pollution effects, missing data, and highly correlated predictors) make modeling all sources of uncertainty challenging. One potential consequence is underestimation of the statistical variance of the site-specific effects to be combined.

In this paper, we investigated the impact of variance underestimation on the pooled relative rate estimate. We focused on two-stage normal-normal hierarchical models and on underestimation of the statistical variance at the first stage. By mathematical considerations and simulation studies, we found that variance underestimation did not affect the pooled estimate substantially. However, some sensitivity of the pooled estimate to variance underestimation was observed when the number of sites is small and underestimation was severe. These simulation results are applicable to any two-stage normal-normal hierarchical model for combining information of site-specific results (including meta-analyses), and they can be

1

easily extended to more general hierarchical formulations.

We also examined the impact of variance underestimation on the national average relative rate estimate from the National Morbidity Mortality Air Pollution Study. We found that variance underestimation as large as 40% had little effect on the national average.

Key words: multi-site time series studies of air pollution and health, meta-analysis, hierarchical models, variance underestimation.

1 Introduction

In multi-site time series studies of air pollution and mortality and morbidity (1; 2), site-specific time series data are assembled under a common framework and analyzed with a uniform analytic approach. Hierarchical modeling is an unified approach for combining evidence across studies, quantifying the sources of variability, and assessing effect modification. Because of the development of computational tools that facilitate their implementation (3; 4), hierarchical models have been recently applied to analysis of multi-site time series data (5; 1; 6; 7; 8; 9). See Dominici (2002) (10) for a more detailed discussion on the use of hierarchical models in multi-site time series studies of air pollution and health.

Hierarchical models (11) for analyses of multi-site time series studies of air pollution and mortality have a multi-stage structure. At the first stage, the association between air pollution and health is described using a site-specific regression model (12; 13; 7) which takes into account potential confounding factors such as trend, season, and climate. Generalized Additive Models (GAM) (14) with non-parametric adjustment for confounding factors for the site (e.g. smoothing splines) or Generalized Linear Models (GLM)(15) with regression splines (e.g. natural cubic splines), are generally used for estimating site-specific relative rates $\hat{\beta}_c$ and their sampling variances v_c. Here $\hat{\beta}_c$ denotes the percentage increase in mortality/morbidity per unit increase in level of the air pollutant, and v_c denotes the statistical uncertainty in $\hat{\beta}_c$ which depends on the number of: 1) days with available air pollution data; 2) adverse health events; and 3) correlated time varying confounders. At the second stage, the information

from multiple sites is combined by assuming that the true city-specific relative rates (β_c) have a common mean α (also called the pooled relative rate) and variance τ^2, which represents the variability across sites of the true relative rates (also called the heterogeneity parameter). Fixed or random effects, empirical Bayesian, or fully Bayesian models (16; 17; 18; 19; 20; 18) are used to estimate α and τ^2.

The nature and characteristics of the time series studies in air pollution and health make estimation of health risk, taking into account of all sources of uncertainty, complex. First, the variability in the mortality and/or morbidity time series explained by air pollution is an order of magnitude lower than the variability in the mortality time series explained by weather, trend, and seasonality. Consequently, the estimates of air pollution effects are sensitive to the method of adjustment for confounding factors. Second, to adequately control for confounding, several highly correlated predictors are included in the site-specific regression model. This can make variance estimation unstable and slow the convergence of fitting algorithms such as the backfitting algorithm in GAM (21). Third, because the confounding effects of climate and season are not linear and slowly varying, these need to be modeled using smooth functions such as smoothing splines or regression splines (22). Non linear modeling increases both the number of nuisance parameters and the computational complexity. In summary, in all these cases, a sound and robust assessment of the statistical uncertainty of β_c can be hard to obtain, and variance underestimation due to model miss-specification is likely to occur.

Problems inherent in standard error estimation of air pollution effects have recently been

pointed out in the literature. For example, Ramsey et al (2002) (23) reported that the inability of the GAM to properly take into account correlation among non linear confounders can lead to underestimation of the standard error of relative rate estimates (see also Chambers and Hastie (1992, pp 303-304) (24) and commentaries by Lumley and Sheppard (2003) and Samet et al. (2003) (25; 26)).

Recently Dominici, McDermott, and Hastie (2003) have developed gam.exact, an extended version of the S-plus function gam used to fit Generalized Additive Models. The S-plus function gam.exact calculates an asymptotically exact covariance matrix of the air pollution effect estimates by properly taking into account the correlation among the non linear predictors.

The re-analyses of the National Morbidity Mortality Air Pollution Study (NMMAPS) (27) empirically confirmed the theoretical results of Ramsey et al. (2002) (23), and showed that the degree of bias in the standard errors was proportional to the size of the standard errors (a form of multiplicative bias). More robust variances than the ones obtained from GAM software can be obtained by using standard statistical theory (28), by bootstrap (29) or by GEE methods (30). However in time series studies of air pollution and health, such methods might be computationally expensive and 'off-the-shelf' statistical software is not always available.

In this paper, we investigate the sensitivity of the pooled estimate $\hat{\alpha}$ with respect to underestimation of the city-specific statistical variances v_c. We focus on underestimation (as

opposed to overestimation) for three reasons. First, underestimation of v_c is a much more serious problem than overestimation because it leads to less conservative conclusions about the statistical significance of a site-specific association between air pollution and health. Second, underestimation of v_c is more common than overestimation as the former generally reflects failing to take into account one or more sources of uncertainty. Third, underestimation of v_c has been identified as a limitation of the statistical software for the implementation of GAM (24; 23). However, this limitation has recently been overcome.

We show that the pooled estimate is unaffected by underestimation of v_c when: 1) the bias is additive, or 2) the bias is multiplicative and the statistical variances are equal across cities. Then by a simulation study, we investigate the case of multiplicative bias and unequal statistical variances. Here we define $3^3 = 27$ scenarios which are a combination of: 1) number of cities; 2) magnitude of the bias; and 3) amount of heterogeneity, and identify under which scenarios the underestimation of v_c affects the estimate of the pooled relative rate estimate $\hat{\alpha}$. We also investigate the impact of variance underestimation on the national average relative rate estimate from the National Morbidity Mortality Air Pollution Study (NMMAPS).

6

2 Methods and Results

We considered the following two stage normal-normal hierarchical model (11)

$$
\begin{aligned}
\hat{\beta}_c &= \beta_c + N(0, v_c), \ c = 1, \ldots, C \\
\beta_c &= \alpha + N(0, \tau^2)
\end{aligned}
\tag{1}
$$

where C is the total number of sites, and $N(a, b)$ denotes the normal distribution with mean a and variance b.

We estimated α and its standard error $SE(\hat{\alpha})$ by using an Empirical Bayes approach (EB)(18). More specifically, we first computed the Restricted Maximum Likelihood Estimate (REML) $\hat{\tau}^2$, and then we calculated $\hat{\alpha}$ and $SE(\hat{\alpha})$ conditional on $\hat{\tau}^2$. Details are in the appendix. The EB estimate of α and its standard error are defined below:

$$
\begin{aligned}
\hat{\alpha} &= \left(\sum_c \hat{\beta}_c / w_c \right) / \left(\sum_c 1 / w_c \right) \\
SE(\hat{\alpha}) &= 1 / \sqrt{\sum_c 1 / w_c}.
\end{aligned}
\tag{2}
$$

where $w_c = v_c + \hat{\tau}^2$. Let v_c be the "true" statistical variance, and let $v_c^\star$ be the "estimated" statistical variance under a miss-specified model. We assumed $v_c^\star < v_c$, and we considered three cases of underestimation of v_c:

1. additive bias: $v_c^\star = v_c - b$, for arbitrary v_c and $b > 0$;

2. multiplicative bias, equal variances: $v_c^\star = k \times v_c$, $v_c \equiv v$ for arbitrary v and $0 < k < 1$;

3. multiplicative bias, unequal variances: $v_c^\star = k \times v_c$, for arbitrary v_c and $0 < k < 1$.

7

In the first and second case, underestimation of the variance does not affect the pooled estimate of α. The mathematical proof is given in the appendix.

In the third case, we explored the impact of multiplicative bias on the estimation of α by a simulation study. We considered $3^3 = 27$ scenarios consisting of three components:

- underestimation of $50\%, 30\%$ and 10% which corresponds to $k = 0.5, 0.7, 0.9$;

- number of sites: $C = 15, 20, 90$;

- amount of heterogeneity: $\tau = 0.05, 0.5, 1$ corresponding, respectively, to small, medium, and large between city standard deviations.

For each scenario, we generated 250 replicates of $\hat{\beta} = (\hat{\beta}_1, \ldots, \hat{\beta}_C)$ from model (1). More specifically, we chose values for α, τ^2, and v_c and we sequentially generated: 1) β_c from $N(\alpha, \tau^2)$, and 2) $\hat{\beta}_c$ from $N(\beta_c, v_c)$. For $C = 90$, we set v_c equal to the estimates obtained from the NMMAPS re-analysis (27). For $C = 15$ and $C = 20$ we took a random sample from the 90 v_c NMMAPS estimates. We also set α equal to 0.21 (the pooled NMMAPS estimate for total mortality at lag 1 (27)). In summary, each scenario (biased $v_c^\star$, sample size C, and amount of heterogeneity τ), led to 250 simulated values of $\hat{\beta}$.

For each replicate $\hat{\beta}$, and for each scenario, we calculated the empirical Bayes estimates of τ^2 and α, using both v_c and $v_c^\star$, by fitting model (1) and using equation (2). This led to two sets of 250 estimates of α, here denoted by $\hat{\alpha}(v_c)$ and $\hat{\alpha}(v_c^\star)$. Figure 1 shows boxplots of the 250 standardized differences between the two α estimates defined as $(\hat{\alpha}(v_c) - \hat{\alpha}(v_c^\star))/SD$, where

8

SD is the standard deviation of the 250 estimates of $\hat{\alpha}(v_c)$. Thus a standardized difference larger than 2 suggests high sensitivity of the pooled estimate to the use of the biased $v_c^\star$ instead of "true" v_c.

For all scenarios, the 95% confidence intervals of the standardized differences were always within 2 standard deviations of 0, suggesting that underestimation of v_c did not affect $\hat{\alpha}$ substantially. In eight scenarios, the distributions of the standardized differences showed more variability with their maximum absolute differences larger than 2 standard deviations. Seven of those scenarios corresponded to extreme underestimation of 50%. The other scenario was characterized by small sample size, small heterogeneity, and a bias of 30%.

We also investigated the impact of variance underestimation on the NMMAPS national average relative rate estimate. We re-calculated the national average relative rate of mortality for a 10 unit increase in PM_{10} by varying the underestimation parameter k from 0.1 (90% variance underestimation) to 1 (no variance underestimation). Figure 2 shows the NMMAPS national average and heterogeneity estimates as function of the underestimation parameter k. Note that for values of k increasing from 0.1 to 0.6 (variance underestimation from 90 to 40 percent), $\hat{\tau}$ decreased, and $\hat{\alpha}$ moved from the un-weighted average of the $\hat{\beta}_c$ (equal to 0.28%, 95% posterior interval from .05 to .51) toward the weighted average of the $\hat{\beta}_c$ defined in equation (2) (equal to 0.21% 95% posterior interval from .10 to .31). The latter occurred because when k was very small (large variance underestimation), $\hat{\tau}^2$ was large relative to $v_c^* = k v_c$, and roughly the same weight was assigned to all the city-specific estimates. For values of k increasing from 0.6 to 1 (variance underestimation from 40 to 0 percent), $\hat{\alpha}$ was

9

constant because $\hat{\tau}^2 = 0$.

In summary, little or no effect was observed when the variance underestimation was less than 40% ($k \geq 0.6$) (national average estimate equal to 0.21%). When underestimation was larger than 40% ($k < 0.6$), $\hat{\alpha}$ gradually increased with the degree of underestimation toward the un-weighted pooled estimate which is equal to 0.28%. However, for all values of k, the t-ratio $\hat{\alpha}/SE(\hat{\alpha})$ remained larger than 2 indicating "statistical significance" of the national average relative rate estimate.

3 Discussion

The results of this paper indicated that in multi-site time series studies of air pollution and health, underestimation of the statistical variances v_c did not affect the estimate of the pooled effect α substantially. Some sensitivity was observed when the number of sites was small (less than 20), the between city variability was close to zero, and the underestimation was larger than 40%. In addition, our results suggest that variance underestimation as much as 40% in the NMMAPS studies would not change the conclusions.

The mathematical results reported in this paper are generalizable to any two-stage normal-normal hierarchical model, and therefore can be applied to any meta-analysis of study results or clinical trials that use such modeling approach. The simulation results reported in this paper can also be applied to any meta-analysis when the size of the overall parameter α with

respect to the heterogeneity parameter τ and the city specific standard deviations, $v_c^{1/2}$ are similar to the ones considered here. However, even in a meta-analysis with very different numerical values for α, τ and v_c, our simulation studies can be easily modified to reflect other situations.

Our underestimation parameter k, here specified a priori, may indeed depend upon site-specific characteristics. In each city, data characteristics may lead to different autocorrelation structures of the residuals, and therefore to different degree of adjustment for confounding factors. Therefore, variance underestimation attributable to model miss-specification might be more or less severe in the different cities, leading to different values of k. Our simulations study in Section 2 consider both minor (10%) and major (50%) standard error underestimation and therefore should encompass the more severe cases of model miss-specification.

The robustness of air pollution effect estimates to variance underestimation applies only to pooled effect α. Empirical Bayes (EB) estimates of the site-specific relative rates, here denoted by $\tilde{\beta}_c$, are affected by the underestimation of the v_c. To see this, consider their definition:

$$\tilde{\beta}_c = \hat{\beta}_c S_c + \hat{\alpha}(1 - S_c)$$

where $S_c = v_c^{-1}/(\hat{\tau}^{-2} + v_c^{-1})$ is the shrinkage factor. Note that S_c is not invariant to underestimation of the v_c. In this case, underestimation of v_c leads to overestimation of S_c

and therefore leads to $\tilde{\beta}_c$ which rely too heavily on $\hat{\beta}_c$. Thus, variance underestimation leads to an overestimation of the heterogeneity of the air pollution effects, and therefore to under-shrinkage of the city-specific EB estimates toward their overall mean $\hat{\alpha}$.

Unfortunately because the true statistical variances are unknown, the distinction between additive versus multiplicative bias is not straightforward. One possibility is to specify two alternative but comparable site-specific regression models and compare the estimates of v_c under the two models. For example Dominici et al. (2002) estimated $\hat{\beta}_c$ and v_c under a GLM with natural cubic splines and under a GAM with smoothing splines and compared the v_c under these two modelling approaches. In time series studies of air pollution and health, a more robust estimate of the statistical variance under a GAM can be obtained by using gam.exact (Dominici, McDermott, Hastie, 2003). In other situations, robust variance estimates can be obtained by using: 1) asymptotic theory (28); 2) bootstrap (29); 3) or sandwich estimates (30). Robust and model-based variance estimates can then be compared for a small number of cities (or sites) to gain insight on the form of bias.

Results in this paper are presented using an empirical Bayesian approach to estimation, where the Bayesian estimate of α is obtained by plugging in a point estimate of τ^2 in Equation (2). However our results apply also if a fully Bayesian version of the above model are fit with non-informative priors on τ^2 and α.

Finally, the results of this paper apply under the assumption that the normal approxima-tion to the likelihood function at the first stage of the hierarchical model is appropriate.

Asymptotically, this approximation has an accuracy proportional to the number of days with available data in each city (31). In time series studies of air pollution and health the asymptotic normal approximation is generally accurate; however additional work is needed to extend such results to distributions other than the normal and to examine the sensitivity of inferences if the normal approximation is not accurate.

Appendix

Details on the estimation of τ^2 and α. The EB estimate of α was obtained by first computing the REML estimate of τ^2, which was then plugged into equation (2). The REML estimate of τ^2 was obtained by maximizing the following likelihood function

$$Lik(\tau^2|\hat{\beta}_c, v_c) = \left(\sum_c w_c(\tau^2)^{-1}\right)^{-1/2} \prod_{c=1}^{C} w_c(\tau^2)^{-1/2} \exp\left\{-\tfrac{1}{2}\sum_{c=1}^{C} w_c(\tau^2)^{-1}\left[\hat{\beta}_c - \hat{\alpha}(\tau^2)\right]^2\right\}$$

(3)

where $\hat{\alpha}(\tau^2) = \left[\sum_c \hat{\beta}_c w_c(\tau^2)^{-1}\right] / [\sum_c w_c(\tau^2)^{-1}]$ and $w_c(\tau^2) = v_c + \tau^2$. This maximization can be done iteratively using the E-M algorithm (18).

Mathematical arguments concerning the robustness of the pooled estimate to variance underestimation under cases 1 and 2.

Under scenario 1, we start by showing that additive bias does not affect estimation of α

for arbitrary v_c. Let's $\tau^{2\star}$ be the maximum likelihood estimate (mle) of τ^2 obtained by maximizing the likelihood (3) with $v_c^\star = v_c - b$ instead of v_c. Note that estimation of α as defined in equation (2) depends on v_c only through $v_c + \tau^2$. Therefore we simply need to show that the estimate for $\tau^{2\star}$ conditional on $v_c^\star$ is b units more than the estimate for τ^2 conditional on v_c. In this case $v_c + \hat{\tau}^2 = v_c^\star + \hat{\tau}^{2\star}$ and therefore $\hat{\alpha}(\hat{\tau}^2)$ is unaffected by underestimation of v_c. We define $\tau^{2\star} = \tau^2 + b$ and we maximize the likelihood (3) (with v_c replaced by $v_c^\star$) with respect to $\tau^{2\star}$. By the invariance property of the maximum likelihood estimates, if $\hat{\tau}^{2\star}$ is mle then $\hat{\tau}^2$ is also mle. Therefore $v_c + \hat{\tau}^2 = (v_c - b) + (\hat{\tau}^2 + b) = v_c^\star + \hat{\tau}^{2\star}$, $\hat{\tau}^{2\star} = \hat{\tau}^2 + b$.

Under scenario 2, we assume multiplicative bias with equal variances $v_c = v$. Here $v_c^\star = k \times v_c$ can be re-written as $v_c^\star = v_c - (1 - k) \times v_c$. Because $v_c = v$, then $(1 - k) \times v_c = (1 - k) \times v = b$ and therefore $v_c^\star = v_c - b$. So the results above for additive bias apply here also.

References

[1] Katsouyanni K, Touloumi G, Spix C, et al. 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, 1997.

[2] Samet JM, Zeger SL, Dominici F, Dockery D, Schwartz J. The National Morbidity, Mortality, and Air Pollution Study Part I: Methods and Methodological Issues. Health Effects Institute, Cambridge, MA., 2000.

[3] Thomas A, Spiegelhalter DJ, Gilks WR. Bugs: A program to perform Bayesian inference using Gibbs sampling. In Bayesian Statistics 4. Proceedings of the Fourth Valencia International Meeting, 837–842. Clarendon Press: Oxford, 1992.

[4] Gilks WR, Richardson S, Spiegelhalter DJ, eds. Markov Chain Monte Carlo in Practice. Chapman and Hall, London, 1996.

[5] Burnett R, Krewski D. Air pollution effects of hospital admission rates: A random effects modelling approach. The Canadian Journal of Statistics, 22:441–458, 1994.

[6] Roemer W, Hoek G, Brunekreef B, Kalandidi A, Pekkanen J. Daily variations in air pollution and respiratory health in a multicenter study: the peace project. European Respiratory Journal, 12:1354–1361, 1998.

[7] Dominici F, Samet JM, Zeger SL. Combining evidence on air pollution and daily mortal-

ity from the twenty largest US cities: A hierarchical modeling strategy (with discussion). Royal Statistical Society Series A with discussion, 163:263–302, 2000.

[8] Zanobetti A, Schwartz J, Dockery D. Airborne particles are a risk factor for hospital admissions for heart and lung disease. Environmental Health Perspective, 108:1071–1077, 2000.

[9] Schwartz J. Assessing confounding, effect modification, and thresholds in the associations between ambient particles and daily deaths. Environmental Health Perspective, 108:563–568, 2000.

[10] Dominici F. Air pollution and health: what can we learn from an hierarchical approach? invited commentary. American Journal of Epidemiology, 1:11–15, 2002.

[11] Lindley DV, Smith AFM. Bayes estimates for the linear model (with discussion). Journal of the Royal Statistical Society Series B Methodological, 34:1–41, 1972.

[12] Samet JM, Zeger SL, Berhane K. The Association of Mortality and Particulate Air Pollution. Health Effects Institute, Cambridge, MA., 1995.

[13] Kelsall J, Samet JM, Zeger SL. Air pollution, and mortality in Philadelphia, 1974-1988. American Journal of Epidemiology, 146:750–762, 1997.

[14] Hastie TJ, Tibshirani RJ. Generalized additive models. Chapman and Hall, New York, 1990.

[15] McCullagh P, Nelder JA. Generalized Linear Models (Second Edition). New York: Chapman & Hall, 1989.

[16] Hedges LV, Olkin I. Statistical Methods for Meta-analysis. Academic Press, 1985.

[17] DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled Clinical Trials, 7:177–188, 1986.

[18] Carlin BP, Louis TA. Bayes and Empirical Bayes Methods for Data Analysis. Chapman & Hall, New York, 1996.

[19] Morris CN, Normand SL. Hierarchical models for combining information and for meta-analysis. In JM Bernardo, JO Berger, AP Dawid, AFM Smith, eds., Bayesian Statistics 4, 321–344, Oxford, 1992. Oxford University Press.

[20] Gelman A, Carlin J, Stern H, Rubin D. Bayesian Data Analysis. Chapman and Hall, London, 1995.

[21] Buja A, Hastie T, Tibshirani R. Linear smoothers and additive models. The Annals of Statistics, 17:453–455, 1989.

[22] Green PJ, Silverman BW. Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach. Chapman & Hall, London UK, 1994.

[23] Ramsay T, Burnett R, Krewski D. The effect of concurvity in generalized additive models linking mortality and ambient air pollution. Epidemiology to appear, 2002.

[24] Chambers J, Hastie T. Statistical Models in S. Chapman and Hall, London, 1992.

[25] Lumley T, Sheppard L. Time series analyses of air pollution and health: Straining at gnats and swallowing camels? Epidemiology, 14:?–?, 2003.

[26] Samet J, Dominici F, McDermott A, Zeger S. New problems for an old design: Time series analyses of air pollution and health. Epidemiology, 14:11–12, 2003.

[27] Dominici F, McDermott A, Zeger SL, Samet JM. Airborne particulate matter and mortality: Time-scale effects in four us cities. American Journal of Epidemiology in press, 2002.

[28] van der Vaart AW. Asymptotic Statistics. Cambridge Series in Statistics and Probalistic Mathematics, Cambridge University Press, 1998.

[29] Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Chapman & Hall, New York, 1993.

[30] Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika, 73:13–22, 1986.

[31] Daniels M, Kass R. A note on first-stage approximation in two-stage hierarchical models. Sankhya Series B, 60:19–30, 1998.

Figures

- Figure 1: Boxplots of the standardized differences between the α's estimates under model (1) with v_c and $v_c^{\star}$. The three digit labels, (i, j, l) on the x-axis, correspond to the simulation scenarios. The first digit corresponds to the underestimation parameter $(l = 1 : k = 0.10; l = 2 : k = 0.30;$ and $l = 3 : k = 0.50)$. The second digit corresponds to the heterogeneity parameter $(j = 1 : \tau = 0.05; j = 2 : \tau = 0.5;$ and $j = 3 : \tau = 1.0)$. The third digit corresponds to the sample size $(i = 1 : C = 15; i = 2 : C = 20;$ and $i = 3 : C = 90)$.

- Figure 2: (a) National average relative rate estimate from NMMAPS and their 95% posterior intervals, denoting the percentage increases in all causes mortality associated with a 10 mug/m^3 increase in PM_{10} at lag 1, as a function of the underestimation parameter. (b) Estimate of the heterogeneity parameter τ, denoting the standard deviation across cities of the "true" city-specific relative rates, as function of the underestimation parameter k.

(a)

(b)

national average estimate of the relative rate of mortality

heterogeneity parameter

underestimation parameter(k)

underestimation parameter(k)

21