

1. Introduction

The purpose of this paper is to re-examine the evidence of an association between ambient ozone and non-accidental all-cause mortality, based in particular on a series of papers by Bell and co-authors that used the NMMAPS database. NMMAPS refers to the National Morbidity, Mortality and Air Pollution Study, in which daily data on mortality, meteorology and various air pollutants were collected for approximately 100 U.S. cities. Bell et al. (2004) calculated estimates for the ozone-mortality coefficient in each city after adjusting for meteorology, seasons, long-term trends and a day of week effect. The results were highly variable from city to city, showing both positive and negative associations, but after a second-stage analysis that combined the results for 95 cities, they claimed a strong overall positive association. Bell et al. (2006) extended this to allow for the possibility of a nonlinear exposure-response association, claiming in particular that the effects of ozone and mortality persist down to quite low levels of ozone. Bell et al. (2007) developed and reinforced the argument first made in Bell et al. (2004), that there is no significant confounding effect due to particulate matter. Bell and Dominici (2008) noted that there is substantial regional and spatial variation in the ozone-mortality coefficient, but claimed that a large part of this could be explained in terms of effects modifiers, in which variation of the ozone-mortality coefficient from city to city is correlated with socio-economic or other environmental variables measured for each city. These papers have been highly influential: for instance, they were extensively cited in the documentation supporting the 2008 decision of the U.S. Environmental Protection Agency (EPA) to revise the ozone standard, and more recently in a National Research Council report on the possible benefits of reducing ambient ozone (NRC (2008)).

In this paper, we re-examine all of these issues with particular attention to the question of sensitivity to alternative analyses. We look extensively at alternative treatments of meteorology and co-pollutants, showing that there are confounding and effect modifier relationships that have been understated or overlooked in previous studies. We also look at the dependence among different ozone metrics – those based on a 24-hour average, on the daily maximum 8-hour average, and on the daily maximum 1-hour value. Most of the work by Bell and co-authors used 24-hour average ozone, but the current EPA standard uses the daily maximum 8-hour average, which itself superseded an earlier standard based on the 1-hour maximum. We argue that conversion between these metrics using standardized ratios is potentially misleading. We provide evidence of regional and spatial variability, and argue that explanations in terms of effect modifiers, as in Bell and Dominici (2008), are at best only part of the answer – however, we provide a new interpretation of the possible effect of air conditioning, noted by them as well as by Levy et al. (2005). We also look closely at the evidence for a non-linear exposure-response association.

2. Data and Statistical Methods

The principal data source for our analysis is the NMMAPS database, available publicly at <http://www.ihapss.jhsph.edu>. We have used daily mortality counts for 1987-2000, excluding accidental deaths, and subdivided them into three age categories (<65, 65-74, 75 and over) for 98 cities. The 97 cities (excepting Honolulu) that are in the mainland US are classified into seven regions (Industrial Midwest, North East, North West, Southern California, South East, South

92 West, Upper Midwest) using a classification that has been used in numerous NMMAPS reports
93 (see e.g. Samet et al. (2000b), page 8). Other variables that are part of the dataset and have been
94 used in the present study are daily temperature, dewpoint, particulate matter of aerodynamic
95 diameter equal to or less than 10 microns (PM_{10}) or 2.5 microns ($PM_{2.5}$), and both hourly and
96 daily values for ozone, in most cases with no detrending and using ordinary rather than trimmed
97 means for combining results across monitors. These variables were used in computing ozone-
98 mortality coefficients for each city under a variety of statistical models. Since, in many cities,
99 ozone data are only available for summer months (April through October), we computed ozone-
100 mortality coefficients that are restricted to those months, as well as those using all available data
101 (henceforth called all-year estimates).

102
103 In addition to the above, our analysis considered numerous variables, defined at the city level,
104 that are used later in the paper for an effects modifier study. These include 77 “city census”
105 variables that are part of the NMMAPS database. Although some are duplicates and not all are
106 suitable for use as demographic covariates, the majority are indicators of population distribution,
107 housing, educational level, race, income, etc. In many cases we have standardized the census
108 data to proportions rather than counts, for example, racial variables have been expressed as
109 proportions of total population. The database includes indicators of the numbers of age 65+
110 residents (subdivided into owners and renters) who last moved during different time periods.
111 These have been consolidated into the proportion of all age 65+ residents who have moved since
112 1995, and similar proportions for age 65+ owners and for age 65+ renters. We also included six
113 “environment” variables, also constructed from the NMMAPS database: mean overall values for
114 each city of ozone (summer only, to ensure comparability among means for different cities),
115 PM_{10} and $PM_{2.5}$, sulfur dioxide (SO_2), daily mean temperature (average of daily maximum and
116 daily minimum), and dewpoint.

117
118 We independently constructed a table of air conditioning use by city. Bell and Dominici (2008)
119 noted that such data are included in two sources of the American Housing Survey (AHS 2006): a
120 “metropolitan” survey which is an extensive survey of selected metropolitan areas conducted
121 every four or six years, and a “national” survey that comprises at least 55,000 houses every two
122 years, that are widely scattered across the United States. However the metropolitan survey covers
123 only about half the NMMAPS cities, which in preliminary analyses proved inadequate to
124 demonstrate a clear-cut association with the city-specific ozone-mortality coefficients. We
125 therefore used national survey estimates from 1997, 1999, 2001, 2003 and 2005, collating
126 reported air-conditioning use with location defined by standard metropolitan statistical area
127 (SMSA). Two particular variables were compiled, one an indicator of central air condition and
128 the other of window air conditioners. Bell and Dominici (2008) combined the two into “all AC”,
129 but also considered “central AC” as a separate variable; they did not consider “window AC” as
130 we do here. Average AC use was computed across all available years for each NMMAPS city,
131 weighted by the number of available survey results in each year. By this method we were able to
132 compute air conditioning use statistics for 79 of the 98 NMMAPS cities. It should be noted that
133 the time period covered by the AC surveys is not the same as that of the NMMAPS data, but the
134 inter-year correlations were high (.96 to .99 for central AC; .82 to .93 for window AC) so it
135 seems likely that AC use during 1997-2005 is also a good proxy for AC use during the earlier
136 time period of the NMMAPS mortality data.

137

138 This paper uses a wide variety of statistical methods. Most analyses used the R statistical
139 package (R Development Core Team (2007)). The initial analysis follows the method described
140 by Bell et al. (2004). It is a two-stage analysis where the first stage is a Poisson regression
141 allowing for overdispersion, with mortality (subdivided into three age groups) as the response
142 variable. We allowed for meteorology by defining covariates: temperature and dewpoint on the
143 current day (lag 0), plus two additional variables representing averages of temperature and
144 dewpoint over lags 1-3, each adjusted for the current day's temperature and dewpoint (these
145 values are included in the NMMAPS database based on a calculation originally described by
146 Curriero et al. (2002)). For reference in future discussion, we call this the NMMAPS
147 meteorology model. The meteorological variables were modeled nonlinearly using natural
148 splines, typically with 6 degrees of freedom for temperature and 3 degrees of freedom for
149 dewpoint, the same as Bell et al. (2004). Long-term trends were also modeled nonlinearly,
150 through natural splines with 7 degrees of freedom per year, and an interaction for the three age
151 groups where we used the public R code cited by Bell et al. (2004). Day of week was also
152 included as a covariate. Finally we included ozone initially through 24-hour averages, but with
153 various combinations of lags. Ozone may be represented as a single day's value at lag 0, 1 or 2,
154 or as the average of lags 0 and 1 (we call this the 0-1 model), or through the "constrained
155 distributed lag" model of Bell et al. (2004). In the latter, ozone is represented through the value
156 at lag 0, the average of lags 1,2 and the average of lags 3,4,5,6, each with its own regression
157 coefficient, and the sum of the three coefficients taken as the "ozone effect". In practice this is
158 rewritten so that one coefficient represents the overall effect and the other two as contrasts,
159 allowing us to estimate a single "ozone-mortality coefficient" and its standard error for each city.
160 An alternative is the "unconstrained distributed lag" model in which each of lags 0 through 6
161 gets a separate regression coefficient; Bell et al. (2004) prefer the constrained version but we
162 have found there is very little difference between the constrained and unconstrained distributed
163 lag models. We have also considered analyses based on 8-hour or 1-hour ozone, where the raw
164 data are hourly values from NMMAPS, and we have defined 8-hour ozone as the maximum of
165 the average of 8 consecutive hourly values in any 24-hour period, and 1-hour ozone as the largest
166 of the hourly values on a given day. In computing 8-hour ozone, we have followed the 75% rule
167 (also used by EPA): 75% of the hourly values in any 8-hour period are needed for the 8-hour
168 average to be counted, and 75% of the 8-hour averages within a 24-hour period are needed for
169 the daily maximum 8-hour average to be counted. All days not meeting these criteria are flagged
170 as missing values. In analyses that also include PM₁₀ or SO₂, we have used the 24-hour average
171 at lag 1, since in earlier NMMAPS analyses (Dominici et al. (2003)) this was found to be the
172 most significant lag for both PM₁₀ and SO₂.

173
174 We considered two alternatives to the NMMAPS meteorology model. One is a simple
175 replacement of the adjusted lagged temperature and dewpoint variables by the unadjusted three-
176 day means of temperature and dewpoint over lags 1-3. The second is what we call the
177 "extended" meteorology model. This is a distributed lag model for meteorology, using separate
178 temperature and dewpoint values for each of lags 0 through 6, each modeled nonlinearly through
179 natural splines of respectively 4 and 3 degrees of freedom. The rationale behind this model is
180 that, if the ozone effect is to be represented by a distributed lag model, then for comparability
181 purposes, meteorology should be represented in a similar way.

182

183 These variables were included in an overdispersed Poisson regression model, also known as a
184 generalized linear model (GLM), and fitted using the “glm” function in R. The result of these
185 analyses was an estimate, together with its standard error, of the ozone-mortality coefficient in
186 each city, expressed in percent rise in mortality per 10 ppb rise in ozone.

187
188 The second phase of the statistical analysis is based on the random effects model

$$189 (1) \theta_c \sim N[\mu, \tau^2],$$

$$190 (2) t_c | \theta_c \sim N[\theta_c, s_c^2]$$

191 where θ_c is the true coefficient in city c , t_c is the estimate from stage 1, s_c is the standard error
192 from stage 1 and N denotes normal distributions (approximate in the case of (2), since the
193 distribution of a GLM estimate is known to be approximately normal). Equations (1)-(2) describe
194 a two-level normal model where the objective is to estimate the hyperparameters μ and τ^2 to
195 obtain posterior distributions for the θ_c . A convenient method for doing this is the Bayesian
196 algorithm *tlmise* that was developed by Everson and Morris (2000) and is available as a
197 downloadable package for R, written by Dr. Roger Peng from the earlier S-Plus code of Everson.
198 However we have also used a restricted maximum likelihood (REML) approach that is described
199 in Appendix 1 and produces very similar results.

200
201 For reasons described in more detail later, we do not accept that μ is the most appropriate
202 parameter to define a “national mean”. An alternative approach is to take the weighted average
203 of the city-based estimates across all cities, using weights proportional to the population in each
204 city. That is natural, because risk assessments in effect try to estimate the total population at risk,
205 so the contribution to the total risk from a given city should be proportional to the population of
206 that city. Mathematically, we formulate this as predicting $\sum w_c \theta_c$ where the weight w_c is
207 proportional to the population of city c . The optimal predictor when τ^2 is known is $\sum (w_c t_c \tau^2 / (s_c^2 + \tau^2)) + M \sum (w_c s_c^2 / (s_c^2 + \tau^2))$, where $M = [\sum t_c / (s_c^2 + \tau^2)] / [\sum 1 / (s_c^2 + \tau^2)]$ is the weighted least
208 squares estimator of μ , and the corresponding mean squared prediction error $\sum (w_c^2 s_c^2 \tau^2 / (s_c^2 + \tau^2)) + (\sum w_c s_c^2 / (s_c^2 + \tau^2))^2 / \sum (s_c^2 + \tau^2)$. In the case where τ^2 is unknown, we use the output from
209 *tlmise* to develop a Monte Carlo sample from the posterior distribution. These formulas are
210 applied to each member of the Monte Carlo sample to derive the mean and variance of the
211 posterior distribution. An outline derivation of these formulas is given in Appendix 2.

212
213
214
215 A major sensitivity issue is whether any other air pollutant could be confounding the ozone
216 effect. Bell et al. (2004) claimed that for the NMMAPS dataset, PM_{10} is not a significant
217 confounder for ozone. Bell et al. (2007) extended that claim to cover also daily $PM_{2.5}$, and gave
218 several other analyses. To the best of our knowledge, this question has not been examined with
219 respect to other possible co-pollutants, though Franklin and Schwartz (2008) showed a 35%
220 reduction in the ozone-mortality coefficient when sulfate particles were included in the model,
221 based on a different and substantially smaller dataset.

222
223 There are several difficulties in making this comparison for the NMMAPS dataset. Ideally we
224 would prefer to use $PM_{2.5}$ (fine particulates) as this is the variable used for the main EPA
225 standard for particulate matter. However, as pointed out by Bell et al. (2007), within the
226 NMMAPS dataset, data for are $PM_{2.5}$ very limited. For the present study we use PM_{10} on grounds
227 of data availability, though bearing in mind that an observed association with PM_{10} may in fact
228 be due to $PM_{2.5}$ or any other component of PM_{10} , such as sulfate particles. Even then, there is a

229 difficulty, because in most cities, PM₁₀ has been measured only once every six days, as required
230 by EPA regulations that were in force at the time the data were collected. Therefore, the only
231 realistic analysis is to repeat the foregoing ozone-only analysis, but restricted to days on which
232 PM₁₀ is also available, and then to run the same analysis with both ozone and PM₁₀ as pollutants.
233 Note that this methodology does not permit a distributed lag model for PM₁₀, since such a model
234 requires daily data, so we have taken PM₁₀ at lag 1, which has been found in previous studies
235 (e.g. Dominici et al. (2003)) to be the single-day lag of PM₁₀ that has the strongest association
236 with mortality. The data deficiency also has an implication for the effects themselves, that it is
237 much harder to detect statistically significant associations based on such a reduced dataset.

238
239 Apart from the possibility that meteorology or co-pollutants may act as confounders, there is the
240 somewhat different question of whether they may be effect modifiers. For example, is the effect
241 of ozone on mortality at high temperatures the same as at low temperatures? In this context, we
242 are considering *within-city* effect modifiers, where we examine the influence of the potential
243 effect modifier first within each city, only then combining the city estimates into an overall
244 population-weighted average. This is different from using *between-city* effect modifiers to
245 explain the variation in ozone-mortality coefficients from city to city, as was done in Bell and
246 Dominici (2008) and is discussed further below.

247
248 For temperature, the analysis is as follows. The “constrained distributed lag” model actually
249 contains three ozone variables, the mean over lags 0-6 and two other variables representing
250 differences in means over subsets of lags 0-6. If these three variables are written x_1 , x_2 , x_3 , then
251 we replace them by six variables, x_1w , $x_1(1-w)$, x_2w , $x_2(1-w)$, x_3w , $x_3(1-w)$, where w is an
252 indicator variable, 1 if the current day’s temperature is greater than the median temperature for
253 that city, 0 otherwise. We also include w itself as a covariate in the regression. Note that the
254 median temperature is different for each city, and in cases where summer-only data is used, the
255 median is also computed for summer-only data. The regression coefficients associated with the
256 first two variables, x_1w and $x_1(1-w)$, represent the overall ozone-mortality effects above and
257 below the temperature median. These coefficients are then combined across cities using a
258 population-weighted average. Similar analyses have been performed for two other potential
259 effect modifiers, PM₁₀ and SO₂. We considered SO₂ because Franklin and Schwartz (2008)
260 showed that sulfate particles may be a confounder of the ozone effect. Sulfate particles do not
261 form part of the NMMAPS dataset, but it is possible that gaseous SO₂ may be a proxy for sulfate
262 particles, or that SO₂ may independently have an effect on mortality. Simultaneous
263 measurements of ozone and SO₂ are available in 81 of the 98 NMMAPS cities, so the analysis is
264 confined to those cities.

265
266 In some of our analyses we also use a more general version of the random effects model,
267 replacing equation (1) by
268 (3) $\theta_c \sim N[x_c^T \beta, \tau^2]$,
269 in which x_c denotes a vector of covariates in city c and β is some vector of regression coefficients
270 (equation (2) remains the same). This can also be fitted using *tlmise*. The main application we
271 have made of this is for regional analyses, in which x_c represents an indicator vector for region
272 (e.g. if city c lies in the j th region out of m regions, x_c is a vector of dimension m consisting of 1
273 in the j th component and 0 in all other components) and the components of β are means for each

274 region. When combined with a uniform prior on β we call this the “regional prior” model, to
275 distinguish it from the “national prior” in equation (1).

276
277 A more general model still replaces the regression function $x_c^T \beta$ by a smooth function of latitude
278 and longitude. In the present paper, this has been implemented by treating the smooth function as
279 a sample from a continuous-parameter Gaussian stochastic process; the resulting
280 estimation/interpolation algorithm is similar to the geostatistics procedure known as kriging.
281 Appendix 3 contains mathematical details of this method.

282
283 For between-city effect modifier analyses, we could again proceed with a Bayesian analysis
284 based on (3), where the covariates x_c include the effect modifier. For the results presented here,
285 we have used a non-Bayesian analysis treating each city coefficient as a random effect whose
286 mean is a linear function of the effect modifier and whose variance is τ^2 , estimated by REML
287 using the method described in Appendix 1. This method is equivalent to the mixed effects meta-
288 regression described by Bell and Dominici (2008). One advantage of the REML approach in this
289 context is that it is very easy and quick to develop a test for statistical significance of the effect
290 modifier, based on the estimated regression coefficient and its standard error; the corresponding
291 Bayesian test, although well defined, requires integrating over a Monte Carlo sample and is
292 therefore less convenient to implement.

293

294 **3. Results**

295

296 The first step was to reproduce the main result of Bell et al. (2004), which used 24-hour ozone in
297 a “constrained distributed lag” model, with adjustments for daily mean temperature and
298 dewpoint, lagged values of temperature and dewpoint (days 1-3) adjusted for current day, plus
299 long-term trends and a day of week effect. The second stage of the analysis used a Bayesian
300 hierarchical model based on (1) and (2) (“national prior”, same as Bell’s analysis). Since the R
301 code to perform the individual city analyses is available on the NMMAPS website, we followed
302 that code to reproduce the Bell result exactly in this case. The resulting posterior estimates and
303 95% prediction intervals are shown in Fig 1, which corresponds to Fig 2 of Bell et al. (2004).

304 The units of all estimates are percent change in mortality associated with a 10 ppb rise in ozone.
305 Also plotted on this figure are the raw maximum likelihood estimates and the posterior means
306 under an alternative “regional prior” based on equation (3). The raw estimates are shown to make
307 the point that there is already a great deal of reduction in variance in passing from the raw to the
308 posterior estimates, even before computing the national average estimate (μ and its 95%
309 prediction interval) at the bottom of the plot. The estimates are, however, sensitive to the prior
310 distribution, as is clear from the fact that for many cities the posterior mean under the regional
311 prior is near one of the endpoints of the 95% prediction interval computed under the national
312 prior.

313

314 The posterior mean of μ (in this instance, 0.52, with a posterior standard deviation of 0.12) has
315 generally been interpreted as a national estimate of the ozone-mortality coefficient, and in many
316 publications, this has been quoted as if it were a truly representative effect. However, it is far
317 from clear what it represents. In this case the population-weighted mean of the raw estimates is
318 0.62, different from the posterior mean of μ . We could also consider the mean of raw estimates
319 weighted by reciprocal of squared standard error (the classical statistical rule for combining

320 estimates), which in this instance is 0.61, but this also does not correspond to μ . The last two
321 estimates should in most cases be very similar, since the variance of the parameter estimate
322 should be proportional to the reciprocal of population. However, the estimated μ is closer to the
323 unweighted mean of the raw estimates, which is 0.50.

324
325 The Bayesian analysis of the population-weighted mean leads to a posterior mean of 0.62, with a
326 posterior standard deviation of 0.10. Further detailed results, for this and a number of other
327 analyses to be described later, are in Table 1.

329 **Confounding and Effect Modification**

330
331 Next, we consider the sensitivity of these estimates to alternative treatments of meteorology, or
332 to co-pollutants. The “extended meteorology” models in Table 1 often lead to lower estimates of
333 the overall ozone-mortality coefficient, especially for all-year estimates, which suggests that they
334 are more efficient at dealing with the confounding effect of temperature. In most cases, the
335 difference in the national or population-weighted average point estimates is less than 10%, which
336 is not of very great practical significance. Later in the paper, we include the extended
337 meteorology model in some of the individual-city comparisons, to make a more specific
338 comparison between the two.

339
340 For the national-average effect (μ), we find a posterior mean ozone-mortality coefficient of 0.40
341 (posterior standard deviation 0.23) when fitted to days on which lag-1 PM_{10} is available but not
342 included in the model, and posterior mean 0.31 (0.23) when the same analysis is repeated
343 including PM_{10} as a co-pollutant. Similar results for the population-weighted average and for
344 summer-only data are included in Table 1, using both the NMMAPS and extended meteorology
345 models. It is difficult to evaluate the statistical significance of these results in view of the large
346 posterior standard deviations (a consequence of the 6-day sampling), but in most cases, the
347 ozone-mortality coefficient is reduced by between 22% and 33% as a result of including PM_{10} as
348 a co-pollutant.

349
350 Further evidence of this PM_{10} confounding effect is in Fig 2. The top plot here shows the raw
351 estimates for the ozone-mortality coefficient computed with PM_{10} , plotted against those without
352 PM_{10} . There is no visual evidence of any effect due to PM_{10} , though in fact 56 of the 93
353 estimates (60%) are below the diagonal straight line. This plot shows the same estimates as Fig 3
354 of Bell et al. (2004). However, the bottom plot shows the same comparison made with the
355 Bayesian posterior estimates. In this case, it is clear that nearly all (89 out of 93) the individual-
356 city coefficients are smaller for the model with PM_{10} than the model without PM_{10} . The most
357 likely explanation why the proportion increases so dramatically is that both sets of posterior
358 estimates (with and without PM) are based on “shrinking towards the mean” (i.e. μ in the
359 notation of equation (1)) but the mean itself is lower in the with-PM analysis (.314) than in the
360 without-PM analysis (.400). The emphasis that Bell et al. (2004) and other authors have given to
361 using the Bayesian hierarchical model to reduce the variability of coefficients would suggest that
362 the bottom plot is more meaningful and appropriate.

363
364 We next discuss results for the “within-city effect modifier” analyses; see Table 2, where we
365 have also included results for 8-hour and 1-hour ozone, as considered later in the paper. For all

366 three ozone metrics, temperature is not an effect modifier in the all-year analyses, but for the
367 summer analyses, it is. In other words, during the summer months, the mortality effect of ozone
368 is statistically significant only when temperature is above the median. We find that SO₂ is a
369 statistically significant effect modifier in the all-year data based on 24-hour ozone (the ozone-
370 mortality effect higher at high SO₂), but not for summer-only data. Moreover, the statistical
371 significance of the all-year result does not extend to the 8-hour and 1-hour ozone metrics.
372 Overall, the evidence that SO₂ is an effect modifier does not seem strong. We also looked for
373 evidence that SO₂ is a confounding variable, but concluded that it was not.

374
375 The result for PM₁₀ is perhaps the most surprising, given the data sparseness issue already
376 discussed, and the inconclusive results over whether PM₁₀ is a confounder. For all three ozone
377 metrics and both all-year and summer-only data, the ozone-mortality effect is statistically
378 significant only when PM₁₀ is above the median. In other words, although there is ambiguous
379 evidence whether PM₁₀ is a confounder, it clearly is an effect modifier.

380
381 **Ozone Metrics**

382
383 We examined whether results calculated using 24-hour averaged ozone remain essentially
384 unchanged when repeated using either the daily maximum 8-hour average or the daily maximum
385 1-hour value.

386
387 Fig 3 (top plot) shows posterior mean estimates by city, under the national prior, plotted against
388 each other based on the 24-hour and 8-hour ozone metrics. The bottom plot shows the same
389 thing comparing the 1-hour and 8-hour metrics. Corresponding plots were drawn for the raw
390 estimates but are not shown here because the overall appearance was similar. These calculations
391 show a correlation of the order of 0.7-0.8 between city-specific estimates under the different
392 metrics, regardless of whether raw or posterior mean estimates are used. This is a moderately
393 high correlation, but it also shows that the correspondence between the different metrics is by no
394 means perfect. Summary results for 8-hour and 1-hour ozone are included in Table 1, based on
395 percent rise in mortality per 10 ppb rise in ozone in the corresponding ozone metric.

396
397 The question naturally arises of how to express ozone-mortality results for the different metrics
398 on a common scale so that they are directly comparable. Bell et al. (2004) provided results for
399 the percent rise in mortality per 15 ppb rise in 8-hour ozone or per 20 ppb rise in 1-hour ozone
400 (compared with 10 ppb for 24-hour ozone), implying without stating directly that this made the
401 estimates roughly comparable.

402
403 This is not an easy question to resolve. One possible solution is to rescale the ozone-mortality
404 coefficients proportionally to the overall mean ozone level in the respective metrics. This would
405 imply, for example, that we multiply the 24-hour ozone-mortality coefficient by 0.646 to convert
406 it to a scale comparable to the 8-hour ozone-mortality coefficient. Here, 0.646 is the ratio,
407 aggregated over all NMMAPS cities, of the mean level of 24-hour ozone to the mean level of 8-
408 hour ozone. By comparison, Bell et al. used an implied ratio of 0.667 (=10/15). However, the
409 value of 0.646 is by no means constant over cities: individual-city estimates of the same ratio
410 vary from 0.515 to 0.801. The issue is similar for the comparison between 1-hour and 8-hour
411 ozone estimates: the overall averages of 1-hour and 8-hour ozone lead to a factor of 1.31 for

412 converting 1-hour ozone-mortality coefficient to corresponding 8-hour values. However, the
413 ratio for individual cities varies from 1.08 to 2.05. Bell et al. used an implied ratio of 1.33
414 (=20/15). Ratios for summer-only ozone are almost the same (0.640 and 1.30 respectively to
415 convert 24-hour and 1-hour ozone-mortality coefficients to their 8-hour equivalents).

416
417 Using the factors 0.646, 1, 1.31 to convert results for all three ozone metrics into an equivalent
418 based on 8-hour ozone, the 24-hour, 8-hour and 1-hour results for all-year data translate into
419 posterior means of 0.40 (standard deviation 0.06), 0.47 (0.06) and 0.39 (0.06). The
420 corresponding result for summer only are 0.30 (0.07), 0.36 (0.07) and 0.33 (0.06). These results
421 support the notion that of the three metrics, 8-hour ozone leads to the strongest association with
422 mortality, but subject to the heavy caveat that they are based on conversion factors that are
423 highly variable over cities.

424
425 Figs 4 and 5 respectively show the individual-city estimates (raw estimates, posterior means
426 under the national prior, and posterior means under the regional prior) for the associations of 8-
427 hour and 1-hour ozone with mortality. Also shown are 95% prediction intervals under the
428 national prior. Comparing these with Fig 1, there is in fact greater homogeneity in the national-
429 prior posterior means than is the case for 24-hour ozone. This might be considered another
430 argument for using 8-hour or 1-hour ozone in preference to 24-hour ozone. However, plotting the
431 raw and regional-prior estimates on the same graph also highlights that the raw estimates are still
432 much more variable than the posterior estimates, and that the regional-prior posterior means are
433 in many cases substantially different from those based on the national prior. Thus, the broad
434 issues about sensitivity of the ozone-mortality associations to alternative specifications of the
435 statistical model are not resolved by switching to 8-hour or 1-hour ozone.

436
437 The analyses of co-pollutants as confounders or within-city effect modifiers were also repeated
438 under 8-hour or 1-hour ozone, with results given in Tables 1 and 2 (no conversion factors were
439 used in computing those tables).

440

441 **Regional and Spatial Variability**

442

443 Earlier studies of the PM₁₀-mortality effect based on the NMMAPS data, such as those of
444 Dominici et al. (2002, 2003), emphasized the variability of the effect across regions, which might
445 be interpretable as due to different constituents of particulate matter. In the case of ozone, as far
446 as we know, the possibility of a regional effect was not discussed anywhere before the paper of
447 Bell and Dominici (2008). However there is in fact a very strong regional effect, as seen in Table
448 3, where we present posterior means and standard deviations for μ and the population-weighted
449 averages for the seven regions of the continental US, using all three ozone metrics, for both all-
450 year and summer data. Table 2 of Bell and Dominici (2008) shows a similar effect based on
451 posterior means of the hyperparameters, using only 24-hour ozone and all-year data. For 24-hour
452 ozone, the table shows a very clear statistically significant effect in the Industrial Midwest and
453 (especially) North East regions, and a smaller effect from the South East (which appears to be
454 primarily due to the results of Houston and Dallas/Fort Worth, which are part of the South East
455 in this classification). The remaining regions do not show a statistically significant effect. The
456 patterns based on 8-hour and 1-hour ozone are broadly similar.

457

458 A natural question is whether the differences among regions are statistically significant. For all
459 six analyses (three ozone metrics, all-year and summer-only), we performed a likelihood ratio
460 test of the null hypothesis that the seven region means are the same, using a standard chi-square
461 approximation to the distribution of the test statistic. The null hypothesis was rejected in every
462 case, with p-values from 0.017 down to 0.0015.

463
464 We can extend this calculation to the construction of spatial maps of the ozone-mortality
465 coefficient. The methodology, which is described in more detail in Appendix 3, is an extension
466 of the hierarchical model represented by equations (1) and (2) to allow for spatial autocorrelation
467 among the city-specific random effects. The technique is similar to kriging, which is a well-
468 known method for spatial interpolation (Cressie 1993, Stein 1999) but allowing for the standard
469 errors of the estimated random effects from the initial GLM part of the analysis. It is necessary to
470 choose a parametric form of spatial correlation function, for which the method was to fit several
471 possible models by the method of restricted maximum likelihood estimation (REML), choosing
472 the best model as the one that gave the largest restricted likelihood. By this approach, we
473 selected the Gaussian correlation function, $\rho = \exp(-(D/R)^2)$, where D is the distance (in
474 kilometers) between two locations and R is a spatial parameter known as the range. In this kind
475 of analysis, it is typically rather difficult to pin down R precisely and this is reflected in the wide
476 range of estimates over the six datasets, and their wide confidence intervals: see Table 4. Some
477 variants on the analysis included adding a nugget parameter to the variance, and accounting for
478 anisotropy by using a linear transformation of the coordinate space in defining D (in the
479 literature on spatial statistics this is called geometric anisotropy). Neither of these variants
480 resulted in statistically significant changes in the model.

481
482 Fig 6 shows maps constructed based on all three ozone metrics for all-year data, and for the 8-
483 hour metric using summer data. (The other two cases, 24-hour and 1-hour ozone and summer
484 data, were also drawn but are not shown here, since they are similar to the maps shown.)

485
486 In drawing these maps, it was considered desirable to use a common color scale. To ensure
487 comparability between different ozone metrics, the interpolated ozone-mortality coefficients for
488 24-hour and 1-hour ozone were converted to an equivalent scale to those for 8-hour ozone, using
489 the factors 0.646 and 1.31 derived previously. As noted earlier, we do not believe any single
490 conversion factor is adequate to compare different metrics, but these seem the most suitable
491 values for the present purpose.

492
493 These maps reinforce the existence of a strong spatial effect. In all four maps, the ozone-
494 mortality coefficient is highest in the region around New York, with other high spots around
495 Chicago, throughout the Industrial Midwest, and down to eastern Texas. Other parts of the
496 country, such as California, Florida, etc., seem to have very low ozone-mortality coefficients.

497
498 The regional analysis was also calculated for the within-city effect modifier results. We noted
499 earlier that temperature is a statistically significant effect modifier for the summer-only data. If
500 this effect is examined regionally, we find that it is actually significant only in the North East and
501 Industrial Midwest regions. For PM₁₀, also, there appears to be a regional effect, the evidence
502 that PM₁₀ is a significant effect modifier being strongest in the North East. In other calculations
503 involving PM₁₀ or SO₂ as an effect modifier, there did not appear to be a regional effect.

504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548

Between-City Effect Modifiers

Bell and Dominici (2008) sought to explain spatial variability among the ozone-mortality coefficients by constructing regression relationships involving numerous covariates defined at the city or community level. They called these covariates effect modifiers; we add the qualifier “between-city” to emphasize that this is a different analysis from the within-city effect modifiers considered earlier, though some of the variables are the same. In their study, they found that high unemployment, the proportion of Blacks or African Americans, use of public transportation, lower mean temperatures and low usage of central air conditioning were all associated with a higher ozone-mortality coefficient. In a similar though not identical context, Krewski et al. (2000) referred to “ecological covariates” in describing factors that could influence city-wide standardized mortality rates in a large prospective study using data from the American Cancer Society.

We have applied a similar analysis based on variables of a demographic or socio-economic nature, and various “environment” variables, as well as air-conditioning (AC) data. At the suggestion of a reviewer, latitude and longitude were also included as possible effect modifiers in this comparison. Results are summarized in Table 5. Of particular concern to us is the robustness of effect modifier analyses across different versions of the ozone-mortality relationship that correspond to the three ozone metrics and all-year and summer-only analyses. Therefore, we have ordered the candidate effect modifiers according to the largest (least significant) p-value across all six analyses. The first ten rows in Table 5 list all the variables that were significant at level 0.05 under all six analyses.

In processing the AC data, Bell and Dominici considered two measures: “central AC” and “any AC”. Bell and Dominici found a statistically significant effect for central AC (p-value 0.02) but not for any AC. We have preferred, instead, to treat the proportion of houses with window AC units as a separate variable. In fact, this turns out to be much more highly significant (and with a positive sign, i.e. increased use of window AC corresponds to an increased ozone-mortality coefficient). In all six analyses, the result for window AC is significant with a p-value <0.001.

Of the other variables that are statistically significant for all six ozone measures, four are related to changing residence: they show that communities where a high proportion of residents have moved since 1995 are associated with a lower ozone-mortality coefficient. The two variables “proportion drive to work” and “proportion public transport to work” are most likely complementary measures of air pollution exposure during commuting and this is reflected in the similar p-values and opposite signs of the effects. We also note that latitude and longitude are statistically significant effect modifiers in most analyses, implying a geographic effect (the ozone-mortality coefficient increases as we move further north or east), though not as strongly significant as the variables related to AC, change of residence or transportation. Finally we see two “environmental” variables that are significant: mean temperature and the mean level of SO₂. It should be noted that other air pollution related citywide variables, such as the mean level of ozone, PM₁₀ or PM_{2.5}, are not significant under this analysis.

549 Bell and Dominici (2008) gave particular attention to two variables, the proportion of blacks and
550 the proportion unemployed, reporting that each was significant with p-value about 0.03. For
551 comparison purposes, these variables are also included in Table 5. Our result for 24-hour ozone,
552 all-year analysis confirms the statistical significance of these two variables, but other analyses do
553 not corroborate this conclusion. Therefore, the relationship between the ozone-mortality
554 coefficient and these two ecological variables may well be spurious.

556 **Non-Linear Exposure-Response Relationships**

557
558 The analyses so far have all been based on a linear relationship between the logarithm of the
559 daily mortality rate and the level of ozone. However, non-linear relationships are of interest for a
560 variety of reasons. If there were a threshold below which there is no relationship between ozone
561 and mortality, this would be of obvious relevance in setting a standard. Even in the absence of a
562 threshold, if the slope of the relationship between log mortality rate and ozone were substantially
563 different over different ranges of ozone, this could affect calculated estimates of the expected
564 benefit associated with a reduction of ozone levels.

565
566 In this analysis, following Bell, we have used “0-1” ozone measure (average of daily ozone at
567 lags 0 and 1), though using any of the 24-hour, 8-hour and 1-hour metrics for computing a daily
568 ozone value. This is used in preference to a distributed lag model because it is not clear how to
569 implement the latter in the case of a nonlinear ozone-mortality association.

570
571 Bell et al. (2006) looked at nonlinear effects in a variety of ways. One approach was the “subset
572 approach”: subset the data by restricting to days below a cutoff, and recompute the ozone-
573 mortality coefficient. They repeated this calculation for several cutoffs down to 15 ppb, showing
574 a statistically significant effect at all cutoffs from 30 ppb and higher. It should be pointed out that
575 their analysis was still based on 24-hour ozone data which, as we have seen, does not translate
576 directly to an equivalent level for 8-hour ozone.

577
578 In Figure 3 of the same paper, they also computed a smooth non-linear exposure-response
579 relationship using splines. Their analysis is complicated, requiring fitting a multiparameter spline
580 function for each city, followed by a combination across cities. They do not provide full details
581 of their methodology and we have not succeeded in reproducing their figure. However, below we
582 propose an alternative “piecewise linear” approach which has a similar effect.

583
584 Let us first consider the subset approach. As noted earlier in this paper, we prefer to combine
585 data across cities using a population-weighted average rather than the posterior distribution of
586 the parameter μ to represent a national average. Bell’s calculation was repeated using this
587 measure, producing the figure at the top of Fig 7. This is very similar to Bell’s Fig 2, and
588 supports Bell’s conclusion that there is a statistically significant relationship for all cutoffs above
589 30 ppb. The result for 25 ppb shows that the 95% posterior predictive interval (PI) just covers 0,
590 which we are interpreting here as not statistically significant, though it is very close to being so.

591
592 The subset analysis shows the effect of ozone at very low levels. However, for policy purposes it
593 may be that ozone levels at a more moderate level, for example between 40 and 80 ppb, are more
594 relevant to risk assessments. Therefore, to make a contrast with the subset approach, we also

595 consider a “reversed subset” approach in which the analysis is restricted to ozone levels above a
596 cutoff, rather than below. We have considered cutoffs in the range 15-60 ppb, showing very little
597 change in the estimates within the range 15-40 ppb but then greatly increasing variability in the
598 estimates as the cutoff gets higher (Fig 7, bottom).
599

600 The piecewise linear approach assumes that the contribution to log mortality rate due to an ozone
601 level x is of the form
602

$$\begin{aligned} 603 \text{ (4) } f(x) &= \beta_1 x \text{ for } x \leq 40, \\ 604 &40 \beta_1 + \beta_2 (x - 40) \text{ for } 40 \leq x \leq 60, \\ 605 &40 \beta_1 + 20 \beta_2 + \beta_3 (x - 60) \text{ for } 60 \leq x \leq 80. \end{aligned}$$

606

607 In other words, the slopes are respectively β_1 , β_2 and β_3 in the ranges 0-40, 40-60 and 60-80 ppb
608 (in this approach, we do not consider ozone levels above 80 ppb). The choice of break points at
609 40, 60 and 80 ppb is to some extent arbitrary though these also seem natural choices, e.g. 40 ppb
610 corresponds to the background ozone level assumed for EPA’s 1997 ozone review and 80 ppb
611 was until recently the ozone standard; 60 is chosen simply as midway between 40 and 80. All
612 estimates based on (4) are for the 8-hour ozone metric.
613

614 The model (4) is estimated first for each city by a direct extension of the method used earlier for
615 linear response (all the other covariates, including meteorology and long-term trend, are kept the
616 same) and then the three slopes β_1 , β_2 , β_3 combined across cities using the *tlmise* software. For
617 each ozone level, we computed a posterior mean and a posterior standard deviation for the
618 population-weighted average response, and based on that computed a pointwise 95% prediction
619 interval. The results are shown in Table 6 and Fig 8, for both the all-year and summer-only
620 estimates.
621

622 Except for the fact that we are plotting a piecewise linear rather than a smooth curve, the
623 appearance of these plots is quite similar to Fig 3 of Bell et al. (2006). However, closer analysis
624 shows that the estimates of the three slopes are quite variable, especially β_3 . Fig 9 shows separate
625 estimates and 95% prediction intervals for β_1 , β_2 , β_3 , both nationally and regionally. There is still
626 substantial regional variation; there is no clear indication that β_3 is the largest of the three
627 coefficients (as one might expect if it were true that ozone toxicity is greater at higher
628 concentrations) and it has by far the largest posterior standard deviation of the three estimates.
629 The Industrial Midwest is the only region with a statistically significant effect in 60-80 ppb,
630 according to this analysis.
631

632 **Individual-city Analyses**

633

634 We select four cities to illustrate different types of results. For each city, the results of 38
635 analyses have been shown, including analyses that represent 24-hour, 8-hour and 1-hour ozone,
636 the NMMAPS and extended meteorology models, all-year and summer data, analyses both with
637 and without PM_{10} as a co-pollutant, and the piecewise-linear as well as linear concentration-
638 response curve. For each analysis, we present the posterior mean ozone-mortality coefficient and
639 the boundaries of a 95% prediction interval. In cases where the prediction interval includes 0, the
640 result should be interpreted as meaning that the ozone-mortality effect is not statistically

641 significant in that city. All results are based on the standard two-stage analysis specified by
642 equations (1) and (2) for the national prior, or (2) and (3) (with x_c taken as a regional indicator
643 variable) for the regional prior.

644
645 Los Angeles is the largest city in the study, with very high ozone levels, and if there were a
646 cause-and-effect ozone-mortality relationship, one would expect it to show up in a wide variety
647 of analyses. Despite this, no single analysis produces a clearly significant result, as shown in Fig
648 10.

649
650 Our previous results have shown New York and Chicago illustrating a strong ozone-mortality
651 relationship. It is therefore no surprise that several analyses show significant results for both
652 those cities (Figs 11 and 12). On the other hand, to take an example of a western city that also
653 has high ozone levels, for Denver (Fig 13), none of the statistical analyses shows a significant
654 relationship. Similar results, showing essentially no significant result for the posterior
655 distributions at the individual-city level, have been obtained for other large western cities (e.g.
656 Albuquerque, Phoenix, Salt Lake City, Seattle) and east coast cities in the south (e.g. Atlanta,
657 Miami). These results reinforce the fact that the ozone-mortality effect is concentrated within
658 certain regions of the US, and cities outside those regions show little if any effect.

659

660 **4. Discussion**

661

662 Multi-city time-series studies have come to be viewed as the gold standard of air pollution health
663 research. They are superior to single-city analyses, which typically exhibit too much statistical
664 variation from city to city for a clear conclusion to be drawn. This is apparent from Fig. 1(a), in
665 which the individual estimates range from -4.1 to +6.3 percent rise in mortality per 10 ppb rise in
666 24-hour ozone, with a median standard error of 1.4. Moreover, single-city results, including
667 which cities emerge with statistically significant effects, are highly sensitive to statistical
668 modeling assumptions.

669

670 Multi-city time-series studies are also widely considered superior to meta-analyses. The
671 distinction is that a multi-city time-series study uses raw data from each city in the study, which
672 it analyzes using uniform statistical methods, whereas a meta-analysis combines data from
673 previously published studies. There are two kinds of potential bias in this approach: publication
674 bias (arising from the tendency of researchers to publish only those results in which they found a
675 statistically significant effect) and model selection bias (selecting the statistical model to
676 maximize the estimated effect, separately for each city). These difficulties are apparent in the
677 three ozone meta-analyses that were published together in *Epidemiology* in 2005 (Bell et al.
678 (2005), Ito et al. (2005), Levy et al. (2005)). In particular, Bell et al. (2005) directly compared
679 NMMAPS multi-city results with a meta-analysis based on previously published results for the
680 NMMAPS cities, in some cases finding effects that were as large as three times greater in the
681 meta-analysis (allowing for differences in the ozone metric and units of measurement), which the
682 authors themselves attributed to the likely effects of publication bias.

683

684 Multi-city time-series studies were first used to study the mortality and morbidity effects of PM₁₀
685 (Dominici et al. (2000, 2002, 2003), Samet et al. (2000a, 2000b)). They have been extensively
686 used in the analysis of ozone and mortality – apart from the papers already cited, a smaller multi-

687 city study was included in Ito et al. (2005), an earlier version of the NMMAPS analysis in Huang
688 et al. (2005), and a similar European study in Gryparis et al. (2004). They were also used in a
689 widely cited study about PM_{2.5} and hospital admissions (Dominici et al. (2006)).
690

691 We believe multi-city time-series studies may also suffer from serious difficulties, as illustrated
692 by our re-analysis of the NMMAPS data on ozone and mortality. One issue is the sensitivity of
693 the results to statistical modeling assumptions, such as the functional form of meteorological
694 adjustment, the inclusion of co-pollutants as confounders and effect modifiers, the choice of
695 ozone metric and the use of a linear concentration-response model. In addition, the final result of
696 a multi-city time-series study is usually quoted in some form of national average. In many cases,
697 this is calculated through a Bayesian hierarchical modeling procedure, though virtually identical
698 results are obtained through a non-Bayesian restricted maximum likelihood (REML) procedure.
699 However, quoting a single value as a national average is misleading if there is substantial
700 heterogeneity. We have investigated these issues by calculating regional estimates, drawing
701 spatial maps, looking at the dependence of individual city results on demographic and
702 environmental covariates (between-city effect modifier analyses), and by looking in depth at
703 some of the posterior analysis results for four individual cities. In the following discussion, we
704 summarize our main conclusions under each of these headings.
705

706 In Fig. 1, we showed the raw (MLE) estimates of the regression coefficients, together with the
707 posterior estimates under two priors (national and regional), and posterior 95% intervals based
708 on the national prior. The results based on the national prior are the same as in Fig. 2 of Bell et
709 al. (2004). The fact that individual-city posterior means are sensitive to which of the two forms
710 of prior is chosen shows that the choice of prior cannot be ignored in evaluating this kind of
711 analysis. Given that we present substantial evidence later in the paper that the regional
712 differences are real, it might be more logical to prefer the regional prior.
713

714 In presenting the different results for national means, we have noted that the posterior mean of
715 the “national average” parameter μ (in (1)) is closer to the unweighted mean of the individual-
716 city estimates than it is to a mean weighted by population or (what is almost the same thing) by
717 the inverses of the individual-city variances. This seems counter-intuitive, since there is large
718 variability in the sizes of individual cities and it seems both logical and inevitable that the largest
719 cities should have a strongest influence on the overall average. However, the result may be
720 explained by noting that, conditionally on τ^2 , the estimate of μ is the weighted average with
721 weights proportional to $1/(s_i^2 + \tau^2)$, in the notation of (1) and (2). If τ^2 is large relative to the
722 individual s_i^2 , the Bayes estimate will be approximately the unweighted mean. This might be a
723 reasonable thing to do if we were selecting cities at random from a very large number of cities
724 (the implicit if unstated assumption in (1)), but that is not valid here because the selection of
725 cities is very clearly not random.
726

727 This discussion highlights the difficulty of interpreting a "national" ozone effect estimate in
728 cases where there is substantial inter-city heterogeneity among estimated ozone-mortality effect
729 coefficients. We caution, again, that any national summary, even a population-weighted average,
730 will necessarily conceal the still-unexplained heterogeneities. Further, we believe that the
731 heterogeneity and sensitivity of ozone effect estimates to a variety of covariates leaves open the
732 issue of whether or not ozone is causally related to mortality. Consequently, the question arises

733 whether any particular ozone-mortality effect estimate can reliably be used to predict mortality
734 reductions that would ensue from specific ozone reductions.

735
736 For the purpose of computing summary results where these are needed, we propose a
737 population-weighted average, but without any implication that this represents a causal
738 association. Some cities, such as Los Angeles and Denver, do not show any statistically
739 significant effect of ozone on mortality. Conversely, the use of a national coefficient for a city
740 like New York or Chicago may underestimate the effect of ozone as compared to use of a city
741 specific coefficient.

742
743 We have examined the sensitivity of the individual-city estimates to alternative specifications of
744 meteorology and the inclusion of co-pollutants. The effect of meteorology has already been
745 examined extensively in previous studies; for example, Ito et al. (2005) used four meteorological
746 adjustments and found that the overall ozone-mortality coefficient varied by a factor of 2 over
747 the models studied; however, the smallest of these was for the model that essentially replicates
748 the meteorology model of Bell et al. (2004). Bell et al. (2004) and Schwartz (2005) argued that
749 the ozone-mortality effect is not changed by the exclusion of hot days. As a comparison, we
750 reran some of our models deleting the hottest 1% of days within each city, with no change in the
751 results. However, one caution we would note is that it is necessary to exclude days with very
752 unusual mortality levels, such as occurred during the Chicago heatwave of July 1995, because
753 such values are highly influential on the regression estimates. In fact Bell et al. (2004) already
754 applied such a correction, as became clear on detailed examination of their computer code,
755 though this was not mentioned in their paper.

756
757 We have considered alternative specifications of meteorology, including an “extended
758 meteorology” model that has a distributed lag form similar to that used for ozone. In some cases,
759 this results in a reduction of the estimated ozone effect, implying a partial temperature
760 confounding that was not accounted for in earlier analyses.

761
762 We looked at the possibility of a confounding effect due to PM_{10} , concluding that the overall
763 ozone-mortality coefficient is reduced between 22% and 33% when PM_{10} is included in the
764 model, despite the fact that in many cities, PM_{10} is only sampled once every six days. It is
765 difficult to make precise statements concerning the statistical significance of this result, given the
766 large posterior standard deviations, but the result has been observed persistently across numerous
767 analyses. It therefore seems doubtful to us that it is spurious, though it is clear that our
768 interpretation of this result is different from that of Bell et al. (2004, 2007). On the other hand,
769 Franklin and Schwartz (2008) found an even larger confounding effect based on a different
770 dataset using sulfate particles.

771
772 Other authors who have considered the confounding effect of PM_{10} as a co-pollutant include
773 Huang et al. (2005), who analyzed a subset of the NMMAPS database; when both ozone and
774 PM_{10} were included in the model at lag 2, there was a decrement in the ozone-mortality effect
775 similar to that reported here, but at other lags, the effect was smaller or non-existent. Schwartz
776 (2005) also looked at PM_{10} confounding in his case-crossover analysis, concluding there was no
777 effect. The meta-analyses (Bell et al. (2005), Ito et al. (2005)) generally reported that inclusion of
778 PM_{10} did not have a great overall effect on the results of the meta-analyses, but Bell et al. (2005)

779 reported that where PM₁₀ was included, its effect was almost always to lower the estimated
780 ozone-mortality effect.

781
782 We also looked at temperature, PM₁₀ and SO₂ as within-city effect modifiers. For temperature,
783 we found that in summer, the ozone-mortality coefficient is higher when temperature is above
784 the median than when it is below. However, a more detailed regional analysis (not tabulated)
785 showed that this effect is statistically significant only in the North East and Industrial Midwest
786 regions. For PM₁₀, there is a statistically significant effect modifier relationship for all three
787 ozone metrics and both all-year and summer-only data. Other calculations have shown that there
788 is also some evidence of a regional effect here, strongest in the North East. The evidence for SO₂
789 being an effect modifier is less clear cut, since the relationship was significant only for 24-hour
790 ozone and all-year data. Taken together, these results imply that the influence of meteorology
791 and co-pollutants on the ozone-mortality association cannot be ignored.

792
793 Our results on temperature as an effect modifier are consistent with those of Ren et al. (2008),
794 who also used the NMMAPS data but treated the ozone, meteorology and long-term trend effects
795 in quite different ways from the analysis here. Ren et al. confined their study to summer data and
796 to two regions: a combination of our Industrial Midwest and North East regions (that they called
797 “northeast”), and the South East region (the same as ours). They stratified temperature into three
798 ranges: below the first quartile, between the first and third quartiles, and above the third quartile.
799 They then computed combined estimates of the ozone-mortality coefficient in each of the three
800 ranges. For their northeast region, they claimed the ozone-mortality effect is by far the largest in
801 the highest range of temperatures (above the third quartile). For the South East, they essentially
802 found no difference in the ozone-mortality coefficient among the three ranges of temperature.
803 These results are consistent with ours, which show temperature is a statistically significant effect
804 modifier for summer data in the North East and Industrial Midwest, but not in the other five
805 regions (including the South East).

806
807 On the more general effect of changing ozone metrics, we found that single-city ozone-mortality
808 coefficients computed under the 24-hour, 8-hour and 1-hour ozone metrics are quite highly
809 correlated (0.7-0.8) in either the raw or posterior mean estimates, but the association is certainly
810 not perfect (Fig 3). Previous authors have used conversion factors to compare results in different
811 metrics; for example, Bell et al. (2005) used the ratios 20:15:8 for the means of 1-hour, 8-hour
812 and 24-hour ozone. However the actual ratios of these means for different cities vary
813 substantially.

814
815 We found significant regional and spatial variability in the ozone-mortality coefficient, using
816 seven regions for the continental US that had earlier been defined for PM₁₀ studies (Samet et al.
817 2000b, Dominici et al., 2002, 2003). The city of Honolulu was excluded from all analyses that
818 involved regional variables. Similar regional results for 24-hour ozone, all-year analyses, were
819 reported by Bell and Dominici (2008); as far as we are aware, this was the first explicit
820 discussion of regional variation in the ozone-mortality coefficient, though such an effect is
821 apparent from close scrutiny of Fig 2 of Bell et al. (2004). Our results showed that this is a
822 persistent effect across all three ozone metrics and both all-year and summer-only data; later (Fig
823 9) we also showed that a regional effect also exists in the piecewise linear analysis. The overall
824 pattern of these results is that the North East and Industrial Midwest regions have by far the

825 largest ozone-mortality associations; there is also a smaller but in many instances statistically
826 significant association for the South-East (that we believe is largely due to the cities in Texas,
827 especially Houston and Dallas/Fort Worth). The piecewise linear analysis (Fig 9) implies that the
828 result for the South East may be entirely due to the ozone-mortality association at low ozone
829 levels, below 40 ppb. The fact that clear differences exist for the individual-city results in
830 different regions (Figs. 10-13) also reinforces the conclusion that the inter-region differences are
831 both statistically and practically significant.

832
833 Spatial patterns have previously been shown for PM₁₀-mortality effects (Dominici et al. 2002).
834 An online figure posted with Bell and Dominici (2008) depicted the spatial variability of the
835 estimated effects, though they claimed there was no evidence of a spatial pattern. They did not,
836 however, make explicit use of spatial correlation and kriging techniques. Our analyses leading to
837 Fig 6 brings out the spatial pattern much more sharply than previous analyses. It naturally raises
838 the question of what could be the cause of such strong spatial effects.

839
840 The main conclusion of Bell and Dominici (2008) was that differences among cities could be
841 explained in terms of (between-city) effect modifiers. Effect modifiers can be classified broadly
842 into three types. Demographic variables (e.g. racial or socioeconomic) might be indicative of
843 vulnerable subgroups of the population. The other two broad classes are those associated with
844 exposure and with possible co-pollutants. The presence of SO₂ among the significant effect
845 modifiers suggests a co-pollutant effect, though we did not find such a strong effect for
846 SO₂ when treated as a within-city effect modifier. In contrast, for PM₁₀ the effect is the other way
847 round: significant as a within-city modifier, but not as a between-city modifier. For temperature,
848 we have observed a positive association (in summer) when treated as a within-city modifier, but
849 there is a negative association when temperature is treated as a between-city effect modifier,
850 which suggests that temperature as a between-city effect modifier could be a proxy for
851 something else (such as AC use). Alternatively, it is possible that temperature (like latitude and
852 longitude, which were also statistically significant effect modifiers) is just a proxy for a broad
853 spatial pattern which, as we saw earlier, leads to significantly higher ozone-mortality coefficients
854 in the North East and Industrial Midwest regions.

855
856 The two variables related to transport presumably reflect exposure - those taking public transport
857 spend more time in the open air and are therefore more exposed to ozone. Air conditioning has a
858 similar interpretation - people who live in air-conditioned houses are less exposed to ambient
859 ozone in summer. However, we found a much stronger effect by separating the air-conditioning
860 variable into one based on central AC and another based on the use of window AC units. For
861 central AC, we found a statistically significant negative association, consistent with Bell and
862 Dominici for the 24-hour ozone all-year analysis but also extending to the other five analyses in
863 Table 2. But there is a stronger *positive* association between the proportion of houses using
864 window AC in a city and the ozone-mortality coefficient in that city. It is known that homes with
865 central AC tend to have lower air exchange rates and hence also lower indoor to outdoor ozone
866 ratios. Since outdoor ozone is the main source of indoor ozone, this would imply that actual
867 ozone exposures are lower in houses with central AC (Weschler 2000, 2006). It therefore seems
868 plausible that houses with window AC have a higher indoor to outdoor ozone ratio and therefore
869 the residents experience higher ozone exposure. Further evidence to support this includes the fact
870 that window AC units contain a 'vent' option which allows addition of outdoor 'make-up' air to

871 recirculated conditioned air, actively increasing room/home air exchange rates although at the
872 expense of reduced cooling and energy efficiency; central AC units usually lack such options and
873 are often installed in relatively tight, well-insulated homes with low air exchange. Also, open
874 windows are found more common (and likely air exchange rates higher) in houses with window
875 AC or no AC than homes with central AC (Johnson and Long (2005)). Other possible
876 confounding issues may arise from indoor temperature and dewpoint being different between
877 houses with and without AC.

878

879 The variables associated with residential mobility could be demographic or exposure-related. In
880 their re-analysis of the American Cancer Society cohort study, Krewski et al. (2000) noted that
881 population change is correlated with standardized mortality rates, and suggested as one possible
882 explanation that healthy people are more likely to move. However it is also possible that this
883 reflects exposure, older houses being less well insulated and perhaps also less likely to have
884 central air conditioning. All four “residential mobility” variables in Table 5 are positively
885 correlated with central AC usage and negatively correlated with window AC use.

886

887 The fact that several variables related to ozone exposure could be significant effect modifiers is
888 consistent with results of Koutrakis et al. (2005) and Sarnat et al. (2006), that demonstrated a low
889 correlation between personal and ambient levels of ozone, in contrast to PM₁₀ or PM_{2.5} where the
890 correlations are much higher. The impact of these findings for regulations is unclear, but it seems
891 to us that there is a need to recognize the different degree of ozone association in cities such as
892 New York and Chicago compared with many others where the associations are much smaller. It
893 is also possible that the fact that ozone-mortality effects have been generally reported higher in
894 European cities (Gryparis et al. (2004)) compared with U.S. may also be associated with
895 variables such as public transport, or with much lower usage of AC, though there are other
896 difficulties in comparing the studies, such as the use of different ozone metrics and statistical
897 methods, and different policies in the US and Europe regarding the siting of ozone monitors.

898

899 Our conclusions from this part of the analysis are as follows. First, the presence of several
900 statistically significant effect modifiers confirms that the variation of the ozone-mortality
901 coefficient over cities is not random, and therefore throws further into question the validity of
902 computing national risk estimates. However, the analysis does not clearly identify any single
903 effect modifier as the one most likely to have a cause and effect interpretation. Variables that are
904 indicators of race or socioeconomic status, as highlighted by Bell and Dominici (2008), seem to
905 us to be spurious. The other variables seem to suggest either a co-pollutant effect, or variations in
906 personal exposure to ozone. In either case, the true causal mechanism that leads to an observed
907 ozone-mortality association is unresolved.

908

909 The final part of our discussion concerns the possibility of a nonlinear exposure-response
910 association, where “response” here is generally interpreted as the logarithm of mortality rate at a
911 given ozone level. Previous discussion of this issue has often been focused on the apparent
912 nonexistence of an ozone threshold (e.g. NRC (2008)), but it is also important to consider the
913 broader implications of a nonlinear association, such as its impact on risk assessments that try to
914 evaluate the benefits of ozone reductions across certain ranges. Bell et al. (2006) presented a
915 “subset analysis” that showed the effect of ozone persists to very low levels. We have essentially
916 replicated their results, but for comparison purposes, we also show a “reverse subset” analysis

917 that confirms the increasing uncertainty of the ozone-mortality association as we restrict to
918 moderate or high levels of ozone.

919
920 The finding of a statistically significant effect at very low ozone levels is mysterious, since direct
921 studies of human lung function response to ozone do not show any effect at 40 ppb (Adams
922 (2002, 2006)). Other authors have questioned the biological plausibility of associations at very
923 low ozone levels (Vedal et al. (2003)). The biologic plausibility argument is delicate in that it
924 refers to human personal exposure as opposed to an ozone metric derived from monitored
925 ambient concentrations. A statistically significant linear response at low ambient ozone might be
926 less surprising if actual personal exposures were biologically meaningful and highly correlated
927 with ambient ozone at low ambient concentrations. However, at this time, the concern of Vedal
928 et al. is relevant and removing the objection of biologic implausibility requires demonstrating
929 that (a) actual personal exposures at low levels are biologically meaningful and (b) substantial
930 personal exposure-ambient correlation exists in the low-ambient range. We believe it is
931 reasonable to hypothesize that the variability across cities of the ozone-mortality coefficient
932 estimates is linked to variability across cities in the relationship between personal exposures and
933 the monitored ozone metric.

934
935 Moreover, although there remains extensive spatial variability (in the range 15-50 ppb) in
936 estimates of policy-relevant background ozone level (Fiore et al. (2002, 2003), Oltmans et al.
937 (2008)), it seems likely that 24-hour ozone levels below 30 ppb or 8-hour ozone levels below 40
938 ppb would be below the level plausibly affected by any currently proposed ozone regulation.

939
940 Plots of total relative risk (compared with zero ozone) across different ozone ranges appear to
941 show a steady increase in risk with ozone level, as in our Fig. 8 or Bell's (2006) Fig. 3. However
942 this could be misleading because it appears that most of the contribution to the overall risk is
943 from relatively low-ozone days, many of them below background level. The ozone range most
944 likely to be affected by regulation is that between 60 and 80 ppb, but here the effect is much
945 more uncertain (Fig 7), and the picture is clouded still more by the regional variation.

946 **5. Summary conclusions**

947
948
949 The basis for the national effect estimates published by Bell and others is questionable in the face
950 of clear evidence that the effect is not homogeneous. We propose the population-weighted
951 average as a more meaningful summary of a heterogeneous effect, but such a summary is
952 specific to the group of cities considered: there is no implication or even expectation that the
953 same numerical result would apply for a different population base, even if the methods of data
954 collection and analysis were essentially the same. Also, the estimates do not translate evenly
955 from one ozone metric to another.

956
957 There is clear evidence of a PM₁₀ co-pollutant effect that has been understated or misinterpreted
958 in previous publications. Given that in many cities PM₁₀ is only available one day in six, and is
959 probably not the most effective measure of a particulate matter effect, the true influence of
960 co-pollutants could be much greater than this. We also find that temperature, particulate matter
961 and sulfur dioxide all act as (within-city) effect modifiers, though the temperature effect is

962 confined to the North East and Industrial Midwest regions in summer, and the effect for sulfur
963 dioxide may be spurious given that it only applies to the 24-hour ozone metric.

964
965 The nonlinear analysis shows that much of the evidence for an ozone-mortality relationship in
966 fact comes from the low-ozone days, but human studies do not support an ozone effect at such
967 low ozone levels. It is possible that the appearance of an association at low ozone levels may be
968 due to the effect of co-pollutants, or an artifact caused by differences between personal and
969 ambient exposure. If we try to estimate the ozone-mortality effect restricted to higher levels of
970 ozone that might be directly affected by new regulations, the uncertainty is much greater and
971 there is still evidence of regional variability.

972
973 The regional estimates and single-city analyses demonstrate that the evidence for any
974 ozone-mortality effect is weak away from a central band across the country that stretches from
975 Houston to New York.

976
977 A number of between-city effect modifiers have been identified that may explain some of the
978 spatial variability. No single effect modifier is primarily responsible; one that turned out highly
979 significant was the proportion of window AC units, and another was residential mobility.
980 However the mechanism by which such factors influence ozone epidemiology is still very
981 unclear. We do believe, however, that this kind of effect modifier has a stronger influence than
982 those based on race or socioeconomic factors.

983
984 There are other methodological issues that have not been discussed in this paper, but that could
985 affect the results. For example, the long-term trend part of the model exactly follows Bell et al.
986 (2004), but it is known from studies of other air pollution effects (Peng et al. (2006) in the case
987 of PM₁₀, Ostro et al. (2006) in the case of PM_{2.5}) that model results may be sensitive to the
988 degrees of freedom and other aspects of the specification of this component. We have not
989 considered specific-cause mortality effects, such as cardiovascular or COPD mortality. It also
990 remains to explore many alternative aspects of co-pollutant effects, including nonlinear
991 specification of the ozone-mortality relationship in the presence of an effect modifier.

992
993 In summary, it is our view that estimates of the association between ozone and mortality, based
994 on time-series epidemiologic analyses of daily data from multiple cities, reveal important still-
995 unexplained inconsistencies and show sensitivity to modeling choices and data selection. These
996 inconsistencies and sensitivities contribute to serious uncertainties when epidemiological results
997 are used to discern the nature and magnitude of possible ozone-mortality relationships or are
998 applied to risk assessment.

1000 **Appendix 1: Derivation of REML estimator.**

1001
1002 Equations (1) and (2) may be combined to imply that $t_c \sim N[\mu, s_c^2 + \tau^2]$ independently for each city
1003 c . The weighted least squares estimator of μ is $M = [\sum t_c / (s_c^2 + \tau^2)] / [\sum 1 / (s_c^2 + \tau^2)]$. If we let T
1004 denote the vector of all the t_c , and if we generalize to the case $T \sim N[X\beta, H]$ for some covariate
1005 matrix X with n rows and $q < n$ columns and of rank q , β a vector of regression parameters of
1006 dimension q , and H a covariance matrix with n rows and n columns, then by equation (3) of
1007 Harville (1974), the logarithm of the restricted likelihood is given by

1008 (5) $-\frac{1}{2}(n-q)\log(2\pi)+\frac{1}{2}\log \det(X^T X) -\frac{1}{2}\log \det(X^T H^{-1}X) -\frac{1}{2}\log \det(H)-\frac{1}{2}R$
1009 where R is the generalized residual sum of squares, $R=(T-Xb)^T H^{-1}(T-Xb)$ and $b=$
1010 $(X^T H^{-1}X)^{-1}X^T H^{-1}T$ is the generalized least squares estimator of β . For the analysis where there is
1011 no covariate, we specialize (5) to the case when $q=1$, X is a column vector of ones, β is the scalar
1012 parameter μ , H is a diagonal matrix whose c 'th diagonal entry is $s_c^2+\tau^2$. Then R is the weighted
1013 sum of squares of residuals $\sum (t_c-M)^2/(s_c^2+\tau^2)$ and if we ignore the first two terms of (5) (which
1014 do not depend on μ and τ^2) then maximizing (5) is equivalent to minimizing
1015 (6) $\log[\sum 1/(s_c^2+\tau^2)] + \sum \log[1/(s_c^2+\tau^2)]+R$
1016 which is a function of τ^2 alone; in other words, we choose τ^2 to minimize (6). This is then called
1017 the REML estimator of τ^2 . Conditionally on τ^2 , we then estimate μ by M as given previously,
1018 with an estimated variance of $1/[\sum 1/(s_c^2+\tau^2)]$. The square root of this then gives the standard
1019 error of μ .

1020
1021 For the effect modifier analysis across cities, we write the regression model in the form
1022 $E(t_c)=\beta_0+\beta_1x_c$ where x_c is the effect modifier in city c ; estimates b_0, b_1 of the regression
1023 coefficients β_0, β_1 are obtained by weighted least squares with weights $1/(s_c^2+\tau^2)$, and we
1024 calculate $R=\sum (t_c-b_0-b_1x_c)^2/(s_c^2+\tau^2)$. We then select τ^2 to maximize (5), which in this instance is
1025 equivalent to minimizing $\log[\sum(1/(s_c^2+\tau^2)) \sum(x_c^2/(s_c^2+\tau^2))-(\sum(x_c/(s_c^2+\tau^2)))^2] + \sum \log(s_c^2+\tau^2)+R$.

1026
1027

1028 **Appendix 2: Estimation of Population-Weighted Means**

1029
1030 We assume the standard hierarchical model defined by equations (1) and (2). Suppose we want
1031 to estimate $\sum w_c\theta_c$ for given weights w_c – in the context of this paper, we are particularly
1032 interested in the choice when w_c is proportional to the population of city c . We can formulate this
1033 as a problem in universal kriging, see e.g. Cressie (1993) or Stein (1999). Suppose Y is a vector
1034 of observations with mean $X\beta$ and covariance matrix V (here $X\beta$ refers to some matrix of
1035 covariates X and an unknown linear regression coefficient β). Suppose Y_0 is some as yet
1036 unobserved scalar quantity with mean $x_0^T\beta$ and variance v_0 (x_0, v_0 known, β unknown). Also
1037 assume the vector of covariances between Y and Y_0 is w_0 . Then the optimal linear predictor of
1038 Y_0 is of the form $\lambda^T Y$ where $\lambda=V^{-1}w_0+V^{-1}X(X^T V^{-1}X)^{-1}(x_0-X^T V^{-1}w_0)$, and the corresponding mean
1039 squared prediction error is $v_0-w_0^T V^{-1}w_0+(x_0-X^T V^{-1}w_0)^T(X^T V^{-1}X)^{-1}(x_0-X^T V^{-1}w_0)$. In the present case,
1040 we identify Y with the vector of city-specific regression coefficients from the first stage of the
1041 analysis, whose means are all μ (i.e. X is a column vector of ones, $\beta=\mu$) and whose covariance
1042 matrix V is diagonal with diagonal entries, $s_1^2+\tau^2, s_2^2+\tau^2, \dots$ where s_1^2, s_2^2, \dots are the individual
1043 standard errors and τ^2 is the inter-city variance. We identify Y_0 with $\sum w_c\theta_c$, x_0 with w (i.e. the
1044 vector of weights w_c) and v_0 with $\tau^2\sum w_c^2$. Routine algebraic rearrangement then shows the
1045 optimal predictor and its mean squared prediction error are of the form given.

1046

1047 **Appendix 3: Spatial Analysis**

1048

1049 In this appendix we briefly outline the method for computing Fig 6.

1050

1051 The approach is an extension of that used in Appendix 1 to allow for spatial dependence in the
1052 prior distribution for θ_c . We extend (2) and (3) to

1053 (7) $T | \Theta \sim N[\Theta, W]$,

1054 (8) $\Theta \sim N[X\beta, H]$
1055 where N represents the multivariate normal distribution, T is the vector of t_c , Θ is the vector of
1056 θ_c , $X\beta$ is the linear mean and W and H are covariance matrices. We assume W is known –
1057 consistent with (2), W is the diagonal matrix with s_c^2 in the c 'th diagonal position – but H is
1058 taken to be a parametric spatial covariance matrix – for the examples given here, this was
1059 assumed to be of the “Gaussian” covariance form with no nugget, after trying several
1060 alternatives, see e.g. Cressie (1993) or Stein (1999) for an introduction to spatial covariances and
1061 their application in spatial prediction.

1062
1063 Combining (7) and (8) into a single equation, we have

1064 $T \sim N[X\beta, H+W],$

1065 from which β and the parameters of H may be estimated by the REML method, using (5) again
1066 with H replaced by $H+W$. Conditional distributions for the components of Θ given T , both at the
1067 actual city sites and elsewhere, are then calculated by the procedure known as universal kriging,
1068 which is described in detail in books on spatial statistics, including Cressie (1993) or Stein
1069 (1999).

1071 Acknowledgements

1072
1073 We are grateful for the advice of a number of individuals who contributed comments and
1074 suggestions, in particular, Dr. Allen S. Lefohn of A.S.L. & Associates, Dr. Will Ollison of the
1075 American Petroleum Institute, Dr. Paul Voss of the University of North Carolina's Odum
1076 Institute for Research in Social Science and Dr. Charles Weschler of the University of Medicine
1077 and Dentistry of New Jersey. We also thank two reviewers for their careful reading of the
1078 manuscript. The study would not have been possible without access to the NMMAPS database,
1079 and we thank the NMMAPS team for making the data publicly available.

1081 Declaration of Interest

1082
1083 The work was supported financially by the American Petroleum Institute. The authors have sole
1084 responsibility for the writing and content of the paper. The views expressed herein do not
1085 necessarily represent the policy position of the American Petroleum Institute.

1088 References

1089
1090 Adams, WC (2002), Comparison of chamber and face-mask 6.6-hour exposures to ozone
1091 on pulmonary function and symptoms responses. *Inhalation Toxicology* **14**, 745-764.

1092
1093 Adams, WC (2006), Comparison of chamber 6.6-hour exposures to 0.04-0.08 PPM
1094 ozone via square-wave and triangular profiles on pulmonary responses.
1095 *Inhalation Toxicology* **18**, 127-136.

1096
1097 AHS (2006), American Housing Survey for the United States: 2005. U.S. Department of
1098 Housing and Urban Development, Office of Policy Development and Research; U.S. Department
1099 of Commerce, Economics and Statistics Administration, U.S. Census Bureau. Available along

1100 with earlier reports and raw data files from <http://www.huduser.org/datasets/ahs.html> or
1101 www.census.gov/prod/www/abs/cons-hou.html#house
1102

1103 Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F (2004) Ozone and Short-term
1104 Mortality in 95 US Urban Communities, 1987-2000, *Journal of the American Medical*
1105 *Association*, **292**(19):2372-2378.
1106

1107 Bell ML, Dominici F, Samet JM (2005) A meta-analysis of time-series studies of ozone and
1108 mortality with comparison to the national morbidity, mortality, and air pollution study.
1109 *Epidemiology* **16**, 436-445.
1110

1111 Bell ML, Peng RD, Dominici F (2006) The exposure-response curve for ozone and risk of
1112 mortality and the adequacy of current ozone regulations. *Environmental Health Perspectives*
1113 **114**, 532-536.
1114

1115 Bell, ML, Kim, JY and Dominici, F (2007), Potential confounding of particulate matter on the
1116 short-term association between ozone and mortality in multisite time-series studies.
1117 *Environmental Health Perspectives* **115**, 1591-1595.
1118

1119 Bell, M. and Dominici, F. (2008), Effect modification by community characteristics on the short-
1120 term effects of ozone exposure and mortality in 98 U.S. communities. *American Journal of*
1121 *Epidemiology* **167**, 986-997.
1122

1123 Cressie, N (1993), *Statistics for Spatial Data*. Second edition, John Wiley, New York.
1124

1125 Curriero, FC, Heiner, KS, Samet, JM, Zeger, SL, Strug, L and Patz, JA (2002), Temperature and
1126 mortality in 11 cities of the eastern United States. *American Journal of Epidemiology* **155**, 80-87.
1127

1128 Dominici F, Samet JM, Zeger SL (2000), Combining evidence on air pollution and daily
1129 mortality from the twenty largest US cities: A hierarchical modeling strategy (with discussion).
1130 *Journal of the Royal Statistical Society A*, **163**, 263-302.
1131

1132 Dominici F, Daniels M, Segel SL and Samet JM (2002), Air pollution and mortality: estimating
1133 regional and national dose-response relationships. *Journal of the American Statistical*
1134 *Association* **97**, 100-111.
1135

1136 Dominici F, McDermott A, Daniels M, Zeger SL and Samet JM (2003), Mortality among
1137 residents of 90 cities. In *Revised Analyses of the National Morbidity, Mortality and Air Pollution*
1138 *Study, Part II*. Health Effects Institute, Cambridge, MA, pp. 9--24.
1139

1140 Dominici F, Peng R, Bell M, Pham L, McDermott A, Zeger, S and Samet, J (2006), Fine
1141 particles air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*
1142 **295** 1127-1135.
1143

1144 Everson PJ and Morris, CN (2000), Inference for multivariate normal hierarchical models.
1145 *Journal of the Royal Statistical Society B*, **62**, 399-412.

1146
1147 Fiore AM, Jacob DJ, Bey I, Yantosca RM, Field BD, Fusco AC and Wilkinson JG (2002),
1148 Background ozone over the United States in summer: Origin, trend, and contribution to pollution
1149 episodes. *Journal of Geophysical Research* **107**, D15, 4275, 10.1029/2001JD000982.
1150
1151 Fiore A, Jacob DJ, Liu H, Yantosca RM, Fairlie TD and Li Q (2003), Variability in surface
1152 ozone background over the United States: Implications for air quality policy. *Journal of*
1153 *Geophysical Research* **108**, D24, 4787, doi:10.1029/2003JD003855.
1154
1155 Franklin M, Schwartz J (2008), The impact of secondary particles on the association between
1156 ambient ozone and mortality. *Environmental Health Perspectives* **116**, 453-458.
1157
1158 Gryparis A, Forsberg B, Katsouyanni K, Analitis A, Touloumi G, Schwartz J, Samoli E, Medina
1159 S, Anderson H R, Niciu EM, Wichmann H-E, Kriz B, Kosnik M, Skorkovsky J, Vonk JM,
1160 Dortbudak Z (2004) Acute effects of ozone on mortality from the air pollution and health: A
1161 European approach project. *Am. J. Respir. And Crit. Care Med.* **170**, 1080-1087.
1162
1163 Harville, DA (1974), Bayesian inference for variance components using only error contrasts.
1164 *Biometrika* **61**, 383-385.
1165
1166 Huang Y, Dominici F, Bell ML (2005) Bayesian hierarchical distributed lag models for summer
1167 ozone exposure and cardio-respiratory mortality. *Environmetrics* **16**, 547-562.
1168
1169 Ito K, De Leon SF, Lippmann M (2005) Associations between ozone and daily mortality,
1170 analysis and meta-analysis. *Epidemiology* **16**, 446-457.
1171
1172 Johnson T and Long T (2005), Determining the frequency of open windows in residences: a pilot
1173 study in Durham, North Carolina during varying temperature conditions. *Journal of Exposure*
1174 *Assessment and Environmental Epidemiology* **15**, 329-349.
1175
1176 Koutrakis P, Suh HH, Sarnat JA, Brown KW, Coull BA, Schwartz J (2005) Characterization of
1177 Sensitive Subpopulations Living in Baltimore and Boston, Health Effects Institute Research
1178 Report, 131:1-65.
1179
1180 Krewski D, Burnett RT, Goldberg MS, Hoover K, Siemiatycki J, Jerrett M, Abrahamowicz M
1181 and White, WH (2000), *Reanalysis of the Harvard Six Cities Study and the American Cancer*
1182 *Society Study of Particulate Air Pollution and Mortality*. Special Report of the Institute's
1183 Particulate Epidemiology Reanalysis Project. Health Effects Institute, Cambridge, MA.
1184
1185 Levy JL, Chemerynski SM, Sarnat JA (2005) Ozone exposure and mortality, an empiric Bayes
1186 metaregression analysis. *Epidemiology* **16**, 458-488.
1187
1188 NRC (2008), Estimating Mortality Risk Reduction and Economic Benefits from Controlling
1189 Ozone Air Pollution. Committee on Estimating Mortality Risk Reduction Benefits from
1190 Decreasing Tropospheric Ozone Exposure. National Research Council, Washington, D.C.
1191

1192 Oltmans SJ, Lefohn AS, Harris JM and Shadwick D (2008), Background ozone levels of air
1193 entering the west coast of the U.S. and assessment of longer-term changes. *Atmospheric*
1194 *Environment* **42**, 6020-6038.

1195
1196 Ostro B, Broadwin R, Green S, Feng W-Y and Lipsett M (2006), Fine particulate air pollution
1197 and mortality in nine California counties: Results from CALFINE. *Environmental Health*
1198 *Perspectives* **114** No. 1, 29-33.

1199
1200 Peng RD, Dominici F and Louis TA (2006), Model choice in time series studies of air pollution
1201 and mortality. *Journal of the Royal Statistical Society, Series A*, **169**, 179-203.

1202
1203 R Development Core Team (2007), *R: A Language and Environment for Statistical Computing*.
1204 R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. [http://www.R-](http://www.R-project.org)
1205 [project.org](http://www.R-project.org).

1206
1207 Ren C, Williams GM, Mengersen K, Morawska L, Tong S (2008), Does temperature modify
1208 short-term effects of ozone on total mortality in 60 large eastern US communities? – An
1209 assessment using the NMMAPS data. *Environment International* **34**, 451-458.

1210
1211 Samet JM, Dominici F, Curreiro FC, Coursac I, Zeger SL (2000a), Fine particulate air pollution
1212 and mortality in 20 US cities 1987-1994 (with discussion). *New England Journal of Medicine*
1213 **343**, 1742-1757.

1214
1215 Samet JM, Zeger SL, Dominici F, Curriero F, Coursac I, Dockery D, Schwartz J and Zanobetti A
1216 (2000b), *The National Morbidity, Mortality and Air Pollution Study (HEI Project No. 96-7):*
1217 *Morbidity and Mortality from Air Pollution in the United States*. Cambridge, MA: Health Effects
1218 Institute.

1219
1220 Sarnat, SE, Coull BA, Schwartz J, Gold DR and Suh HH (2006) Factors Affecting the
1221 Association between Ambient Concentrations and Personal Exposures to Particles and Gases.
1222 *Environmental Health Perspectives*, 114(5):649-654.

1223
1224 Schwartz J (2005), How sensitive is the association between ozone and daily deaths to control
1225 for temperature? *Am. J. Respir. Crit. Care Med.* **171**, 627-631.

1226
1227 Stein, M.L. (1999), *Interpolation of Spatial Data: Some Theory of Kriging*.
1228 Springer Verlag, New York.

1229
1230 Vedal S, Brauer M, White R, Petkau J (2003) Air pollution and daily mortality in a city with low
1231 levels of pollution. *Environmental Health Perspectives* **111**, 45-51.

1232
1233 Weschler, CJ (2000), Ozone in indoor environments: concentration and chemistry. *Indoor Air*
1234 **10**, 269-288.

1235
1236 Weschler, CJ (2006), Ozone's impact on public health: Contributions from indoor exposures to

1237 ozone and products of ozone-initiated chemistry. *Environmental Health Perspectives* **114**, 1489-
1238 1496

Tables and Figures

Ozone	Lags	Met	PM ₁₀	Original tlnise				Pop.-weighted Average			
				All-yr	RMSE	Summ	RMSE	All-yr	RMSE	Summ	RMSE
24-hr	CDL	NMMAAPS	No	0.522	0.124	0.404	0.133	0.618	0.096	0.459	0.107
24-hr	CDL	Ext	No	0.478	0.129	0.406	0.143	0.568	0.105	0.472	0.117
24-hr	CDL	NMMAAPS	Part	0.400	0.231	0.367	0.257	0.479	0.192	0.400	0.216
24-hr	CDL	NMMAAPS	Yes	0.314	0.230	0.265	0.261	0.360	0.193	0.267	0.223
24-hr	CDL	Ext	Part	0.379	0.248	0.376	0.281	0.434	0.212	0.410	0.240
24-hr	CDL	Ext	Yes	0.279	0.253	0.268	0.293	0.310	0.215	0.275	0.249
8-hr	CDL	NMMAAPS	No	0.411	0.080	0.323	0.084	0.466	0.063	0.361	0.069
8-hr	CDL	Ext	No	0.402	0.082	0.366	0.091	0.438	0.069	0.405	0.076
8-hr	CDL	NMMAAPS	Part	0.383	0.150	0.325	0.162	0.416	0.125	0.353	0.138
8-hr	CDL	NMMAAPS	Yes	0.292	0.153	0.258	0.168	0.314	0.129	0.208	0.145
8-hr	CDL	Ext	Part	0.311	0.165	0.313	0.180	0.337	0.140	0.342	0.154
8-hr	CDL	Ext	Yes	0.212	0.171	0.244	0.190	0.227	0.144	0.269	0.163
1-hr	CDL	NMMAAPS	No	0.292	0.054	0.239	0.057	0.299	0.043	0.250	0.047
1-hr	CDL	Ext	No	0.260	0.051	0.236	0.059	0.266	0.045	0.249	0.050
1-hr	CDL	NMMAAPS	Part	0.296	0.098	0.259	0.108	0.299	0.083	0.266	0.092
1-hr	CDL	NMMAAPS	Yes	0.213	0.102	0.200	0.112	0.213	0.086	0.206	0.097
1-hr	CDL	Ext	Part	0.257	0.105	0.244	0.116	0.261	0.091	0.250	0.100
1-hr	CDL	Ext	Yes	0.170	0.112	0.183	0.125	0.172	0.095	0.192	0.107
8-hr	0-1	NMMAAPS	No	0.234	0.053	0.242	0.057	0.260	0.041	0.265	0.045
8-hr	0-1	Ext	No	0.244	0.050	0.250	0.054	0.256	0.043	0.261	0.047
8-hr	0-1	NMMAAPS	Part	0.266	0.092	0.268	0.101	0.270	0.080	0.275	0.089
8-hr	0-1	NMMAAPS	Yes	0.196	0.098	0.190	0.111	0.191	0.084	0.190	0.096

Table 1: Summary of main hierarchical model results. First four columns indicate ozone metric (24-hour, 8-hour, 1-hour); lag structure (CDL=constrained distributed lag model; 0-1=average of lags 0 and 1); meteorology (NMMAAPS or extended) and whether PM₁₀ was included (“No” means PM₁₀ not included; “Part” means analysis with ozone only but only calculated for days where PM₁₀ was also available; “Yes” means PM₁₀ was included as a co-pollutant. For each combination of ozone, meteorology and PM₁₀, we tabulate the all-year TLNISE estimate (μ) and its root mean squared error (RMSE); the summer TLNISE estimate and its RMSE; the all-year population-weighted average estimate and its RMSE; and the summer population-weighted average estimate and its RMSE. All estimates in units of percentage rise in mortality per 10 ppb rise in the relevant metric of ozone.

Description	TLNISE				Pop.-weighted Average			
	All-yr	RMSE	Summ	RMSE	All-yr	RMSE	Summ	RMSE
24-hour ozone, below temp. median	0.449	0.162	0.02	0.163	0.531	0.136	-0.019	0.144
24-hour ozone, above temp. median	0.582	0.138	0.581	0.162	0.655	0.106	0.699	0.126
8-hour ozone, below temp. median	0.478	0.120	0.151	0.109	0.509	0.099	0.138	0.099
8-hour ozone, above temp. median	0.378	0.090	0.357	0.105	0.435	0.069	0.433	0.080
1-hour ozone, below temp. median	0.360	0.094	0.149	0.083	0.352	0.074	0.129	0.070
1-hour ozone, above temp. median	0.278	0.062	0.264	0.072	0.289	0.048	0.296	0.054
24-hour ozone, below SO ₂ median	0.509	0.154	0.353	0.177	0.605	0.119	0.392	0.141
24-hour ozone, above SO ₂ median	0.726	0.163	0.527	0.179	0.807	0.119	0.570	0.138
8-hour ozone, below SO ₂ median	0.406	0.098	0.314	0.111	0.446	0.078	0.330	0.091
8-hour ozone, above SO ₂ median	0.493	0.103	0.391	0.115	0.543	0.076	0.436	0.088
1-hour ozone, below SO ₂ median	0.230	0.064	0.186	0.080	0.236	0.052	0.188	0.062
1-hour ozone, above SO ₂ median	0.323	0.071	0.273	0.074	0.326	0.053	0.282	0.059
24-hour ozone, below PM ₁₀ median	-0.020	0.293	-0.111	0.318	0.100	0.235	-0.072	0.273
24-hour ozone, above PM ₁₀ median	0.569	0.228	0.564	0.275	0.603	0.200	0.581	0.242
8-hour ozone, below PM ₁₀ median	0.070	0.206	-0.004	0.243	0.163	0.160	0.068	0.188
8-hour ozone, above PM ₁₀ median	0.271	0.139	0.382	0.205	0.234	0.117	0.367	0.173
1-hour ozone, below PM ₁₀ median	0.077	0.135	-0.018	0.165	0.112	0.109	0.021	0.128
1-hour ozone, above PM ₁₀ median	0.332	0.105	0.350	0.119	0.324	0.089	0.347	0.104

Table 2: Within-city effect modifier results. For each of three ozone metrics (24-hour, 8-hour, 1-hour) and each of three potential effect modifiers (temperature, SO₂, PM₁₀), we show the estimated ozone-mortality relationships in a model that allows for separate linear ozone effects when the potential effect modifier is below or above its median, within a single overall regression model that includes adjustments due to temperature, dewpoint, seasonal and day of week effects, and long-term trends. Tabulated are point estimates (posterior means) and root mean squared errors, for all-year and summer-only data, using both the traditional TLNISE model and the alternative population-weighted average.

Region	Ozone	TLNISE				Pop.-weighted Average			
		All-yr	RMSE	Summ	RMSE	All-yr	RMSE	Summ	RMSE
Industrial Midwest	24-hour	0.755	0.263	0.816	0.232	0.682	0.279	0.723	0.253
North East	24-hour	1.507	0.268	1.669	0.216	1.289	0.294	1.450	0.246
North West	24-hour	0.071	0.413	0.058	0.394	0.240	0.487	0.218	0.471
Southern California	24-hour	0.198	0.301	0.153	0.235	-0.339	0.325	-0.348	0.259
South East	24-hour	0.405	0.199	0.482	0.174	0.284	0.216	0.340	0.194
South West	24-hour	-0.051	0.356	0.002	0.333	-0.044	0.393	0.017	0.371
Upper Midwest	24-hour	-0.039	0.485	-0.036	0.463	0.088	0.556	0.100	0.538
Industrial Midwest	8-hour	0.630	0.173	0.666	0.158	0.522	0.181	0.541	0.169
North East	8-hour	1.001	0.183	1.075	0.151	0.881	0.199	0.973	0.167
North West	8-hour	0.126	0.276	0.120	0.266	0.224	0.309	0.214	0.301
Southern California	8-hour	0.249	0.177	0.232	0.132	0.017	0.191	0.024	0.148
South East	8-hour	0.335	0.126	0.373	0.112	0.242	0.135	0.267	0.122
South West	8-hour	0.170	0.249	0.176	0.240	-0.044	0.274	-0.045	0.267
Upper Midwest	8-hour	-0.238	0.361	-0.209	0.354	0.112	0.408	0.141	0.402
Industrial Midwest	1-hour	0.419	0.128	0.424	0.122	0.349	0.136	0.350	0.129
North East	1-hour	0.664	0.124	0.674	0.108	0.585	0.137	0.614	0.118
North West	1-hour	0.370	0.180	0.362	0.178	0.304	0.199	0.299	0.197
Southern California	1-hour	0.103	0.090	0.095	0.065	0.065	0.102	0.057	0.074
South East	1-hour	0.253	0.087	0.259	0.082	0.199	0.096	0.209	0.089
South West	1-hour	0.046	0.181	0.044	0.179	-0.026	0.204	-0.025	0.201
Upper Midwest	1-hour	0.155	0.287	0.166	0.285	0.261	0.328	0.283	0.326

Table 3: Regional estimates. Similar to Table 1, except that results are subdivided into each of seven regions of U.S. All results for controlled distributed lag model with NMMAPS meteorology and without PM_{10} .

Model	Estimated Range (km.)	95% Confidence Interval
24-hour ozone, all year	327	(114, 938)
24-hour ozone, summer	1394	(378, 5134)
8-hour ozone, all year	427	(103, 1769)
8-hour ozone, summer	1387	(450, 4270)
1-hour ozone, all year	1084	(323, 3633)
1-hour ozone, summer	1264	(376, 4250)

Table 4: Range parameters for the Gaussian spatial covariance functions used in drawing the spatial maps, together with 95% confidence intervals. Estimates and standard errors were calculated on a logarithmic scale using the method of restricted maximum likelihood; confidence intervals were also calculated on a logarithmic scale and transformed back to the original scale for the results presented here.

Variable	24-hour AY	24-hour Summ	8-hour AY	8-hour Summ	1-hour AY	1-hour Summ
Window AC	3.5	3.2	2.1	2.1	1.1	1.1
SE	0.7	0.8	0.5	0.5	0.3	0.3
<i>p</i> -value	0.00001	0.0001	0.00002	0.00003	0.0006	0.0003
<i>R</i> ²	0.22	0.17	0.21	0.20	0.14	0.16
Proportion in different house in 1995	-5.9	-6.0	-3.7	-4.1	-2.7	-2.5
SE	1.7	1.9	1.1	1.2	0.8	0.8
<i>p</i> -value	0.0008	0.0018	0.0014	0.0004	0.0005	0.0024
<i>R</i> ²	0.11	0.10	0.10	0.11	0.12	0.09
Proportion 65+ moved since 1995	-7.0	-7.7	-4.0	-5.3	-3.2	-3.0
SE	2.0	2.3	1.4	1.5	0.9	1.0
<i>p</i> -value	0.0009	0.0009	0.0047	0.0004	0.0006	0.0028
<i>R</i> ²	0.11	0.11	0.08	0.12	0.12	0.09
Proportion public transport to work	3.5	3.0	1.8	1.8	1.0	1.1
SE	0.9	1.0	0.6	0.5	0.4	0.3
<i>p</i> -value	0.0001	0.0039	0.0017	0.0016	0.0049	0.0018
<i>R</i> ²	0.14	0.08	0.10	0.10	0.08	0.10
Proportion renters 65+ moved since 1995	-4.5	-4.1	-2.3	-2.3	-1.4	-1.2
SE	1.1	1.2	0.7	0.7	0.4	0.4
<i>p</i> -value	0.00005	0.0008	0.0013	0.0010	0.0023	0.0068
<i>R</i> ²	0.16	0.11	0.10	0.11	0.09	0.07
Mean SO2	0.115	0.101	0.062	0.056	0.049	0.042
SE	0.032	0.035	0.021	0.021	0.011	0.012
<i>p</i> -value	0.0007	0.006	0.004	0.010	0.00004	0.001
<i>R</i> ²	0.14	0.09	0.10	0.08	0.19	0.12
Proportion drive to work	-2.9	-2.4	-1.5	-1.4	-0.9	-0.8
SE	0.8	0.9	0.5	0.5	0.3	0.3
<i>p</i> -value	0.0002	0.006	0.002	0.005	0.006	0.013
<i>R</i> ²	0.14	0.08	0.09	0.08	0.08	0.06

Table 5, Part 1: Summary results for between-city effect modifiers. For each potential effect modifier, we show the regression coefficient (increase in mortality/ozone coefficient per 1-unit rise in effect modifier) together with standard error, *p*-value and *R*², in a random effects regression model fitted by the method of restricted maximum likelihood. The effect modifiers are ordered according to the largest *p*-value (shown in bold) over the six analyses of each effect modifier.

Variable	24-hour AY	24-hour Summ	8-hour AY	8-hour Summ	1-hour AY	1-hour Summ
Latitude	0.046	0.061	0.035	0.045	0.028	0.031
SE	0.021	0.023	0.014	0.015	0.010	0.010
<i>p</i> -value	0.033	0.011	0.015	0.003	0.005	0.002
<i>R</i> ²	0.05	0.07	0.06	0.09	0.08	0.09
Proportion owners 65+ moved since 1995	-5.5	-5.7	-2.8	-3.8	-2.4	-2.1
SE	1.9	2.1	1.3	1.4	0.8	0.9
<i>p</i> -value	0.005	0.008	0.036	0.006	0.005	0.022
<i>R</i> ²	0.08	0.07	0.04	0.07	0.08	0.05
Mean temperature	-0.029	-0.038	-0.022	-0.028	-0.018	-0.019
SE	0.014	0.015	0.009	0.009	0.006	0.006
<i>p</i> -value	0.037	0.012	0.015	0.003	0.004	0.004
<i>R</i> ²	0.04	0.06	0.06	0.09	0.09	0.08
Longitude	0.014	0.018	0.009	0.010	0.008	0.007
SE	0.007	0.008	0.005	0.005	0.002	0.002
<i>p</i> -value	0.043	0.025	0.056	0.039	0.001	0.006
<i>R</i> ²	0.04	0.05	0.04	0.04	0.11	0.08
Central AC	-1.2	-1.1	-0.9	-0.8	-0.6	-0.4
SE	0.5	0.5	0.3	0.3	0.2	0.2
<i>p</i> -value	0.014	0.039	0.004	0.015	0.006	0.080
<i>R</i> ²	0.08	0.05	0.11	0.07	0.09	0.04
Proportion black/African American	1.5	1.2	0.6	0.6	0.8	0.6
SE	0.8	0.8	0.5	0.5	0.3	0.4
<i>p</i> -value	0.04	0.15	0.20	0.23	0.02	0.12
<i>R</i> ²	0.04	0.02	0.02	0.01	0.06	0.02
Proportion unemployed	12.8	4.8	5.0	3.4	3.1	2.6
SE	5.8	6.5	3.7	4.0	2.6	2.7
<i>p</i> -value	0.03	0.46	0.19	0.40	0.23	0.35
<i>R</i> ²	0.05	0.01	0.02	0.01	0.01	0.01

Table 5, Part 2.

Range	Original tlnise				Pop.-weighted Average			
	All-yr	RMSE	Summ	RMSE	All-yr	RMSE	Summ	RMSE
0–40 ppb.	0.157	0.094	0.128	0.123	0.194	0.081	0.147	0.113
40–60 ppb.	0.331	0.107	0.313	0.108	0.340	0.089	0.327	0.096
60–80 ppb.	0.325	0.178	0.369	0.178	0.306	0.157	0.350	0.159

Table 6: Piecewise linear results based on 8-hour ozone, controlled distributed lag model, NMMAPS meteorology, no PM₁₀.

OZONE-MORTALITY COEFFICIENTS AND 95% PIs 24-HOUR OZONE – BELL (2004) MODEL

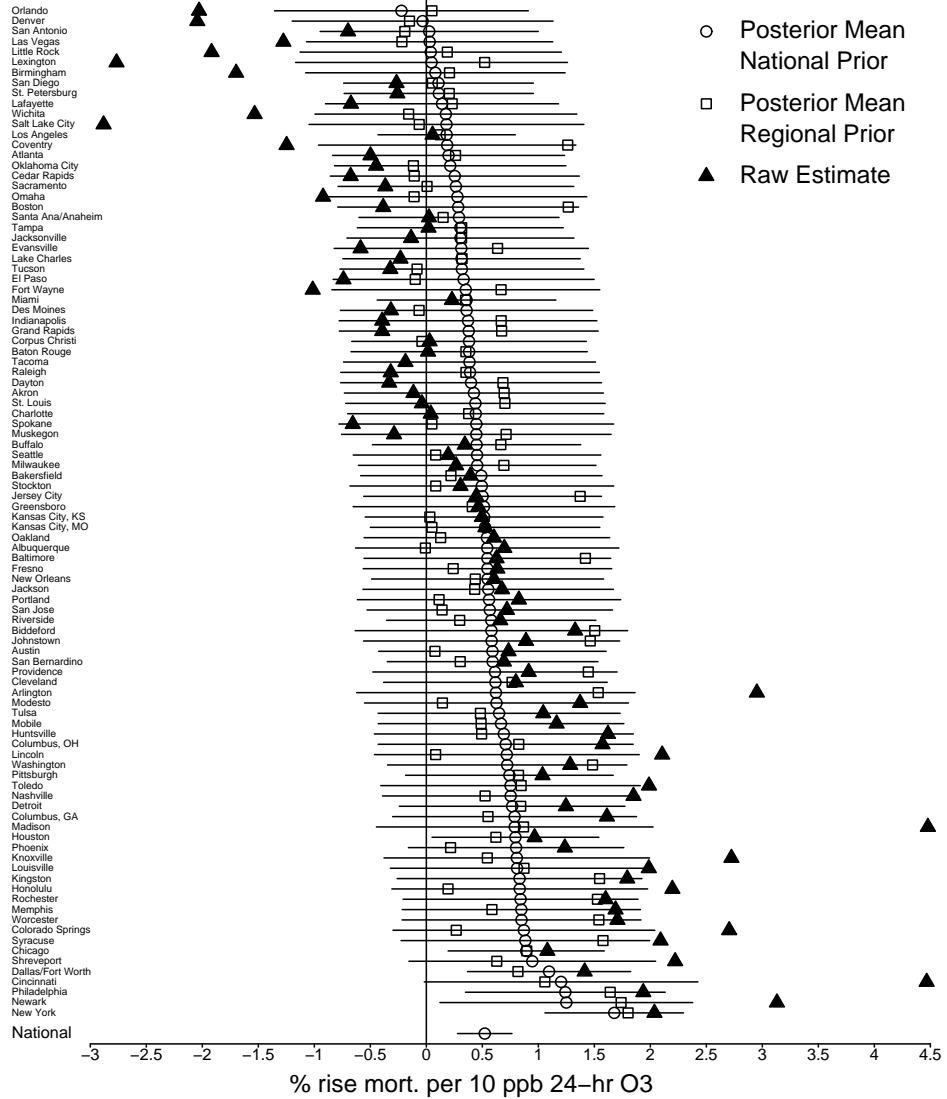


Fig. 1: 95% posterior intervals for the ozone-mortality coefficients, all-year data, by the hierarchical Bayesian method as in Fig. 2 of Bell et al. (2004). The Bayesian posterior estimates under the “national prior” (circles) are shown alongside those for the “regional prior” (squares) and the raw maximum likelihood estimates (triangles).

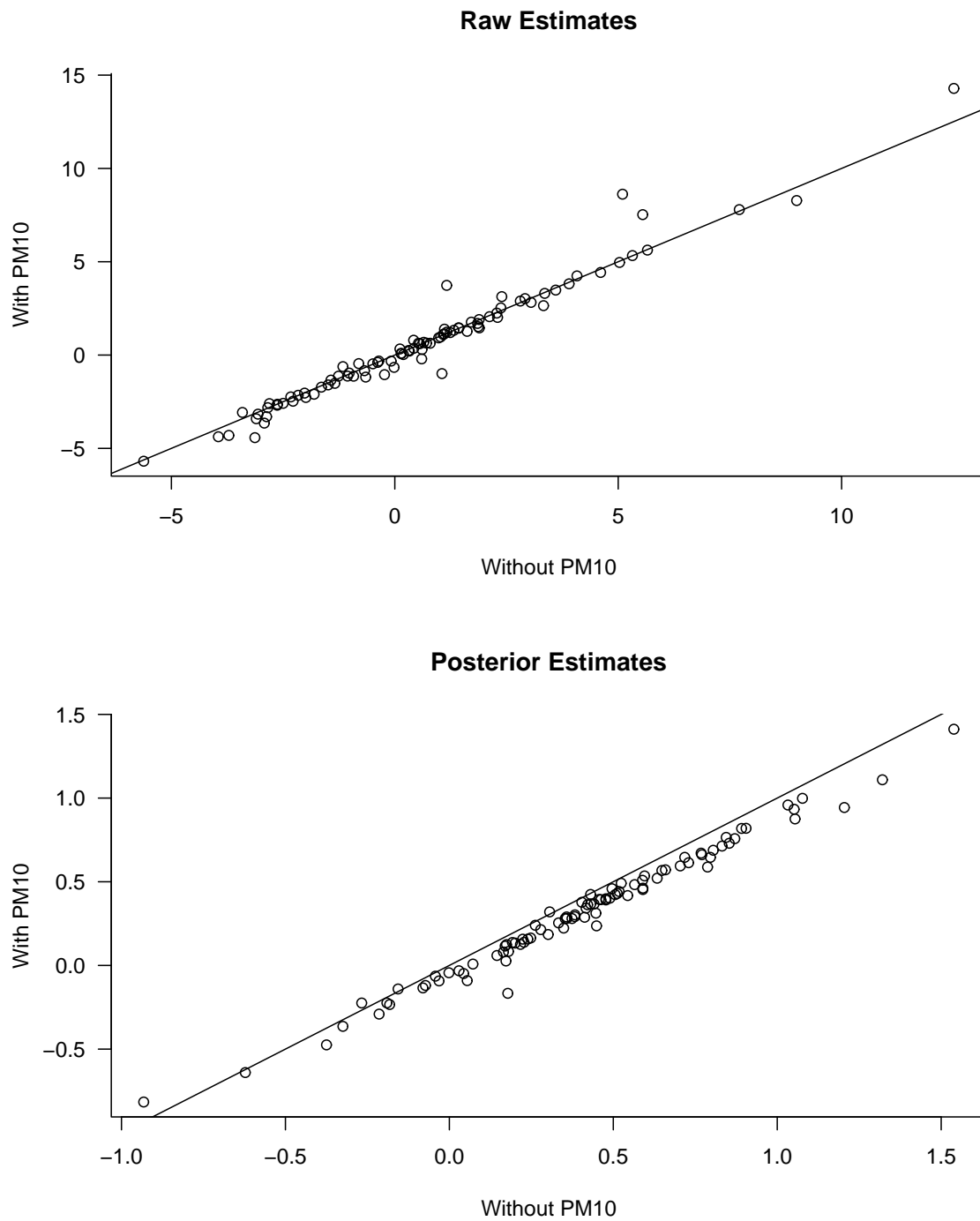


Fig 2: Plot of estimates including PM_{10} against those without. Top plot: based on raw estimates (this is the same as Fig. 3 of Bell (2004)). Bottom plot: based on posterior estimates. The posterior estimates are nearly all below the diagonal straight line, indicating that the PM-adjusted ozone coefficient is closer to 0 than the unadjusted ozone coefficient.

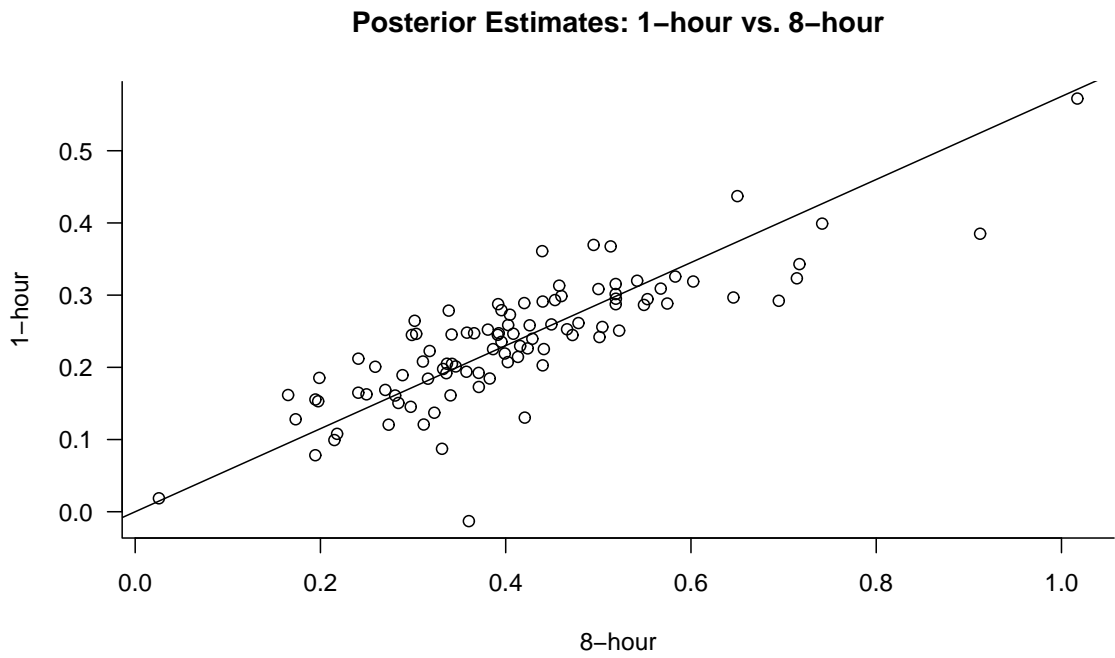
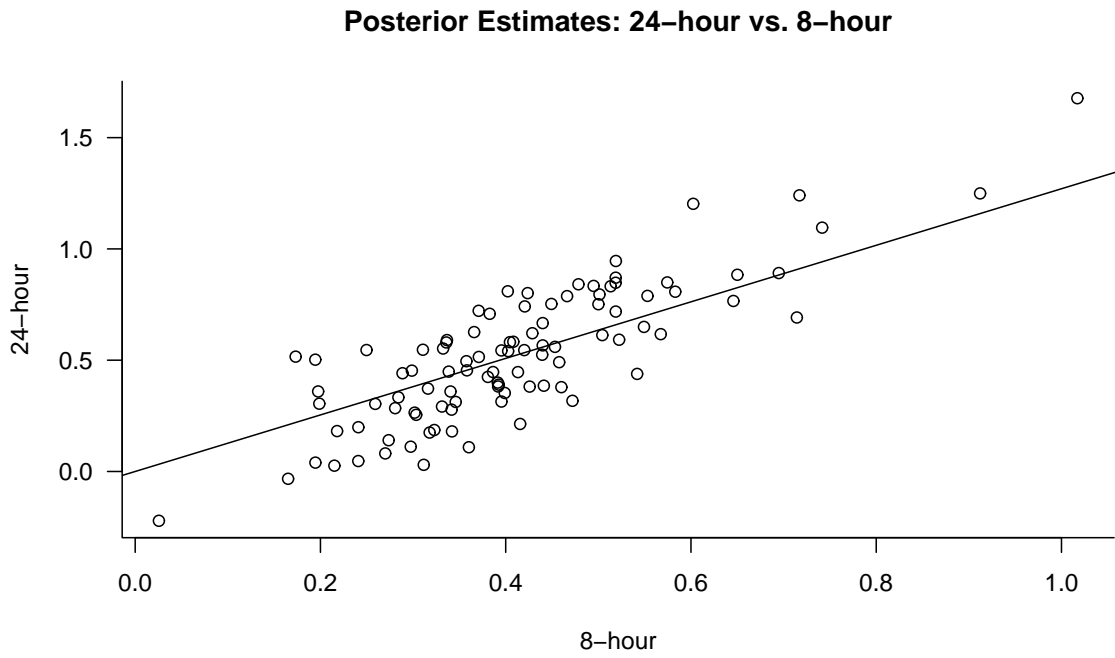


Fig 3: Scatterplots of posterior estimates corresponding to 24-hour, 8-hour, 1-hour ozone

OZONE-MORTALITY COEFFICIENTS AND 95% PIs 8-HOUR OZONE

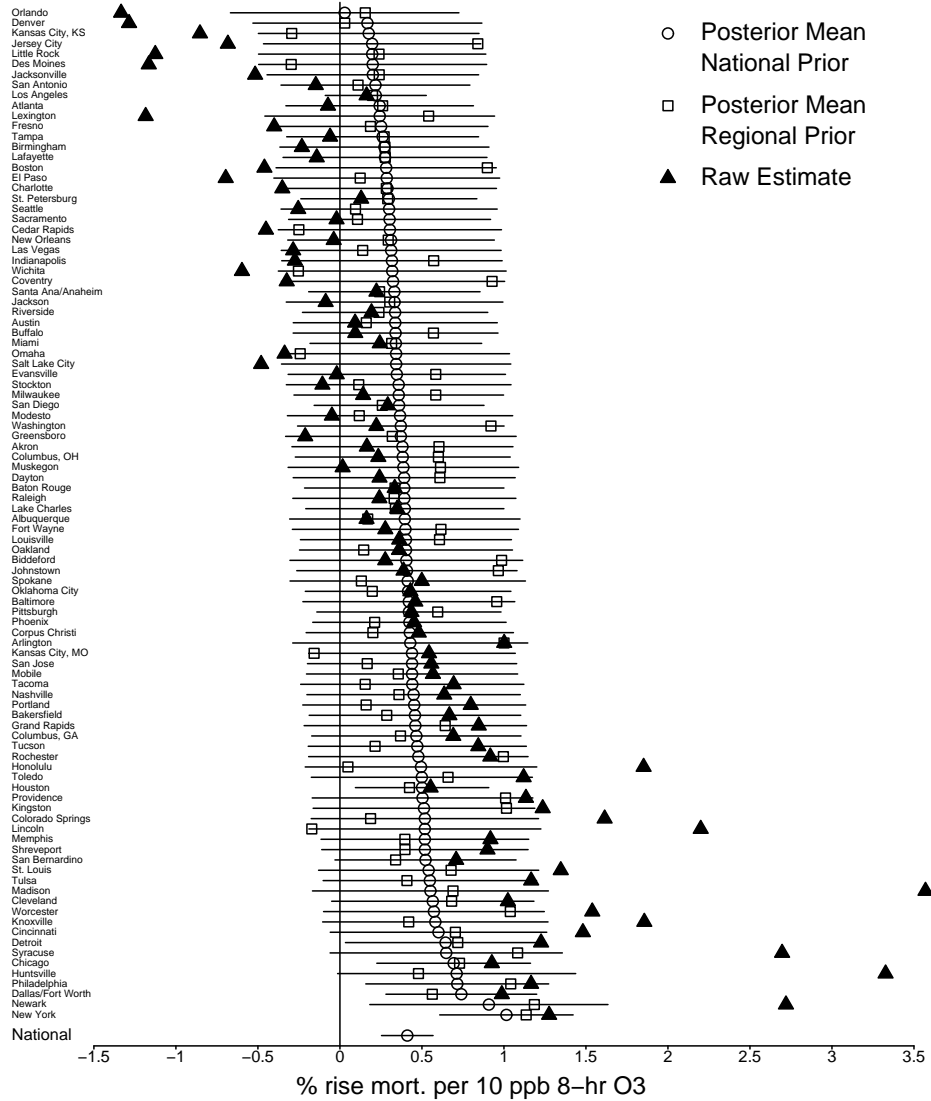


Fig 4: 95% posterior intervals for the ozone-mortality coefficients, based on 8-hour ozone, all-year data. The Bayesian posterior estimates under the “national prior” (circles) are shown alongside those for the “regional prior” (squares) and the raw maximum likelihood estimates (triangles).

OZONE-MORTALITY COEFFICIENTS AND 95% PIs 1-HOUR OZONE

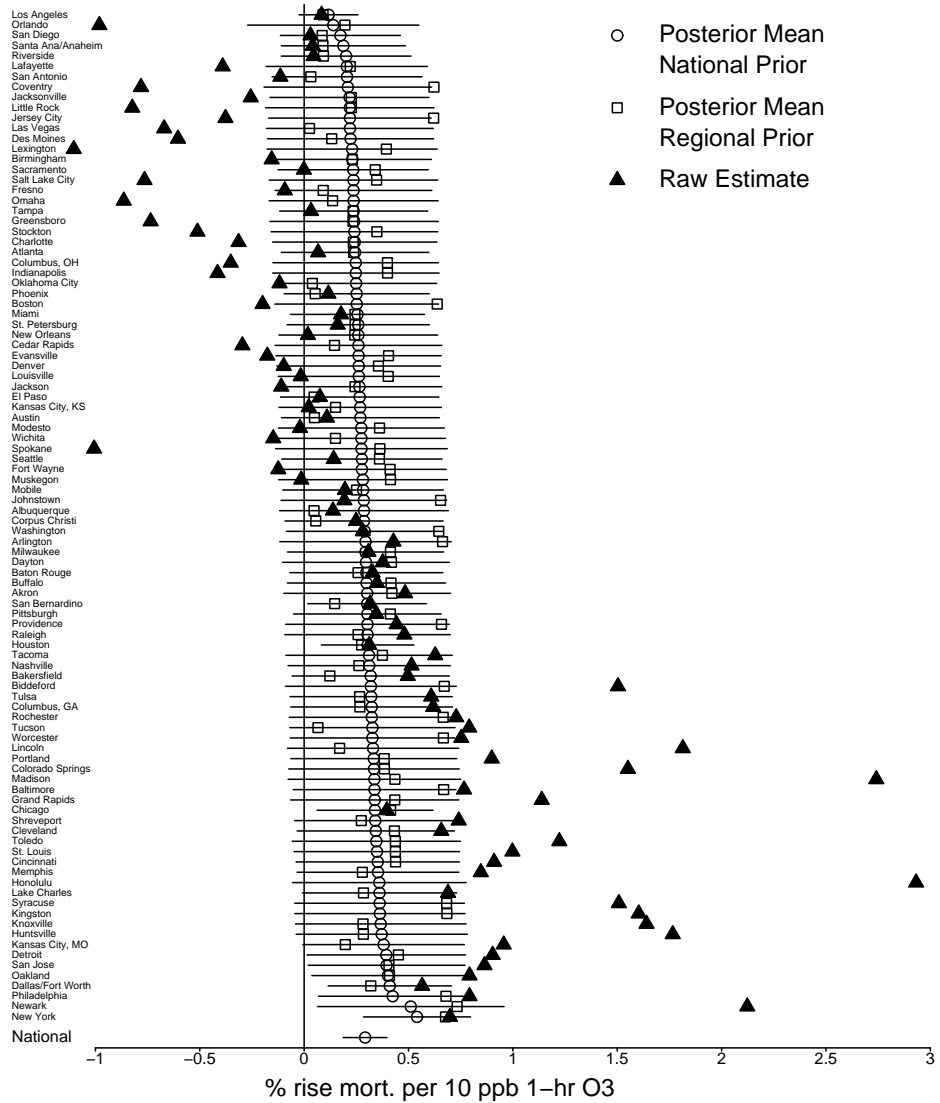


Fig 5: 95% posterior intervals for the ozone-mortality coefficients, based on 1-hour ozone, all-year data. The Bayesian posterior estimates under the “national prior” (circles) are shown alongside those for the “regional prior” (squares) and the raw maximum likelihood estimates (triangles).

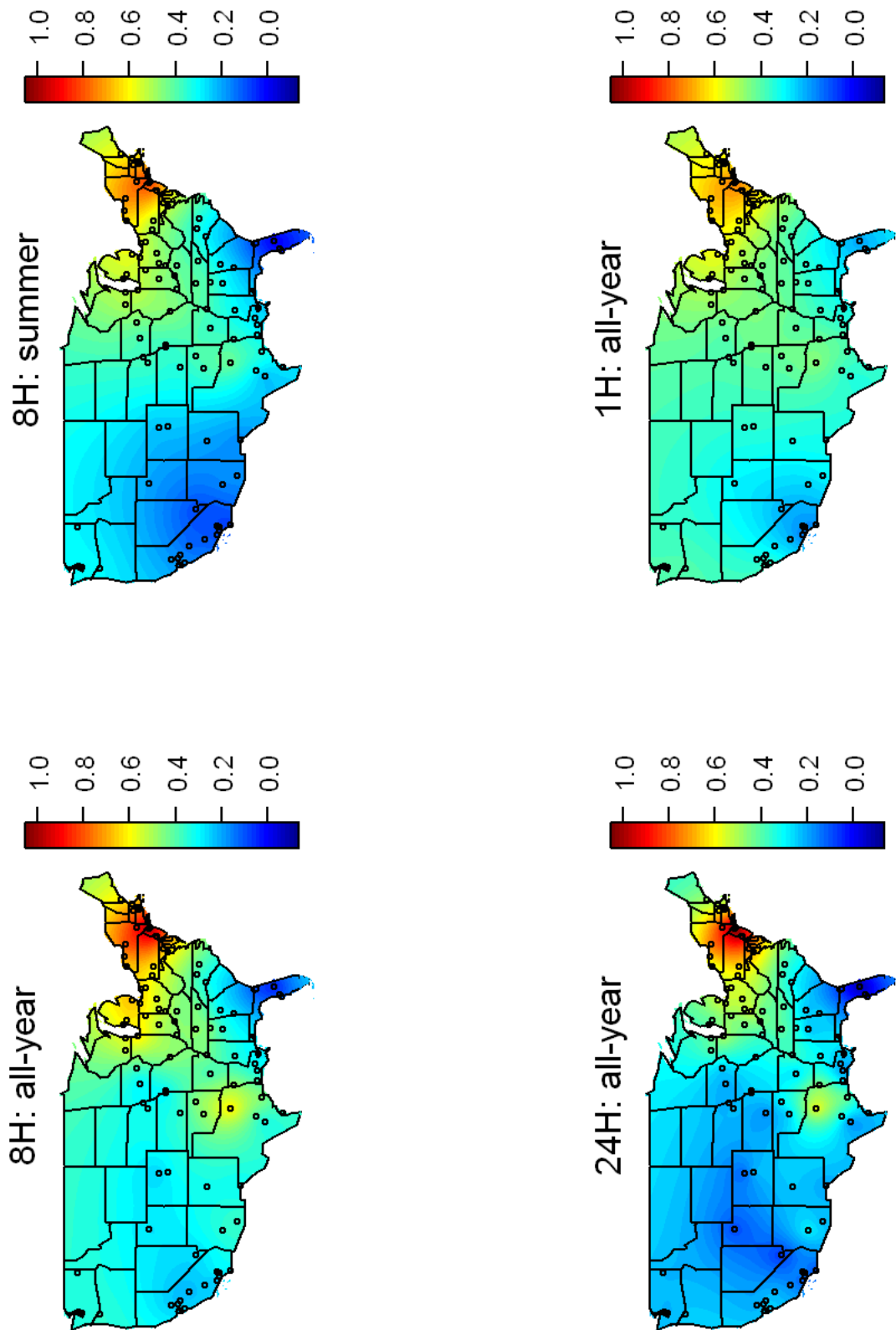


Fig 6: Map of spatially dependent ozone-mortality coefficient for 8-hour ozone (all-year data), 8-hour ozone (summer data), 24-hour ozone (all-year data) and 1-hour ozone (all-year data).

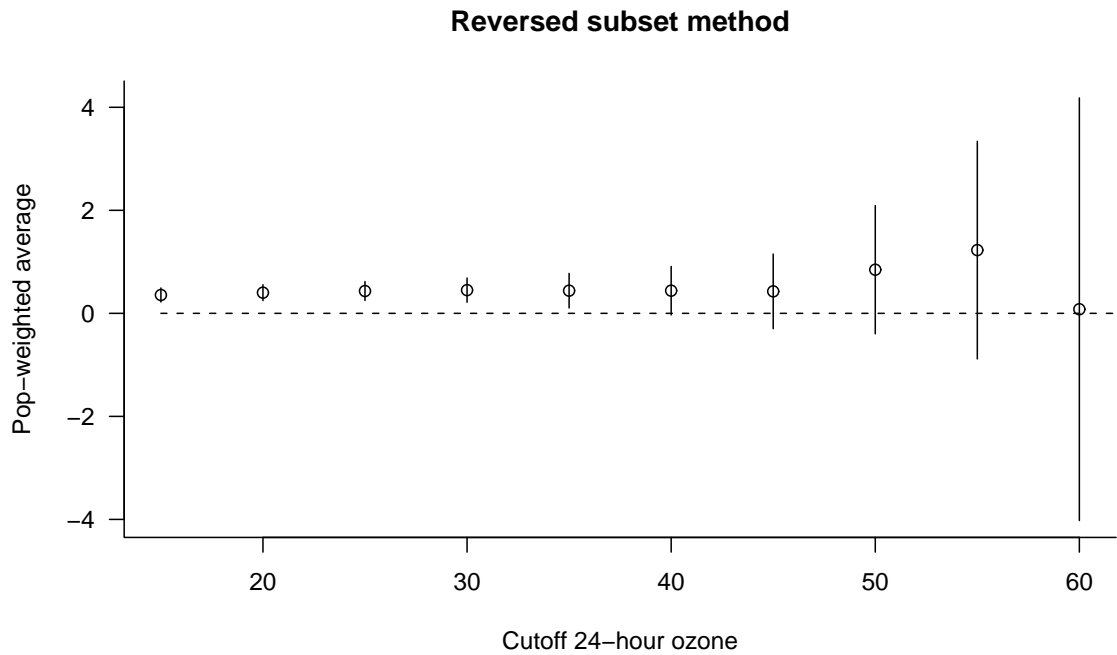
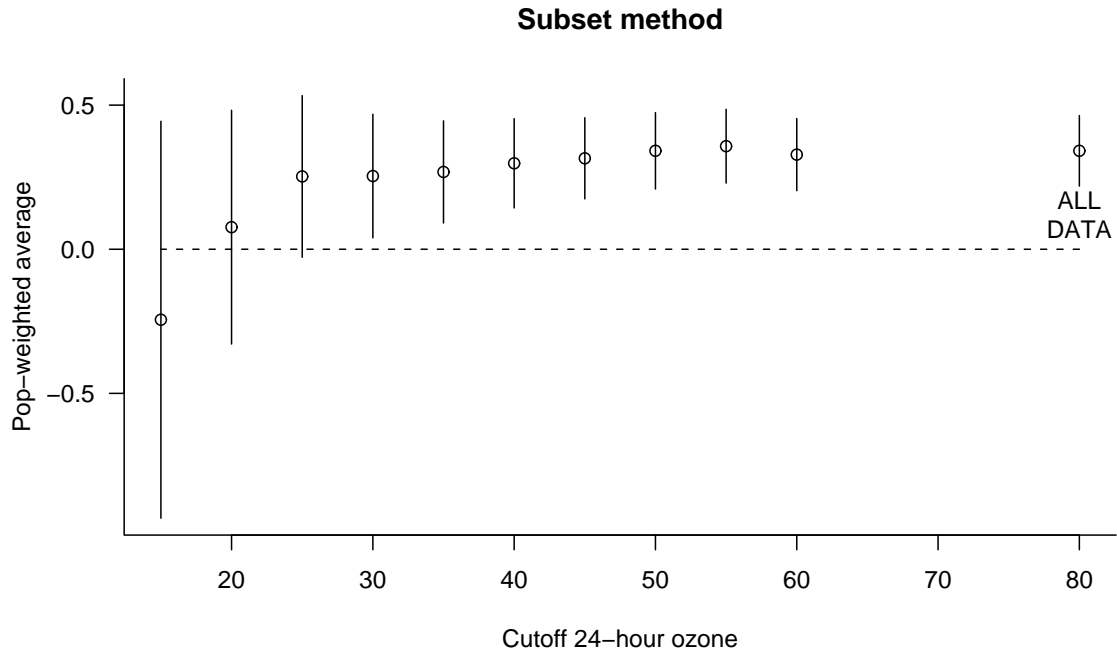


Fig 7. Top: The “subset approach” in which the 24-hour ozone-mortality model is restricted to a subset of days below 15, 20,...,60 ppb, and the unrestricted estimate. This is the same as Fig. 2 of Bell (2006) except that the calculation is in terms of population-weighted average rather than “national average” (but the appearance is very similar). Bottom: the same with the “reversed subset” approach, in which the data are subsetted according to all days greater than a given threshold, rather than all days less than a given threshold. The increasing width of the PIs as the threshold increases shows the uncertainty of estimating an ozone-mortality coefficient at high ozone levels.

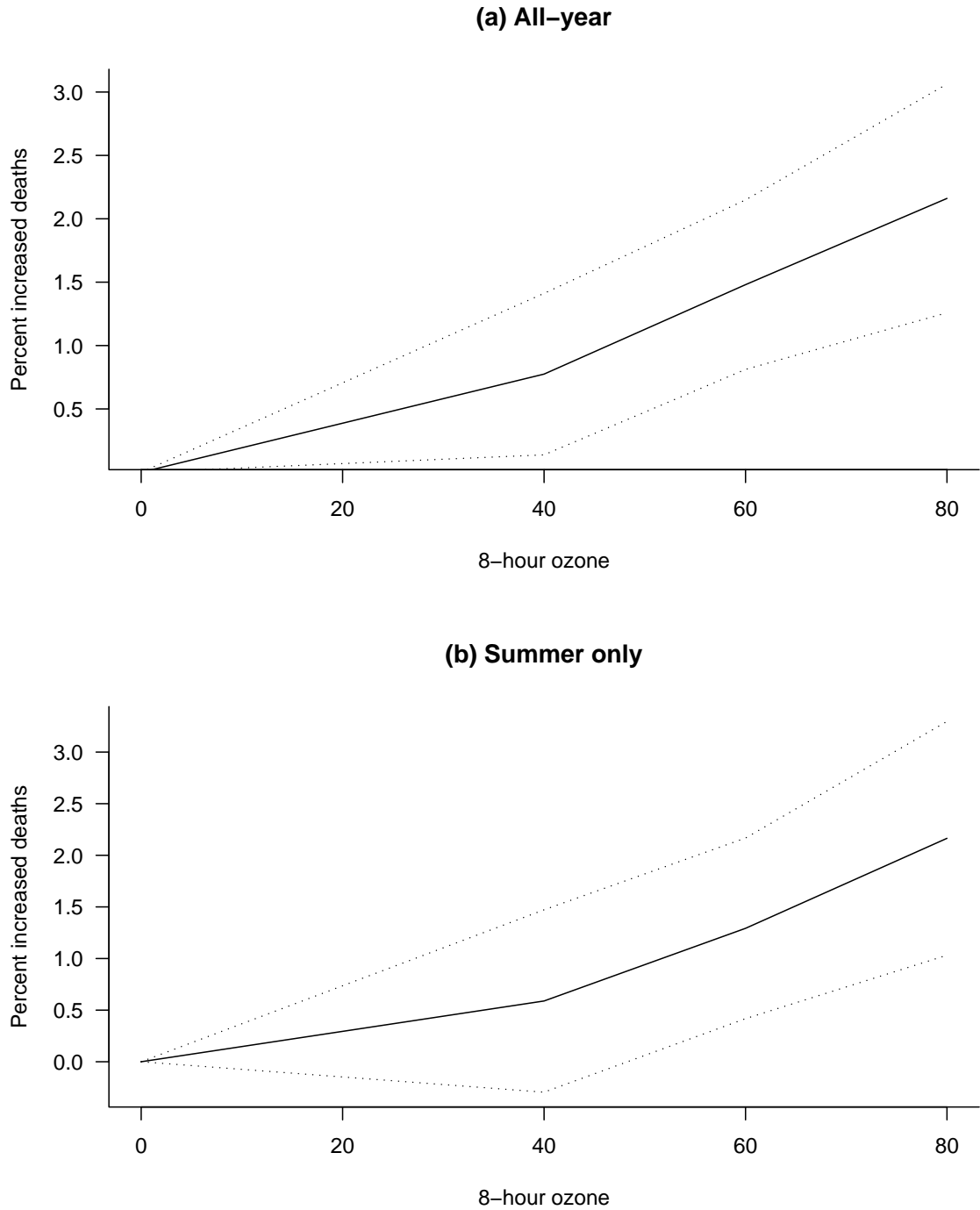


Fig 8: Estimates and pointwise 95% PIs assuming a piecewise-linear exposure-response relationship with breaks at 40 ppb and 60 ppb 8-hour ozone.

REGIONAL WEIGHTED AVERAGES, 8-HOUR OZONE PIECEWISE LINEAR CONCENTRATION-RESPONSE

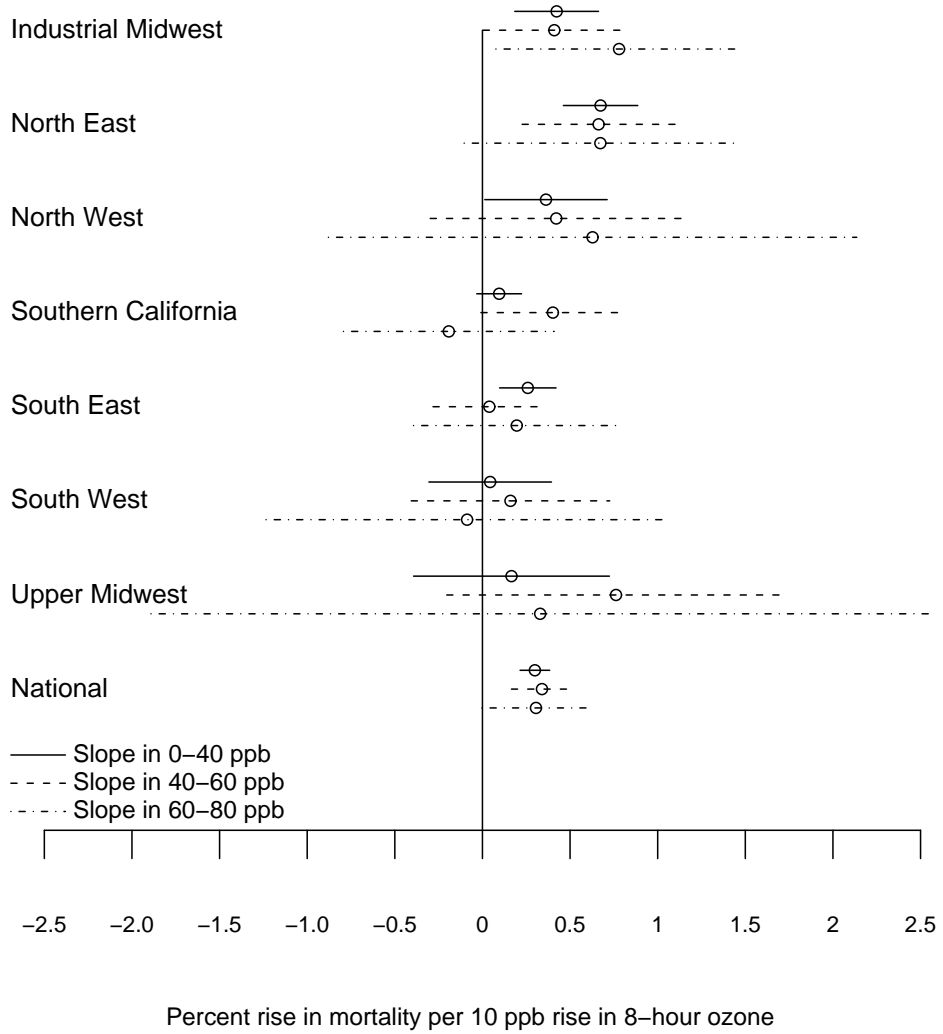
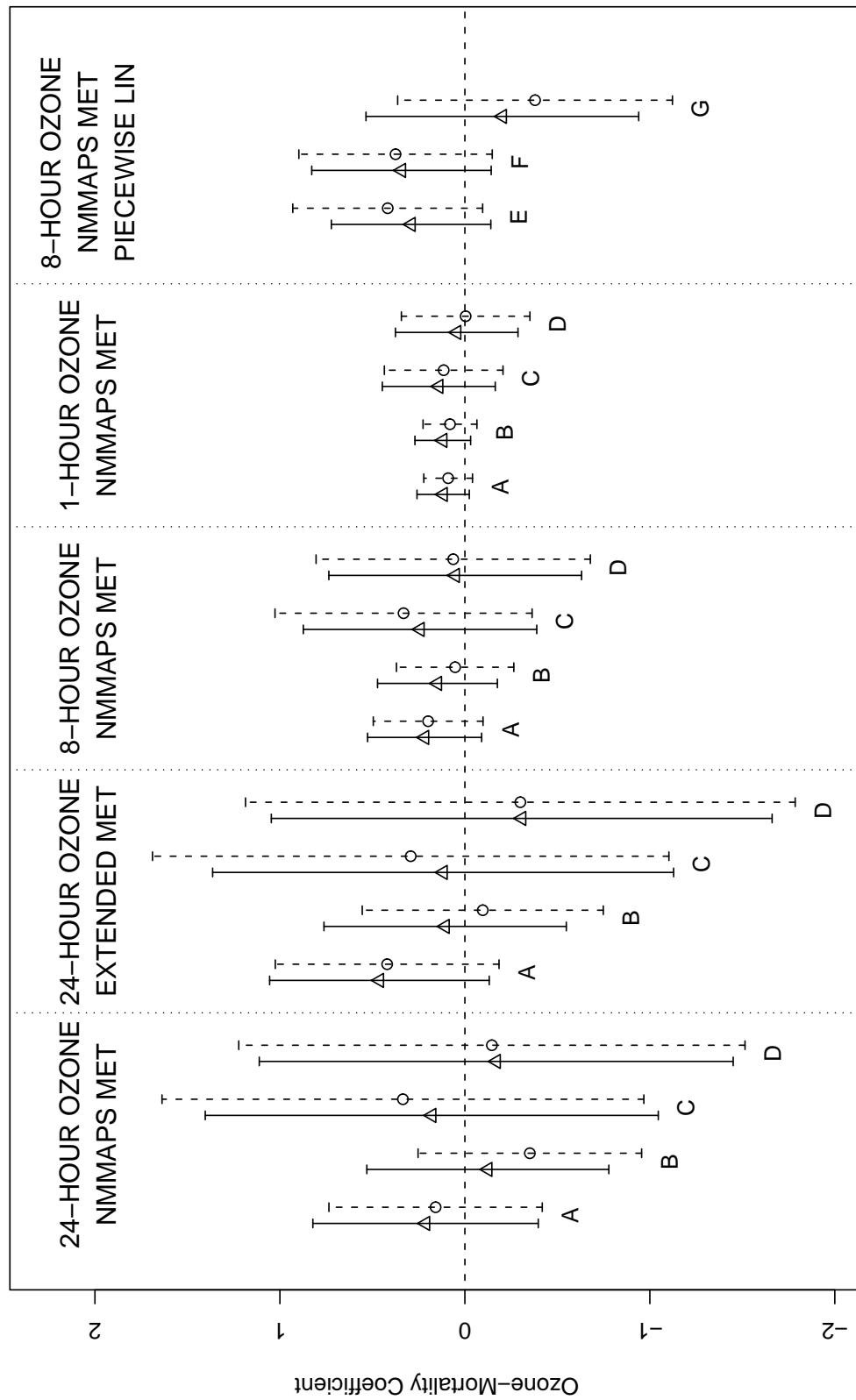


Fig 9: Regional slopes for piecewise-linear approach

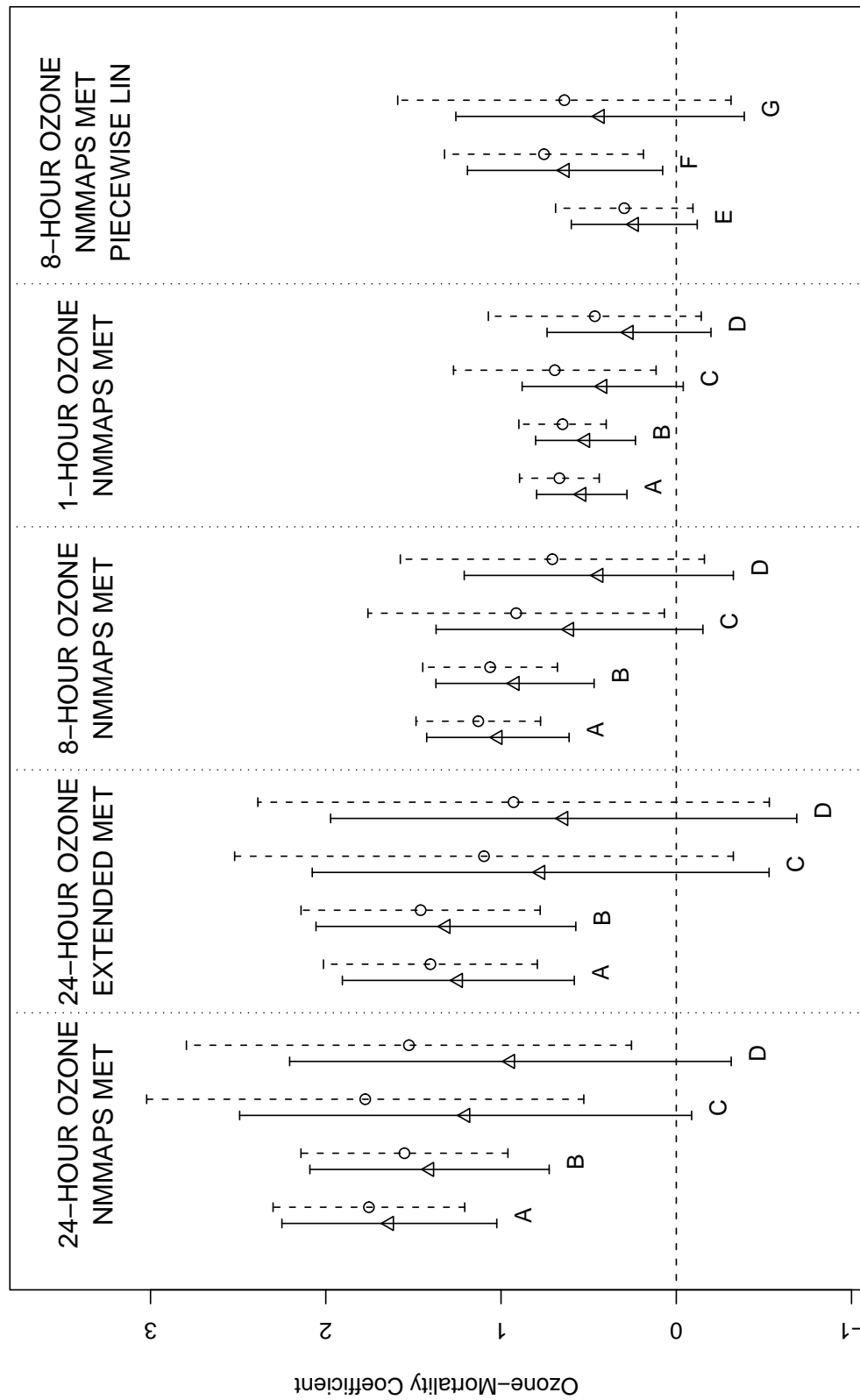
Los Angeles



Point estimates and 95% Prediction Intervals for 3 ozone metrics (24-hour, 8-hour, 1-hour), and 2 meteorologies (NMMAPS, extended). Triangles/solid lines: national prior. Circles/dashed lines: regional prior.
 A: Full dataset; B: Summer only; C: PM10 not included but analysis restricted to days on which PM10 available;
 D: PM10 included; EFG: Piecewise linear model slopes for 0-40 ppb (E), 40-60 ppb (F), 60-80 ppb (G)

Fig 10: City plot for Los Angeles

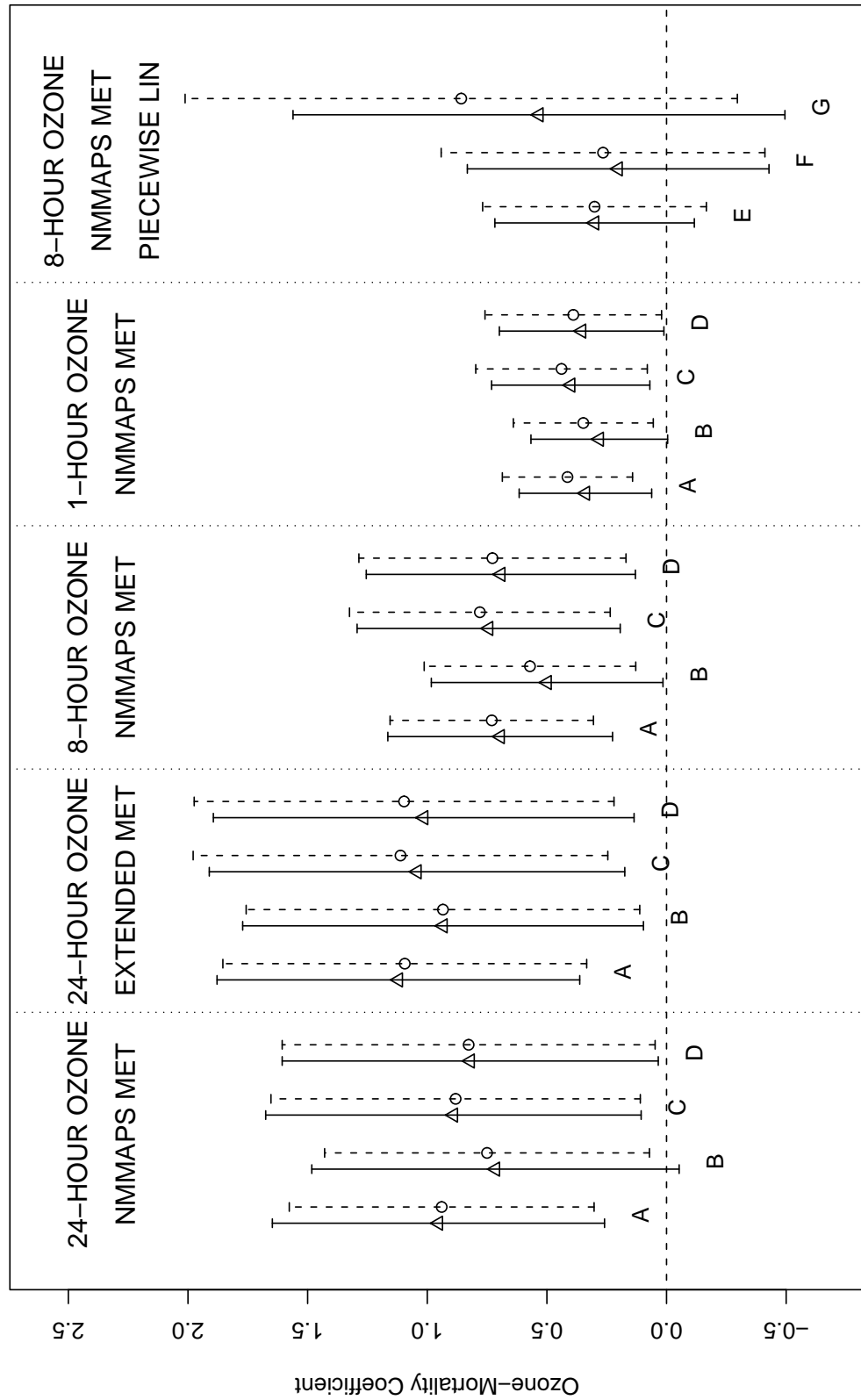
New York



Point estimates and 95% Prediction Intervals for 3 ozone metrics (24-hour, 8-hour, 1-hour), and 2 meteorologies (NMMAPS, extended). Triangles/solid lines: national prior. Circles/dashed lines: regional prior.
 A: Full dataset; B: Summer only; C: PM10 not included but analysis restricted to days on which PM10 available;
 D: PM10 included; EFG: Piecewise linear model slopes for 0-40 ppb (E), 40-60 ppb (F), 60-80 ppb (G)

Fig 11: City plot for New York

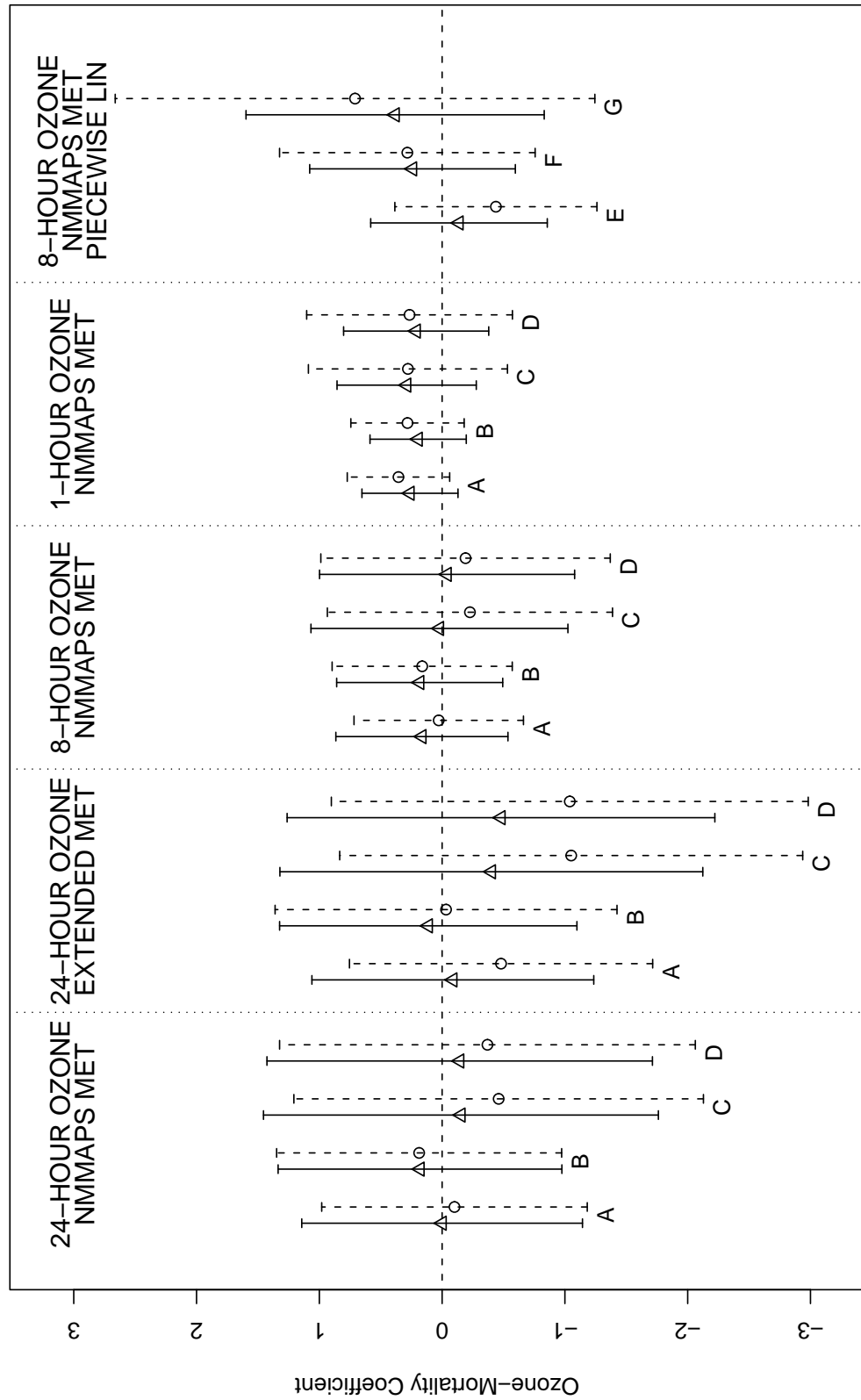
Chicago



Point estimates and 95% Prediction Intervals for 3 ozone metrics (24-hour, 8-hour, 1-hour), and 2 meteorologies (NMMAPS, extended). Triangles/solid lines: national prior. Circles/dashed lines: regional prior.
 A: Full dataset; B: Summer only; C: PM10 not included but analysis restricted to days on which PM10 available;
 D: PM10 included; EFG: Piecewise linear model slopes for 0-40 ppb (E), 40-60 ppb (F), 60-80 ppb (G)

Fig 12: City plot for Chicago

Denver



Point estimates and 95% Prediction Intervals for 3 ozone metrics (24-hour, 8-hour, 1-hour), and 2 meteorologies (NMMAPS, extended). Triangles/solid lines: national prior. Circles/dashed lines: regional prior.
 A: Full dataset; B: Summer only; C: PM10 not included but analysis restricted to days on which PM10 available;
 D: PM10 included; EFG: Piecewise linear model slopes for 0-40 ppb (E), 40-60 ppb (F), 60-80 ppb (G)

Fig 13: City plot for Denver