REASSESSING THE RELATIONSHIP BETWEEN OZONE AND SHORT-TERM MORTALITY IN U.S. URBAN COMMUNITIES

Richard L. Smith and Baowei Xu
(UNC STOR Department)
and Paul Switzer
(Stanford University)

UNC Department of Biostatistics
October 3 2007

Background

- Ground-level ozone is a pollutant that has long been regulated by the Environmental Protection Agency (EPA).

- In June this year, EPA proposed a reduction in the current ozone standard, from 80 ppb to a level between 70 and 75 ppb (based on max 8-hour ozone in any given day). We are now in the middle of a public comment period. The standard has to be finalized no later than March 2008.

- Critical in the EPA's scientific review were two papers from the "NMMAPS group":
  - Bell, McDermott, Zeger, Samet, Dominici (*JAMA* 2004)
  - Bell, Peng, Dominici (*Environ. Health Perspectives* 2006).

- The present paper is a reassessment of the results in these two papers, using data that the authors have made publicly available (http://www.ihapss.jhsph.edu)

What is NMMAPS?

- National Morbidity Mortality and Air Pollution Study
- A large multi-city air pollution study based at Johns Hopkins since 1997.
- Original focus on particulate matter (PM_{10}) but since 2004, main focus has been on ozone.
- 108 cities (98 used in ozone study), collected daily mortality data for 1987–2000, together with data on meteorology (principally temperature and dewpoint) and air pollution (ozone, PM_{10}, SO_2, NO_2, CO).
- Objective to study how mortality is affected by the air pollutant of interest when controlling for all the other known factors (including other air pollutants)

The NMMAPS Methodology: Stage I

- GLM (or GAM) model fitted in each city, based on overdispersed Poisson model for mortality
- One model fitted to all three age groups (< 65, 65–74, ≥ 75) but with interaction terms to allow for different long-term trends in the three groups
- Covariates
  - Temperature, dewpoint (nonlinear)
  - Long-term trends (via splines, typically 7 knots per year)
  - Day of week
  - Co-pollutants (in some analyses)
- Different lags for O_3. The 2004 paper emphasized the *constrained distributed lag model*, in which separate regression coefficients are fitted to each lagged day from 0 through 6, but the final result for each city is expressed as a single ozone-mortality coefficient (% increase in mortality for each 10 ppb rise in O_3)

The NMMAPS Methodology: Stage II
Combining Results Across Cities

Suppose θ_c is the unknown "true" coefficient in city c.

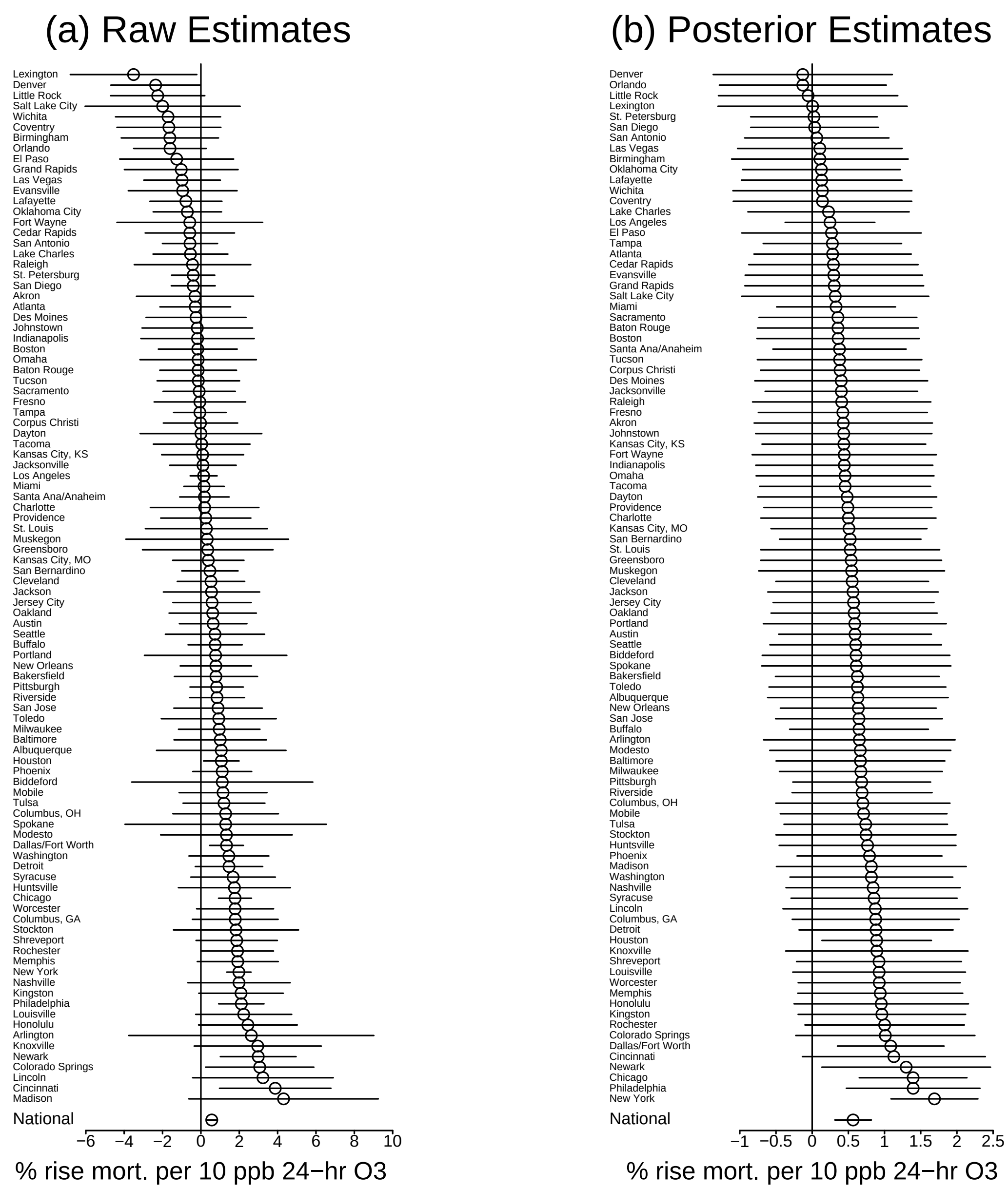
$\widehat{\theta}_c$ is the estimate in city c, with standard error s_c (treated as an exact standard deviation throughout the subsequent analysis)

Statistical model:

$$\begin{aligned} \theta_c &\sim N[\mu, \tau^2], \\ \widehat{\theta}_c \mid \theta_c &\sim N[\theta_c, s_c^2]. \end{aligned}$$

With an additional prior distribution for (μ, τ^2), they are able to compute posterior distributions for (μ, τ^2) and for each θ_c, using either Gibbs sampling or the TLNISE algorithm of Everson and Morris (2000).

OZONE−MORTALITY COEFFICIENTS AND 95% PIs



The "raw" estimates (derived from the individual-city GLMs) are very scattered, and have very wide 95% confidence intervals. But the "posterior" estimates are much closer together, and the "national" estimate (posterior mean and 95% PI for μ) is much narrower again.

Bell *et al.* (2004) quoted $\widehat{\mu} = 0.52$, with a 95% PI from 0.27 to 0.77. We get $\widehat{\mu} = 0.57$, 95% PI 0.31 to 0.82. Allowing for inevitable minor differences in the fine details of the analysis, this is excellent agreement.

However it's worth pointing out the huge reduction in apparent uncertainty levels that this involves — from the very wide range of raw estimates, to a much reduced but still wide range of posterior estimates, to a "national" estimate with apparently quite narrow PI.

Comment on methodology

Although the TLNISE method is elegant and easy to apply, it is not the only way of making these calculations. A non-Bayesian calculation based on restricted maximum likelihood (REML) produces almost the same answers.

Combined National Estimates

Posterior (or REML) estimates of μ:

TLNISE — 0.566 (posterior SD=0.129)
REML — 0.569 (SE=0.123)
Bell *et al.* (2004) — 0.52 (posterior SD=0.13)

But it's not clear what μ actually means. It would be better to estimate a weighted mean effect over the 98 cities, using weights that are proportional to populations. This leads to:

TLNISE — 0.681 (posterior SD=0.097)
REML — 0.681 (SE=0.097)

However, later we shall argue that even this estimate has limited practical meaning.

Alternative Meteorological Model

The NMMAPS meteorological model ("regular") uses nonlinear functions of temperature and dewpoint at lag 0, and of the average of temperature and dewpoint across lags 1–3, but no contribution from lags 4–6.

But with a distributed lag model for ozone, there may be meteorological confounding, especially at lags 4–6.

We therefore tried a "distributed lag" meteorological model, ("extended") using temperature and dewpoint at each of lags 0 through 6, modeled nonlinearly through splines with respectively 4 and 3 df.

Results

Posterior (or REML) estimates of μ:

TLNISE — 0.428 (posterior SD=0.133)
REML — 0.435 (SE=0.125)

Weighted mean effect over the 98 cities:

TLNISE — 0.520 (posterior SD=0.107)
REML — 0.519 (SE=0.106)

The point estimates are about 25% smaller than in the original distributed lag model, though they are still larger than from any single-lag or two-lag model.

PM_{10} as a Co-Pollutant

Bell *et al.* claim that putting in PM_{10} as a co-pollutant (as well as ozone) does not affect the ozone coefficient.

The comparison is not so straightforward because in most cities, PM_{10} is measured only every sixth day. We can only include days on which PM_{10} measurement exists, and also, this restricts us to single-lag models for PM_{10} (but not ozone).

We continue to measure ozone through the constrained distributed lag model, and PM_{10} at lag 1 (as in several earlier NMMAPS papers). Initially we run the ozone model without PM_{10}, but only on days for which PM_{10} measurement exists. Then we repeat the analysis with both ozone and PM_{10}.

Also, this reduces the count of available cities from 98 to 93.

Results (extended met model)

Estimate	Without PM_{10}	With PM_{10}
μ (Bayes)	.420	.312
(Posterior SD)	(.254)	(.257)
Weighted mean (Bayes)	.456	.323
(Posterior SD)	(.235)	(.237)

The ozone coefficient seems to be reduced on average about 25% if PM_{10} is included. Combined with the use of the distributed lag met. model, we have already reduced the estimated weighted mean by over 50% compared with the original estimate.

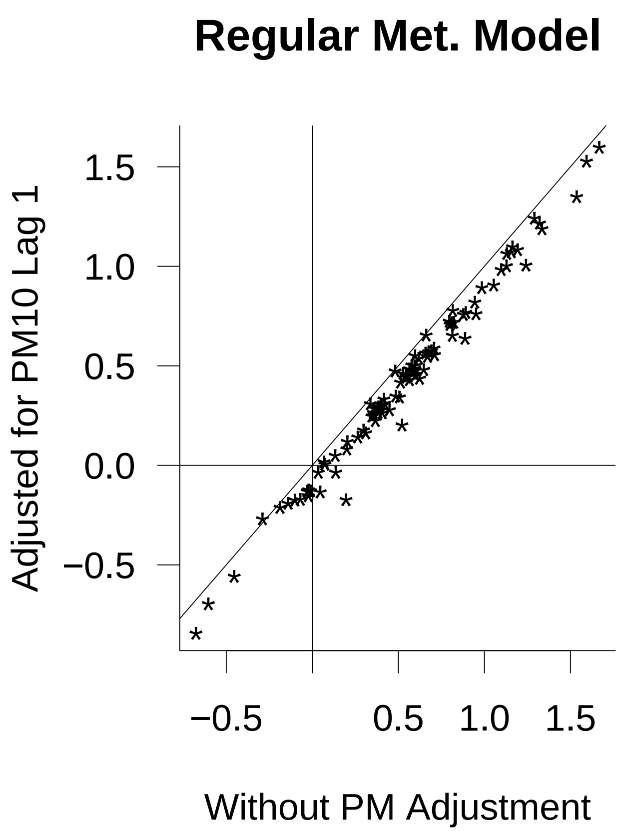
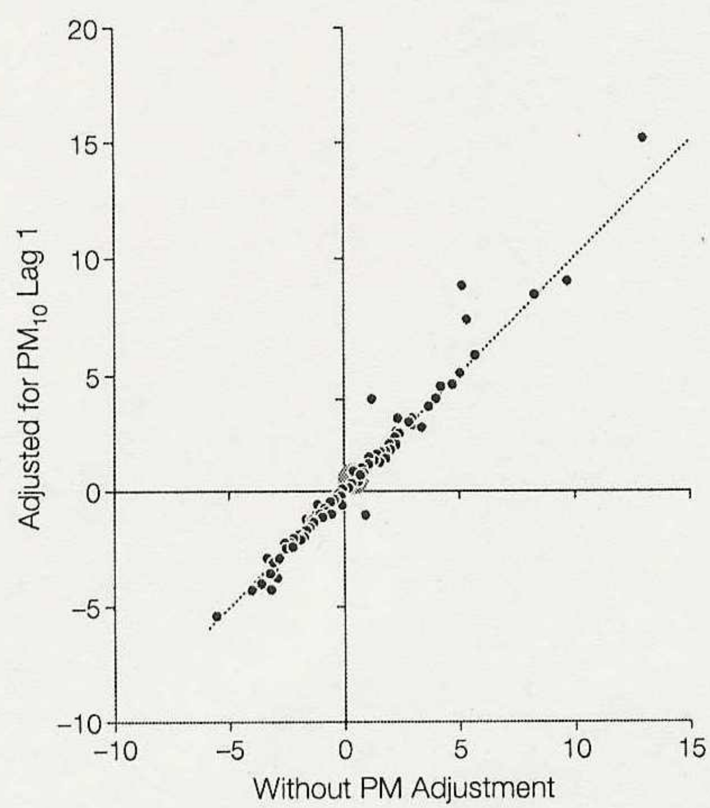


Figure 3. Community-Specific Maximum Likelihood Estimates of the Short-term Effects of Ozone on Mortality, With and Without Adjustment for PM_{10}



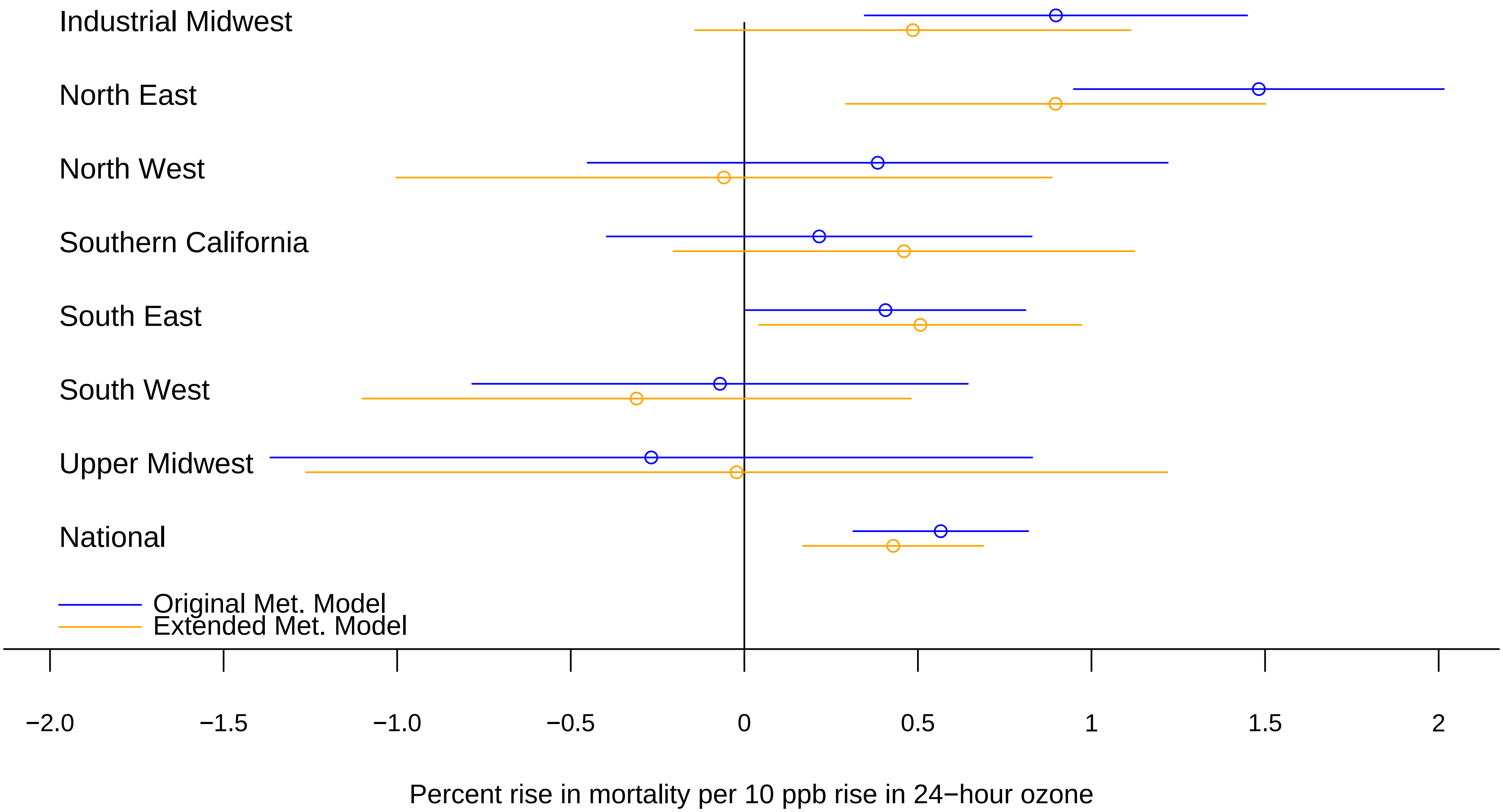
Remark. This is the one place in the entire talk where we got a result that appears to be in direct contradiction to Bell *et al.* (2004). The left plot (individual-city posterior estimates, regular met. model) is clearly not the same as Bell's Figure 3 (right plot).

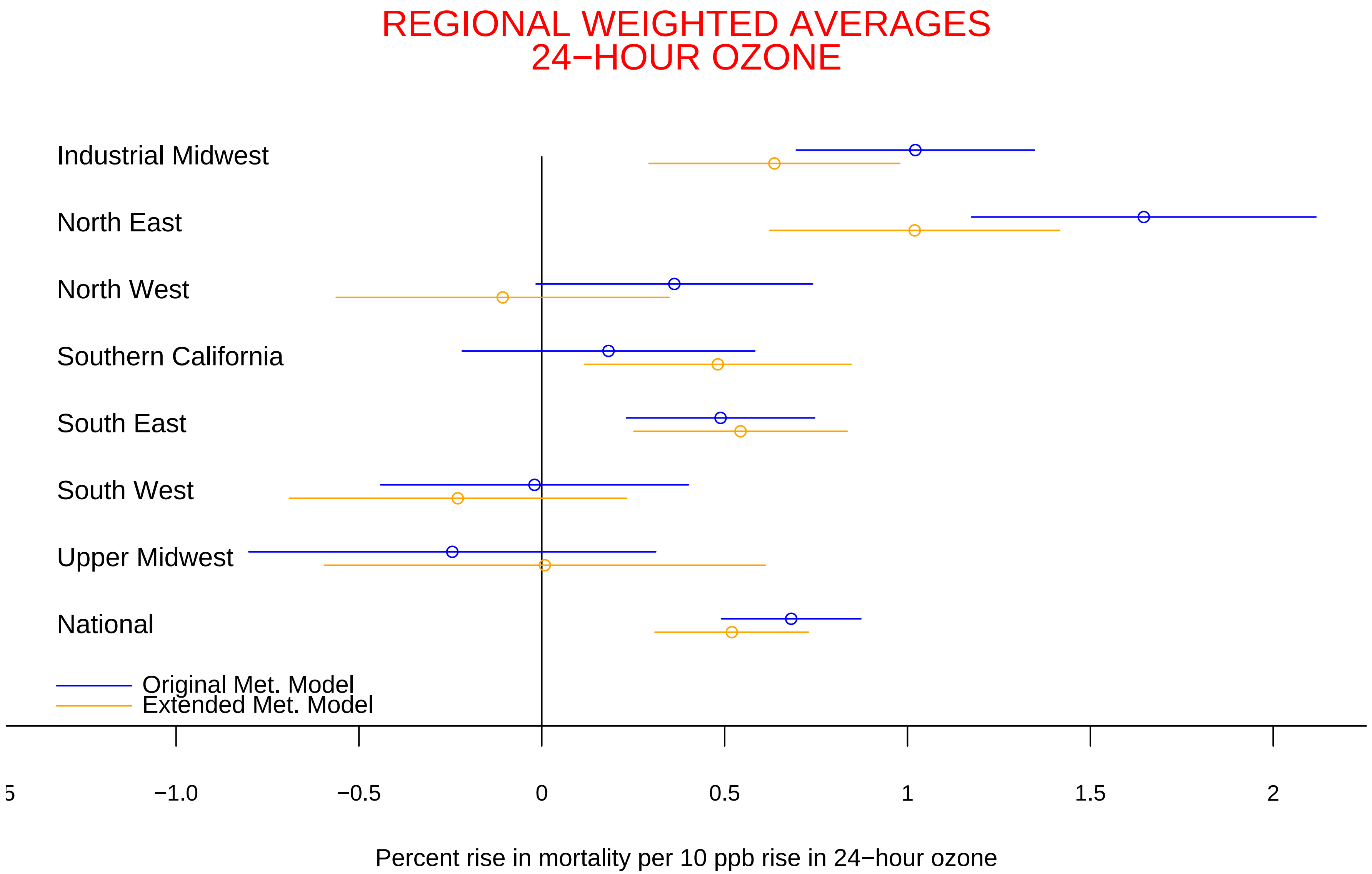
Regional Estimation

In several of the NMMAPS PM_{10} papers, the calculation of a "national" estimate and 95% PI was followed by the corresponding calculation for 7 regions of the US. Bell and Dominici (2006) contained similar calculations for ozone.

We compute regional estimates by including an indicator for region as a covariate in the TLNISE analysis, both regular and extended met. model.

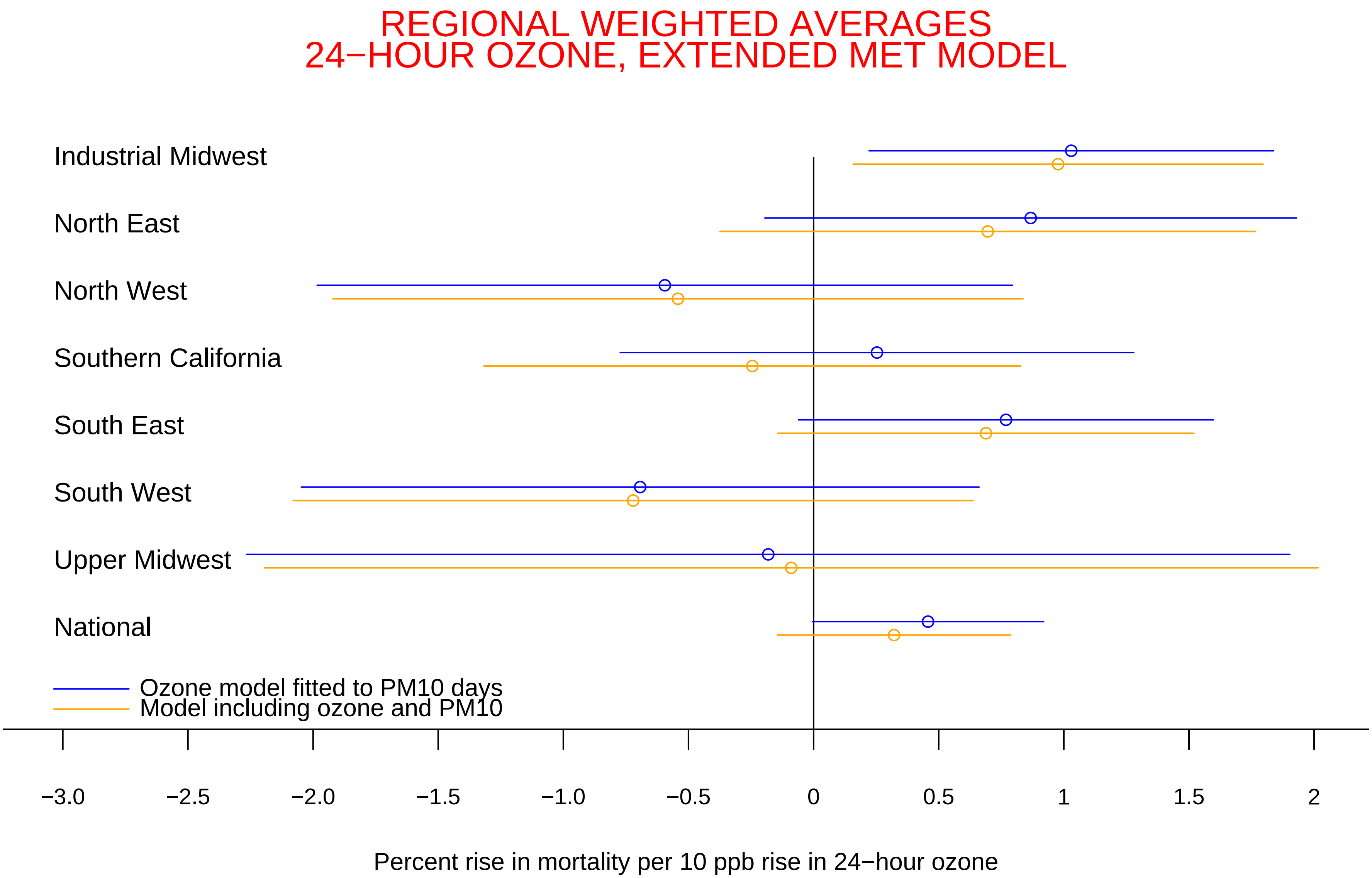
REGIONAL ESTIMATES
24−HOUR OZONE





REGIONAL WEIGHTED AVERAGES
24−HOUR OZONE
Industrial Midwest
North East
North West
Southern California
South East
South West
Upper Midwest
National
Original Met. Model
Extended Met. Model
−1.0
−0.5
0
0.5
1
1.5
2
Percent rise in mortality per 10 ppb rise in 24−hour ozone

We also make the same calculations for the extended meteorology model with and without PM_{10}


REGIONAL WEIGHTED AVERAGES
24−HOUR OZONE, EXTENDED MET MODEL
Industrial Midwest
North East
North West
Southern California
South East
South West
Upper Midwest
National
Ozone model fitted to PM10 days
Model including ozone and PM10
−3.0
−2.5
−2.0
−1.5
−1.0
−0.5
0
0.5
1
1.5
2
Percent rise in mortality per 10 ppb rise in 24−hour ozone

Interpretation of Regional Estimates

The regional analyses imply strong effects in the Northeast and Industrial Midwest, less strong but still significant effects in the Southeast (including Texas) and possibly Southern California, and insignificant or negative estimates in the other regions.

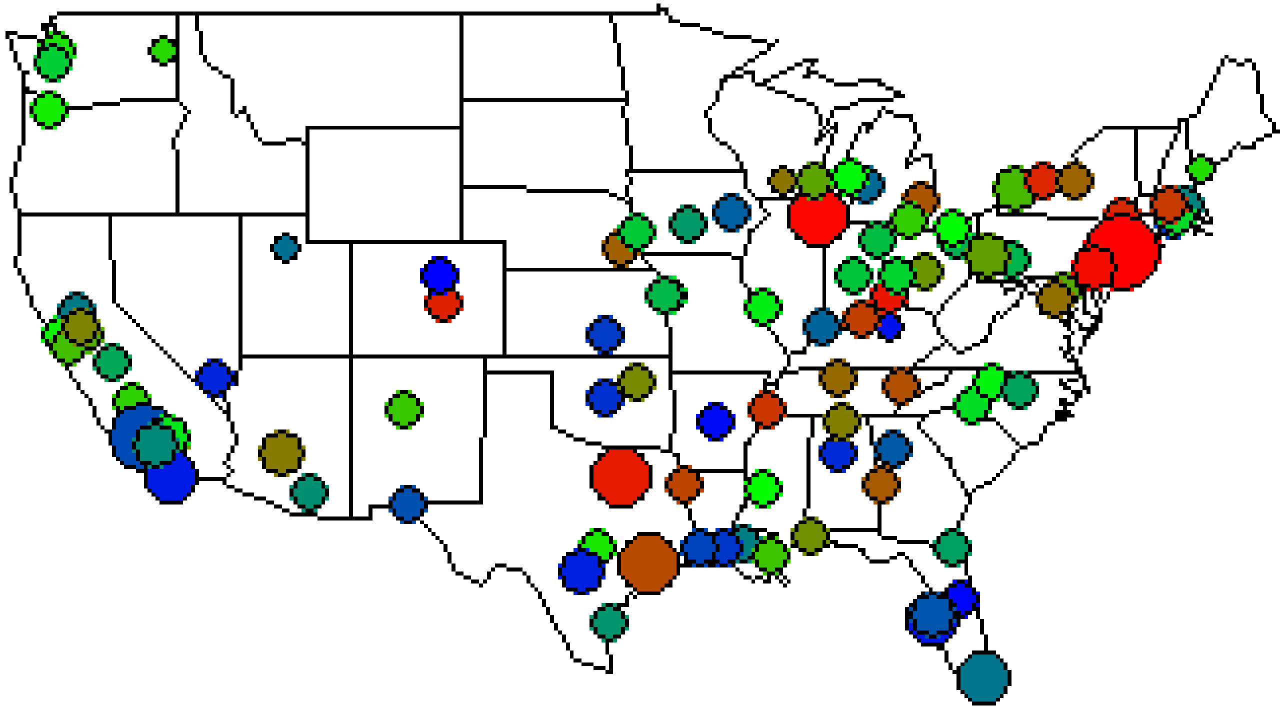
The NMMAPS PM_{10} analysis (e.g. Dominici *et al.* 2003) also showed regional variation, but not nearly so much inter-region variability.

Moreover, in the case of PM_{10} there are natural reasons for the variability (e.g. different chemical constituents in different places), which are not evident here.

Spatial Analysis

The next figure shows a spatial representation of the variability of the ozone-mortality coefficient. You can see a similar picture for PM_{10} in the IHAPPS home page.

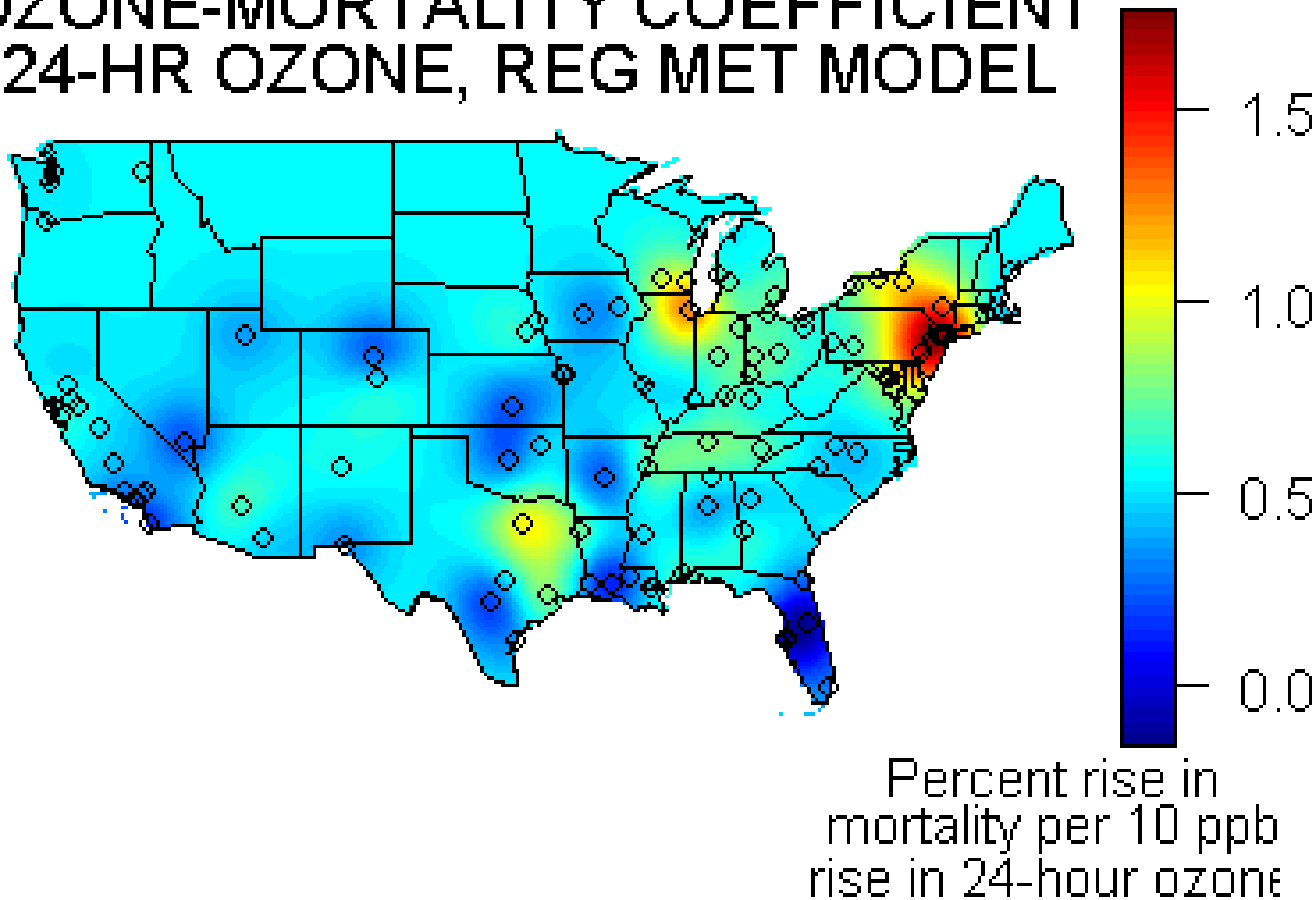
EFFECT ESTIMATES BASED ON 24-HOUR OZONE



Map of posterior city-by-city ozone-mortality coefficients on an ordinal scale (blue=smallest; green=median; red=largest). Areas of circles are inversely proportional to posterior variances.

This suggests we explore a more rigorous spatial analysis, using spatial interpolation methods analogous to kriging.

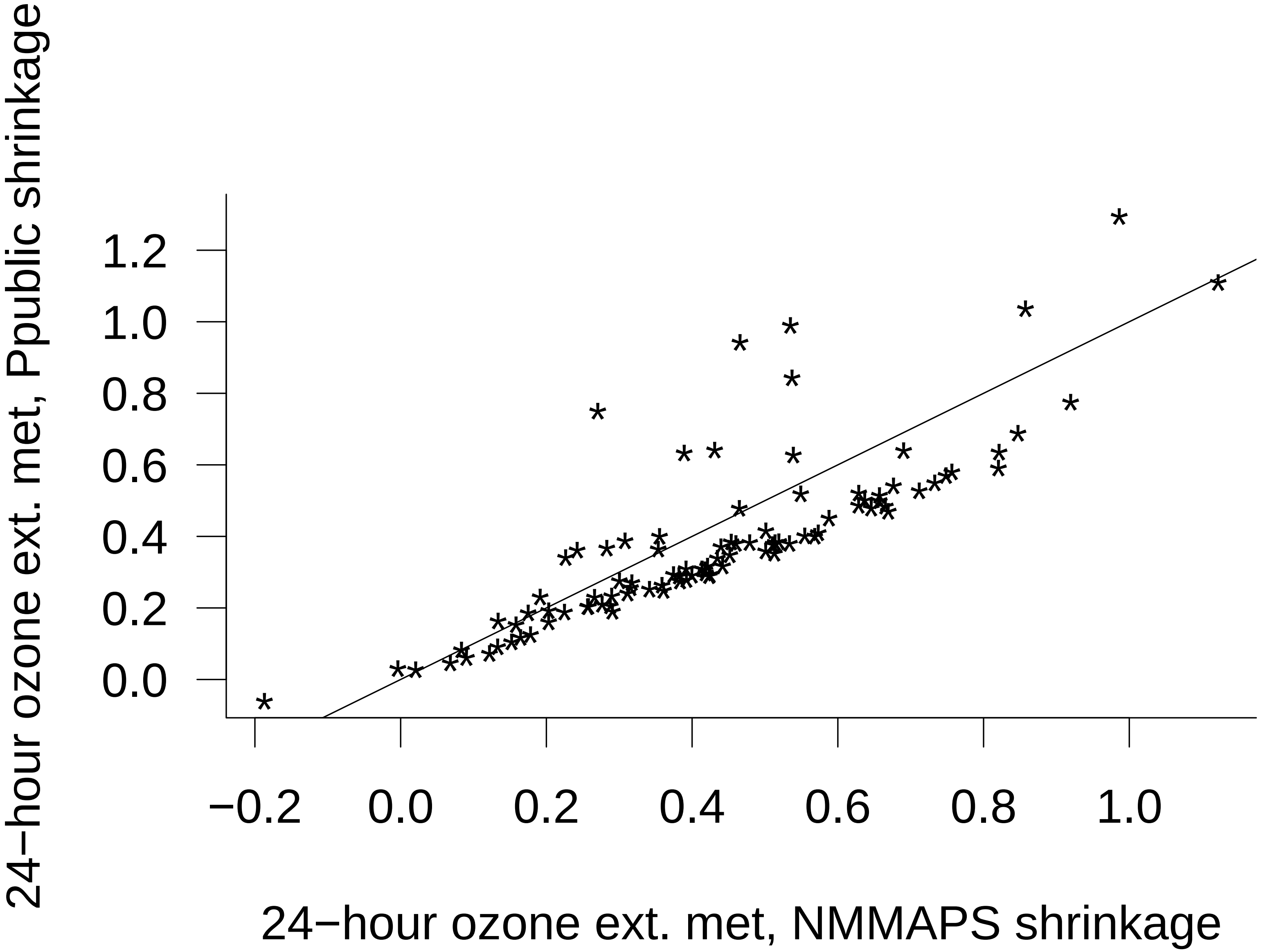
OZONE-MORTALITY COEFFICIENT
24-HR OZONE, REG MET MODEL



If we cannot explain the spatial variability through different constitutents of air pollution, an alternative explanation may be demographics, i.e. the most vulnerable people live in the places with the highest coefficients.

The NMMAPS dataset also contained 77 variables defined at the city level that are derived from the US census. Several of these are correlated with the ozone-mortality coefficient. Two of these are *Ppublic*, defined as "Proportion public transport to work", and *p006003*, the proportion who are solely of black or African-American race. Bell and Dominici (2006) also identified these two variables as the most significant "effect modifiers".

We would caution against overinterpreting any single covariate of this nature. However, such covariates may help to explain the inter-city heterogeneity and to improve the overall precision of the estimates.



24−hour ozone ext. met, Ppublic shrinkage
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
24−hour ozone ext. met, NMMAPS shrinkage

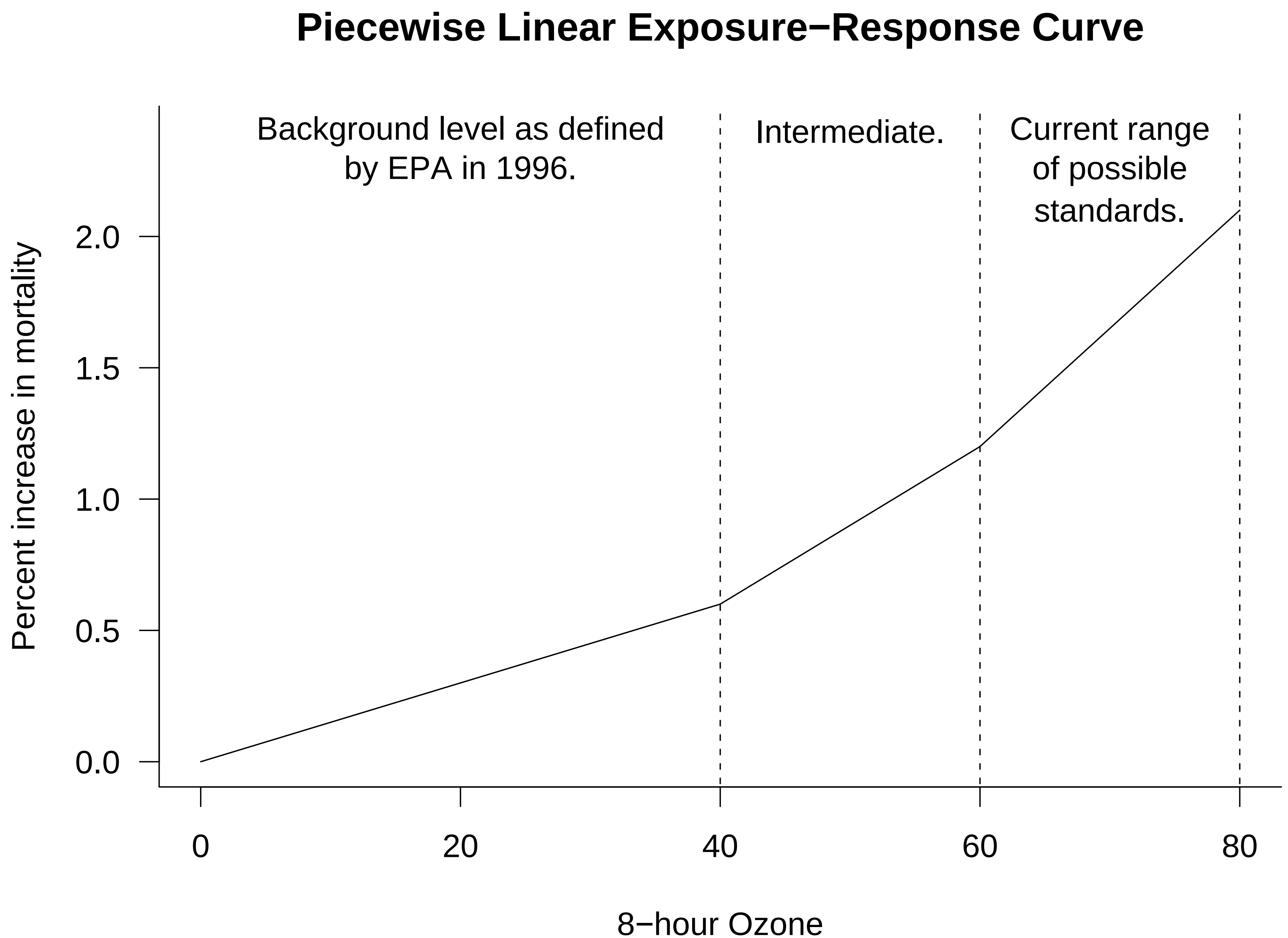
So far, all the analyses have used 24-hour averages as the ozone metric. However, there are two good reasons for looking also at 8-hour averages:

1. The standard is based on 8-hour averages, so results based on this metric should be more directly relevant to the standards-setting process.

2. Exposure-pattern considerations also suggest 8-hour daily max ozone should be more relevant to health effects than 24-hour averages. Therefore, we might expect to see a more clearly defined signal if we use 8-hour ozone. The following results bear this out to some extent.

When using the 8-hour ozone metric, we also want to take into account possible nonlinearity of the exposure-response relationship (as in Bell *et al.* 2006). However there seem to be difficulties in estimating nonlinear relationships, combined across many cities, using standard methods such a kernels and splines.

As an alternative, we take a *piecewise linear* approach, using the average of lags 0 and 1 (as in Bell *et al.* 2006).


Piecewise Linear Exposure−Response Curve
Background level as defined by EPA in 1996.
Intermediate.
Current range of possible standards.
Percent increase in mortality
0.0
0.5
1.0
1.5
2.0
0
20
40
60
80
8−hour Ozone

NATIONAL WEIGHTED AVERAGES IN PIECEWISE LINEAR MODEL, 8-HOUR OZONE

0–40 ppb	40–60 ppb	60–80 ppb
.233 (.085)	.250 (.091)	.359 (.166)

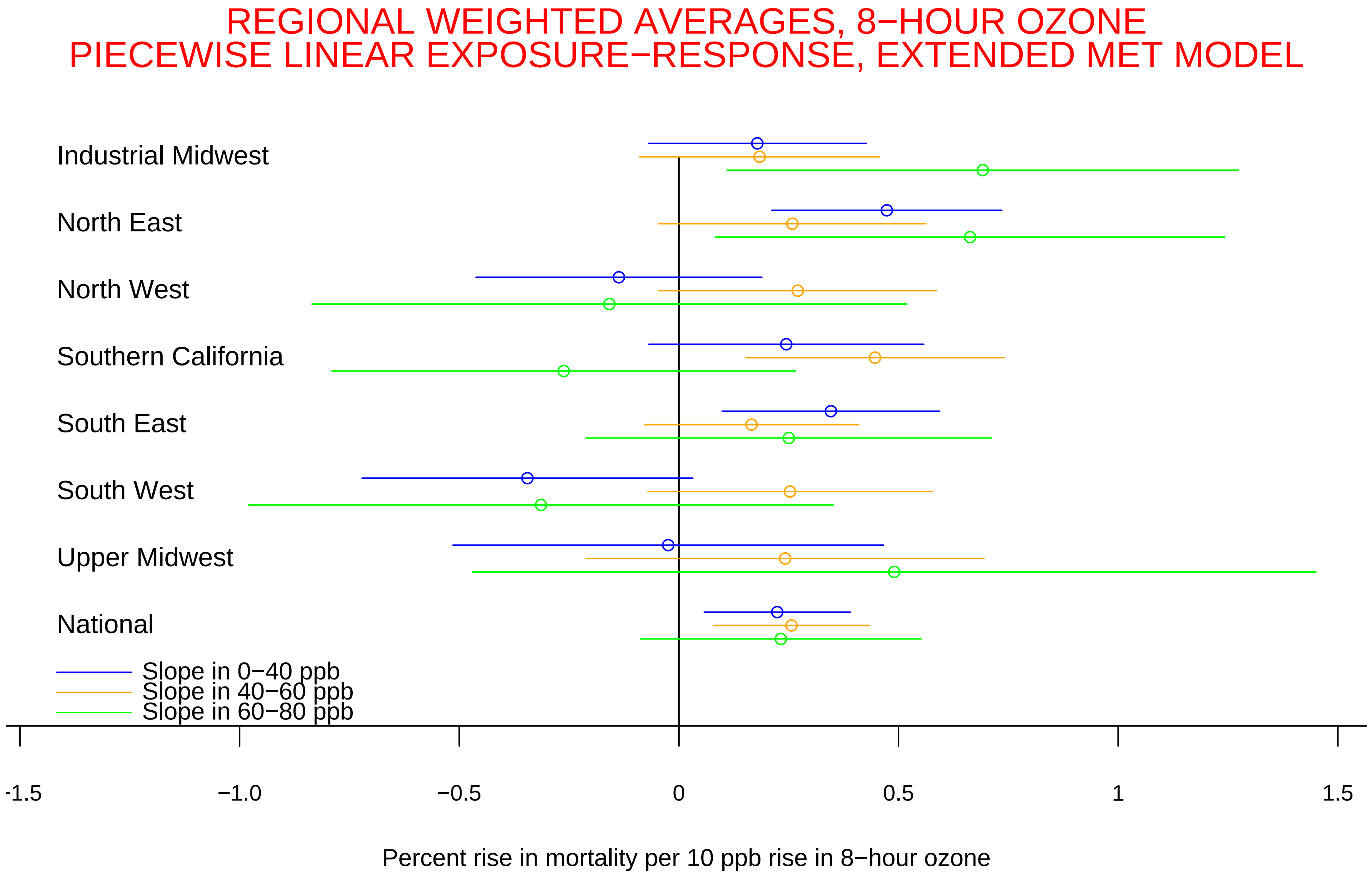
NATIONAL WEIGHTED AVERAGES IN PIECEWISE LINEAR MODEL, 8-HOUR OZONE

0–40 ppb	40–60 ppb	60–80 ppb
.233 (.085)	.250 (.091)	.359 (.166)

Omit July 9-16 1995 in Chicago...

0–40 ppb	40–60 ppb	60–80 ppb
.224 (.085)	.257 (.091)	.232 (.163)

However, we can also look at regional results here.


REGIONAL WEIGHTED AVERAGES, 8−HOUR OZONE
PIECEWISE LINEAR EXPOSURE−RESPONSE, EXTENDED MET MODEL
Industrial Midwest
North East
North West
Southern California
South East
South West
Upper Midwest
National
Slope in 0−40 ppb
Slope in 40−60 ppb
Slope in 60−80 ppb
-1.5
−1.0
−0.5
0
0.5
1
1.5
Percent rise in mortality per 10 ppb rise in 8−hour ozone

The "Chicago heatwave correction" is essential to avoid serious bias in these and the following results.

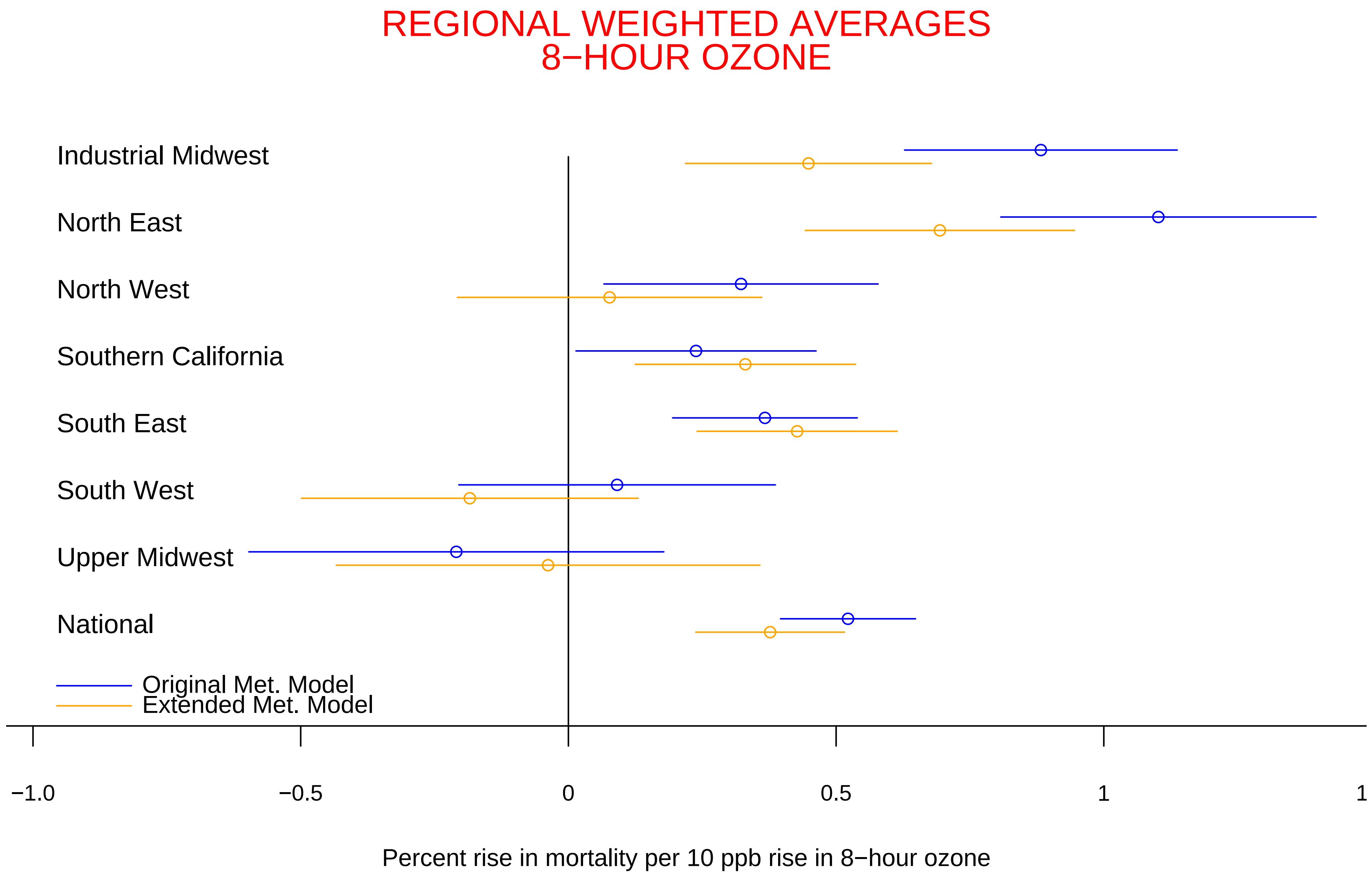
With this correction, we can no longer see a statistically significant result in the 60-80 ppb range for the national result, though we can for two regions.

Based on the national average results, we might also conclude that the three slopes are the same (implying a linear exposure-response curve). However, some of the regional results suggest a different conclusion.

Therefore, for the rest of the present analysis we revert to the linear exposure-response model based on distributed lag ozone, though we don't consider this a satisfactory resolution of the issue.

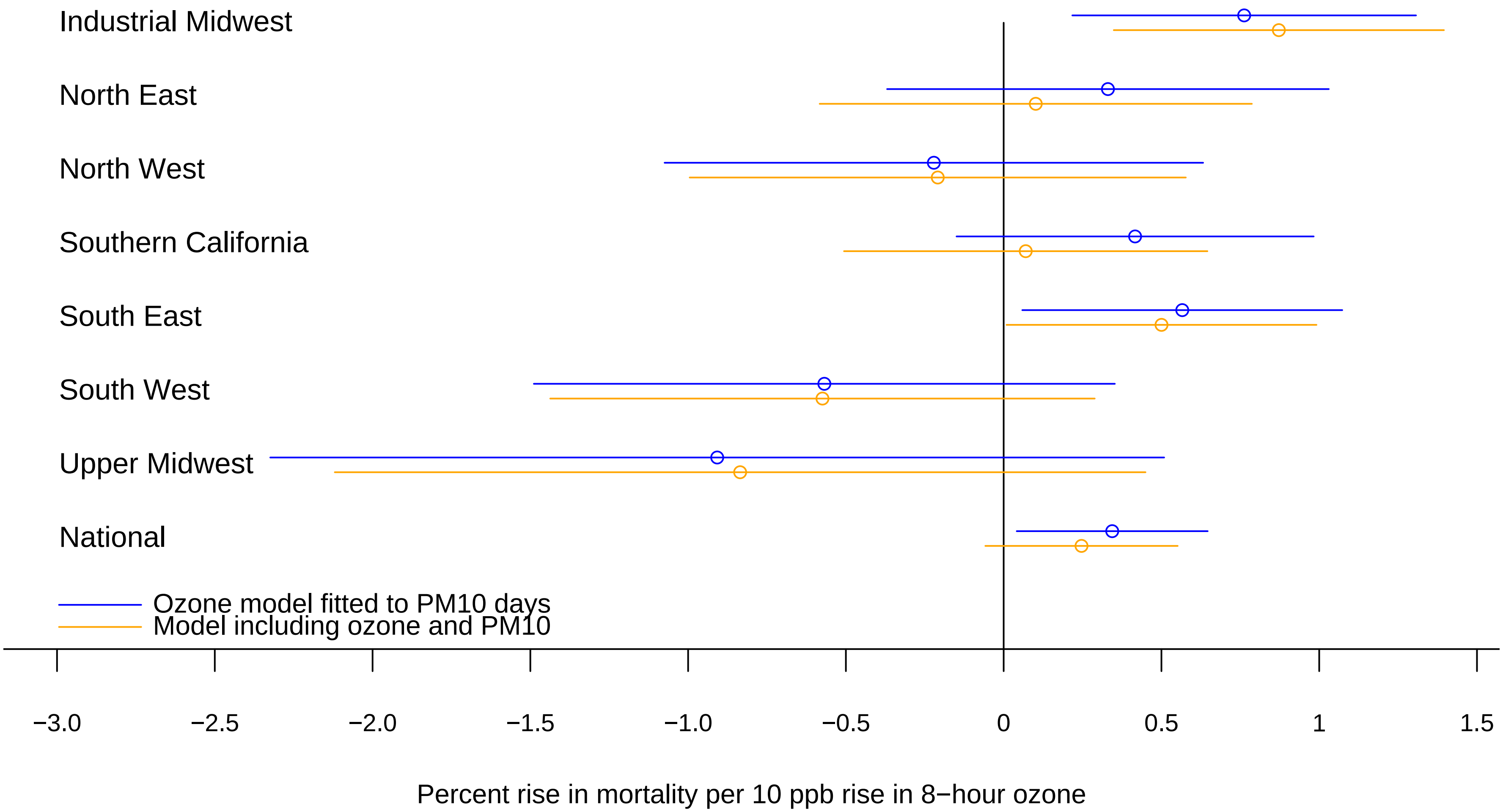
With 8-hour ozone and linear exposure-response curve, we can repeat many of the earlier analyses for 24-hour ozone.

The extended met. model leads to smaller ozone-mortality estimates than the regular met. model, though still clearly significant in four of seven regions, and nationally.

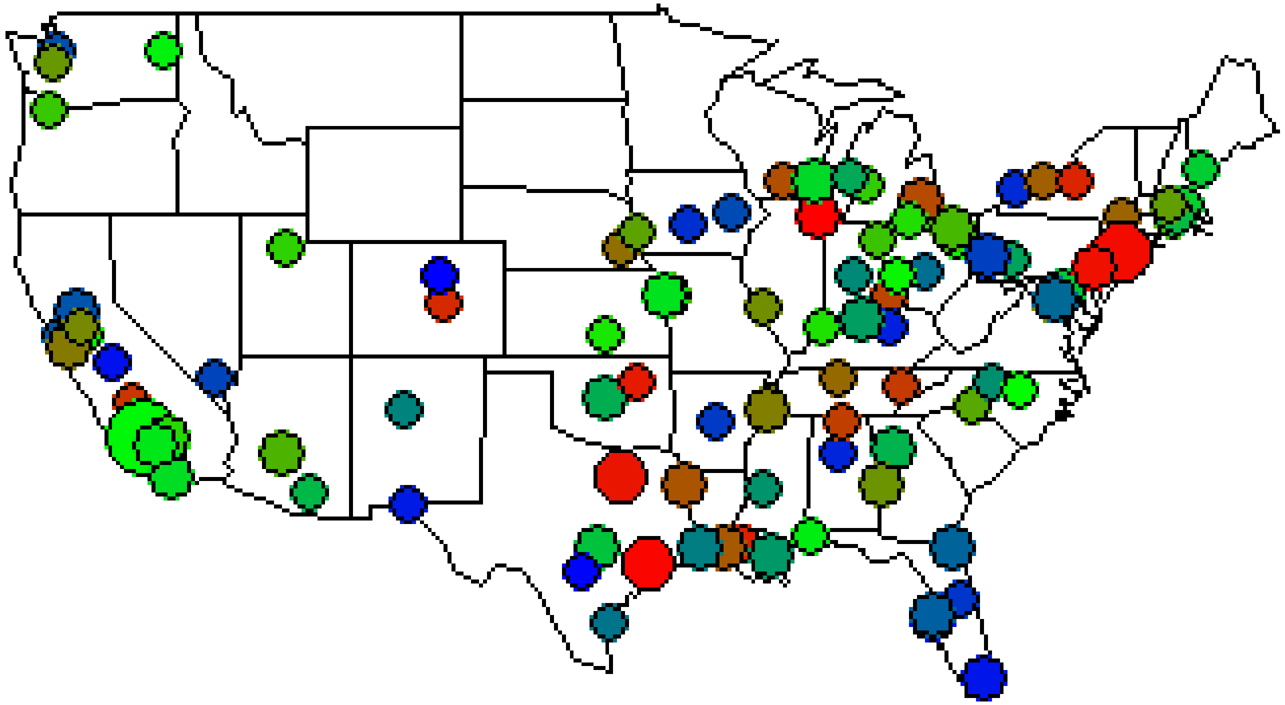

REGIONAL WEIGHTED AVERAGES
8−HOUR OZONE
Industrial Midwest
North East
North West
Southern California
South East
South West
Upper Midwest
National
Original Met. Model
Extended Met. Model
−1.0
−0.5
0
0.5
1
1
Percent rise in mortality per 10 ppb rise in 8−hour ozone

Confounding by PM_{10} may still be an issue..

REGIONAL WEIGHTED AVERAGES
8−HOUR OZONE, EXTENDED MET MODEL

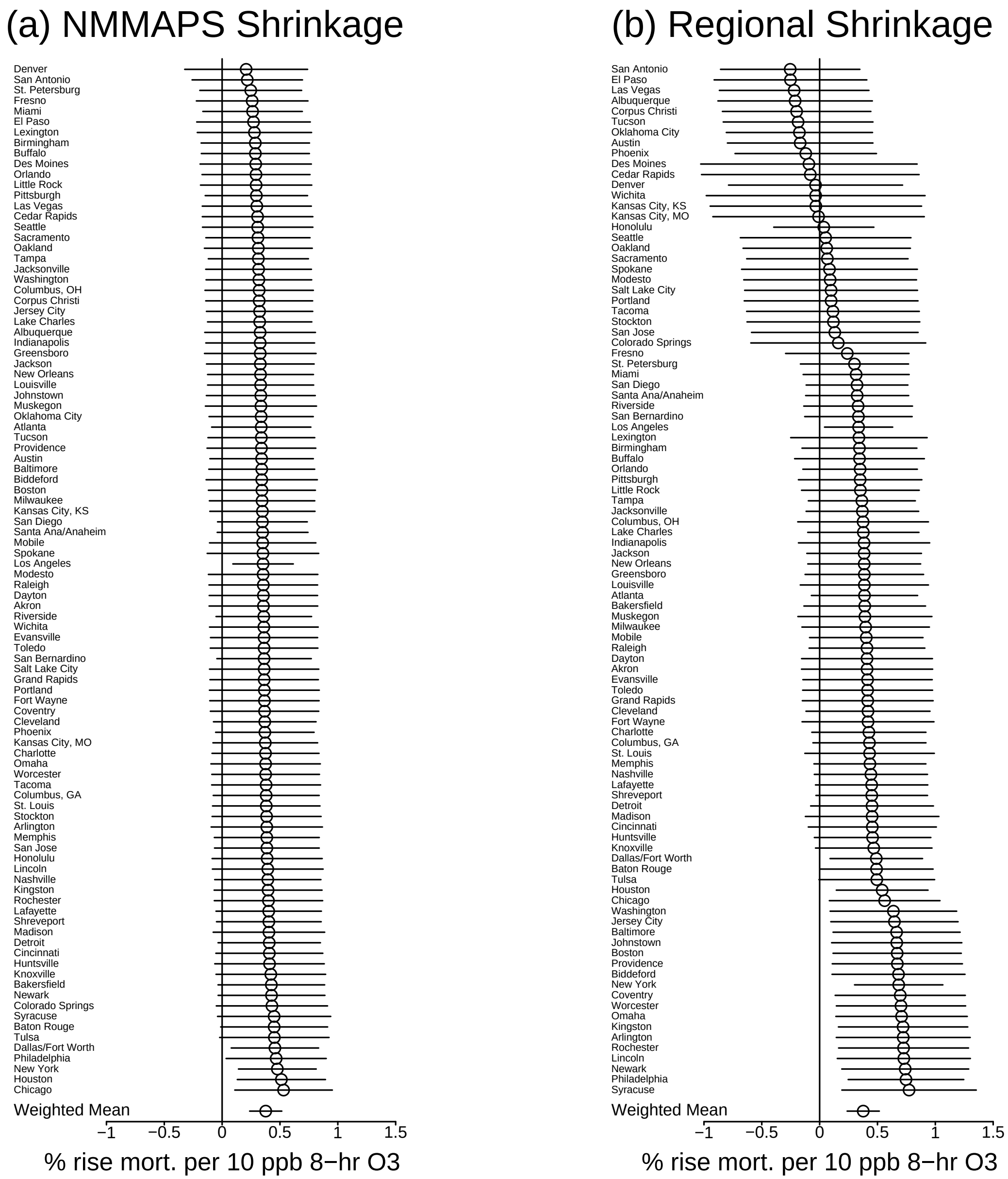


EFFECT ESTIMATES BASED ON 8-HOUR OZONE



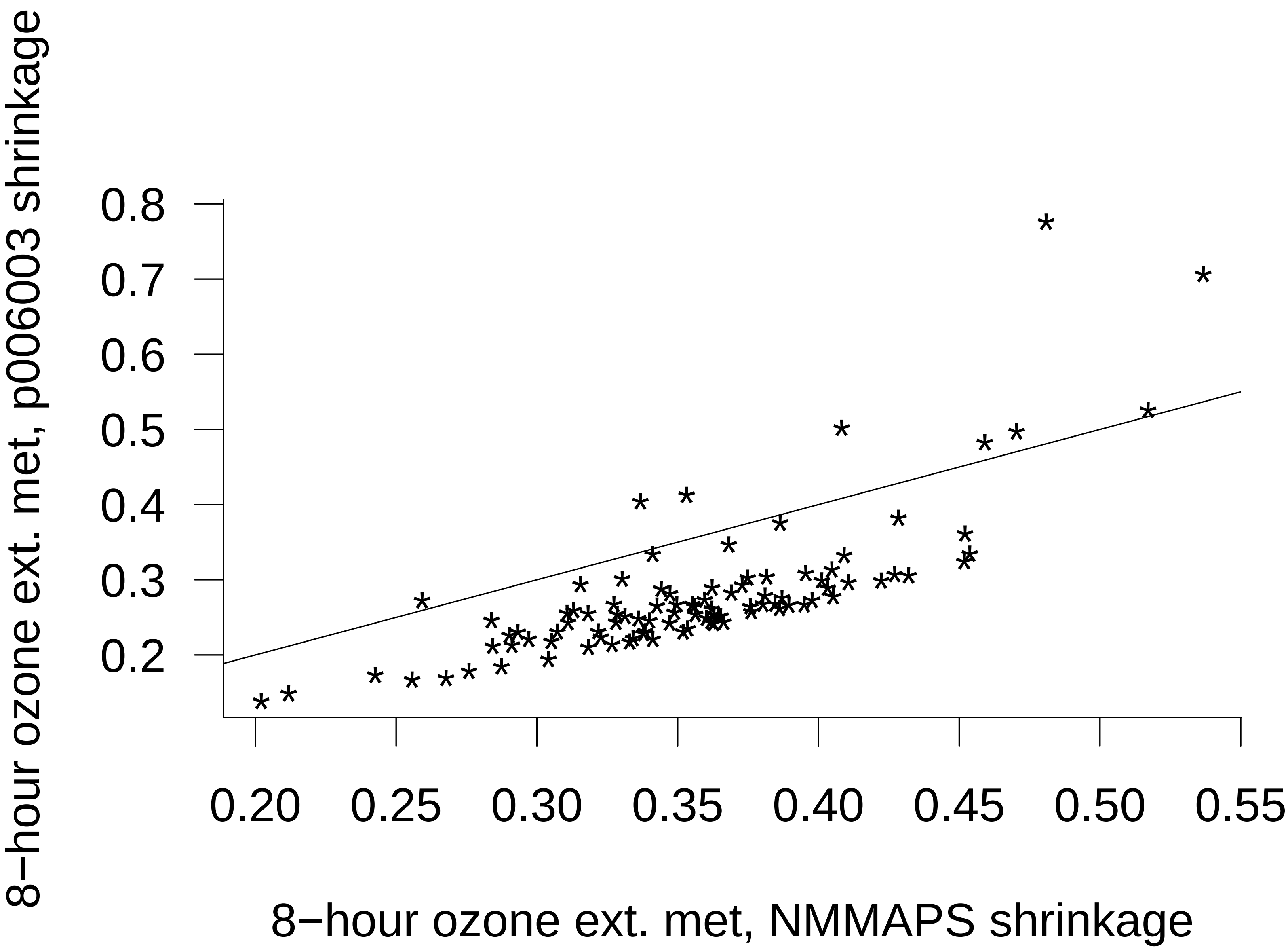
Map of posterior city-by-city ozone-mortality coefficients on an ordinal scale (blue=smallest; green=median; red=largest). Areas of circles are inversely proportional to posterior variances.

OZONE−MORTALITY COEFFICIENTS AND 95% PIs



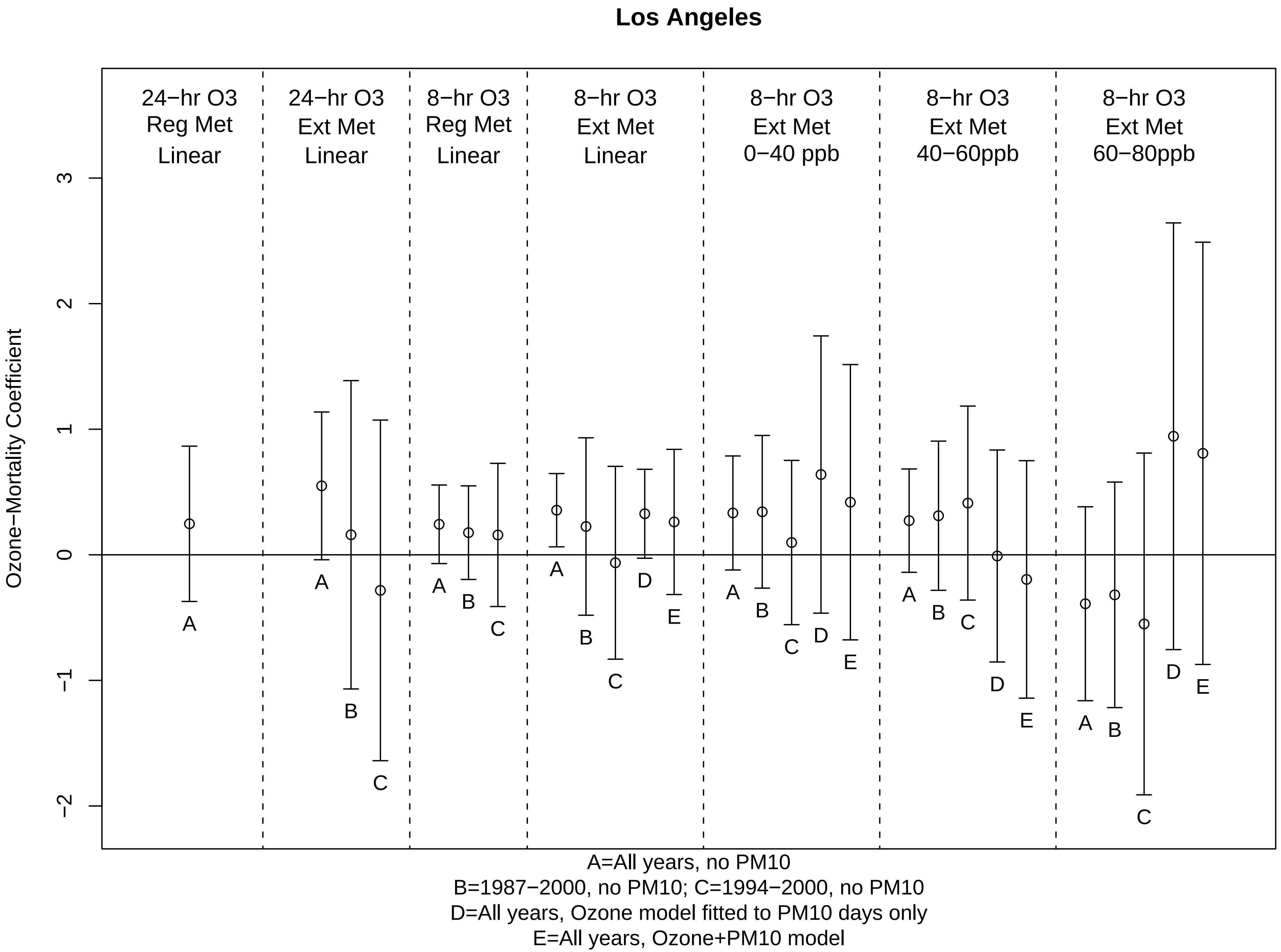
The evidence for regional/spatial variability is less strong than in the 24-hour analyses, though the overall pattern of results is similar (e.g. the North East and Industrial Midwest are still the two regions with the strongest effects).

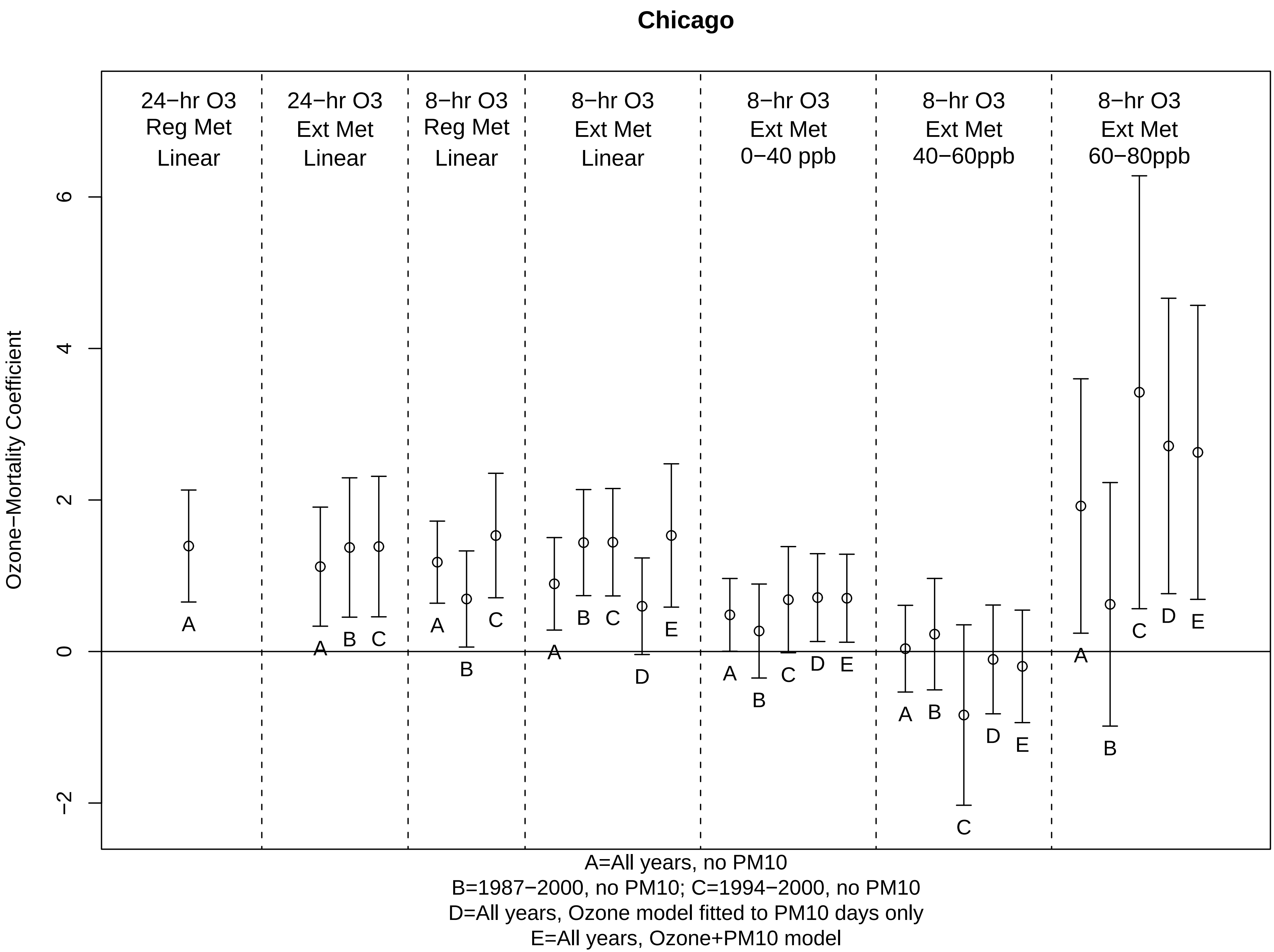
We can also look for a correlation with demographic variables. In this case *Ppublic* is not significant, but *p006003* still is. Once again, we urge caution not to over-interpret this.



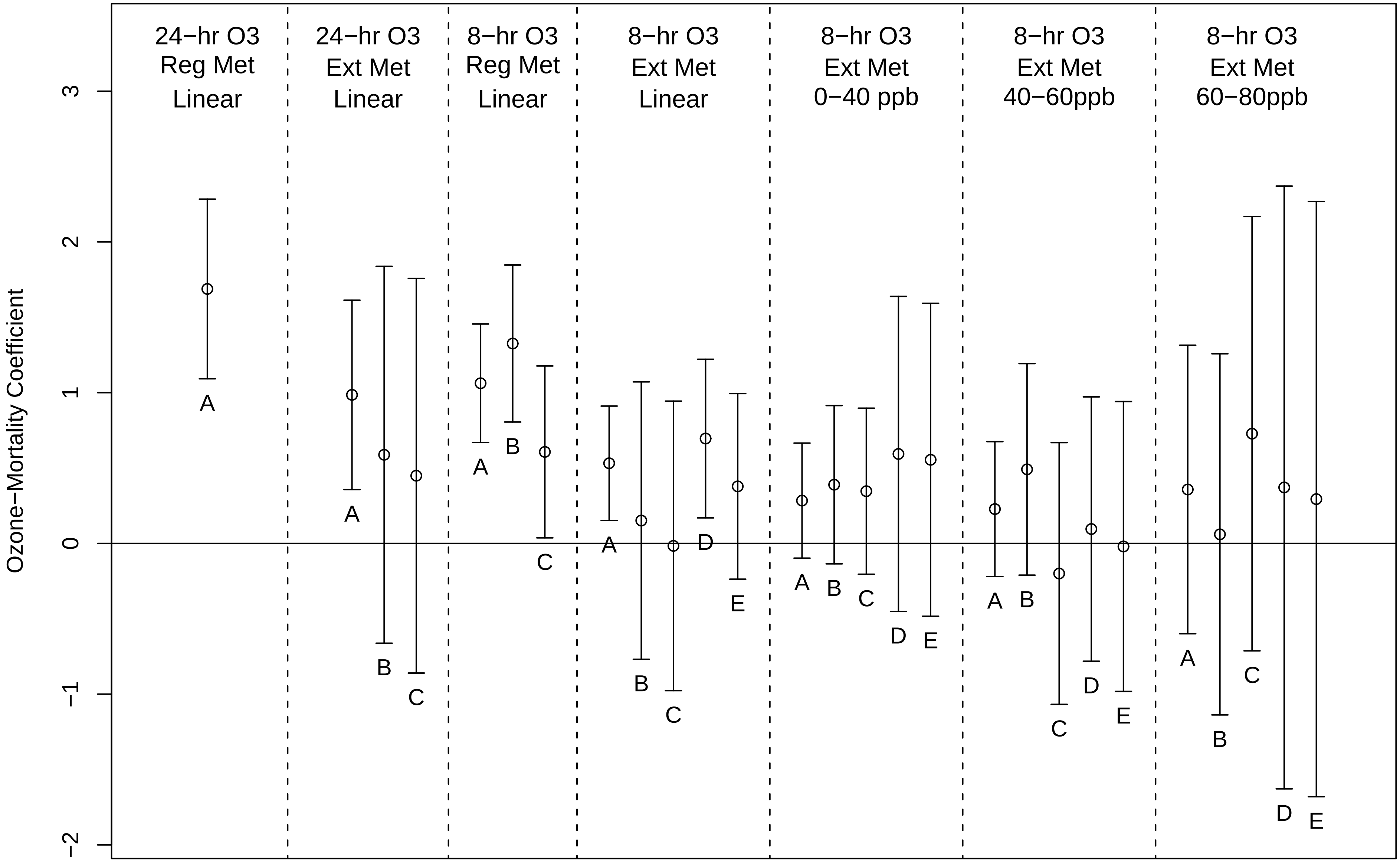
8−hour ozone ext. met, p006003 shrinkage
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
8−hour ozone ext. met, NMMAPS shrinkage

We can also look at individual city results.


Los Angeles
Ozone−Mortality Coefficient
24−hr O3 Reg Met Linear
24−hr O3 Ext Met Linear
8−hr O3 Reg Met Linear
8−hr O3 Ext Met Linear
8−hr O3 Ext Met 0−40 ppb
8−hr O3 Ext Met 40−60ppb
8−hr O3 Ext Met 60−80ppb
A=All years, no PM10
B=1987−2000, no PM10; C=1994−2000, no PM10
D=All years, Ozone model fitted to PM10 days only
E=All years, Ozone+PM10 model

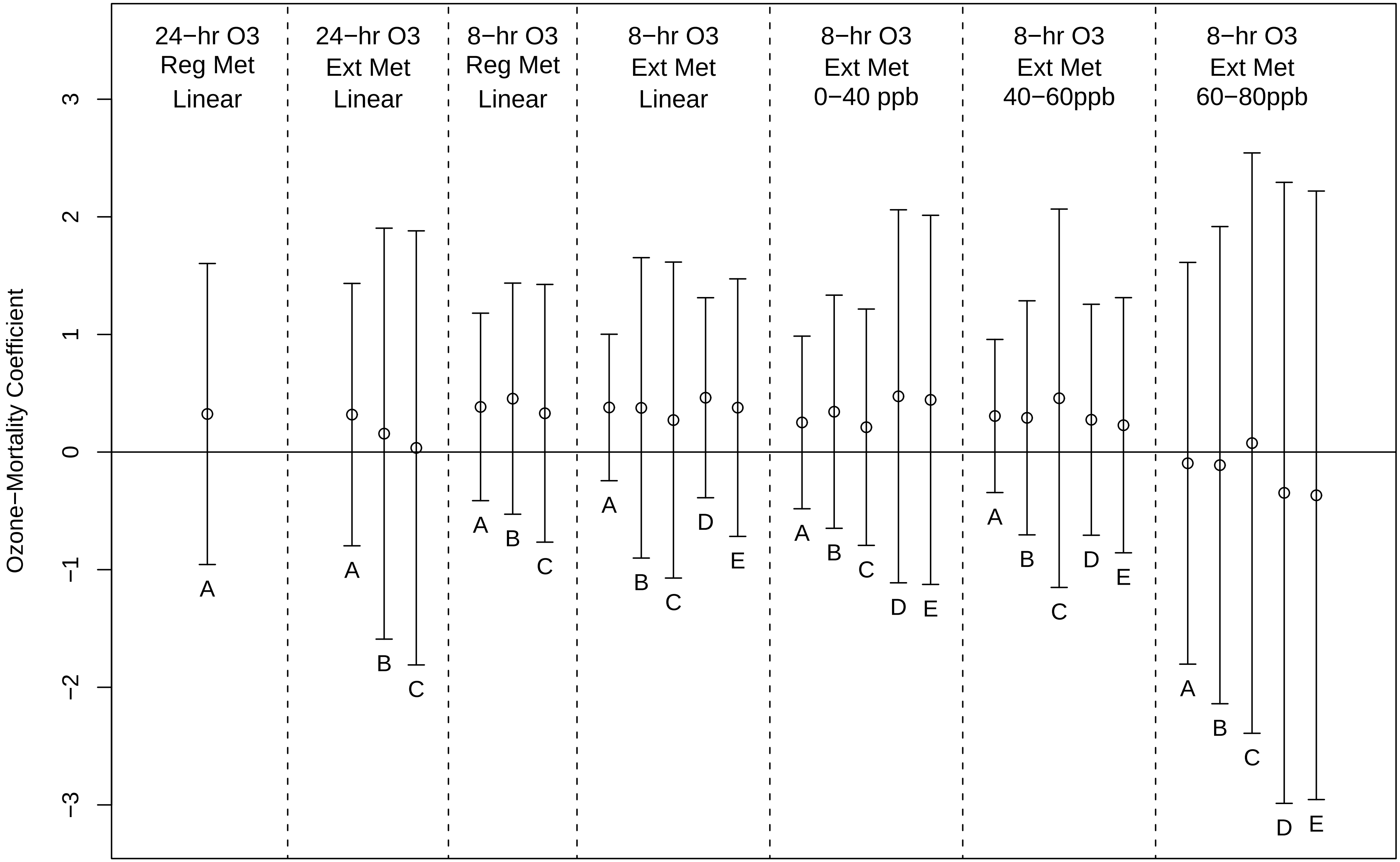

Chicago
24−hr O3 Reg Met Linear
24−hr O3 Ext Met Linear
8−hr O3 Reg Met Linear
8−hr O3 Ext Met Linear
8−hr O3 Ext Met 0−40 ppb
8−hr O3 Ext Met 40−60ppb
8−hr O3 Ext Met 60−80ppb
Ozone−Mortality Coefficient
−2
0
2
4
6
A
A B C
A B C
A B C D E
A B C D E
A B C D E
A B C D E
A=All years, no PM10
B=1987−2000, no PM10; C=1994−2000, no PM10
D=All years, Ozone model fitted to PM10 days only
E=All years, Ozone+PM10 model

New York



A=All years, no PM10
B=1987−2000, no PM10; C=1994−2000, no PM10
D=All years, Ozone model fitted to PM10 days only
E=All years, Ozone+PM10 model

Salt Lake City



A=All years, no PM10
B=1987−2000, no PM10; C=1994−2000, no PM10
D=All years, Ozone model fitted to PM10 days only
E=All years, Ozone+PM10 model

SUMMARY AND CONCLUSIONS

1. The results based on 24-hour ozone show a number of difficulties, such as sensitivity to different meteorological models, to the inclusion of PM_{10}, and especially, spatial variability of the ozone-mortality coefficient, which may be due to lifestyle/exposure issues.

2. Results based on the 8-hour ozone metric with the "extended meteorology" are more uniform, provided we make the "Chicago heatwave correction", but still show some spatial variability or dependence on demographic covariates.

3. Taken altogether, the results imply that the ozone-mortality effect is concentrated in large cities in the north and east, such as New York, Philadelphia and Chicago, and to a lesser extent Houston and Dallas (and possibly Los Angeles). There is no evidence of any effect in places such as Salt Lake City, Denver and Albuquerque, though all of these will be in violation of the new EPA standard if implemented as proposed.

4. However, all these results rely on the linear ozone-response curve and we still feel this issue is unresolved. Preliminary results based on the piecewise linear model suggest further regional variability which needs to be explored further.

5. Future EPA risk assessments should take account of these and other sensitivities in epidemiological analyses.