Probability Analysis for the BAAQMD Multi-Pollutant Evaluation Method

David Fairley, BAAQMD

Bay Area Air Quality Management District
939 Ellis Street
San Francisco, CA 94109

March, 2010

TABLE OF CONTENTS

TABLES AND FIGURES

PROBABILITY ANALYSIS FOR BAAQMD MULTI-POLLUTANT EVALUATION METHOD

Abstract

The Bay Area Air Quality Management District (District) has prepared the Bay Area 2009 Clean Air Plan (CAP) to address four types of pollutants: ground-level ozone, particulate matter, air toxics, and greenhouse gases (GHGs). The District developed a multi-pollutant evaluation method (MPEM) to analyze and rank potential emission control measures on a multi-pollutant basis for the 2009 CAP. The MPEM has a number of steps and assumptions, with some degree of uncertainty related to each.

This document analyzes the uncertainty in the results and develops a probabilistic model for each step of the MPEM calculations. A Monte Carlo simulation based on the probablistic models is run to estimate the distribution of uncertainty in the valuation of each of the emission control measures. These distributions are combined to yield the distribution of uncertainties for the plan as a whole and to rank the control measures.

Brief Overview of Multi-Pollutant Evaluation Method

The District has prepared the Bay Area 2009 Clean Air Plan (CAP) to update its current ozone plan (the 2005 Ozone Strategy), as required by the California Health & Safety Code. The District decided to include the analysis and evaluation of three additional pollutants – particulate matter (PM), air toxics and green house gasses (GHGs) – to maximize reductions in the four types of pollutants. These pollutants differ in fundamental ways in terms of their emission sources, atmospheric formation, chemical composition and health effects.

The District developed a multi-pollutant evaluation method (MPEM) to analyze and rank potential emission control measures on a multi-pollutant basis for the 2009 CAP. The MPEM is fully described in a June 2009 *Multi-Pollutant Evaluation Method Technical Document* which is available on the District's website: www.BAAQMD.gov and briefly summarized below:

The purpose of the MPEM:

- Estimate how reductions of each pollutant for a given control measure will affect ambient concentrations, population exposures, and health outcomes related to that pollutant.
- Monetize the value of total health benefits and greenhouse gas reductions for all pollutants that would be reduced by each potential control measure.
- Evaluate and rank the estimated benefit of potential control measures based on the value of each measure in reducing health costs from air pollutants and environmental/social impact related to climate change.

The MPEM involves several steps:

Step 1. Emissions: Estimate how much a given control measure reduces (or increase) emissions of each of the pollutants.

Step 2. Concentrations: Estimate how a change in emissions of each pollutant affects its ambient concentrations and other pollutants related to it, i.e., the response of ambient concentrations to emission reductions.
Step 3. Population Exposure: Estimate how a change in ambient concentrations affects the exposure of Bay Area residents to each pollutant.
Step 4. Health Impacts: Estimate how a reduction in population exposure impacts various health endpoints, projecting changes in the incidence of endpoints such as asthma emergency room visits, lower respiratory symptoms, and deaths.
Step 5. Health/Social Benefits: Estimate avoided costs of each emission control measure by estimating the cost of the health and climate impacts from each pollutant. For each pollutant except GHGs, the change in the number of incidents of each health endpoint the pollutant affects is multiplied by an estimate of the per-incident social cost of that endpoint. For GHGs, the change in tons of GHG emissions is multiplied by the estimated social cost per ton of GHGs.

The output of the MPEM (Steps 1-5) is an estimated dollar value of the health and social benefits of each potential control measure, based on the amount of reduction (or increase) in each pollutant.

Description of Probability Analysis

Some uncertainty is involved with each step above. But the uncertainties can be quantified then combined in a Monte Carlo simulation. In this way we can simulate the overall uncertainty of the process and obtain an estimate of the likely range of benefits of each control measure.

Section 0 discusses types of uncertainties encountered and statistical assumptions and methods used in the uncertainty analysis. Sections 1-5 present how uncertainties were calculated for steps 1-5 of the MPEM. Section 6 presents examples and the uncertainty of the overall plan benefits.

Some of the MPEM incorporates the uncertainty methodology of other researchers and some of it is new. There has been considerable work done to evaluate the uncertainties in the last two steps of the MPEM – the health effects and monetary valuations. (See, e.g., BenMAP Appendices F-I, Hall *et al.* 2006, or Ostro *et al.* 2006.) We have followed these precedents, where they exist. For the uncertainties in the first three stages on the MPEM – emissions estimates, emissions to concentrations, and population – we developed our own probability distributions.

0. Statistical Considerations

There are a number of statistical considerations: What is meant by "uncertainty"? Which components are statistically independent? How should uncertainty distributions be chosen? How should overall uncertainty be summarized?

0.1 Statistical definitions and setup

We use the Monte Carlo method to gain an understanding of uncertainty in the estimated benefits (or disbenefits) of a control measure by randomly simulating each of the steps. The MPEM has many *parameters* – emissions to concentrations conversion factors, population for each grid square, air pollutant health impact functions, health effect incidence rates, and monetary valuations. When we apply the MPEM, we use *point estimates* (that is, our best estimates) for each of these parameters. But, as discussed below, these parameters are known to various levels of precision, i.e., certainty. The parameter uncertainties can be captured with probability distributions. That is, the point estimate might be set as the mean of the probability distribution, with the distribution representing the range of likely values for this parameter.

For the Monte Carlo analysis, we randomly sample parameter values from each of these distributions and use these simulated values in the MPEM rather than the point estimates.

The simulation of a set of parameters followed by their application in the MPEM is called a *realization* of the Monte Carlo simulation for a given emission control measure. Each realization results in a "bottom line" dollar benefit (or disbenefit) estimate for the measure. Thus, repeated Monte Carlo realizations lead to a probability distribution of a measure's overall dollar benefits.

0.2 Uncertainty categories

The various statistical uncertainties in the MPEM can be divided into three categories:

- uncertainties within our probability model
- uncertainties in data inputs
- uncertainties in assumptions

Uncertainties in the model: The first category includes the uncertainties within our choice of probability model. For example, many of the health effects estimates derive from epidemiological studies where the estimate is the coefficient from a regression. Typically, the estimate and its standard error are published. We assume, based on statistical theory,[1] that the coefficient has a normal distribution with the estimate as its

[1] In linear regression, the regression coefficient estimates are linear combinations of the observations. With some regularity conditions, the Central Limit Theorem implies that these estimates are approximately normally (Gaussian) distributed. In the general linear models, the Gaussian distribution is a reasonable first order approximation to the true distribution. We might note that the degree to which this is not true is an example of the third category of uncertainties.

mean and the estimate's standard error as its standard deviation. We can term this distribution *within-model uncertainty*, that is, the uncertainty given a particular set of assumptions about the distribution of the data.

Other examples include many of the incidence rates. For example, we can use the past few years of hospitalization counts to project future counts assuming the year-to-year counts are independent, that the system is steady-state (without trend) and the counts are normally distributed. The uncertainty in this projection can be assumed to be captured by the year-to-year variability observed in the counts over the past few years. Another area where we have within-model uncertainties are the valuation distributions we've incorporated from EPA's BenMAP.

Uncertainties in the data: In the second category are uncertainties in the emissions reductions from many of the proposed control measures, uncertainties in the total Bay Area emission estimates, uncertainties in the population projections, and uncertainties in some of the incidence rates. In the absence of knowledge of actual uncertainties, we've assumed the values have a Gaussian error with mean 0 and standard deviation of 10%. We're calling these *placeholder uncertainties* in that they serve as an acknowledgement that some uncertainty exists but await a better characterization of that uncertainty.

Uncertainties in assumptions: These include some of the key assumptions of the MPEM, such as the non-existence of a $PM_{2.5}$ threshold, the 50 ppb ozone threshold; the assumption of 100% certainty that the correlation between $PM_{2.5}$ and mortality and between ozone and mortality are causal; the assumption of 100% certainty that global warming won't have catastrophic consequences; the assumption that people are in their backyards 24/7.

We deal with these categories differently. The first category is clear-cut, with a known or reasonably assumed probability model that can be simulated. For the second category, all we know is that we don't know the values for certain, except that we believe that the deviations could be substantial. For MPEM values in this category, we simulate a modest Gaussian deviation of 10%. For the third category we do sensitivity analyses, where we alter the assumptions and rerun the Monte Carlo analysis.

0.3 Statistical Independence

It appears reasonable to assume that the uncertainties in the variables being estimated at a step of the MPEM are statistically independent of the uncertainties at the other steps: whether ABAG over- or under-estimated Bay Area population has nothing to do with whether a health effects study over- or under-estimated the individual risk from a given health effect; and these have nothing to do with whether an economist over- or under-estimated the amount that a case of a given health effect is valued. And so on.

It also seems reasonable to assume independence within some of the stages. For example, most of the health effects estimates were produced in different studies and different populations. When the estimates do come from the same study, such as the

estimated rates of cardiovascular-related hospitalizations among 65+ and 18-64 year-olds, it seems likely that they would be positively associated, suggesting that assuming independence would lead to a less variable result. We have, in fact, ignored this source of dependence, thereby underestimating the uncertainty to some extent.

Valuations also appear independent, although the methodology might result in systematic over- or under-estimation which would fall in the third category of uncertainty. There may be some weak dependence between incidence of various effects in the Bay Area, but we will assume independence.

It is likely, however, that spatial parameter estimates will be statistically dependent. For example, if the estimated concentration in a particular grid cell is higher than the actual concentration, then it's likely that the estimated concentrations in neighboring cells will also be higher than the actual. The same argument applies for the relation between estimated and actual population exposures.

We deal with spatial correlation in a couple of simple ways. For the population projections, it seemed reasonable that certain errors, such as underestimating or overestimating economic growth, would lead to systematic under- or over-estimates of the population. For this case, we applied the same simulated error to each population value. See Section 3 for details.

Another instance of spatial correlation was localized errors in the model, e.g., over- or under-estimating wind speed in an area leading to under- or over-estimated $PM_{2.5}$ concentrations. For this we simulated clumps of grid cells – rectangles where we applied the same random simulated percentage error. See Section 2.

0.4 MPEM goal: Estimate projected benefits or disbenefits

The emission control measures proposed in the CAP will not take effect immediately; they must be adopted, implemented, and enforced via a variety of mechanisms. Thus, we are considering future effects, that is, the MPEM estimates are predictions or projections to future years. Thus, each aspect of the calculation is a projection: what the control measure emissions changes will be, what the total District-wide emissions will be, what the population will be, what the incidence rates will be, what the value of the dollar will be.

0.5 Simulation dynamics and comparison between control measures

For both computational reasons and to facilitate comparisons between control measures, the simulation was performed as follows: For each parameter, a set of 1,000 values was simulated according to the appropriate distribution, resulting in 1,000 vectors of simulated parameter values.

For each control measure, the same sets of random values were used (with the exception of the uncertainties related to an individual control measure's emissions). Thus, for every

control measure, the ith simulation would use the same random population figures, the same simulated betas, the same simulated valuations and so on. Effectively, with this approach, each simulated set of values is the *same* for every control measure being evaluated. In the extreme case where one control measure is a scalar multiple of another – e.g., 2 NOx-only control measures – then one set of simulated values will be nearly a scalar multiple of the other set; e.g., if one control measure reduced NOx by an estimated 50 tons/year and the other reduced NOx by 75 tons/year, then the latter's simulated values would be 1.5 x the former's simulated values, except for the relatively small errors associated in the estimates of the control measures' NOx reductions.

In essence, for each simulation, we are creating a new set of MPEM parameters then applying that new set of parameters to all the control measures. This means that, in practice, if any given pair of control measures is compared, their simulated values will be (possibly highly) positively correlated. Thus, even if there is considerable uncertainty in the "bottom line" for each of the control measures, it may be possible to determine that one's "bottom line" is significantly greater than the other's. Section 6 shows that this results in being able to distinguish among the overall valuations of the control measures in most cases.

1. Simulating Emissions

Emissions estimates for a given pollutant are used in the MPEM as a ratio of a control measure's estimated change in the emissions of that pollutant to the District total emissions for that pollutant. We assume a Gaussian error of 10%, e.g., suppose that a control measure reduces ROG by r tons per day. Then its uncertainty would be simulated by

$r * (1 + 0.1 Z)$,

where Z is a standard normal pseudo-random variable.

We don't have estimates for the overall uncertainty in District emissions totals, so we've assumed a 10% Gaussian error for these also. Thus, to continue the example, suppose a total of R tons/day of ROG is emitted in the District. Then the fraction of ROG reduced by the control measure would be simulated by:

$$\frac{r(1+0.1Z)}{R(1+0.1Y)}$$

where Y is a standard normal pseudo-random variable.

2. Simulating Concentrations

The MPEM includes a grid of factors for estimating pollutant concentrations from the emissions of each precursor pollutant. These factors are derived from runs of grid-based pollution models. The evidence available for evaluating the precision of these factors is a

set of observations collected at a limited number of points sometimes on a limited number of days.

We can look at the differences between these observations and the modeled values for the grid cells where the observations were collected, but in using this to evaluate model uncertainty we need to keep two points in mind:

i. Not only are the observations measured with error but, probably more significantly, an observation represents an individual point whereas a grid represents an area of 1 to several square kilometers. The true variation in precursor concentrations across the grid may be large, so that some of the discrepancy between observation and model may be because the pollutant concentrations at the observed point may be unrepresentative of the whole grid.

ii. The error in the model may result from errors in the underlying emissions inventory estimates. But this error is already accounted for in 1.

Nevertheless, the model is subject to other errors, including errors in the wind field being used, and errors in the chemistry. There is also a substantial uncertainty in extrapolating from the modeled period to the entire year. In what follows, we attempt to evaluate the magnitude of the errors in i. and use this to better model the remaining model concentration error.

2.1 Evaluating concentration uncertainties for the PM2.5 model

To evaluate model error, we considered the potential error in the observed values. An analysis of the difference in the means of measured $PM_{2.5}$ during the modeled period suggests that, if the monitoring site is considered as representing the mean $PM_{2.5}$ concentration, then it's deviation from the mean of the surrounding grid is approximately normal with standard deviation 1.7 $\mu g/m^3$ (See Appendix A).

Figure 2.1 shows the modeled vs. measured mean $PM_{2.5}$ from Bay Area and Sacramento area $PM_{2.5}$ monitors. The 95% confidence intervals show that, in the majority of cases,

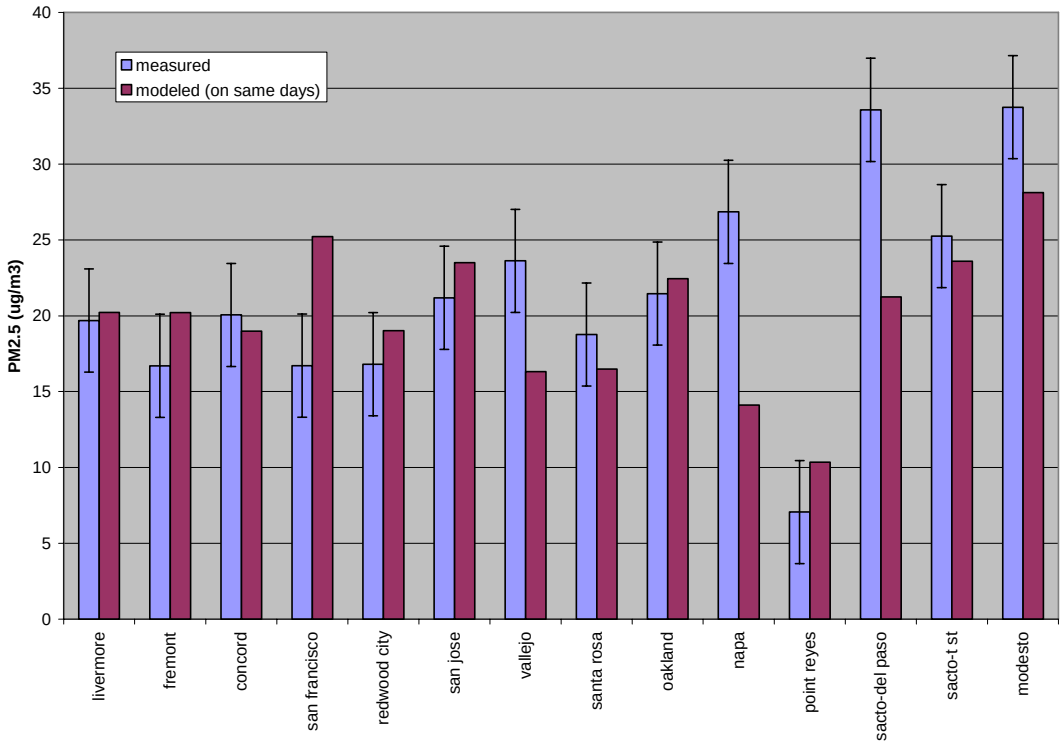


Figure 2.1. Mean measured $PM_{2.5}$ for December, 2006-January, 2007 compared with the 4x4 grid cell containing it. Also shown are 95% confidence intervals for the measured values.

the modeled values are within or close to the range of the ambient data that one would expect just by chance, i.e., there is little or no evidence of model error. For San Francisco, our analysis suggests that the emissions inventory may have overestimated off-road emissions, assigning construction emissions not to the construction site but to the office of the construction company. Seeing as this is an emissions error, we may discount it.

For the other sites with large discrepancies – Napa, Vallejo, Sacramento-Del Paso, & Modesto – the problem again may be emissions. Comparing hourly modeled PM vs. ambient, we noted patterns at several sites suggestive of woodburning, with large ambient peaks in the evening and night not matched in the model, but reasonable model-ambient agreement during the daylight hours. Thus, these sites also may be impacted by emissions.

On the other hand, the pattern of errors may also indicate that the meteorology has not been well-captured. Specifically, the lower modeled values in the Sacramento/Modesto area may derive from overestimated wind speeds. The meteorological modeling for the Napa/Vallejo area may also not adequately reflect the micro-climate of the area.

2.1.1 Modeling concentration errors

The limited analysis above suggests that overall, the modeled $PM_{2.5}$ concentrations are not substantially biased and that where substantial errors exist, they are clumped into localized areas. To simulate this, we decided on a combination of a set of individual, independent 10% errors applied to each grid cell, and clumped errors of a random size occurring with low probability in randomly selected locations. The concentration, c, for a given grid cell, was simulated as:

$$c = c_0 *(1 + .1Z + .2X) \quad (1)$$

where c_0 is the original modeled concentration for the cell, Z is a standard normal random variable, and X is the summation of clumped values landing on the grid square, if any. The clumped values were simulated as follows. For every grid cell, a clump was started with probability 0.02. The clump was grown (N, S, E, W) with probability 0.4 in each direction. In other words, in each direction, a sequence of coin flips was performed with probability 0.4 of heads, until tails occurred. The clump was extended in that direction by that number of heads. The clump was defined as the rectangle of cells extending to the number of heads in each direction. With this choice of probabilities, about 10% of the cells are covered by clumps. The average clump area is 5.4 cells. (See Appendix B.)

As with individual cells, the clump is multiplied by an independent standard normal random variable. As shown in formula (1), these values are multiplied by 0.2, so that we assume the clumps have a standard deviation of 20%.

Lacking a similar analysis for other pollutants, we've adopted the identical error structure for them. We have modeled the errors as independent sets between different pollutants although it's likely that there is perhaps substantial correlation in some cases. For example, benzene and 1,3-butadiene are both largely emitted from motor vehicles. Thus, if meteorological modeling causes an error in one, then it's likely to have a similar effect on the other.

2.1.2 NOx and ammonium nitrate

The MPEM treats the impact of NOx on ammonium nitrate differently, assuming the same scalar conversion factor applies everywhere. There is considerable uncertainty both in the factor itself and also the degree that the true factor varies spatially. Thus, there are likely both District-wide errors and local errors. To account for these possibilities, we've modeled this uncertainty as:

$$c = c_0 *(1 + .2Z + .2D) \quad (2)$$

where Z is a standard normal random variable picked independently for each grid cell, as in equation (1), and D is a standard normal random variable that is the same for all grid

cells. In other words, we assume a 20% uncertainty for each grid cell, but an additional 20% District-wide uncertainty.

3. Simulating population uncertainty

Population data are projections made by the Association of Bay Area Governments. Information about the uncertainties in these projections is not available, so we fall back on using placeholder uncertainties. As in equation (2) we use a combination of individual uncertainties and a global uncertainty with the idea that an error in the projection assumptions could be Bay Area-wide (e.g. a regional economic forecast). So we use the following formula to simulate the population, p, in a census tract:

$$p = p_0*(1 + .1Z + .1D) \quad (3)$$

where p_0 is the population ABAG projects for that census tract. For a given MPEM simulation run, these same simulated population values are used everywhere. In other words, we don't use one simulated population for estimating $PM_{2.5}$ exposures and another for ozone exposures.

4. Simulating health effects and incidence

The evaluation of each health effect uses an impact function with a parameter, beta, that has an estimated value, b, that converts a change in pollutant concentrations to a fractional change in health effect.

4.1 $PM_{2.5}$ health effects

In most cases, the published confidence intervals suggest that the underlying calculation was $b \pm ks_b$, with $k \cong 2$. From statistical theory, b is approximately normal, so it is not unreasonable to simulate betas as normal random variables with mean b and standard deviation s_b.

4.1.1 Beta for $PM_{2.5}$ mortality

The beta for $PM_{2.5}$ mortality is an exception. Because this is such a key element of the total health impact, the USEPA performed an evaluation based on the opinions of 12 experts in the field (EPA 2006). Each expert provided not just a point estimate of the effect, but a probability distribution representing the range where they expected the true effect would be. The benefit of such an expert elicitation is that the experts will attempt to correct for biases in the studies that served as the basis for their judgements.

The median of their estimates was 1.0 (% increase per 1 $\mu g/m^3$ increase in $PM_{2.5}$), with a 90% confidence interval of 0.3 to 2.0 (medians of their 5^{th} and 95^{th} percentiles, respectively). There are an infinite number of distributions with median 1.0, 5^{th} percentile 0.3 and 95^{th} percentile 2.0. Somewhat arbitrarily, but with the advantage of being a short-tailed distribution like the normal, we simulated b so that $y = (b^a + c)^{1/a}$ is

normal, where a and c and the mean and standard deviation of y are chosen so that b will have the above percentiles. Then $b = (y^a - c)^{1/a}$. We set $b = 0$ if $y^a \le c$.

Actually, a and c are not unique, although in the present case, values of $a \ge 0.6$ have no solution. One combination that works is $a = .5$ and $c = 3.95$. Then the mean of y is $\mu = (1^a + c)^{1/a} = 4.95^2 = 24.5$, and the standard deviation is $\sigma = [(2^a + c)^{1/a} - (.3^a + c)^{1/a}]/(2x1.645) = 2.597$. The probability density function for b is $f(x) = (1/\sigma)\phi[((x^a+c)^{1/a} - \mu)/\sigma]\, x^{a-1}(x^a+c)^{1/a-1}$, for $x>0$, where $\phi(z)$ is the standard normal pdf. For $a = .5$, this simplifies somewhat to $f(x) = \frac{1}{\sigma}\phi(\frac{x + 2c\sqrt{x} + c^2 - \mu}{\sigma})(1 + \frac{c}{\sqrt{x}})$, $x>0$. This density integrates to less than 1, specifically, it integrates to $P(Y > c^{1/a})$. For the remainder, we set $B = 0$. In the case at hand, $P(B=0) = P(Y<c^{1/a}) = P(Y<3.95^2) = \Phi[(15.6 - 24.5)/2.597] = 0.0003$, where $\Phi(z)$ is the standard normal cdf. Figure 4.1 shows the density for beta.

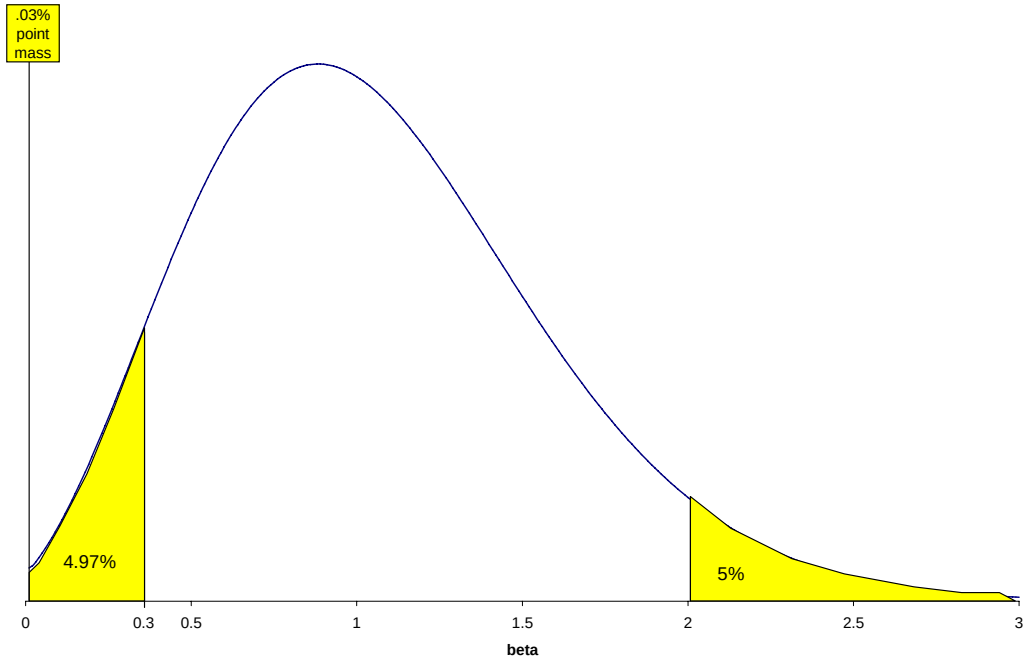


Figure 4.1. Assumed distribution of $PM_{2.5}$ mortality coefficients.

4.1.2 Betas for other PM effects

For the other PM effects, the uncertainties were expressed simply with a standard error. Knowing nothing more about the error distributions, we will assume that the distributions are Gaussian, except that any simulated value less than zero will be set to zero, under the assumption that air pollutants can't be beneficial.

4.2 Ozone health effects

We take our estimates of uncertainty from Ostro et al. (2006).

4.2.1 Ozone Mortality

Ostro considered a number of meta-analyses in developing their estimates. Their central estimate was 0.4% change in mortality per change in 10 ppb in 1-hour max. ozone, but they also included a lower estimate of 0.2% based on the NMAPS study and a European study, and an upper estimate of 0.6% based on studies by Thurston and Ito and a European study limited to summer months. They used a weight of 0.5 for the central estimate and 0.25 for the upper and lower estimates.

We take these weights as a Bayesian prior and use standard deviations of 0.05% for the central estimate, 0.15% for the lower and upper estimates, with the deviations simulated with Gaussian random variables. If the simulated coefficient is negative, it is set to zero.

4.2.2 Other ozone health effects

Ostro et al. (2006) used a similar approach for other health effects, using weights of 0.25, 0.50 and 0.25 for lower, midpoint and upper values. Their rationale was to give greater weight to the extremes to represent the additional uncertainties not accounted for within the studies cited. We follow their approach, simulating low, mid-point and high coefficients with probabilities 0.25, 0.50 and 0.25.

4.2.3 Uncertainty in threshold adjustment

Our method does not use the 1-hour maximum per se, but rather the excess of the 1-hour maximum above a presumed threshold of 50 ppb. To adjust for bias caused by this assumption, we multiply the betas by 2. This adjustment factor is approximate. Although the degree of uncertainty is itself uncertain, we believe it is easily 10%. Thus, we simulate the adjustment factor using a standard deviation of 0.2.

4.3 Incidence

For several health effects, the actual recent-year incidence rates are known. We use the mean and standard deviation of the incidence in recent years. Thus, for the health effects where Bay Area county incidence rates are known – mortality, hospital admissions, and asthma emergency room visits – we use the standard error of the 2005-07 mean, namely $s/\sqrt{3}$. Specifically, we simulate these incidence rates as normal random variables with μ= 05-07 mean and $\sigma = s/\sqrt{3}$.

For the other health effects, we have incidence rates from studies of other areas. In these cases, we have used placeholder uncertainties, namely 10% Gaussian errors.

5. Simulating Valuation

5.1 Mortality

Mortality valuations use an EPA analysis where the values from 26 "value of a statistical life" studies were modeled with a Weibull distribution (EPA 2008). Adjusting to 2009 dollars, the cumulative distribution function is $F(x) = 1 - \exp(-ax^b)$, where a = .0473, b = 1.5, and x is the VSL in millions of dollars. The mean of this distribution is $6.9 million. The figure shows the distribution. 90% of the simulated values fall between the two yellow areas, ranging from $1 million to about $16 million, with 5% falling below $1 million and 5% falling above $16 million.

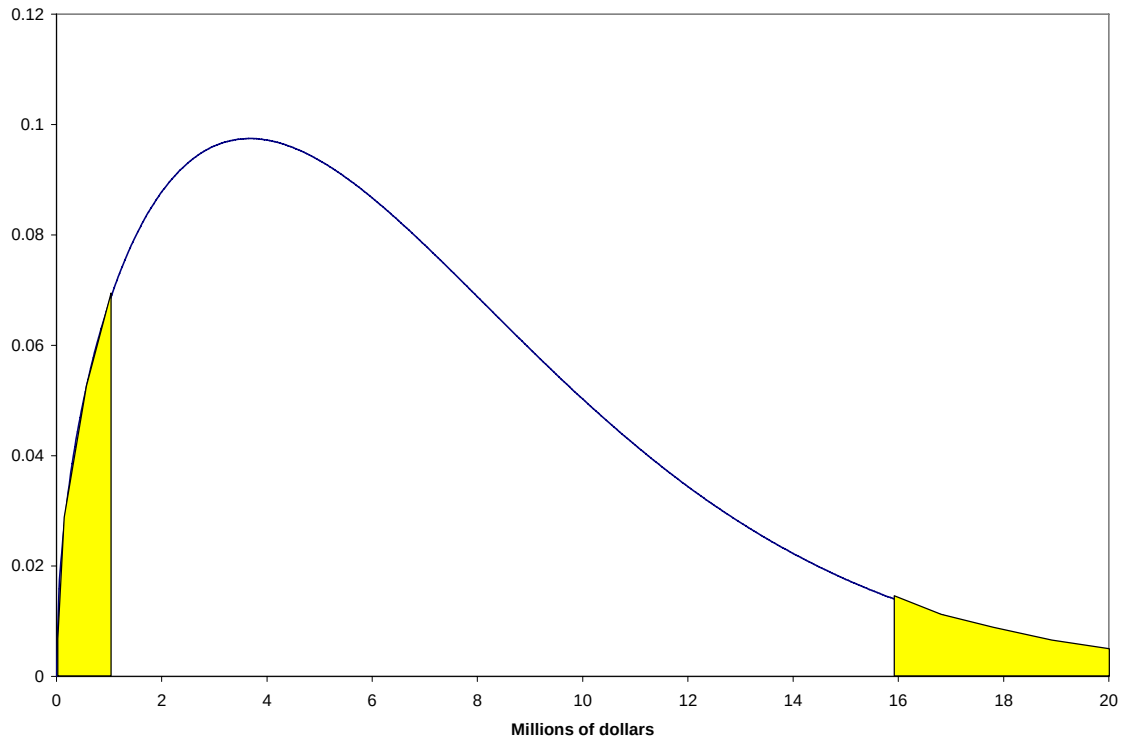


Figure 5.1. Assumed distribution of mortality valuation.

5.2 GHG Valuation

We assume a 95% confidence interval from $15 million to $85 million, with a median value of $28 million. Using the transform y = ln(beta + k), we look for the value of k so that

$$\ln(85 + k) - \ln(28 + k) = \ln(28 + k) - \ln(15 + k)$$

That is, a transform that makes the confidence interval symmetric around the median. Solving for k yields k = - 491/44 = 1.159. If we model y as a normal random variable, then its mean is ln(28 – 491/44) = 2.8238, and its standard deviation is approximately (ln(85-491/44)-ln(28-491/44))/2 = .73905. This implies the assumption that beta has a shifted lognormal distribution. See Figure 5.2.

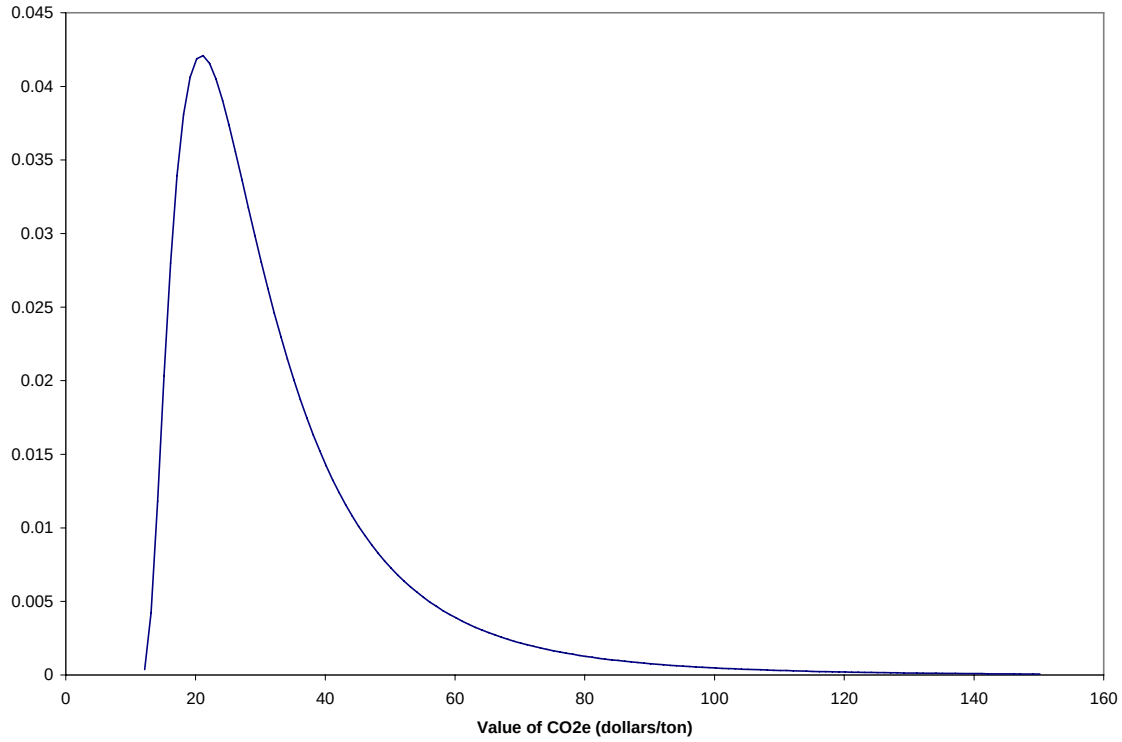


Figure 5.2. Distribution of GHG valuation per ton CO2e.

5.3 Chronic Bronchitis

We use the approach in BenMAP (2008), based on a study of severe bronchitis by Viscusi (1991), which established WTP values from a survey. The BenMAP approach models WTP using the Viscusi figures reduced to account for a lower valuation of less severe bronchitis. The Viscusi figures are modeled as WTP_{13}, with equal weights on the 9 deciles in his Table V. See table. The willingness to pay estimates are simulated as:

$$WTP = WTP_{13} e^{-\beta(13-x)}$$

where β is simulated as a normal random variable N(.18, .0669), and x has a triangular distribution with endpoints 1 and 12.

Table 2. WTP_{13}: Dollar Value ($100,000)

Decile	.1	.2	.3	.4	.5	.6	.7	.8	.9
1990 $	1.5	3.0	3.5	4.0	4.57	5.33	6.4	8.0	20.0
2009 $	2.55	5.09	5.94	6.79	7.76	9.04	10.86	13.58	33.94

5.4 Other health valuations

Table 3 shows the assumed distributions of health effects valuations, taken from BenMAP. In the cases labeled "No distribution available", we've used the placeholder assumptions of 10% Gaussian errors.

Table 3. Distributions of health effect valuations

Health Effect	Unit Value	Type of Measure	Distribution
Mortality (all ages)	$6,900,000	WTP	Weibull: $F(x) = 1 - e^{-.0473x^{1.5}}$
Chronic Bronchitis Onset	$409,189	WTP	$WTP_x = WTP_{13} * e^{-\beta * (13-x)}$. x ~ Triang(1,12), β ~ N(.18, .0669) For WTP13, see text
Respiratory Hospital Admissions	Age 65 < : $35,228 Age 65 > : $33,375	WTP + Third Part COI	No distribution available
Cardiovascular Hospital Admissions	Age 65 < : $43,889 Age 65 > : $38,759	WTP + Third Part COI	No distribution available
Non-Fatal Heart Attacks	$84,076	COI	No distribution available
Asthma Emergency Room Visits	$468	COI	Triangular Distribution where the minimum value is $348 and maximum value is $647 (asymmetric)
Acute Bronchitis Episodes	$534, for a 6 day illness period	WTP	Uniform Distribution where the minimum value is $157 and maximum value is $909.
Upper Respiratory Symptom Days	$35	WTP	Uniform Distribution where the minimum value is $13 and maximum value is $58.
Lower Respiratory Symptom Days	$22	WTP	Uniform Distribution where the minimum value is $13 and maximum value is $32.
Work Loss Days	Daily Median Wage by County; Alameda & Contra Costs: $202 Marin, San Francisco, & San Mateo: $228 Santa Clara: $243 Napa: $174 Solano: $168 Sonoma: $179	COI	No distribution available
School Absence Days	$91	COI	No distribution available
Minor Restricted Activity Days	$61	WTP	Triangular Distribution where the minimum value is $33 and maximum value is $202
Cancer	$1.75	WTP+COI	Uniform Distribution where the minimum value is $.7 million and maximum value is $2.8.

6. Examples

To illustrate the analysis, we take TCM C-5 Public Outreach and Education for Smart Driving/Speed Moderation. The estimated emissions reductions were: 336.6 lb/day NOx, 153.8 lb/day ROG, 18 lb/day $PM_{2.5}$, 1.6 lb/day acetaldehyde, 3 lb/day benzene, 0.62 lb/day 1,3-butadiene, 2.2 lb/day formaldehyde, and 59,602 lb/day CO_2. Applying the MPEM yielded an estimated benefit of $3,752,900 per year.

A Monte Carlo analysis was done. The screen shot, Figure 6.1, shows a typical simulation. Column B has the original estimated emissions reductions. Column C has the simulated reductions, e.g., PM2.5 reduced by 17.72 lb/day rather than 18, NOx reduced by 340.72 lb/day rather than 336.6, and so on. Column D has the simulated District total emissions, e.g., 44.38 tons/day direct non-diesel, carbonaceous PM2.5 rather than 49 tons/day, 487.926 tons/day NOx rather than 521 tons/day, and so on. In column G are the estimated Bay Area population exposures, where random factors have been used for the conversion of emissions to concentrations, and also to the population projections. For example, the estimated change in direct carbon PM2.5 exposure is 0.4905 ng/m3 compared with .5109 ng/m3 in the original (unrandomized) MPEM analysis. Column J shows the estimated number of cases, applying random factors for the health effects. For example, the simulated number of deaths from PM2.5 is 0.4487 compared with 0.2899 in the original analysis. Column K shows the simulated $/Case, e.g., mortality is valued at $12,057,198/case rather than $6,900,000. The bottom line for this simulation was $8,906,973 considerably more than the original MPEM.

Figure 6.2 shows a histogram of the simulated valuations. A few of the simulated values are much larger than the original $3.75 million per year; the smallest simulated value was over $1,000,000 per year.

Figure 6.3 shows a smoothed version of the cumulative distribution of the simulated valuations (mathematically, the integral of the histogram). The 10^{th} and 90^{th} percentiles represent a range of probable valuations, $1.9 million to $7.1 million per year. Thus, although there is considerable uncertainty at every step, the true "bottom line" is probably within a factor of 2 of the original $3.75 million per year estimate.

The median simulated value is somewhat below the original value. This is the general pattern. There is no reason the median value should equal the original, as shown in a simple example in Appendix C.

Simulation Number:	12			Simulation Size:	12				
	Emissions			Exposure Change		Health Effects		Valuation	
Precursor Pollutants	Reductions	Simulated Reductions	Bay Area Totals	Ambient Pollutants	Reduction / (Increase)	Effects	Cases	$/Case	Valuation
Particulate Matter	(lbs/day)		(tons/day)		(ng/m3)	Mortality	0.4487	$12,057,198	$5,409,739
Diesel PM2.5	0	0.00	13.97	Direct Carbon. PM2.5	0.4905	Chronic Bronchitis	0.1849	$119,793	$22,152
Direct PM2.5 except Diesel	18	17.72	44.38	ammonium nitrate	0.2226	Cardiovascular Hospital Admissio	0.0583	$39,028	$2,276
				ammonium sulfate	0.0000	Respiratory Hospital Admissions	0.0171	$38,660	$660
NOx	336.60	340.72	487.926	Total PM2.5	0.7131	Asthma Emergency Room Visits	0.1465	$595	$87
SO2	0.00	0.00	56.4672			Acute Bronchitis	0.5121	$376	$192
Ammonia	0.00	0.00	79.1063			Upper Respiratory Symptoms	42.4375	$16	$670
Direct Sulfate	0.00	0.00	2.92407			Lower Respiratory Symptoms	4.0099	$30	$121
ROG	154.80	139.27	522.516			Nonfatal Heart Attacks	0.2463	$85,437	$21,040
						Work Loss Days	29.7103	$219	$6,507
						Minor Restricted Activity Days	226.7153	$92	$20,948
						PM Total			**$5,484,393**
Ozone	(lbs/day)				(ppt)	Mortality	-0.0055	$12,057,198	-$65,944
NOx	336.60	340.72			-0.174	Respiratory Hospital Admissions	-0.0331	$38,660	-$1,278
ROG	154.80	139.27				Asthma Emergency Room Visits	-0.0207	$595	-$12
						Minor Restricted Activity Days	-111.1265	$92	-$10,268
						School Absences	-110.7951	$82	-$9,123
						Ozone Total			**-$86,624**
Toxics	(lbs/day)		(tons/day)		ng/m3				
Diesel	0.00	0.00	13.97		0.00	Cancer Incidence/Non-fatal	0.0000	$1,387,370	$0
Acetaldehyde	1.60	1.74	5.8025		0.09	Cancer Incidence/Mortality	0.0000	$10,456,723	$253
Benzene	3.00	2.74	4.58239		0.16	Cancer Incidence/Mortality	0.0005	$10,456,723	$5,152
1,3-Butadiene	0.62	0.55	1.10369		0.02	Cancer Incidence/Mortality	0.0004	$10,456,723	$4,319
Formaldehyde	2.20	2.27	7.11106		0.11	Cancer Incidence/Mortality	0.0001	$10,456,723	$723
						Toxics Total			**$10,446**
Greenhouse Gases	(tons/year)		(tons/day)					($/ton)	
CO2 equivalent	59602	47636	279918			Environmental/Economic Benefit		58.70	**$3,498,759**
Total						**Grand Total**			**$8,906,973**

Figure 6.1 Example screen shot (TCM C-5) from the uncertainty analysis template

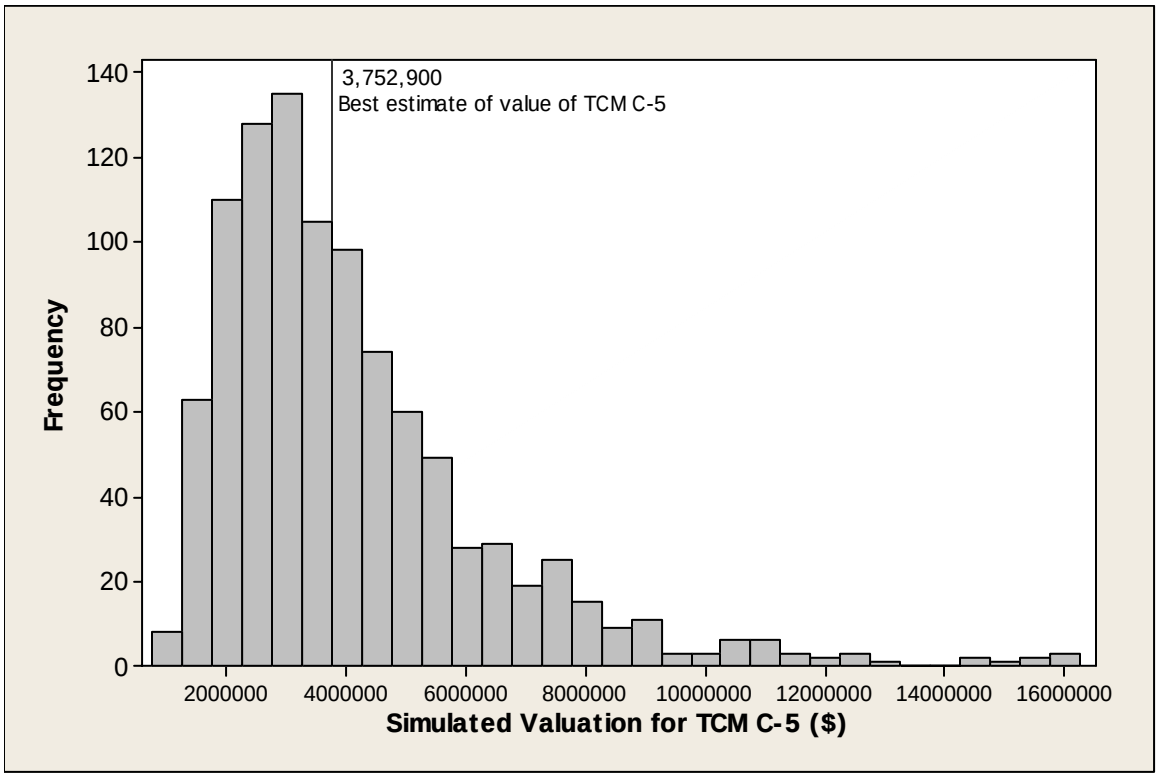


Figure 6.2. Histogram of simulated total valuation in dollars for TCM C-5.

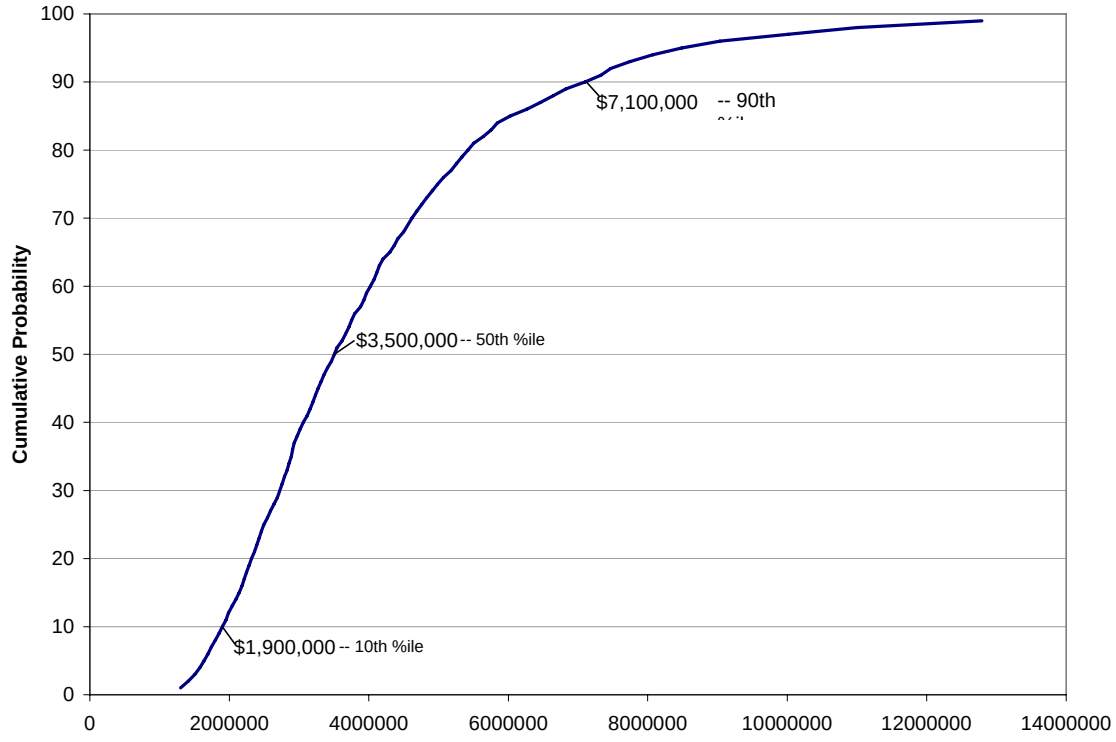


Figure 6.3. Cumulative distribution of simulated valuations for TCM C-5.

6.1 Comparison of two control measures

We can use the uncertainty analysis to compare the valuations of pairs of control measures. As an example, consider another measure, TCM E-3 Implement Transit Pricing Reform. Its estimated benefit is $5,561,00 per year, but the uncertainty analysis yields a 10% to 90% range of $2.5 million to $10.8 million per year, that is, the range overlaps with the range for TCM C-5. But if we look at the valuations run by run (Figure 6.4), we see that in *every* run the valuation of TCM E-3 is greater than that of TCM C-5, so that it's virtually certain that the dollar benefit of TCM E-3 is greater than that of TCM C-5.

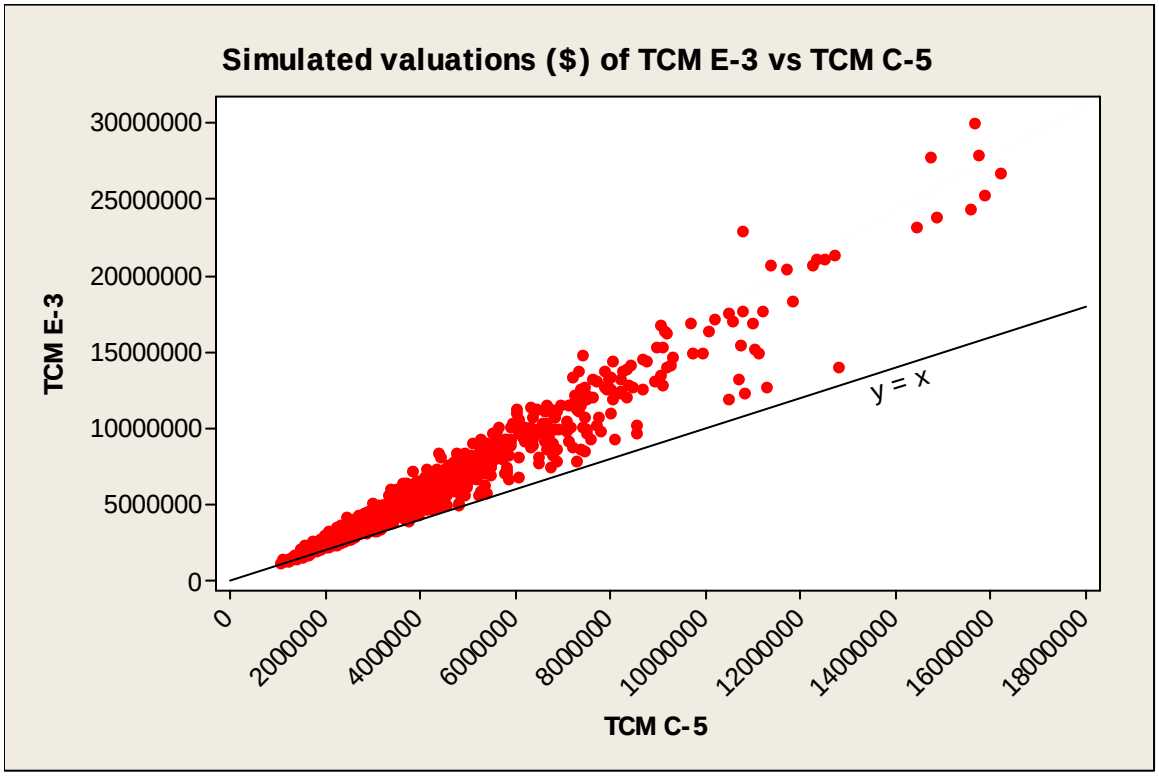


Figure 6.4. Pairwise comparison of TCM E-3 and TCM C-5 based on 1,000 simulations.

In other cases, there is overlap. For example, TCM C-5 has a larger point estimate valuation than ECM 3 Urban Heat Island Mitigation, $3.9 million vs. $3.1 million, but Figure 6.5 shows that there were a substantial fraction of runs (16%), where the simulated valuation of ECM 3 was greater than that of TCM C-5.

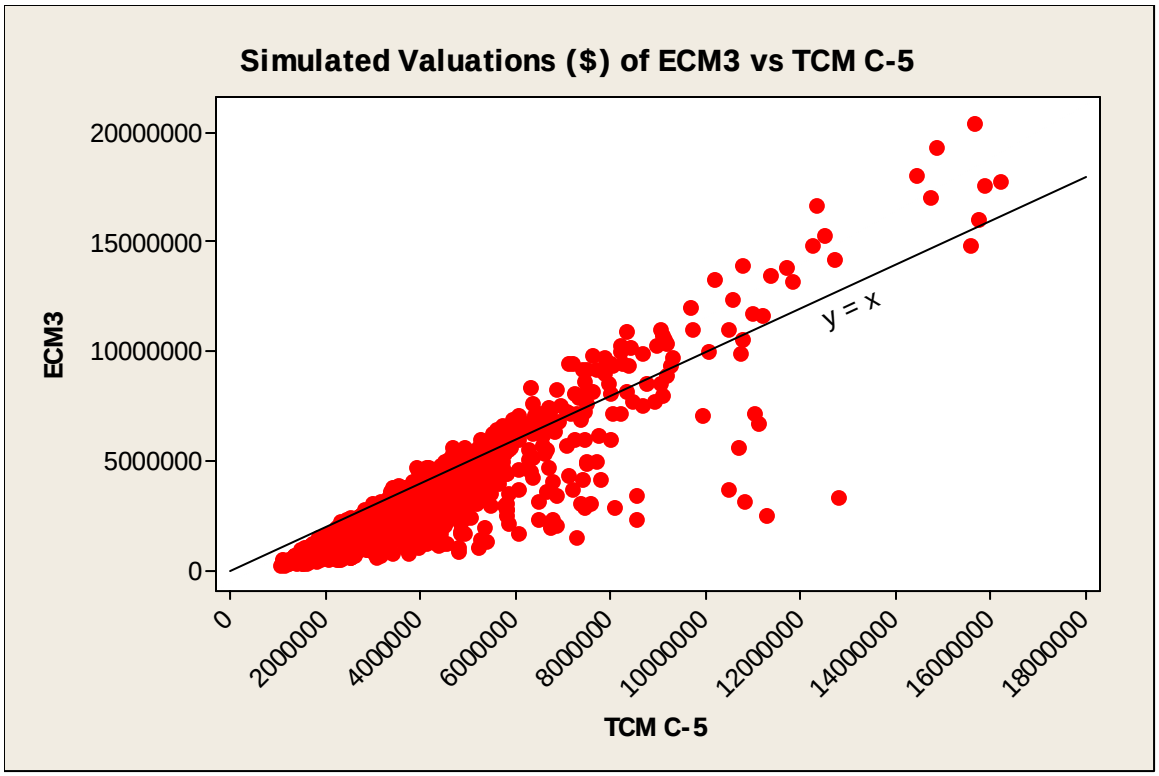


Figure 6.5. Pair-wise comparison of ECM 3 and TCM C-5 based on 1,000 simulations.

Figure 6.6 shows pair-wise comparisons for all the measures. The measures in the rows and columns are sorted in order of decreasing point estimate valuation. The colors indicate whether or not there is strong evidence that the valuation of the column measure is greater than that of the corresponding row measure. Red indicates strong evidence – at most 10 runs (i.e., 1%) where the valuation of the row measure was greater. Orange indicates that it's likely that the column measure is greater – between 10 and 50 runs where the row measure had a greater valuation. Blue indicates that the valuations of measures can't be well distinguished, pairs where there were > 50 runs with r > c.

More than eighty percent of the pairs can be distinguished; most measures are indistinguishable only from a few others. LUM2, for example, can be distinguished from all measures with smaller point valuations except for SSM6 and TCM D-3, and from all measures with larger point valuations except for LUM1 and TCM B-1.

A set of measures does stand out for having greater uncertainties – SSM 9 through SSM 14 and MSM B-2. These are all NOx-only measures. That these can't be distinguished from many measures with much smaller valuations indicates the large uncertainty about the impact of NOx reductions for the Bay Area.

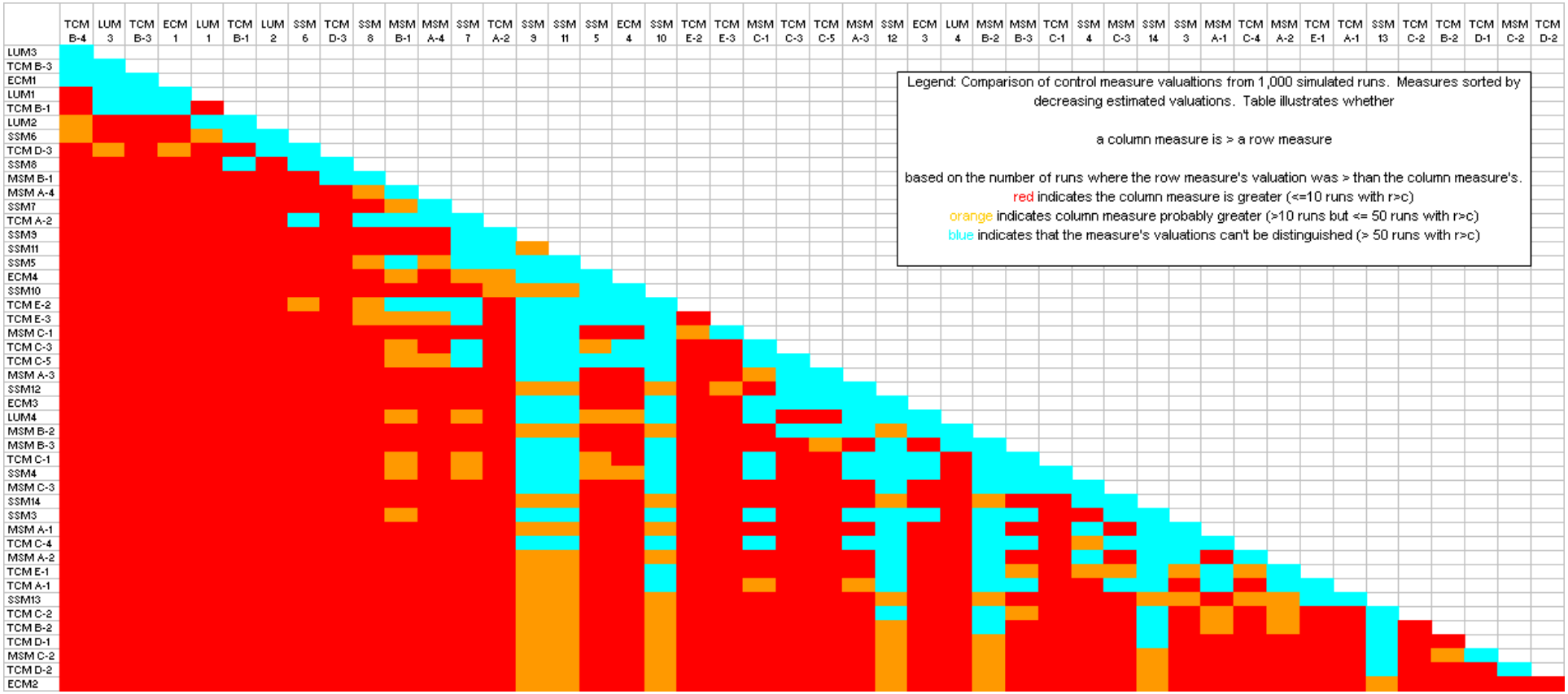
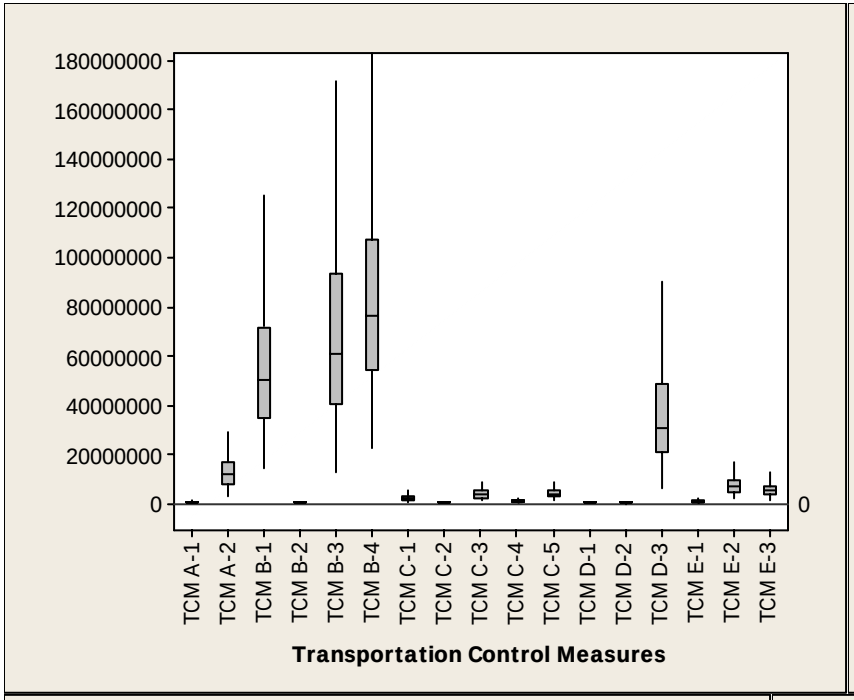
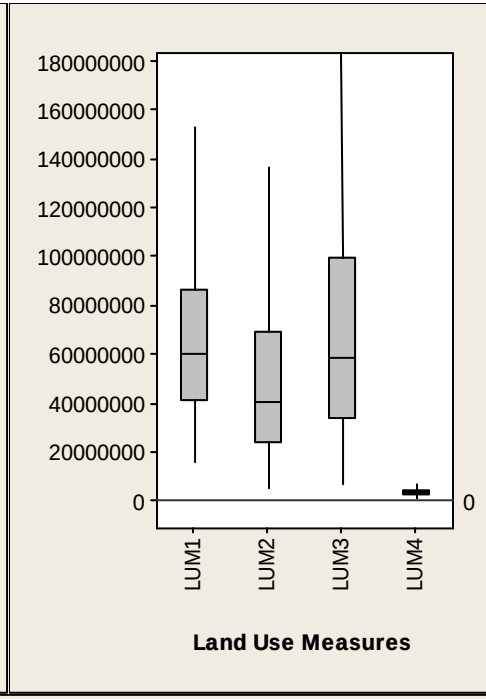
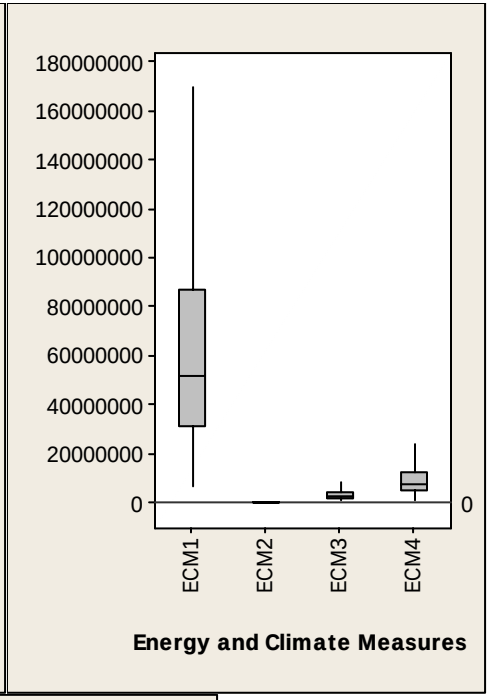
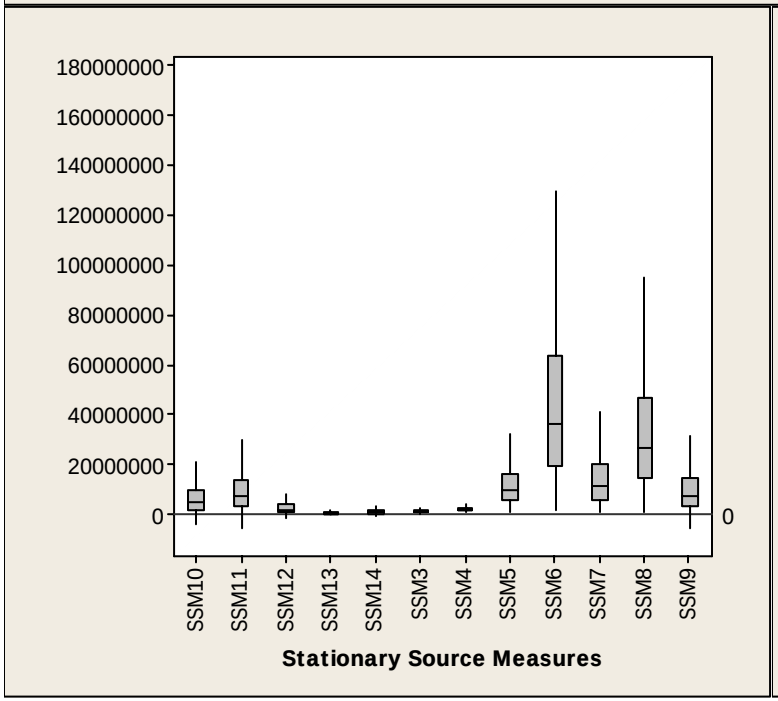


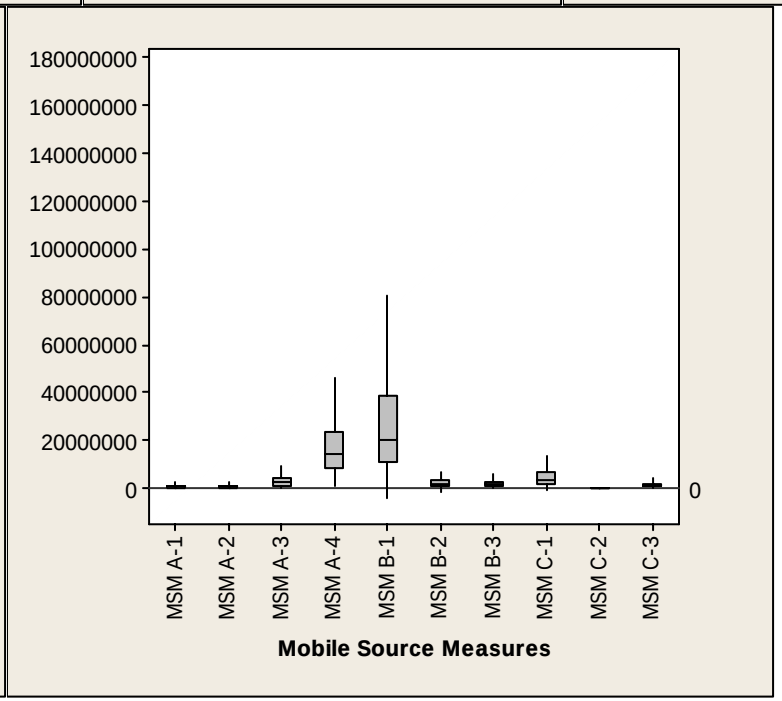
Figure 6.6. Pairwise comparison of the valuations of the plan measures. The colors indicate whether the pairs are distinguishable, based on the Monte Carlo analysis.


180000000
160000000
140000000
120000000
100000000
80000000
60000000
40000000
20000000
0
0
TCM A-1
TCM A-2
TCM B-1
TCM B-2
TCM B-3
TCM B-4
TCM C-1
TCM C-2
TCM C-3
TCM C-4
TCM C-5
TCM D-1
TCM D-2
TCM D-3
TCM E-1
TCM E-2
TCM E-3
Transportation Control Measures


180000000
160000000
140000000
120000000
100000000
80000000
60000000
40000000
20000000
0
0
LUM1
LUM2
LUM3
LUM4
Land Use Measures


180000000
160000000
140000000
120000000
100000000
80000000
60000000
40000000
20000000
0
0
ECM1
ECM2
ECM3
ECM4
Energy and Climate Measures


180000000
160000000
140000000
120000000
100000000
80000000
60000000
40000000
20000000
0
0
SSM10
SSM11
SSM12
SSM13
SSM14
SSM3
SSM4
SSM5
SSM6
SSM7
SSM8
SSM9
Stationary Source Measures


180000000
160000000
140000000
120000000
100000000
80000000
60000000
40000000
20000000
0
0
MSM A-1
MSM A-2
MSM A-3
MSM A-4
MSM B-1
MSM B-2
MSM B-3
MSM C-1
MSM C-2
MSM C-3
Mobile Source Measures

6.2 Uncertainty of the Sum Total of Valuations

The sum of the point estimate valuations from all the plan control measures is $769 million per year. An uncertainty range for this estimate is obtained by summing the simulated valuations for each of the 1,000 runs. Figure 6.7 shows the cumulative distribution of these sums. The estimated overall benefits of the plan range from $270 million to $1,550 million per year. Thus, even incorporating the uncertainties from every step of the MPEM, the overall uncertainty is still between a third and twice the estimated value. In other words, the overall benefits of the plan aren't exactly $739 million per year, but they are certainly in the hundreds of millions, possibly over a billion dollars, and certainly less than two billion dollars per year.

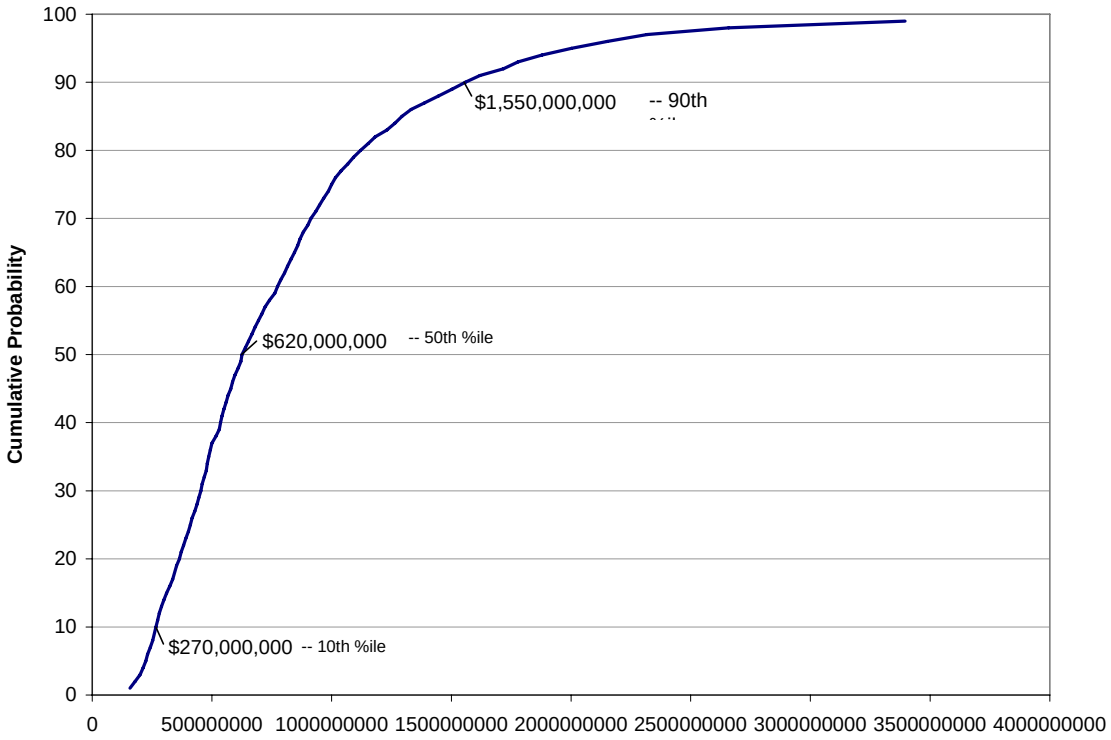


Figure 6.7. Cumulative distribution of total plan benefits ($).

7. References

Abbey, D.E., B.D. Ostro, F. Petersen, and R.J. Burchette. 1995. Chronic respiratory symptoms associated with estimated long-term ambient concentrations of fine particulates less than 2.5 microns in aerodynamic diameter (PM2.5) and other air pollutants. *Journal of Exposure Analysis and Environmental Epidemiology* 5(2):137-159.

Dockery, D.W., J. Cunningham, A.I. Damokosh, L.M. Neas, J.D. Spengler, P. Koutrakis, J.H. Ware, M. Raizenne, and F.E. Speizer. 1996. Health effects of acid aerosols on North American children – Respiratory symptoms. *Environmental Health Perspectives* 104(5):500-505.

EPA 2006 "Expanded expert judgment assessment of the concentration-response relationship between $PM_{2.5}$ exposure and mortality", prepared for OAQPS-EPA by Industrial Economics Inc, September 21, 2006.

EPA 2008 *BENMAP Users Manual Appendices*, September, 2008. Prepared for the Office of Air Quality Planning and Standards, USEPA by Abt Associates, Inc.

Hall JV, Brajer V, and Lurmann, FW 2006 "The benefits of meeting clean air standards in the South Coast and San Joaquin Valley air basins," November, 2008.

Ito K. (2003) Associations of particulate matter components with daily mortality and morbidity in Detroit, Michigan, in: *Revised Analyses of Time-Series Studies of Air Pollution and Health*, Special Report, Health Effects Institute, Boston, MA.

Moolgavkar, S.H. 2000a. Air Pollution and Hospital Admissions for Chronic Obstructive Pulmonary Disease in Three Metropolitan Areas in the United States. Inhalation Toxicology. Vol. 12 (Supplement 4): 75-90.

Moolgavkar, S.H. 2000b. Air pollution and hospital admissions for diseases of the circulatory system in three U.S. metropolitan areas. J Air Waste Manag Assoc. Vol. 50 (7): 1199-206.

Moolgavkar, S.H. 2003. Air Pollution and Daily Deaths and Hospital Admissions in Los Angeles and Cook Counties. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Health Effects Institute. Boston, MA. May.

Ostro, B. D. 1987. Air Pollution and Morbidity Revisited: A Specification Test. Journal of Environmental Economics and Management. Vol. 14: 87-98.

Ostro, B. D. and S. Rothschild. 1989. Air Pollution and Acute Respiratory Morbidity - an Observational Study of Multiple Pollutants. Environ Res. Vol. 50 (2): 238-247.

Ostro BD, Tran H, Levy JI (2006) "The health benefits of reduced tropospheric ozone in California", *Journal of the Air and Waste Management Association*, 56: 1007-1021.

Peters A., D.W. Dockery, J.E. Muller and M.A. Mittleman. (2001) " Increased Particulate Air Pollution and the Triggering of Myocardial Infarction" *Circulation* 103(23): 2810-2815.

Schwartz, J. and L. M. Neas. 2000. Fine particles are more strongly associated than coarse particles with acute respiratory health effects in schoolchildren. Epidemiology. Vol. 11 (1): 6-10.

Viscusi W.K., W.A. Magat and J. Huber (1991) Pricing environmental health risks: survey assessments of risk-risk and risk-dollar trade-offs for chronic bronchitis *Journal of Environmental Economics and Management* **21**(1), 32-51.

Appendix A – Differences in $PM_{2.5}$ Means Between Pairs of Sites as a Function of Distance

Using ambient data to evaluate the performance of grid models is made more complex because the ambient data are collected at points whereas the grid models estimate a volume average. This appendix analyses how to account for this difference.

The model simulated hourly concentrations in a 4x4 km grid for a 2-month period, December 2006 through January 2007 for this study.

The key issue is the use of modeled values in the multi-pollutant method. For this method, only the average value for each grid square is used. The question is the extent that this model average, M, deviates from the true average, T.

The evidence available are daily measurements from 14 Bay Area $PM_{2.5}$ monitoring sites for some or all of the days. Let the average value at a given site be denoted, S. One approach would be to look at M – S, and assume that any difference is due to modeling error. But generally speaking, $S \neq T$, so it seems reasonable to discount discrepancies between M and S to account for this.

The difference between S and T will derive almost exclusively from actual spatial variation in the ambient concentrations as opposed to day-to-day variation: Correlations of daily $PM_{2.5}$ concentrations at nearby sites are near 1, making it likely that daily values measured at a site would correlate highly with the average in a 4x4 km grid containing the site. But we would expect considerable variation in mean concentrations spatially, with measurements near sources like roadways and near woodburners being higher than in the middle of a park.

We can get a sense of how much variation there could be by looking at the differences in the means between pairs of sites and seeing how this depends on distance. Figure A1 shows pairwise differences among 14 sites plotted versus the distance between the sites. There is considerable scatter but, generally, the differences increase with distance. A linear regression has an intercept of 2.0 $\mu g/m^3$ and a slope of 0.07 $\mu g/m^3$ / km.

If one or both the measurements were made with a BAM unit, there is the potential for bias. In fact, it's likely that a good part of the 4 $\mu g/m^3$ difference between the Oakland and San Francisco means (furthest left point on the graph) is measurement bias. However, the 6 $\mu g/m^3$ difference between San Jose and Fremont is likely real since both sets of measurements are made with the same scale at the District office.

What do these differences tell us about the relation of S and T? We can think of the regression predictions for distances < 4km as representing the expected differences between sites located in different randomly selected locations within a grid square, that is $S_1 - S_2$, where each S_i has the same distribution as S.

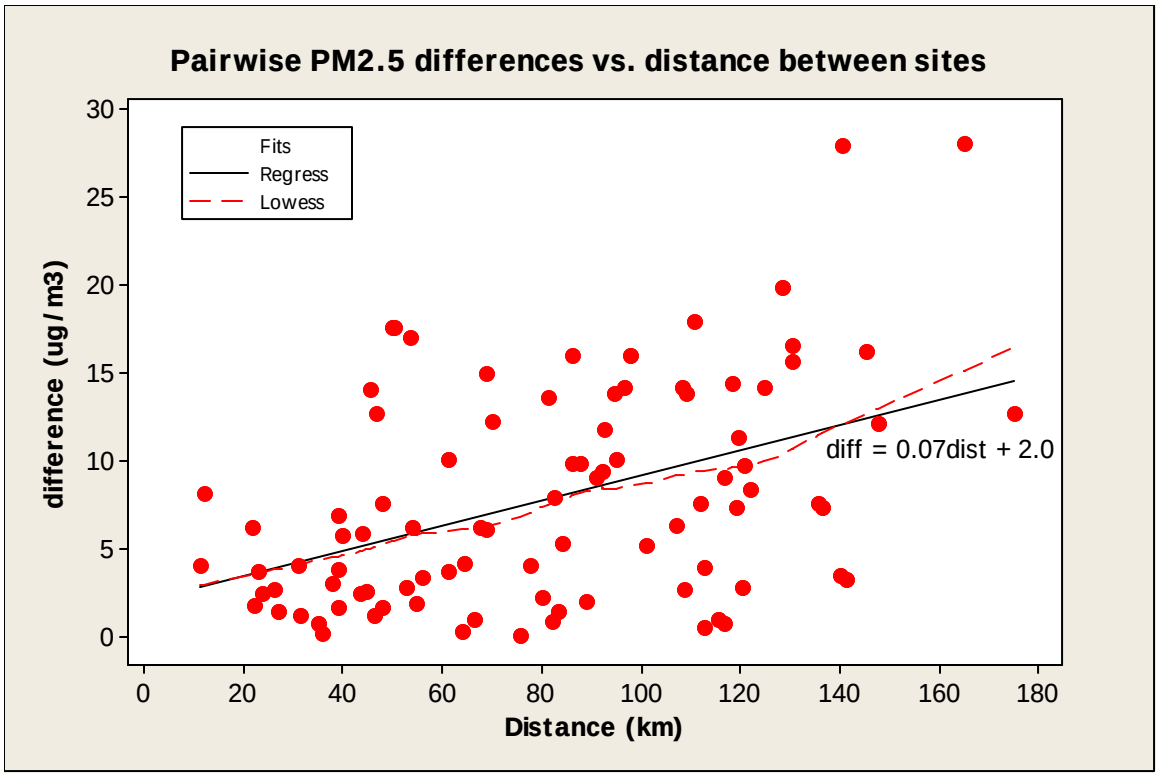


Figure A1. Absolute values of difference in means between pairs of 11 Bay Area sites + 2 Sacramento sites + Modesto plotted vs. distance between sites. Regression lines and non-parametric smooth also shown. Data from the PM2.5 modeling period: December 2006 through January 2007. Note that for each difference, the mean difference was calculated for days when both sites had data.

Although the S_i are sample means, suggesting a normal distribution, they also have a random mean, depending on where they fall within the grid. The variation of concentrations within the grid is not necessarily normal. To look further, I took the values in Figure 1 and divided by the regression prediction. See Figure A2.

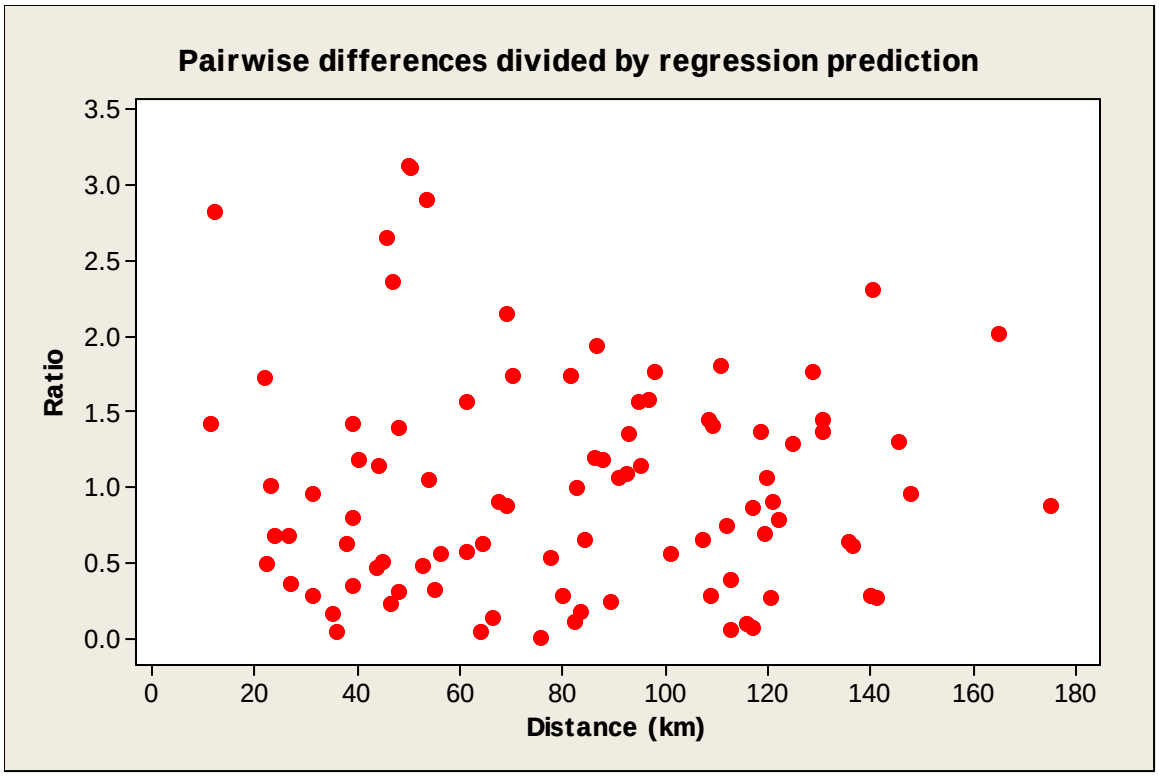


Figure A2.

Visually, this transformation seems to stabilize the variability, so that the distribution of the ratio is approximately independent of distance between sites. If the differences do behave like normally distributed random variables, then the ratios should follow a folded normal distribution – the distribution of the absolute value of a normal random variable.

Figure A3 shows a comparison of the sorted ratios versus the quantiles from a folded normal. The plot is quite straight, indicating a good fit. Thus, the assumption that $S_1 - S_2$ is normally distributed seems reasonable.

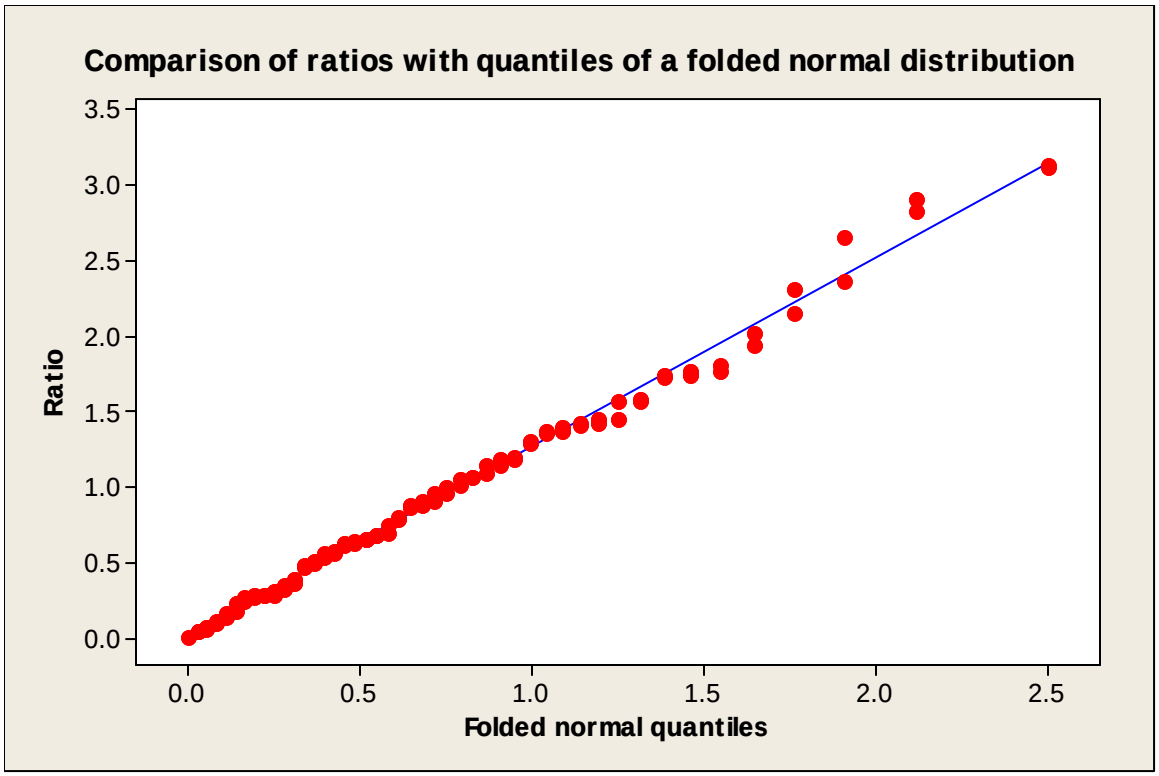


Figure A3.

The expectation of the absolute value of a normal random variable with mean 0 and variance σ^2 is $\sqrt{\frac{2}{\pi}}\sigma$. As shown in the equation in Figure 1, the mean regression prediction for nearby sites is about 2 μg/m^3. Thus, $\sqrt{\frac{2}{\pi}}\sigma \cong 2$, or $\sigma \cong \sqrt{2\pi}$. The variance of $S_1 - S_2$ is twice the variance of S, so the standard deviation of S is $\cong \sqrt{\pi}$ μg/m^3 or about 1.7 μg/m^3.

Figure A4 shows a comparison of the mean measured $PM_{2.5}$ concentrations at various sites compared with the modeled concentrations in the grid square containing the sites.

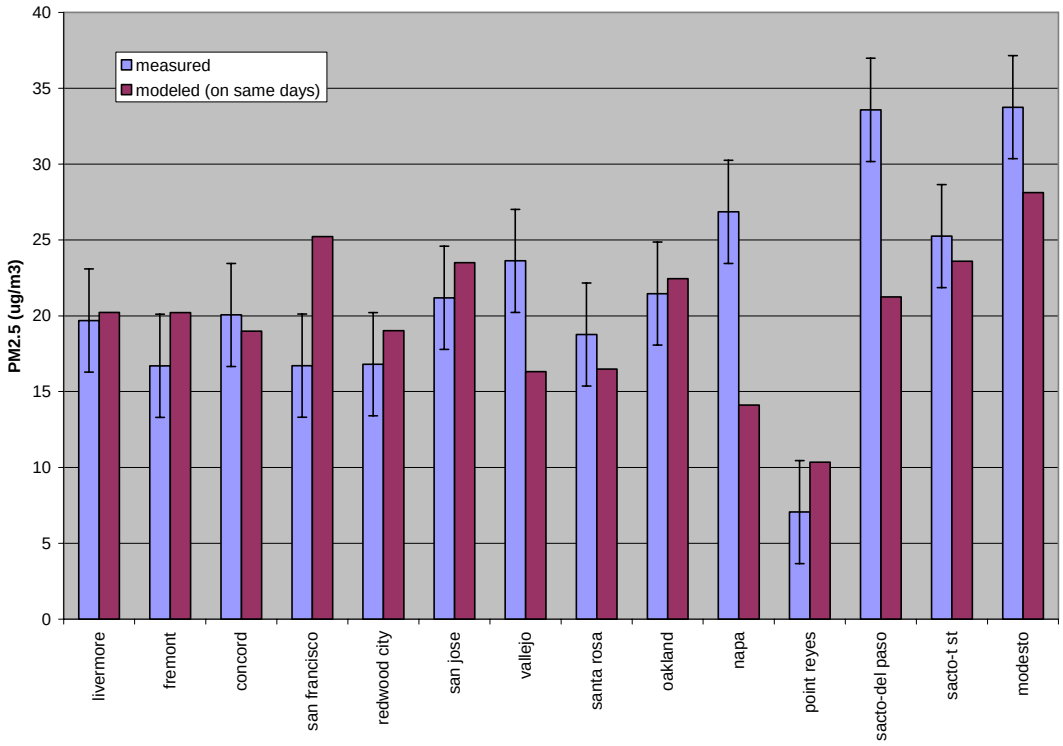


Figure A4. Comparison of measured and modeled $PM_{2.5}$ concentrations. Also shown are error bars representing the uncertainty between the measured point estimated concentration and the true average concentration over the grid.

For many of the sites, the model values are within the error bars of the measured values so, although not identical, the differences may well be due to the site not being representative of the grid as a whole.

Appendix B – Calculations of clump size and frequency

The simulation of grid cell uncertainty includes a term for errors in random clumps of cells. This appendix analyzes the distribution of the number of clumps and their size.

The setup is that a given grid cell is the clump seed with a small probability, p. In the simulation, p = 0.02. Once a clump is started, it is expanded in 4 directions with probability, r, say. In the simulation, r = 0.4.

Let's first consider the expected number of clumps covering a given grid cell. Figure B1 shows the number of steps to a given cell from the cells surrounding it. There are 4 cells 1 step away, 8 cells 2 steps away, 12 cells 3 steps away, and so on.

5	4	3	4	5
4	3	2	3	4
3	2	1	2	3

2	1	0	1	2
3	2	1	2	3
4	3	2	3	4
5	4	3	4	5

Figure B1. Centered on a given grid cell showing the number of steps from other cells to that cell.

The probability a given cell is the seed of a clump is 0.02. The probability it's covered by a clump from the cell just to the left, above, right, or below is 0.02 x 0.4. The probability it's covered by a clump from a cell that's 2 away is 0.02 x 0.4 x 0.4, and so on. Thus, the expected number of clumps covering a cell (that's far from the edge of the grid) is:

$$\begin{aligned} p + 4pr + 8pr^2 + 12pr^3 + \ldots &= p[1 + 4\sum_{j=1} jr^j] \\ &= 0.02[1 + 4\sum_{j=1} j0.4^j] \\ &= 0.02[1 + 4(0.4)/0.6^2] \\ &\cong 0.109 \end{aligned}$$

Thus, we expect about 10% of the cells to be covered by a clump. We can also look at the expected area of a clump. For a given clump, its area is:

(1 + N + S)(1 + E + W),

where N, S, E, and W are the numbers of cells the clump extends in those directions. Each has a geometric distribution with probability r. The mean of this distribution is r/(1–r). Because the numbers in each direction are independent, the expected value of the product is the product of the expected values, so that

$$\text{E(Area of clump)} = (1 + \frac{r}{1-r} + \frac{r}{1-r})(1 + \frac{r}{1-r} + \frac{r}{1-r}) = (1 + \frac{2r}{1-r})^2$$

For r = 0.4, E(Area) ≅ 5.4.

Appendix C – An example where the median simulated value is less than the original

The valuations from the Monte Carlo simulation can be thought of as a function of a set of random terms, $Y = f(X_1, X_2, X_3,\ldots)$, where the X_i are initial emissions estimates, total emissions estimates, emissions-concentrations factors, population estimates, and so on. The original estimate is the function evaluated at the means of the X_i: $f(\mu_1, \mu_2, \mu_3,\ldots)$, where μ_i = the expected value of X_i.

In general, there's no reason that the median of Y should be equal to this original estimate. The function, f, contains many terms – additive, multiplicative and exponential. A simple example illustrates that multiplicative functions of random variables with symmetric distributions don't have this property.

Consider a very simple case where $Y = X_1X_2$, and X_1 and X_2 each take the values 0.6 and 1.4 with equal probability. Then X_1 and X_2 each have expected value 1. Here $f(x_1,x_2) = x_1*x_2$, so that $f(\mu_1, \mu_2) = f(1,1) = 1*1 = 1$. But if we look at the joint distribution, there is a ½ * ½ = ¼ chance each equals 0.6, with a product of 0.36, a ½ chance that one equals 0.6 and the other equals 1.4, with a product of 0.6 x 1.4 = 0.84, and a ¼ chance that both equal 1.4, with a product of 1.96. The expected value of Y is $E(X_1X_2) = 0.25*0.36 + 0.50*0.84 + 0.25*1.96 = 1$, but there is a 75% chance that the product is ≤ 0.84. So the original value and the expected value of the Monte Carlo valuations are both 1.00, but the median is 0.84.

For a slightly more realistic example, suppose that X_1 and X_2 are uniformly distributed between 0.6 and 1.4 (e.g. modeling errors with a range of ± 40%). If we consider values of X_1X_2 where the product is > 0.6 x 1.4, then the probability the product is greater than a given value, r, is

$$P(X_1X_2 > r) = [\ 1.4^2 - r*\ln(1.4) - r + r\ *\ln(r/1.4)]/0.64$$

With some trial and error, the median, that is the value of r such that $P(X_1X_2 > r) = 0.5$, is r = 0.952.

In general, if the errors are symmetrically distributed around a central value, then the fractional error that falls below the central value will average somewhat greater than the fractional error that falls above: Let the central value = c, and the error = kc. Then the fraction error below the central value is (c – kc)/c = 1 – k and the fractional error above is (c + ck)/c = 1 + k. If the errors are additive, then the errors tend to cancel (1 – k) + (1 + k) = 0. But if the errors are multiplicative, then:
$(1 - k)*(1 + k) = 1 - k^2 < 1$. To see it another way, suppose k = 0.1. Then the positive error is 1.1 times c, but the negative error is c divided not by 1.1, but by 1/.9 = 1.1$\overline{1}$