ESTIMATING UNCERTAINTY IN THE SECOND 812 PROSPECTIVE ANALYSIS OF THE BENEFITS AND COSTS OF THE CLEAN AIR ACT

SUMMARY OF THE ISSUE

The estimation and representation of uncertainty in results is an integral part of developing a sound analytical plan for the next Section 812 Prospective Analysis of the Clean Air Act. A well-conducted and well-presented uncertainty analysis can provide decision makers with valuable context concerning the range and likelihood of various outcomes, as well as characterize the state of scientific knowledge in areas critical to developing benefit and cost estimates.

IEc previously developed for OPAR a chapter in the May 2003 analytical plan (Chapter 9) devoted to uncertainty analysis of benefit and cost results. The chapter outlined how uncertainty was addressed in the first prospective, described plans for addressing uncertainty in the second prospective, listed the major uncertainties from the first prospective study, and indicated how our proposed plan for the second prospective might affect those uncertainties.

Much has changed since IEc completed the uncertainty chapter in May 2003. EPA has made significant progress in some areas of uncertainty analysis (e.g., expert elicitation of PM-related mortality), while shifting its focus away from other approaches being explored in 2003 (e.g., influence analysis). The 812 project team has updated the alternative scenarios it plans to evaluate in the study. Finally, we recently became aware of additional efforts within EPA that may be useful for addressing options for uncertainty analysis within the context of benefit/cost analysis of air pollution regulations.

In light of these new developments, IEc has reassessed options that currently appear most promising for updating how uncertainty is assessed and presented in the next 812 analysis. This white paper describes the approaches IEc would recommend implementing in the analysis. It begins with an overview of the objective of a revised uncertainty analysis. It then discusses specific recommended approaches to particular types of uncertainties (scenario, model, parameter) associated with costs, emissions and AQM, health effects estimation, and valuation. Finally, it discusses recommendations for presenting uncertainty information. We also include three attachments to this white paper: a memo from Sonoma Technologies proposing an approach for addressing uncertainties in emissions and air quality modeling; a brief description of the approach for addressing uncertainty analysis in the air toxics case study of benzene reductions in Houston that is being conducted as part of this 812 analysis; and a memo describing an alternative method for estimating EGU emissions in the without-CAAA scenario.

In brief, we recommend the following:

- **Scenario Analyses:** We propose to run the full-scale 812 analysis for two alternative future scenarios for economic and population growth in the U.S., in addition to the primary analysis. Analysis of the high growth and low growth alternative scenarios will allow for an integrated treatment of the effect of temporal uncertainty on both the benefit and cost results of the 812 analysis. In addition, we propose to use scenarios featuring incremental changes in emissions from specific sectors to better understand the sensitivity of the analytical models to these changes. These results along with current knowledge of uncertainties in the data for different emission sectors can help EPA identify key areas for future research in the emissions portion of the analytical chain.

- **Expanded Benefits Uncertainty Analysis:** We propose to significantly expand consideration of uncertainty in benefits estimates beyond the Monte Carlo analysis of statistical uncertainty in health effect C-R functions and valuation that was performed for the first prospective. The proposed changes include an additional, more expansive uncertainty analysis for PM-related mortality using results from EPA's 2006 expert elicitation study; evaluation of the impact of model uncertainty related to the cessation lag for reductions in premature mortality; evaluation of the sensitivity of monetized benefits to alternative assumptions about the income elasticity of willingness-to-pay based valuations of health effects; and exploring alternative treatments of uncertainty in VSL.
- **Offline Uncertainty Analyses:** We propose additional "off-line" analyses for elements less easily integrated into the 812 analytical chain, including a sensitivity analysis of key uncertainties in emissions and air quality modeling using a limited number of CMAQ runs and available reduced-form alternatives to CMAQ; and a break-even analysis that explores both the magnitude of error in emissions reduction estimates necessary to produce negative net benefits for particular sectors and the likelihood of those scenarios. Finally, we propose a comprehensive qualitative uncertainty analysis similar to that performed for the first prospective.
- **Enhanced Presentation Tools:** We propose enhanced use of graphics, such as box and whisker plots, probability density functions, and cumulative density functions to complement the tabular presentation of quantitative uncertainty results. These types of graphs can be particularly useful for displaying the results of scenario and sensitivity analyses, representing parameter uncertainty, comparing alternative distributions, and conveying the relative likelihoods of outcomes (e.g., positive versus negative net benefits).

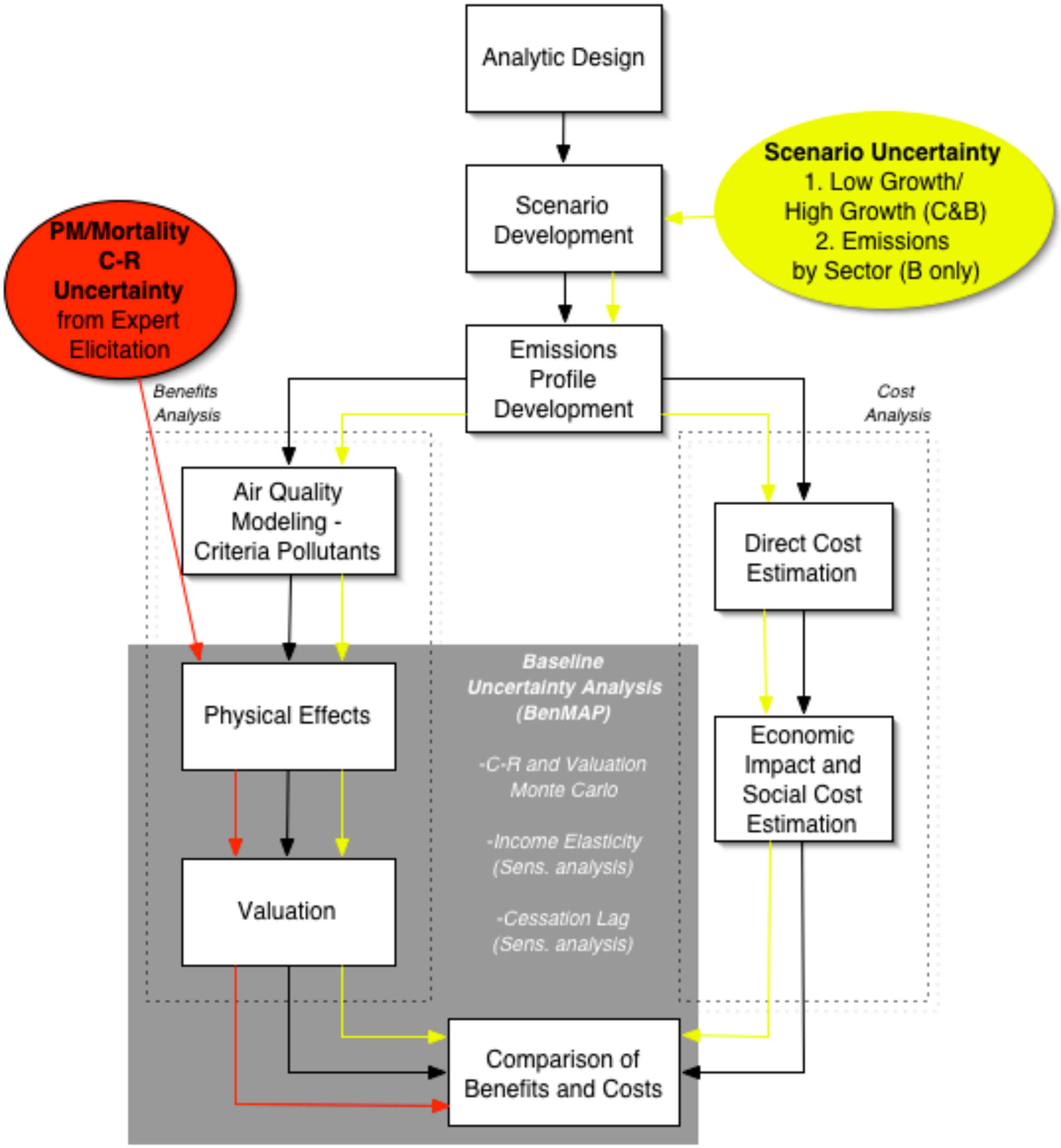
OVERVIEW AND OBJECTIVES OF UNCERTAINTY ANALYSIS

A comprehensive uncertainty analysis for a national-scale study with a scope as expansive as the Section 812 Benefit-Cost Analysis represents an immensely challenging task. The complexity of the models involved precludes use of Monte-Carlo style statistical sampling to analyze the impact of upstream uncertainties in emissions and air quality modeling. Uncertainty analysis in previous 812 analyses has been limited to Monte Carlo analysis of statistical uncertainty in concentration-response and valuation for health outcomes in the U.S. population. Both the NAS in its 2002 report and the 812 Council in numerous advisories have encouraged more comprehensive analysis of uncertainties in benefits analyses for air quality regulations. While the NAS report presents ambitious long-tem goals for Agency analysis, much of what has been recommended is simply not practical to apply in the 812 analysis given available time and resources, and our current state of knowledge about uncertainties in key analytical elements such as air quality modeling. Our objective in designing the second 812 prospective is to recommend reasonable incremental adjustments to the uncertainty analysis that provide policy-relevant insights into the impacts of alternative assumptions on benefits and costs of the Clean Air Act.

IEc's recommended modifications to the 812 uncertainty analysis include both "online" analyses, that feed information on uncertainty into the analytical chain at various points and propagate it through the remaining steps in the chain, and separate "offline" analyses and research that will provide insights into the uncertainty, sensitivity, and robustness of results to alternative assumptions that are currently most easily modeled outside the main analytical process.

Exhibit 1 shows the basic analytical flowchart underlying the section 812 analysis and superimposes on top of this both the existing and new uncertainty elements that feed into the flowchart (the "online" analyses). The gray box in the lower left corner of the figure indicates current capabilities in quantitative uncertainty analysis available with no new investment in methods. Current capabilities utilize EPA's BenMAP model to perform Monte Carlo-style statistical sampling to model uncertainty in Physical Effects, represented by the standard errors of concentration-response (C-R) coefficients, and uncertainty in Valuation, using distributions of unit values (e.g., value of a statistical life or VSL) for mortality and other health effects associated with air pollution exposure. In addition, we are aware of new features in BenMAP that can be used to: 1) model sensitivity of results to assumptions about the income elasticity of VSL and other willingness-to-pay based valuation estimates; and 2) model the impact on mortality-related benefits of alternative functional forms for the cessation lag. Both of these can be applied in the valuation step of the benefits analysis.

EXHIBIT 1 “ONLINE” QUANTITATIVE UNCERTAINTY ANALYSIS FOR SECOND 812 PROSPECTIVE



The items in color in Exhibit 1 represent potential additions to the uncertainty analysis beyond the baseline approach. The yellow oval includes analyses of alternative scenarios for economic growth projections, as well as scenarios exploring the sensitivity of the modeling to marginal additional changes in criteria pollutant emissions from particular source categories. The impact of these assumptions will cascade through the entire analytical chain on both the benefit and cost sides of the analysis. The red oval represents an alternative treatment of uncertainty in the C-R function for the mortality impacts of exposure to $PM_{2.5}$, a major portion of the benefits of Clean Air Act regulations. This alternative treatment will be based on the results of the recent EPA-sponsored expert elicitation (EE) study of the mortality effects of $PM_{2.5}$ (IEc, 2006) and is expected to capture a broader range of uncertainty than the current approach applied in EPA's BenMAP model. The application of the EE study results will be applied in Physical Effects modeling and propagate through the rest of the benefits analysis.

Exhibit 2 lists additional "off-line" research and analysis we recommend incorporating into the current 812 study. As in the first prospective, starting with emissions profile development, each analytical element will feature a comprehensive qualitative evaluation of key uncertainties, presented in a summary table at the end of each chapter of the report. In addition, we recommend conducting a literature review assessing the current state of knowledge concerning uncertainty analysis for emissions profile development and for large-scale air quality models such as CMAQ, coupled with a phased sensitivity analysis using CMAQ and/or offline reduced form models; and a break-even analysis, which would attempt to identify the magnitude of error in emissions reductions necessary to generate negative net benefits and then characterize the likelihood of such a scenario.

We discuss each of the online and offline elements further below, roughly in the order of the elements they affect in the analytical chain. We begin with the treatment of scenario uncertainty, whose impacts cascade through the all parts of the analysis, and then focus on efforts to address uncertainty in specific elements. We then turn to an alternative approach to addressing uncertainty, a break-even analysis.

EXHIBIT 2 ADDITIONAL RECOMMENDED "OFFLINE" UNCERTAINTY ANALYSIS

ISSUE	RECOMMENDED APPROACH	ANALYTICAL ELEMENTS AFFECTED	OUTPUT
Emissions/Air Quality Parameter Uncertainty	Identification of key factors coupled with scalable sensitivity analysis	Emissions and air quality modeling	Characterization of current state of knowledge concerning uncertainty assessment for large-scale air quality modeling applications, and limited quantitative information on the influence of those factors on AQ outcomes.
Emissions Scenario uncertainty	Break-even analysis by sector using RSM and BenMAP, coupled with interviews of emissions modeling experts	Benefits side elements	Assessment of the likelihood of emissions reduction scenarios (or errors in emissions estimation) that would result in costs exceeding benefits for particular sectors.
Unquantified uncertainties	Comprehensive qualitative uncertainty analysis	All	Summary tables at the end of each report chapter describing key uncertainties and the size and direction of their likely impact on results (if known).

SCENARIO UNCERTAINTY

One of the most significant sources of uncertainty in the 812 analyses is temporal prediction uncertainty – that is, limitations in the ability of analysts to predict the future values of key model variables. As Resources for the Future (RFF) notes in its recent report on uncertainty analysis, “the future is inherently unknowable,” despite the best efforts of analysts to extrapolate based on past data trends and assumptions about human and market behavior (RFF, 2006). While the 812 analysis does not need to predict as far into the future as some analyses, the temporal scope is sufficiently large for temporal prediction errors to have significant impacts on the analysis results.

Many of these future variables, such as projected population, can be characterized by a continuous range or distribution of potential future values, but it is likely to be impractical for EPA to analyze such prediction uncertainty using Monte Carlo-style analysis. In addition, many future model parameters are inter-related and likely to co-vary (e.g., economic and population growth). Scenario analysis provides a more practical and convenient way to evaluate the potential impacts of alternative sets of related future assumptions. RFF’s recent report suggests scenario analysis is useful “for characterizing parameter uncertainties when the relative weights of the different parameter states cannot be estimated through expert judgment or a reasonable averaging method” (2006).

Economic growth/Population scenarios

The Project Team has previously proposed and finalized a set of emissions scenarios for use in the Second Prospective.[1] In addition to the base case, central growth scenario that will support the primary results, the Project Team will also generate full emissions scenarios for a low and high variant on future economic and population growth. Full benefit and cost analyses of these variants will be completed for the 2010 and 2020 results target years. The alternative low and high economic growth and population scenarios are integrated inputs, developed at the regional level, and derived from the same U.S. Department of Energy National Energy Modeling System (DOE NEMS) used to generate the primary results; they correspond to the Annual Energy Outlook (AEO) 2005 low and high growth scenarios. In practice, the emissions and population-dependent benefits results will be derived using alternative economic growth and population data for both the with-CAAA and without-CAAA; then, alternative economic benefit, cost, and net benefit results will be generated.

The results of these analyses will be useful for evaluating the sensitivity of the primary results to uncertainty in the key driver data that affects both benefit and cost-side estimates. The need to evaluate the uncertainty in economic growth and population projections was an important suggestion of the SAB Council in its review of the analytical plan for the Second Prospective.

Energy input scenarios

The Project Team has discussed two possible alternative scenarios to address uncertainty in energy inputs: 1) a "high renewables penetration" case; and 2) alternative natural gas price case(s). We discuss each in turn below.

The CAA-related motivation for a high renewables case is that shifts in energy production away from fossil fuels and toward non-polluting renewables could reduce the need for local controls to meet NAAQS requirements. However, we were unable to identify a scenario that is both reasonably likely and indicates a significant shift toward renewables by 2020. For example, we investigated the use of the AEO 2005 high renewables scenario, a scenario that assumes much faster than baseline reductions in costs for installed renewable capacity, but discovered that only between one and two percent of fossil fueled EGU or distributed generation capacity would be shifted to renewable production under this scenario by 2020. As a result, we currently have no plans to evaluate a high renewables penetration analysis.

The Project Team also considered analysis of the impact of alternative natural gas prices on costs and benefits. As noted in SAB Council comments on the Second Prospective Analytical Plan, natural gas price projections affect both costs and benefits, particularly for the EGU sector. That effect was recently demonstrated in the RFF (2006) analysis of a hypothetical EGU NO_x reduction program. RFF applied the same 30 percent reduction, 30 percent increase, and 70 percent reduction to both baseline and "with controls" emissions analyses, using the HAIKU EGU sector model and the Tracking and Analysis

---

[1] See memorandum from Jim DeMocker to the 812 Prospective II Files, "Scenario Specifications," August 3, 2005.

Framework (TAF) benefits assessment tool. Their results showed that both costs and benefits (measured from the difference in emissions between baseline and "with controls" scenarios) varied with the natural gas price projection, in roughly proportional fashion, therefore the overall effect on net benefits was insignificant. We subsequently confirmed with the lead RFF researcher that additional analysis using the 812 scenarios would likely yield similar results - that is, roughly proportional changes in benefits and costs and a small effect on net benefits.

In addition, EPA has conducted analyses of high natural gas price scenarios as part of their effort to assess alternative legislative proposals for Congress. EPA applied the IPM model to two alternative natural gas price trajectories: 75 percent and 150 percent "scale up" relative to the base case AEO-derived trajectory. The range of costs assessed in the high price cases were therefore between $5 and $8 per mmBTU (average Henry Hub prices, 2004$), compared to a base case of roughly $3.50 per mmBTU. The results indicated that costs of compliance for Clear Skies decreased about $200 million in 2010 and increased between $300 and $400 million in 2020.[2] EPA also reported that emissions of sulfur dioxide increased between 600,000 and 800,000 tons nationwide in 2010, and decreased between 200,000 and 300,000 tons in 2020.

Based on the results of these two efforts, and in the interest of allocating our resources to the areas of most urgent need, we do not recommend additional natural gas price sensitivity analysis for the purposes of this uncertainty analysis. We believe that referencing the results of these existing analyses, however, will provide important information to the readers of the Second Prospective. We continue to consider whether it might be useful to conduct a screening level analysis of the air quality and human health implications of the EPA Clear Skies high natural gas price scenarios. It may be possible, at relatively low cost, to apply the RSM to the results of the IPM sensitivity runs to assess air quality implications of those runs, and to link those results to BenMAP assessments of health effects and valuation to yield an estimate of the effects on benefits to compare to the above referenced cost results. We welcome your comments on this option.

Emissions scenario uncertainty by emitting sector

The Second Prospective analytical plan includes analysis of scenarios designed to disaggregate the effect of emissions regulation by major emitting sector. There are two types of analysis that have been proposed: 1) Parse the primary cost and benefit results by emitting sector, by "turning off" the CAAA regulations for each of the five major sectors one at a time; and 2) Estimate the marginal benefits of additional emissions reductions, beyond those mandated by the current CAAA, for each major emitting sector. The first analysis will involve the full primary model set - including runs of EPA's reduced –form CMAQ Response-Surface Metamodel (RSM) coupled with BenMAP. The second analysis will involve a potentially large number of RSM/BenMAP runs. The policy relevance of the latter analysis is clear - it will provide an estimate of what benefits might

---

[2] Of the three legislative proposals analyses, Clear Skies is closest to the CAIR scenario used in the Second Section 812 Prospective analysis. A fact sheet and briefing charts on the results of the IPM runs can be found at: www.epa.gov/airmarkets/progsregs/cair/multi.html

be gained, in absolute and relative terms, by additional emissions reductions in each of the major sectors.

Neither of these analyses were originally designed as uncertainty analyses. The latter analysis, however, does provide a sensitivity test of benefits relative to emissions inputs that, if coupled with other information, could be useful in evaluating the impact of key uncertainties. By itself, it provides an answer to the question of how benefits might vary if we have underestimated the impact of the CAAA regulations on emissions reductions "on the ground". The likelihood of that scenario, however, is unknown without further information. For example, as noted above, if we have information that encouraging use of renewable energy might reduce emissions by some number of tons of sulfur and nitrogen oxides beyond those reductions required by the CAAA, the marginal impact analysis could be designed to assess the benefits of that scenario. In the next section, we discuss how this emissions scenario sensitivity analysis might be coupled with other available information to assess the effect of some emissions uncertainties on benefits estimates.

EMISSIONS AND AQM UNCERTAINTY

In this section of the white paper, we first discuss a broad, phased approach to identifying and characterizing uncertainties in emissions and air quality modeling proposed by Sonoma Technology (Sonoma). We then discuss a specific proposed sensitivity analysis aimed at addressing potential errors identified in IPM model predictions of EGU emissions for the 2000 target year.

PHASED APPROACH FOR CHARACTERIZING EMISSIONS AND AQM UNCERTAINTIES

A systematic assessment of the impact of all uncertain emissions parameters on benefits and costs for a study of the scope of the Second Prospective is not feasible at this time. Several basic issues in characterizing the uncertainty of input parameters and building an appropriate covariance matrix, for example, remain unresolved for such a large-scale analysis. There are many examples in the existing literature, however, of the impact of air quality model inputs in smaller scale studies, particularly when the scope of the analysis is limited to air quality outcomes. In addition, there are a few examples of parameter uncertainty analyses for larger scale analyses (e.g., for the OTAG region, see Hanna et al. (2000)).

In September 2006, Sonoma updated its previous 2004 literature review that summarized much of the existing uncertainty literature related to the possible application of existing approaches to the Second Prospective. (The Sonoma memo is included as an attachment to this white paper). The result of Sonoma's assessment was a recommendation to pursue a four-component framework for estimating the uncertainties in the integrated air quality modeling system (IAQMS), which includes the emissions, meteorological, and air quality models:

1. **Identify and quantify internal and external uncertainties in the IAQMS** – Summarize known internal (model formulation, parameterizations etc.) and external (model input) uncertainties based on the current literature. Prioritize uncertainties based on the expected IAQMS sensitivity to those uncertainties. In cases where a particular uncertainty is poorly defined or the literature is out of

date, literature could be supplemented by expert opinion. Uncertainties will be ranked based on their potential to affect the specific model with which they are associated and their overall effect on the IAQMS response to emission changes. The results of this phase would be a tabular summary describing key uncertainties that indicates the type of uncertainty (e.g. measurement, input, model), an estimate of the direction of potential bias for the net benefits estimate, and an assessment of its likely significance (e.g., Potentially Major) These tables and rankings would be informed by both literature review and a proposed 2-day workshop with experts in the relevant areas from government, academia, and industry. A key task will be to structure the workshop effectively so that participants are well-conditioned as to the objectives of the process, key terminology, as well as the analytical model linking uncertainties in emissions and AQM, so they can provide well-informed judgments about the ranking of uncertainties.

2. **Assess importance of uncertainties** – Identify the elements from Phase 1 that have the greatest impact on model predictions. Uncertainties that are shown to significantly affect model predictions during this phase will be explored further in Phase 3 to determine whether or not they affect model response to emissions changes The Sonoma memo reviews existing literature on techniques for sensitivity and process analysis that might be applicable, as well as existing applications reported in the literature to the emissions/air quality modeling analytic suite, including EPA's CMAQ Response Surface Metamodel (RSM). A key task will be reviewing which of these techniques might be applicable to current model sets and reduced-form tools. In addition, EPA/ORD may have some insights in this area.
3. **Assess uncertainties in IAQMS output** – Estimate the effects of propagating uncertainties through an IAQMS to quantify uncertainties in projected changes in air quality. Goals for Sonoma's proposed third phase would be to identify which uncertainties affect the IAQMS's response to emission changes and provide a range of air quality outcomes for use in the rest of the analytic chain. Doing so could involve limited additional CMAQ runs or past literature analyses of the impacts of emissions uncertainties, meteorological uncertainties, chemistry uncertainties, and model formulation uncertainties, as well as runs of EPA's RSM to assist in evaluating uncertainties in emissions.
4. **Final Synthesis** – Sonoma's proposed final phase would attempt to summarize and synthesize individual estimates of sensitivity and uncertainty in the various elements studied in prior phases to ideally produce estimates of the range of air quality outcomes in terms of relative changes in ozone and $PM_{2.5}$ for the scenarios being modeled, which could be carried through to the subsequent analytical steps of health effects estimation and valuation. The likelihood of successfully developing a semi-quantitative "bottom-up" estimate of uncertainty is unclear at this point. At a minimum, this phase will produce an updated tabulation and prioritization of uncertainties describing potential sources,

indicating direction of potential bias for the net benefits estimate and likely significance relative to key uncertainties in the net benefit estimate.

Sonoma noted that the assessment could be carried out at varying levels of detail, each requiring a commensurate level of effort, implying that the framework could be scaled to available resources.

We propose to follow the framework suggested by Sonoma to assess uncertainties in the emissions and AQM analytical steps. To the extent the project is successful in propagating uncertainties through application of the RSM, the results in terms of AQ outcomes might also be coupled with BenMAP. The result would provide new information on the impact on benefits estimates of a range of plausible emissions and air quality modeling uncertainties. Even if the analysis can only be conducted at a regional or local scale, the results ought to prove useful in identifying key areas to focus in reducing the overall uncertainty of future benefit-cost assessments.

ASSESSING THE IMPACT OF MODEL UNCERTAINTY ON EGU EMISSION ESTIMATES

In reviewing the draft emissions projections for the EGU sector, questions were raised within the 812 Project Team regarding the validity and reliability of the year 2001 IPM validation run being proposed for adoption as the 812 study's target year 2000 results. The results of that run are provided in Chapter 3 of the document, *Emission Projections for the Clean Air Act Second Section 812 Prospective Analysis Draft Report*, June 21, 2006. Appendices B and C of that document provide further detail comparing the results of the with-CAAA scenario for the year 2001 with actual historical emissions rates, fuel prices, and allowance prices.[3] Differences between the spatial distribution of emissions as modeled by IPM and the actual spatial distribution from continuous emissions monitor (CEM) data, and differences in modeled versus actual fuel and allowance prices for the historical (with-CAAA) case, have led us to consider the possibility of an alternative approach for modeling the effect of the CAAA on the EGU sector in the year 2000 or 2001.

In August 2006, EPA presented to the Air Quality Modeling Subcommittee (AQMS) an alternative approach based on a counterfactual without-CAAA scenario for EGU sector sulfur oxide emissions. This alternative approach closely follows an approach developed by Dr. A. Denny Ellerman of the Massachusetts Institute of Technology. The details of that approach can be found in an IEc memo to Jim DeMocker of EPA/OAR/OPAR dated July 24, 2006 and attached to this white paper (Attachment #3). The AQMS agreed that the approach would be a suitable basis for a sensitivity analysis, but did not recommend the approach be used in as the primary basis for assessing emissions and costs for the 2001 EGU analysis. Based on that advice, we propose to conduct a sensitivity analysis that employs CEM data for the 2000/2001 with-CAAA case, and the Ellerman-method counterfactual for the 2000/2001 without-CAAA case, for EGU sector sulfur dioxide emissions.

---

[3] See, in particular, Exhibits B-1 through B-3 in Appendix B for emissions comparisons and Exhibit C-4 in Appendix C for fuel and allowance price comparisons.

HEALTH EFFECTS UNCERTAINTY

The primary focus of uncertainty analysis for health effects is the uncertainty in premature mortality associated with changes in exposure to $PM_{2.5}$, because the deaths avoided dominate all other categories in terms of monetized benefits. We describe below our proposed alternative treatment of uncertainty in the concentration-response (C-R) function for this endpoint, followed by a discussion of our analysis of alternative model structures for the cessation lag that may accompany reductions in mortality risk due to reductions in PM.

PM/MORTALITY C-R FUNCTION

Since 2003, EPA has been working with IEc to apply expert elicitation (EE) methods to generate improved estimates of the uncertainty concerning the change in annual adult mortality associated with a 1 $\mu g/m^3$ change in ambient annual average $PM_{2.5}$ concentrations. The goal of this effort is to elicit from the experts probabilistic distributions of the mortality C-R coefficient that reflect potential sources of uncertainty not captured in the standard errors reported in epidemiological studies (and thus typically not captured in EPA's uncertainty analysis for health effects). Between 2003 and 2004, IEc conducted a pilot study for EPA in which we interviewed 5 experts selected from the membership of two relevant NAS committees. In September 2006, IEc completed a full-scale elicitation of 12 peer-nominated experts with expertise in epidemiology, toxicology, and medicine. The study yielded 12 probabilistic distributions of the impact of a 1 $\mu g/m^3$ decrease in annual average $PM_{2.5}$ in the U.S. on annual, adult mortality.

The final EE study results not only provide twelve alternative representations of parameter uncertainty in the C-R function coefficient; it also reflects model uncertainty in the mortality C-R function across experts. The elicitation protocol for the full-scale study has been revised to allow experts greater freedom in specifying alternative shapes for the concentration-response function, including the option to include a threshold. These alternative model specifications add another dimension to uncertainty analysis for mortality effects not captured in EPA's current approach.

EPA's Office of Air Quality Planning and Standards applied the pilot study results as supplemental alternative assessments of uncertainty in benefits for EPA's Non-Road Diesel Engine and CAIR rules (EPA, 2004 and EPA, 2005).[4] In both cases, the preferred approach for applying the EE results to benefits assessment was to generate separate probabilistic estimates of mortality benefits using each expert's C-R distribution and then pool those estimates using equal weighting to obtain a combined benefits distribution. In the application in the CAIR rule, EPA presented the distributions based on individual expert responses in the main text of the benefits chapter and then presented a pooled estimate in an appendix to that document. Most recently, EPA has applied the results of the full-scale EE study in the benefits analysis of the PM National Ambient Air Quality Standard (PM NAAQS, EPA, 2006). The full-scale EE uncertainty analysis for the PM

---

[4] See Appendix 9B to the EPA's 2004 Nonroad Diesel RIA (http://www.epa.gov/nonroad-diesel/2004fr/420r04007j.pdf) and Chapter 4 and Appendix B of EPA's 2005 CAIR RIA (http://www.epa.gov/interstateairquality/pdfs/finaltech08.pdf) for more information.

NAAQS did not include a pooled estimate; it presented 12 separate benefit distributions based on application of the individual expert C-R distributions from the study.

We recommend using the results from the full-scale study, because of the significant improvements it features over the pilot study, including:

- A revised protocol with a more systematic approach for cataloguing and assessing expert concerns related to confounding, effect modification, exposure issues, and other potential sources of error or bias in published estimates of effect;
- A protocol designed to allow experts greater flexibility in specifying the shape of the C-R function as well as their distribution of mortality effect estimates;
- Larger expert panel of 12 selected from an unrestricted pool of potential experts via a peer nomination process (versus five experts selected from a restricted pool of experts in the pilot);
- Improved "briefing book" materials for experts, including a CD containing over a hundred relevant studies, plus data on air quality in the US, population demographics, health status, summaries of published effect estimates, and data on other factors (air conditioning use, housing stock, PM composition) that may affect experts' judgments;
- A new real-time feedback system using spreadsheet models, Crystal Ball™ probabilistic modeling software, and WebEx internet teleconferencing to provide experts with graphs and data during the elicitation. System allows experts to visualize their distributions, assess the effect of judgments about causality and threshold, compare their results against published mortality effect estimates, and estimate the change in deaths associated with $PM_{2.5}$ reductions;
- A pre-elicitation workshop held in January 2006 that better prepared experts in advance of the interviews by providing training in the elicitation process; by reviewing and discussing the elicitation protocol; and by allowing experts to share and critique data and analyses they believe are relevant to the questions in the protocol; and
- A post-elicitation workshop held in June 2006 following completion of the interviews that allowed for a final discussion of themes that emerged in expert responses, differences in interpretation of key studies used to support responses, and any areas of confusion that arose during the interviews..

Differences in the content and administration of the elicitation process cited above make it inappropriate to pool the results of experts from the pilot study with those from the full-scale study for purposes of uncertainty analysis. Therefore, we recommend that only the results of the full-scale study be used in any uncertainty application.

We further recommend that the full-scale EE results be applied individually to develop 12 separate benefit distributions for the 812 analysis, as was done for the PM NAAQS RIA. We recommend against calculating a pooled estimate of benefits from these 12 distributions. Our primary concern is that a combined estimate masks inter-expert differences of opinion, condenses uncertainty, and likely produces a distribution with which none of the experts would agree. The full scale study did not present combined responses, based on feedback from several sources, including peer reviewers of the pilot study, the Health Effects Subcommittee of EPA's Science Advisory Board Council on Clean Air Compliance Analysis, and health experts invited to EPA's April 2005 EE Symposium (IEc, 2006). Furthermore, we are unaware of any combination approach that would yield a non-arbitrary blending of effect estimates, including the use of equal weights. The study did not include calibration questions, such as those that have been used by Cooke (1991), for weighting purposes because of concerns both about implementation and about whether such questions are a fair measure of expertise in the subject area. The panel of experts was not selected to represent a statistical sample of expert opinion and thus cannot be weighted on that basis, nor do we think weighting by prevalence of opinion would necessarily lead to a more accurate result. Ultimately, we believe that presenting separate results is both most informative and least controversial.

EPA's recent update to the BenMAP model includes the 12 expert distributions from the full-scale study, so no additional programming or modeling would be required. This approach would satisfy NRC recommendations to incorporate expert judgment to address data issues while also, as they suggest, distinguishing it from empirically-derived estimates of uncertainty.

DIFFERENTIAL TOXICITY OF PM COMPONENTS

One element of uncertainty in EPA's estimates of reductions in PM-related mortality that was not covered in the current expert elicitation is the potential for differential toxicity of specific PM components. This issue was not addressed in the EE study questioning of experts because a number of PM health effect experts who attended EPA's April 2005 EE Symposium concurred that the literature on differential toxicity is insufficient to support knowledgeable estimates of differential impacts of PM components (IEc, 2006). We concur with this opinion and would not advocate pursuing sensitivity or other analysis of the effects of modeling differential toxicity uncertainty at this time.

We did review a draft exploratory sensitivity analysis conducted by EPA and submitted to the PM NAAQS RIA docket that included differential toxicity of PM components as one of the factors evaluated (Abt Associates, 2006). We found that particular analysis of differential toxicity to be handicapped by the lack of available data on differential toxicity. It was unable to distinguish the likelihood of alternative differential toxicity scenarios and therefore provides only limited insight into the impact of this uncertainty on mortality benefit estimates at this time. We believe that the only differential attribution of benefits that can be supported at this time would be an analysis such as that performed for the PM NAAQS RIA and included in Chapter 5 of that report (EPA, 2006). That analysis apportioned benefits by the major PM component species, including ammonium sulfate, ammonium nitrate, elemental carbon, organic carbon, and crustal material. For each attainment strategy, EPA calculated the percent of benefits associated with each species based on the fraction of the total expected population-weighted change in total $PM_{2.5}$ accounted for by each component species, as shown in Exhibit 4.

EXHIBIT 4 EXAMPLE ANAYSIS OF BENEFITS BY PM COMPONENT

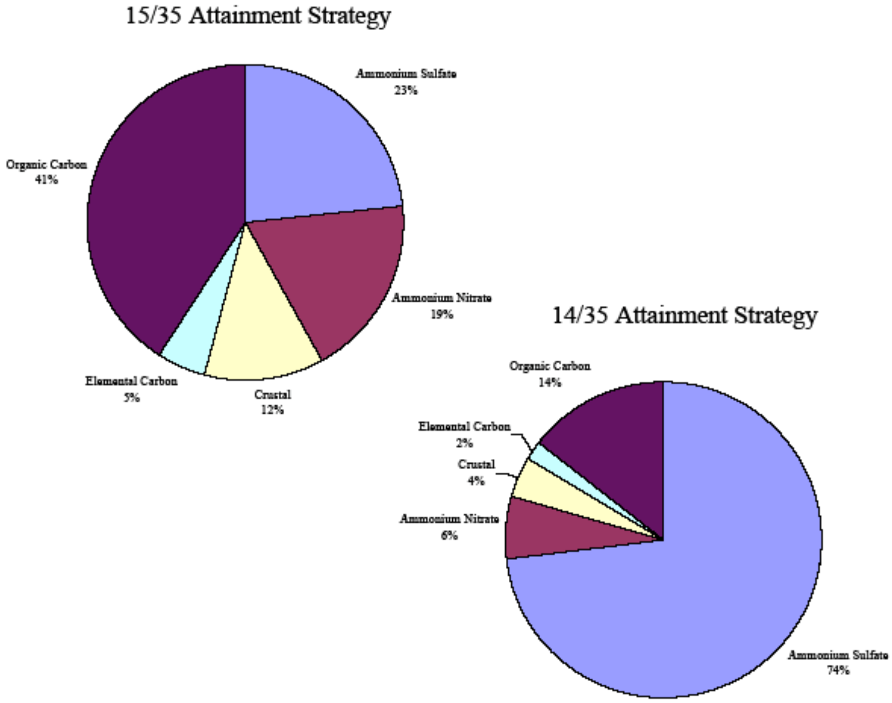


Figure 5-16. Proportion of Population-weighted Reduction in Ambient Annual PM2.5 Associated with PM2.5 Components for Modeled Attainment Strategies

Source: U.S. EPA (2006) Regulatory Impact Analysis for the 2006 Revision to Particulate Matter National Ambient Air Quality Standards. http://www.epa.gov/ttn/ecas/ria.html.

CESSATION LAG

EPA's treatment of a cessation lag for the effect of PM reductions on mortality outcomes was recently re-evaluated. Where the lag was formerly modeled over five years, with half of the effect being manifest in the first two years and the remainder in equal amounts in the ensuing three years, the current approach reflects a longer and more complex lag structure with short-term, medium-term, and long-term effects. At the time the SAB Council's Health Effects Subcommittee recommended this change, however, they also noted the lack of a definitive study to model cessation lags, and suggested conducting additional uncertainty analyses to gauge the impact of alternative lag structures.

Since that time, the Second Prospective Project Team has developed a population simulation model that provides a flexible and nimble approach to assessing the impact of alternative lag structures on mortality, age structure, life expectancy, and life-years saved/lost. In addition, OAQPS has updated BenMAP to allow for alternative lag structure assessment. We believe both of these tools may provide useful perspective on the impacts of uncertainty about the cessation lag structure on monetized estimates of avoided mortality.

VALUATION UNCERTAINTY

In previous analyses valuation uncertainties have been evaluated within BenMAP, relying largely on cross-study variability in key input parameters such as the value of statistical life (VSL) and unit values for morbidity. For the second prospective, we propose to follow existing Agency practice for evaluation of uncertainty in the VSL, and to assess the impact of uncertainty in applicable longitudinal income elasticity for the first time.

VSL

We see three potential options for analyzing uncertainty in VSL:

- **Apply approach from First Prospective (Weibull distribution).** Under this option, EPA would use the 26-study Weibull distribution with a mean of $6.1 million in 1999 dollars that was employed the First Prospective and is described in EPA's *Guidelines for Preparing Economic Analyses* (2000). Although EPA and OAR have changed their estimates of VSL since the First Prospective, the revised distribution has not yet been reviewed by the SAB. The Weibull distribution represents longstanding EPA policy and has the benefit of having been exhaustively reviewed by the SAB.
- **Apply current EPA approach (Normal distribution).** In recent OAR analyses, such as the March 2005 RIA for the final Clean Air Interstate Rule (CAIR) , EPA has used an ad hoc approach for premature mortality valuation based loosely on two VSL meta-analyses. The central VSL used in the CAIR RIA was $5.5 million in 1999 dollars (reflecting 1990 income levels). A distribution of values was also presented, for use in uncertainty analysis - the central value is the mean of a normal distribution, with a 5th percentile at $1 million and a 95 percentile at $10 million. The implied 90 percent confidence interval is based on two meta-analyses of the wage-risk VSL literature: $1 million represents the lower end of the interquartile range from the Mrozek and Taylor (2002) meta-analysis and $10 million represents the upper end of the interquartile range from the Viscusi and Aldy (2003) meta-analysis. While this approach has been applied in some of OAR's recent RIAs, it has yet to undergo review by the SAB or other outside expert peer review process, and its distribution may not be consistent with past SAB advice. As noted in the SAB Council's 2004 Advisory, "If called upon to recommend just a single meta-analysis at this point, the Council Panel would recommend a primary focus on the Viscusi-Aldy estimates based on U.S. wage studies."

- **Develop New Approach.** EPA could propose a new distribution based on an in-depth analysis of two VSL meta-analyses – the Viscusi and Aldy analysis, which was favorably reviewed by SAB as noted above, and the Kochi et al. meta-analysis of VSL (2006) which has been published since the time of the last SAB review of VSL studies.

Whichever approach EPA chooses, we propose to include adjustments for changes in willingness-to-pay (WTP) over time as described below.

INCOME ELASTICITY

As noted above, the \$5.5 million estimate for a VSL has been interpreted to reflect WTP consistent with 1990 income levels. Consistent with current regulatory analysis practice and at least some past SAB Council and EEAC advice, EPA proposed in the May 2003 Second Prospective Analytical Plan to adjust unit values for health effects and recreational visibility using available estimates for income elasticity (culled from a 1999 literature review).[5] In its review of the Second Prospective Analytical Plan, however, the SAB Council did not support including this type of income adjustment in the primary analysis. It did, however, leave an opening for including an income adjustment, stating that if "an adjustment of this type is considered essential even at this stage in the analytic process, the Agency should be especially prudent in qualifying it and present the results in a format that is as transparent as possible."

In response to the SAB Council advice, EPA updated the 1999 literature review for relevant income elasticity values in September 2004.[6] That review recommended low, central, and high income elasticity values for three categories of health effects - mild effects, chronic and severe effects, and mortality. For the Second Prospective analysis, we propose to adopt values from the updated literature review, but to ensure consistency with the broader analysis, to use projections of GDP/income and population growth from AEO 2005. EPA has recently updated BenMAP to perform sensitivity analysis with alternative income elasticity values.

BREAK-EVEN ANALYSIS

A well-constructed and well-communicated uncertainty analysis should give the reader or decision-maker a sense of the likelihood of an adverse outcome – in this case, the likelihood that CAA regulations may result in negative net benefits to society. Because of the significant challenges involved in modeling upstream uncertainties in emissions and AQM, it is extremely difficult to take a "bottom-up" approach to answering this question. An alternative approach would be to address this question using break-even analysis, exploring how large changes in emissions for certain sectors would have to be for benefits to fall below costs, and then characterizing the likelihood of those scenarios.

---

[5] See page 4-18, Table 4-3 of USEPA, 2005, "Regulatory Impact Analysis for the Final Clean Air Interstate Rule," for a summary of central elasticity estimates used in that analysis. The categories for adjustment of valuation estimates included the following: premature mortality; severe and chronic health effects; minor health effect; and visibility.

[6] Memorandum to Nona Smoke and James DeMocker, EPA/OAR/OPAR, from James Neumann and Sarah Brennan, Industrial Economics, "Responding to SAB Council Comments on the May 2003 Draft Analytical Plan for the Section 812 Second Prospective - Options for Adjusting WTP Estimates to Reflect Changes in Real Income", September 30, 2004. See Exhibit 6 on page 19 of that memo for a table of low, central, and high values for health effects valuation adjustments.

The development of new tools such as EPA's RSM makes it possible to attempt this type of analysis. By coupling the RSM with BenMAP, we can perform iterative reduced form analyses of the impacts of increasing estimates of potential error in emissions estimates for individual sectors in order to identify potential break-even points. We will need to develop a baseline without-CAAA case for RSM. To facilitate the analysis, we will focus on impacts on premature mortality related to changes in $PM_{2.5}$, which constitute the majority of monetized CAA-related benefits. We will begin by reviewing published literature and available data on emissions uncertainty to identify potential starting points for this analysis. We will work closely with EPA and Pechan to ratchet up or down emissions of PM and its precursors in ways that are scientifically and technologically plausible for a given sector. We will then run these emissions scenarios through RSM and BenMAP, generating monetized mortality benefit estimates and comparing these with cost estimates from the main analysis. We will attempt to identify scenarios where more than half of the resulting distribution of mortality benefits is less than the associated CAA costs.

Once we have identified scenarios that result in negative net benefits for each sector, we will consult with emissions experts on the 812 project team and at EPA, reviewing the assumptions associated with the scenarios and attempting to qualitatively characterize the conditions under which and/or the likelihood that these scenarios might materialize.

Recognizing the limitation that this will address the impact of only one upstream uncertainty (emissions), we will present by sector the error in emissions necessary to produce negative net benefits at least 50 percent of the time along with descriptions characterizing the plausibility and likelihood of these scenarios occurring. The results should provide some policy relevant insights into the robustness of benefit/cost results for particular sectors, even in the absence of a start-to-finish quantitative propagation of uncertainty in the 812 analytical chain.

PRESENTING UNCERTAINTY RESULTS

Uncertainty analysis can yield an array of alternative estimates and probabilistic distributions that can be complex and bewildering to interpret. A well-conducted uncertainty analysis is useless without careful consideration for how best to convey the story told by the data and associated probabilities. In this section of the white paper, we present our approach to presenting qualitative assessments of uncertainty in the 812 analysis and our recommendations for presenting results of the quantitative assessments described above.

QUALITATIVE UNCERTAINTY

We support the continued use of summary tables for qualitative assessment of key uncertainties in each analytical step. IEc will maintain draft lists of uncertainties in each element, revising them as necessary as the analysis progresses. These lists will describe the source of error, indicate most likely direction of potential bias (if known) and categorize the error as to its significance. Upon completion of the analysis, IEc will circulate the lists to relevant experts on the 812 project team and within EPA for critical review of IEc's assessment of the key uncertainties.

QUANTITATIVE UNCERTAINTY

We recommend presenting quantitative uncertainty analysis results using a combination of tables and graphs that may include box and whisker plots, cumulative density functions (CDFs), probability density functions (PDFs), or tornado diagrams. We focus on these graphs because they are among the most straightforward and elegant ways to convey information about uncertain results. Since they have different strengths, they can be used in combination, each helping to put the results in context.

We believe the tables in the first prospective analysis serve as good models for summarizing quantitative results. (See Exhibit 5 for an example.) The table in Exhibit 5 presents estimates of monetized benefits from the primary 812 analysis. This table satisfies a key recommendation of the NRC panel; it avoids the appearance of an unwarranted degree of certainty. Instead of a single point estimate, the table presents a range of estimates of monetized benefits from the primary analysis: a central estimate, a low estimate, and a high estimate. In addition, values are rounded to two significant figures to avoid false precision. The table also carefully labels the estimates as Primary Low, Primary Central, and Primary High, while acknowledging in a footnote that the low and high estimates are the 5th and 95th percentiles, respectively, of a partial uncertainty distribution of CAA-related benefits. We believe this approach avoids giving the impression that the bounds of the range are the 5th and 95th percentiles of the entire uncertainty distribution, while satisfying the need for transparency about the statistical modeling performed. We believe this presentation is a reasonable template for presenting results in the second 812 prospective analysis.

The tables of results can be complemented with one or more graphs to explain the influence of alternative assumptions and uncertainties on results. We have excerpted examples from the recent RFF report (2006) as well as EPA's Non-Road Diesel RIA to illustrate some available options.

EXHIBIT 5 EXAMPLE TABLE OF UNCERTAINTY RESULTS

CRITERIA POLLUTANT HEALTH AND WELFARE BENEFITS IN 2010			
	MONETARY BENEFITS (IN MILLIONS 1990$)*		
BENEFITS CATEGORY	PRIMARY LOW	PRIMARY CENTRAL	PRIMARY HIGH
MORTALITY			
Ages 30+	14,000	100,000	250,000
CHRONIC ILLNESS			
Chronic Bronchitis	360	5,600	18,000
Chronic Asthma	40	180	300
HOSPITALIZATION			
All Respiratory	76	130	200
Total Cardiovascular	93	390	960
Asthma-Related ER Visits	0.1	1.0	2.8
MINOR ILLNESS			
Acute Bronchitis	0.0	2.1	5.2
URS	4.2	19	39
LRS	2.2	6.2	12
Respiratory Illness	0.9	6.3	15
Mod/Worse Asthma[1]	1.9	13	29
Asthma Attacks[1]	20	55	100
Chest Tightness, Shortness of Breath, or Wheeze	0.0	0.6	3.1
Shortness of Breath	0.0	0.5	1.2
Work Loss Days	300	340	380
MRAD/Any-of-19	680	1,200	1,800
WELFARE			
Decreased Worker Productivity	710	710	710
Visibility - Recreational	2,500	2,900	3,300
Agriculture (Net Surplus)	7.1	550	1,100
Acidification	12	50	76
Commercial Timber	180	600	1,000
AGGREGATE RANGE OF BENEFITS[2]	26,000	110,000	270,000

Note:

* The estimates reflect air quality results for the entire population in the US.

[1] Moderate to worse asthma, asthma attacks, and shortness of breath are endpoints included in the definition of MRAD/Any of 19 respiratory effects. Although valuation estimates are presented for these categories, the values are not included in total benefits to avoid the potential for double-counting.

[2] The Aggregate Range reflects the 5th, mean, and 95th percentile of the estimated credible range of monetary benefits based on quantified uncertainty, as discussed in the text.

Source: U.S. EPA (1999) The Benefits and Costs of the Clean Air Act: 1990 to 2010.

Box and Whisker Plots

These "top-down" views of uncertainty distributions can be very effective for comparing alternative results. Exhibits 6 and 7 show examples of the use of these plots. In both cases the ends of the "whiskers" represent the 5th and 95th percentiles of the distribution, the box represents the interquartile range between the 25th and 75th percentiles (where half of the possible values are found) and a symbol inside the box indicates the median. (Both figures also superimpose the mean value for each distribution, though this is not required.) Exhibit 6 from the RFF analysis compares the influence of alternative source-receptor assumptions on the distribution of benefits. From a glance at this graph, it is clearly evident that these alternative assumptions have minimal impacts on benefits uncertainty. Another useful feature of this graph is the line indicating the magnitude of costs, which are unaffected by these assumptions. Superimposing the cost line makes it easy to see the fraction of the distribution for which benefits exceed costs, and whether the mean or median estimate would have positive net benefits. The other example in Exhibit 7 shows the use of box plots to illustrate the influence of pooling the mortality benefit results from the EE pilot study, and how the individual and combined results compare to uncertainty represented by the standard error of the Pope et al., 2002 estimate. The first 812 analysis report included a simplified box plot graph (Figure 8-2 in that report) showing the contribution of key parameters to overall uncertainty. We believe box and whisker plots can be used in this analysis to compare the results of scenario analysis, describe the results of applying the full-scale PM/mortality EE results, and assess the contribution of the various uncertain elements evaluated to overall quantified uncertainty.

PDFs and CDFs

The use of the more traditional representations of probability distributions can be more familiar to some readers and can provide more detail concerning the shape of the distributions than can be seen from the box and whisker plot. Exhibits 8 and 9 show an example of the use of PDFs and CDFs respectively. In our opinion, PDFs are of more limited usefulness compared to CDFs, though the two can complement each other. Exhibit 9 shows the use of CDFs to compare results of alternative policy options under three different scenarios. The CDF can provide a very clear way of distinguishing among policies under each scenario; for example, in this exhibit it is easy to see that compulsory belts dominate under scenario "a" but become inferior to airbags as one moves to scenarios b and c. These figures can also be very effective for illustrating that "X" percent of values have positive net benefits, a useful way of presenting uncertainty to a decision-maker. We acknowledge that the recent RFF study (2006) found CDFs were less helpful as communication tools for uncertainty; nonetheless, we believe CDFs are still a valuable element of a suite of graphical tools for presenting uncertainty. For example, as noted in the RFF study, PDFs tend to focus respondents on the tails of the distribution, while CDFs provide a more neutral presentation of uncertainty.

Tornado diagrams

Tornado diagrams are special bar charts typically used to neatly summarize the results of sensitivity analysis. Exhibit 10 shows an example. In this case, the bars represent correlation coefficients of the output with the various inputs, sorted on the absolute value of the coefficient, from highest to lowest. The tornado diagram can also be used to show the change in output associated with a specific change in each input (e.g., 10 percent, one standard deviation). By centering the graph on zero, it elegantly conveys not only the magnitude of the sensitivity to a given input, but also the directionality of the effect. A limitation of these diagrams, however, is that they do not show impacts of simultaneous adjustments to multiple inputs. For a more complex sensitivity analysis, a response surface would be necessary; however, this is likely beyond the scope of the 812 analysis.

The usefulness of the tornado diagram for presenting results in the 812 analysis depends on the ease with which we can compare results across the online and offline analyses being proposed. For example, sensitivity to the choice of model used to represent cessation lag could not be included in a tornado diagram. Sensitivity to assumptions about economic growth may be better captured using CDF diagrams, since we are not assessing the effect of incremental changes in growth on benefits. The diagram may be most useful for demonstrating the sensitivity of benefits results to marginal changes in emissions across different sectors, because a similar size change can be used for each sensitivity test, allowing for easier comparison of results.

EXHIBIT 6 BOX AND WHISKER PLOT EXAMPLE 1

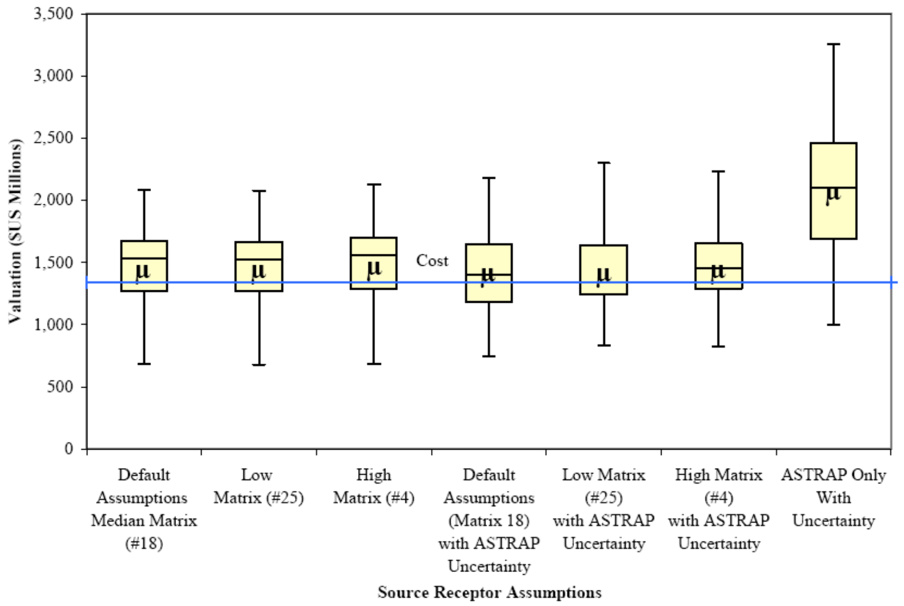
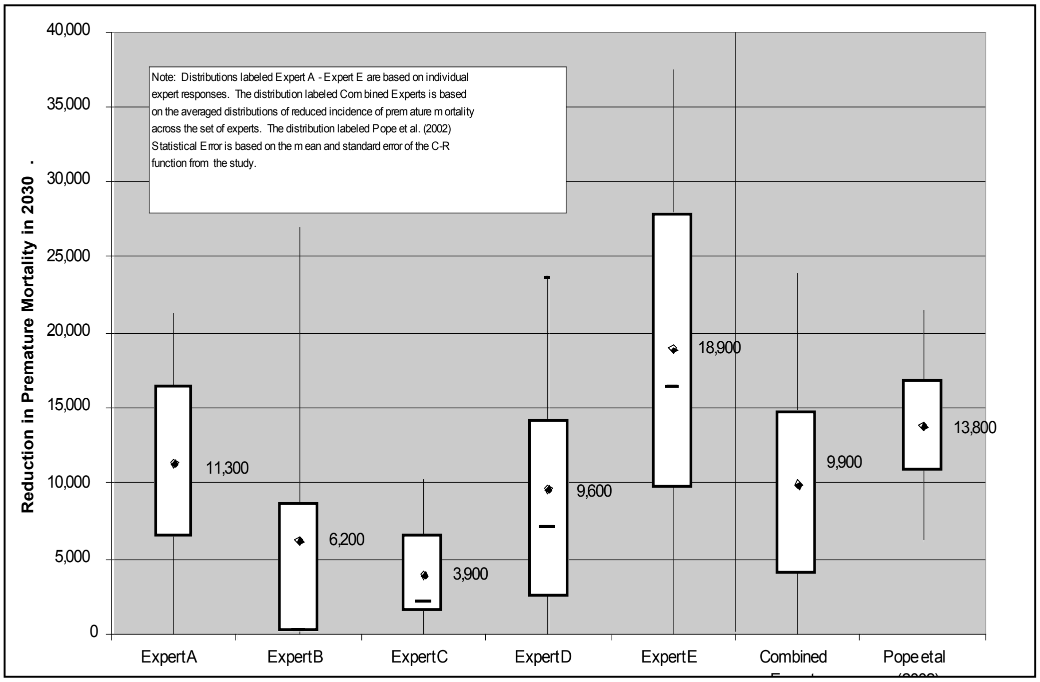


Figure 3-4. Benefits and Costs of Large NO_x Reduction Policy from CAIR Baseline: Using Alternative URM and ASTRAP Source–Receptor Assumptions, 2025

Source: Resources for the Future (2006) *Not a Sure Thing: Making Regulatory Choices Under Uncertainty* , http://rff.org/rff/Documents/RFF-Rpt-RegulatoryChoices.pdf

EXHIBIT 7 BOX AND WHISKER PLOT EXAMPLE 2

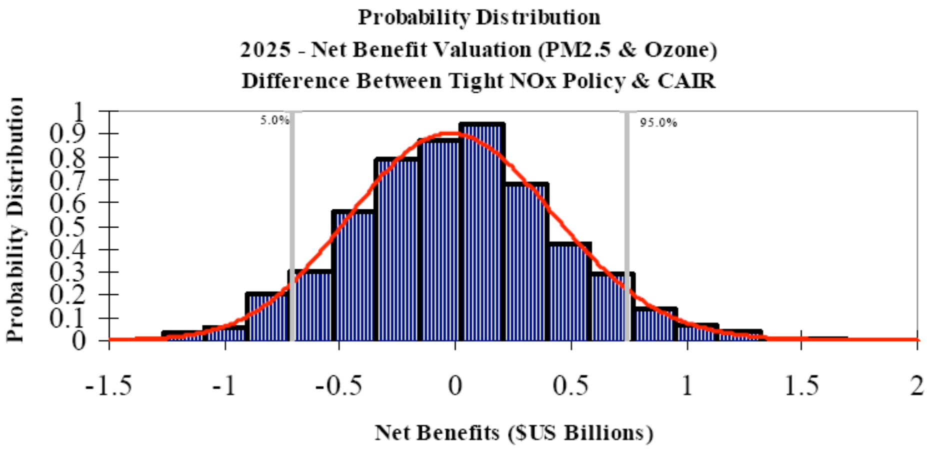


Note: Distributions labeled Expert A - Expert E are based on individual expert responses. The distribution labeled Combined Experts is based on the averaged distributions of reduced incidence of premature mortality across the set of experts. The distribution labeled Pope et al. (2002) Statistical Error is based on the mean and standard error of the C-R function from the study.

Source: EPA (2004) *Final Regulatory Analysis: Control of Emissions from Nonroad Diesel Engines*

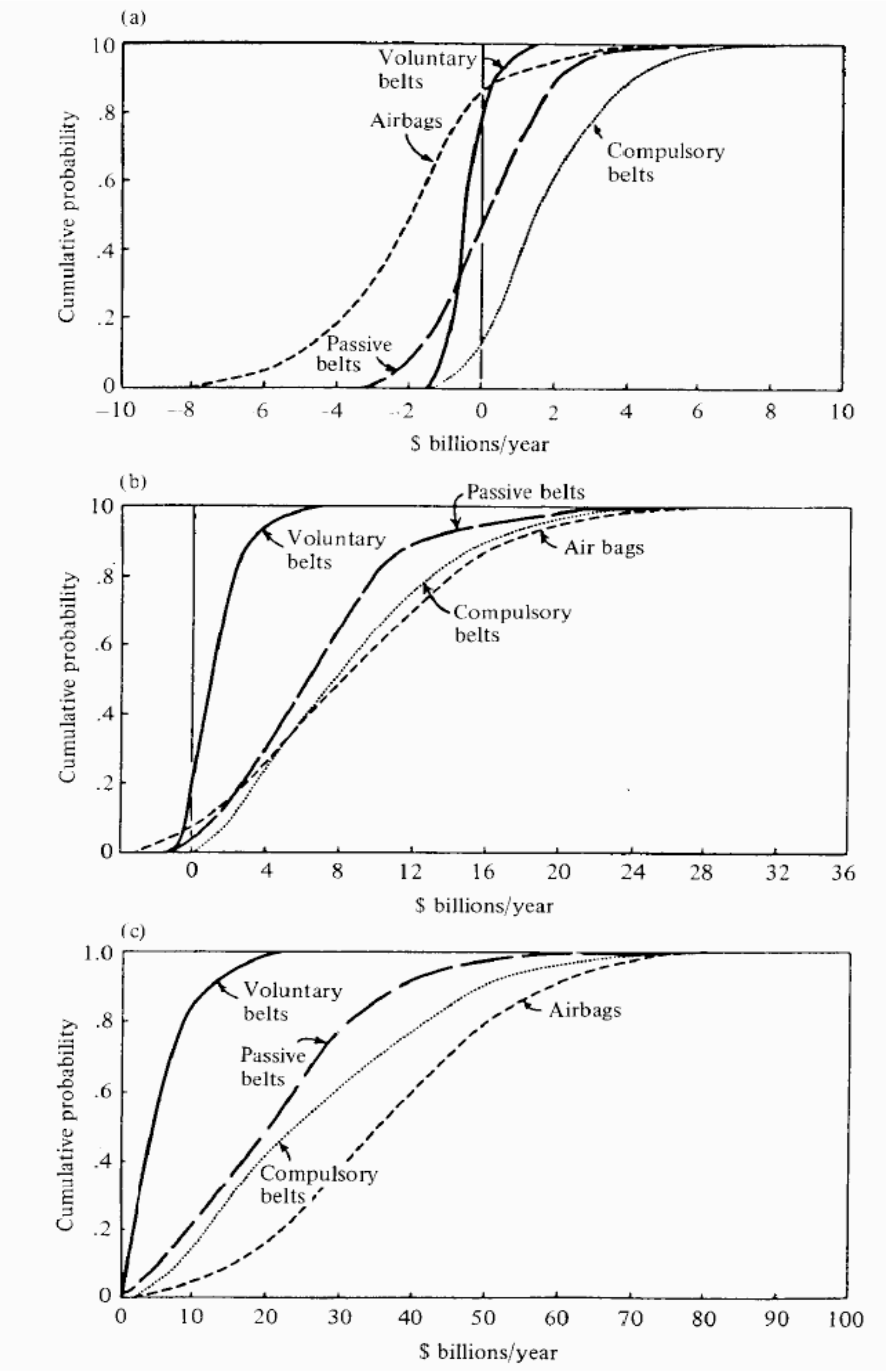
EXHIBIT 8 PDF EXAMPLE

Figure 4-4.



Source: Resources for the Future (2006) *Not a Sure Thing: Making Regulatory Choices Under Uncertainty* , http://rff.org/rff/Documents/RFF-Rpt-RegulatoryChoices.pdf

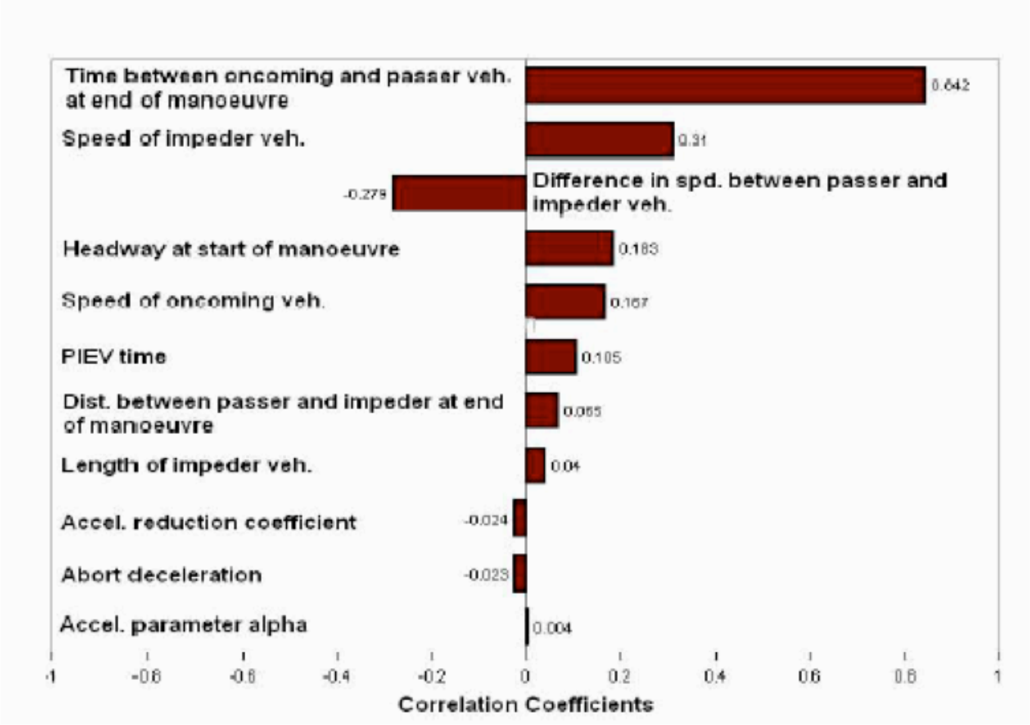
EXHIBIT 9 CDF EXAMPLE



Source: Resources for the Future (2006) *Not a Sure Thing: Making Regulatory Choices Under Uncertainty* , http://rff.org/rff/Documents/RFF-Rpt-RegulatoryChoices.pdf

EXHIBIT 10 TORNADO DIAGRAM EXAMPLE

Source: Roozenberg and Nicholson 2003



Source: Resources for the Future (2006) *Not a Sure Thing: Making Regulatory Choices Under Uncertainty* , http://rff.org/rff/Documents/RFF-Rpt-RegulatoryChoices.pdf

REFERENCES

Abt Associates, Incorporated. 2006. *Sensitivity/Uncertainty Analysis of the Relationship Between Premature Mortality and Long-Term Exposure to $PM_{2.5}$.* Prepared for: National Center for Environmental Economics, U.S. Environmental Protection Agency, Washington, DC. April.

Cooke, R.M. *Experts in Uncertainty: Opinion and Subjective Probability in Science*; Oxford University Press: New York, NY, **1991**.

Emission Projections for the Clean Air Act Second Section 812 Prospective Analysis: Draft Report; Office of Air and Radiation, US Environmental Protection Agency, Washington, DC. Prepared by E.H. Pechan & Associates, Inc. and Industrial Economics, Inc. June 2006. Available at www.epa.gov/oar/sect812

Hanna S.R., Lu Z., Frey H.C., Wheeler N., Vukovich J., Arunachalam S., Fernau M., and Hansen D.A. (2000) "Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG Domain". *Atmospheric Environment* 34.

Industrial Economics, Incorporated (IEc). 2006. *Expanded Expert Judgment Assessment of the Concentration-Response Relationship Between $PM_{2.5}$ Exposure and Mortality*. Prepared for: Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC. September.

Industrial Economics, Incorporated (IEc). 2007. *Expanded Expert Judgment Assessment of the Concentration-Response Relationship Between $PM_{2.5}$ Exposure and Mortality*. Prepared for: Office of Policy Analysis and Review, U.S. Environmental Protection Agency, Washington, DC. February.

Kochi, I., B. Hubbell, and R. Kramer. 2006. An Empirical Bayes Approach to Combining Estimates of the Value of Statistical Life for Environmental Policy Analysis. *Environmental and Resource Economics*. 34: 385-406.

National Research Council. 2002. *Estimating the public health benefits of proposed air pollution regulations.* The National Academies Press: Washington, D.C.

Memorandum from James DeMocker, EPA/OAR/OPAR, to the 812 Prospective II Files, "Scenario Specifications," August 3, 2005.

Memorandum from Jim Neumann, Industrial Economics, to Jim DeMocker, EPA/OAR/OPAR, "Alternative Methodology for Estimating Emissions for the Year 2001 Without-CAAA Scenario for the Electric Generating Unit Sector," July 24, 2006.

Memorandum from James Neumann and Sarah Brennan, Industrial Economics, to Nona Smoke and James DeMocker, EPA/OAR/OPAR, "Responding to SAB Council Comments on the May 2003 Draft Analytical Plan for the Section 812 Second

Prospective - Options for Adjusting WTP Estimates to Reflect Changes in Real Income," September 30, 2004.

Memorandum from Neil Wheeler, Sonoma Technologies, to Jim Neumann, Industrial Economics, Incorporated, "Section 812 Second Prospective Plan for Uncertainty Analysis – Emissions and Air Quality Analyses". September 29, 2006

Mrozek, J.R., and L.O. Taylor. 2002. "What Determines the Value of Life? A Meta-Analysis." *Journal of Policy Analysis and Management* 21(2):253-270.

Resources for the Future (RFF). 2006. *Not a Sure Thing: Making Regulatory Choices Under Uncertainty*, Available at http://rff.org/rff/Documents/RFF-Rpt-RegulatoryChoices.pdf

U.S. Environmental Protection Agency (EPA). 1999. *The Benefits and Costs of the Clean Air Act, 1990–2010.* Prepared for U.S. Congress by U.S. EPA, Office of Air and Radiation/Office of Policy Analysis and Review, Washington, DC, November; EPA report no. EPA-410-R-99-001.

U.S. Environmental Protection Agency, 2004. *Final Regulatory Analysis: Control of Emissions from Nonroad Diesel Engines*. EPA420-R-04-007. Prepared by Office of Air and Radiation. Available at http://www.epa.gov/nonroad diesel/2004fr/420r04007.pdf

U.S. Environmental Protection Agency, 2004. *Final Regulatory Analysis: Control of Emissions from Nonroad Diesel Engines*. EPA420-R-04-007. Prepared by Office of Air and Radiation. Available at http://www.epa.gov/nonroad diesel/2004fr/420r04007.pdf

U.S. Environmental Protection Agency, 2000. *Guidelines for Preparing Economic Analyses*. EPA 240-R-00-003. Prepared by the National Center for Environmental Economics. September

U.S. Environmental Protection Agency, 2006. *Regulatory Impact Analysis for the 2006 Revision to Particulate Matter National Ambient Air Quality Standards*. Prepared by Office of Air and Radiation. Available at: http://www.epa.gov/ttn/ecas/ria.html

U.S. Environmental Protection Agency Science Advisory Board. 2004. *Review of the Draft Analytical Plan for EPA's Second Prospective Analysis – Benefits and Costs of the Clean Air Act, 1990-2020: An Advisory by the Advisory Council for Clean Air Compliance Analysis*. EPA-SAB-COUNCIL-ADV-04-004. December.

Viscusi, V.K., and J.E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World." *Journal of Risk and Uncertainty* 27(1):5-76.

ATTACHMENT 1

MEMORANDUM FROM SONOMA TECHNOLOGY, INC.

" SECTION 812 SECOND PROSPECTIVE PLAN FOR UNCERTAINTY ANALYSIS – EMISSIONS AND AIR QUALITY ANALYSES"

MEMORANDUM

Sonoma Technology, Inc.
1360 Redwood Way, Suite C
Petaluma, CA 94954-1169
707/665-9900
FAX 707/665-9800
www.sonomatech.com

September 29, 2006

To: James Neumann, Industrial Economics, Inc. STI-906032.02-3044

From: Neil Wheeler

Re: Section 812 Second Prospective Plan for Uncertainty Analysis – Emissions and Air Quality Analyses

Under Section 812 of the Clean Air Act Amendments (CAAA), the U.S. Environmental Protection Agency (EPA) is requested to periodically conduct and submit to Congress a report on economic benefits and costs of all provisions of the Act and its Amendments. The EPA delivered the first of these reports, a retrospective analysis covering provisions of the original Clean Air Act during the period 1970-1990, in 1997, and the second report, a prospective analysis covering provisions of the Amendments during the period 1990-2010, in 1999.

The EPA is currently working on the third report to be developed under Section 812. This "Second Prospective" report will estimate benefits and costs for provisions of the Amendments as they are expected to be implemented during the period 1990-2020. An analytical plan for the Second Prospective was completed in May 2003, and comments from the Science Advisory Board (SAB) Council reviewing the plan were received by the EPA in May 2004.[1]

This memorandum addresses one of the key issues raised in the SAB Council comments: clarifying the strategy for uncertainty analyses for the emissions estimation and air quality modeling steps of the analytical chain. In September 2004, Sonoma Technology, Inc. (STI) completed a literature review that summarized much of the existing uncertainty literature and assessed the possible application of existing approaches to the Second Prospective.[2] The result of STI's assessment was a recommendation to pursue a four-phase process for estimating the uncertainties in the integrated air quality modeling system (IAQMS), which includes the emissions, meteorological, and air quality models:

---

[1] The May 2003 Analytical Blueprint for the second prospective study, along with a complete copy of the first prospective Report to Congress, can be found on the EPA's web site at: http://www.epa.gov/oar/sect812/. The SAB Council's comments on the May 2003 Analytical Blueprint can be found at: http://www.epa.gov/sab/pdf/council_adv_04004.pdf and http://www.epa.gov/sab/pdf/council_adv_04_001.pdf.

[2] See September 30, 2004 memorandum to Nona Smoke and James DeMocker, EPA/OAR/OPAR, from Neil Wheeler and Kiren Baum, Sonoma Technology, Inc., "Response to SAB Council Comments on the May 2003 Draft Analytical Plan for the Section 812 Second Prospective – Options for Uncertainty Analysis for Emissions and Air Quality Analyses".

1. **Identify and quantify internal and external uncertainties in the IAQMS** – Summarize known internal (model formulation, parameterizations, numerical solvers, etc.) and external (model input) uncertainties based on the current literature. Prioritize uncertainties based on the expected IAQMS sensitivity to those uncertainties.
2. **Assess importance of uncertainties in IAQMS predictions** – Determine the IAQMS's sensitivity to the identified uncertainties.
3. **Assess uncertainties in IAQMS response to emission changes** – Estimate the effects of propagating uncertainties through an IAQMS to quantify uncertainties in projected changes in air quality.
4. **Final Synthesis** – Update prioritization of uncertainties based on assessment components and identify key uncertainties. Summarize the key uncertainties associated with air quality modeling by potential sources, indicating direction of potential bias for the net benefits estimate and likely significance relative to key uncertainties in the net benefit estimate.

The proposed uncertainty assessment is national in scope and would focus on particles less than 2.5 micrometers in diameter ($PM_{2.5}$) and ozone. For $PM_{2.5}$, uncertainties in total $PM_{2.5}$ mass, not individual species, would be assessed because species-specific toxicity is not yet possible to estimate and will not be addressed in the Second Prospective Analysis.

This memorandum describes uncertainties in the IAQMS and a recommended plan for estimating those uncertainties as part of the Second Prospective analysis. Each phase of this plan is described in the following sections, followed with a recommended schedule and estimates of the resources required to complete the assessment. The approach for estimating uncertainties from the IAQMS should be as rigorous as possible within the practical constraints of time and resources. The recommended approach for assessing uncertainties offers the ability to provide better estimates as each phase of the assessment is performed.

UNCERTAINTIES IN THE IAQMS

Sources of Uncertainty

Uncertainty in estimated values for future air quality arises from at least three sources: (1) inherent or stochastic variability in the observations; (2) errors in model physics and chemistry assumptions; and (3) errors due to uncertainties in model input variables. For prospective analyses, we need to focus on uncertainty in the context of model response to future-year emissions. For example, an air quality model (AQM) may be very sensitive to a particular input without affecting its response to emission changes. Alternatively, an AQM may show little sensitivity to an input under current conditions (e.g., boundary conditions) but become increasingly sensitive to that input in future years as anthropogenic emissions are reduced.

Measurement Uncertainty

While measurement uncertainty is less important when using relative reduction factors (RRFs) and linear cost-response functions, it can affect the ability to evaluate model performance and gain confidence that a model is getting the right answer for the right reason. For gases, instruments can be calibrated using gases of known concentrations, and the uncertainty in the measurement is reasonably well-known. However for PM, this is not the case. Uncertainties in PM mass and speciation can be significant, which limits our ability to critically evaluate model performance and reduce uncertainty in model simulations.

Hogrefe et al. (2000) developed an approach to gain insight into the distribution of future air quality predictions attributable to variability in currently observed air quality at a given location. The procedure is to fit a theoretical statistical distribution to the tail of a set of daily observations at a monitoring site (e.g., over a three-year period) and compute a design value consistent with the form of the National Ambient Air Quality Standards (NAAQS). The next step is to perform a bootstrapping operation several hundred times to obtain different sets of air quality data. For each instance, a design value is determined from the resulting data. The result is a distribution of current design values, which can be translated into a distribution of future air quality estimates using the RRF approach recommended in EPA guidance. While work so far has focused on the 1-hr and 8-hr NAAQS for ozone, it may be possible to apply the methodology to PM-related applications.

Model Uncertainty

Emissions, meteorological, and air quality models are mathematical representations of the physical world, and as such, have inherent uncertainties associated with their formulation, assumptions, and implementation. Some of the uncertainties are due to the limitations of our scientific knowledge. Other uncertainties are a result of simplifications or approximations needed to make the model practical. At the present time, we do not see a way to completely quantify uncertainty due to inherent limitations in a model. However, methods and a body of research are available to help us understand the importance of uncertainty in individual model components. We can also reduce uncertainty by using models whose scientific basis is fully and satisfactorily explained in its accompanying documentation.

In some cases, it is necessary to use a simplified “engineering” or “reduced-form” version of a model. Uncertainty inherent in such results may be reduced if it has been shown that the engineering and more complete versions of a model produce similar results under the conditions that are of greatest interest for a particular application.

Input Uncertainty

The best-formulated and least uncertain models are only as good as their inputs. Model input uncertainty has been explored extensively in past decades and has driven research to improve these model inputs. In some cases, these inputs are based on measurements, which may be available only at limited temporal or spatial resolutions. In other cases, the input for one model may be the output of another model (i.e., the use of a mobile source emissions model to provide input to an AQM).

Methods for Quantifying Uncertainty in Model Inputs and Options

The first step in uncertainty analysis is to estimate the uncertainties in model input variables and options. Model options may include alternative techniques for solving model equations or alternative physical or chemical submodels. The two primary methods available for the Second Prospective Analysis are literature reviews and expert elicitation. For longer-term efforts in assessing uncertainty, these methods could be supplemented with specific applications of methods already discussed in the literature and in new research.

Literature Reviews

Past and current literature can provide estimates of uncertainties in model inputs based on measurement and sensitivity studies. Because models and measurements are constantly evolving, care must be taken to ensure that estimates of uncertainty in the literature are still valid.

Emission Inventories

Table 1 provides an overview of methods reviewed for the Emission Inventory Improvement Program (EIIP) in its final report on evaluating the uncertainty of emission estimates (Emission Inventory Improvement Program, 1996). While many of the studies cited are now out-of-date, the report provides a good summary of the methods available for quantifying uncertainty.

More recent research has been performed to develop and demonstrate improved methods for quantifying uncertainty in emission inventories. A complete review of research on quantifying uncertainty in emission estimates was not possible within the scope of this work assignment. However, the following discussion provides some examples of the methods being used and the results obtained.

In the area of mobile source emissions, Kini and Frey (1997) developed quantitative estimates of uncertainty associated with Mobile5b emission factor model estimates of light-duty gasoline-vehicle base emissions and speed-corrected emissions and found that the uncertainty in average emissions is often ±20% or more. Pollack et al. (1999) performed a similar study on California’s EMFAC7G highway vehicle emission factor model. Frey et al. (1999) revisited the earlier analysis of Mobile5b emission factor estimates to include uncertainties associated with temperature corrections. Rhodes and Frey (1997) quantified variability and uncertainty in AP-42 emission factors using a bootstrap simulation method.

Table 1. Overview of methods for evaluating the uncertainty of emission estimates.

Method	Description	References
Qualitative Discussion	Sources of uncertainty are listed and discussed. General direction of bias and relative magnitude of imprecision are given if known.	Steiner et al., 1994
Subjective Data Quality Ratings	Subjective rankings based on professional judgment are assigned to each emission factor or parameter.	U.S. EPA, 1995 Saeger, 1994
Data Attribute Rating System (DARS)	Numerical values representing relative uncertainty are assigned through objective methods.	Beck et al., 1994
Expert Estimation Method	Emission distribution parameters (i.e., mean, standard deviation, and distribution type) are estimated by experts. Simple analytical and graphical techniques can then be used to estimate confidence limits from the assumed distributional data. In the Delphi method, expert judgment is used to estimate uncertainty directly.	Linstene and Turoff, 1975 SCAQMD, 1982 Horie, 1988 Horie and Shorpe, 1989
Propagation of Errors Method Direct Simulation Method	Emission parameter means and standard deviations are estimated using expert judgment, measurements, or other methods. Standard statistical techniques of error propagation typically based on Taylor's series expansions are then used to estimate the composite uncertainty.	Mangat et al., 1984 Benkovitz, 1985 Benkovitz and Oden, 1989 Balentine et al., 1994 Environment Canada, 1994
Direct Simulation Method	Monte Carlo, Latin hypercube, bootstrap (resampling), and other numerical methods are used to estimate directly the central value and confidence intervals of individual emission estimates. In the Monte Carlo method, expert judgment is used to estimate the values of the distribution parameters prior to performance of the Monte Carlo simulation. Other methods require no such assumptions.	Freeman et al., 1986 Iman and Helton, 1988 Oden and Benkovitz, 1990 Efron and Tibshirani, 1991 Environment Canada, 1994 Gatz and Smith, 1995a Gatz and Smith, 1995b
Direct or Indirect Measurement (Validation) Method	Direct or indirect field measurements of emissions are used to compute emissions and emission uncertainty directly. Methods include direct measurement such as stack sampling and indirect measurement such as tracer studies. These methods also provide data for validating emission estimates and emission models.	Pierson et al., 1990 Spellicy et al., 1992 Fujita et al., 1992 Peer et al., 1992 Mitchell et al., 1995 Claiborn et al., 1995
Receptor Modeling (Source Apportionment) Method	Receptor modeling is an independent means to estimate the relative contribution of specific source types to observed air quality measurements. The method works best for nonreactive pollutants for which unique emission composition "fingerprints" exist for all significant source categories. The method provides a measure of the relative contribution of each source type but not absolute emission estimates.	Watson et al., 1984 Lowenthal et al., 1992 Chow et al., 1992 Scheff et al., 1995
Inverse Air Quality Modeling Method	Air quality simulation models are used in an inverse, iterative approach to estimate the emissions that would be required to produce the observed concentrations fields.	Hartley and Prinn, 1993 Chang et al., 1993 Chang et al., 1995 Mulholland and Seinfeld, 1995

Frey and Bammi (2002) estimated uncertainty in the emission factors for lawn and garden (L&G) equipment. For 2-stroke L&G engines, the 95% confidence intervals for the mean emission factors for total hydrocarbon (THC) and NO_x emissions were -30% to +41% and -45% to +75%, respectively. For 4-stroke L&G engines, the confidence intervals were -33% to +46% for THC and -27% to +35% for NO_x.

Frey and Li (2003) applied quantitative methods for characterizing variability and uncertainty to case studies of emission factors from AP-42 for stationary natural gas-fueled internal combustion engines. The approximate range of uncertainty in mean emission factors varies from as little as ±10% to as much as -60% to +80%, depending on the pollutant, control technology, and nature of the available data.

Frey and Zheng (2002a) developed a probabilistic methodology for quantifying variability and uncertainty in highway vehicle emission factors based on data used in MOBILE5b. Empirical distributions of emissions measurement data were used to characterize variability, while the bootstrap simulation method was used to characterize uncertainty. Inter-vehicle variability in emissions was found to span 2 or 3 orders of magnitude. The uncertainty in fleet average emission factor range from +/- 10% to as much as -90% to +280%.

Frey and Zheng, (2002b) quantified the variability and uncertainty in emission factors and activity factors for power plant NO_x emissions using the Monte Carlo and bootstrap simulation. The uncertainties were then propagated through an emission inventory to produce a probabilistic power plant NO_x emission inventory for North Carolina.

Frey and Bammi (2003) estimated variability and uncertainty in NO_x and total hydrocarbon emission factors for construction, farm, and industrial (non-road) engines. Bootstrap simulations were used to develop confidence intervals for the mean. The 95% confidence intervals for the mean emission factors were as small as –10 to +11% and as large as –48 to +49%, with an average range of –26 to +27%.

Abdel-Aziz and Frey (2003a) used univariate stochastic time series models, and ordinary least-squares regression models were employed to quantify hourly uncertainty in capacity emission factors and heat rate, respectively. The models were used to develop an hourly probabilistic power plant NO_x emission inventory for a four-day period. Abdel-Aziz and Frey (2003b) used multivariate time series models (time series approach) to account for the dependence between emissions from correlated units.

Zhao and Frey (2004) developed probabilistic toxic emission inventories for 1,3-butadiene, mercury, arsenic, benzene, formaldehyde, and lead for Jacksonville, Florida. Parametric bootstrap simulation and empirical bootstrap simulation were used to quantify the uncertainty in urban air toxic emission factors. The emission inventory 95% uncertainty ranges were as small as -25% to +42% for chromium to as large as -75% to +224% for arsenic with correlated surrogates. Uncertainty was dominated by only a few source categories.

Chi et al. (2004) used bootstrap sampling, expert elicitation, and Monte Carlo (MC) simulations to characterize uncertainty of nonroad emissions for Georgia from the EPA NONROAD model. Tools used were a bootstrap resampling technique and a parametric bootstrap analysis method in Zheng and Frey's Analysis of Uncertainty and Variability Tool (AuvTool). Overall uncertainty ranged from -23 to +33%; however, fuel consumption, growth factors, equipment age distributions, PM and HC speciation profiles, temporal activity adjustments, fuel sulfur effects, and evaporative emissions were not accounted for in the analysis.

Meteorological and Air Quality Models

Derwent and Hov (1988) made estimates of uncertainty in photochemical model inputs based on "best judgments" for an application of sensitivity and analysis techniques. They estimated uncertainties to be ±50% for concentrations aloft; ±30% for emissions and deposition velocities, and hydroxyl radical sinks; ±20% for boundary layer depth; and ±10% wind speed. In preparation for an MC uncertainty analysis of Ozone Transport Assessment Group (OTAG) (1997) modeling, Frey (1998) developed estimates of uncertainty in the AQM inputs based on expert elicitation. Frey reported the uncertainty range, which includes 95% of the data, to be a factor of 5 for initial VOC and NO_x concentrations; a factor of 3 for initial ozone concentrations, boundary conditions of VOC and NO_x, and vertical diffusivity above 1000 m and at times other than 8:00 a.m. to 6:00 p.m.; and a factor of 2 for photolysis rates, cloud liquid water content, rainfall amounts, and emissions except major point sources. The range of uncertainty for chemical reactions in the Carbon Bond IV chemical mechanism varied, by reaction, from a factor of 1.01 to 3.02. The least uncertain model inputs were major point source emissions (±50%), horizontal boundary condition for ozone (±50%), concentrations aloft (±50%), wind direction (±40 degrees), cloud cover (±30%), vertical diffusivity below 1000 m from 8:00 a.m. to 6:00 p.m. (±30%), relative humidity (±30%), and ambient temperature (±3 °C).

While formal estimates of uncertainty are not typically made of the meteorological model outputs used as inputs to AQMs, some information about uncertainty can be gained from the performance evaluations of these models. Often statistical comparisons of the model predictions to observations are provided. While these statistics provide a first-order estimate of the uncertainty, it must be kept in mind that model estimates and observations may not be spatially and temporally commensurate. Model predictions represent grid-cell volume averages of the predicted parameters at a particular time while observations are most often for a point location and may be averaged over various periods of time. Therefore, model performance-based estimates of uncertainty are likely to be larger than the actual uncertainty.

Olerud et al. (2000) performed meteorological modeling with MM5 for all of 1996 on a grid covering the entire continental United States at 36-km resolution. The results of this modeling have been used by EPA and regional planning organizations (RPOs) in subsequent air quality modeling studies with REMSAD, UAM-V, CAMx, and Community Multiscale Air Quality (CMAQ) model. The root mean square errors for the entire domain were reported by season and ranged from 1.15 to 1.47 m/s for wind speed, 35.2 to 38.5 degrees for wind direction, 2.3°C to 4.2°C for temperature, and 0.8 to 1.7 g/Kg for humidity. Doty et al. (2002) reported on meteorological modeling with the RAMS model for the Southern Appalachian Mountains Initiative (SAMI). They found that for their 12-km domain, over all days modeled, the root mean square error for wind speed was 2.18 m/s, the gross error for wind direction was 39 degrees, the gross error for temperature was 1.9°C with a bias of –0.8°C, and the gross error for humidity was 0.8 g/kg with a bias of –0.1 g/kg.

Moore and Londergan (2001) used a modification of the basic Monte Carlo to determine uncertainty. The computationally intensive aspects of the full methodology are replaced by a highly restricted sampling approach that exploits the spatial persistence found in predicted concentration fields. The approach was tested in an application of UAM-IV to assess the uncertainty in the differences in predicted maximum ozone concentration between the base-case and control scenarios. Uncertainty in model inputs and parameters were simulated using stochastic models driven by Latin Hypercube Sampling (LHS). They propagated uncertainty in 168 model inputs for emissions, chemistry, meteorology, and boundary conditions.

A probabilistic hourly NO_x emission inventory was developed for 32 units of nine coal-fired power plants in the Charlotte, North Carolina, region for 1995 (Abdel-Aziz and Frey, 2003a,b). The uncertainty was then propagated through the MAQSIP model to estimate the uncertainty in maximum 1-hr and 8-hr concentrations for the Charlotte, North Carolina, modeling domain using an MC simulation (Abdel-Aziz and Frey, 2004). Statistical dependencies between power plant units (inter-unit variability), as well as temporal autocorrelation for each individual unit (intra-unit variability) were accounted for. A total of 50 simulations were performed in order to represent the ranges of uncertainty in hourly emissions and predicted ozone levels. The range of uncertainty in predicted peak 1-hr ozone concentrations solely attributable to utility NO_x emissions was a large as 25 ppb. Uncertainties in peak ozone concentrations at specific locations could be pinpointed to emissions from a specific power plant. Exceedances of the 8-hr standard were more widespread and not attributable to any one plant.

Expert Elicitation

Quantifying the uncertainties in model input variables may be difficult because there is little specific information on this subject in the literature for the complete spectrum of inputs (e.g., initial and boundary conditions, emissions components, meteorological variables, model parameterization constants, photolysis rates, and chemical rate constants). When quantifying the uncertainties is difficult, Morgan and Henrion (1990) suggest that it is appropriate to carry out an expert elicitation where "experts" are asked to give estimates of uncertainties based on their experience. To combine information from a number of different experts, each expert can be assigned a subjective weight indicating the relative extent of the individual's expertise with respect to the other experts participating in the elicitation (National Council on Radiation Protection and Measurements [NCRP], 1996). In many instances, each expert may be given equal weight, but in those areas where the degree of expertise differs markedly, unequal weights may be assigned to each expert.

Hanna et al. (1998) estimated uncertainties in model inputs by taking the median of the uncertainty values (expressed as a plus and minus percentile that would include 95% of the variability) suggested by ten modelers (experts) who responded to questionnaires. That is, each expert was given equal weight. In that study, no attempt was made to carry out a comprehensive survey of modelers (experts) or to encourage discussions among modelers.

Hanna et al. (2000) improved upon this process by attempting to reach about 100 experts via a web page where the experts could enter their estimates of input uncertainties. The 100 experts included 10 or 20 from each major category of input data (e.g., emissions, boundary and initial conditions, chemical rate constants, and meteorology). However, only about 20 experts responded to the request. It was found that better information could be obtained by meeting with groups of experts at several different laboratories. One reason for the difficulty is that many photochemical modeling experts have not thought much about uncertainties in input parameters and, therefore, the estimates are largely based on intuition and compromise. Hanna et al. suggested that future expert elicitations should be more thorough, including workshops where experts come together to discuss the uncertainties. Experts should also assign weights to themselves based on their degree of expertise. The problem with the approach is that it is time-consuming and resource-intensive (two or three weeks of effort over a time period of about six months plus travel costs for two or three meetings for each of about 20 experts).

Methods for Assessing the Effects of Uncertainty

Sensitivity analysis is the most widely used method for assessing the effects of uncertainty on future-year air quality outcomes. Process analysis has been used in more recent AQM applications to identify those processes in the AQM that contribute the most to predicted pollutant concentrations and, thus, may be most affected by uncertainty. These methods and their use are discussed in greater detail below.

Sensitivity Analysis

The response of AQM predictions to changes of input parameters or model options can provide valuable information about uncertainties in model predictions. Such information can be obtained by sensitivity analysis, the systematic calculation of sensitivity coefficients, to quantitatively measure these dependencies. Basic sensitivity analysis may involve perturbing input parameters or model options one at a time or in combinations.

Beck et al. (1997) provide an overview of evaluations and uncertainties of environmental models, with emphasis on water quality models. They stress the need to specify a hypothesis or question to be answered by the model, and describe three alternatives to basic sensitivity analysis: (1) brute-force MC uncertainty analysis; (2) response surface evaluation; and (3) first-order error analysis, which is sometimes called sensitivity or "small perturbation" analysis. Each technique is discussed below.

Basic Sensitivity Analysis

Because of its ease of use and interpretation, there exist many examples of basic sensitivity analysis applied to AQMs. For example, Seigneur et al. (1981) estimated the sensitivities of an urban model to variations in input data. Winner et al. (1995) and Dabdub et al. (1999) showed that ozone predictions are especially sensitive to the inflow boundary conditions in Los Angeles and the San Joaquin Valley, respectively. Hass et al. (1997) carried out a sensitivity study of four European long-range transport and dispersion models, finding factors of 2 to 3 differences in the sensitivities of the different models to variations in emissions. Our review of these sensitivity studies suggests that the results are applicable only to a narrow range of conditions associated with the specific scenario. Since photochemical processes are often non-linear, the magnitude and even the sign of the sensitivity coefficients may vary as the scenario varies.

While meteorological parameters are undoubtedly important in photochemical grid models, it is not easy to decide how to account for variations in meteorology, especially wind speed and direction. The problem is that it is necessary for the wind field to always satisfy mass-continuity, so that it is not correct to simply randomly vary the winds in each grid square of the model. Photochemical grid models make use of meteorological preprocessors, which may adjust the wind fields so they are mass-consistent. Hanna et al. (1998) avoided this problem by assuming that the perturbations in wind speed and direction applied uniformly across all grid squares. Schere and Coates (1992) suggested a more elegant (and time-consuming) method of accounting for uncertainties or variations in winds. Bergin et al. (1999) attacked the problem by generating a small number of alternate wind fields based on systematically “withdrawing” data from the meteorological preprocessor. This method is a useful first estimate but will underestimate the total uncertainty because of the limited number of runs and the failure to account for the full range of wind uncertainty.

Meteorologists have accounted for variability in weather forecasts by applying the “ensemble” method in which several forecast models (i.e., an ensemble) are run for the same scenario, and the best-guess forecast is assumed to be given by the mean of the several forecasts. These methods have been applied to air quality models by Straume et al. (1998), who showed that the ensemble method produced improved forecasts of tracer concentrations for the long-range ETEX tracer experiment in Europe. It is implied that the uncertainty would be given by the variability of the forecasts. These methods have also been extended to regulatory air quality modeling by using and evaluating alternative AQMs. For example, Ozone Transport Assessment Group (1997) modeling used of multiple meteorological models (SAIMM and RAMS) and multiple AQMs (UAM-V and CAMx) for some episodes. However, it is clear that the full range of possible input conditions can not be covered by these ensemble methods.

The EPA guidance documents on attainment demonstrations (U.S. Environmental Protection Agency, 1999, 2001) identify three sensitivity tests that may be useful for assessing uncertainty in AQM predictions. The first of these, which has been proposed by Reynolds, et al., (1996), is to prepare “alternative base-case” emission estimates, reflecting reasonable alternative assumptions about current emissions which lead to comparable or better model performance. A second test is to assume alternative (reasonable) growth assumptions. This could reflect using differing growth rates or placement of new sources in different, equally probable, locations. Combinations of these first two tests are also possible. A third test involves simulating a future-year case with an alternative grid resolution or with different (reasonable) meteorological assumptions. For example, due to resource constraints, it might be necessary to perform modeling using a grid with 36-km grid cells (horizontal dimension). Differences in projected air quality obtained with a grid having 12-km or 4-km cells could then be evaluated.

The EPA guidance documents on modeling for attainment demonstrations were influenced by earlier guidance developed at the California Air Resources Board (CARB) which specifically addressed uncertainty (DaMassa, 1992). CARB applied this guidance in a series of uncertainty analyses to support the development of California's State Implementation Plans (SIPs). This program included analyses of uncertainty associated with future-year boundary conditions (Wagner and Wheeler, 1988), meteorology (Wagner and Wheeler, 1989; Wheeler, 1992), emission inventory bias (Wagner et al., 1992), horizontal advection solvers (Odman et al., 1996), chemical mechanisms (Whitten and Killus, 1998), and photolysis rates (Vuilleumier et al., 2000).

Monte Carlo Uncertainty Analysis

MC methods are the most widely used means for uncertainty analysis. These methods involve random sampling from the distribution of inputs and successive model runs until a statistically significant distribution of outputs is obtained. There has been a rapid growth in the use of MC uncertainty analysis with photochemical AQMs in recent years. This "brute-force" method is computer-intensive because it requires 50 to 100 or more model runs for each base-year and future emission scenario. However, because of the exponential growth of computer speed and storage, it is now possible to carry out Monte Carlo runs with a complex photochemical grid model applied to large domain.

One of the first applications of MC uncertainty analysis to photochemistry was the study of relationships between stratospheric ozone and chlorine reported by Solarski et al. (1978). Alcamo and Bartnicki (1987) used MC methods to study the uncertainties in sulfur deposition predicted by the EMEF-W model in Europe. They found that it is more important to specify the width (i.e., the standard deviation) rather than the shape of the probability density function of the input variables. Gao et al. (1996) applied MC uncertainty analysis to the chemical rate parameters. Deuel et al. (1998) studied the uncertainties of the UAM-V model using MC methods; however, the uncertainty ranges that they assumed for the input variables (vertical resolution, vertical diffusivity, plume-in-grid method, land-use, chemical reaction rates, and emissions) were a third or less than those recommended by the experts in the studies by Hanna et al. (1998; 2000). Bergin et al. (1999) applied MC methods with LHS to a Lagrangian photochemical AQM (i.e., not a grid model) in Southern California. They accounted for meteorological variability by using several solutions of a mass-consistent wind model, run with random data-withholding assumptions. This method has been widely used in other environmental fields (e.g., water pollution modeling), as described in the reviews by International Atomic Energy Agency (IAEA) (1989), NCRP (1996), and Beck et al. (1997).

Frey (1992) discusses the decision process followed in applications of MC uncertainty analysis, stressing the importance of good estimates of input data uncertainties. Conover (1971) provides guidance concerning the computation of statistical tolerance limits from a simple random sample. Bergin et al. (1999) discuss the use of LHS, which they believe provides a better coverage of the data distribution than Simple Random Sampling (SRS). However, the advantage of LHS comes with a price—only with SRS can the confidence in the results be interpreted through statistical tolerance limits.

From a practical standpoint, Hanna et al. (2000) demonstrated that MC methods could be applied to larger photochemical modeling studies (i.e., OTAG) by performing 100 simulations each for a base-case and three emission reduction scenarios. Recent work by Houyoux et al. (2003) simplified the use of AQMs for assessing emission inventory uncertainties by generating multiple realizations of model-ready emissions with the Sparse Matrix Operator Kernel Emissions (SMOKE) processing system (Coats and Houyoux, 1996) by modifying SMOKE to accept parametric and empirical probability distributions to describe the uncertainty about them. This approach allows emissions modelers to assign uncertainty information about an existing inventory without having to change the actual inventory files. The same inventories can be used for both deterministic (i.e., without uncertainty) modeling and stochastic modeling (i.e., with uncertainty), and the type of modeling that is performed depends only on the presence of the additional inventory uncertainty file.

Response Surface Analysis

At the other extreme from simple one-at-a-time sensitivity studies, the response surface method (Tatang et al., 1997) attempts to fit orthogonal polynomials to the input conditions and the predictions of numerical geophysical models. For this approach, it is necessary to run the models a sufficient number of times to have enough data to develop the response surfaces. It is claimed that 25 to 60 times fewer runs are needed than for a MC SRS exercise. However, in a Response Surface Model (RSM) pilot study, Hubbell (2003) reported that 144 REMSAD runs were required to characterize a second order polynomial surface to develop a RSM for $PM_{2.5}$.

Nevertheless, the response surface simply amounts to a model of a model and therefore, is susceptible to problems associated with scenarios outside of the range of parameters used to generate the data for deriving the model.

First-Order Sensitivity Analysis

Sensitivity analysis has not been used as extensively as desired because of implementation complexity and computational limitations. As a result, the simple "brute-force" method has been used most frequently to determine model sensitivities, especially in multidimensional chemistry transport models. By this method, a separate simulation is required to calculate the effects of each parameter or emission rate in the model. However, this approach rapidly becomes impractical when a large number of sensitivity coefficients need to be computed.

A number of other approaches have been developed to calculate sensitivity coefficients. One method of reducing this effort is determining the equations governing the sensitivity coefficients and solving them directly. In this method, the sensitivity equations are derived from the model equations and solved simultaneously with the model equations. This method proved to be unstable and inefficient when applied to stiff equations found in many air quality problems (Dunker, 1984). Other techniques rely on Green's function (Rabitz et al., 1983; Cho et al., 1987; Harley et al., 1997) or the adjoint method, in which the sensitivity coefficients are computed from integrals of the Green's function of sensitivity equations derived from the model equations.

The automatic differentiation of Fortran (ADIFOR) technique (Bischof et al., 1992) automatically translates large FORTRAN codes to a subprogram that includes the original functions as well as those for the desired sensitivity coefficients. This method has been used in past studies for sensitivity analysis of the advection equation as used for atmospheric modeling (Hwang et al., 1997), and initial concentrations and reactions rates in photochemical models (Carmichael et al., 1997). Because ADIFOR is designed for general-purpose sensitivity analysis, the expanded codes do not take advantage of the program structure and re-use of calculations. Also, computing some sensitivity coefficients, such as those with respect to the subdomain emissions or the boundary conditions, requires additional modifications which can be cumbersome.

Another approach for computing sensitivity coefficients is the decoupled direct method (DDM) (Dunker, 1981; 1984), in which the sensitivity equations are derived from the model equations, but solved separately from the model equations. DDM does not share the instability problem found with the direct and adjoint methods. Further, the implementation of this method is more straightforward than the coupled direct or adjoint methods since the sensitivity equations are linear, even though they are functions of concentrations. Therefore, the calculations of sensitivity coefficients are much less computationally demanding. Milford et al. (1992) and Seefeld and Stockwell (1999) also applied the DDM to study variations in chemical rate constants. A recently developed technique for sensitivity study is DDM-3D (decoupled direct method in three dimensions), which has been successfully implemented in the CIT, CAMx, and CMAQ photochemical AQMs. This approach is highly computation-efficient and capable of calculating a full set of model sensitivity in a three-dimensional domain.

Process Analysis

A technique called process analysis (PA) has been used to assess relative importance of various model assumptions as well as simulated physical and chemical phenomena contributing to an ozone concentration at a particular time and location (Jeffries, 1997, Jeffries et al., 1996, Jang et al., 1995, Lo and Jeffries, 1997). Since models used to simulate ozone and secondary particulate matter are similar, process analysis should also be useful for addressing $PM_{2.5}$ issues. The technique works by breaking down a modeled simulation into a sequence of physical and chemical processes that lead to a predicted concentration at a given location and time and by tracking the contributions of those processes. PA has been implemented in CMAQ and CAMx but not REMSAD.

While PA requires a substantial amount of expertise to be interpreted to full advantage, useful insights are possible with less detailed analyses. PA takes advantage of numerical grid models that address physical and chemical factors affecting ozone in a sequential manner. For example, a typical sequence followed in a model for each time step might be (1) advection of $PM_{2.5}$ components and precursors present at the beginning of the time step, (2) $PM_{2.5}$ and precursor emissions added during the time step, (3) vertical diffusion of the advected material and fresh emissions, (4) estimated cloud cover and its effects on photolysis rates, (5) atmospheric chemistry involving advected and diffused material with fresh emissions, and (6) deposition of certain compounds. PA examines incremental effects on changes in component and/or $PM_{2.5}$ predictions from hour to hour attributable to each of the processes described above. In this way, one gets a sense of how important each process is as a contributor to predicted air quality at a specific time and location.

Methods Proposed for Assessing the Effects of Uncertainty

Uncertainties in IAQMS will be assessed using the CMAQ model (National Exposure Research Laboratory, 1999) and EPA's Response Surface Metamodels (RSMs) for ozone (U.S. Environmental Protection Agency, 2006a) and particulate matter (U.S. Environmental Protection Agency, 2006b).

The CMAQ modeling system has been designed to approach air quality as a whole by including state-of-the-science capabilities for modeling multiple air quality issues, including tropospheric ozone, fine particles, toxics, acid deposition, and visibility degradation. In this way, the development of CMAQ involves the scientific expertise from each of these areas and combines the capabilities to enable a community modeling practice. CMAQ was also designed to have multiscale capabilities so that separate models were not needed for urban- and regional-scale air quality modeling.

The RSMs are based on an approach known as air quality metamodeling that aggregates numerous pre-specified individual air quality modeling simulations into a multi-dimensional air quality "response surface". Simply, this metamodeling technique is a "model of the model" and has been shown to reproduce the results from an individual modeling simulation with little bias or error. The RSM incorporates statistical relationships between model inputs and outputs to provide a real-time estimate of air quality changes. The RSM provides a wide breadth of model outputs, which we can utilize to assess the impact of emission uncertainties. This approach allows for the rapid assessment of air quality impacts of different combinations of emission levels.

The RSMs will be used to assess IAQMS response to emission uncertainties within the RSMs' operating range. For emission uncertainties outside the RSM's operating range and non-emission uncertainties (e.g., meteorological uncertainties), CMAQ will be used.

UNCERTAINTY ASSESSMENT PLAN

Phase 1: Identify and Quantify Internal and External Uncertainties in the IAQMS

The first phase of this process will be to summarize the uncertainties in the IAQMS and their relative importance based on information available in the literature. Uncertainties will be separated into broad categories for types of models (i.e., emissions, meteorological, and air quality) and sub-categories for measurement, input, and model uncertainties. In cases where a particular uncertainty is poorly defined or the literature is out of date, the opinions of experts will be relied upon to refine the available information. Uncertainties will be ranked based on their potential to affect the specific model with which they are associated and their overall effect on the IAQMS response to emission changes. The results of this phase will be summarized in tables of uncertainties that indicate the type of uncertainty, relevant references, significant findings from the literature and expert elicitation, comments on its importance, and its importance ranking. The principal deliverable from this phase will be a tabular summary, similar to that prepared in the First Prospective Analysis, which provides a list of key uncertainties (see **Tables 2 and 3**). For each key uncertainty, an estimate of the direction of potential bias for the net benefits estimate and likely significance relative to key uncertainties in the net benefit estimate will be provided. In addition, the tabular summary will also include the ranges of uncertainties identified in the literature.

Table 2. Uncertainties associated with emissions modeling identified in the First Prospective Analysis.

Page 1 of 2

Key Uncertainties Associated with Emissions Estimation Potential Source of Error	Direction of Potential Bias for Net Benefits Estimate	Likely Significance Relative to Key Uncertainties in Net Benefit Estimate*
PM2.5 emissions are largely based on scaling of PM10 emissions.	Overall, unable to determine based on current information, but current emission factors are likely to underestimate $PM_{2.5}$ emissions from combustion sources, implying a potential underestimation of benefits.	Potentially major. Source-specific scaling factors reflect the most careful estimation currently possible, using current emissions monitoring data. However, health benefit estimates related to changes in PM2.5 constitute a large portion of overall CAAA-related benefits.
Primary PM2.5 emissions estimates are based on unit emissions that may not accurately reflect composition and mobility of the particles. For example, the ratio of crustal to primary carbonaceous particulate material likely is high.	Underestimate. The effect of overestimating crustal emissions and underestimating carbonaceous when applied in later stages of the analysis, is to reduce the net impact of the CAAA on primary PM2.5 emissions by underestimating PM2.5 emissions reductions associated with mobile source tailpipe controls.	Potentially major. Mobile source primary carbonaceous particles are a significant contributor to public exposure to PM2.5. Overall, however, compared to secondary PM2.5 precursor emissions, changes in primary PM2.5 emissions have only a small impact on PM2.5 related benefits.
The Post-CAAA scenario includes implementation of a region-wide NOx emissions reduction strategy to control regional transport of ozone that may not reflect the NOx controls that are actually implemented in a regional ozone transport rule.	Unable to determine based on current information.	Probably minor. Overall, magnitude of estimated emissions reductions is comparable to that in expected future regional transport rule. In some areas of the 37 state region, emissions reductions are expected to be overestimated, bur in other areas, NOx inhibition of ozone leads to underestimates of ozone benefits (e.g., some eastern urban centers).
VOC emissions are dependent on evaporation, and future patterns of temperature are difficult to predict.	Unable to determine based on current information.	Probably minor. We assume future temperature patterns are well characterized by historic patterns, but an acceleration of climate change (warming) could increase emissions.
Use of average temperatures (i.e., daily minimum and maximum) in estimating motor-vehicle emissions artificially reduces variability in VOC emissions.	Unable to determine based on current information.	Probably minor. Use of averages will overestimate emissions on some days and underestimate on other days. Effect is mitigated in Post-CAAA scenarios because of more stringent evaporative controls that are in place by 2000 and 2010.

Table 2. Uncertainties associated with emissions modeling identified in the First Prospective Analysis.

Page 2 of 2

Key Uncertainties Associated with Emissions Estimation Potential Source of Error	Direction of Potential Bias for Net Benefits Estimate	Likely Significance Relative to Key Uncertainties in Net Benefit Estimate*
Economic growth factors used to project emissions are an indicator of future economic activity. They reflect uncertainty in economic forecasting as well as uncertainty in the link to emissions.	Unable to determine based on current information.	Probably minor. The same set of growth factors are used to project emissions under both the Pre-CAAA and Post-CAAA scenarios, mitigating to some extent the potential for significant errors in estimating differences in emissions.
Uncertainties in the stringency, scope, timing, and effectiveness of Post-CAAA controls included in projection scenarios.	Unable to determine based on current information.	Probably minor. Future controls could be more or less stringent, wide reaching (e.g., NOx reductions in OTAG region - see above), or effective (e.g., uncertainty in realizing all Reasonable Further Progress requirements) than projected. Timing of emissions reductions may also be affected (e.g., sulfur emissions reductions from utility sources have occurred more rapidly than projected for this analysis).

* The classification of each potential source of error reflects the best judgment of the section 812 Project Team. The Project Team assigns a classification of "potentially major" if a plausible alternative assumption or approach could influence the overall monetary benefit estimate by approximately five percent or more; if an alternative assumption or approach is likely to change the total benefit estimate by less than five percent, the Project Team assigns a classification of "probably minor."

Table 3. Key uncertainties associated with air quality modeling from the First Prospective Analysis.

Page 1 of 4

Potential Source of Error	Direction of Potential Bias for Net Benefits Estimate	Likely Significance Relative to Key Uncertainties in Net Benefit Estimate*
PM10 and PM2.5 concentrations in the East (RADM domain) are based exclusively on changes in the concentrations of sulfate and nitrate particles, omitting the effect of anticipated reductions in organic or primary particulate fractions.	Underestimate.	Potentially major. Nitrates and sulfates constitute major components of PM, especially PM2.5, in most of the RADM domain and changes in nitrates and sulfates may serve as a reasonable approximation to changes in total PM10 and total PM2.5. Of the other components, primary crustal particulate emissions are not expected to change between scenarios; primary organic carbon particulate emissions are expected to change, but an important unknown fraction of the organic PM is from biogenic emissions, and biogenic emissions are not expected to change between scenarios. If the underestimation is major, it is likely the result of not capturing reductions in motor vehicle primary elemental carbon and organic carbon particulate emissions.
The number of PM2.5 ambient concentration monitors throughout the U.S. is limited. As a result, cross estimation of PM2.5 concentrations from PM10 (or TSP) data was necessary in order to complete the "monitor level" observational dataset used in the calculation of air quality profiles.	Unable to determine based on the current information.	Potentially major. PM2.5 exposure is linked to mortality, and avoided mortality constitutes a large portion of overall CAAA benefits. Cross estimation of PM2.5, however, is based on studies that account for seasonal and geographic variability in size and species composition of particulate matter. Also, results are aggregated to the annual level, improving the accuracy of cross estimation.
Use of separate air quality models for individual pollutants and for different geographic regions does not allow for a fully integrated analysis of pollutants and their interactions.	Unable to determine based on current information.	Potentially major. There are uncertainties introduced by different air quality models operating at different scales for different pollutants. Interaction is expected to be most significant for PM estimates. However, important oxidant interactions are represented in all PM models and the models are being used as designed. The greatest likelihood of error in this case is for the summer period in areas with NOx inhibition of ambient ozone (e.g., Los Angeles).

Table 3. Key uncertainties associated with air quality modeling from the First Prospective Analysis.

Page 2 of 4

Potential Source of Error	Direction of Potential Bias for Net Benefits Estimate	Likely Significance Relative to Key Uncertainties in Net Benefit Estimate*
Future-year adjustment factors for seasonal or annual monitoring data are based on model results for a limited number of simulation days.	Overall, unable to determine based on current information.	Probably minor. RADM/RPM and REMSAD PM modeling simulation periods represent all four seasons and characterize the full seasonal distribution. Potential overestimation of ozone, due to reliance on summertime episodes characterized by high ozone levels and applied to the May-September ozone season, is mitigated by longer simulation periods, which contain both high and low ozone days. Also, underestimation of UAM-V western and UAMIV Los Angeles ozone concentrations (see below) may help offset the potential bias associated with this uncertainty.
Comparison of modeled and observed concentrations indicates that ozone concentrations in the western states were somewhat underpredicted by the UAM-V model, and ozone concentrations in the Los Angeles area were underestimated by the UAM-IV model.	Unable to determine based on current information.	Probably minor. Because model results are used in a relative sense (i.e., to develop adjustment factors for monitor data) the tendency for UAM-V or UAM to underestimate absolute ozone concentrations would be unlikely to affect overall results. To the extent that the model is not accurately estimating the relative changes in ozone concentrations across regulatory scenarios, the effect could be greater.
Ozone modeling in the eastern U.S. relies on a relatively coarse 12 km grid, suggesting NOx inhibition of ambient ozone levels may be under represented in some eastern urban areas. Coarse grid may affect both model performance and response to emissions changes.	Unable to determine based on current information.	Probably minor. Though potentially major for eastern ozone results in those cities with known NOx inhibition, ozone benefits contribute only minimally to net benefit projections in this study. Grid size affects chemistry, transport, and diffusion processes which in turn determine the response to changes in emissions, and may also affect the relative benefits of low-elevation versus high-stack controls. However, the approach is consistent with current state-of-the-art for regional-scale ozone modeling.

Table 3. Key uncertainties associated with air quality modeling from the First Prospective Analysis.

Page 3 of 4

Potential Source of Error	Direction of Potential Bias for Net Benefits Estimate	Likely Significance Relative to Key Uncertainties in Net Benefit Estimate*
UAM-V modeling of ozone in the western U.S. uses a coarser grid than the eastern UAM-V (OTAG) or UAM-IV models, limiting the resolution of ozone predictions in the West.	Unable to determine based on current information.	Probably minor. Also, probably minor for ozone results. Grid cell-specific adjustment factors for monitors are less precise for the west and may not capture local fluctuations. However, exposure tends to be lower in the predominantly non-urban west, and models with finer grids have been applied to three key population centers with significant ozone concentrations. May result in underestimation of benefits in the large urban areas not specifically modeled (e.g., Denver, Seattle) with finer grid.
Emissions estimated at the county level (e.g., area source and motor vehicle NOx and VOC emissions) are spatially and temporally allocated based on land use, population, and other surrogate indicators of emissions activity. Uncertainty and error are introduced to the extent that area source emissions are not perfectly spatially or temporally correlated with these indicators.	Unable to determine based on current information.	Probably minor. Potentially major for estimation of ozone, which depends largely on VOC and NOx emissions; however, ozone benefits contribute only minimally to net benefit projections in this study.
The REMSAD model underpredicted western PM concentrations during fall and winter simulation periods.	Unable to determine based on current information.	Probably minor. Because model results are used in a relative sense (i.e., to develop adjustment factors for monitor data) REMSAD's underestimation of absolute PM concentrations would be unlikely to significantly affect overall results. To the extent that the model is not accurately estimating the relative changes in PM concentrations across regulatory scenarios, or the individual PM components (e.g., sulfates, primary emissions) do not vary uniformly across seasons, the effect could be greater.
Lack of model coverage for acid deposition in Western states.	Underestimate	Probably minor. Because acid deposition tends to be a more significant problem in the eastern U.S. and acid deposition reduction contributes only minimally to net monetized benefits, the monetized benefits of reduced acid deposition in the western states would be unlikely to significantly alter the total estimate of monetized benefits.

Table 3. Key uncertainties associated with air quality modeling from the First Prospective Analysis.

Page 4 of 4

Potential Source of Error	Direction of Potential Bias for Net Benefits Estimate	Likely Significance Relative to Key Uncertainties in Net Benefit Estimate*
Uncertainties in biogenic emissions inputs increase uncertainty in the AQM estimates.	Unable to determine based on current information.	Probably minor. Potentially major impacts for ozone outputs, but ozone benefits contribute only minimally to net benefit projections in this study. Uncertainties in biogenics may be as large as a factor of 2 to 3. These biogenic inputs affect the emissions-based VOC/NOx ratio and, therefore, potentially affect the response of the modeling system to emissions changes.

* The classification of each potential source of error reflects the best judgment of the section 812 Project Team. The Project Team assigns a classification of "potentially major" if a plausible alternative assumption or approach could influence the overall monetary benefit estimate by approximately five percent or more; if an alternative assumption or approach is likely to change the total benefit estimate by less than five percent, the Project Team assigns a classification of "probably minor."

Phase 1 will be carried out in seven steps:

1. Review, summarize, and update uncertainties identified in the First Prospective Analysis – provide the foundation for further revisions.
2. Perform an updated literature review – ensure that the most recent information on uncertainties is considered. This review will be extended to include state, regional, and federal agency reports.
3. Develop an annotated bibliography – provide a summary of important literature and facilitate further discussions.
4. Prepare an initial summary of key uncertainties – provide an initial summary of IAQMS uncertainties for the Second Prospective Analysis.
5. Hold a two-day workshop to solicit additional information. Selected government, academic, and industry scientists will be invited to review materials developed in Steps 1–4 and provide more input on additional sources of relevant literature; opinions on importance and magnitude of uncertainties; and recommendations for Phase 2 and 3 assessments. The focus of this workshop will be to develop consensus on the range and importance of uncertainties summarized in Step 4.
6. Prepare revised summary of key uncertainties. Based on input received during the two-day workshop, a revised summary of IAQMS uncertainties for the Second Prospective Analysis will be developed.
7. Prepare recommendations for assessments. Based on input received during the two-day workshop, recommendations for Phase 2 assessments will be prepared along with time and resources estimates.

The products of Phase 1 will be

- an annotated bibliography of relevant literature;
- a summary of the two-day workshop;
- a revised summary of IAQMS uncertainties suitable for inclusion in the Second Prospective Analysis report; and
- a recommendation for Phase 2 assessments.

Phase 2: Assess Importance of Uncertainties in IAQMS Predictions

The second phase of the process will be to assess the importance of key uncertainties identified in the first phase as requiring further assessment with the actual modeling system(s) to be used in the prospective analysis. Sensitivity analysis and process analysis will be performed to identify which processes and their uncertainties have the largest effect on predicted pollutant concentrations. The purpose of this phase is to identify uncertainties that have the largest impact on model predictions. Uncertainties having little or no effect on model predictions are assumed to have little effect on model response to emissions changes. However, a model prediction may be sensitive to input uncertainties without affecting its response to emission changes. Therefore, those uncertainties that are shown to significantly affect model predictions during this phase will be explored further in Phase 3 to determine whether or not they affect model response to emissions changes.

While some of these assessments may require additional simulations with the CMAQ model, many of the needed simulations may have already been performed as part of previous air quality modeling studies. Reviews of past studies will expedite the assessment process and reduce the resources needed to complete the assessment. For example, past modeling studies have been carried out by states (e.g., SIP modeling), RPO s (e.g., VISTAS), and EPA (e.g., the Clean Air Interstate Rule [CAIR]) that have included simulations involving alternative models, model configurations, chemistry and advection solvers, chemical mechanisms, and grid configurations.

Phase 2 will be carried out in six steps:

1. Assess the importance of emission uncertainties.
    - Estimate the range of uncertainties in base emissions (magnitude, composition, and spatial and temporal distribution). These estimates will be based on a synthesis of the literature review and workshop completed in Phase 1. Particular attention will be paid to range and uncertainty in measurements that have been used to develop emission factors and speciation profiles.
    - Determine which uncertainties have been assessed in previous modeling studies.
    - Assess air quality (AQ) sensitivity to uncertainties in base emissions using the EPA response surface model (RSM).

- Estimate range of uncertainties in projected emissions.
- Assess AQ sensitivity to uncertainties in projected emissions using RSM.
- Assess AQ sensitivity to selected uncertainties (e.g., emissions reactivity and spatial distribution) using limited sensitivity simulations with CMAQ.

2. Assess the importance of meteorological uncertainties.
    - Estimate range of uncertainties in meteorological inputs to CMAQ.
        - Advection and diffusion (wind and vertical diffusion inputs).
        - Clouds and precipitation.
        - Temperature.
    - Determine which uncertainties have been assessed in previous modeling studies.
    - Assess AQ sensitivity to meteorological uncertainties using limited sensitivity simulations with CMAQ.
3. Assess the importance of chemistry uncertainties.
    - Estimate range of uncertainties in model chemistry.
        - General uncertainty in chemical reaction rates and mechanisms.
        - Chemical mechanism selection.
        - Photolysis rates.
        - Initial and boundary conditions.
    - Determine which uncertainties have already been assessed in previous modeling studies.
    - Assess AQ sensitivity to uncertainties using limited sensitivity simulations with CMAQ.
4. Assess the importance of model configuration.
    - Describe range of options in model configuration.
        - Horizontal grid resolution.
        - Vertical grid structure and resolution.
        - Chemistry solver.
        - Advection solver.
        - Diffusion scheme
    - Determine which configurations have been assessed in previous modeling studies.
    - Assess AQ sensitivity to configuration options using results of past studies and limited sensitivity simulations with CMAQ.
5. Summarize model sensitivities to uncertainties.
6. Develop recommendations for Phase 3, including time and resources estimates.

The products of Phase 2 will be

- a summary of past model sensitivity analyses;
- a technical report on any additional CMAQ or RSM sensitivity analyses performed;
- a revised summary of IAQMS uncertainties suitable for inclusion in the Second Prospective Analysis report; and
- recommendation for Phase 3 assessments.

Phase 3: Assess Uncertainties in IAQMS Response to Emission Changes

In Phase 2, the IAQMS's sensitivity to model inputs within their range of uncertainty will be assessed. However sensitivity to an input does not mean that the sensitivity will influence the IAQMS's relative response to emission changes. Goals for this third phase are to identify which uncertainties affect the IAQMS's ***response to emission changes*** and provide a range of air quality outcomes for use in the rest of the analytic chain. Sensitivity analyses will be performed one at a time, in combinations, or by the DDM as appropriate. Some of the sensitivity simulations proposed may have already been performed by EPA or others. Only those uncertainties evaluated in Phase 2 and resulting in significant model sensitivity will be assessed in this phase. While some of these assessments may require additional simulations with the CMAQ model, many of the simulations may have already performed as a part of previous air quality modeling studies. Reviews of past studies will expedite the assessment process and reduce the resources needed to complete the assessment.

Phase 3 will be carried out in five steps:

1. Identify previous studies in which the impact of uncertainties on model responses to emission changes has been investigated.
2. Assess the impact of emission uncertainties on model response to emission changes:
    - assess changes in AQ relative response to uncertainties in emissions using RSM; and
    - assess changes in AQ relative response to selected uncertainties (e.g., emissions reactivity and spatial distribution) using limited sensitivity simulations with CMAQ.
3. Assess the impact of meteorological uncertainties on model response to emission changes using results of past studies or limited sensitivity simulations with CMAQ.
4. Assess the impact of chemistry uncertainties on model response to emission changes using results of past studies or limited sensitivity simulations with CMAQ.
5. Assess the impact of model configuration on model response to emission changes using results of past studies or limited sensitivity simulations with CMAQ.

The products of Phase 3 will be:

- a summary of past model sensitivity and uncertainty analyses;
- a technical report on any additional CMAQ and RSM sensitivity analyses performed; and

- a revised summary of IAQMS uncertainties suitable for inclusion in the Second Prospective Analysis report.

Phase 4: Final Synthesis

Because performing a comprehensive analysis of all key sensitivities and uncertainties, and their effect on IAQMS response will not be practical, at the end of the third phase, there will be many individual estimates of these effects. In this final phase, the results of all the preceding phases, with a focus on summarizing the most important uncertainties, will be synthesized. Estimates of the range of air quality outcomes in terms of relative changes in ozone and $PM_{2.5}$ will be developed for the scenarios being modeled for the Second Prospective Analysis. These estimates could then be carried through to the subsequent health effects estimation and valuation steps of the analytical process.

The products of Phase 4 will be

- a summary of the range of air quality outcomes in terms of relative changes in ozone and $PM_{2.5}$ estimated for the scenarios being modeled for the Second Prospective Analysis; and
- a final summary of IAQMS uncertainties suitable for inclusion in the Second Prospective Analysis report.

REFERENCES

Abdel-Aziz A. and Frey H.C. (2003a) Development of hourly probabilistic utility NOx emission inventories using time series techniques: Part I - univariate approach. *Atmos. Environ.* **37** (38), 5379-5389.

Abdel-Aziz A. and Frey H.C. (2003b) Development of hourly probabilistic utility NOx emission inventories using time series techniques: Part II: multivariate approach. *Atmos. Environ.* **37** (38), 5931-5401.

Abdel-Aziz A. and Frey H.C. (2004) Propagation of uncertainty in hourly utility NOx emissions through a photochemical grid air quality model: A case study for the Charlotte, NC, modeling domain. *ES&T*, **38**(7), 2153-1260.

Alcamo J. and Bartnicki J. (1987) A framework for error analysis of a long-range transport model with emphasis on parameter uncertainty. *Bulletin of the American Meteorological Society* **21** (10), 2121-2131.

Balentine H.W., Dickson R.J. and Oliver W.R. (1994) Development of uncertainty estimates for the Grand Canyon Visibility Transport Commission Emissions Inventory. Radian Corporation Technical Memorandum, Sacramento, California. December.

Beck L., Peer R., Bravo L., and Yan Y. (1994) A data attribute rating system. Presented at *Air & Waste Management Association's Emission Inventory Conference*, Raleigh, North Carolina.

Beck M.B., Ravetz J.R., Mulkey L.A. and Barnwell T.O. (1997) On the problem of model validation for predictive exposure assessments. *Stochastic Hydrology and Hydraulics* **11**, Springer-Verlag, 229-254.

Benkovitz C.M. (1985) Framework for uncertainty analysis of the NAPAP emissions inventory. U.S. Environmental Protection Agency, Air and Energy Engineering Research Laboratory. EPA/600/7-85/036. Research Triangle Park, North Carolina.

Benkovitz C.M. and Oden N.L. (1989) Individual versus averaged estimates of parameters used in large scale emissions inventories. *Atmos. Environ.* **23**, 903-909.

Bergin M.S., Noblet G.S., Petrini K., Dhieux J.R., Milford J.B., and Harley R.A. (1999) Formal uncertainty analysis of a Lagrangian photochemical air pollution model. *ES&T* **33** (7), 1116-1126.

Bischof C.A., Carle A., Corliss G., Griewank A., and Hovland P. (1992) ADIFOR: Generating derivative codes from Fortran programs. *Scientific Programming*, 1(1):11-29.

Carmichael G.R., Sandu A., and Potra F. (1997) Sensitivity analysis for atmospheric chemistry models via automatic differentiation. *Atmos. Environ.* **31**, 475-489.

Chang M., Cardelino C., Chameides W., and Chang W. (1993) An iterative procedure to estimate emission inventory uncertainties. Presented at *The Regional Photochemical Measurement and Modeling Studies Meeting of the Air & Waste Management Association at San Diego, California. November.*

Chang M., Cardelino C., Hartley D., and Chang W. (1995) Inverse Techniques for developing historical and future year emission inventories. Presented at *The Emission Inventory: Programs and Progress Specialty Conference of the Air & Waste Management Association, Research Triangle Park, North Carolina. October 11-13.*

Chi T.R., Unal A.D., and Tian R.A. (2004) Uncertainty of NONROAD emissions in Georgia. *13th Annual Emission Inventory Conference, Clearwater, FL.*

Cho S.-Y., Carmichael G. R., and Rabitz H. (1987) Sensitivity analysis of the atmospheric reaction-diffusion equation, *Atmos. Environ*., **21**, 2589.

Chow J.C., Watson J.G., Lowenthal D.H., Soloman P.A., Magliano K.L., Ziman S.D., and Richards L.W. (1992) PM_{10} Source apportionment in California's San Joaquin Valley. *Atmos. Environ.* **26A**, 3335-3354.

Claiborn C., Mitra A., Adams G., Bamesberger L., Allwine G., Kantamaneni R., Lamb B., and Westberg H. (1995) Evaluation of PM_{10} Emission rates from paved and unpaved roads using tracer techniques. *Atmos. Environ.* **29**, 1075-1089.

Coats C.J. Jr., and Houyoux M.R. (1996) Fast emissions modeling with the Sparse Matrix Operator Kernel Emissions modeling system. Paper presented at *The Emission Inventory: Key to Planning, Permits, Compliance, and Reporting*, *Air and Waste Management Association, New Orleans, La., Sept. 4- 6.*

Conover W.J. (1971) *Practical nonparametric statistics*. 2nd Ed., Wiley, New York.

Dabdub D., DeHaan L.L., and Seinfeld J.H. (1999) Analysis of ozone in the San Joaquin Valley of California. *Atmos. Environ.* **33**, 2501-2514.

DaMassa J. (1992) Technical guidance document for photochemical modeling. Technical Support Division, California Air Resources Board, Sacramento, CA.

Derwent R. and Hov Ø. (1988) Application of sensitivity and uncertainty analysis techniques to a photochemical ozone model. *J. Geophys. Res.* **93** (D5), 5185-5199.

Deuel H.P., Douglas S.G. and Burton C.S. (1998) Estimation of modeling system noise for two applications of the UAM-V modeling system: one-hour ozone, eight-hour ozone, and ozone exposure. SYSAPP-98/27, Systems Applications International., Inc., 38 pages.

Doty K., Tesche T.W., McNally D.E., Timin B., Mueller S.F. (2002) Meteorological modeling for the Southern Appalachian Mountains Initiative (SAMI). Final report prepared by University of Alabama in Huntsville, Huntsville, AL, July.

Dunker A.M. (1981) Efficient calculation of sensitivity coefficients for complex atmospheric models. *Atmos. Environ.* **15** (7), 1155-1161.

Dunker A.M. (1984) The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. *J. Chem. Phys.* **81**, 2385-2393.

Efron B., and Tibshirani R. (1991) Statistical data analysis in the computer age. *Science* **253**, 390-395.

Emission Inventory Improvement Program (1996) Evaluating the uncertainty of emission estimates - Final Report. Prepared for the Quality Assurance Committee, Emission Inventory Improvement Program by Radian Corporation, Post Office Box 13000, Research Triangle Park, NC, July.

Environment Canada (1994) Uncertainties in Canada's 1990 greenhouse gas emission estimates, a quantitative assessment. Prepared by T.J. McCann & Associates. Unpublished report. Ottawa, Ontario, March.

Freeman D.L., Egami R.T., Robinson N.F., and Watson J.G. (1986) A method for propagating measurement uncertainties through dispersion models. *J. Air Pollution Control Assoc.* **36**, 246-253.

Frey H.C. (1992) Quantitative analysis of uncertainty and variability in environmental policy making. Directorate for Science and Policy Programs, American Association for the Advancement of Science, Washington, DC.

Frey H.C. (1998) Estimates of uncertainty in air quality model inputs based upon expert elicitation. Prepared by North Carolina State University for Hanna Consultants, Kennebunkport, Maine, September 28, 1998.

Frey H.C. and Bammi S. (2002) Quantification of variability and uncertainty in lawn and garden equipment NO_x and Total hydrocarbon emission factors. *J. Air and Waste Manag. Assoc.* **52**(4), 435-448.

Frey H.C. and Bammi S. (2003) Probabilistic nonroad mobile source emission factors. *J. Envir. Engrg.* **129**(2), 162-168.

Frey H.C. and Li S. (2003) Methods for quantifying variability and uncertainty in AP-42 emission factors: case studies for natural gas-fueled engines. *J. Air and Waste Manag. Assoc.* **53** (12).

Frey H.C. and Zhao Y. (2004) Quantification of variability and uncertainty for air toxic emission inventories with censored emission factor data. *ES&T*, **38** (22), 6094-6100.

Frey H.C. and Zheng J. (2002a) Probabilistic analysis of driving cycle-based highway vehicle emission factors. *ES&T*, **36** (23), 5184-5191.

Frey H.C. and Zheng J. (2002b) Quantification of variability and uncertainty in air pollutant emission inventories: method and case study for utility NO_x emissions. *J. Air and Waste Manag. Assoc.* **52** (9), 1083-1095.

Frey H.C., Bharvirkar R., and Zheng J. (1999) Quantitative analysis of variability and uncertainty in emissions estimation. Prepared by North Carolina State University for the U.S. Environmental Protection Agency, July.

Fujita E.M., Croes B.E., Bennett C.L., Lawson D.R., Lurmann F.W., and Main H.H. (1992) Comparison of emission inventory and ambient concentration ratios of CO, NMOG, and NO_x in California's South Coast Air Basin. *J. Air and Waste Manag. Assoc.* **42** (3), 264-276.

Gao D., Stockwell W.R., and Milford J.B. (1996) Global uncertainty analysis of a regional-scale gas-phase chemical mechanism. *J. Geophys. Res.* **101**, 9107-9119.

Gatz D.F., and Smith L. (1995a) The standard error of a weighted mean concentration - I. bootstrapping vs other methods. *Atmos. Environ.* **29**, 1185-1193.

Gatz D.F., and Smith L. (1995b). The standard error of a weighted mean concentration - I. estimating confidence intervals. *Atmos. Environ.* **29**, 1195-1200.

Hanna S.R., Chang J.C., and Fernau M.E. (1998) Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmos. Environ.* **32**, 3619-3628.

Hanna S.R., Lu Z., Frey H.C., Wheeler N., Vukovich J., Arunachalam S., Fernau M., and Hansen D.A. (2000) Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG Domain. *Atmos. Environ.* **34**.

Hanna S.R., Lu Z., Frey H.C., Wheeler N., Vukovich J., Arunachalam S., Fernau M., and Hansen D.A. (2001) Uncertainties in predicted ozone concentration due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmos. Environ.* **35**, 891-903.

Harley R.A., Sawyer R.F., and Milford J.B. (1997) Updated photochemical modeling for California's South Coast Air Basin: comparison of chemical mechanisms and motor vehicle emission inventories. *Environ. Sci. Technol.* **31**, 2829-2839.

Hartley D., and Prinn R.J. (1993) Feasibility of Determining Surface Emissions of Trace Gases Using an Inverse Method in a Three-Dimensional Chemical Transport Model. *J. Geophys. Res.* **98**, 5183-5197.

Hass H., Builtjes P.J.H., Simpson D. and Stern R. (1997) Comparison of model results obtained with several European regional air quality models. *Atmos. Environ*. **31**, 3259-3279.

Hogrefe C., Rao S.T., Zurbenko I.G., and Porter P.S. (2000) Interpreting the Information in Ozone Observations and Model Predictions Relevant to Regulatory Policies in the Eastern United States. *Bulletin of the American Meteorological Society* **81**, 2083-2106 (9).

Horie Y. (1988) Handbook on procedures for establishing the uncertainties of emission estimates. Prepared for the California Air Resources Board Research Division, by Valley Research Corporation, ARB Contract A5-184-32.

Horie Y. and Shorpe A.L. (1989) Development of procedures for establishing the uncertainties of emission estimates. *82nd Annual Air and Waste Management Association Meeting, Anaheim, California, June* (Paper 89-24.7).

Houyoux M.R., Loughlin D.H., Holland A.P., Frey H.C., and Abdel-Aziz A. (2003) Design, application, and recommendations for including inventory uncertainties in emission inventory preparation for modeling. *12th International Emission Inventory Conference, Emission Inventories - Applying New Technologies, San Diego, April 29 - May 1.*

Hubbell B. (2003) Identification of cost effectiveness measures using response-surface modeling. Presented at *Symposium on Cost-Effectiveness Analysis for Multiple Benefits, Washington, DC, September 9.*

Hwang D., Byun D., and Odman M.T. (1997) An automatic differentiation technique for sensitivity analysis of numerical advection schemes in air quality. *Atmos. Environ.*, **31**(6), 879-888.

Iman R.L., and Helton J.C. (1988) An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis* **8**, 71-90.

International Atomic Energy Agency (1989) Evaluating the reliability of predictions made using environmental transfer models. IAEA Safety Series No. 100, International Atomic Energy Agency, Vienna, Austria.

Jang J.C.C., Jeffries H.E., and Tonnesen S. (1995) Sensitivity of ozone to model grid resolution-I. Detailed process analysis for ozone chemistry. *Atmos. Environ.* **29** (21), 3085-3100.

Jeffries H.E. (1997) Use of integrated process rate analyses to perform source attribution for primary and secondary pollutants in Eulerian air quality models. Presented at *U.S. EPA Source Attribution Workshop, Research Triangle Park, NC, July 16-18.*

Jeffries H.E., Keating T., and Wang Z. (1996) Integrated process rate analysis of the impact of NO_x emission height on UAM-modeled peak ozone levels. Final Topical Report prepared for Gas Research Institute, Chicago, IL.

Kini M.D. and Frey H.C. (1997) Probabilistic evaluation of mobile source air pollution. Prepared for the Center for Transportation and Environment by North Carolina State University, Raleigh, NC.

Linstene H.A., and Turoff M. (1975) *The Delphi Method, Techniques and Applications*. Edited by Linstene and Turoff. Addison-Wesley Publishing Company.

Lo C.S. and Jeffries H.E. (1997) A quantitative technique for assessing pollutant source location and process composition in photochemical grid models. Presented at the *Annual Air & Waste Management Association Meeting, Toronto, Ontario*.

Lowenthal D.H., Chow J.C., Watson J.G., Neuroth G.R., Robbins R.B., Shafritz B.P., and Countess R.J. (1992) The Effects of collinearity in the ability to determine aerosol contributions from diesel- and gasoline-powered vehicles using the chemical mass balance model. *Atmos. Environ.* **26A**, 2341-2351.

Mangat T.S., Robinson L.H., and Switzer P. (1984) *Medians and percentiles for emission inventory totals.* Presented at the *77th Annual Meeting of the Air Pollution Control Association, San Francisco, California. June*.

Milford J.B., Gao D., Russell A.G., and McRae G.J. (1992) Use of sensitivity analysis to compare chemical mechanisms for air-quality modeling. *ES&T* **26**, 1179-1189.

Mitchell W.J., Suggs J.C., and Streib E.W. (1995) A statistical analysis of stationary source compliance test audit data. *J. Air & Waste Manag. Assoc.* **45**, 83-88.

Moore G.E. and Londergan R.J. (2001) Sampled Monte Carlo uncertainty analysis for photochemical grid models. *Atmos. Environ.* **35**, 4863-4876.

Morgan M.G., and Henrion M. (1990) *Uncertainty*; Cambridge University Press: New York.

Mulholland M., and Seinfeld J.H. (1995) Inverse Air Pollution Modelling of Urban-Scale Carbon Monoxide Emissions. *Atmos. Environ.* **29**, 497-516.

National Council on Radiation Protection and Measurements (1996*)* A guide for uncertainty analysis in dose and risk assessments related to environmental contamination. (ed., F.O. Hoffman), NCRP Commentary Number 14, Nat. Council on Radiation Protection and Meas., Bethesda, MD.

National Exposure Research Laboratory (1999) Science algorithms of the EPA Models-3 Community Multiscale Air Quality (CMAQ) modeling system. Report prepared by the National Exposure Research Laboratory, Research Triangle Park, NC, EPA/600/R-99/030 (peer reviewed), March.

National Research Council (1991) *Rethinking the ozone problem in urban and regional air pollution*. National Academies Press, Washington, DC.

National Research Council (2004) Air quality management in the United States. Committee on Air Quality Management in the United States, National Academy Press, Washington, DC, January.

Oden N.L. and Benkovitz C.M. (1990) Statistical implications of the dependence between the parameters used for calculations of large scale emissions inventories. *Atmos. Environ.* **24A** (3), 449-456.

Odman M. T., Wilkinson J.G., McNair L.A., Russell A.G., Ingram C.L., and Houyoux M.R. (1996) Horizontal advection solver uncertainty in the Urban Airshed Model. Final Report. Environmental Programs, MCNC, North Carolina Supercomputing Center, Research Triangle Park. Prepared for California Air Resources Board and California Environmental Protection Agency, April 30.

Olerud D., Alapaty K., and Wheeler N. (2000) Meteorological modeling of 1996 for the United States with MM5. Final report for Task Order Number CAA689805 prepared for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards by MCNC Environmental Programs, Research Triangle Park, NC, September.

Oreskes N., Shrader-Frechette K., and Belitz K. (1994) Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, 263, 641-646.

Ozone Transport Assessment Group (1997) OTAG regional and urban scale modeling, Version 1.1. Modeling Report.

Peer R.L., Epperson D.L., Campbell D.L., and von Brook P. (1992) Development of an empirical model of methane emissions from landfills. United States Environmental Protection Agency, Air and Energy Engineering Laboratory, EPA-600/R-92-037, Research Triangle Park, North Carolina.

Pierson W.R., Gertler A.W., and Bradow R.L. (1990) Comparison of the SCAQS Tunnel Study with other on-road vehicle emission data. *J.Air & Waste Manag. Assoc.* **40** (11), 1495-1504.

Pollack A.K., Bhave P., Heiken J., Lee K., Shepard S., Tran C., Yarwood G., Sawyer R.F., and Joy B.A. (1999) Investigation of emission factors in the California EMFAC7G model. Prepared for the Coordinating Research Council by ENVIRON International Corporation, Novato, CA, PB99-149718NZ.

Rabitz H., Kramer M. and Docol D. (1983) Sensitivity analysis in chemical kinetics. *Annual Review of Physical Chemistry* 34, 419-461.

Reynolds S.D., Michaels H.M., Roth P.M., Tesche T.W., McNally D., Gardner L. and Yarwood G. (1996) Alternative base cases in photochemical modeling: their construction, role and value, *Atmos. Environ.* **30** (12), 1977-1988.

Rhodes D.S. and Frey H.C. (1997) Quantification of variability and uncertainty in AP-42 emission factors using bootstrap simulation. Presented at the *Emission Inventory: Planning for the Future, Air and Waste Management Association: Pittsburgh, Pennsylvania, October*, 147-161.

Saeger M. (1994) Procedures for Verification of emissions inventories, Final Report. Prepared for Emissions Inventory Branch, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina (EPA Contract No. 68-D3-0030).

Scheff P.A., Wadden R.A., Kenski D.M., and Chang J. (1995) Receptor model evaluation of the SEMOS ambient NMOC measurements. Paper 95-113C.03. Presented at the *88th Annual Meeting of the Air Pollution Control Association, San Antonio, Texas. June*.

Schere K. L. and Coats C. J. (1992) A stochastic methodology for regional wind-field modeling. *J. Appl. Meteor.* 31, 1407-1425.

Seefeld S. and Stockwell W.R. (1999) First-order sensitivity analysis of models with time-dependent parameters: an application to PAN and ozone. *Atmos. Environ.* **33**, 2941-2953.

Seigneur C., Tesche T.W., Roth P.M., and Reid L.E. (1981) Sensitivity of a complex urban air quality model to input data. *Journal of Applied Meteorology* **20**, 1020-1040.

Solarski R.S., Butler D.M., and Rundel R.D. (1978) Uncertainty propagation in a stratospheric model, 2. Monte Carlo analysis of imprecisions due to reaction rates. *Journal of Geophysical Research* **83**, 3074-3078.

South Coast Air Quality Management District (SCAQMD) (1982) Uncertainty of 1979 emissions data, Chapter IV: 1983 Air Quality Management Plan Revision, Appendix IV-A; 1979 Emissions Inventory for the South Coast Air Basin, Revised October 1982. Planning Division, El Monte, California. October.

Spellicy F.L., Draves J.A., Crow W.L., Herget W.F., and Buchholtz W.F. (1992) A Demonstration of optical remote sensing in a petrochemical environment. Presented at the *Air & Waste Management Association Specialty Conference on Remote Sensing, Houston, Texas. April 6-9.*

Steiner C.K.R., Gardner L., Causley M.C., Yocke M.A., and Steorts W.L. (1994) Inventory quality issues associated with the development of an emissions inventory for the Minerals Management Service Gulf of Mexico Air Quality Study. In: The Emission Inventory: Perception and Reality. Proceedings of an *International Specialty Conference, VIP-38. Air & Waste Management Association, Pittsburgh, Pennsylvania*.

Straume A.G., N'Dri Koffi E., and Nodop K. (1998) Dispersion modeling using ensemble forecasts compared to ETEX measurements. *J. Applied Meteorology*, 37, 1444-1456.

Tatang M.A., Pan W., Prinn R.G., and McRae G.J. (1997) An efficient method for parametric uncertainty analysis of numerical geophysical models. *J. Geophys. Res.* **102**, 21925-21932.

U.S. Environmental Protection Agency (1995) Compilation of air pollutant emission factors, volume i: stationary point and area sources, Fifth Edition, AP-42. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards. Research Triangle Park, North Carolina.

U.S. Environmental Protection Agency (1999) Draft guidance on the use of models and other analyses in attainment demonstrations for the 8-hour ozone NAAQS. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, EPA-454/4-99-004.

U.S. Environmental Protection Agency (2001) Draft guidance for demonstrating attainment of air quality goals for $PM_{2.5}$ and regional haze. Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, Draft 2.1, January.

U.S. Environmental Protection Agency (2006a) Technical support document for the proposed Mobile Source Air Toxics Rule: Ozone Modeling. Prepared by U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, February.

U.S. Environmental Protection Agency (2006b) Technical support document for the proposed PM NAAQS Rule: Response Surface Modeling. Prepared by U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC, February.

Vuilleumier L., Brown N. J., Orlando E., Harley R.A., and Reynolds S.D. (2000) Review and improvement of methods for estimating rates of photolysis in photochemical models, Final Report for Contract No. 96-335, prepared for California Air Resources Board and the California Environmental Protection Agency by Lawrence Berkeley National Laboratory, Berkeley, CA., December.

Wagner K.K., and Wheeler N.J. (1989) Variability of photochemical modeling emission sensitivity with respect to meteorology. Preprints, *Sixth Joint Conference on Applications of Air Pollution Meteorology, January 30 - February 3, 1989, Anaheim, CA*, Amer. Meteor. Soc., 222-223.

Wagner K.K., and Wheeler N.J.M. (1988) 1995 Ozone modeling boundary condition sensitivity, Modeling & Meteorology Branch Memorandum, California Air Resources Board, July 5.

Wagner K.K., Wheeler N.J.M., and McNerny D.L. (1992) The effect of emission inventory uncertainty on Urban Airshed Model sensitivity to emission reductions, Tropospheric Ozone: Nonattainment and Design Value Issues, J.J. Vostal, Ed., *Transactions of an international conference, TR-23, Air and Waste Management Assoc., 123-134.*

Watson J.G., Cooper J.A., and Huntzicker J.J. (1984) The effective variance weighing for least squares calculations applied to the mass balance receptor model. *Atmospheric Environment* 18:1347-1355.

Wheeler N.J.M. (1992) Urban Airshed Model sensitivity to horizontal transport paths. Tropospheric Ozone and the Environment II: Effects, Modeling and Control, R.L. Berglund, Ed., *Transactions of an international conference, Atlanta, GA, November 5-7, 1991, TR-20, Air and Waste Management Assoc., 400-410.*

Whitten G., and Killus J. (1998) Final Report: Effect of chemical mechanism uncertainty on airshed model results, SYSAPP-97/57, Systems Applications International, Inc. February, Prepared for California Air Resources Board, Technical Support Division, Copyright 1997 by Systems Applications International, Inc.

Winner D.A., Cass G.R., and Harley R.A. (1995) Effect of alternative boundary conditions on predicted ozone control strategy performance: A case study in the Los Angeles area. *Atmos. Environ.* **29**, 3451-3464.

Zhao Y. and Frey H.C. (2004) Development of probabilistic emission inventories of air toxics for Jacksonville, Florida, USA. J. Air and Waste Manage. Assoc. 54 (11), 1405-1421.

ATTACHMENT 2

APPROACH FOR CHARACTERIZING UNCERTAINTY AND VARIABILITY IN SECTION 812 CASE STUDY OF THE BENEFITS OF CLEAN AIR ACT-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON METROPOLITAN AREA

Approach For Characterizing Uncertainty And Variability In Section 812 Case Study Of The Benefits Of Clean Air Act-Related Reductions In Benzene In The Houston Metropolitan Area

The main method for characterizing uncertainty in the benzene case study will be through the use of several sensitivity analyses. A central concentration-response estimate was selected from a cohort study linking benzene and leukemia mortality (the "Pliofilm Cohort study"). In order to reflect the uncertainty in this central estimate and its underlying assumptions, we plan to perform the following sensitivity analyses:

- Model uncertainty in the epidemiologic studies will be accounted for by running the lifetable model with estimates from the Pliofilm Cohort study derived using different assumptions about the level of exposure experienced by the cohort than the central estimate. The model will also be run with estimates from a separate cohort study (the "Chinese Worker Cohort study").

- Uncertainty in the length of the latency period between benzene exposure and leukemia mortality will be characterized by running the model with multiple risk estimates, each derived using a different assumption about latency (i.e, different patterns of weights for past exposures).

- Uncertainty in the mortality endpoint will be assessed by running the model with estimates based on deaths from a specific type of leukemia (acute myeloid/monocytic leukemia) in addition to the central estimate, which is based on all leukemias.

In addition to sensitivity analyses, the benzene case study analysis will account for variability in exposures of individuals both within and across age groups. EPA's HAPEM6 produces a distribution of exposure values for each age group. In order to reflect this variability, the benzene lifetable model will be run with 5th percentile exposure values for each age group and again with 95th percentile exposure values. While we recognize that the combination of these values across age groups will produce more extreme scenarios, they will allow us to bound the variability in exposure based on differences in time-activity patterns within age groups.

ATTACHMENT 3

IEc MEMORANDUM ON ALTERNATIVE METHODOLOGY FOR ESTIMATING EMISSIONS FOR THE YEAR 2001 WITHOUT-CAAA SCENARIO FOR THE ELECTRIC GENERATING UNIT SECTOR

MEMORANDUM | 24 July 2006

TO Jim DeMocker, EPA OAR/OPAR

FROM Jim Neumann, IEc

SUBJECT Alternative Methodology for Estimating Emissions for the Year 2001 Without-CAAA Scenario for the Electric Generating Unit Sector

INTRODUCTION

The purpose of this memorandum is to provide a brief outline of an alternative method for generating an emissions inventory for one component of the Section 812 Second Prospective: the year 2001 without-CAAA scenario for the electric generating unit (EGU) sector. In reviewing the draft emissions projections for the EGU sector, questions were raised within the 812 Project Team regarding the validity and reliability of the year 2001 IPM validation run being proposed for adoption as the 812 study's target year 2000 results. The results of that run are provided in Chapter 3 of the document, *Emission Projections for the Clean Air Act Second Section 812 Prospective Analysis Draft Report*, June 21, 2006. Appendices B and C of that document provide further detail comparing the results of the with-CAAA scenario for the year 2001 with actual historical emissions rates, fuel prices, and allowance prices.[1] Differences between the spatial distribution of emissions as modeled by IPM compared to the actual spatial distribution from continuous emissions monitor (CEM) data, and differences in modeled versus actual fuel and allowance prices for the historical, with-CAAA case, have led us to consider the possibility of an alternative approach for modeling the effect of the CAAA on the EGU sector in the year 2000 or 2001.

One of the key questions in determining whether an alternative approach is feasible is whether we can construct a defensible counterfactual without-CAAA scenario for the historical year 2001, for comparison to the historical CEM data. This memo focuses on a method for developing an EGU sector sulfur oxide emissions counterfactual that closely follows an approach developed by Dr. A. Denny Ellerman. The method is described in some detail in a series of published papers and working papers.[2]

---

[1] See, in particular, Exhibits B-1 through B-3 in Appendix B for emissions comparisons and Exhibit C-4 in Appendix C for fuel and allowance price comparisons.

[2] The working papers are: "The Sources of Emission Reductions: Evidence from U.S. SO2 Emissions from 1985 through 2002," (Ellerman and Dubroeucq, 2004), "Ex Post Evaluation of Tradable Permits: The U.S. Cap-and-Trade Program," (Ellerman, 2003a), and "Lessons from Phase II Compliance with the U.S. Acid Rain Program" (Ellerman, 2003b); all of which are available at http://web.mit.edu/ceepr/www/workingpapers.htm. The basic methodology for development of a counterfactual case, however, was first published in A. Denny Ellerman, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth M. Bailey, Markets for Clean Air: The U.S. Acid Rain Program, Center for Energy and Environmental Policy Research, Massachusetts Institute of Technology (Cambridge, MA: 2000).

The remainder of this memorandum consists of three sections. First, we provide some general background on the sulfur oxide provisions of Title IV of the Clean Air Act. Amendments Second, we describe the basic elements of constructing a counterfactual for sulfur oxide emissions as outlined in Ellerman et al. (2000) and Ellerman (2003a). Third, we provide a short summary of the steps that would be necessary to develop a unit-level counterfactual emissions inventory consistent with Ellerman's approach, for both sulfur oxide and nitrogen oxides.

SECTION 1 BACKGROUND ON SULFUR OXIDE COMPONENTS OF TITLE IV OF THE CAAA

Title IV of the CAAA imposed certain limitations on the overall amount of sulfur dioxide (SO_2) emissions from fossil fuel fired electric generating units. The restrictions were imposed in two phases. Phase I lasted from 1995 through 1999 and applied to generating units of capacity over 100 MWe and base period (an average of 1985 to 1987) emissions in excess of 2.5 pounds SO_2 per million BTU (#SO_2/mmbtu) of heat input. Phase I units were required to reduce emissions from approximately 10.0 million tons of SO_2 in the 1985 to 1987 period to approximately 6.9 million tons in 1999. This level is equivalent to 2.5 #SO_2/mmbtu of heat input for Phase I facilities.

Phase II began in 2000 and continues indefinitely. In Phase II, virtually all fossil fuel fired generating stations are included in the limits, which are set based on 1.2 # SO_2/mmbtu of heat input to all units in the base period. By 2010, electric generators must reduce SO2 emissions to 8.9 million tons per year, compared to 1985-1987 emissions of some 16 million tons.

To implement these restrictions, the U.S. EPA issues allowances to generating units to emit SO_2. Allowances are generally issued each year to those generating units that were in existence during the base period, in proportion to each unit’s baseline heat input. In addition, the EPA sells a certain number of allowances in a public auction on an annual basis. The allowances may be bought and sold between generating units (and other parties), and may be “banked” for future use.

SECTION 2 KEY ELEMENTS OF A COUNTERFACTUAL CASE

In the absence of Title IV, it is plausible to assume that SO_2 emissions rates would have been more like those seen just prior to adoption of the CAAA in 1990. This premise is the basis for the "simple counterfactual" scenario developed in Ellerman et al. (2000) (see Chapter 5, pages 110-113). The simple counterfactual relies on unit-level heat input data and unit level SO_2 emission rates from 1993; the counterfactual sulfur oxide emission estimates are the product, at each unit, of the heat input (in mmBTU) and the emissions rate at that unit in 1993 (in pounds SO_2/mmBTU).

Ellerman notes that three key assumptions are made in construction of this simple counterfactual. First, that none of the emissions reduction observed by 1993 was attributable to Title IV. Second, that unit emission rates would not have changed between 1993 and our target year in the absence of Title IV. Third, that heat input would not be influenced at the unit-level by Title IV.[3] Ellerman acknowledges the potential for error in

---

[3] Ellerman et al. (2000) describes a few obvious exceptions to this basic approach, including opt-in units and retirements, but concludes that the effect of the first can be modeled by the allocation of allowances to these units in the factual case, and the effect of the second is very small.

this estimate, and then develops a much more rigorous econometrically estimated counterfactual scenario to test the potential for error. In subsequent work, however, Ellerman concludes that the likely effect of Title IV on unit-level heat input is probably small, while acknowledging that the potential for errors in emissions rates may be significant.[4] We review each of these two important factors below.

ELECTRICITY DEMAND, PLANT DISPATCH, AND UNIT-LEVEL HEAT INPUT

Imposing limits on SO_2 emissions necessarily impacts the cost of generating electricity. For traditional rate-of-return regulated electric utility generating stations, that cost increase is generally passed on to ratepayers through regulated rates. For generating stations that are in competitive generation markets, however, the cost of the programs may be reflected in rates paid, leading to the potential for effects on electricity demand. Ellerman concludes that this effect is nonetheless likely to be small:

> "This counterfactual assumption has the effect of making the estimated emission reduction equal to the heat-input-weighted changes in observed emission rates at affected units and to assume that no emission reduction can be attributed to changes in demand, either at individual units or in the aggregate. Since the demand for electricity is price inelastic, the cost of SO_2 controls is relatively small on a kilowatt-hour basis, and the major element determining the dispatch, or utilization, of individual generating plants is the cost of fuel, the error arising from assuming no effect on demand is probably small. Nevertheless, to the extent that the added costs from the program reduce the demand for electricity or change the order of dispatch of generating units in meeting that demand, the effect of the program is under-estimated." (Ellerman 2003a, pg 6-7)

By way of background, overall U.S. electricity demand has continued to grow since the base period used for allocating allowances (1985 to 1987), although at a somewhat slower pace than the overall economy, indicating that the electricity intensity of the economy is declining modestly. From 1985 to 2004, domestic electricity consumption has increased by an average of 2.5 percent per year compared to GDP growth of 3.0 percent. Combined with the assumption of a constant emissions rate, then, an increasing trend in electricity demand yields a slightly upward sloping trajectory for aggregate counterfactual emissions.

SULFUR EMISSIONS RATE

Counterfactual sulfur emissions rate is determined by the sulfur content of the fuel used in the scenario and the emissions control technologies assumed (e.g, scrubbers). Fossil fuel-fired generating stations are affected by a variety of environmental restrictions that affect both of these parameters. For the purpose of developing the simple counterfactual, however, it appears reasonable to assume that the SO_2 emissions rate per unit of fossil fuel consumed would have remained constant from the base period levels in the counterfactual scenario. As noted above, the approach used by Dr. Ellerman was to adopt 1993 emissions rates; yet as Ellerman notes:

---

[4] See "Ex Post Evaluation of Tradable Permits: The U.S. Cap-and-Trade Program," Working Paper, A Denny Ellerman, 2003a, provided as an attachment to this memo.

> "A more likely source of error arises from the assumption about the counterfactual emission rate. To the extent that other environmental regulations, or changes in relative fuel prices, cause the emission rate at affected units to fall during the period of evaluation, the effect of the SO_2 program is over-estimated. Increases in the true counterfactual emission rates would have the opposite effect, but the scope for these is limited since all units face emission rate limits under the pre-existing command and control regulation and those limits are rarely, if ever, increased." (Ellerman 2003a, pg 7)

As noted above, the assumption of a 1993 emissions rate as the baseline rate implies that none of the emissions reduction observed by 1993 was attributable to Title IV. Because Title IV was passed in 1990, and with the knowledge that many EGUs complied early with the requirements of Title IV, this may seem like a poorly justified assumption. One of the major events that is important to keep in mind, however, is the rapid change in coal economics during the 1990 to 1993 period owing to the ready availability, at low costs, of low-sulfer Powder River Basin (PRB) coal. This change, coupled with the deregulation of railroad transport that greatly reduced the transport cost of these coals, meant that many utilities were able to switch to these low sulfur coals at no cost or even at a savings relative to the costs they faced for higher sulfur coal prior to this period. Ellerman et al. (2000) conducts an econometric analysis to support this assumption and concludes that "the effect of changing coal economics is clearly more important in the aggregate than are factors related to early compliance with Title IV." (see especially pages 99 to 105 in Ellerman et al. 2000). Predicting EGU activity in the absence of Title IV is never conclusive, but it is not unreasonable to assume that much of the compliance activity that occurred prior to 1993, at least that involving fuel switching rather than scrubber installation, would have occurred regardless of Title IV.

SECTION 3 NEXT STEPS IN DEVELOPING A COUNTERFACTUAL WITHOUT-CAAA SCENARIO FOR EGUS FOR THE YEAR 2001

The preceding discussion suggests that Title IV may have had only a relatively small impact on the construction and operation of coal-fired power plants during the relevant period. Nevertheless, because electricity generation from coal has increased even with Title IV, it is reasonable to assume that SO_2 emissions would also have increased, but for the adoption of Title IV, associated with this increase in coal generation. We therefore propose to construct a counterfactual sulfur emissions scenario using a constant emissions rate assumption, reflecting an assumption that Title IV had relatively little effect on fuel choice but an important effect on emissions rates from existing units.

Consistent with the analysis prepared by Dr. Ellerman, if this alternative method for characterizing emissions in 2001 is chosen we will set the baseline emissions rate for Phase I units at the actual 1993 levels, and for Phase II units at the actual 1998 levels. We will also examine any significant changes in emissions rate prior to those dates, to evaluate whether any major change resulted from Title IV. As noted above, this scenario implies that, but for Title IV, overall SO_2 emissions would have grown modestly because overall heat input and electric output at existing coal-fired generating plants increased over this period. Developing the unit level counterfactual estimates will be relatively straightforward, as IEc has already developed a database of electric generating units that includes actual SO_2 emissions and fuel heat inputs from 1985 to the present to support the analyses summarized in Appendices B and C of the draft emissions report. This database also allows us to employ hybrid approaches, such as using 1993 emissions rates for all facilities except those Phase I facilities that installed scrubbers in the 1990 to 1993 period, or Phase II facilities that installed scrubbers between 1990 and 1998. Unlike the fuel switching decision, which is greatly affected by the availability of inexpensive low sulfur PRB coal, the decision to install a scrubber could reasonably be attributed to Title IV and therefore represents an exception to the broad use of 1993 or 1998 emissions rates; for those facilities, we propose to use the 1990 rates as the baseline emissions rate.

The unit-level results for the counterfactual scenario will then be compared to the unit-level CEM data; the latter would constitute the with-CAAA scenario. These unit-level data would then be used as the EGU emissions inputs for the subsequent air quality modeling step in the overall analysis.

We have identified one potentially important concern in applying the simple Ellerman approach. Most of Ellerman's work has focused on overall cost estimates for Title IV. As a result, emissions outcomes may have been of secondary importance. In the Second Prospective analysis, however, emissions outcomes are the basis not only for cost estimates, but also for benefits estimates. As a result, the spatial distribution of differences between the with-CAAA and without-CAAA results are more important than they might be for an analysis focused only on cost estimates. Some of Ellerman's work applying the simple counterfactual does, however, evaluate whether emissions trading outcomes were optimal - the trading analysis might demand unit-level resolution of emissions outcomes. We propose to discuss with Dr. Ellerman whether adoption of the simple counterfactual might present any known biases in the spatial distribution of emissions outcomes. In particular, we hope to discuss with Dr. Ellerman whether and how adjustments to the simple counterfactual results might be applied to incorporate some of the unit-level shifts in dispatch that he estimates in his econometrically estimated counterfactual.

Once the emissions results for sulfur oxides are complete, remaining work would include the following tasks:

1. *Development of cost estimates for the SO_2 reductions.* We believe this task could be relatively straightforward. Costs could be generated either directly from the published Ellerman estimates, perhaps through direct collaboration with Dr. Ellerman, or by an offline application of IPM cost functions.

2. *Development of counterfactual estimates for NO_x emissions from EGUs.* The Ellerman work does not address nitrogen oxide emissions. We have not yet had the opportunity to develop a detailed approach, but a screening level estimate might involve a direct analog of the simple counterfactual approach outlined above, using 1993 emissions rates and historical heat input data at the unit level.

3. *Development of cost estimates for NO_x reductions.* Cost estimates for nitrogen oxide reductions might be based on cost functions available in the AirControlNET database for EGU NO_x controls, or could developed by an offline application of IPM cost functions.

Ex Post Evaluation of Tradable Permits:

The U. S. SO_2 Cap-and-Trade Program[1]

A. Denny Ellerman
Massachusetts Institute of Technology

Introduction

A Brief Description of the Program

The U.S. SO_2 cap-and-trade program was established as a result of the enactment of the 1990 Clean Air Act Amendments (1990 CAAA) under the authority granted by Title IV, which included several measures to reduce precursor emissions of acid deposition.[2] The SO_2 component consisted of a two-phase, cap-and-trade program for reducing SO_2 emissions from fossil-fuel burning power plants located in the continental forty-eight states of the United States. During Phase I, lasting from 1995 through 1999, electric generating units larger than 100 MW^e in generating capacity with an annual average emission rate in 1985 greater than 2.5 pounds of SO_2 per million Btu of heat input in 1985 (hereafter, $\#SO_2/mmBtu$) were required to reduce emissions to a level that would be, on average, no greater than 2.5 $\#SO_2/mmBtu$. In Phase II, beginning in 2000 and continuing indefinitely, the program was expanded to include fossil-fuel electricity generating units greater than 25 MW^e, or virtually all fossil-fuel power plants in the United States. Emissions from these affected units are limited, after accounting for any allowances banked from Phase I, to an annual cap of 8.9 million tons, or about half of total electric utility SO_2 emissions in the early 1980s. The Phase II cap is equivalent to an

---

[1] This paper was prepared as a case study report under the program for the "Ex Post Evaluation of Tradable Permits: Methodological and Policy Issues" being conducted by the National Policies Division of the Environmental Directorate of the OECD.

[2] The most important of the other measures reduced NO_x emissions by two million tons by imposing technology-based, maximum average annual NO_x emission rates on affected sources. In meeting these standards, utilities were allowed to average emission rates among the units they controlled, but not to trade NO_x emissions among utilities.

average emission rate of 1.2 #SO_2/mmBtu, when divided by the mid-1980s level of heat input at fossil-fuel burning power plants.

This cap on national SO_2 emissions was implemented by issuing tradable allowances—representing the right to emit one ton of SO_2 emissions—equal in total to annual allowed emissions from affected units in each year after 1995, and by requiring that the owners of these units surrender an allowance for every ton of SO_2 emitted. Allowances not used in the year for which they are allocated can be carried over or banked for future use by the original owner or by any party to whom the banked allowance is sold. Allowances are allocated to owners of affected units free of charge for the next thirty years, generally in proportion to each unit's average annual heat input during the three-year baseline period, 1985-87. A small percentage (2.8 percent) of the allowances allocated to affected units are withheld for sale through an annual auction conducted by the EPA to encourage trading and to ensure the availability of allowances for new generating units. The revenues from this auction are returned on a *pro rata* basis to the owners from whose allocations the allowances were withheld.

The SO_2 cap-and-trade program also contained several provisions that allowed generating units not subject to the cap until Phase II to opt-in to Phase I and to receive allowances for the year in which the unit participated. These units were then subject to the same compliance requirements as the 263 units that were mandated to be part of Phase I, namely, that they must surrender allowances equal to emissions in that year. Also, SO_2-emitting industrial sources not otherwise affected by Title IV could establish baselines and be allocated allowances and participate like any other unit in Phases I and II.

The Political and Regulatory Context of Title IV

Three features of the political and regulatory context are important in evaluating the SO_2 cap-and-trade program. The first is that the cap-and-trade system is not the only means, nor the first means, of controlling SO_2 emissions from electric utility power plants in the United States. The cap-and-trade system supplements an extensive set of command-and-control regulations that has been in effect since the early 1970s. These regulations take two principal forms according to whether power plants were in existence when the regulations implementing the 1970 Clean Air Act Amendments became effective. Plants already in existence or under construction in 1971 must meet emission rate limits imposed by State Implementation Plans (SIPs), which the individual states are required to develop in order to bring all areas of the country into compliance with National Ambient Air Quality Standards (NAAQS) for six "criteria" pollutants (including SO_2). New units constructed after the effective date of the 1970 Amendments are required to meet the New Source Performance Standard (NSPS), which is a technology-based, uniform national requirement that, in the case of SO_2, effectively requires new coal-fired generating plants to install flue gas desulfurization equipment (or a scrubber).[3] New sources have additional requirements if they are to be located in areas not in attainment with the NAAQS (non-attainment areas). Sources locating in areas that are in attainment may also face prevention of significant deterioration (PSD) requirements, which are intended to ensure that areas in attainment do not slip into non-attainment status. Finally, any source located near a national park or other pristine (Class I) area may be required to meet additional limits, such as those aimed at preserving visibility. Typically, all of these pre-existing regulatory requirements impose either emission rate limits or technology mandates on individual units. This complex and comprehensive, underlying command-and-control structure means that Title IV is not burdened with meeting all environmental objectives. Other regulatory mechanisms are available to ensure that adverse local health effects are avoided and that other environmental values, such as visibility, are preserved. Another consequence of this regulatory context is that the ability of individual power plants to participate in emissions trading can be, and often is, limited by these other requirements.

The second notable feature of the political and regulatory context is that the motives lying behind enactment of Title IV are mixed, as is the case for most legislation. The ostensible purpose and most commonly cited motive is to reduce the effects of acid deposition, a cumulative environmental problem, the effects of which are experienced

---

[3] The scrubber mandate for new units was added by the 1977 Amendments to the Clean Air Act. The original NSPS provisions of the 1970 Clean Air Act required only that emissions from new coal-fired power plants be limited to 1.2 #SO_2/mmBtu. This standard was achievable either by installing a scrubber or switching to a limited sub-set of coals (thereafter known as compliance coals) that emitted less than 1.2 #SO_2/mmBtu without scrubbing. Ackerman and Hassler (1981) provide the now classic account of the interest group politics and other considerations leading to the redefinition of the NSPS.

mainly in the Northeast in large part as a result of SO_2 emissions originating from the heavy concentration of coal-fired power plants in the Mid-West. Yet, SO_2 emissions from power plants located in other parts of the country, such as Florida, that have little effect on the Northeast or other areas suffering from acidic deposition are included in the Acid Rain Program; and emissions from these sources are considered, for the purposes of emissions trading, as completely equivalent to emissions from power plants located in areas that are far more likely, given the prevailing patterns of atmospheric transportation, to have an affect on sensitive receptor areas. Two other motives operated at the time of enactment. The first concerned fine particulates, which research on health effects was beginning to implicate as a threat to public health. Although considerable controversy surrounded the origin of fine particulates—and such questions would need to be resolved in order to revise the appropriate NAAQS—SO_2 emissions from coal-fired electric power plants were considered a likely contributor. A second, and probably more important, motive was a desire to narrow the disparity between the emission limits imposed on new sources by the NSPS and the limits imposed on existing sources by State Implementation Plans. If SO_2 emissions were to be reduced for any of these reasons, something more than the existing regulatory structure would be needed since nearly all areas of the United States were in compliance with the SO_2 NAAQS by the 1980s. Moreover, the use of tall stacks to loft SO_2 emissions high above ground to avoid violating the ambient standard exacerbated the acidic deposition in more distant down-wind regions. A fifty percent reduction in the aggregate level of SO_2 emissions came to be viewed as a measure that would at once significantly reduce the amount of SO_2-originated deposition in the Northeast, contribute to some reduction of fine particulates, and largely close the disparity between the emission requirements imposed on new and existing sources. It is telling with respect to this last motive that the emission rate standard used to decide the cap and to allocate allowances in Phase II is identical to the original New Source Performance Standard enacted in the 1970 Amendments to the Clean Air Act.

The third and final feature of the political and regulatory context surrounding enactment of the SO_2 cap-and-trade program is that it ended a decade of debate concerning additional controls on existing coal-fired power plants. Earlier proposals would have achieved a similar 50% reduction of total SO_2 emissions by mandating scrubbers on the largest power plants and mandating switching to lower sulfur coal with limited trading. These earlier proposals were viewed as very costly, they faced the adamant opposition of the Reagan Administration, and they failed to gain a legislative majority in several sessions of Congress. The willingness of the new Bush (père) Administration to back significant SO_2 emission reductions, so long as they were achieved by market-based mechanisms, and of some environmental lobbying groups, notably the Environmental Defense Fund, to experiment with new and potentially more effective means for achieving environmental goals broke the stalemate and allowed a legislative majority to coalesce around a proposal that would reduce aggregate SO_2 emissions significantly and achieve the disparate goals that motivated various actors in the political process.

Institutional Location and Methodology

Unless otherwise noted, this paper is based on the continuing ex post evaluation of the U.S. SO_2 cap-and-trade program that faculty and students associated with the Center for Energy and Environmental Policy Research (CEEPR) at the Massachusetts Institute of Technology (MIT) have conducted since 1995. This effort was initially funded by the National Acid Precipitation Assessment Program (NAPAP) to support the 1996 Quadrennial Report to the U.S. Congress and the research has received continued funding through grants from the U. S. Environmental Protection Agency and from the underlying financial support provided to CEEPR by a number of corporate sponsors. This evaluation has been a major focus of CEEPR's research program, which aims to inform the public policy process by providing the results of objective, theoretically sound, and empirically rigorous research through publications and less formal presentations to interested audiences.

The results of the first years of this research are presented comprehensively in *Markets for Clean Air: The U.S. Acid Rain Program* (Ellerman et al., 2000), which is cited by Smith (2001) as an example for conducting ex post evaluations. This paper updates *Markets for Clean Air*, and it incorporates more of the work of other researchers who have since published on various aspects of the program.

In specifying the requirements of an ex post evaluation, Smith (2001) seconded the reinforced the admonition of Frondel and Schmidt (2001) that “the essential task of any evaluation analysis is the construction of a credible counterfactual situation—a precise statement of what economic agents would have done in the absence of the policy intervention.” With this in mind, the rest of this section describes the counterfactuals used in evaluating the SO_2 emissions trading program.

Two counterfactuals are involved in assessing any emissions trading program: one to assess the amount and cost of the emission reduction and the other to assess the cost savings and other effects of trading. The counterfactual for assessing the emission reduction requires assumptions about basic economic drivers, such as the demand for electricity and the relative price of fuels, and about other environmental regulations that may limit emissions. These factors can be observed and used in formulating this first counterfactual. In the case of the SO_2 program, the observed utilization of individual units provides a reasonably close estimate of the effect of the basic economic drivers in any given year. The effect of the pre-existing regulatory regime can be captured in the emission rate observed shortly before the start of the cap-and-trade program. Accordingly, the counterfactual used in this paper, as in previous work by the author and colleagues, is based on the heat input observed at affected units in each year and an unchanging pre-Title IV emission rate at those units.

This counterfactual assumption has the effect of making the estimated emission reduction equal to the heat-input-weighted changes in observed emission rates at affected units and to assume that no emission reduction can be attributed to changes in demand, either at individual units or in the aggregate. Since the demand for electricity is price inelastic, the cost of SO_2 controls is relatively small on a kilowatt-hour basis, and the major element determining the dispatch, or utilization, of individual generating plants is the cost of fuel, the error arising from assuming no effect on demand is probably small. Nevertheless, to the extent that the added costs from the program reduce the demand for electricity or change the order of dispatch of generating units in meeting that demand, the effect of the program is under-estimated.[4] A more likely source of error arises from the assumption about the counterfactual emission rate. To the extent that other environmental regulations, or changes in relative fuel prices, cause the emission rate at affected units to fall during the period of evaluation, the effect of the SO_2 program is over-estimated. Increases in the true counterfactual emission rates would have the opposite effect, but the scope for these is limited since all units face emission rate limits under the pre-existing command and control regulation and those limits are rarely, if ever, increased.

The other counterfactual, that used to assess trading, is much harder to specify. This other counterfactual requires a hypothetical, equally effective, alternative program without emissions trading. Estimates of cost savings are necessarily more subjective since they depend directly on the degree of inefficiency assumed in the imagined alternative regime. In this paper, a source-specific, quantity limit equal to the allowance allocation to specific units is used. This assumption conforms with the well-established propensity to source-specific limits (although rarely on total emissions from an individual plant), but it is relatively benign in not having a technology mandate similar to that characterizing much of the existing regulatory structure and to that contained in earlier, failed legislative proposals.

Economic efficiency

Two aspects of economic efficiency need to be distinguished in evaluating cap-and-trade programs. The first concerns trading among firms subject to the cap and the extent to which they realize the full cost savings attainable through emissions trading. The second aspect of economic efficiency concerns the broader welfare effects from the tax and regulatory interactions resulting from the treatment of abatement costs and the scarcity rents generated by the environmental constraint. From the standpoint of this second aspect, it has been argued that Title IV did not achieve full economic efficiency because first, allowances were not auctioned and the proceeds used to reduce

---

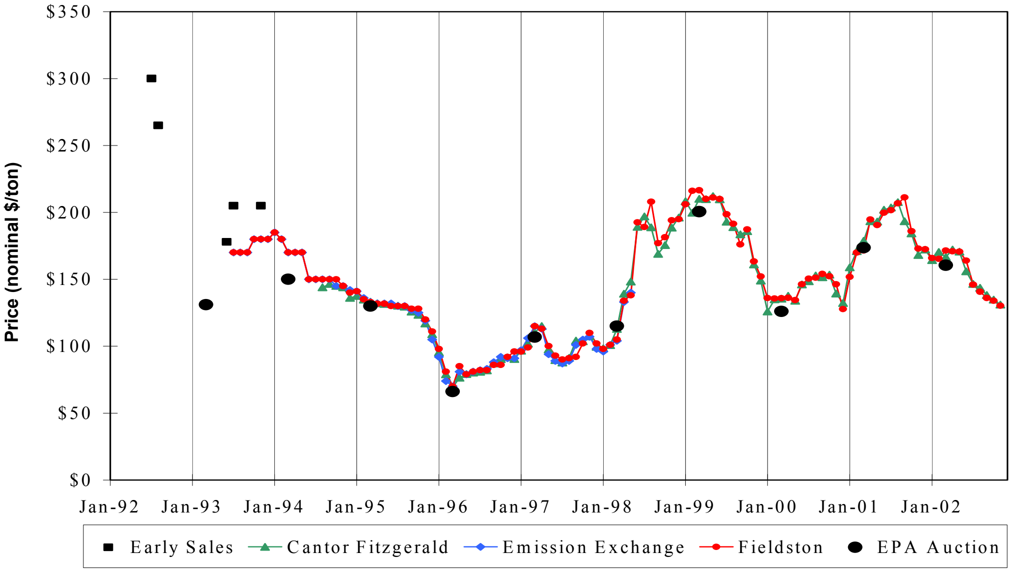
[4] The appendix to *Markets for Clean Air* contains an econometric estimation of the extent to which Title IV requirements changed the dispatch of generating units during Phase I. In brief, the demand placed on unscrubbed units subject to Title IV was shifted to affected, scrubbed units and to non-affected, Phase II units. Both effects are relatively small and the latter did not increase emissions perceptibly since the emission rates for unscrubbed units under the cap in Phase I were generally higher than the emissions rates for non-affected units, all of which were exempt from Phase I because of a lower emission rate.

distortionary taxes on labor and capital, and second, the average cost rules applying to units remaining under public utility cost-of-service regulation prevent the full marginal cost of abatement from being passed on to customers in the price of electricity (Goulder *et al*., 1997). A full discussion of this aspect of the economic efficiency of Title IV would involve consideration of the practical likelihood of economically efficient recycling, of equitable concerns, and how public utility regulation is applied in practice: all topics that are beyond the scope of this paper. Henceforth, all references to economic efficiency in this paper refer to the conventional use in emissions trading, that is, to the cost savings resulting from the flexibility provided by emissions trading without regard to the larger welfare issues reflecting allocative inefficiencies that may result from the existing regulatory and tax system.

The primary evidence for the economic efficiency of the SO_2 cap-and-trade system lies in the early emergence of an allowance market and the significant amount of trading that has occurred since before the program started. Figure One depicts the movement of allowance prices from the earliest observations through late 2002 as reported monthly by various brokers and in the annual EPA auction.

Figure 1: SO_2 Allowance Prices, 1992-2002

(Spot, Current Vintage Prices)

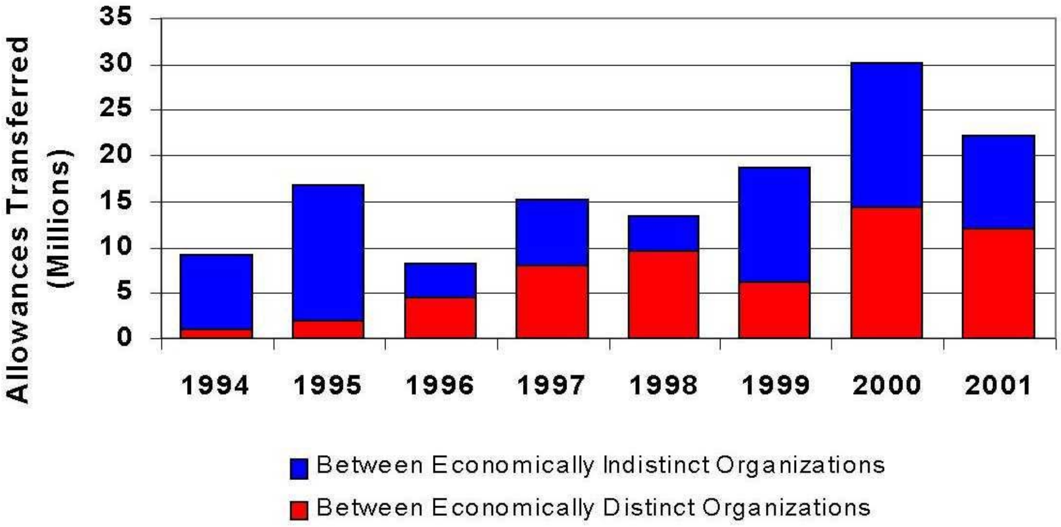


Source: Compiled by the author from monthly broker reports and the annual EPA auction reports

Prices have varied substantially over time—from an all-time low of $65 in early 1996 to highs slightly above $200 in 1999 and again in 2001—but at any one moment in time a single price prevails. The earliest reported trades took place at widely disparate prices, which were higher than the clearing price in the first EPA auction, held in March 1993. At this time, it would be hard to say that a market existed; however, by mid-1994, approximately six months before Phase I entered into effect, a market seems to have formed and the law of one price has prevailed since then.

Since allowances are readily substitutable for abatement, this single price provides a common point of reference and a coordinating mechanism for all owners of affected sources in deciding whether to abate more or less at any one time and thereby to equalize the marginal cost of abatement. Moreover, the significant and increasing volume of trading between economically distinct organizations, as illustrated in Figure Two, suggests that utilities are taking advantage of the cost-saving opportunities provided by emissions trading.

Figure 2: Annual Allowance Trading Activity



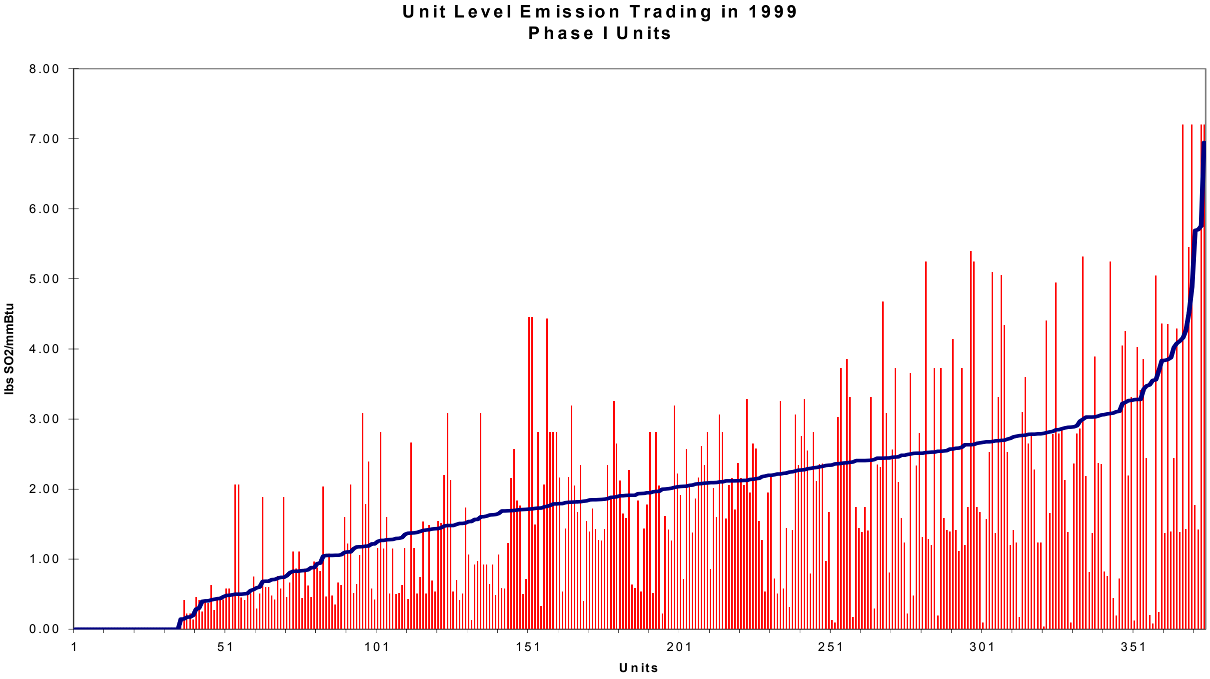
Source: US EPA

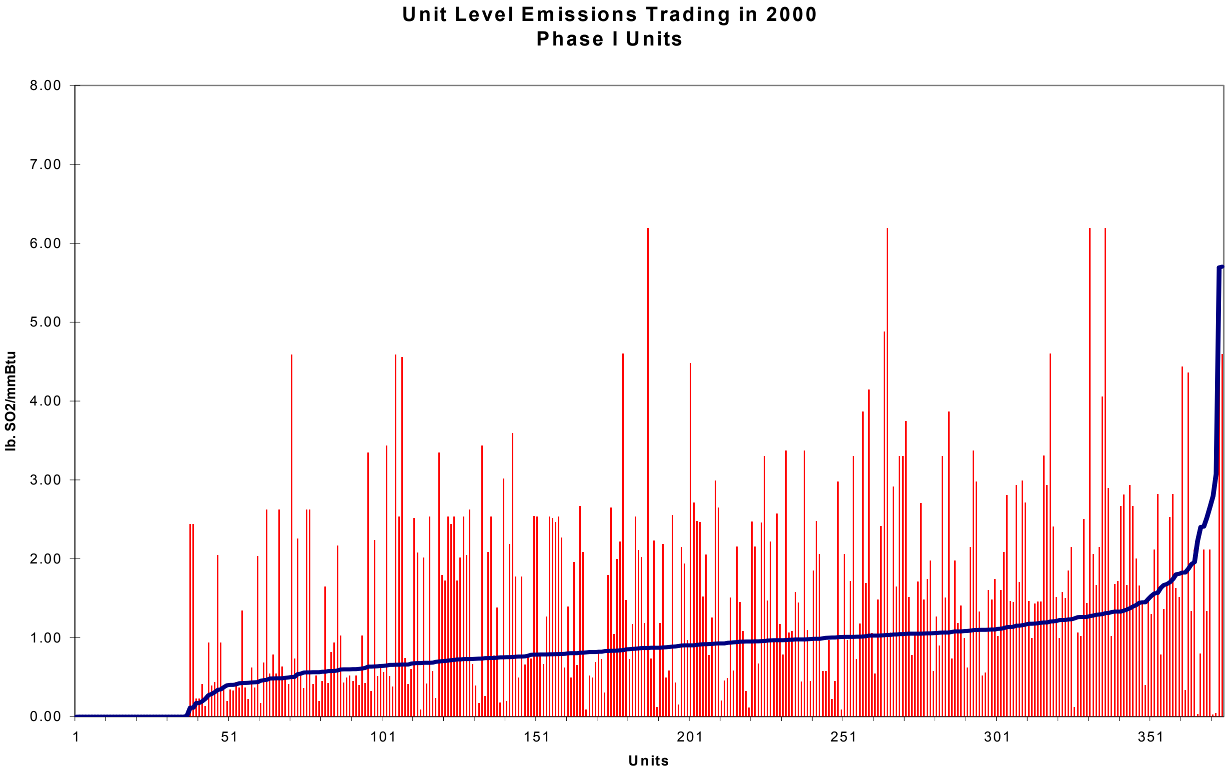
Since the equalization of marginal costs presumes a common price and trading among sources facing different costs, the preconditions for cost-effective abatement are being observed. An argument that the efficiency goals of the program are not being achieved would require an alternative hypothesis to explain the existence of a market and the observed volume of trade. In fact, no observer argues that observed trades are motivated by other than expected cost savings. As will be discussed later in this paper, the only disagreement among analysts concerning the economic efficiency of the SO_2 cap-and-trade program concerns the extent to which the full potential cost savings have been achieved.

Further evidence to support the argument for economic efficiency can be observed in the unit-level differences between allowances and emissions. The two panels of Figure Three show for Phase I affected units in 1999 and 2000 the emission rate that would be observed with no trading (the solid line) and the actual rate (the columns), given the heat input at each unit in these years. Few units are along the solid line, where they would have to be in the absence of trading, either when the allowance allocation is relative generous in 1999 or when the significantly reduced Phase II allocation went into effect. The average difference between observed emission rates and the no-trading rate is about 50% of the mean emission rate: 0.81 #SO_2/mmBtu over 1.64 #SO_2/mmBtu in 1999 and 0.86 #SO_2/mmBtu over1.48 #SO_2/mmBtu in 2000.

A further indication of economic efficiency is given by the relatively small change in average emission rates (-10%) when the allowed emission rate declined by 53%, from 1.85 #SO_2/mmBtu in 1999 to 0.87 #SO_2/mmBtu in 2000, when Phase II began. This smaller change in emission rates could occur only with banking; and in fact these 375 units went from banking 1.8 million allowances in 1999 to drawing the accumulated bank down by 1.5 million tons in 2000. This pattern of aggregate abatement over time is characteristic of an optimal banking program with certainty, in which firms take future required abatement and prices into account in formulating current abatement plans. In turn, this behavior implies that allowance prices rise at the interest rate and abatement increases gradually over the entire banking period. Such a pattern is observed in the transition from Phase I to Phase II among the units affected in both years. Moreover, despite all the stochastic variation in allowance prices since early 1994, as shown in Figure One, a definite upward trend can be observed.

Figure 3


Unit Level Emission Trading in 1999
Phase I Units
lbs SO2/mmBtu
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1
51
101
151
201
251
301
351
Units


Unit Level Emissions Trading in 2000
Phase I Units
lb. SO2/mmBtu
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1
51
101
151
201
251
301
351
Units

Recent research by Ellerman and Montero (2002) confirms that in the aggregate banking has been surprisingly optimal. The surprise resides in the general consensus, voiced in *Markets for Clean Air* as well as elsewhere, that too much banking had occurred in Phase I. The explanation of the surprise lies in the discount rate applicable to SO_2 allowances. The prices shown on Figure One allow a discount rate to be derived for SO_2 allowances by application of the capital asset pricing model to determine the amount of *undiversifiable* risk associated with holding SO_2 allowances. This risk is expressed by the correlation of returns from holding allowances (i.e., the monthly change in allowance prices) with returns from a well-diversified portfolio of equities over the same period of time. This correlation is zero, which makes SO_2 allowances zero-beta assets that should be discounted at the risk-free rate for comparable holding periods.

Compliance Costs and Savings from Emissions Trading

While the emergence of an SO_2 allowance market and the concomitant growth in the volume of SO_2 allowance trading suggests strongly that cost savings are being realized, these data alone provide no estimates of the magnitude of the cost savings, nor of the relation of these savings to actual or avoided, command-and-control compliance costs. In the case of the Acid Rain Program, many assertions have been made about the cost savings, but only two rigorous ex-post evaluations of compliance cost have been made [Carlson et al., 2000; Ellerman et al, 2000; hereafter, CBCP (for the initials of the authors) and MCA (for *Markets for Clean Air*)]. These two studies agree in finding the more extreme claims of cost savings unfounded, and their estimates of actual compliance costs are approximately the same, but they differ concerning the extent of the cost savings in the early years, as well as in methodology.

Ex Post Estimates of Compliance Cost

In reviewing the debates about the cost savings from Title IV, two distinctly different definitions must be kept in mind: one, loosely defined but more repeated; the other, more rigorous but less frequently cited. The former defines the cost savings as the difference of actual observed costs from predicted costs. The difference is loosely attributed to emissions trading even though other factors can and did intervene to cause actual costs to be lower. The second definition, used by the two studies cited above, relies upon a more rigorously defined no-trading alternative that incorporates identifiable cost-reducing exogenous factors. Accordingly, the following discussion will discuss first the findings of the two studies on actual compliance cost, then compare them with earlier estimates, and finally address the differences between the two studies concerning the magnitude of the cost savings.

CBCP and MCA agree roughly on the cost of compliance in the early years of the Acid Rain Program. The latter estimates the cost of compliance at $726 million in 1995 and about $750 million in 1996, while the former places the cost at $832 million in 1995 and $910 million in 1996, all stated in 1995 dollars. These estimates are not as far apart as they would seem. Complete comparability is not possible because of differences in methodology; however, both treat scrubber expense in the same manner.[5] Although they largely agree on the fixed cost of scrubbers ($375 million in MCA and $382 million in CBCP), they differ significantly on the variable costs associated with scrubbers ($89 million and $274 million, respectively).[6] CBCP uses scrubber data that reflect pre-1995 estimates of the variable cost of scrubbing, but the actual performance of the Phase I scrubbers has been much better than predicted, as will be discussed more fully in the section of this paper concerning dynamic aspects. Correction of this item alone largely removes the disparity in cost estimates between these two ex post evaluations. As an approximate figure, $750 million is probably a good estimate of the annual cost of abatement in the first years of Phase I.

---

[5] MCA provides a bottom-up, plant-by-plant analysis based on reported capital costs and observed sulfur premia. CBCP conducts an econometric estimation of a translog cost function and share equations of unit-level data for 734 *non-scrubbed* units over the 1985-94 period and then takes the resulting parameter values to form marginal abatement cost functions for individual units, which are then used to estimate actual costs based on observed 1995-96 emission levels. Scrubbed units are handled separately on a cost accounting basis using identical cost of capital and depreciation assumptions as in Ellerman *et al.* (2000). It should be noted that the estimation of 1995-96 cost in CBCP is almost an aside to the main purpose of the article which is to explain the reduction in abatement cost from pre-1995 estimates and to provide updated estimates of the cost of compliance in 2010.

[6] The numbers cited from CBCP are from their break-out of the costs of 2010 compliance. This estimate will be approximately the same as the scrubber costs in 1995-96 since the fixed costs are annualized over 20 years, fuel costs are assumed not to change after 1995, the number of scrubbers remains unchanged, and costs are stated in 1995 dollars.

Comparison with Ex Ante Estimates of Cost

The important difference, however, is not the minor one between CBCP and MCA concerning actual costs in 1995-96, but the larger one between these two careful ex post estimates and ex ante estimates of the same Phase I cost, as well as of predicted costs in Phase II. Most of the disparity between ex ante and ex post estimates reflects very different assumptions about the nature of proposed acid rain controls, the demand for electricity, and the relative availability and cost of low sulfur coal. For instance, the total annual costs associated with some of the early proposals to control acid rain precursor emissions were estimated at amounts ranging from \$3.5 to \$7.5 billion. Although the details of these earlier proposals varied, they generally mandated scrubbers at a significant number of units and allowed very limited emissions trading. Once the proposal that ultimately became Title IV was proposed (in 1989) and enacted (in 1990), the ex ante cost estimates for the fully phased-in program with trading fell to a range from \$2.3 billion to \$6.0 billion, with most of this variation reflecting varying assumptions about the extent to which emissions trading would be used.[7] The now current estimates for compliance costs in 2010, as provided by CBCP and MCA, are significantly lower still, \$1.0 billion and \$1.4 billion, respectively, for what is the same program but updated to reflect more current market conditions.

CBCP provides a very helpful quantification of the causes of the change between the early estimates of Title IV and the current estimates. In examining the changes over the period of their panel regression, 1985-94, they find that the marginal cost of abatement for a representative unit reduction has been approximately halved and that 80% of the reduction in cost is attributable to falling price of low-sulfur coal relative to the price of high sulfur coal and that the remaining 20% is attributable to technological change. The change in the relative price of low sulfur coal is discussed in more detail in Ellerman and Montero (1998), who attribute the change to reduced rail rates, made

---

[7] MCA includes (pp. 231-235) a discussion of the few ex ante estimates of Phase I costs and compares them with the MCA estimate of actual cost. Most of the variation in these estimates, made only a few years before Phase I began, reflects differing assumptions about the extent to which utilities made full use of the flexibility afforded by emissions trading. When compared on an average cost basis to account for differences in assumptions about the quantity of abatement, the MCA estimate of actual cost in 1995 was slightly above (3-15%) ex ante estimates assuming full use of emissions trading and 20-35% below estimates that assumed relatively little use of emissions trading.

possible by rail deregulation, for transporting distant, but cheap western coal to mid-western markets where local, high-sulfur coal had predominated. They estimate that the switching of mid-western high sulfur coal units, most of whom were mandated to be subject to Title IV in Phase I, to lower sulfur western coal because the latter had become cheaper reduced the amount of abatement required to meet the Phase I cap by about 1.7 million tons, or by about half of that predicted by early estimates of required abatement.

Table 1 provides CBCP's quantification of the effects of these exogenous changes on estimates of compliance costs for a fully phased-in Title IV program.

Table 1: Total Cost of Compliance with Title IV in 2010
(billion 1995 dollars)

Cost Assumptions	Command-and-Control	Efficient Trading
1989 Prices and Technology	\$2.67	\$1.90
1995 Prices and Technology	\$2.23	\$1.51
1995 Prices and 2010 Technology	\$1.82	\$1.04

Source: Carlson et al. (2000), Table 2, p. 1313

The changes in relative fuel prices and technology between 1989 and 1995 lowered costs by about 20% and CBCP's preferred estimate for 2010, which maintains 1995 relative fuel prices but extrapolates the 1985-94 rate of technological progress to 2010, reduces predicted costs by another third. The assumption of continued technological change also explains the difference between the CBCP and MCA estimates of Phase II annual cost, since the latter does not make any allowance for this factor.

To summarize, most of the explanation for the lower than expected cost of Title IV is attributable to changes in the nature of the proposed controls, from prescribing technology to the flexibility of a cap-and-trade system, and to changes in related sectors of the economy that were reducing SO_2 emissions anyway. As can be seen by comparing cells in Table 1, the difference in total cost between a relatively benign command-and-control alternative and fully efficient trading accounts for a relatively small part of the difference from the earliest cost estimates, which remained for better or worse stuck in many observers' mind. Moreover, the impression of dramatically lower costs was reinforced by the price of SO_2 allowances, which has been the most visible manifestation of cost to most observers. No one predicted the allowance prices of $100 and even less that occurred in late 1995 and for most of 1996. Most predictions of early Phase I allowance prices ranged between $250 and $400, prices that have yet to be realized. Furthermore, many casual observers remembered only the predictions of Phase II prices, usually after the bank had been drawn down, which ranged from $500 to as much as $1000. The very low, early 1996 allowance prices may have reflected an over-reaction to the correction in early expectations of market conditions; but, with eight years of experience with SO_2 allowance trading, there seems little doubt now that changes in technology and the availability of low sulfur coal fundamentally changed the quantity and cost of abatement that would be required to comply with Title IV and shifted allowance prices commensurately lower.

The Extent of Cost Savings from Trading

The principal area of disagreement among analysts about the economic efficiency of the program concerns whether the full cost savings potential of emissions trading is being achieved. The point in dispute concerns the effect of cost-of-service regulation on the incentives of electric utilities to engage in trading with each other. The argument takes two forms: first, that conventional cost-of-service regulation provides no incentives to trade in the external market, since the gains would be passed on to rate-payers and losses might not be recoverable; and second, that public utility commissions have adopted policies that encourage sub-optimal choices by individual utilities, such as to scrub local high-sulfur coal in order to protect in-state jobs (Bohi and Burtraw, 1997; Rose, 1995; Rose, 2000). Research that simulates the effect of several of these disincentives suggests that compliance costs might be as much as doubled (Fullerton *et al*., 1997; Winebrake *et al*., 1995).

Empirical research tending to confirm this effect has been published. The most striking result was that in CBCP which found that the actual cost of compliance with Title IV in 1995 and 1996 was slightly higher than the cost of compliance under a benign command-and-control alternative (quantity caps equal to allowances at each affected unit). Moreover, their estimate of total cost with fully efficient trading was some $200-

$250 million lower. This finding indicated that the unrealized cost savings were substantial and implied that emissions trading had not resulted in any cost savings in the first two years of the program. The authors were quick to note that the volume of emissions trading was increasing and to state that they did not expect the apparent forsaking of the gains from emissions trading to last. More recently, Arimura (2002) has published research supporting the view that public utility commission regulation influenced abatement choices and contributed to low allowance prices.

The contrasting point of view is associated with researchers at MIT and is stated most completely in MCA, although also published in earlier articles and working papers (Joskow *et al.*,1998; Schmalensee *et al.*,1998; Ellerman and Montero, 1998; and Bailey, 1996). Here, the findings are that a reasonably efficient allowance market emerged as early as mid-1994; trading volumes have increased significantly, even in the early years; the effect of state PUC rulings on trading activity is insignificant; and that cost savings have been realized.

Much of the contrast between these two interpretations is a matter of tone, although substantive differences exist concerning the effect of PUC regulation on emissions trading. It is beyond the scope of this paper to explore these differences in any detail, but a reader not already familiar with this debate should keep several points in mind.

First, the argument on cost savings is as much one of whether the glass is half full or half empty. The MIT group makes no estimate of what the full cost savings might be and allows that some cost savings are undoubtedly unrealized, but they emphasize that cost savings have been realized and that no market is perfect. The MCA estimate of the cost savings in the early years of Phase I ($350 million, about half the observed cost of compliance) is derived from observed data assuming that the data reflect nearly efficient choices by abaters. In other words, this particular estimate assumes away the problem insisted on by the other school. This particular estimate was developed to discourage the then current views that the cost savings from emissions trading under Title IV were much greater. With the exception of the CBCP finding, the other camp does not dispute the existence of cost savings from Title IV. For instance, Bohi and Burtraw (1997) refer to the “puzzle” of cost savings with limited trading and Rose (2000) concludes that Title IV shows that “trading mechanisms appear to be robust enough to allow substantial savings…to occur even when faced with less than ideal conditions.” The problem with the accuracy of the scrubber costs in the CBCP finding has already been mentioned, but even setting this aside, the focus in CBCP is more on quantifying the extent of unrealized cost savings as it is insisting that their less costly CAC alternative is realistic.[8] Thus, one camp tends to emphasize the short-fall, while the other stresses the achievement. Still, a difference remains concerning magnitude. The difference is perhaps more aptly whether the glass is nearly full or only half full.

A second point to be kept in mind is that the debate about regulatory influence is at bottom one about how public utility regulation works in practice. Although not so far publicly stated, the MIT group would not dispute the theoretical effect of the alleged influences; their contention would be that the theory of regulation applied is over-simplified and not representative of the performance-based, rate-making as practiced in the 1990s. The only direct empirical test of the hypothesis of significant regulatory influence on emissions trading is Arimura (2002), which is unsatisfactory in attributing a difference found between the abatement decisions at Phase I units owned by the Tennessee Valley Authority, a publicly owned utility, and those owned by PUC-regulated utilities to test a hypothesis concerning differences between profit-maximizing firms and regulated electric utilities.

Environmental effectiveness

The arguments in favor of emissions trading programs always assume that trading will not jeopardize environmental effectiveness, and this is invariably the main concern of environmental groups and those who tend to be skeptical of emissions trading. The experience with Title IV has provided no grounds for concern about environmental effectiveness; in fact, the experience suggests that environmental performance may be better than that experienced with command-and-control analogues. This section of the

---

[8] Still, their CAC counterfactual is identical to the one assumed in MCA, which is found to cost about 50% more than the observed cost of compliance. Also, the methodology adopted by CBCP would attribute the same change in scrubber cost to the CAC alternative so that the finding of no cost savings would still hold.

paper addresses this point, adduces the evidence indicating greater environmental effectiveness, and provides some tentative explanations for this result.

An important first issue in evaluating environmental effectiveness is identifying the appropriate metric. The acid rain motivation of this program would suggest that an appropriate one would be the amount of wet deposition, or even the acidity of lakes and forests in sensitive regions; however, the most obvious and easily measured metric, total emissions, is the one typically used.

No doubt surrounds the issue of whether SO_2 emissions have been reduced.[9] The two panels of Figure Four show actual emissions, the caps, and an estimate of counterfactual emissions for the 375 units first subject to Title IV in 1995 and for the much larger cohort of units that have been subject to Title IV since 2000. For both the Phase I and Phase II cohorts of units, the largest annual emission reduction is made in the first year, when the affected units first incur a cost for every ton of emissions. Given the phased-in nature of the requirement facing the Phase I units and the ability to bank, the annual reduction by these units was much greater than required. The annual reduction of emissions in 1995 was 3.9 million tons and that quantity of abatement has increased steadily and now stands at 6.3 million tons in 2001. Banking implies that emissions in the first years of Phase II will be greater than the allowances issued for these years, but the appropriate metric is the cumulative reduction since 1995, which has been 33.7 million tons, about 29% more than the 26.1 million tons that would have been required as of 2001 without banking. By the end of Phase I, the actual cumulative reduction was twice what was required, and that ratio will now decline steadily to 1.0 when the accumulated Phase I bank will be exhausted, probably in the second half of this decade.

---

[9] Suggestions to the contrary, such as those contained in *Darkening Skies*, a publication of the New York Public Interest Research Group, are misleading in citing specific plants and comparing 1999 emissions with 1995 emissions.

Figure 4a. Phase I Unit Emissions, Caps, and Counterfactuals

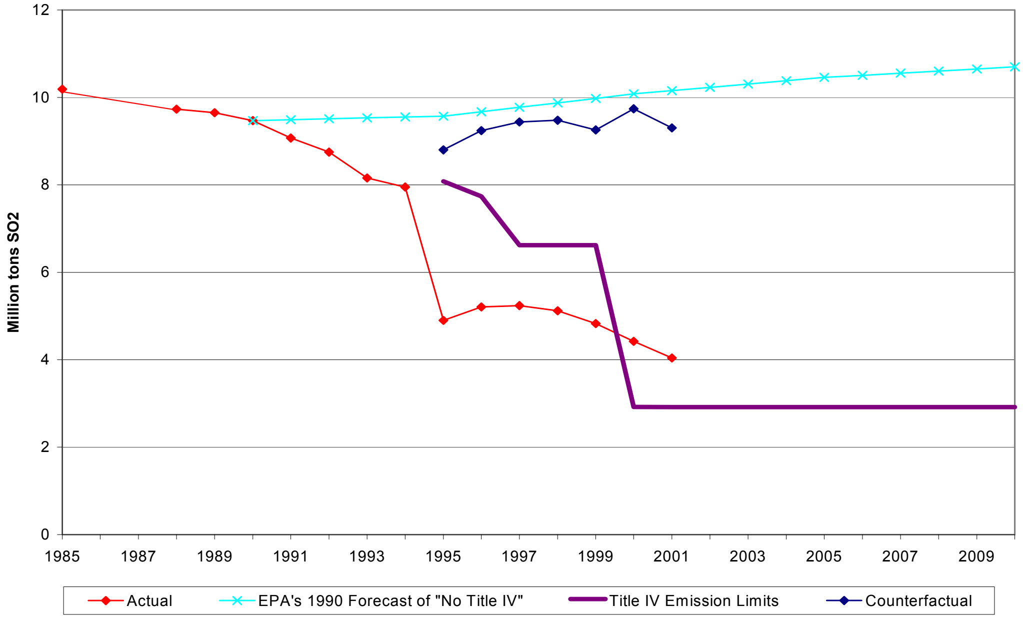
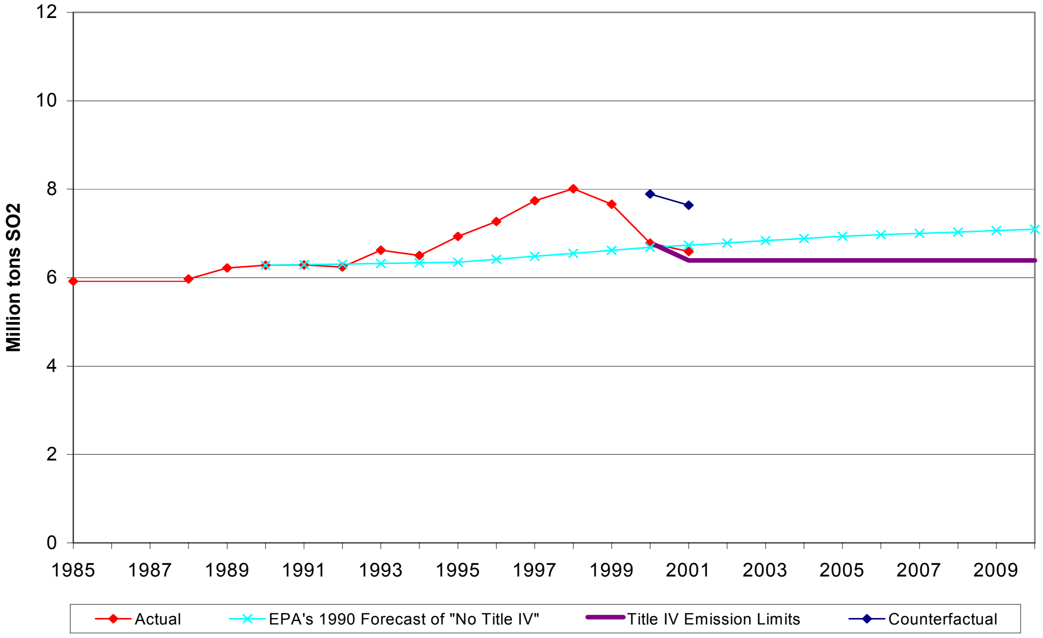
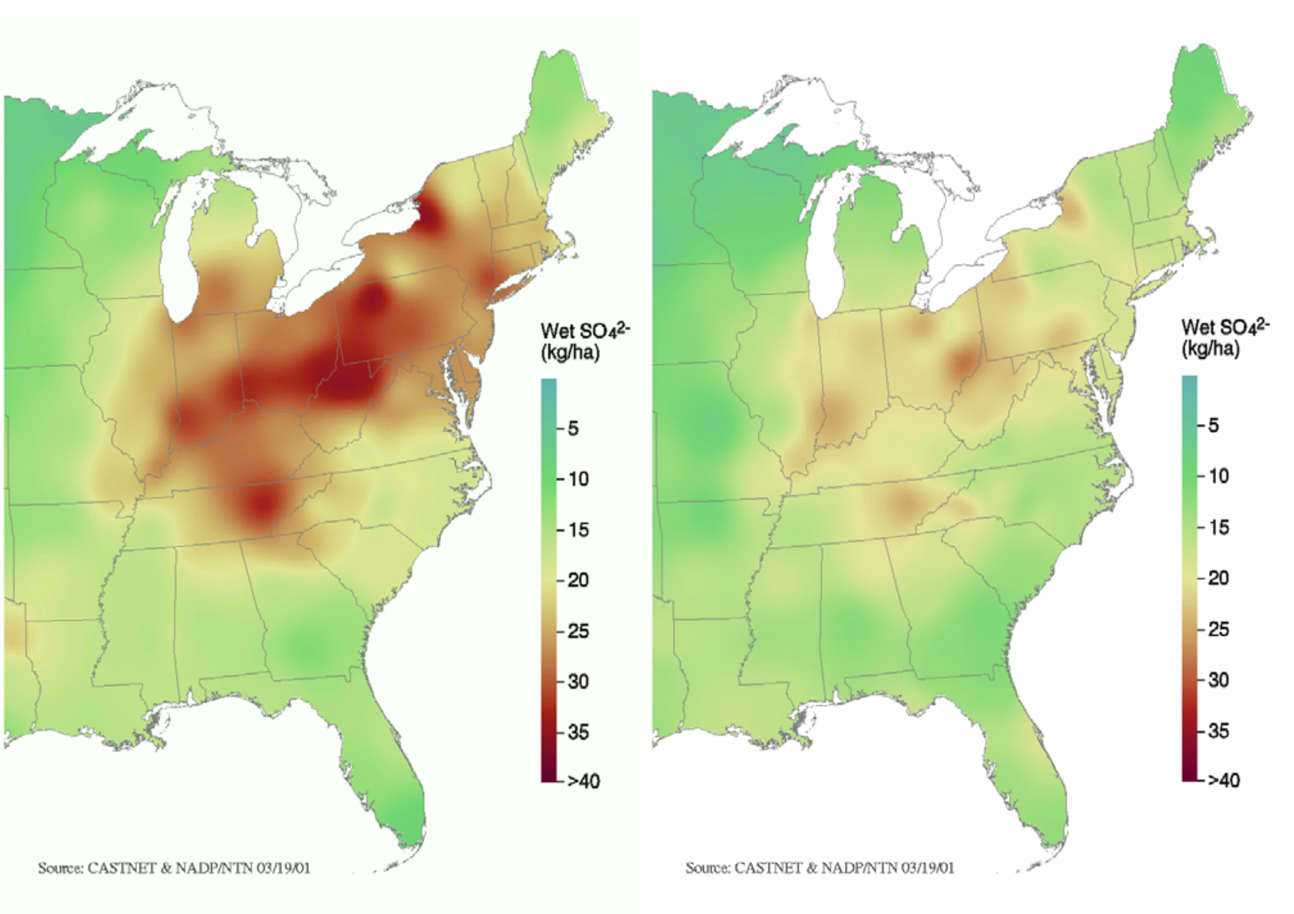


Figure 4b. Phase 2 Unit Emissions, Caps, and Counterfactuals



The significant and accelerated reduction of emissions implies that the deposition of acidic particles has also fallen. The latest progress report from the U.S. Environmental Protection Agency (USEPA, 2002) reports that all of the conventional indicators relative to SO_2 have declined markedly because of the Acid Rain Program. Figure 5 provides a graphic illustration of the change in wet sulfate deposition in the eastern U.S. between the late 1980s and the late 1990s.

Figure 5: Monitored Reduction in Wet Sulfate Deposition



Source: CASTNET & NADP/NTN 03/19/01

Source: CASTNET & NADP/NTN 03/19/01

Similar diagrams could be shown for ambient concentrations of SO_2 and sulfate concentrations in the atmosphere, both of which have fallen generally across the Northeast and mid-Atlantic regions and in some places by as much as 50%. Sulfate concentrations in lakes and streams have declined significantly in all monitored regions of the Eastern United States, except Virginia, and in some areas, notably Pennsylvania and the Adirondacks, the acid neutralizing capacity of the soil has begun to increase, which is an indication of the beginning of recovery in ecosystems suffering from acidification.

Another aspect of the environmental effectiveness of the Acid Rain Program is the extent of compliance. With the exception of a few very small, new gas units in 2000, all generating units have been in compliance with Title IV requirements in all years. This record of virtually 100% compliance is not encountered with command-and-control regulation under which sources not infrequently receive various forms of dispensation that have the effect of delaying and sometimes permanently relaxing the applicability of the standard. The reason is that a single standard imposes greater costs of some than on others because of differing site-specific considerations and these firms pleading unique hardship petition for administration relief that is often granted. Although such relief is may be justified in the interest of equity, compensating tighter standards are not imposed on firms facing relatively less onerous costs and these latter never step forward to assume a greater cost burden in the interest of equity, nor are regulators able to identify who they are and thus to impose compensating, more stringent standards on them. The information asymmetries between regulator and regulated in CAC systems effectively lead to a form of adverse selection that makes the standard less effective than it otherwise would be.

This problem is avoided in a cap-and-trade system for two reasons. First, the market removes the rationale of unique hardship since the greatest burden borne by any is the price of an allowance; and, in a market with many buyers, no single one can claim to be uniquely disadvantaged. Second, the market provides at once a cheaper means of relief and the offset that preserves environmental integrity. Nothing prevents a firm from petitioning for relief from the requirement to surrender allowances, even if the grounds for doing so are weak; however, doing so can be costly and a market makes it cheaper simply to pay another to make the compensating reduction. In a sense, the ability to trade, and the market that it implies, renders special pleading uneconomic.

A frequently voiced worry about the environmental effectiveness of emissions trading programs concerns “hot spots.” This phrase refers to the potential in a trading system for emission reductions to be transferred away from areas where emissions cause greater damage to those where the emissions cause less damage. Well-designed programs would not have this problem since emissions would not be traded unless they had equal environmental effect; however, real programs contain unavoidable compromises and the SO_2 program is no exception. The enabling myth of the acid rain program is that location does not count, when in fact from the standpoint of acid rain effects, location obviously does. The fear in the acid rain program is that emissions in the Midwest would not be reduced if utilities in this region could pay others located in parts of the country with little impact on the Northeast to reduce on their behalf.

This fear has proved to be unfounded (Swift, 2000). Sources in the Midwest have provided about 80% of the emissions reduction achieved in Title IV while accounting for about 55% of emissions in 2000. It may be argued that emissions from the Midwest are still too high, but it can hardly be argued that emissions trading has allowed sources in the Midwest to avoid abating. A tendency to autarkic compliance in initial planning and a program incentive to scrub early also encouraged reductions in this region, but the more important reason appears to be that the cheapest abatement is to be found where the largest sources are located.

This happy result is not accidental. Most deep abatement technology, like scrubbing, is capital intensive and the per-ton cost depends how many tons are removed per MW^e of capacity. Higher utilization and higher sulfur content of the coal being burned means more tons of abatement over which the fixed capital cost can be spread and lower total cost per ton. Thus, where capital-intensive, deep-abatement technology is an option, market systems will direct abatement to relatively larger and more heavily utilized sources with relatively high sulfur coal. And, if these sources are the most damaging from an environmental standpoint, the experience with Title IV suggests they will be cleaned up first and that hot spots will not appear.

Voluntary Aspects of Title IV[10]

Title IV had several provisions that allowed sources of SO_2 emissions outside of the cap to opt-in to the program. Such features are attractive as a further means of

---

[10] The discussion of this section is based largely on the work of Juan-Pablo Montero (Montero, 1999, and Montero, 2000), which is summarized in chapter seven of *Markets for Clean Air*.

lowering program costs if sources that are excluded from the cap are able to provide cheaper abatement. In the case of Title IV, certain utility sources that were not required to be under the cap until Phase II could opt-in to Phase I, and non-utility sources that were otherwise not a part of the program could do so in either phase.[11] The response of these two groups was very different: many eligible utility sources opted-in, while few industrial sources did so. The response of the utility sources also revealed an unavoidable trade-off between the economic and environmental objectives of the basic program.

The theory underlying voluntary features is obvious enough: if the aggregate cap is set optimally and non-capped sources can reduce emissions at lower marginal cost than the price of allowances traded among capped sources, then costs are reduced without harm to the environmental objective by allowing non-capped sources with lower marginal costs to opt-in. In Title IV, sources opting-in received allowances equal, in theory, to what emissions would have been without participation and were then held to the same compliance requirements as capped sources.[12] The manner of opting-in implied both that emissions were monitored and that the opt-in unit's counterfactual emissions would be accurately determined. This meant that continuous emissions monitoring systems, or an equivalent system, would have to be in place and that counterfactual baselines would have to be established. The differing responses and the revealed trade-off can be traced back to these two problems of implementation.

Over 200 electric utility units opted-in for one or more years of Phase I, and 110 of them participated in all five years. In contrast, only a few industrial sources chose to opt-in to the program. The different response is largely explained by the differences in transaction costs for each category of participant (Atkeson, 1997). Industrial sources that considered participation but decided not to do so cited the costs of monitoring as the largest consideration. Moreover, the few that did participate already had monitoring

---

[11] The legislative and regulatory provisions for industrial units are known as the Industrial Opt-in Program and utility units fell under the substitution or compensation provisions; however, all are referred to here as voluntary or opt-in participants. Electric utility units eligible for opting-in to Phase I were those owned by utilities with other units mandated to be part of Phase I. Provisions were included for opting-in units owned by a utility without Phase I units through contract with a utility having Phase I units, but these contract provisions were little utilized.

[12] Note that this mode of voluntary participation is different from many instances in which tradable credits are issued only for the emissions avoided. Thus, most of the allowances issued to opt-in units were needed to cover emissions from these units.

equipment in place as a result of other environmental requirements or otherwise did not need to install monitors.[13] This obstacle was not faced by eligible electric utility units because all sources subject to the Acid Rain Program were required to install a continuous emission monitoring system by 1995 regardless of whether the unit was required to participate in Phase I beginning in 1995 or in Phase II beginning in 2000. Also, the utilities owning the units eligible for becoming substitution and compensation units in Phase I were already incurring the overhead costs of managing emissions and accounting for allowances. Finally, electric utility units did not need to establish a baseline. The number of allowances that would be granted to eligible electric utility units was pre-determined by a set of mathematical formulae that were similar to those used for units required to participate in Phase I. As a result of all these factors, the additional costs of participation were very low for eligible electric utility units and a significant number of them volunteered.[14] In contrast, industrial sources would have had to incur the costs of monitoring emissions in addition to those of establishing a baseline and keeping track of allowances and emissions. These transaction costs were greater than the potential gains from trading that would have been possible through voluntary participation.

While the voluntary participation in the Acid Rain Program was heartening, an analysis of which eligible units opted in and which did not reveal a strong element of adverse selection, which resulted from the impossibility of specifying a true contemporaneous baseline (Montero, 1999). The pre-specified baseline, which greatly reduced transaction costs, relied mostly on 1989-90 data; however, changes in coal markets and in the utilization of electric generating units in the intervening years caused the true counterfactual emissions for eligible units in 1995-99 to be different. Thus, units that had already switched to lower sulfur coal for purely economic reasons because of changes in coal markets tended to opt-in and to receive some allowances in excess of

---

[13] For instance, in one case, an electric utility subject to the program undertook to provide steam and power to an industrial facility thereby allowing that facility to shut down the boilers it had previously used to generate electricity and steam. Allowances equal to what the closed down facilities would have produced in supplying the ongoing needs of the industrial facility were then awarded to the electric utility providing the facility's power and steam needs.

[14] A further consideration was motivating electric utility participation was the NO_x grandfathering provision. Units with certain types of boilers could be grandfathered from Title IV's Phase II NO_x emission limits if they participated in the SO_2 program in 1995. While many did, these units generally did not receive excess allowances and were not part of the adverse selection problem that characterized most electric utility opt-in units.

what might be considered the true baseline. And those who might have had low cost abatement to offer but whose emissions had risen above the pre-specified baseline tended not to opt-in since they would incur the costs of reducing emissions to the baseline before they would receive any benefits from emissions trading. The end result was that the units opting in were not so much low cost abaters, although some may have been, as they were units that were abating anyway.

This problem of adverse selection was exacerbated by allowing the owners of eligible units to wait until November 30 of each year to decide whether to opt-in for that year and to take the unit out of the program in the following year if opting-in would be disadvantageous. While many eligible units remained in the program for the entire five years of Phase I, a number of units can be observed opting in and out according to whether emissions were higher or lower than the allowances they would receive by opting in.

While the evidence of this selection bias is very strong, the environmental effects from the loosening of the Title IV cap must be kept in perspective. The number of allowances that could be considered excess amounted to only 3% of the total issued during 1995-99 and the inflation of the cap during Phase II, when these allowances will be used is only about 2%. These magnitudes are not great and they cannot be said to have threatened the overall integrity of the SO_2 cap. In addition, many of the units opting in also abated emissions in response to allowance prices and thereby contributed some cost savings to the program. Whether these cost savings were greater than the reduced environmental benefit depends greatly on the assumption about the true but unobservable baseline. In summary, it is hard to avoid the conclusion that the environmental damage was not great, but neither was the economic benefit, and that on balance, the voluntary features of Title IV were not worth the extra administrative effort.[15]

---

[15] See Ellerman *et al*. (forthcoming) for an argument that this conclusion, which results from a balancing of costs and benefits, ought not to be carried over to potential applications of emissions trading for the control of greenhouse gases.

Dynamic effects

Theoretical work has long predicted that market-based instruments, such as a cap-and-trade program, would provide greater impetus to innovation than command-and-control regulation, and thus add another cost-reducing attribute to these instruments (Magat, 1978; Milliman and Prince, 1989). Title IV has provided the occasion for testing this theoretical prediction and there is plenty of anecdotal evidence of what could be interpreted as innovation. Nevertheless, there is only one study that has attempted to address this issue rigorously and its results provide some confirmation, but not much (Popp, 2001). It may be still too early to be able to test the hypothesis confidently; and, under the best of circumstances, the difficulty of disentangling the effects of the regulatory instrument from exogenous technological change is great. Accordingly, in ths section, the term, dynamic effects, is interpreted broadly to encompass factors other than the direct trading of emission rights that contribute to lower compliance cost.

In considering dynamic effects, it is natural to focus of flue gas desulfurization, or scrubbers, since they are capable of removing 95% or more of SO_2 emissions from the stack, they are commercially available and widely used, and they are costly. Moreover, the total costs of scrubbing for the Title IV scrubbers installed at the beginning of Phase I has been less than predicted and a second cohort of Title IV scrubbers that have come on line at the start of Phase II have shown even lower cost. The key components of this change in cost are given in Table 2.

Table 2: Evolution of Scrubber Costs

	Ex Ante Phase I	Ex Post Phase I	Phase II
Initial Capital Cost ($/KW^e)	$240	$249	$150
Tons SO_2 Removed per MW^e	99	137	137
Per ton Fixed Cost ($/ton)	$273	$206	$124
Fixed O & M Cost ($/ton)	$75	$15	$15
Variable O & M Cost ($/ton)	$116	$65	$65
Total Cost per ton ($/ton)	$464	$286	$204

Source: MCA, Table 9.3 at p. 236 and discussion on p. 240.

The costs of scrubbing can be broken down into three components: 1) the initial capital cost, conventionally expressed as dollars per kilowatt of capacity, 2) the tons of SO_2 removed per unit of capacity over some period, which depends on the sulfur content of the coal and the utilization of the scrubber, and 3) the O&M costs, which are often expressed as cents per kilowatt-hour but are more properly stated as dollars per ton removed. Ex ante estimates for the cost of scrubbing a retrofitted Phase I unit typically fell between \$400/ton and \$500/ton, but ex post average cost has been below \$300 a ton, well above allowance prices, but not as uneconomic as often assumed. And this average masks huge variation, from a few units with apparent costs higher than \$500/ton to several with costs around \$200/ton. As shown in Table 2, the calculated 33% reduction in average cost was due not to lower initial capital costs, which were as expected, but to 25% higher utilization of the retrofitted units and a halving of operating and maintenance costs from what had been predicted. Operating costs were lower mostly because of improved instrumentation and control, which reduced the parasitic loss of power and manpower requirements, and it is probable that this improvement was a reflection of broader changes in information technology that were occurring throughout the economy.

The more interesting change from the standpoint of the effects of Title IV was the increase in utilization from 65% of total hours to 85%. This shift in dispatch reflected the effects of the sulfur premium that appeared in coal markets across the entire sulfur gradient and which tended to be equal (when appropriately converted) to the price of allowances. Whereas the only coal receiving a sulfur premium prior to Title IV was "compliance" coal, that required in generating units meeting the pre-1978 NSPS by burning coal with less than 1.2 $\#SO_2$/mmBtu, a sulfur premium now extended across the entire range of sulfur content. This differentiation in the prices of coals having more than 1.2$\#SO_2$/mmBtu had other consequences that will be discussed below, but it had two effects that influenced the utilization of units with retrofitted scrubbers. Since the sulfur premium and allowance prices tend to equality and allowance prices were higher than the variable cost of scrubbing, a scrubbed unit would have lower marginal cost for generating electricity than an unscrubbed unit, if all else were equal. The second effect, and undoubtedly the more important one, reflected the change in fuel cost, the major component in the variable cost of generating electricity, due to the new sulfur premium.

Unscrubbed units, typically burning mid- to low-sulfur coals, found themselves facing not only higher marginal abatement costs, but also higher fuel costs relative to scrubbed units, which would typically burn the higher sulfur coals that were now cheaper relative to coals with lower sulfur content.[16] Thus, the lower cost of scrubbing observed in Phase I is not the result of new technology but of the new requirement that the cost of emitting sulfur dioxide to be incorporated into operating costs in a systematic way.[17]

After the first cohort of Phase I scrubbers, vendors touted a reduction in capital cost for follow-on scrubbers, and these claims became real in 1998 when allowance prices rose to $200 and scrubber retrofits were announced for eight additional units, which are now online. Many of these units came in with initial capital costs around \$100/KW^e (which implied total costs below $200/ton), but these units were able to achieve cost savings because of previously installed scrubbers at other units at the same generating plant. The total cost indicated for Phase II scrubbers provides a good estimate of the long-run marginal cost of SO_2 removal by scrubbing, but that cost will rise as the scrubbers are retrofitted to units that are less utilized and burning lower sulfur fuels (Ellerman and Joskow, forthcoming). Nevertheless, it is clear that there has been a large reduction in the cost of scrubbing, and the question is whether this can be attributed to Title IV.

The only research so far to address this question explicitly is Popp (2001) who compared patents relating to scrubbers from the early 1970s through 1997 with scrubber performance as reported in annual submissions to the Energy Information Administration. He finds that the passage of the 1990 Clean Air Act Amendments did not increase the level of innovative activity, and that in fact it fell somewhat, but that the nature of innovation did change in a more environmentally beneficial way. Throughout the period, the continuing level of innovative activity led to lower operating cost, but the patents granted after 1990 are associated with an improvement in removal efficiency that had remained constant previously. Popp's finding conflicts in part with those of two other

---

[16] This effect applies only to scrubbed units. Unscrubbed units burning higher sulfur coals would pay less for fuel but require more allowances and on balance enjoy no advantage over unscrubbed units burning lower sulfur coals.

[17] As exemplified by the compliance coal phenomenon, the costs of complying with environmental regulations often entered into marginal cost decisions, but it was not systematic as it became after the introduction of SO_2 allowances.

studies of changes in scrubber technology (Bellas, 1998; and Taylor *et al*., 2001). Bellas examined the same cost data as Popp but only through 1992 and found "no significant progress…in abatement technology," which he associated with "the small incentives for innovation [associated with] the form of regulation typically used in the U.S." Taylor *et al*. (2001) examine a slightly different question in seeking to determine the relative efficacy of R&D spending and regulatory constraint in inducing innovative activity related to scrubbers, and in doing so they find the same decline in patent activity as Popp but a continual increase in removal efficiency as well as a steady decline in capital cost, both of which are attributed to "learning by doing."[18] These interesting but conflicting results concerning the trend in scrubber costs do not provide very solid ground for attributing dynamic effects, as usually defined, to Title IV.

While scrubbing can be considered the backstop technology for SO_2 abatement, it is not the only way, and it accounts for relatively less (40%) of the total reduction in SO_2 emissions in Title IV than switching to lower sulfur coal. Cost reductions in switching are not as easy to document, since switching does not attract the same attention as installing a scrubber, but cost-reducing changes can be inferred, most of all in the ability of boilers built to fire bituminous Mid-western coals to accommodate lower sulfur, sub-bituminous coal from the West. It was always recognized that these units could be converted to the use of sub-bituminous coals, but the higher water and ash content of the latter would lead to a significant derating, or reduction, in the generating capacity of the unit. As a result, it was expected that the predominantly high-sulfur burning units in the Midwest would either install scrubbers or switch to low-sulfur bituminous coal produced in the Appalachian region. As the effects of rail deregulation increasingly reduced the significant transportation component in the cost of *western* low sulfur, sub-bituminous coals delivered to the Midwest, power plant engineers began to experiment with blending these coals with locally produced high-sulfur bituminous coals. While a 100% conversion to a sub-bituminous would result in a derating, it was equally evident that a 1% blend would have little effect and the operational question became at what mixture did the unit start to experience a reduction in operating efficiency. In what must be seen as a triumph

---

[18] Popp (2002) and Taylor *et al*. (2001) use the patent data in different ways. Popp constructs a "stock of knowledge" using various diffusion and decay assumptions as the independent variable while Taylor *et al*. rely on the annual count of patent grants.

of continuous thinking, the answer emerged that, depending on the unit and the coals being blended, mixtures of up to 60% of low-sulfur, sub-bituminous coal (and sometimes higher) could be used without significant derating in the generating capacity of the unit.

This re-engineering of existing bituminous coal-burning units to accommodate significant blends of low-sulfur, sub-bituminous coal could be considered an innovation. It was not observed before and not expected, but it can be seen also as diffusion of already known techniques for which there was previously no incentive to apply. It is clear that the previous regulatory instruments, which either mandated scrubbers or low sulfur coal, removed any incentive for experimenting with these blends which resulted in a coal of lower sulfur content (without being low sulfur coal) at much less cost than scrubbing or switching to a low-sulfur bituminous coal from Appalachia. The net effect was a lower sulfur premium for Appalachian low sulfur coal, consequent lower costs for switching in regions to the east beyond the economic frontier for western low-sulfur coals, and a lower allowance price.

Other cost-reducing changes that might be termed innovations can be observed upstream of the power plant in response to the sulfur premium. Mid-sulfur coal mines were developed in the Midwest where none existed before. These could supply a local coal at a price competitive with western blends, but when the only sulfur premium paid was for coal less than 1.2#SO_2/mmBtu, these mines could not compete with the lower cost but higher sulfur mines in the Midwest and were therefore not developed before. A similar shifting downward of the average sulfur content of coal being supplied was observed in Northern Appalachia, the other high-sulfur coal-producing region. These changes in coal supply to somewhat lower sulfur coals, which would still be considered mid- or high-sulfur coals, account for about 36% of the total reduction attributable to switching, or somewhat more than one fifth of the total. The causes were new mines now made economic in local markets, changes in mining practices that reduced the sulfur content of coal being already mined, and increased sulfur removal in coal preparation plants. The incentive for all of these changes was that premium now paid for lower sulfur content across the entire sulfur gradient. Whether these opportunities were known before to geologists, mining engineers, and prep plant operators and only needed the incentive to bring them forth awaits further research, but the answer will determine whether these innovative changes can be considered a change in the menu of technological options induced by Title IV or simply the diffusion of known techniques once the incentive was in place.

One further contribution of Title IV to lower cost that does not involve innovation is noted in Burtraw (1996) and labeled cost savings without emissions trading. Burtraw noted that giving plants the ability to choose between scrubbing, switching, and purchasing allowances created a competition among suppliers of abatement that was not present before. The threat to purchase allowances implies some trading to be credible, but it would not require a fully developed market and even without this threat, the ability to choose between switching and scrubbing increased competition and contributed to lower costs.

What emerges from the experience with Title IV is that costs are lower for reasons beyond the ability to trade emission reductions among sources. Improvements in productivity were occurring throughout the American economy during this period and Carlson et al. (2000) find that unspecified, exogenous productivity improvement applied to SO_2 abatement as well and accounted for as much as 20% of the reduction in the cost between 1985 and 1994. Quite aside from this background trend, a variety of industry sources indicate that the ability to trade emissions, and actual trading, have had effects in upstream markets and on the choice of technique that can be directly attributable to the flexibility that is inherent in market-based approaches to air emission regulation. Whether these changes, which often look like innovation, are true changes of the technical choices facing firms or simply the diffusion of known technology in response to the right incentive awaits further research. It is clear that costs are lower than expected for reasons beyond the extent of actual trading and that these changes were not expected.

Other Costs and Effects

All air emission control programs involve costs and effects beyond the directly observable abatement costs and the concomitant reduction in emissions. In the Acid Rain Program, administrative costs for both the regulator and the regulated are believed to have been less than in conventional regulatory programs, but no comprehensive study has been conducted on this subject. The more important aspect of the program’s administration concerns the revolutionary change in the nature of the tasks that are now required of the regulator and the regulated (Kruger, McLean, and Chen, 2000).

The shift of regulatory instrument from site-specific mandates to cap-and-trade has been accompanied by a corresponding shift in enforcement from relatively labor-intensive but intermittent inspection to data-intensive but continuous measurement and accounting. When what each source is doing to abate matters to the regulator, a corps of inspectors is needed to check periodically on the performance of the regulated. In a cap-and-trade system, the requirement that allowances be surrendered for all emissions permits the regulator to be indifferent about each source's abatement, and therefore to do without the corps of inspectors (except for the monitors); however, the *quid pro quo* is continuous measurement and reporting of emissions. In turn, this requires the handling of more data and a greater focus on accounting than was true of more conventional regulation.

The hallmark of the new system of regulation is continuous emissions monitoring and these monitors impose a non-negligible cost on operators that is estimated at 7% of direct compliance cost (MCA, pp. 248-50). As shown by Atkeson (1997) in her study of Title IV opt-in candidates, this cost can be a significant deterrent to voluntary participation. In the case of electric utility units subject to Title IV, continuous emissions monitoring and reporting was mandated for SO_2, NO_x and CO_2. To the extent that the information from these systems is used for the implementation of other air emission control programs, such as the Title IV NO_x averaging program or the Northeastern NO_x Budget Program, or that the data provide benefits aside from compliance uses, this cost should be shared with those other uses. Nevertheless, the experience with Title IV makes clear that the cost of this prerequisite for emissions trading is not negligible.

The administrative costs incurred by EPA are recognized as being less although of a different nature. Kruger, McLean and Chen (2000) describe the significant data handling requirements that are now faced and they suggest that this would not have been much more costly before recent advances in computing and data management. Despite this change of the nature of regulatory activity, the number of people involved in administering the program is a third [get McLean quote] of what would be required for a more conventional air emission control program.

Although no researcher has attempted to address the issue, the administrative costs of the cap-and-trade program for the regulated are not as clearly less than with conventional regulatory means. The cost of continuous emissions monitoring is the main item in this accounting. As is the case for the regulator, corporate administrative resources are shifted to emissions reporting and allowance management, but a good comparison of how these costs compare with what is required for dealing with inspectors and reporting under conventional command-and-control systems has not been made. It may not be any greater, but it is not clearly less. Whatever the case, regulated firms seem to be unanimous in expressing their preference for this type of regulation, presumably because the gains in reduced, direct compliance costs more than offset whatever additional costs are involved in monitoring and allowance management.

Another notable achievement in the realm of other costs is the notable reduction in the transaction costs involved in trading. The creation of a standard unit of account in allowances and the lack of any review requirement for trading has avoided the very large transactions costs that limited EPA's earlier experiments with emissions trading (Ellerman *et al*., forthcoming; Kruger, McLean, and Chen, 2000). The right to emit has been made into a readily tradable commodity and broker commissions are correspondingly low. This feature has, of course, greatly facilitated the development of a market and the concomitant cost savings.

Two effects of the Acid Rain Program that are not related to ancillary costs are also important. The first has been the creation of institutions with a continuing interest in emissions trading. The emergence of intermediaries, such as brokers, banks, and others who can offer trading and risk-management services, has already been mentioned. And, as is perhaps inevitable for any economic activity of note, an association has been formed, the Emissions Marketing Association, to promote emissions trading through a variety of educational, lobbying, informational, and other out-reach programs. Finally, there seems to be no end to the conferences, meetings, and workshops that bring participants from the private and public sectors and academia together to discuss one aspect or another of emissions trading.

While this institutionalization of emissions trading has occurred, somewhat of a backlash has also emerged recently as represented by Clear the Air (2002) and Moore (2002). The latter succinctly states the position of these groups: "trading ought to be rejected when proposed and repealed where it now exists" (p. 2). Both of these purported studies are lobbying documents occasioned by the Bush Administration's Clear Skies Proposal, which in addition to lowering the SO_2 cap by two-thirds and instituting national NO_x and mercury caps would effectively exempt units subject to these proposed caps from the best available control technology requirements of the existing Clean Air Act. Based on the experience with Title IV, one might conclude that this is a good trade-off, as advanced by some academics (Ellerman and Joskow, 2000) and as suggested by the publications of some environmental organizations (Goffman and Dudek, 1995; Environmental Defense, 2000) and researchers at some environmentally oriented research organizations (Swift, 2000; Swift, 2001), but this is far from a universally shared view among the environmental community. The reasons for rejecting emissions trading are beyond the scope of this paper but disdain for pollutant trading as morally reprehensible and concern for the loss of administrative discretion (and its many uses for non-environmental purposes) are always present. Although these attitudes may be viewed as a rear-guard reaction to an increasingly dominant consensus, they do find an echo on the editorial page of the New York Times and they have been translated into a law in New York that would restrict emissions trading. In what is perhaps an example of the new institutions, this state law has been struck down in the federal court as a violation of the interstate commerce clause of the U.S. Constitution in a motion for summary judgment brought by members of the Emissions Trading Association. [Get references on above]

Conclusions and Implications

The experience with Title IV and, to a lesser extent, other cap-and-trade programs marks a turning point in the regulation of air emissions in the U.S. This experience has shown that market-based incentive systems can reduce emissions as effectively, and even more so, and at considerably less cost than through conventional command-and-control mandates. As it result, it has become virtually obligatory that any legislative proposal to limit air emissions in the U.S. include emissions trading. While the agreement of left and right in the political spectrum is not as complete as it may appear on the surface, there seems little doubt that emissions trading will play an increasing role in the regulation of air emissions in the U.S. and probably elsewhere.

The conventional wisdom is that emissions trading will be necessary for new emission control initiatives and that the existing structure of command-and-control regulation is sacrosanct. Hence, all legislative proposals granting new authority to regulate air emissions include emissions trading; yet, their passage has been no faster for this reason. The same issues of cost and benefit and the same imperatives of building a viable political consensus remain. While legislative proposals that include emissions trading do not appear to be going anywhere fast, a less noticed and potentially more important change is occurring. Cap-and-trade systems are being adopted as a preferred means for achieving environmental goals for which ample legislative and regulatory authority already exists. The RECLAIM and Northeastern NO_x Budget Programs, as well as the NO_x SIO call, are instances of cap-and-trade programs being implemented within existing regulatory authority. This trend is in keeping with the reliance on market forces that has become manifest in one regulatory domain after another and it indicates that the increased use of cap-and-trade programs may occur as much through such incremental changes in the existing command-and-control structure than through bold new advances in the legislative domain.

References

Ackerman, Bruce A. and William T. Hassler. 1981. *Clean Coal and Dirty Air*, New Haven, CT: Yale University Press.

Arimura, Toshi H. (2002). “An Empirical Study of the SO_2 Allowance Market: Effects of PUC Regulation,” *Journal of Environmental Economics and Management* 44, pp. 271-289.

Atkeson, Erica. 1997. *Joint Implementation: Lessons from Title IV’s Voluntary Compliance Programs.* MIT-CEEPR Working Paper 97-003 (May).

Bailey, Elizabeth M. (1996). *Allowance Trading and State Regulatory Rulings: Evidence from the U.S. Acid Rain Program*, MIT-CEEPR Working Paper 96-002, Massachusetts Institute of Technology.

Bellas, Allen S. 1998. “Empirical Evidence of Advances in Scrubber Technology,” *Resource and Energy Economics*, 20:4 (December): 327-343.

Burtraw, Dallas 1996. “The SO_2 Emissions Trading Program: Cost Savings without Allowance Trades,” *Contemporary Economic Policy*, 14 (April): 79-94.

Carlson, Curtis P., Dallas Burtraw, Maureen Cropper, and Karen Palmer (Carlson *et al*, 2000, abbreviated to CBCP). ”SO_2 Control by Electric Utilities: What are the Gains from Trade?” *Journal of Political Economy,* Vol. 108, No. 6: 1292-1326.

Clear the Air: National Campaign against Dirty Power. 2002. “Darkening Skies: Trends toward increasing power plant emissions,” Washington, DC. Available at: http://cta.policy.net/fact/darkening_skies/.

Ellerman, A. Denny and Paul L. Joskow, 2000. “Clear Skies op-ed…” The New York Times, May 1, p. xxx. [Title and page no to be added]

Ellerman, A. Denny and Paul L. Joskow. (forthcoming). “To Retrofit or to Replace: The Economics of Scrubbing and Switching to Natural Gas,” MIT-CEEPR Working Paper.

Ellerman, A. Denny and Juan-Pablo Montero. 1998. “The Declining Trend in Sulfur Dioxide Emissions: Implications for Allowance Prices,” *Journal of Environmental Economics and Management* 36: 26-45. (This article is substantively reproduced as chapter 4 of Ellerman et al. (2000)

Ellerman, A. Denny and Juan-Pablo Montero. 2002. *The Temporal Efficiency of SO_2 Emissions Trading*, MIT-CEEPR Working Paper 02-003 (September)

Ellerman, A. Denny, David Harrison and Paul L. Joskow. (Ellerman *et al*., forthcoming). *Emissions Trading: Experience, Lessons, and Considerations for Greenhouse Gases*. Washington, D.C.: Pew Center for Global Climate Change.

Ellerman, A. Denny, Paul L. Joskow, Richard Schmalensee, Juan-Pablo Montero, and Elizabeth Bailey (Ellerman *et al*., 2000; abbreviated to MCA). *Markets for Clean Air: The U.S. Acid Rain Program*. Cambridge University Press, 2000.

Environmental Defense 2000. *From Obstacle to Opportunity: How Acid Rain Emissions Trading is Delivering Cleaner Air*. (September)

Frondel, Manuel and Christoph M. Schnmidt (2001). *Evaluating Environmental Programs: The Perspective of Modern Evaluation Research*, ZEW Discussion Paper 01-59. Mannheim, Germany: Centre for European Economic Research (ZEW).

Fullerton, Don, Shaun P. McDermott and Jonathan P. Caulkins (Fullerton *et al*., 1997). "Sulfur Dioxide Compliance of a Regulated Utility," *Journal of Environmental Economics and Management* 34, pp. 32-53.

Goffman, Joseph and Daniel J. Dudek 1995. "The Clean Air Act Acid Rain Program: Lessons for Success in Creating a New Paradigm." Print copy of presentation at 88th Annual Meeting of the Air and Waste Management Association, San Antonio, TX, June 18-23.

Goulder, Lawrence H., Ian W. H. Parry and Dallas Burtraw. 1997. "Revenue-raising vs. Other Approaches to Environmental Protection: The Critical Significance of Pre-existing Tax Distortions," *RAND Journal of Economics*, 28:4 (winter), 708-731.

Joskow, Paul L., Richard Schmalensee, and Elizabeth M. Bailey (Joskow *et al*., 1998). "The Market for Sulfur Dioxide Emissions," *American Economic Review* 88:4, pp. 669-685.

Kruger, Joseph A., Brian J. McLean, and Rayenne Chen (2000). "A Tale of Two Revolutions: Administration of the SO_2 Trading Program," in Richard Kosobud (ed.), *Emissions Trading: Environmental Policy's New Approach*. New York: John Wiley & Sons, Inc.

Magat, Wesley A. 1978. "Pollution Control and Technological Advance: A Dynamic Model of the Firm," *Journal of Enviornmental Economics and Management*, 5, pp1-25.

Milliman, Scott R. and Raymond Prince (1989). "Firm Incentives to Promote Technological Change in Pollution Control," *Journal of Environmental Ecoinomics and Management*, 17, pp. 247-265.

Montero, Juan-Pablo. 1999. "Voluntary Compliance with Market-based Environmental Policy: Evidence from the US Acid Rain Program," *Journal of Political Economy*, 107 (October), 998-1033.

Montero, Juan-Pablo, 2000. "Optimal Design of a Phase-in Emissions Trading Program," *Journal of Public Economics*, 75:2, 273-291.

Moore, Curtis A. 2002. *Marketing Failure: The Experience with Air Pollution Trading in the United States*. (January 2003 draft) Unpublished study being circulated at the time of this writing.

Organization for Economic Cooperation and Development, Environment Directorate. (OECD, 2001) "Ex Post Evaluations of Tradable Permits Programmes." ENV/EPOC/WPNEP (2001) 25.

Popp, David (2001). *Pollution Control Innovations and the Clean Air Act of 1990*, NBER Working Paper 8593. Cambridge, MA: National Bureau of Economic Research (November)

Rose, Kenneth J., Jr. (1995). "Twelve Common Myths of Allowance Trading: Improving the Level of Discussion," *The Electricity Journal* (May), pp. 64-69.

Rose, Kenneth J., Jr. (2000). "Electric Industry Restructuring and the SO_2 Emissions Trading Program: A Look Ahead by Looking Back," in *Emissions Trading: Environmental Policy's New Approach*, (Richard F. Kosobud, ed.), New York: John Wiley & Sons, Inc., pp. 209-215.

Schmalensee, Richard, Paul L. Joskow, A. Denny Ellerman, Juan-Pablo Montero and Elizabeth M. Bailey (Schmalensee *et al.*, 1998). *Journal of Economic Perspectives* 12 (summer), pp. 53-68.

Swift, Byron 2000. "Allowance Trading and SO_2 Hot Spots: Good News from the Acid Rain Program," *Environment Reporter,* 31:19 (May 12), pp. 954-59. Washington, DC: BNA, Inc.

Swift, Byron 2001. "How Environmental Laws Work: An Analysis of the Utility Sector's Response to Regulation of Nitrogen Oxides and Sulfur Dioxide under the Clean Air Act," *Tulane Environmental Law Journal*, 14:309 (summer)

Taylor, Margaret R., Edward S. Rubin, and David A. Hounshell (Taylor *et al.*, 2001), "The Effect of Government Actions on Technological Innovation for SO_2 Control," *Proceedings of the EPA/DOE/EPRI/AWMA MegaSymposium, August20-23, 2001.* Pittsburgh, PA: Air & Waste Management Association (October).

U.S. Environmental Protection Agency (USEPA 2002). *EPA Acid Rain Program: 2001 Progress Report*. EPA-430-R-2-009 (November) (available at: http://www.epa/gov/airmarkets)

Winebrake, James J., Alexander E. Farrell, and Mark A. Bernstein (Winebrake *et al.*, 1995). "The Clean Air Act's sulfur dioxide emissions market: Estimating the costs of regulatory and legislative intervention," *Resource and Energy Economics* 17, pp. 239-260.