

Supporting information for  
CABOT-O<sub>3</sub>: An Optimization Model for Air  
Quality Benefit-Cost and Distributional  
Impacts Analysis

David A. Bielen,\* Alexander J. Macpherson, Heather Simon, and Neal Fann

*Office of Air Quality Planning and Standards, Office of Air and Radiation, United States  
Environmental Protection Agency*

E-mail: [bielen.david@epa.gov](mailto:bielen.david@epa.gov)

# Mathematical Model Formulation

In this section, we describe the optimization model mathematically. First, we restate the emissions control strategy cost optimization model presented in Macpherson et al.,<sup>1</sup> as that model forms a basis for the enhancements that are the focus of this paper. Second, we describe the mathematical formulation of the health impacts module. Third, we present the mathematical formulation that enables the examination of distributional impacts of changes in ozone levels. Fourth, we present how the components are integrated to examine a range of scenarios.

## Control Strategy Cost Model

### Ozone Design Values

Air quality monitors that measure ozone concentrations are located across the contiguous United States and are indexed by  $m = 1, \dots, M$ . Each monitor has a projected future baseline ozone level, which is called a design value, or  $dv_m^{base}$ . The design value is the air quality metric subject to ambient air quality-related regulatory constraints. Impacts on ozone concentrations at the monitors depend on changes in emissions of ozone precursor pollutants,  $p = 1, \dots, P$ , within the regions of the monitors and in upwind regions. To model the source-receptor relationship between changes in emissions and design values at monitors, we employ a system of linear air quality response factors,  $a_{p,r,m}$ , which estimate the impact of emissions reductions in region,  $r = 1, \dots, R$ , on ozone concentrations at monitor  $m$  in terms of parts-per-billion (ppb) per ton emissions reduction in pollutant,  $p$ . Emissions reductions within each region  $r$ , written as  $er_r$ , are multiplied by the air quality response factors associating each region to each monitor to obtain a revised ozone level  $dv_m^{new}$  at each monitor, or:

$$dv_m^{new} = dv_m^{base} + \sum_{p=1}^P \sum_{r=1}^R a_{p,r,m} er_{p,r} \text{ for all } m. \quad (1)$$

Policy scenarios then require design values at each monitor to be reduced to or below an ozone target,  $dv_m^{goal}$ :

$$dv_m^{new} \leq dv_m^{goal} \text{ for all } m. \quad (2)$$

## Emissions Controls

There is a set of potentially controllable emissions sources,  $i = 1, \dots, I$ . For many sources, there are multiple applicable control measures, each of which may achieve different levels of emission reductions at different costs than other applicable measures. Decisions whether to control emissions at these sources are defined by a vector of binary decision variables,  $x_{i,t}$ , where  $i$  indexes the source and  $t$  indexes the applicable emissions control measure for that source, with  $T_i$  denoting the set of applicable control measures for source  $i$ . The variable equals one if the control is applied, zero otherwise. There are as many binary decision variables as there are applicable control measures per controllable source. Here we limit the control application to a single control, while in practice more than one measure may be feasibly applied to the source. In this application, the sum of the  $x_{i,t}$  must not exceed one, or:

$$\sum_{t \in T_i} x_{i,t} \leq 1 \text{ for all } i. \quad (3)$$

The vector of annualized costs associated with each application of these control measures is identified by  $c_{i,t}$ .

As in Macpherson et al.,<sup>1</sup> we include a backstop emissions control measure within each region, which can be implemented if, for example, achieving an ozone target requires more emissions reductions than can be accomplished with the control measures within the emissions control measure database. We model each regional backstop measure as a non-negative, continuous decision variable  $\hat{x}_{p,r}$  indicating the amount of emissions reductions from backstop technologies in region  $r$  at the assumed cost of  $\hat{c}_{p,r}$ . The total annual cost,  $TC$ , of the emission control strategy is the sum of the costs of decisions over specified and backstop

47 technologies:

$$TC = \sum_{i=1}^I \sum_{t \in T_i} c_{i,t} x_{i,t} + \sum_{p=1}^P \sum_{r=1}^R \hat{c}_{p,r} \hat{x}_{p,r}. \quad (4)$$

48 To obtain total emissions reductions of each pollutant within each region,  $er_{p,r}$ , the  
 49 emissions reductions from the application of control measures are added to the reductions  
 50 produced by backstop technologies, which are obtained from the decision vector  $\hat{x}_{p,r}$ :

$$er_{p,r} = \sum_{i=1}^I \sum_{t \in T_i} e_{p,r,i,t} x_{i,t} + \hat{x}_{p,r} \text{ for all } p \text{ and } r, \quad (5)$$

51 where  $e_{p,r,i,t}$  represents the reduction in emissions of pollutant  $p$  in region  $r$  from source  $i$  from  
 52 the application of control  $t$ . While we do not index emission control measures by region, the  
 53 emissions reductions produced through the applications of controls belong in regions. This  
 54 is equivalent to saying each application of a control produces reductions within the region of  
 55 its source, but not in other regions. Also, as it is not possible to reduce emissions in a region  
 56 more than the level recorded in the region's anthropogenic emissions inventories, regional  
 57 emissions reductions are subject to an upper-bound constraint for each pollutant and region:

$$er_{p,r} \leq er_{p,r}^{max} \text{ for all } p \text{ and } r. \quad (6)$$

## 58 Complete Control Strategy Cost Model

59 Combining (1) through (6) into an optimization framework yields a mixed integer linear  
60 programming problem:

$$\begin{aligned}
& \underset{x_{i,t}, \hat{x}_{p,r}}{\text{minimize}} & TC &= \sum_{i=1}^I \sum_{t \in T_i} c_{i,t} x_{i,t} + \sum_{p=1}^P \sum_{r=1}^R \hat{c}_{p,r} \hat{x}_{p,r} \\
& \text{subject to} & dv_m^{base} + \sum_{p=1}^P \sum_{r=1}^R a_{p,r,m} er_{p,r} &\leq dv_m^{goal} \text{ for all } m \\
& & \sum_{i=1}^I \sum_{t \in T_i} e_{p,r,i,t} x_{i,t} + \hat{x}_{p,r} &\leq er_{p,r}^{max} \text{ for all } p \text{ and } r \\
& & x_{i,t} &\in \{0, 1\} \text{ for all } i \text{ and } t \\
& & \sum_{t \in T_i} x_{i,t} &\leq 1 \text{ for all } i \\
& & \hat{x}_{p,r} &\geq 0 \text{ for all } p \text{ and } r.
\end{aligned} \tag{7}$$

61 This model chooses controls to meet design value targets while minimizing total cost.

## 62 Health Impacts and Benefits Module

### 63 Daily Maximum 8-Hour Average Ozone Concentrations

64 The assessment of changes in ozone concentration for health impacts analysis differs from  
65 the assessment of changes in ozone design values for evaluating compliance with ambient  
66 air quality standards. Whereas the analysis of design values focuses on changes at specific  
67 monitor locations, the assessment of ozone changes for health impacts analysis is performed  
68 where people might experience the air quality changes, which are measured by the change in  
69 daily maximum eight-hour average ozone (MDA8) concentrations across all days from May  
70 through September.

71 Similar to health impacts analysis performed using BenMAP,<sup>2</sup> we perform analysis using  
72 geospatial grid cells that contain population data. We index these grid cells by  $g = 1, \dots, G$ .

Like we do with the regulatory ozone metric for monitoring stations, we leverage source-receptor relationships between emissions changes in regions to ozone concentrations in receptor grid cells to estimate the change from baseline summertime average MDA8 concentrations,  $mda8_g^{base}$ , to concentrations under the policy,  $mda8_g^{new}$ . Using a system of linear air quality response factors,  $b_{p,r,g}$ , which estimate the impact of emissions reductions of precursor pollutant  $p$  in region  $r$  on summertime MDA8 concentrations at grid cells in terms of parts-per-billion (ppb) per ton emissions reduction, the equation describing changes in summertime average MDA8 concentrations is

$$mda8_g^{new} = mda8_g^{base} + \sum_{p=1}^P \sum_{r=1}^R b_{p,r,g} er_{p,r} \text{ for all } g, \quad (8)$$

where  $s$  denotes a summertime average and  $er_{p,r}$  is calculated as in (5) above. For convenience in later equations, we define the change in summertime mean MDA8 concentrations under a policy as  $mda8_g^\Delta$ , where

$$mda8_g^\Delta = \sum_{p=1}^P \sum_{r=1}^R b_{p,r,g} er_{p,r} \text{ for all } g. \quad (9)$$

## Ozone Health Impact Functions and Valuation

The approach to the ozone health impact analysis in the benefits module requires estimating the change in population-level exposure from changes in summertime MDA8 concentrations. From this change in population exposure, we calculate health impacts by applying concentration-response (C-R) relationships drawn from the epidemiological literature.<sup>3</sup> While there are numerous functional forms for various health impact endpoints in the literature, in presenting the mathematical framework of the this model, we focus upon a common C-R functional form that is used to estimate the change in premature deaths from ozone pollution. We index health impacts by  $h = 1, \dots, H$  and the baseline incidence rate for health impact  $h$  in grid cell  $g$  as  $y_{h,g}^0$ . For expositional convenience, we assume that the

94 population in each grid cell is demographically homogeneous so that we do not need to index  
 95 sub-groups that may have different baseline incidence levels within a grid cell. In the nu-  
 96 merical representation of the model, there are different demographic groups within grid cells  
 97 that have different baseline incidence levels. Where the C-R function parameter is indicated  
 98 by  $\beta_h$ , the change in the health impact incidence rate,  $y_{h,g}^\Delta$ , is

$$y_{h,g}^\Delta = y_{h,g}^0 \times \left( e^{\beta_h \times mda8_g^\Delta} - 1 \right) \text{ for all } h, g. \quad (10)$$

99 The change in health impact incidence is then the population subject to the health impact  
 100  $h$  in grid cell  $g$ , indexed by  $pop_{h,g}$ , multiplied by the change in the incidence rate:

$$z_{h,g}^\Delta = y_{h,g}^\Delta \times pop_{h,g} \quad (11)$$

101 To estimate the monetary value of the change in incidence,  $benefits_{h,g}$ , we multiple a valu-  
 102 ation parameter,  $valuation_h$ , by the change in incidence:

$$benefits_{h,g} = valuation_h \times z_{h,g}^\Delta \text{ for all } h, g. \quad (12)$$

Where total benefits is represented by TB, the complete ozone benefits module combines (8)  
 through (12):

$$TB = \sum_{h=1}^H \sum_{g=1}^G benefits_{h,g},$$

103 with

$$\begin{aligned}
er_{p,r} &\leq er_{p,r}^{max} && \text{for all } p \text{ and } r \\
mda8_g^\Delta &= \sum_{p=1}^P \sum_{r=1}^R b_{p,r,g} er_{p,r} && \text{for all } g \\
y_{h,g}^\Delta &= y_{h,g}^0 \times \left( e^{\beta_h \times mda8_g^\Delta} - 1 \right) && \text{for all } h, g \\
z_{h,g}^\Delta &= y_{h,g}^\Delta \times population_{h,g} && \text{for all } h, g \\
benefits_{h,g} &= valuation_h \times z_{h,g}^\Delta && \text{for all } h, g.
\end{aligned} \tag{13}$$

104 There are two ways to use the tool to estimate the health-related impacts and monetized  
105 benefits of a change in ozone concentrations, as well as the net benefits,  $NB$ , of a policy. First,  
106 the health impacts module can be used to "post-process" the emissions reductions drawn  
107 from a least cost solution of a model (or any hypothetical change in emissions) as presented in  
108 (7). This post-processing approach yields an estimate of health impacts, monetized benefits,  
109 and net benefits even though health impact considerations were not directly incorporated  
110 into the objective function of the model. Alternatively, the health impacts module can  
111 be directly incorporated into the optimization problem, such that emissions reductions are  
112 selected to maximize net benefits (or benefits alone) subject to constraints.

## 113 **Distributional Impacts**

114 We next describe how the risk of ozone-related premature death is distributed across the  
115 U.S. population using the Atkinson Index (AI). This measure expresses quantitatively the  
116 degree to which air pollution attributable risks are shared equally throughout the U.S. The  
117 AI measures risk inequality on a scale of 0 to 1, with higher values indicating greater in-  
118 equality in risk. Among the measures of risk inequality, the literature frequently identifies  
119 the AI as being most appropriate because it is both subgroup decomposable (i.e., it can  
120 express inequality within and between subgroups) and because it satisfies the Pigou-Dalton  
121 transfer principle (i.e., the index decreases when risks transfer from high-risk to low-risk  
122 populations).<sup>4</sup>



In the interest of brevity, we calculate the AI for the total population for the baseline and each of five policy scenarios. We do not explore here how each scenario affects the distribution of risk within and between subgroups. The AI is calculated as follows:

$$A = \begin{cases} 1 - \frac{1}{\sum_{g=1}^G pop_g} \left( \sum_{g=1}^G pop_g \left( \frac{mda8_g}{\overline{mda8}} \right)^{1-\epsilon} \right)^{1/(1-\epsilon)} & \text{if } \epsilon \neq 1 \\ 1 - \frac{\sum_{g=1}^G pop_g \overline{mda8_g}}{\overline{mda8}} & \text{if } \epsilon = 1 \end{cases} \quad (14)$$

where  $pop_g$  is the total adult population in grid cell  $g$ ,  $\overline{mda8}$  is the summertime average MDA8 concentration across grid cells, and  $\epsilon$  is a sensitivity parameter describing tolerance for risk inequality, with greater values expressing less tolerance for inequality. Consistent with recent analyses, we present results for epsilon values of 0.75 and 3 (e.g., Fann et al. <sup>5</sup>).

## Implementation

The model is implemented in the General Algebraic Modeling System (GAMS) version 25.2. The cost minimization version of the model is a linear mixed-integer program (MIP) solved with the CPLEX 12 solver, which uses a branch and cut algorithm to solve MIPs. The linearity of the program is due to the assumed linear form of the cost function and in the relationship between emission reductions and ozone concentrations. The net benefit maximization version of the model is a mixed-integer nonlinear program (MINLP) solved with DICOPT. The nonlinearity of the model is due to the form of the concentration-response functions, as illustrated in Equation (10). The MINLP algorithm iterates between NLP and MIP sub-problems. The NLP sub-problems are solved using the CONOPT 3 solver, which uses a generalized reduced gradient method, and the MIP sub-problems are solved with CPLEX. For the case study considered here, the solution to the MINLP appears very stable. For example, the same solution is reached when the model variables are initialized from the default values in GAMS (zeros) and when they are initialized using the solution to the *CostMin-65* scenario. Model inputs are pre-processed in R, which reads data from Microsoft

145 Excel 2013 workbooks, comma-separated values (.csv) files, and the BenMAP-CE database  
146 and exports .csv files to be read into GAMS. Model outputs are exported from GAMS as  
147 GAMS data exchange (.gdx) files and read into R using the GDXRRW version 1.0.4 package,  
148 where they are post-processed and summarized as figures and tables.

## Air Quality Modeling

Here we describe the modeling and methods used to derive the air quality response factors ( $a_{p,r,m}$  for ozone design values and  $b_{p,r,g}$  for summertime mean MDA8 concentrations used in the health impacts analysis). First we describe the underlying photochemical modeling simulations. Then we describe how air quality model results and ambient air quality data are processed to project future year design values. Next, we describe how ambient ozone measurement data and model results are used to project ozone season average gridded fields for the future year of 2023. Finally, we describe how additional data derived from the instrumented model simulations using the source apportionment technique are applied in concert with projected 2023 ozone values to derive the air quality response factors.

### Baseline Air Quality Modeling Simulations

Air quality modeling results for a 2011 base year and a 2023 projected year were obtained from EPA's modeling to support the 2018 – Final Cross-State Air Pollution Rule Close-Out.<sup>6</sup> The modeling was performed using the CAMx v6.4 photochemical grid model using a 12 km-by-12 km resolution grid covering the contiguous United States.<sup>7</sup> Both base year and projected future year modeling simulations used the same 2011 meteorological and boundary condition inputs. Meteorological inputs representing 2011 were derived from a simulation of the Weather Research Forecasting model version 3.4.<sup>8</sup> Boundary conditions were created from a global chemistry model simulation using GEOS-Chem standard version 8-03-02 with 8-02-01 chemistry.<sup>9</sup> Base year emissions (2011) and projected future year emissions (2023) are described in detail in U.S. EPA.<sup>10</sup>

Seasonal average 8-hour daily maximum (MDA8) modeled ozone concentrations for 2011 were compared to measured values at over 1,000 monitors across the United States. Aggregate model performance statistics were calculated by region for the summertime ozone season (May–September). The model evaluation, which is described in detail in U.S. EPA,<sup>6</sup> found

that model bias and error were generally within the range of model performance statistics reported in other recent state-of-the-science air quality modeling simulations. Overall, the comparison with measured values showed that the model does a good job of replicating the magnitude and variability of ozone across the United States and is suitable for regulatory and scientific research applications.<sup>6</sup>

## Projection of 2023 Design Values at Monitor Locations

MDA8 design values (DVs) for 2023 were calculated in accordance with EPA guidance.<sup>11</sup> First, a five-year weighted base year DV was calculated using 2009–2013 data at each monitoring site. While DVs for the purpose of showing compliance with the NAAQS are calculated as the three-year average of the annual 4<sup>th</sup> highest MDA8 at a particular monitor, five years of ambient data were used to project values into the future with the intent of providing a more stable value due to the impact of year-to-year meteorological changes. Next, the 2011 and 2023 model values were used to calculate a relative response factor (RRF), which is the ratio of the future year ozone divided by the current year ozone. The procedure used to calculate the RRF for each monitor was as follows:

- Start with a  $3 \times 3$  matrix of 12 km-by-12 km grid cells surrounding each monitor.
- Eliminate any grid cells that are predominantly covered by water.
- Calculate the RRF for each monitor location using the following equation:

$$RRF_m = \frac{\overline{MDA8}_{2023,m,highdays}}{\overline{MDA8}_{2011,m,highdays}}, \text{ where} \quad (15)$$

- $RRF_m$  is the RRF for monitor  $m$
- $\overline{MDA8}_{2011,m,highdays}$  is the average 2011 modeled MDA8 on high days near monitor  $m$ , calculated as follows:

- \* On each day in 2011, find the highest modeled MDA8 values in the  $3 \times 3$  matrix of grid cells surrounding the monitor excluding water cells.
  - \* If more than 10 days have MDA8 values  $\geq 60$  ppb in the  $3 \times 3$  matrix, then take the average of the MDA8 values on the 10 days with the highest modeled 2011 MDA8 ozone values.
  - \* If at least 5 days but fewer than 10 days have modeled MDA8 values of  $\geq 60$  ppb, then take the average of all days  $\geq 60$  ppb.
  - \* If fewer than 5 days have modeled MDA8 values  $\geq 60$  ppb, then no value RRF can be calculated and no DV is projected.
- $\overline{MDA8}_{2023,m,highdays}$  is the average 2023 modeled MDA8 on high days near monitor  $m$ , calculated as follows:
- \* For each monitor find the days and grid cells used to calculate  $\overline{MDA8}_{2011,m,highdays}$ .
  - \* For those days and grid cells, average the 2023 modeled MDA8 values together.

Using the RRF, the 2023 design value was calculated using the following equation:

$$DVF_m = RRF_m \times DVB_m, \text{ where} \quad (16)$$

- $DVF_m$  is the projected future-year (2023) DV at monitor  $m$ , and
- $DVB_m$  is the 5-year weighted average (2009–2013) base year DV at monitor  $m$ .

## Projection of Gridded Seasonal Average MDA8 Values

Gridded seasonal average 2011 MDA8 ozone values were created by fusing gridded 2011 model fields with observed (2009–2013) data at monitor locations, using enhanced Voronoi Neighbor Averaging (eVNA). This technique starts with an interpolated field of observed MDA8 ozone values using Voronoi Neighbor Averaging with inverse distance weighting. Then interpolated

values in model grid cells away from monitors are adjusted based on the modeled ozone spatial gradient.<sup>12–15</sup> The technique allows for accurate values at monitor locations where ozone concentrations are known but still takes advantage of the extra information provided by the model in locations with no observed data. The eVNA processing was performed using the Software for EPA’s Modeled Attainment Test – Community Edition (SMAT-CE) software package.<sup>16</sup>

## **Use of Source Apportionment Modeling to Calculate Air Quality Response Factors**

Source apportionment is a technique within the CAMx photochemical model that allows predefined groups of emissions to be tracked through the model’s chemical and physical processes. When applying this method, the model is instrumented with capabilities that track the source of ozone precursors, determine how much of the ozone formed through chemical reactions results from each of the predefined emissions categories, and track the ozone formed from each emissions category (called a tag) as it is transported from one location to another and lost from the atmosphere through chemical reactions and deposition to surfaces. For this analysis, we used outputs created with the Anthropogenic Precursor Culpability Analysis (APCA) source apportionment method within CAMx as part of the EPA 2023 modeling to support 2015 NAAQS transport State Implementation Plans.<sup>17</sup> This modeling tagged ozone from anthropogenic NO<sub>x</sub> and VOC emissions by state resulting in estimates of contributions from anthropogenic NO<sub>x</sub> and VOC in each state to ozone at every grid cell in the modeling domain for every hour in 2023. These hourly contributions were then aggregated up to MDA8 values matching the hours that were used to calculate the total 2023 MDA8 values for each grid cell and day. This information was used to determine both state-level NO<sub>x</sub> and VOC contributions to 2023 DVs at each monitor location and state-level NO<sub>x</sub> and VOC contributions to 2023 fused surfaces of summertime average MDA8 values.

The following equation was used to calculate each state’s NO<sub>x</sub> and VOC contribution to

243 2023 projected DVs at monitor locations:

$$DVC_{t,p,m} = \frac{\sum_{d=1}^n MDA8_{t,p,m,d}}{\sum_{d=1}^n \sum_{t=1}^T \sum_{p=1}^2 MDA8_{t,p,m,d}} \times DVF \quad (17)$$

244 •  $DVC_{t,p,m}$  represents the DV contribution (ppb) from tag  $t$  and pollutant  $p$  to monitor  
 245  $m$ . Note that here we use subscript  $t$  for tag, rather than  $r$  for region, because there were  
 246 some additional source apportionment tags that did not track emissions by regions,  
 247 such as a tag to track ozone coming from all biogenic emissions across the U.S. All  
 248 ozone sources are included in exactly one of the source apportionment tags, and there  
 249 are  $T$  tags in total.

250 •  $MDA8_{t,p,m,d}$  represents the 2023 MDA8 contributions from tag  $t$  and pollutant  $p$  to  
 251 monitor  $m$  on day  $d$ , with  $n$  days included in the DV calculation.

252 Linear air quality DV response factors  $a_{p,r,m}$  were then calculated by dividing  $DVC_{t,p,m}$  by  
 253 the annual tons of emissions from tag  $t$  (i.e., region  $r$ ) and pollutant  $p$ .

254 To calculate the response factors (denoted  $b_{p,r,g}$ ) from tag  $t$  to summertime average MDA8  
 255 at grid cell  $g$ , we started with the 2023 eVNA value ( $eVNA_{g,2023}$ ) at grid cell  $g$ . Similar to  
 256 the method used to determine  $DVC_{t,p,m}$ , we used:

$$eVNAC_{t,p,g} = \frac{\sum_{d=1}^{153} MDA8_{t,p,g,d}}{\sum_{d=1}^{153} \sum_{t=1}^T \sum_{p=1}^2 MDA8_{t,p,g,d}} \times eVNA_{g,2023} \quad (18)$$

257 where

258 •  $eVNAC_{t,p,g}$  is the contribution (ppb) of emissions from tag  $t$  and pollutant  $p$  to the  
 259 2023 eVNA seasonal fused surface at grid cell  $g$ .

260 •  $MDA8_{t,p,g,d}$  represents the 2023 MDA8 contributions from tag  $t$  and pollutant  $p$  to  
 261 grid cell  $g$  on day  $d$ , with 153 days occurring in summertime.

262 Linear air quality summertime mean MDA8 response factors  $b_{p,r,g}$  are then calculated by

263 dividing  $eVNAC_{t,p,g}$  by the annual tons of emissions from tag  $t$  (i.e., region  $r$ ) and pollutant  
264  $p$ .



## Backstop Cost Sensitivity

In this section, we describe the results of a sensitivity analysis on the cost of the backstop emissions controls. The results that we present in the main text are based on a backstop cost of \$15,000 per ton of emissions reductions. Here we summarize results for alternative cost scenarios with backstop costs of \$10,000 and \$20,000 per ton of emissions reductions, respectively, in addition to the default \$15,000 per ton scenarios. The \$10,000 and \$20,000 per ton cost levels reflect low and high cost scenarios that were explored in the 2015 Ozone NAAQS RIA.<sup>18</sup> To keep the presentation tractable, we focus our sensitivity analysis on cost minimization with a 65 ppb ozone constraint and net benefits maximization without an ambient air quality constraint. For illustrative purposes, we also show results assuming a backstop cost of \$5,000 per ton for the unconstrained net benefits scenario, since they vary significantly when the backstop cost parameter is set at a low value. Thus, in total we compare results for seven scenarios (three backstop levels each for both program configurations, plus one additional net benefits scenario), denoting them by the objective function and the assumed backstop cost level. The scenarios denoted as *CostMin-15k* and *BenMax-15k* in this section are identical to the *CostMin-65* and *BenMax-Free* scenarios, respectively, described in the main text.

Aggregate cost and benefit values for the contiguous United States are depicted in Figure 1. The panels in Figure 1 represent total control costs (top left), total ozone reduction benefits (top right), net ozone benefits (bottom left), and total PM<sub>2.5</sub> benefits (bottom right). Each bar within a panel represents a different scenario.

Figure 1 illustrates how the results depend on the backstop cost assumption. For the cost minimization scenarios, total costs are heavily influenced by the assumed backstop cost parameter, ranging from \$6.0 to \$11 billion. Total costs for the *CostMin-20k* scenario are 1.27 times higher than for the *CostMin-15k* scenario, which makes sense given the one-third increase in the assumed backstop cost. In contrast to cost levels, the total ozone (and PM<sub>2.5</sub> benefits levels are largely the same across scenarios. Accordingly, the net benefits vary

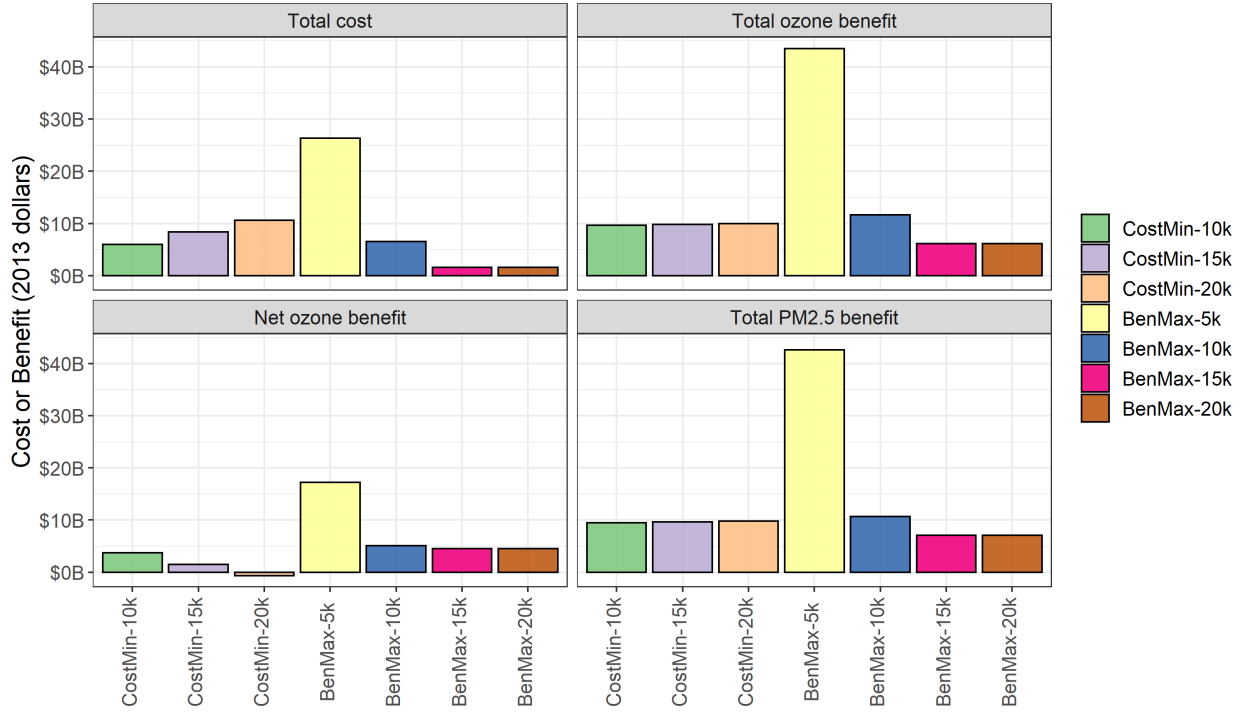


Figure 1: National cost and benefit metrics for different backstop cost assumptions.

greatly, from \$3.7 to \$-0.6 billion.

For the net benefits maximization scenarios, Figure 1 provides additional context for the health benefits value of reducing emissions. The results for the *BenMax-15k* and *BenMax-20k* scenarios are identical, which is an anticipated result since no backstop reductions were leveraged in the solution for the *BenMax-15k* scenario (see Figure 4 in the main text). Reducing the backstop costs to \$10,000 results in a relatively small increase in net ozone benefits, while reducing the backstop costs to \$5,000 results in an enormous increase in net ozone benefits, demonstrating significant convexity in the relationship between backstop cost levels and net ozone benefits.

The emissions reductions underlying costs and benefits are depicted in Figure 2. The left panels in Figure 2 represent emissions reductions for the *CostMin-15k* and *BenMax-15k* scenarios. The right panels represent the differences in reductions from those scenarios for the other cost minimization and net benefits maximization scenarios, respectively.

The cost minimization scenarios illustrate how trade-offs between backstop and known

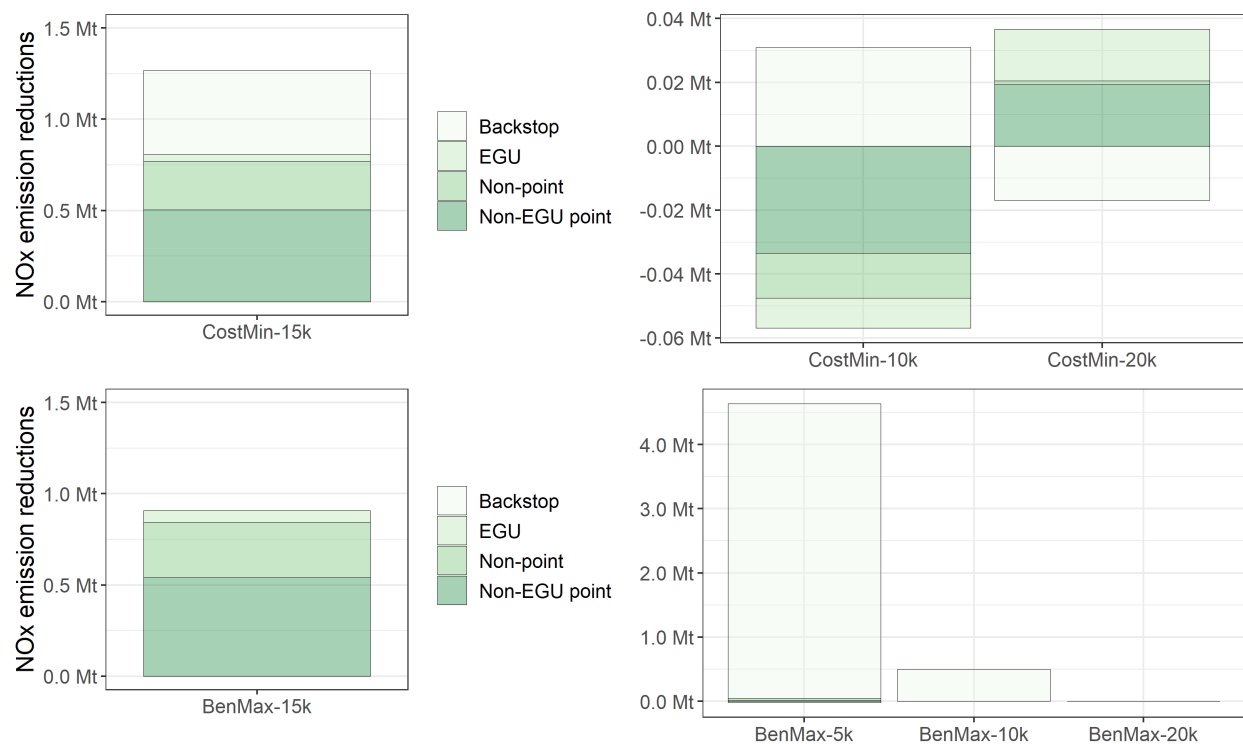


Figure 2: National NO<sub>x</sub> emissions reductions by type for the *CostMin-15k* and *BenMax-15k* scenarios (left) and difference in reductions from those scenarios for scenarios with alternative backstop cost assumptions (right).

controls depend on the assumed backstop cost. Roughly 4% of the backstop cost reductions leveraged in the *CostMin-15k* scenario are displaced by reductions from known sources in the *CostMin-20k* scenario, while roughly 7% more backstop control is used in the *CostMin-10k* scenario than in the *CostMin-15k* scenario. In both cases, differences in emissions reductions from backstop control are smaller in magnitude than the corresponding differences in emissions reductions from known sources. Since backstop control is allowed to be used in any state for the same cost, known sources from further away tend to be displaced by local backstop control as the backstop cost decreases.

The net benefits maximization scenarios demonstrate that, under our modeling assumptions, there is significant value in reducing NO<sub>x</sub> emissions when backstop costs are inexpensive. We find that an additional five million tons of emissions reductions are net beneficial when moving from the *BenMax-15k* scenario to the *BenMax-5k* scenario, which is more than five times the total emissions reductions from all sources in the *BenMax-15k* scenario.

## References

- (1) Macpherson, A. J.; Simon, H.; Langdon, R.; Misenheimer, D. A Mixed Integer Programming Model for National Ambient Air Quality Standards (NAAQS) Attainment Strategy Analysis. *Environmental Modelling & Software* **2017**, *91*, 13–27.
- (2) U.S. EPA, Environmental Benefits Mapping and Analysis Program - Community Edition (BenMAP-CE) User's Manual. [https://www.epa.gov/sites/production/files/2015-04/documents/benmap-ce\\_user\\_manual\\_march\\_2015.pdf](https://www.epa.gov/sites/production/files/2015-04/documents/benmap-ce_user_manual_march_2015.pdf), 2018.
- (3) Hubbell, B.; Fann, N.; Levy, J. I. Methodological considerations in developing local-scale health impact assessments: balancing national, regional, and local data. *Air Quality, Atmosphere & Health* **2009**, *2*, 99–110.
- (4) Levy, J. I.; Chemerynski, S. M.; Tuchmann, J. L. Incorporating concepts of inequality

and inequity into health benefits analysis. *International Journal for Equity in Health* **2006**, 5.

(5) Fann, N.; Coffman, E.; Timin, B.; Kelly, J. T. The estimated change in the level and distribution of PM<sub>2.5</sub>-attributable health impacts in the United States: 2005–2014. *Environmental Research* **2018**, 167, 506–514.

(6) U.S. EPA, *Air Quality Modeling Technical Support Document for the Updated 2023 Projected Ozone Design Values*; 2018; [https://www.epa.gov/sites/production/files/2018-06/documents/aq\\_modelingtsd\\_updated\\_2023\\_modeling\\_o3\\_dvs.pdf](https://www.epa.gov/sites/production/files/2018-06/documents/aq_modelingtsd_updated_2023_modeling_o3_dvs.pdf).

(7) Ramboll Environ, WRF-CAMX version 4.3 Release Notes. <http://www.camx.com>, 2014.

(8) Skamarock, W. C.; Klemp, J. B.; Dudhia, J.; Gill, D. O.; Barker, D. M.; Duda, M. G.; Huang, X.-Y.; Wang, W.; Powers, J. G. *A Description of the Advanced Research WRF Version 3*; 2008; <https://opensky.ucar.edu/islandora/object/technotes:500>.

(9) Yantosca, B. GEOS-CHEMv7-01-02 User's Guide. Atmospheric Chemistry Modeling Group, Harvard University: Cambridge, MA, 2004.

(10) U.S. EPA, *Technical Support Document (TSD): Additional Updates to Emissions Inventories for the Version 6.3, 2011 Emissions Modeling Platform for the Year 2023*; 2017; [https://www.epa.gov/sites/production/files/2017-11/documents/2011v6.3\\_2023en\\_update\\_emismod\\_tsd\\_oct2017.pdf](https://www.epa.gov/sites/production/files/2017-11/documents/2011v6.3_2023en_update_emismod_tsd_oct2017.pdf).

(11) U.S. EPA, *Modeling Guidance for Demonstrating Air Quality Goals for Ozone, PM<sub>2.5</sub> and Regional Haze*; 2018.

(12) Gold, C. M.; Remmele, P. R.; Roos, T. In *Algorithmic Foundations of Geographic Information Systems*; van Kreveld, M., Nievergelt, J., Roos, T., Widmayer, P., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1997; Chapter 2, pp 21–35.

- (13) U.S. EPA, *Technical Report on Ozone Exposure, Risk, and Impact Assessments for Vegetation*; 2007.
- (14) Ding, D.; Zhu, Y.; Jang, C.; Lin, C.-J.; Wang, S.; Fu, J.; Gao, J.; Deng, S.; Xie, J.; Qiu, X. Evaluation of health benefit using BenMAP-CE with an integrated scheme of model and monitor data during Guangzhou Asian Games. *Journal of Environmental Sciences* **2016**, *42*, 9–18.
- (15) U.S. EPA, *Regulatory Impact Analysis for the Proposed Emission Guidelines for Greenhouse Gas Emissions from Existing Electric Utility Generating Units; Revisions to Emission Guideline Implementing Regulations; Revisions to New Source Review Program*; 2018.
- (16) U.S. EPA, Support Center for Regulatory Atmospheric Modeling (SCRAM). <https://www.epa.gov/scram/photochemical-modeling-tools>, 2018.
- (17) U.S. EPA, *Information on the Interstate Transport State Implementation Plan Submissions for the 2015 Ozone National Ambient Air Quality Standards under Clean Air Act Section 110(a)(2)(D)(i)(I)*; 2018.
- (18) U.S. EPA, *Regulatory Impact Analysis of the Final Revisions to the National Ambient Air Quality Standards for Ground-level Ozone*; 2015; <https://www3.epa.gov/ttn/ecas/docs/20151001ria.pdf>.