NO$_x$ emissions from large point sources: variability in ozone production, resulting health damages and economic costs

Denise L. Mauzeralla,b,*, Babar Sultan$c,d$, Namsoug Kim$a$
David F. Bradford$a,c,e,f,g$

*a Science, Technology and Environmental Policy program, Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544, USA
b Department of Geosciences, Princeton University, USA
c Economics Department, Princeton University, USA
d Harvard Medical School, Boston, MA, USA
e National Bureau of Economic Research, Cambridge, MA, USA
f New York University School of Law, USA
g CESifo, Munich, Germany

Received 30 August 2004; accepted 10 December 2004

Abstract

We present a proof-of-concept analysis of the measurement of the health damage of ozone (O$_3$) produced from nitrogen oxides (NO$_x = NO + NO_2$) emitted by individual large point sources in the eastern United States. We use a regional atmospheric model of the eastern United States, the Comprehensive Air quality Model with Extensions (CAMx), to quantify the variable impact that a fixed quantity of NO$_x$ emitted from individual sources can have on the downwind concentration of surface O$_3$, depending on temperature and local biogenic hydrocarbon emissions. We also examine the dependence of resulting O$_3$-related health damages on the size of the exposed population. The investigation is relevant to the increasingly widely used “cap and trade” approach to NO$_x$ regulation, which presumes that shifts of emissions over time and space, holding the total fixed over the course of the summer O$_3$ season, will have minimal effect on the environmental outcome. By contrast, we show that a shift of a unit of NO$_x$ emissions from one place or time to another could result in large changes in resulting health effects due to O$_3$ formation and exposure. We indicate how the type of modeling carried out here might be used to attach externality-correcting prices to emissions. Charging emitters fees that are commensurate with the damage caused by their NO$_x$ emissions would create an incentive for emitters to reduce emissions at times and in locations where they cause the largest damage.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Surface ozone; NO$_x$ emissions; Point sources; Health impacts; Mortality; Morbidity; Cap-and-trade

1. Introduction

Pollution in economic theory is essentially defined by externalities, negative impacts of one economic agent on others that are not reflected in market prices. An important input to corrective policy is the size of the
externality, an estimation of the amount of damage done by the externality-causing action. In this paper, we lay out a method of estimating the damage to human health due to exposure to ozone ($O_3$) formed as a result of emission of nitrogen oxides ($NO_x = NO + NO_2$) from individual large stationary sources in the eastern United States. We also present quantitative estimates for cases selected to draw attention to sources of variation in health damage that the present thrust of $NO_x$ emissions regulation neglects. We apply the modifier “proof of concept” to the estimates to reflect the need for further refinement of the modeling components and for incorporating effects other than health damages resulting from $O_3$-exposure. The longer term objective is to associate an estimate of external costs due to emissions from any point source as a function of the prevailing weather conditions, other emissions in the region, and populations, ecosystems and materials exposed. The initial calculations suggest the possibility of substantial gains (i.e. lower total damages for a given quantity of emissions) from regulatory redesign.

$NO_x$ contributes to a wide range of environmental effects including the formation of acid rain (via formation of nitric acid), PM2.5 (via formation of secondary particulates such as ammonium nitrate) with resulting health impacts and contributions to regional haze, eutrophication of aquatic ecosystems (via addition of excess nitrogen), and elevated $O_3$ concentrations (via reaction with hydrocarbons and carbon monoxide) with resulting impacts on health and agriculture. Due to recent regulatory emphasis on controlling $O_3$ through $NO_x$ “cap-and-trade” programs, we focus here on the contribution of $NO_x$ emissions from large point sources to ambient $O_3$ concentrations and the resultant damage to human health.

“Cap and trade” is a mechanism for regulating $NO_x$ emissions, whereby a limit, or cap, is placed on total emissions from large point sources (typically during the “O3 season” from May to October each year). Regulated sources (mostly electric power generating plants) are obliged to have an allowance for each unit (ton) of $NO_x$ emitted. In practice, allowances are initially allocated to sources based on their historical operating levels. A source that does not own sufficient allowances for its needs can purchase them from others at a price agreed upon by the trading parties. The trading of allowances is expected to lower the total cost of whatever level of emissions control is embodied in the cap relative to a prescriptive type of regulatory control.

A premise of such a system is that, because the total emissions are unaffected, such an exchange has no environmental significance. However, this premise may not be realized in any given situation and when it fails an exchange of allowances may have either positive or negative environmental effects. Our damage estimates allow us to isolate instances where a shift in $NO_x$ emissions can have large effects on environmental damage.

To estimate the damages to human health resulting from $NO_x$ emissions we use a model of the chemical and transport processes in the troposphere to calculate the increase in $O_3$ concentration, at each time and location, that results from a given incremental quantity of $NO_x$ emitted from a power plant under different conditions. These $O_3$ concentration changes are combined with geographically specific demographic data to estimate the increase in mortality and respiratory illness that results from that increase in $O_3$. Specifically, we examine how temperature variations within a single summer month can have a significant effect on the quantity of $O_3$ produced from a unit of $NO_x$. We also examine how the same quantity of $NO_x$ emissions from a power plant can result in very different $O_3$ production depending on the quantity of biogenic hydrocarbons emitted in the surrounding region. Finally, we show how the health damage that elevated $O_3$ concentrations cause increases as the size of the exposed population increases.

In the section relating these exercises to the design of policy, we draw attention to the standard economic argument for presenting firms with the “right” prices of the resources they use. In this case the resource used is $NO_x$ emissions, for which, in the locations we study, the price charged is a combination of the price of emission allowances and the implicit price effect of applicable environmental regulations. The current regulatory regime diverges from one that would connect $NO_x$ emission choices with their economic consequences in at least two respects. First, the present regulatory regime gives little systematic incentive to shift $NO_x$ emissions during the summer from times and places where it has a large negative externality in the form of health impacts to times and places where the externality is smaller. Our analysis describes the benefits of considering this variability. Second, the regulatory system is focused on avoiding exceedences of limits on ambient concentrations in order to comply with the National Ambient Air Quality Standards (NAAQS). The epidemiological studies on which we base our estimates of damages indicate that health impacts increase linearly with increasing concentration (EPA, US, 2003; Steib et al., 2003). A threshold limit on ambient $O_3$ concentration is a related but distinct objective to the avoidance of mortality and morbidity that a standard economic approach would imply because it is not concerned with damages which result from exposures to concentrations below the NAAQS.

In Section 2, we provide background on the science of $O_3$ formation from $NO_x$ emissions and on the regulatory history of $NO_x$ and $O_3$. In Section 3, we describe CAMx (Comprehensive Air Quality Model with Extensions), the regional air quality model we use to quantify the $O_3$ produced from $NO_x$ emitted from point sources,
evaluate its ability to reproduce observed O3 concentrations, and describe the method we use to quantify morbidities and mortalities resulting from O3 exposure. Section 4 uses CAMx to demonstrate the differences in O3 production resulting from variability in local temperature and biogenic hydrocarbon emissions, as well as the importance of the location of high population densities relative to emission sites in determining resulting mortality and respiratory morbidity. These exercises demonstrate the potential environmental impact of an exchange of emissions from one time, place, or weather condition to another, expressed in terms of lives lost or sickness experienced. In the fifth section we estimate the monetary value of these environmental impacts, using valuation factors taken from the literature. Section 6 discusses the policy advantages of a damage based incentive system to control NOx emissions from point sources. Section 7 summarizes our conclusions and highlights needed future research.

2. Scientific and regulatory background

2.1. Ozone chemistry

O3 is a pollutant that is formed in the lower atmosphere (troposphere) from a complex series of sunlight-driven reactions between NOx, carbon monoxide (CO), volatile organic compounds (VOCs) which are largely hydrocarbons, and methane (CH4). The primary source of NOx to the troposphere is fossil-fuel combustion. In 1998 in the United States, electric utilities accounted for 25% of total NOx emissions, transportation for 53% and industrial sources for 12% (EPA, US, 2000). VOCs are emitted from a range of human activities, including fossil-fuel combustion (excepting high-temperature combustion), direct evaporation of fuel and solvents, and chemical manufacturing. Terrestrial vegetation also provides a large natural source of hydrocarbons in summer with isoprene being the biogenic hydrocarbon emitted in the largest quantity. The emission of biogenic hydrocarbons varies greatly by location as indicated in Fig. 4. Difficulty regulating O3 occurs because in regions of high NOx (primarily urban centers and power plant plumes), O3 formation is limited by the availability of hydrocarbons. In regions of low NOx (primarily rural areas with abundant emission of natural hydrocarbons), O3 formation is limited by the availability of NOx (Sillman et al., 1990).

O3 production from NOx emitted in power plant plumes varies depending on the availability of hydrocarbons, the magnitude of the NOx emission rate, and meteorological conditions. A power plant plume contains large quantities of NOx, little CO and virtually no VOCs. Hence O3 production from NOx in a power plant plume is limited by the availability of VOCs in the surrounding environment as it dilutes. Data from aircraft transects of power plant plumes indicate that higher availability of biogenic hydrocarbons, principally isoprene, result in higher quantities of O3 produced from the NOx in the plume (Ryerson et al., 2001). In addition, as the size of the plume increases, the O3 production efficiency (the number of molecules of O3 produced per molecule of NOx consumed (Liu et al., 1987)) decreases. Therefore, large concentrated plumes can result in less net O3 formation than small plumes per NOx molecule emitted because of the more rapid conversion of NOx in concentrated plumes to HNO3 (Ryerson et al., 2001). Hence, in our later analysis all of our cases use large power plants which emit similar quantities of NOx.

In addition, O3 production occurs more rapidly under conditions of stagnant high pressure which are frequently accompanied by high temperatures and sunny conditions (Logan, 1989; Vukovich, 1995). As a result, the probability of exceeding the NAAQS for O3 increases with daily maximum temperature (Lin et al., 2001). In our analysis temperature is used as a proxy for stagnant meteorological conditions.

2.2. NOx and O3 regulatory history

Legislative and regulatory efforts to control ambient O3 concentrations started with the passage of the 1970 Clean Air Act (CAA). In 1971 EPA established a NAAQS for O3 of a 1-h maximum daily average concentration not to exceed 0.08 ppm which in 1979 was revised to 0.12 ppm (EPA, US, 2004a). Based on scientific evidence of the adverse effect of O3 on health, in 1997 EPA again revised the O3 NAAQS and required that a 3-year average of the annual fourth highest daily 8-h average concentration not exceed 0.08 ppm (EPA, US, 2004a; NRC, 2004). Under the CAA, a state with monitoring stations that fail to meet the NAAQS is regarded as “out of attainment” and is obliged to specify in its State Implementation Plan (SIP) how it will attain the NAAQS in the future. No incentive presently exists to reduce O3 further if a state is in compliance. O3 has a lifetime of approximately 2 days in the boundary layer of the eastern United States in summer (Fiore et al., 2002) and hence may be transported across state boundaries. Inter-state transport can make it difficult for states that fail to meet the NAAQS to become compliant by simply reducing in-state emissions.

Economists have been encouraging the use of market-based policies for controlling pollution as an alternative to more traditional command-and-control approaches (such as uniform emission standards or specific technology requirements) for some time. In the late 1960s, discussion of the concept of markets in pollution rights began in academic circles and entered the United States.
policy world in 1975 (Atkinson and Tietenberg, 1987). Economic theory predicts that tradeable emission permits will induce emitters to find a cost-minimizing emission allocation when a cap on total emissions is specified (NRC, 2004). Emitters that can reduce emissions easily do so and sell permits to those for whom reducing emissions is more costly.

In an attempt to control regional O₃ concentrations and to address inter-state transport of O₃ and its precursors to facilitate the compliance of down-wind states with the O₃ NAAQS, a coalition of 13 north-eastern states formed the Ozone Transport Commission (OTC) and created a regional NOₓ cap-and-trade program. The OTC “NOₓ Budget program” ran from 1999 to 2002 during which time the total NOₓ emissions from large point sources within the 13-state coalition were reduced approximately 50% from 1990 baseline levels (EPA, US, 2004b). The period of NOₓ emission reduction from power plants coincided with a period of economic growth during which emissions from other sources (ex. vehicles) may have increased. We compared surface NO concentrations from the EPA-AIRS measurement network between 10 a.m. and 2 p.m. over the summer season pre- and post- the 1999 cap, and do not find the expected decrease in NO concentrations. Nor does a similar comparison of O₃ concentrations over the summer season pre- and post- the 1999 cap show a decrease in O₃ concentrations resulting from a decrease in NOₓ emissions from the power plants as a result of the cap (analysis not shown).

In 2004 the OTC program was replaced by the NOₓ SIP Call budget trading program. This program developed in response to the EPA’s call for SIPs to reduce the transport of O₃ over large regions of the country. Under the SIP Call program the trading region was expanded to 19 eastern states and total summer emissions from electric power generators are planned to be reduced by approximately 80% from 1990 levels (EvoMarkets, 2003).

The theoretical advantages of an emissions cap and trade system depend, however, on the assumption that only the total amount emitted matters. Under this condition, a shift of emissions from one time or place to another will have no environmental consequences and to address inter-state transport of O₃ and its precursors in the lower troposphere by solving the continuity equation for each chemical species in a three-dimensional grid. The Eulerian continuity equation relates the time dependency of the average species concentration in each grid cell volume to the sum of all physical and chemical processes operating on that volume.

We use the CAMX model in a configuration previously used by the Ozone Transport Assessment Group (OTAG) to solve for the concentration distribution of O₃ and its primary precursors. Model resolution is 0.33° latitude by 0.5° longitude extending in 5 vertical layers (100, 500, 1500, 2500, and 4000 m above ground level) from the surface to 4 km. CAMX includes a plume-in-grid sub-model to simulate the chemistry of NOₓ plumes emitted from point sources such as power plants. The plume-in-grid sub-model tracks individual plume segments, accounting for dispersion and inorganic chemical evolution, until the puff mass can be adequately resolved by the grid box in which it resides at which time the remaining NOₓ is released to the grid. This CAMX configuration uses the carbon bond IV (CB-IV) chemical mechanism with updated radical termination and isoprene oxidation mechanisms based on (Carter, 1996) including 25 gas species and 96 reactions. Winds are generated by the regional atmospheric modeling system (RAMS) with horizontal advection provided by the diffusive-corrected algorithm of (Smolarkiewicz, 1983). Boundary conditions are set to clean background concentrations for all constituents on all four sides and at the top of the domain (θₐ = 35 ppbv) with a day/night diurnal cycle imposed on short-lived species. Hourly surface emissions are included from elevated point sources, gridded low-level point, mobile, biogenic (from the BEIS2 model) and anthropogenic area emissions in the eastern United States. The BEIS2 model includes the dependence of isoprene emissions on temperature, solar radiation, ecosystem extent and leaf area (Geron et al., 1994; Guenther et al., 1994; Pierce et al., 1998). However (Palmer et al., 2003), inferred from GOME satellite measurements of formaldehyde that BEIS2 under-predicts isoprene emissions by a factor of two. CAMX also includes wet and dry deposition (Wesley, 1989) of gases. Emissions and meteorological fields for the model were obtained from the New York State Department of Environmental Conservation (NYSDEC) (Sistla and Hao, personal communication, 2001).
Simulations were conducted for 7–17 July 1995, with the first 2 days of the simulation used for model spin-up.

3.2. CAMx model/EPA-AIRS O3 data comparison

To evaluate the ability of this version of CAMx to accurately simulate ambient O3 concentrations within the region, we compared concentrations of O3 measured as part of the EPA-AIRS surface measurement network with O3 concentrations obtained from the CAMx simulation. Locations of the EPA-AIRS O3 measurement sites for July 1995 are shown in Fig. 1. Fig. 2 shows a comparison of 1-h daily maximum O3 concentrations calculated by the CAMx model for 9–17 July 1995 and measured on the same dates by the EPA-AIRS network. A comparison of 1-h maximum concentrations is shown because the epidemiological dose-response functions used later in this paper to estimate both mortality and morbidity use this value.

Statistical analysis of the results in Fig. 2, with \( P_i \) indicating the predicted CAMx O3 values and \( O_i \) indicating the observed EPA-AIRS O3 concentrations, result in the following summary statistics:

Root mean square error (RMSE)

\[
\text{RMSE} = \left[ \frac{1}{N} \sum_{i=1}^{N} (P_i - O_i)^2 \right]^{1/2} = 19.6 \text{ ppb},
\]

Mean bias (MB)

\[
\text{MB} = \frac{1}{N} \sum_{i=1}^{N} (P_i - O_i) = 0.62 \text{ ppb},
\]

Mean normalized bias (MNB)

\[
\text{MNB} = \frac{1}{N} \sum_{i=1}^{N} \frac{(P_i - O_i)}{O_i} = 6.0\%.
\]

This evaluation indicates that although there is random error in the maximum daily O3 concentrations calculated by CAMx the model is without significant bias. A least squares regression of the data in Fig. 2 has a slope of 0.65 (0.01, 56) and intercept of 26.02 (0.91, 29) with standard error and \( t \) value indicated in parentheses and with \( R^2 = 0.42 \).

3.3. Model simulations

The two goals of the CAMx simulations presented in this paper are first to demonstrate the spatial and temporal heterogeneity of tropospheric O3 production and the variability in resulting health damages (morbidity and mortality) that may result from a given amount of NOx emitted from the same location under different meteorological conditions or from different locations with varying biogenic hydrocarbon emissions or downwind populations. The second objective is to provide a proof-of-concept prototype of the feasibility of conducting a full integrated assessment that links NOx emissions to O3 concentrations to human exposure, to morbidity, mortality and damage valuations.

To demonstrate the spatial and temporal heterogeneity of tropospheric O3 production, we conduct a set of
CAMx regional simulations for seven cases. First, we conduct a standard simulation that includes all regional emissions. This is followed by six perturbation simulations. In order to quantify the impact of NO\textsubscript{x} emissions from individual power plants on O\textsubscript{3} concentrations, we reduce the NO\textsubscript{x} emissions from each plant by 62.5 tons (125,000 lbs) in a 24-h period (1.77 \times 10^6 moles where NO\textsubscript{x} emissions from power plants were assumed to be 90% NO and 10% NO\textsubscript{2}) and calculate the difference in maximum O\textsubscript{3} concentration between the standard simulation and each perturbation simulation. The power plants used for this experiment are among the largest 5% of power plants in the United States each of which emits approximately 100 tons of NO\textsubscript{x} per day; 62.5 tons of NO\textsubscript{x} was chosen for the analyses because it is sufficiently large to generate a discernable effect and sufficiently small that a single large power plant may emit this quantity or more in a 24-h period. The perturbation simulations were chosen to demonstrate the variability in O\textsubscript{3} production resulting from identical NO\textsubscript{x} emissions from power plants produced during conditions of high and low temperature at the same location and in locations of high and low biogenic hydrocarbon emissions. They were also chosen to demonstrate the dependence of total health effects on the population density in the regions downwind of the emissions. Figs. 3, 5 and 7 show the increase in maximum surface O\textsubscript{3} concentrations over the model domain between the standard simulation and perturbation simulations summed over the day of emission and the subsequent 3 days. After 4 days the power plant emissions from the 24-h period have a non-discernable effect on regional O\textsubscript{3} concentrations. The total increase in 1-h maximum O\textsubscript{3} is shown because it is incorporated in the dose-response functions used to estimate the mortality and morbidity that result from exposure to O\textsubscript{3}.

3.4. Estimating mortality and morbidity

The relationship between O\textsubscript{3} exposure and resulting health effects can be estimated from epidemiological studies. Most epidemiological research has made use of time-series analyses that have focused on the effects of short-term exposures to a pollutant by using records of hospital admissions and deaths that occur following periods of elevated concentrations. These studies are unable to include the negative effects of chronic or long-term exposure and hence presumably underestimate the total health damage due to exposure. Cohort studies, which follow a population over time and do include both acute and chronic impacts, have been conducted for PM2.5 (Dockery et al., 1993; Pope et al., 1995, 2002) but have not focused on O\textsubscript{3}. Hence, in relying in this analysis on time-series studies only (specifically, on meta-analyses that pool the results of many regional time-series studies (Stieb et al., 2003) we use an arguably conservative estimate of the mortality and respiratory morbidity that results from elevated O\textsubscript{3} exposure.

The change in the incidence of mortality or morbidity due to exposure to air pollution (here O\textsubscript{3}) is estimated with the following function (EPA, US, 1999)

$$\Delta M = Y_0(1 - e^{-\beta \Delta O_3}) \times \text{population.} \quad (1)$$

Here $\Delta M$ is the change in the number of mortalities during a 24-h period, or change in incidence of hospital admissions for respiratory disease, relative to a baseline number of deaths or hospital admissions for respiratory disease. For mortality, $Y_0$ is the baseline incidence of daily non-accidental deaths per person of any age. For morbidity, $Y_0$ is the daily average hospital admission rate for diseases of the respiratory system. Epidemiological studies often report a relative risk (RR) (the percentage increase in mortality or morbidity) due to a specific increase in pollutant concentration. The concentration-response coefficient, $\beta$, can be derived from the reported RR and the change in pollutant concentration relative to a baseline (here $\Delta O_3$), by solving:

$$\beta = \frac{\ln(RR)}{\Delta O_3}. \quad (2)$$

We discuss the $\beta$ values we use for mortality and morbidity in the following subsections. We obtain the population in Eq. (1) from 1995 world population data gridded at 0.04167 latitude by longitude resolution (CIESIN et al., 2000). Population data for the eastern United States is regridded at the CAMx resolution of 0.3° latitude by 0.5° longitude (approximately (36 km)\textsuperscript{2} in the United States). The population within each CAMx grid box is used to calculate the mortality and morbidity that result from the change in O\textsubscript{3} concentrations in each grid box. Mortality and morbidity incidence is summed over all affected grid boxes over all days in which the O\textsubscript{3} concentration was enhanced due to (or, in some locations, reduced by) the NO\textsubscript{x} emitted from the individual power plants.

3.4.1. O\textsubscript{3} mortality

To obtain the RR of mortality for an increase in O\textsubscript{3} concentration resulting from individual power plant emissions, we use a pooled estimate of percent excess mortality derived from a compilation of 25 epidemiological time-series analyses of daily 1-h maximum O\textsubscript{3} concentrations during the warm season (Stieb et al., 2003). The pooled estimate of RR for single pollutant studies is 1.014 for a 31.2 ppb increase in 1-h maximum daily O\textsubscript{3} concentration, which translates into a $\beta$ value of 0.000446 (Stieb et al., 2003). We estimate $Y_0$ from the total population in 1995 and calculate an average daily mortality rate of 2.4 \times 10^{-5} (2,312,132 deaths/261,638,000 persons/365 days) (CDC-NCHS, 2003; USCB, 2000).
Fig. 3. $\text{O}_3$ produced by an identical amount of NO$_x$ emitted from the same power plant in eastern Pennsylvania in a 24-h period on (a) 9 July 1995, a cool day ($296 \pm 4.3 \, ^\circ\text{K}; 72 \pm 7.7 \, ^\circ\text{F}$) and (b) 14 July 1995, a warm day ($302 \pm 3.5 \, ^\circ\text{K}; 83.5 \pm 6.3 \, ^\circ\text{F}$) and the subsequent 3 days until $\text{O}_3$ production from these NO$_x$ emissions cease. The “X” in each figure marks the power plant location. The ‘average increase in 1-h maximum $\text{O}_3$’ is the total increase in maximum surface $\text{O}_3$ concentrations over the model domain between the standard simulation and perturbation simulations on the day of emission and the subsequent 3 days. Beyond 3 days the power plant emissions cease to have a significant influence. Here (c) and (d) show the mortalities and (e) and (f) show the morbidities that are estimated to result from the increase in daily maximum $\text{O}_3$ occurring in (a) and (b), respectively.
3.4.2. \(O_3\) morbidity

Hospital admissions for respiratory ailments have been found to be substantially and consistently associated with \(O_3\) exposure. A synthesis of \(O_3\) hospital admission time-series studies, compiled by Thurston and Ito (1999) is used here to estimate the effect of \(O_3\) on morbidity. Data on hospital admissions for respiratory disease in the United States were not available. Therefore, to calculate \(Y_0\) we assume that the average daily rate of hospital admissions for respiratory disease is approximately equal to the rate of hospital discharge following a diagnosis of a respiratory ailment which was 3.41 \(\times 10^{-5}\) discharges per day in 1998 (124.6 discharges/10,000 population/365 days), the year for which such data are available (CDC-NCHS, 1998). This value is of similar magnitude to values obtained in regional studies of daily hospital admission for respiratory diseases (EPA, US, 1999). The estimate for hospital admissions for major respiratory ailments due to \(O_3\) exposure for all ages was a RR = 1.18 per 100 ppb increase in daily 1-h maximum \(O_3\) (95% CI, 1.10–1.26) (Thurston and Ito, 1999) which translate to a \(\beta\) value of 0.00166. In cases where studies in the meta-analysis used other measures of \(O_3\) exposure besides a daily 1-h maximum concentration (such as the 8-h average \(O_3\) concentration) the ratio of the daily 1-h maximum to the other measures used in the original analyses were used to compute the increment comparable to a 100 ppb change in daily 1-h maximum \(O_3\) concentrations (Thurston and Ito, 1999).

4. Results

Our results demonstrating the impacts of variability in temperature and biogenic hydrocarbon emissions on \(O_3\) production, and the importance of population density in determining resulting mortalities and respiratory morbidity are summarized below. These factors were chosen for examination because of the expected size of their influence and because of the ability of our model to simulate them. Additional factors may also influence \(O_3\) production from large point sources such as the concentration of \(NO_x\) in the plume, cloudiness, time of day of emission, proximity of the \(NO_x\) emission to other anthropogenic emissions, and other factors. Analysis of these additional factors is left for the future.

4.1. Temperature variability—effect on \(O_3\) production and resulting mortality and morbidity

First, to examine the temporal heterogeneity of tropospheric \(O_3\) production due to temperature variability within a short period of time, a point source was selected which experienced a range of temperatures between 9 and 17 July 1995. The effects on \(O_3\) production of identical emissions of \(NO_x\) from a power plant in eastern Pennsylvania on 9 July 1995 and 14 July 1995 when the 24-h period temperature ranged over 72 ± 7.7°F and 83.5 ± 6.3°F, respectively, are shown in Figs. 3a and b. The average increase in 1-h maximum surface \(O_3\) over the model domain between the standard simulation and perturbation simulations on the day of emission and the subsequent 3 days, was 0.021 ppbv during the cool period and 0.036 ppbv during the warm period. The model predicts that the same amount of \(NO_x\) emitted in the warmer period results in more \(O_3\) which leads to twice the mortalities as in the cooler period—0.185 versus 0.354 persons as shown in Figs. 3c and d. Likewise, the number of cases of respiratory morbidity due to \(O_3\) exposure was also nearly twice as high during the warm period—0.866 versus 1.654 cases as is shown in Figs. 3e and f. Due to the use of the concentration response function for \(O_3\), calculated mortalities and morbidities isolate the effect of increased \(O_3\) and do not include possible direct effects of increased temperature on mortality and morbidity.

These simulations show that, in violation of the preconditions for the use of a cap and trade system, a shift of a fixed amount of \(NO_x\) emissions from a single source between two closely separated time points could have a substantial effect on the quantities of \(O_3\) that are produced, and the consequent health effects. In fact, since energy demand is usually larger when temperatures are high, emissions are also likely to be larger during periods of high temperature when the model indicates that more \(O_3\) is produced per unit \(NO_x\) emitted. An emitter using emission allowances at two different times presently has no reason to favor the time when the resulting damage is lower. Trading allowances (or variability in the timing of the use of allowances from a single source) to emit at a price that does not reflect the difference in damage will lead to too much emission in the high damage conditions relative to the low damage conditions. The question is naturally raised, whether an alternative regulatory system might be developed that provides an appropriate incentive for emitters to reduce their emissions during times of elevated temperature within the O3 season.

4.2. Biogenic hydrocarbon emission variability—effect on \(O_3\) production and resulting mortality and morbidity

Isoprene is a hydrocarbon that is naturally emitted from deciduous tree species, particularly oak, poplar and gum trees. In the United States, these species are most abundant in the south-eastern part of the country and are absent in most of the west. This results in large spatial variability in biogenic VOC emissions in the United States, as is shown in Fig. 4. In addition,
isoprene emissions increase with increasing temperature and sunlight. Biogenic hydrocarbon emissions reach a maximum at approximately 40 °C and decrease following a bell-shaped curve on either side of 40 °C (Pierce et al., 1998). A unit of NOx emitted from a power plant in a region of high isoprene emissions can result in larger O3 production than the same amount emitted in a region of low isoprene emissions (Ryerson et al., 2001). To examine the impact that biogenic hydrocarbon emissions, particularly isoprene, can have on O3 production, we examined the quantity of O3 produced from identical emissions of NOx by two different power plants, one in the south (Mississippi) and the other in the northern mid-west (Indiana) of the United States. To control for the effect of temperature on isoprene emissions and on O3 production, we chose 24-h periods during which the mean temperature at the two sites was nearly identical: 87.2 °F ± 5.9 °F in the high-isoprene region and 87.9 °F ± 6.3 °F in the low-isoprene region. The regional average increase in 1-h maximum O3, as shown in Figs. 5a and b, was 0.033 ppbv in the region with low emissions and over three times higher at 0.109 ppbv in the region of high emissions. Due to the apparent factor of two under-prediction of isoprene emissions by BEIS2 (Palmer et al., 2003), the quantity of O3 produced in our simulation in the high-isoprene region may be under-predicted.

In this example, the region of higher isoprene emissions was also a region of lower population density (see Fig. 6). Despite this fact, the 3.5 times larger quantity of O3 produced in the region of higher isoprene emissions resulted in both 30% more mortalities (0.166 versus 0.121 deaths) and morbidities (0.774 versus 0.568 cases) than resulted from the same increase in NOx emissions in the region of lower isoprene emissions (see Figs. 5c–f). By choosing locations where the region of high isoprene corresponds to a region of lower population, this scenario is intentionally conservative and is hence understating the potential importance of variability of isoprene concentration on resulting O3 mortalities.

A regulatory scheme that ignores the spatial variability of biogenic hydrocarbon emissions when regulating NOx may miss an opportunity to decrease O3 concentrations and could inadvertently increase mortalities and morbidities by trading from a region of relatively low isoprene emissions (north-east) to high emissions (south-east). In particular, trades between regions under a cap-and-trade program in which all NOx emissions from large point sources in a region are treated identically, could have a substantial effect (up or down, depending on the direction of the exchange) on the quantities of O3 that are produced, and the consequent health effects. A regulatory system that provides incentives for emitters to reduce their emissions in regions of high isoprene emissions would be more effective at reducing ambient O3 concentrations (Ryerson et al., 2001). If these reductions coincided with high population densities the reductions would also be more effective at reducing adverse effects on exposed populations than would a regulatory program that assumes emissions in all locations have identical effects. We discuss the effect of downwind population density next.

4.3. Downwind population density—effect on resulting mortality and morbidity

Population density in the eastern United States is shown in Fig. 6. The north-eastern Washington DC to Boston corridor has among the highest population densities in the United States. To examine the impact that equivalent reductions of NOx would have from point sources upwind of regions of high and low population density, point sources in the high population density northeast (Maryland) and lower-population density mid-Atlantic region (North Carolina) were selected. Average isoprene emissions are substantially larger in North Carolina than in Maryland, however. Thus, despite controlling for temperature, the average increase in 1-h maximum O3, as shown in Figs. 7a and b, was more than twice as great (0.081 ppbv versus 0.035 ppbv) when emissions originated from lowly populated North Carolina (Fig. 7a) than from highly populated Maryland (Fig. 7b). Despite higher maximum O3 concentrations in the low population region, the resulting incidence of mortalities shown in Figs. 7b and d (0.592 versus 0.268 deaths) and respiratory morbidity shown in Figs. 7c and e (2.771 versus 1.252 cases), is much lower. Thus, for the same NOx emission and a smaller O3 increase, the damage (i.e., higher mortality and morbidity) was substantially greater in the high population region.

Regulation of NOx emissions that ignores the variability of downwind population density misses an opportunity to decrease the health effects of emissions.
As in the case of variations in biogenic hydrocarbon emissions from place to place, trades between regions under a cap-and-trade program in which all NO\textsubscript{x} emissions from large point sources in a region are treated identically could have a substantial effect (up or down, depending on the direction of the exchange) on...
health outcomes. A regulatory system that provides incentives for emitters to reduce their emissions in regions of high downwind population density would be more effective at reducing the adverse effects on exposed populations than would a regulatory program that assumes emissions in all locations will have identical effects. Variations in daily wind direction can determine whether a particular NO\textsubscript{x} source is upwind or downwind from a large population center.

Ideally, a chemical weather forecasting system integrated with a health damage model could be used to predict a few days in advance what the likely damage that NO\textsubscript{x} emitted today would cause. Emitters could then take that information into account when making decisions as to whether to control their emissions, import power from elsewhere, or purchase or sell permits to cover the damage that their emissions were projected to cause.

5. Monetizing health costs resulting from O\textsubscript{3} produced from point source NO\textsubscript{x} emissions

In the standard economic analysis of externalities, the problem that externalities cause is an inefficient allocation of resources. Much as a producer of electricity would be expected to use “too much” coal if it were provided free of charge (and as a result, in a competitive system, the price of electricity would be “too low”), a producer of electricity who is charged nothing for NO\textsubscript{x} emissions may be expected to emit “too much” NO\textsubscript{x}. A standard corrective in the economic theory of pollution is to determine the “right” price for the unpriced input (here NO\textsubscript{x} emissions), and to charge the firms that use NO\textsubscript{x} emissions in production via a fee or tax. The argument is that the producing firm “should” pay the price for these inputs, just as the firm pays for other inputs, not as a moral matter but as a condition for avoiding a use of resources that result in a loss of potential value to the ultimate consumers of goods and services and to the general public.

Our analysis indicates potentially large differences in the damage done by incremental NO\textsubscript{x} emissions from sources at different locations and under different regional VOC and meteorological conditions, where the impacts have been expressed in terms of change in mortality and morbidity risks. The logic of correcting a negative externality calls for expressing these effects in monetary terms, so that they can be compared with the cost of corrective actions. In the case of health impacts, the needed valuation is the willingness of people to pay for reduced risks of suffering them.

A large literature addresses the estimation of the monetary value of health effects. A deeper analysis than the one we undertake here would consider issues such as whether the willingness to pay for reduced risk of health effects depends on the age of the affected people, their incomes, number of dependents, etc. For purposes of the present analysis, we take off-the-shelf figures for the value of a statistical life or an illness. Mortality valuation estimates, obtained from 26 different studies that use both contingent valuation and labor market valuation methods to approximate the willingness to pay to avoid the occurrence of a premature death, give values of a statistical life that range from $0.6 million–$13.5 million with a mean of $4.8 million in 1990 US dollars (EPA, US, 1999; Viscusi, 1992). The model simulations represent O\textsubscript{3} production during July 1995, so in order to calculate the value of the damages caused by the NO\textsubscript{x} emissions in our 1995 simulations in contemporary terms, we use the 17% increase in the consumer price index (CPI) between 1990 and 1995 to adjust the mean value of a death to 1995$5.6 million (DoL, 2004). Due to the increase in CPI between 1995 and 2004 these same deaths, had they occurred in 2004, would have a valuation 24% higher than the 1995 value (DoL, 2004).

For the specified identical quantities of NO\textsubscript{x} emitted from individual power plants over 24-h periods, mortalities resulting from the O\textsubscript{3} produced over the subsequent several days ranged from a minimum of approximately 0.1 deaths in the low-isoprene region to 0.6 deaths when emitted upwind from a region of high population density. Using the mean inflation-adjusted value of 1995$5.6 million/life, the monetary cost of the deaths in our six cases varies between a low of $670,000 to a high of $3.3 million. The NO\textsubscript{x} emission change (62.5 tons) that gives rise to these impacts is comparable to shutting down a large power plant for a day. For an
Fig. 7. O₃ produced by an identical amount of NOₓ emitted on 10 July 1995 from a power plant (a) upwind from a relatively low population region and (b) upwind of a high population region. The mean temperature at the two sites was nearly identical: 83.9 ± 6.3 °F in the high population region and 83.4 ± 5.8 °F in the low population region. The rest of the notation is as described in Fig. 3.
emission of this magnitude, the estimated value of the impact per incremental ton emitted ranges in the six cases from $10,700 to $52,800.

The literature on the cost of morbidity due to an illness indicates substantial variation depending on the illness. Chronic bronchitis is usually estimated to have the highest cost and a restricted activity day among the lowest. We utilize a morbidity dose-response function to estimate the RR of a hospital admission due to an increase in daily 1-h maximum O\textsubscript{3} concentrations. We therefore use a health valuation for all respiratory hospital admissions of 1990$6100/case (EPA, 1997) which adjusted for inflation is 1995$7020/case. Thus, for the identical quantities of NO\textsubscript{x} used in our exercises, emitted from individual power plants over 24-h periods, the calculated morbidities range from 0.57 cases for the low-isoprene region to 2.8 cases for the high population region. The implied estimated costs of total respiratory effects range from 1995$19,700 or from $64 to $315 per ton NO\textsubscript{x} emitted. Given valuations found in the literature, the morbidity impacts of NO\textsubscript{x} emissions are tiny when compared with the mortality risk effects.

6. Policy implications

The implications of our investigation for policy design are potentially wide-ranging. Focusing only on the mortality effects, which our calculations suggest are of greater policy significance, the estimated impact of an incremental ton of emitted NO\textsubscript{x} ranges from 0.0019 to 0.00095 fatalities (a factor of nearly 5) depending on ambient temperature and emission location. By shifting 11 tons of NO\textsubscript{x} emissions per day from a (relatively) high damage to lower damage location over a 10-day period one could avoid the loss of approximately one life (on average).

The figures also suggest a significant influence of meteorological conditions on the health impact of incremental NO\textsubscript{x} emissions. The mortality impact of emissions from a single location may vary by a factor of nearly two as the temperature varies within a short span of time. In our simulations we calculated 0.00566 and 0.00296 deaths per incremental ton of NO\textsubscript{x} emitted for the high and low temperature scenarios, respectively. Reducing O\textsubscript{3} generation and thus saving human lives might be brought about, for example, by reallocating electricity production in space to reduce NO\textsubscript{x} emissions under high-temperature stagnant weather conditions and by utilizing pollution control devices more intensely in locations of high population exposure. The use of a chemical weather forecast system would assist in predicting damages given current weather conditions.

Placing monetary value on the damages that result from incremental NO\textsubscript{x} emissions (taking into account, as we do here, only the contribution to downwind O\textsubscript{3} formation and neglecting the health damages from secondary PM formation) suggests that the magnitudes are significant in relation to the costs of abating the emissions. Still considering only fatality effects, and using mean off-the-shelf estimates of the value of a statistical life saved, the estimated damage per incremental ton of emissions in our scenarios ranges from 1995$10,700 to 1995$52,800. To get a sense of what these figures mean, we compare them to the price of NO\textsubscript{x} emission allowances under the SIP Call cap and trade system. One allowance permits an affected source to emit 1 ton of NO\textsubscript{x} between 1 May and 30 September for a given year. As reported at http://www.evemarkets.com/ on 22 July 2004, the last price for NO\textsubscript{x} emission allowances for 2004 was $2550, for 2005, $3750, and for 2006, $3200. Using a CPI adjustment of 0.8 (DoL, 2004), these prices are equivalent to 1995$2040, 1995$3000, 1995$2560, respectively. These figures can be read as measures of the cost of incremental NO\textsubscript{x} abatement during the O\textsubscript{3} season (given the existing regulations governing power plants, which have complex implications for the marginal cost of abatement). Economic theory predicts that the marginal cost of abatement of NO\textsubscript{x} will be driven to the market price of allowances, with the important proviso that the marginal cost may exceed the allowance price for all levels of abatement. Although the allowance prices cannot be directly compared with our estimate of the marginal damage due to NO\textsubscript{x} emissions (since both the year and the underlying emissions picture is different), the figures certainly suggest the possibility that net gains are possible from more stringent, and as significantly, more effective, NO\textsubscript{x} control.

The textbook remedy for a defect in the allocation of resources that results in an environmental externality is a corrective tax or fee (usually called a Pigovian tax after the British economist A.C. Pigou, who first elaborated the idea) on the action that gives rise to the externality. The damage done by a marginal unit of NO\textsubscript{x} emissions under a variety of conditions would, under ideal competitive conditions, be exactly the emission fees called for in theory to correct for the polluting effect. Taking those fees into account would induce NO\textsubscript{x} emitters to make decisions about how much, when, and where to emit that are “correct” in the usual sense used in economics, that is, in the same sense that the wages of workers and prices of fuel induce correct decisions about the quantity of these goods and services to use. The marginal damages from our calculations can be thought of as preliminary estimates of the needed Pigovian fees.

An important question, which we do not address here, is how much difference it would make if NO\textsubscript{x} emitters had to pay for the damage they cause. In economists’ jargon, a lot depends on elasticities. If, for example, raising the price of electricity to reflect the health costs
implied by our analysis would not affect the amount of electricity demanded (the price elasticity of demand is low); if raising the cost to producers of emitting \( \text{NO}_x \), perhaps varying it substantially over short periods of time, would have little effect on their decisions about how much or where to produce (the price elasticity of producer demand for \( \text{NO}_x \) emissions as an input to producing electricity is low), then the substitutions of the sort described above as possible tradeoffs would not occur, or would not occur to significant degree. If, however, as experience suggests is true in most areas of economic activity, there is substantial flexibility in the time and place of production, and substantial flexibility in the commercial and household use of electricity, especially given time to adjust to a different price regime, then large gains are to be had from a regulatory regime that provides the “right” incentives to respond to differences at the margins we have identified. These differences include: meteorological conditions from the perspective of a given source, differences in the biogenic VOC emissions in source areas, and differences in the population density downwind of emission sources.

The tendency in the recent evolution of \( \text{NO}_x \) regulation has been to rely on cap and trade systems. In particular, under the recent OTC and \( \text{NO}_x \), SIP Call trading regimes implemented in the northeastern United States, fixed totals of emissions are set for the entire \( \text{O}_3 \) season. Economic theory predicts that trading among sources will tend to equalize the marginal cost of abatement across emitters, where abatement is measured by reductions in emissions at any point in the \( \text{O}_3 \) season. This system misses the distinctions identified in this paper that determine the downstream health impacts of marginal emissions: differences in meteorological conditions within the \( \text{O}_3 \) season, variations in biogenic hydrocarbon emissions and downwind population densities. Variations on cap and trade systems that would address these shortcomings are possible. For example, a modified cap-and-trade program with “zones” could be adopted in which trading in only one direction is permitted (e.g. from regions upwind to downwind of high population centers, from regions of high to low emissions of isoprene, etc.). The same sort of modeling work as is carried out in this paper could be used in the design of such alternatives to guide producers to reduce their emissions at times and in places where the emissions cause the most damage.

7. Conclusions and future work

In this paper, we present a proof-of-concept use of an integrated assessment that combines an atmospheric model with economic and demographic information to estimate the health damage externality due to ozone (\( \text{O}_3 \)) formation from nitrogen oxide (\( \text{NO}_x = \text{NO} + \text{NO}_2 \)) emissions from large point sources. We use a regional atmospheric model of the eastern United States (CAMx) to examine the variation in the amount of \( \text{O}_3 \) produced as a result of locating, at times and places with strategically chosen differences in temperature and local biogenic hydrocarbon emissions, a fixed reduction in the nitrogen oxides emitted from power plants within a 9-day period in July 1995. Results of the CAMx model indicate that for the same \( \text{NO}_x \) emission the \( \text{O}_3 \) produced can vary by more than a factor of five. In addition, we showed that the variation in health damages that the resulting \( \text{O}_3 \) causes depends strongly on the size of the exposed population. Within our span of days and conditions examined, variation in \( \text{O}_3 \) production and downwind population can together result in a factor of six difference in resulting mortalities for an identical change in the quantity of \( \text{NO}_x \) emitted.

The increasingly widely adopted emission cap-and-trade approach has been very successful at reducing total \( \text{NO}_x \) emissions from large point sources (EPA, US, 2004b). However, because it does not control for the location or time during the summer that emissions take place nor for the resulting damages, it is less successful at minimizing the damages that result from the emissions permitted under the cap. To reduce total damage, consideration could be given to a system of fees for emitters to provide incentives for them to reduce \( \text{NO}_x \) emissions at times and in locations where health damages are greatest. This could be achieved by coupling a regional atmospheric chemistry model that calculates \( \text{O}_3 \) concentrations obtained from \( \text{NO}_x \) emitted from individual point sources with estimates of resulting health damages and their economic cost. Such a system would involve a transformation of air quality regulation as it would create an incentive system to reduce total damages in an economically efficient manner rather than simply requiring a reduction in total emissions or the achievement of a particular uniform air quality goal regardless of location or resulting damage.

Although we have suggested the possibility of using model-derived estimates for purposes of setting emission fees, the details of such a system merit examination. Since the fees are to depend on weather and other conditions, would they be set after the fact (making the emitters, in effect, use the model in forecast mode to make their decisions), or would forecasts be used by the regulator to set fees in advance (based on expected damages)? Regardless of which approach were used, refinements in chemical weather forecasting systems would clearly be beneficial to enable the regulatory system to estimate more accurately appropriate corrective fees, providing emitters with the information they need to adjust their power production and \( \text{NO}_x \) emission plans to minimize total costs which include mortality and morbidity risks as well as abatement measures.
In addition, to account more fully for the health effects of power plant NOx emissions it is necessary to include the formation of secondary particulates (PM) and their impact on mortality and morbidity. Secondary PM, in addition to O3, are formed from NOx emissions in winter as well as summer. Using an integrated assessment model without chemistry, analyses of the benefits of an annual (rather than just summer) NOx emissions cap were found to yield substantial net benefits due to reduction of PM2.5 concentrations and its detrimental health effects (Burtraw et al., 2001, 2003). Additional analyses that include both gas and aerosol chemistry and an examination of the effects that spatial and temporal variability in emissions have on PM2.5 formation and resulting health damages and costs would be useful. Such an analysis would permit a determination of the total health damage NOx emissions cause which would in turn allow a calculation of the fees that an emitter might be charged for the total damage its emissions cause.

Acknowledgments

We thank Winston Hao and Gopal Sistla of the New York State Department of Environmental Conservation for providing us with the version of the CAMx model used by the Ozone Transport Commission. We thank Arlene Fiore, Sarah Fox, Don Fullerton, Michael Greenstone, Michael Herz, Larry Horowitz, Jesse Rothstein, Wolfram Schlenker, Quansong Tong, and Jason West for helpful comments on earlier versions of this manuscript. We are pleased to acknowledge the Woodrow Wilson School of Public and International Affairs and the Center for Economic Policy Studies at Princeton University for financial support.

References


