Designing markets for pollution when damages vary across sources: What are the gains from differentiation?

Meredith Fowlie and Nicholas Muller*

December 9, 2011

Abstract

A majority of the air pollution currently regulated under U.S. emissions trading programs is non-uniformly mixed, meaning that health and environmental damages depend on the location and dispersion characteristics of the sources. Most emissions trading programs ignore this fact. Emissions are penalized at a single permit price, regardless of the location of the source. In theory, differentiated policies can be designed to accommodate non-uniformly mixed pollution using emissions penalties that vary with emissions damages. We present a simple framework to illustrate the theoretical gains from differentiation in first-best and second-best settings. This serves as foundation for a detailed analysis of the gains from differentiation in a major U.S. emissions trading program. We take two complementary approaches to estimating these gains. Our preferred estimate, which is generated using an econometrically estimated model of firms' compliance choices, is surprisingly small given the extent of the variation in damages across sources. A more standard approach to simulating of policy outcomes, one that assumes strict cost minimization on behalf of all firms, predicts larger gains. A comparison of the two approaches provides insights into the determinants of the gains from policy differentiation, some of which are inadequately captured by standard policy simulation models.

Keywords: Market-based Policy, NO_x Budget Program, Policy Instrument Choice.

JEL Classifications: Q54, Q53, Q58

*Fowlie: Department of Agricultural Economics, UC Berkeley, and NBER. Muller: Department of Economics and the Environmental Studies Program, Middlebury College. We thank Larry Goulder, Jacob LaRiviere, Steven Puller, Christian Vossler, and seminar participants at Carnegie Mellon, Resources for the Future, University of Pittsburgh, the University of California Energy Institute, the University of Tennessee, and the US EPA for helpful discussions and comments. We thank Dilek Uz for invaluable research assistance. Muller thanks the United States Environmental Protection Agency for support under award: EPA-OPEI-NCEE-08-02.

1 Introduction

Economists have long advocated for market-based approaches to pollution regulation (Montgomery, 1972; Baumol and Oates, 1988). The past three decades have witnessed large scale experimentation with emissions trading implementation. By many measures, this experimentation has been very successful. Targeted emissions reductions have been achieved or exceeded, and it is estimated that total abatement costs have been significantly less than what they would have been in the absence of the trading provisions (Keohane, 2008; Stavins, 2005)

In terms of economic efficiency, however, many existing cap-and-trade programs likely fall short of the theoretical ideal. Efficiency requires that the marginal cost of pollution reduction be set equal to the marginal damage caused by pollution (Baumol and Oates, 1988; Montgomery, 1972). If pollution is "non-uniformly mixed", (i.e. health and environmental damages from pollution depend on the location of the source), marginal costs of pollution abatement should vary across sources according to the degree of damage caused (Muller and Mendelsohn, 2009). Most policies are currently implemented as spatially uniform, "undifferentiated" emissions trading programs meaning that regulated emissions are penalized at the same permit price. Under an undifferentiated program, marginal abatement costs should be set equal across sources, thereby minimizing the costs of meeting the emissions cap. Importantly, this outcome cannot be efficient when pollution is non-uniformly mixed.

Some of the most pernicious air quality problems in the United States involve non-uniformly mixed pollutants. Examples include nitrogen oxides (NOx) and sulfur dioxide, two criteria pollutants currently subject to undifferentiated market-based regulation. As policy makers work to design the next generation of policies to control these pollutants, the question of how to design market-based emissions trading programs when damages from pollution depend significantly on the location of the source has emerged as a major point of contention.[1] This paper investigates the potential efficiency gains from implementing so-called "differentiated" policies with emissions penalties that vary across damage-differentiated sources.

In theory, market-based policies can be designed to perfectly accommodate non-uniformly mixed pollution (Montgomery, 1972; Tietenberg, 1980; Muller and Mendelsohn, 2009). Baumol and Oates (1988) use a general equilibrium model to depict optimal pollution taxes in a setting with heterogeneous costs and damages. The optimal tax rate is calibrated to the marginal damage caused by emissions. When damages vary by source, so do the tax rates. Others have proposed so-called "differentiated" emissions permit market designs wherein differences in marginal social cost are reflected in different compliance requirements and incentives (Mendelsohn, 1986; Tietenberg, 1995; Muller and Mendelsohn, 2009). In a differentiated trading program, relatively high (low) damage sources must pay relatively more (less) to offset uncontrolled emissions. For any binding emissions cap, a differentiated policy will allocate a greater proportion of the permitted emissions to sources where they cause less harm.

[1] In 2008, a federal district court vacated the Clean Air Interstate Rule and the associated regional NO_x trading program. The rule was to intended to provide a cost-effective, market-oriented approach to regulating a non-uniformly mixed pollutant (subject to compliance feasibility). On July 11, 2008, in North Carolina v. EPA, the U.S. Court of Appeals for the D.C. Circuit vacated CAIR, in large part due to policy's failure to adequately accommodate spatial transport of pollution and the associated variation in damages across sources. *State of North Carolina v. Environmental Protection Agency, No. 05-1244,* slip op. (2008), District of Columbia Court of Appeals.

We begin in section 2 by introducing a conceptual framework useful for analyzing the welfare implications of damage-based policy differentiation. As emphasized by Mendelsohn (1986), the net gains from differentiation will depend on both the extent of the variation in damages across sources and the steepness of the marginal abatement cost functions. We extend the established theoretical literature in order to accommodate two practical considerations that complicate the design and implementation of differentiated policy designs. First, we consider a setting in which a regulator seeks to minimize pollution damages plus abatement cost subject to an exogenously set cap on emissions. Importantly, this cap may not reflect the optimal aggregate limit. As such, the analysis explores constrained optimality in differentiated allowances programs. Second, we note that the policy maker is rarely, if ever, fully informed. It is typically the case that policy makers must design and implement policies in the presence of significant uncertainty surrounding estimates of pollution damages and limited information about abatement costs.

The conceptual framework serves as foundation for a detailed analysis of the gains from policy differentiation in the context of a landmark emissions trading program. Section 3 introduces the NOx Budget Program which limits nitrogen oxide (NOx) emissions from large point sources in the Eastern United States. In the design stages of this program, policy makers considered imposing restrictions on interregional trading (FR 63(90): 25902). Ultimately, it was decided that the potential benefits from this additional complexity would not justify the costs (US EPA, 1998). The program was therefore implemented as a single jurisdiction, spatially uniform trading program in which all emissions are traded on a one-for-one basis.

In the years since the NOx Budget Program was designed and implemented, debates over how to regulate non-uniformly mixed pollution have become more heated and complex.[2] With the benefit of hindsight, we revisit the decision to forego spatially differentiated NOx trading in favor of the simpler undifferentiated alternative. Section 4 describes our analytical approach. First, source-specific marginal damage estimates are generated using a stochastic integrated assessment model, AP2 (Muller, 2011).[3] These estimates are used to define the terms of compliance in counterfactual differentiated policy regimes. We model firms' compliance choices under the observed and counterfactual policy regimes. These simulated compliance choices define source-specific NO_x emissions. The integrated assessment model is re-introduced for the purpose of estimating the aggregate health and environmental impacts of the simulated emissions associated with each policy scenario. Comparisons of damages and costs across observed and counterfactual policy regimes yields an estimates of the net gains from policy differentiation.

This approach has three distinguishing features. First, we use a detailed integrated assessment model to estimate marginal damages for each facility in the NBP. We document significantly more spatial variation in damages as compared to what earlier studies of spatially differentiated NOx trading have assumed. All else equal, greater variation in damages implies increased potential gains from policy differentiation.Moreover, we find that almost half of the variation occurs within (versus between) states. Previous

[2] In 2008, a federal district court vacated the Clean Air Interstate Rule which was to subsume the NOx Budget Program, in large part due to policy's failure to impose regulation accommodates spatial variation in contribution to non-attainment.The D.C. Circuit Court determined that the proposed trading programs were unlawful because they did not connect upwind states' emission reductions to any measure of their contribution to nonattainment in other (downwind) states.

[3] AP2 is the stochastic version of the APEEP model (Muller, Mendelsohn, 2007;2009).

analysis of spatially differentiated NOx trading considers multi-state zonal approaches to policy differentiation. This blunt approach which would obscure a significant portion of the spatial variation in damages. (Krupnick et al., 2000; USEPA, 1998)[4].

Second, we use an econometrically estimated model to simulate facility-level compliance choices in the NOx Budget Program (Fowlie, 2010). This allows us to address an important limitation of earlier work that has investigated the benefits of policy differentiation. In the workhorse models that are typically used to to inform the design and implementation of market-based emissions regulation, it is standard to assume that firms adhere to efficient, cost-minimizing compliance strategies. Emissions market outcomes are often simulated using a deterministic, compliance cost-minimization algorithm. Policy analysts have duly noted that this approach may fail to capture salient features of the real world decision processes that drive emissions abatement decisions (Krupnick et al. 2000). We conduct two sets of policy simulations. One incorporates a fairly stylized cost-minimization-based model of firms' compliance choices. Parameters are calibrated to mimic the simulation models that are routinely used in policy analysis. The second incorporates a similarly structured model, the parameters of which are econometrically estimated using detailed data from the NOx Budget Program. Notably, the two models predict substantially different firm-level responses to damage-differentiated policy incentives. This has important implications for our estimated gains from policy differentiation.

Finally, in addition to estimating the expected gains from differentiation, we conduct a formal assessment of the parameter uncertainty inherent in marginal damage estimates. Detailed concerns have been raised with respect to the treatment of uncertainty in the analysis of the benefits from environmental regulation (NRC 2002; GAO 2006 ; Krupnick et al. 2006; OMB, 2003).[5] We explicitly track uncertainty in emissions, atmospheric transport and chemistry, population exposure, exposure-response, and valuation of health impacts by propagating statistical or parameter uncertainty through the integrated assessment model in order to produce estimates of the source-specific marginal damage distributions. This allows us to characterize uncertainty in both our marginal damage estimates and the estimated gains from policy differentiation.

When we use the calibrated cost-minimization model of firm-level compliance choices to simulate outcomes, our estimated gains from policy differentiation are larger than those reported in past studies. Intuitively, this is because we capture variation in emissions damages to a much greater extent. Annual damages from the permitted pollution are reduced by 30 percent relative to the undifferentiated policy. Net benefits of the trading program (i.e. avoided damages less abatement costs) increase by more than 30 percent.

When we replace the cost-minimization algorithm with the econometrically estimated model of firms' compliance decision, our estimated gains from policy differentiation are smaller. We find annual damages reductions of only 3-8 percent. Net benefits are also small, or even negative, depending on our assumptions

[4] For example, policy makers considered dividing the regulated region into two or three subregions in an effort to make a distinction among the States that may contribute the most to the ozone transport problem and those where the wind patterns may be less likely to affect air quality in the other states.

[5] The report specifically requested the development of probabilistic, multiple-source uncertainty models based not only on available data but also on expert judgment.

about how the differentiated policy is implemented. The striking differences between the two sets of simulations has everything to do with costs. All else equal, the steeper are firms' abatement cost curves, the more costly it is to reallocate permitted emissions across sources, the smaller the benefits from policy differentiation. These cost curves that rationalize observed compliance choices are much steeper than those implied by the standard cost minimization model. Taken together, our results suggest that the efficiency lost as a consequence of regulating a non-uniformly mixed pollutant with an undifferentiated policy were likely small.

2 Theoretical framework

Consider a group of N firms emitting a non-uniformly mixed pollutant. The extent of the damage caused by emissions of this pollutant depends not only on the level of emissions, but also how the emissions are distributed across sources.[6] We define abatement cost functions in terms of emissions: $C_i(e_i)$. We assume that $C_i'(e_i) \leq 0 \leq C"_i(e_i)$.

We define pollution damage functions in terms of emissions: $D_i(e_i)$. For each source, we define a marginal damage parameter $D_i'(e_i) \equiv \delta_i$. We capitalize on a series of empirical simulations, which are discussed in appendix A6, to impose structure on the damage function. Specifically, we assume that the damage function is linear and additively separable. This implies that the product of the marginal damage times emissions is equal to the total damages: $D(e) = \sum_i \delta_i e_i$. We begin by assuming that policy makers know these marginal damage parameters across all (N) soruces with certainty prior to implementing the emissions policy.

Suppose that the policy maker's objective is to minimize the total social cost (TSC) associated with emissions of this pollutant:

$$TSC = \sum_{i=1}^{N} \left(D_i(e_i) + C_i(e_i)\right) \tag{1}$$

The first component in Eq. (1) measures damages from pollution. The second term measures the costs of reducing emissions levels below unconstrained "business as usual" levels. To minimize the total social costs, one differentiates Eq. (1) with respect to source-level emissions. As is well-known, assuming an interior solution, first-order conditions for total cost-minimization imply:

$$-C_i'(e_i^*) = \delta_i \quad \forall\ i. \tag{2}$$

Intuitively, marginal costs are set to equal marginal damages across all sources. The * superscript denotes efficient emissions levels. The efficient level of aggregate emissions is thus $E^* = \sum_{i=1}^{N} e_i^*$.

.

[6] In this analysis, we will focus exclusively on the spatial heterogeneity in damages. See Joskow, Martin and Ellerman (CITE) for an analysis of the implications of temporal variation in damages.

2.1 Market-based regulation of non-uniformly mixed pollution

Having characterized the first-best emissions outcome, we next evaluate the performance of alternative market-based policy designs against this benchmark. We will consider emissions-based policy designs exclusively. Although we are ultimately concerned about limiting the *damages* associated with pollution exposure, we will limit our attention to policies that regulate *emissions*. Ambient permit systems could be used to limit damages in theory (see, for example, Montgomery 1972). But these policies are too complicated to implement in practice. Moreover, the Clean Air Act explicitly stipulates that emissions targets be used to achieve the National Ambient Air Quality Standards (NAAQS).

We are primarily interested in understanding how the market-based policies that are currently in place could be modified to better accommodate variation in damages across sources. Consequently, our analysis focuses almost exclusively on emissions trading programs (versus emissions taxes). Under the auspices of the Clean Air Act, emissions trading programs have been used to regulate non-uniformly mixed point source pollution at the federal and regional level. That said, it is instructive to briefly consider how a tax regime could be designed to achieve the first best outcome.

Consider a differentiated tax regime in which each source is required to pay δ_i per unit of emissions. In theory, cost minimizing firms would choose the level of emissions that minimizes:

$$\min_{e_i} \; C_i(e_i) + \delta_i e_i.$$

Intuitively, each source reduces emissions until the marginal abatement cost equals the imposed tax. This delivers the optimal outcome defined by (2). We will revisit this differentiated tax regime in the applied analysis. It will serve primarily as a basis for comparison.

In the emissions trading programs we consider, tradeable emissions permits corresponding to the total emissions cap are allocated to participating sources either by auction or a gratis using some allocation rule that does not depend on production decisions going forward. Any free allocation of permits to firm (i) is represented by the initial allocation A_i. As a point of departure, we assume that the emissions cap is set optimally at E^* and that markets are efficient and free of distortions. We subsequently refer to this as the "first-best" case. To keep the analytics simple and intuitive, we consider a case with only two price taking firms. It is straightforward to generalize the following analysis to $N > 2$, although this complicates notation. Producers are denoted h and l to indicate high and low damage areas, respectively, and in order to clean-up the exposition of our results, the marginal damage for firm (h), $D'_h(e_h)$, is denoted (δ_H), and the marginal damage for firm (l), $D'_l(e_l)$, is denoted (δ_L).

We first consider an undifferentiated market-based policy design that implicitly assumes that marginal damages are equal across all sources. We then contrast this with policy regimes in which the terms of compliance are designed to reflect heterogeneity in damages.

2.2 Undifferentiated Permit Market

Most existing and planned emissions trading programs feature undifferentiated permits: meaning that firms are required to hold a permit to offset each unit of emissions, regardless of where the emissions occur. Trading occurs on a ton-for-ton basis. We assume that each firm chooses emissions (e_i), emissions permit purchases (e_{bi}), and permit sales (e_{si}), both valued at the market-determined price (τ), to minimize the costs of complying with this emissions-based trading program.

$$\begin{aligned} \min_{e_i, e_{si}, e_{bi}} \quad & C_i(e_i) + \tau\left(e_{bi} - e_{si} - A_i\right) \\ s.t. \quad & e_i \leq A_i - e_{si} + e_{bi} \\ & e_i, e_{si}, e_{bi} \geq 0, \end{aligned} \tag{3}$$

If we assume an interior solution, cost-minimization implies that marginal abatement costs are set equal across all sources:

$$C_h'(e_h^u) = C_l'(e_l^u) = \tau, \tag{4}$$

where the u superscript denotes the undifferentiated trading equilibrium.

Figure 1 illustrates these first order conditions in the simple two firm case. The width of this figure, measured in units of emissions, is equal to the total quantity of permitted emissions $\overline{E}$. We first consider a case in which the cap has been optimally set to E^*. At the left origin, all emissions occur at the low damage firm (i.e. $e_l = \overline{E}$) and emissions at the high damage firm are driven to zero ($e_h = 0$). The upward sloping solid line, moving from left to right, represents the marginal abatement costs at the low damage firm: $C_l'(e_l)$. At the right origin, the high damage firm emits E^* (i.e. $e_h = \overline{E}$) and the low damage firm emits nothing ($e_l = 0$).The solid line increasing from right to left measures marginal abatement costs at the high damage firm $C_h'(e_h)$.

Equilibrium emissions under the undifferentiated trading regime are given by $\{e_l^u, e_h^u\}$. This equilibrium occurs at the intersection of $C_h'(e_h)$ and $C_l'(e_l)$ which is congruent with the first-order condition for cost-minimization depicted above. This allocation of permitted emissions minimizes the total abatement costs required to meet the emissions cap $\overline{E}^*$. However, this is not the optimal outcome since, by construction, damages do not enter into the constrained optimization problem in (3). Total social welfare could be improved by shifting some of the permitted emissions away from the high damage source to the low damage source (Muller and Mendelsohn, 2009). As long as the policy facilitates trading and compliance based on one permit price (τ), and $\delta_h \neq \delta_l$, an undifferentiated market cannot achieve allocative efficiency. This motivates the consideration of differentiated designs.

In Figure 1, we assume that the relatively high damage firm also faces relatively higher costs of abatement. If the reverse were true, the optimal cap E^* would change, but it would still be the case that too much of the permitted emissions would be allocated to the high damage firm under the undifferentiated policy.

2.3 Differentiated Permit Market

We now consider how this market-based policy design can be modified so as to achieve the socially optimal allocation of permitted emissions. There is a growing literature that examines "differentiated" policies that are designed to reflect variation in pollution damages (Teitenberg, 1995; Farrow et al., 2004; Horan and Shortle, 2005; Muller and Mendelsohn, 2009). To date, work in this area has focused on the construction of trading ratios based on the ratio of marginal damages between each pair of regulated sources.[7] It is straightforward to operationalize these ratios within our simple analytical framework.

Let $\overline{\delta}$ represent the average of the marginal damage across all sources in a trading program. In this simple two firm case, $\overline{\delta} = \frac{\delta_l + \delta_h}{2}$. We construct firm-specific damage ratios r_i, normalizing each firm's marginal damage by the mean damage parameter : $r_i = \frac{\delta_i}{\overline{\delta}}$, and therefore, $\frac{r_i}{r_j} = \frac{\delta_i}{\delta_j}$. Because this policy recognizes that emissions produced by different firms produce different degrees of harm, to remain in compliance, each firm must hold r_i permits to offset each unit of uncontrolled emissions. For the case of an undifferentiated policy, the constraint that defines the terms of compliance (at the facility-level) was given by: $e_i \leq A_i - e_{si} + e_{bi}$. In the case of a differentiated policy, this compliance constraint is modified in the following manner:

$$r_i e_i \leq A_i - e_{si} + e_{bi}.$$

All else equal, the more damage caused by emissions at a given source, the more permits that source needs to hold to offset its emissions.

The first order conditions with respect to (e_i) for cost minimization in this differentiated regime imply that the ratio of marginal damages will be set equal with the ratios of marginal costs across the two firms:

$$\frac{C_j'(e_j^r)}{C_i'(e_i^r)} = \frac{\delta_j}{\delta_i}, \; i \neq j. \tag{5}$$

where the r superscript denotes the equilibrium outcome under a regime that incorporates these trading ratios.

In this first-best setting, (5) delivers the socially optimal allocation of permitted emissions across sources (see Appendix 1). Figure 1 illustrates this result graphically. The broken lines represent the marginal abatement cost schedules scaled by the inverse of the corresponding marginal damage: $C_i'(e_i)\frac{1}{\delta_i}$, $i = l, h$. By (5), the allocation of emissions across these two sources occurs where these broken lines intersect. This allocation of the permitted emissions achieves the optimal trade off between abatement costs and benefits from reduced damages.

This differentiated policy design will welfare-dominate the undifferentiated system if benefits (in the form of avoided damages) exceed the increase in abatement costs. In Figure 1, gross benefits from differentiation

[7] It is worth pointing out that the phrase "trading ratio" is somewhat misleading insofar as these ratios affect emissions trading activities only indirectly via the effect on compliance requirements.

are represented by area ABCE The increase in abatement costs is equal to area ACD. The net benefits from differentiation, defined by the shaded areas ABD + CDE, are positive.

To fix ideas, we derive a more general expression for these net benefits. We maintain our assumption that the damage function is linear and additively separable in source-specific damages. We now assume the following functional form for the source-specific abatement cost functions: $C_i(e_i) = \alpha_{0i} - \alpha_{1i}e_i + \beta_i e_i^2$. We accommodate heterogeneity in abatement costs by allowing the parameters of this function to vary across sources. The first derivative of $(C_i(e_i))$ serves as a linear approximation to the marginal abatement cost at each source. In practice, these cost curves may be discontinuous as pollution abatement often involves lumpy investments in emissions reducing capital equipment.

Solving for the optimal cap yields the following:

$$\overline{E}^* = \frac{(\alpha_{1h}\beta_l + \alpha_{1l}\beta_h - \delta_h\beta_h - \delta_l\beta_h)}{2\beta_l\beta_h}. \tag{6}$$

Note that this optimal limit depends not only on the damage parameters $\{\delta\}$ and the cost parameters$\{\alpha_1, \beta\}$, but also on the correspondance between these two sets of parameters. More precisely, rich information regarding source-specific damage and cost parameters is required to set the cap optimally.

Solving for equilibrium emissions under the emissions-based and differentiated policy designs, respectively, we obtain an expression for the change in source-specific emissions induced by this differentiation:

$$e_h^r - e_h^u = \frac{\delta_l - \delta_h}{2\left(\beta_h + \beta_l\right)} \tag{7}$$

$$e_l^r - e_l^u = \frac{\delta_h - \delta_l}{2\left(\beta_h + \beta_l\right)} \tag{8}$$

Intuitively, differentiation shifts some share of permitted emissions from the high damage source to the low damage source. The extent of this reallocation depends on the difference in damage parameters and the steepness of the marginal abatement cost curves.

It is then straightforward to derive an expression for the net benefits from differentiation in terms of the model parameters (see Appendix 2 for complete derivation):

$$TSC^u(\delta, \beta) - TSC^r(\delta, \beta) = \frac{(\delta_l - \delta_h)^2}{4(\beta_l + \beta_h)} \geq 0, \tag{9}$$

The u superscript denotes the undifferentiated design; r denotes the differentiated design that incorporates ratios. The vectors of damage parameters and abatement cost coefficients are denoted δ and β, respectively.

We can now make two observations based on equation (9):

(1) The extent to which differentiation reduces pollution damages (via a reallocation of permitted emissions) is increasing with the variation in damages across sources and decreasing with the slope of the marginal abatement cost functions.

(2) The net benefits conferred by differentiation in this first best setting are increasing with the variation in damages across sources and decreasing with the slope of the marginal abatement cost functions.

These findings should be intuitive. If damages do not vary, there is no advantage to differentiated policy. Accordingly, the more heterogeneous the damages, the greater the benefits from differentiation, all else equal. Finally, if marginal abatement costs are steeply increasing in abatement, it will be relatively more costly to shift emissions from the high damage to the low damage source.

Note that the benefits from differentiation do not depend on the correlation between source-specific abatement costs and source-specific damage parameters. This contrasts with the findings of Mendelsohn (1986) who finds that positive covariance between abatement cost parameters and emissions damages increase the relative effectiveness of differentiated policy designs. Given our maintained assumptions regarding the linearity of the damage function, this relationship disappears.

The above analysis is predicated on some strong simplifying assumptions that are likely to be violated in practice. In what follows, we investigate optimal policy of jurisdictional constraints and uncertainty.

2.3.1 Exogenously determined emissions constraint

In the theoretical literature that considers the design and implementation of spatially differentiated emissions policies, it is standard to assume that the cap can be optimally set (Muller and Mendelsohn, 2009). In fact, this is unlikely to be a safe assumption. It is often the case that the emissions constraint is (explicitly or implicitly) determined by a superseding authority. The implementing agency must therefore determine the optimal policy design conditional on this cap. Additionally, even if the regulator does set the emissions cap, it is not clear that they could do so without firm-specific estimates of marginal social cost and marginal abatement costs.

To examine this more common situation, we now assume that the policy maker seeks to minimize total social costs associated with a given emissions constraint $\overline{E}$:

$$\begin{aligned} \min_{e_h, e_l} TSC &= D_h(e_h) + D_l(e_l) + C_h(e_h) + C_l(e_l) \\ s.t. e_h + e_l &\leq \overline{E} \end{aligned}$$

First order conditions with respect to (e_l) and (e_h) for constrained total cost minimization imply[8]:

[8] We derive (10) using the Lagrangian method. The first-order condition with respect to (e_l) is: $\delta_l + C_l'(e_l) = \lambda$. Since (λ), the shadow value of relaxing the emission constraint, also appears in the same way in the first-order conditions with respect to (e_h), (λ) drops out of (10).

$$C_h'(e_h) - C_l'(e_l) = \delta_h - \delta_l \tag{10}$$

Intuitively, Eq. (10) implies that the benefits associated with moving a unit of pollution from the relatively high damage source to the relatively low damage source ($\delta_h - \delta_l$) are set equal to the costs associated with this incremental reallocation. Note that equations (2) and Eq. (10) are simultaneously satisfied when $C_i'(e_i) = \delta_i$ across all sources. However, if the cap has not been set optimally, the market fails to clear when marginal abatement costs are set equal to marginal damages. In a regime that incorporates first best trading ratios, marginal abatement costs will exceed (fall below) marginal damages if the cap has been set too stringently (loosely). In either case, Eq. (10) is not met if the first-best ratios are used to design policy.

Figure 2 illustrates the case in which the emissions constraint is too stringent. As in Figure 1, the intersection of the broken lines defines the equilibrium allocation of permitted emissions under a policy regime that employs first-best trading ratios $\{e_h^r, e_l^r\}$.At this allocation, the ratio of marginal abatement costs equals the ratio of marginal damages. Because the cap has been set too stringently, the difference in marginal abatement costs exceeds the difference in marginal damages. Contrast this with the allocation of emissions denoted by the superscript ^{w} which satisfies Eq. (10). This allocation minimizes the total social cost associated with the emissions constraint $\overline{E}$. Net benefits vis a vis the emissions-based equilibrium are given by the value of avoided damages (area DEFG) less increased costs (area ABC).

Can the constrained optimal outcome be achieved using a decentralized, market-based policy intervention? The theory literature has investigated the use of "second best" ratios which are intended to minimize the total social costs associated with a given emissions constraint (see, for example, Horan and Shortle, 2005; Muller, 2011). As compared to the first-best ratios introduced above, these second-best ratios are more difficult to construct. For example, to implement the second-best ratios developed by Horan and Shortle (2005) the policy maker must correctly anticipate the shadow value of the imposed emissions constraint.

We introduce an alternative approach to achieving the constrained optimum. Let w_i represent a source-specific difference (or "wedge") between the firm-specific marginal damage and the expected marginal damage across sources: $w_i \equiv \delta_i - \overline{\delta}$. As in the emissions-based regime, firms must hold permits to offset uncontrolled emissions in order to remain in compliance. But we now introduce an additional compliance requirement. In addition to holding permits, each firm must pay a compliance payment equal to: $e_i($ $\delta_i - \overline{\delta}_i)$. This is simply their compliance wedge times their emission level. It reflects either the additional (or subtracted) damage relative to an average firm.

The firm's compliance cost minimization problem can now be written:

$$\min_{e_i, e_{si}, e_{bi}} \quad C_i(e_i) + (\tau + w_i)\,(e_{bi} - e_{si} - A_i) \tag{11}$$

$$s.t. \quad e_i \leq A_i - e_{si} + e_{bi}$$

$$e_i, e_{si}, e_{bij} \geq 0,$$

This policy design modification drives a wedge between the market clearing permit price and the price paid by firms whose marginal damage exceeds, or falls below, the average damage value. As in the first-best trading ratio regime, relatively high (low) damage firms face a higher (lower) cost of offsetting emissions using permits. This tends to reallocate emissions from sources that cause high marginal damage to sources whose discharges cause lower damages.

The first order conditions for emissions (e_i), purchases (e_{bij}), and sales of allowances (e_{si}) under this spatially differentiated regime indicate that the firm's cost minimizing behavior is aligned with the socially optimal outcome in the constrained setting:

$$C_h'(e_h) - C_l'(e_l) = (\delta_h - \delta_l). \tag{12}$$

A policy regime that incorporates these wedges will achieve the constrained optimum because the cap is not optimally set. The net welfare gains (relative to the undifferentiated design subject to the same emissions cap) do not depend on the exogenously set emissions cap. Simply stated:

(3) When the emissions cap is not set optimally (as will often be the case), differentiation based on ratios of marginal damages will not yield the optimal allocation of the permitted emissions. Differentiation based on differences in marginal damages is preferable.

The advantage of using wedges (versus second-best trading ratios) to define compliance requirements is that the information requirements are not as onerous. Only source specific measures of marginal damages are required. A potential disadvantage is that the wedge-based policy is not public revenue neutral. Under the benchmark emissions-based design and policies that incorporate trading ratios, the implementing agency does not pay out or take in funds while administering program compliance. In contrast, under the wedge-based regime, program compliance requires high damage firms to pay ($w_h e_h$) to the implementing agency, while the agency is required to pay ($w_l e_l$) to low damage firms. Depending on how the permitted allocations are allocated across high and low damage firms in equilibrium, the implementing agency may be net long or short after compliance requirements have been satisfied.

2.3.2 Uncertain damages

Another important practical consideration is the uncertainty that pervades the policy design process. Policy makers are confronted with significant uncertainty regarding both costs and damages from pollution. We will focus exclusively on uncertainty surrounding the source-specific damage parameter estimates. More often than not, uncertainty about the benefits of pollution abatement is of substantially greater magnitude than abatement costs (Stavins, 1996). This is certainly true of the policy application we analyze below.

Estimates of source-specific marginal damages, which form a key aspect of the design of efficient policy, are highly uncertain. This uncertainty arises from random variation in data (variability), lack of knowledge

about an empirical quantity (parameter uncertainty), incorrect model specification (model uncertainty), and modeling choices that reflect implicit decisionmaker judgment (decision uncertainty), (Burtraw et al., 2006). We will model the uncertainty that derives from stochastic data inputs, parameter uncertainty, and decision uncertainty using the joint probability distribution $f(\delta)$. Our analysis will not account for either model uncertainty.

We consider a very basic decision theoretic framework to study the policy design implications of parameter uncertainty. We assume that the joint distribution of the marginal damage parameters $f(\delta)$ is known ex ante. The risk neutral policy maker seeks to minimize the expected social costs of pollution. The timing of the design decision is quite simple. The policy maker must commit to a policy design at the outset of the program. Compliance parameters are fixed over the duration of the program.

Maintaining the simple two-firm set-up as above, the policy maker seeks to minimize expected total social costs from emissions subject to the exogenously determined emissions constraint:

$$\begin{aligned} \min_{e_h,e_l} TSC &= C_h(e_h) + C_l(e_l) + \int\int(\delta_h e_h + \delta_l e_l) f(\delta_l, \delta_h) \\ s.t. \quad e_h + e_l &= \overline{E} \end{aligned} \tag{13}$$

Substituting in the constraint, the first order condition for cost minimization yields:

$$\begin{aligned} -(C_h'(e_h) - C_l'(e_l)) &= \int\int(\delta_h - \delta_l) f(\delta_l, \delta_h) \\ -(C_h'(e_h) - C_l'(e_l)) &= E[\delta_h] - E[\delta_l] \end{aligned} \tag{14}$$

The constrained optimum in the presence of uncertainty about the marginal damage parameters equates differences in marginal abatement costs with differences in expected marginal damages. If the w_i parameters are constructed using expected source-specific marginal damages $E[\delta_h]$, $E[\delta_l]$, Eq. (14) will be satisfied in expectation.

We note some important qualifications. First, our assumption regarding the linear form of the damage function is important here. In the literature that examines the implications of uncertain damages on optimal policy, researchers have argued that the optimal trading ratio between two sources with equal expected damages but varying degrees of uncertainty should penalize the more uncertain damages (e.g. Horan, 2001; Horan and Shortle, 2005;). In our case, linearityin damages eliminates the covariance term that gives rise to this penalty.

Second, our assumed policy objective function is also important. We assume that the regulator seeks to minimize the expected social costs of pollution. If instead the regulator wants to meet an ambient target probabilistically, varying degrees of uncertainty will matter because otherwise identical firms will have differential marginal effects on the probability the target is violated. Alternatively, the regulator could

place a priority on minimizing the deviation from the status quo, undifferentiated design. The regulator might seek to aggregate sources with similar damage estimates so as to implement a zonal policy design.

Before looking into how uncertain damages affect the expected benefits from differentiation, it is instructive to consider how net benefits vary across draws from the distribution $f(\delta)$. Consider a policy regime that incorporates wedges defined using $E[\delta_l]$ and $E[\delta_h]$. Let $\delta' = \{\delta'_h, \delta'_l\}$ denote a particular draw from the joint distribution of marginal damages. We can think of this δ' as the marginal damages that were actually realized. The difference in total social costs under the compliance wedge design (TSC^w), relative to an undifferentiated policy (TSC^u) is derived in Appendix 3. This expression reduces to:

$$TSC^w - TSC^u = \frac{1}{2}\frac{\left(\delta'_l - \delta'_h\right)\left(E[\delta_l] - E[\delta_h]\right)}{(\beta_l + \beta_h)} - \frac{1}{4}\frac{\left(E[\delta_l] - E[\delta_h]\right)^2}{(\beta_l + \beta_h)}. \tag{15}$$

The first argument in Eq. (15) represents the difference in damages across emissions-based and differentiated policy regimes. The second argument captures the increase in abatement costs associated with a move to the differentiated policy design (see Appendix 3).

When damages are uncertain, note that it is possible for the realized net benefits of differentiation to be negative. Consider an extreme case where the realized damage values $\{\delta'_h, \delta'_l\}$ are negatively correlated with the expected damage values, $E[\delta_l]$ and $E[\delta_h]$. In this case, the source that was expected to be associated with relatively low damages is actually the relatively high damage source. The differentiated policy will incorrectly penalize the low damage source vis a vis the high damage source. Given this realization of damages, the uniform, emissions-based regime welfare dominates the (misguided) differentiated regime.

The expected net welfare gain associated with the spatially differentiated policy is obtained by integrating over the entire distribution of damages (see Appendix 4):

$$E[TSC^w(f(\delta), \beta) - TSC^u(f(\delta), \beta)] = \frac{(\delta'_h - \delta'_l)^2}{4\left(\beta_h + \beta_l\right)}\left(2cov(\delta_l, \delta_h) - var(\delta_l) - var(\delta_h)\right) \tag{16}$$

Uncertain damage parameter estimates introduces some additional determinants of the benefits from policy differentiation:

(4) Benefits from differentiation are increasing in the variance of damages across sources, decreasing with the variance of the source-specific damage distributions, and increasing with the covariance in damages across sources.

Intuitively, if the damage parameters are very precisely estimated, the policy maker can design the policy to very accurately reflect the ex post realized damages. In contrast, if there is a lot of within source variance in marginal damage estimates, then it is more likelty that the parameters used to define the terms of compliance ($E[\delta_l]$ and $E[\delta_h]$) will innacurately reflect the damages that are actually realized. If there is strong positive correlation in damage realizations across sources, the average damage parameters which are used to define the terms of compliance will be more positively correlated (in expectation) with the ex post realized damages.

3 The NOx Budget Program

The NOx Budget Program (NBP) is a market-based emissions trading program created to reduce the regional transport of NOx emissions in the eastern United States. The program establishes a region-wide cap on emissions of NOx from large stationary sources in twenty eastern states during ozone season (May-September). The NBP was primarily designed to help Northeastern and Mid-Atlantic states attain Federal ozone standards. When the NBP was promulgated, significant portions of the Northeast, Mid-Atlantic, and parts of the Midwest were failing to meet Federal standards (Ozone Transport Assessment Group (OTAG), 1997).

Although the precise contribution of individual sources to the non-attainment problems in this region was difficult to estimate at the time of the rulemaking, there was plenty of evidence to suggest that marginal damages varied significantly across sources. The EPA received over 50 responses when, during the planning stages of the NOx SIP Call, it solicited comments on whether the program should incorporate trading ratios or other restrictions on interregional trading in order to reflect the significant differential effects of NOx emissions across states(FR 63(90): 25902). Most commentors supported unrestricted trading and expressed concerns that "discounts or other adjustments or restrictions would unnecessarily complicate the trading program, and therefore reduce its effectiveness" (FR 63(207): 57460). These comments and some accompanying analysis (US EPA, 1998a) led regulators to design a single jurisdiction, undifferentiated trading program. There are no spatial restrictions on trading within the program. All emissions are treated symmetrically for compliance purposes.[9]

In 2008, a federal district court vacated the rule that was to succeed the NOx Budget Program due to policy's failure to adequately accommodate regional transport of pollution and associated spatial variation in damages.[10] Since that time, debates surrounding this program, and how it should be redesigned have become increasingly contentious. In what follows, we revisit the decision to forego a differentiated policy design in favor of the simpler, undifferentiated alternative.

Our analysis will focus exclusively on the coal-fired generating units in the program. Although gas- and oil-fired generators and other industrial point sources are also included in the NBP, coal-fired units represent approximately 94 percent of the NOx emissions regulated under the program and at least 94 percent of the NOx emissions reductions over the first five years (U.S. EPA, 2005; US E.P.A. 2008). Natural gas and oil-fueled plants tend to have much lower uncontrolled NOx emissions rates.[11] By exempting these units from our analysis, we are implicitly assuming that operating decisions at these units would not be differentially affected under an emissions-based design and the counterfactual policy designs we consider. Future versions of the paper will test this assumption explicitly.

[9] The US EPA has also investigated the potential to use weather and atmospheric chemistry forecasts to vary the NOx permit price over time (US EPA, 2007).

[10] The court found that the CAIR regulation "does not prohibit polluting sources within an upwind state from preventing attainment of National ambient air quality standards in downwind states."*State of North Carolina v. Environmental Protection Agency, No. 05-1244,* slip op. (2008), District of Columbia Court of Appeals.

[11] Whereas the average pre-retrofit NOx emissions rate among coal plants exceeded 5.5 lbs/MWh, average NOx emissions rates among marginal electricity producers are estimated to range between 0.3 to 2.2 lbs NOx/MWh (NEISO, 2006; Keith et al., 2003).

4 Estimating the benefits from differentiation

We use the theory model presented in section 2 as a conceptual framework for a detailed analysis of the benefits from differentiation in the context of the NOx Budget Program. This applied analysis proceeds in several steps, each of which are described in detail in the following five subsections.

4.1 Estimating damages from pollution

NOx emissions affect health and environmental outcomes through two main pathways: ozone formation and particulate matter formation.[12] Specifically, emitted NO_x interacts with ambient ammonia to form ammonium nitrate, a constituent of ambient $PM_{2.5}$. And NO_x also forms tropospheric O_3 through a series of chemical reactions (Seinfeld, Pandis, 1998). Both $PM_{2.5}$ and O_3 are criteria air pollutants regulated under Title I of the Clean Air Act. As such, exposures to these two pollutants have been shown to have a number of adverse effects on human health and welfare. Prior research has shown that the majority of damages due to exposures to both $PM_{2.5}$ and O_3 are premature mortalities and increased rates of illness (USEPA, 1999; Brunekreef and Holgate, 2002; WHO, 2003 Muller and Mendelsohn, 2007;2009).

The extent to which NO_x emissions react with precursors to form ozone or particulate matter depends upon prevailing meteorological conditions, pre-existing precursor emissions and concentrations, and other factors that vary across time and space. Furthermore, the health impacts associated with a change in ozone and/or particulate matter at a particular location will depend on the human populations at that location. For these reasons, the damage caused by a given quantity of NOx emissions will depend significantly on the spatial distribution of the emissions.

The integrated assessment models that are used to estimate marginal damages from air pollution incorporate many imprecisely estimated parameters (such as dose-response parameters and population exposure estimates) and stochastic inputs (such as wind direction and humidity). In what follows, we characterize both variability and uncertainty in NOx emissions damages in unprecedented detail. With respect to the former, we estimate the extent of the variation in marginal damage estimates across sources in the NOx Budget Program. With respect to the latter, it is important to emphasize that our uncertainty analysis is not comprehensive. We formally quantify the parameter uncertainty inherent in source-specific damage estimates. But we make no attempt to capture modeling uncertainty.

4.1.1 Source-specific damage parameters

The source-specific (δ_i) parameters capture the estimated effect of an incremental change in NOx emissions at source i on health and environmental impacts across the airshed. We use a stochastic integrated assessment model, AP2, to estimate these source-specific damage parameters (Muller, 2011). The AP2 model is comprised of six modules; emissions, air quality modeling, concentrations, exposures, physical effects, and monetary damages. The emissions data used in AP2 is provided by the US EPA's National Emission Inventory for 2005 (US EPA, 2009). These data encompass emissions of NO_x, $PM_{2.5}$, sulfur

[12] NOx emissions also contribute to acid rain in some mountain regions, and exacerbate eutrophication problems.

dioxide (SO_2), volatile organic compounds (VOCs), and ammonia (NH_3). AP2 attributes these data to both the appropriate source location and source type. Specifically, AP2 models emissions from 656 individual point sources (mostly large EGUs). Emissions from the remaining point sources are decomposed according to height of emissions and the county in which the source is located. For ground-level emissions (these are produced by cars, residences, and small commercial facilities) AP2 attributes these discharges to the county in which they are reported (by US EPA) to occur.

The approach to air quality modeling used in AP2 relies on the Gaussian Plume model (Turner, 1994). This approach uses a reduced form statistical model to capture the processes that connect emissions (e) to concentrations (C). The relationship between emissions of nitrogen oxides released at source i and the concentration of pollutant s (ozone or particulate matter) at receptor point r is captured by $C_{sri}(e_i)$. The predicted pollutant concentrations generated using the AP2 model have been tested against the predictions made by a more advanced air quality model (see the appendix in Muller, 2011). The agreement between the county-level surfaces produced by the two models is quite strong.

AP2 then connects ambient concentrations to physical impacts using peer-reviewed dose-response functions. Let β^s_{kp} represent the dose response coefficient which captures the effect of an incremental change in the concentrations of pollutant s on health outcome k in population cohort p. In order to model impacts of exposure to $PM_{2.5}$ on adult mortality rates, this analysis uses the findings reported in Pope et al., (2002). The impact of $PM_{2.5}$ exposure on infant mortality rates is modeled using the results from Woodruff et al., (2006). For O_3, we use the findings from Bell et al., (2004). In addition, this analysis includes the impact of exposure to $PM_{2.5}$ on incidence rates of chronic bronchitis (Abbey et al., 1995).

The final modeling step in connecting emissions to damages translates the physical effects predicted by the dose-response functions into monetary terms. Let α_k represent the valuation coefficient that is used to translate the health outcome k into dollar terms. We rely on valuation methodologies used in the prior literature. In order to value the risk of premature mortalities due to pollution exposure, we employ the Value of a Statistical Life (VSL) method. (See Viscusi and Aldy, 2004 for a summary of this literature.) In particular, we employ a VSL of approximately $6 million; this value, which is used by US EPA, results from a meta-analysis of nearly 30 studies that compute VSLs using both stated and revealed preference methods. Further, each case of chronic bronchitis is valued at approximately $300 thousand which is also the value used by US EPA.

The marginal ($/ton) damage for NO_x for the 632 coal-fired EGUs regulated by the NBP are estimated using the marginal damage algorithm used in Muller (2011) which is based on the routine developed in Muller and Mendelsohn (2007; 2009) . This algorithm includes the following steps. First, baseline emissions are constructed from detailed emissions data collected by the US EPA in the years immediately preceding the introduction of the NOx Budget Program. These emissions reflect the NO_x controls required for all sources in non-attainment areas. AP2 computes total national damages associated with these baseline levels of NOx emissions. Next, one ton of NO_x is added to baseline emissions at a particular EGU. AP2 is then re-run. Concentrations, exposures, physical effects, and damages are recomputed. Since the only difference between the baseline run and the "add-one-ton" run is the additional ton of

NO_x, the change in damages is strictly attributable to the added ton. This design is then repeated over all of the EGUs encompassed by the NBP.

The marginal damage calculation in the context of statistical uncertainty involves the following steps. First AP2 makes a random draw (denoted the mth draw) from the input distributions. Next, AP2 computes m^{th} realization for emissions, concentrations, exposures, physical effects, and damages based on the realized draw from each input distribution. AP2 then adds one ton of NO_x to source i. Again, AP2 tabulates concentrations, exposures, physical effects, and damages conditional on the added ton of NO_x at source (i), (using the same m^{th} realization from the input distributions) . AP2 computes the difference between damages with baseline emissions and after adding the ton of NO_x to (i). This is repeated 4,999 times to estimate the empirical distribution of marginal damages for NO_x emitted from facility i. This process is then repeated for each EGU in the analysis.

Equation [17] provides a very parsimonious description of the marginal damage estimates used in our analysis:

$$\delta_i = \sum_{ri} \sum_{k} \sum_{s} \alpha_k \beta^s_{kp} P_{ri} \frac{dC_{si}(e_i)}{de_i}. \tag{17}$$

Given the stochastic nature of AP2, the parameters of the atmospheric model, the population estimates P_{ri} , the dose response parameters β^s_{kp} and valuation parameters α are treated as being uncertain. Even the emissions levels at individual sources cannot be predicted with certainty. These multiple sources of uncertainty beget significant uncertainty in the marginal damage estimates. As described above, the AP2 model is used to compute δ_{im}, where m indexes draws. This exercise yields an empirical distribution for each δ_i parameter.

The extent to which marginal damage estimates vary across draws is striking. Figure 3 summarizes the distribution of a single marginal damage parameter. This source, a single coal-fired electricity generating unit in Ohio, was chosen because the variance and skewness of the corresponding empirical distribution are very close to the median values across all units. The point estimate, or expected value, of the damage caused by an incremental change in emissions at this source is \$1496/ton NOx. The standard deviation is \$1796/ton. Muller (2011) finds that most of this within source variation stems from uncertainty in the air quality modeling component, adult mortality dose-response parameter estimates, and mortality valuation parameters. The skewness of the distribution stems from the multiplicative nature of the process that links emissions to damages.

Figure 4 illustrates the extent to which the expected values of source specific damage parameters $E[\delta_i]$ vary across sources. The average parameter value (averaged across all sources) is \$1711/ton of NOx. In the subsequent discussion, we classify any source with estimated damages exceeding (falling below) \$1711/ton NOx as "high" ("low") damage. Notably, a significant amount of the inter-source variation (approximately 45 percent) occurs within (versus between) states. This suggests that a zonal trading regime that employs state-level trading ratios (and permits one-for-one trading within states) is a fairly blunt policy tool to capture heterogeneity in emissions damages.

For five of the 632 units in our data, we find that the expected value of the marginal damage parameter δ is negative. This suggests that a decrease in NOx emissions at these sources leads to increased overall damages. This result is driven by the complex, non-linear photochemical reactions that transform NO_x and VOCs into ozone. Daily ozone concentrations are non-linear and monotonic functions of NOx and the ratio of volatile organic compounds (VOCs) and NOx. At sufficiently low ratios, the conversion of NOx to ozone is limited by the availability of VOCs. In these VOC limited conditions, reductions of NOx can increase peak ozone levels until the system transitions out of a VOC-limited state (Seinfeld and Pandis, 1998).

4.1.2 Parameterizing a damage-differentiated policy

The unit-specific damage parameter estimates summarized by Figure 4 can be used to define the compliance requirements imposed in a differentiated emissions trading program. Eq. (14) implies that the constrained optimum is obtained by equating differences in marginal abatement costs with differences in expected marginal damages. To construct these source-specific "wedges", we subtract the average damage parameter ($1711) from the source-specific expected damage measures $E[\delta_i]$. Relatively "high damage" units are required to pay an amount that exceeds the market clearing permit price for each unit of emissions, whereas relatively "low damage" units pay less than the permit price per ton. We assume that incentivizing pollution at facilities with negative damage parameter estimates would be politically unpopular. Instead, we exempt any units with negative expected damage parameters.

We will also simulate outcomes under a differentiated tax and an emissions trading program that uses the first-best trading ratios defined in section 2.3 to define the terms of compliance. Under the differentiated tax, each source is required to pay $E[\delta_i]$ per unit of uncontrolled emissions. To construct the ratios, the expected value of the source-specific damage measure δ_i is divided by the average expected value (averaged across all sources in the program). Relatively "high damage" units are required to hold $r_i > 1$ permit per ton of emissions under the spatially-differentiated trading counterfactual, whereas relatively "low damage" units are required to hold $r_i < 1$ permit per ton.

4.2 Estimating source-specific NOx abatement costs

The NBP mandated a dramatic reduction in average NOx emissions rates.[13] In the period between when the rule was upheld by the US Court of Appeals (March 2000) and the deadline for full compliance (May 2004), firms had to make costly decisions about how to comply with this new regulation. We will assume perfect compliance on behalf of all units. In fact, compliance has been close to 100 percent for the duration of the program (US EPA, 2008).

To comply, firms can do one or more of the following: purchase permits to offset emissions exceeding their allocation, install NOx control equipment, or reduce production at dirtier plants during ozone season. For the coal-fired units in our analysis, we rule out reduction in ozone season output as a compliance

[13] Pre-retrofit emissions rates at affected coal plants were, on average, three and a half times higher than the emissions rate on which the aggregate cap was based (0.15 lbs NOx/mmbtu).

strategy and assume that firm-level production and aggregate output are exogenously determined and independent of the environmental compliance choice. Coal-fired units are typically inframarginal due to their relatively low fuel operating costs. Consistent with this observation, Fowlie (2010) finds that the introduction of the NBP reduced profit margins at operating units, but not production levels.

We take as given the population of sources in the NOx Budget Program. That is, we rule out the possibility that one or more of the differentiated policy designs we consider would cause an electricity generating to exit the market prematurely. The coal plants in our analysis are long-lived. The average retirement age of a coal plant in the United States is 49 years. We do observe a small number of coal-fired boilers retiring during the study period. These are units with decades of service stretching as far back as the end of World War II. We assume that these retirement decisions are unaffected by the policy design; we exempt these units from the analysis.

The specific NOx control options available to a given unit vary across units of different vintages and boiler types. In general, the more capital intensive the compliance option, the greater the emissions reductions. Compliance options that incorporate Selective Catalytic Reduction (SCR) technology, a very capital intensive post-combustion control, can reduce emissions by up to ninety percent. NOx emissions rates can be reduced by thirty-five percent through the adoption of Selective Non-Catalytic Reduction Technology (SNCR). Pre-combustion control technologies such as low NOx burners (LNB) or combustion modifications (CM) require much smaller upfront investments and can reduce emissions by fifteen to fifty percent, depending on a boiler's technical specifications and operating characteristics. Some of these technology options are physically additive. For example, SCR technologies can be combined with pre-combustion control technologies to deliver even greater emissions reductions.

Three factors that are likely to significantly influence a manager's choice of environmental compliance strategy are the up-front capital costs K, the anticipated variable operating costs V, and the expected emissions rate m. The capital costs, variable operating costs, and emissions reduction efficiencies associated with different compliance alternatives vary significantly, both across NOx control technologies and across generating units with different technical characteristics. We do not directly observe the variable compliance costs and fixed capital costs or the post-retrofit emissions rates that plant managers anticipated when making their decisions. We can, however, generate detailed, unit-specific engineering estimates of these variables.

In the late 1990s, to help generators prepare to comply with market-based NOx regulations, the Electric Power Research Institute[14] developed software to identify all major NOx control options (including combinations of control technologies) available to coal-fired boilers, conditional on unit and plant level characteristics. The software has been used not only by plant managers, but also by regulators to evaluate proposed compliance costs for the utilities they regulate (Himes, 2004; Musatti, 2004; Srivastava, 2004). This software was used to generate the unit-specific cost estimates used in this analysis (EPRI, 1999b). This cost estimation exercise is described in detail in Fowlie (2010).

[14] The Electric Power Research Institute (EPRI) is an organization that was created and is funded by public and private electric utilities to conduct electricity industry relevant R&D.

Table 1 presents summary statistics for unit-level operating characteristics that significantly determine NOx emissions levels. To construct this table, units are classified as either "high damage" (above average) or "low damage" (below average) units. Overall, these unit-level characteristics are very similarly distributed similarly across the two groups. Table 2 summarizes the unit-specific estimates of the capital and variable costs for the most commonly adopted NOx control technologies. These estimates will be used in our modeling of firm-level compliance decisions.

4.3 Simulating facility-level compliance decisions

In the policy modeling that informs the design and implementation of market-based emissions regulations, it is standard to assume strict cost minimization on behalf of all firms (IPM, ISIS). This is the assumption we maintain in the theory model introduced in Section 2. However, previous studies have noted that cost minimization algorithms provide a very crude and possibly inaccurate approximation of real-world decision-making (Krupnick et al. 2000). If firm's environmental compliance choices deviate significantly from cost-minimization algorithms, standard policy simulation models will innacurately predict the gains from differentiation.

In the interest of exploring the extent of this inaccuracy, we conduct two sets of policy simulations. We first specify a stylized cost minimization model that is intended to capture the salient features of workhorse policy simulation models. The second approach incorporates the same basic modeling structure, but replaces the cost minimization algorithm with an econometrically estimated model of the compliance choice.

4.3.1 Cost-minimization algorithm

Our first approach to simulating firms' compliance decisions under the observed and counterfactual policy designs uses a simple algorithm to find the combination of NOx control options that minimizes the cost of meeting the emissions cap. Let $j = 1...J_i$ index the NOx control technology options available to the ith electricity generating unit. Let K_{ij} represent the engineering cost estimate of the required capital investment specific to unit i and technology j. Let V_{ij} represent the corresponding variable operating cost estimate (per kWh). Let m_{ij} represent the corresponding post-retrofit emissions rate. Let e_{i0} represent the pre-retrofit emissions rate; this is the amount of NOx the ith unit emits per kWh of electricity generated if it installs no new pollution controls.

In the baseline, undifferentiated policy regime, we calculate the ex ante expected annual compliance cost associated with unit i and compliance strategy j as follows:

$$\begin{aligned} \min_{j} \quad &: \quad C_{ij} = v_{ij}Q_i + l_iK_{ij} \\ where \quad v_{ij} \quad &= \quad (V_{ij} + \tau m_{ij})Q_i. \end{aligned}$$

Capital investments K_{ij} are converted to annual costs using a levelized annual cost factor l_i. Unit-specific

time horizons are constructed by subtracting the unit age from the assumed life span (50 years).[15] The average investment time horizon is 13 years. We assume a discount rate of 5.34 percent. This rate, which was derived from financial data for electric utilities, is used in the modeling of investments in environmental retrofits conducted by the US EPA.[16] Expected annual operating costs v_{ij} are obtained by multiplying estimated per kWh operating costs by expected seasonal production Q_i. Historic electricity production during the ozone season, Q_i, is used to proxy for expected ozone season production.[17] To estimate the annual variable compliance cost v_{ij}, the technology operating costs are added to the expected costs of holding permits to offset any remaining emissions. NOx permits were trading throughout the period of time that these compliance decisions were being made. We use the average permit price that prevailed during the period prior to the NBP compliance deadline as a proxy for what managers' expected cost of offsetting uncontrolled emissions.

To simulate outcomes under the differentiated tax, the τ parameter in (**??**) is replaced with the source-specific expected damage parameter δ_i. In the differentiated regime that incorporates wedges, the variable compliance costs are redefined to be: $(V_{ij} + \tau m_{ij} + w_i m_{ij})Q_i$. Finally, to simulate outcomes under the counterfactual policy that incorporates first-best ratios, variable compliance costs are redefined as $(V_{ij} + \tau r_i m_{ij})Q_i$.

The process for simulating permit market clearing is as follows. The cap is set equal to the seasonal NOx emissions associated with the compliance choices we actually observe (597,000 tons).[18] Beginning with an initial permit price τ^0, we find the compliance option j^0 at each electricity generating unit that minimizes (**??**). The ozone season NOx emissions associated with these choices are summed across units. If these aggregate emissions exceed (fall below) the cap, the permit price is incrementally increased (reduced) and the process is repeated until the aggregate emissions constraint is just satisfied.[19]

The top two panels in Table 3 summarize the compliance choices associated with the simulated equilibrium under the observed (i.e. undifferentiated) policy regime. The top row shows that a majority of units chose to rely on the permit market exclusively for compliance (the "no retrofit" option). A majority of the mandated emissions reductions were achieved using highly capital intensive selective catalytic reduction (SCR) technologies. The middle panel shows that the cost minimization algorithm poorly predicts the

[15] Note that, in addition to treating the retirement decision as exogenous, we are not attributing any costs to NOx reductions from the new plants replacing these units once they retire. This is equivalent to assuming that the new capacity investment will comply with new source standards, and that the cap will cease to bind as these new plants make up a larger share of the fleet.

[16] For more complete documentation of this model and its assumptions, see http://www.epa.gov/airmarkt/progsregs/epa-ipm/. This discount rate is slightly lower than the 6 percent assumed by Carlson et al. (2000).

[17] Anecdotal evidence suggests that managers used past summer production levels to estimate future production (EPRI, 1999a). We adopt this approach and use the historical average of a unit's past summer production levels ($\bar{Q}_n$) to proxy for expected ozone season production.

[18] The estimated emissions associated with observed compliance choices exceed the emissions levels that were actually observed. In 2004, the first year of full compliance, NOx emissions from coal units were 564,000 tons. Emissions levels dropped below 500,000 tons in later years (US EPA, 2007). One possible explanation for this discrepancy is that many units that made no capital investment in abatement equipment were able to make extensive small-scale improvements to reduce emissions intensity. Linn (2008) estimates that 10-15 percent of emissions reductions were the product of these small process changes and modifications.

[19] If this iterative procedure arrives at a point where it is vascillating around the cap, the price that delivers the quantity of emissions just below the cap is chosen to be the equilibrium price. Equilibrium emissions are calculated and the simulation stops.

compliance choices that firms actually made. The cost minimization model correctly predicts compliance choices at only 24% of the EGUs covered in the analysis. In particular, the model overestimates the share of less capital intensive combustion modifications and underestimates the share of capital intensive SCR retrofits.

In sum, we find that observed outcomes in the NOx Budget Program depart markedly from those consistent with strict cost minimization of ex ante expected abatement costs. Carlson et al. (2000) document similar discrepancies in the context of the Acid Rain Program. If the cost minimization-based policy simulations incorrectly predict how firms respond to the observed (undifferentiated) policy, this model cannot be relied upon to accurately simulate firms' response t o differentiated policy incentives. For this reason, we pursue a second (preferred) approach to modeling firms' compliance decisions which is designed to more accurately capture the real world distortions and idiosyncracies that determine firms' environmental compliance choices.

4.3.2 An econometric model of the compliance decision

Fowlie (2010) estimates an econometric model of the compliance choices made by plant managers in the NBP. We use this model to simulate the compliance decisions that plant managers in the NBP under the observed (undifferentiated) and counterfactual (differentiated) regimes. The decision maker at unit i is assumed to choose the compliance strategy that minimizes the unobserved latent value C_{ij} :

$$\begin{aligned} C_{ij} &= \alpha_j + \beta_m^v v_{ij} + \beta_m^K K_{ij} + \beta^{KA} K_{ij} \cdot Age_{ij} + \varepsilon_{ij}, \\ where \quad v_{ij} &= (V_{ij} + \tau m_{ij}) Q_i \end{aligned} \tag{18}$$

Note that this specification is very similar to Eq. (18). The primary difference is that parameters are econometrically estimated versus calibrated. The deterministic component of C_{ij} is a weighted sum of expected annual compliance costs v_{ij}, the expected capital costs K_{ij} associated with initial retrofit and technology installation, and a constant term α_j that varies across technology types. The technology fixed effects are intended to capture average biases for or against particular types of NOx control equipment. An interaction term between capital costs and demeaned plant age is included in the model because older plants can be expected to weigh capital costs more heavily as they have less time to recover these costs. Expected annual compliance costs are obtained by multiplying estimated per kWh variable costs by expected seasonal production Q_i. We maintain the assumption that expected seasonal electricity production (Q_n) is independent of the compliance strategy being evaluated.

With some additional assumptions, this model can be implemented empirically as a random-coefficients logit (RCL) model. More specifically, the ε_{nj} are assumed to be *iid* extreme value and independent of the covariates in the model. The variable cost coefficient (β^v) and the capital cost coefficient (β^K) are allowed to vary randomly in the population according to a bivariate normal distribution, thereby accommodating any unobserved heterogeneity in responses to changes in compliance costs.[20] The econometric model is

[20] It is common in the literature to assume that cost coefficients are lognormally distributed, so as to ensure the a priori expected negative domain for the distribution (with costs entering the model as negative numbers). Model specifications that assumed a log-normal distribution for cost coefficient failed to converge.

estimated separately for units serving restructured wholesale electricity markets versus publicly owned units and units subject to cost-of-service regulation. A more detailed description of the econometric specification and estimation results can be found in Fowlie (2010).

An electricity generating facility or "plant" can consist of several physically independent generating units, each comprising of a boiler (or boilers) and a generator. The 632 boilers in our data represent 221 power plants. Presumably, the same plant managers make compliance decisions for all boilers at a given plant. To accommodate correlation across choices made by the same plant managers, the β_m coefficients are allowed to vary across managers according to the density $f(\beta|b, \Omega)$, but are assumed to be constant across choices made by the same manager. Estimates of the parameters of the distribution of β^v and β^K in the population of managers can be combined with information about observed choices in order to make inferences about where in the population distribution a particular decision maker most likely lies (Allenby and Rossi, 1999; Revelt and Train, 2000; Train, 2003). We use the means of these plant-manager specific distributions, versus the population means, to parameterize our policy simulation model. This should improve our ability to simulate the choices that these plant managers would have made in counterfactual policy scenarios.

Table 4 summarizes the the parameter estimates that define the policy simulations. The top panel reports the estimated technology specific fixed effects. These are all negative, suggesting that the average plant manager was biased against emissions abatement technology retrofits (vis a vis the compliance option that relies exclusively on purchasing permits). The bottom panel reports the means of the manager-specific distributions of the two cost coefficients (β^K and β^v).The ratio $(\beta^K + \beta^{KA} Age) : \beta^v$ is of particular interest as it can be interpreted as a measure of how a plant manager trades off fixed capital costs (i.e. investments in NOx control equipment) and variable compliance costs (including the cost of holding permits to offset uncontrolled emissions each year). Point estimates of this ratio (computed using the estimated means of the manager specific conditional distributions) are 0.48 and 0.21 among managers of deregulated and regulated units, respectively. As compared to the cost minimization model, these econometric estimates imply that plant managers were more strongly biased against more capital intensive compliance options.[21]

With this econometrically estimated compliance choice model in hand, our approach to simulating permit market outcomes is mechanically very similar to cost minimization exercise described above. We implicitly assume that the fundamental structure of the firm-level compliance decisions we model would not change under a differentiated regime. This seems very plausible. We see no reason why managers willingness to trade off annual operating costs and upfront capital investment, and/or managers' preferences for or against particular pollution control technologies, should be impacted by policy differentiation.
The bottom panel of Table 3 summarizes the equilibrium choice probabilities under the observed (undifferentiated) policy regime. For the purpose of comparing choice model predictions against observed compliance choices at the unit-level, we define the choice with the highest simulated choice probability as the simulated choice. Whereas the cost minimization model predicts that the emissions cap would be met with hundreds of relatively small investments in pre-combustion controls and modifications, the

[21] The high discount rate among plants serving restructured electricity markets likely reflects significant credit rating downgrades which affected several firms during the time period in which plant managers were having to make their compliance decision.

econometric model correctly predicts that a significant portion of the mandated emissions reductions is achieved using more capital intensive compliance strategies.[22]

4.4 Estimating the costs of compliance

To meaningfully assess the welfare impacts of policy differentiation, the gains (in the form of reduced damages) must be weighed against any increase in compliance costs. This requires estimates of the abatement costs that were actually incurred under the undifferentiated policy and the costs that would have been incurred under the counterfactual, differentiated policies we consider. [23]

Accurate measurement of the costs of complying with environmental regulation is notoriously difficult. Some regulations involve costs that are not readily observable in market transactions. When costs are observable in principle (e.g. investment in pollution abatement equipment), it is difficult to obtain information on the actual expenditures of the plants subject to the regulation. Moreover, the compliance costs that are attributable to the policy of interest can be difficult to parse out. Plant activities undertaken to comply with one environmental regulation can increase (or decrease) production efficiency, or make it more (or less) difficult to comply with other regulations. In light of these difficulties, retrospective analyses of environmental regulations often under-emphasize, or ignore completely, the costs of compliance.(Morgenstern, 2011).

In light of these constraints and limitations, our estimates of the costs of regulatory compliance are ineluctably imperfect. We take two different approaches to constructing our cost estimates. The first is intended to estimate the costs as perceived by the policy maker. The second is intended to estimate the compliance costs from the perspective of the firm. In theory, these two measures can differ if pre-existing distortions, market failures, cognitive biases and/or optimization errors on the part of the plant manager a wedge between private and social costs. We first explain how these cost estimates are constructed. We then discuss some key limitations.

4.4.1 Compliance costs from the perspective of the policy maker

To estimate the social costs associated with the simulated compliance choices, we use the same cost assumptions and accounting parameters that were used to parameterize equation (**??**). The levelized annual cost of compliance under policy regime r is defined to be:

$$LAC_r^{CM} = \sum_i V_{ir} Q_i + l_i K_{ir},$$

[22] One possible explanation for the apparent over-investment in more capital intensive compliance options could be the regulatory incentives faced by public power authorities and plants operating under rate-of-return regulation (Fowlie, 2010; Sotkiewitz, 2008),

[23] These counterfactual costs present even more of a challenge because they involve compliance choices that these plants did not actually make. Using data on realized abatement costs to impute the costs of abatement options that were not chosen requires either a valid exclusion restriction or strong assumptions regarding the conditional independence of observed costs and unobserved determinants of the compliance decision (Keohane, 2004).

where V_{ir} and K_{ir} are the boiler-specific, technology-specific cost estimates associated with the compliance option chosen by firm i under policy regime r.

When the econometric model is used to simulate compliance choice probabilities, this cost estimate is defined to be:

$$LAC_r^{EST} = \sum_i \sum_j P_{irj}(V_{ir}Q_i + l_i K_{ir}),$$

where P_{irj} denotes the simulated choice probability associated with unit i and choice j in regime r.

These cost measures are intended to measure costs as anticipated by the policy maker. The V_{ij} and K_{ij} values reflect the cost and emissions reduction efficiency information that was available in the years immediately preceding the NBP. The levelized cost factor is calibrated to match the assumptions underlying a standard policy simulation model (IPM).

4.4.2 Compliance costs from the perspective of the regulated firm

An alternative approach identifies the cost estimates are most consistent with the econometrically estimated model of the compliance choices that firms actually made. Figure 6 provides an intuitive illustration of how these cost estimates are constructed. Both the cost minimization model and the econometrically estimated choice model can be used to simulate permit market outcomes, and equilibrium permit prices in particular, over a range of emissions limits or caps. Plotting the simulated permit prices against the corresponding emissions reductions traces out an aggregate marginal abatement cost (MAC) curve. The horizontal axis in Figure 6 measures emissions abatement (in millions of tons of NOx per ozone season). The vertical axis measures marginal abatement costs (in \$/ton). The vertical line corresponds to the emissions cap we impose in our policy simulations. The lower MAC curve is generated using the model that assumes strict cost minimization. The more steeply sloped MAC curve is generated using the econometric model of the compliance choice.

Note that differences in the two models we use to simulate compliance choices beget significant differences in the corresponding MAC curves. The econometric model captures negative biases against specific NOx control technologies (in the negative technology fixed effects) and a reduced willingness to take on large capital investments in exchange for reducing annual variable compliance costs (captured by the ratio of β_k and β_v coefficients). The econometric model thus predicts a relatively muted response (in terms of a firms' choice of emissions level) to a given change in the permit price. This manifests as a more steeply sloped abatement cost curve.

Integrating under the higher marginal abatement cost curve in Figure 6 over the range of abatement activities required by the cap obtains an estimate of the aggregate cost of compliance under the undifferentiated policy regime that is consistent with the compliance choices we simulate using the econometrically estimated model. This estimate more accurately captures the compliance costs as perceived by the regulated

firms[24] As compared to (4.4.1), these cost estimates presumably capture some additional information about the costs of the compliance choices that were actually made.

As noted above, neither of these approaches are perfect. Both rely to a significant extent on the boiler-specific, technology-specific cost estimates V_{ij} and K_{ij} to estimate the true costs of installing and operating the full range of feasible compliance options. These assumed costs are almost certainly an imperfect proxy for the costs that would actually manifest. There is evidence to suggest that the costs of complying with the NOx Budget Program were lower than expected. Figure 5 plots the NOx permit price movements over the pre-compliance period and the first three years of the program. Prices in the pre-compliance period reflect the price expectations that informed major compliance and investment decisions. Once the program actually took effect, the permit price dropped and stabilized at a lower value .[25] This suggests that our cost estimates likely overestimate the realized and counterfactual compliance costs we seek to measure.

4.5 Pollution damages

AP2 is also used to quantify the change in damages due to the various policy scenarios explored in the study. In this context, rather than systematically perturbing NO_x emissions one source at-a-time, NO_x emissions change simultaneously at many of the regulated EGUs in response to the different modeled policies. Since the aggregate emissions level is held fixed across scenarios, any difference in simulated damages across scenarios is attributable to the spatial redistribution of the permitted emissions (rather than a change in the overall stringency of the policy).

Once the compliance choices associated with a particular policy scenario have been simulated, the corresponding vector of unit-level, ozone season emissions is processed through the Monte Carlo machinery. This yields a distribution of estimates of the damages stemming from each policy design. The simulations feature all input parameters (emissions, transfer coefficients in the stochastic air quality model, population, dose-response, and valuation) as random variables consisting of 5,000 possible realizations. Conditional on a policy emissions vector, one realization is selected from each input distribution, and the total exposures, physical effects, and monetary damages are computed. This is repeated 5,000 times for each policy scenario. When comparing damages across policy scenarios, resultant damages are compared for the same draws from the input distribution.

As an additional robustness check for our assumption in section 2. regarding the additively separable and linear NO_x damage function, we also compute total damages by simply multiplying marginal damages times emissions at each facility. This is repeated 5,000 times since the marginal damages also consist of 5,000 estimates for each EGU. In comparing the estimated welfare impacts across policy scenarios using

[24] Note that, by construction, these two alternative approaches to estimating compliance costs yield equivalent results when the cost minimization model is used to simulate compliance choices. Intuitively, the very same cost estimates define both the compliance cost minimization routine and the ex post computation of costs. When the econometric model is used, these two approaches yields very different results.

[25] Firms' choices regarding investment in NOx abatement equipment were made in the years leading up to the compliance deadline. A subset of states imposed a compliance deadline of May 2003. Full compliance began in May 2004.

the margin-times-emission approach against the simulations that allow emissions to change simultaneously, we are careful to align the 5,000 draws such that the same "states-of-the-world" are compared to one another.

5 Results and synthesis

This section is divided into three subsections. The first characterizes the simulated outcomes (including the allocation of permitted emissions across sources, abatement costs incurred, and total damages avoided) under the observed, undifferentiated policy regime. The second section estimates the benefits and costs of differentiation relative to this benchmark. The third examines the uncertainty in these estimates.

5.1 Benchmark case: Undifferentiated trading

Table 5 summarizes the simulated outcomes under the observed, undifferentiated regime. Results generated using the calibrated cost minimization algorithm and econometrically estimated model are reported in column (1) and (2), respectively. These outcomes serve as a benchmark for the subsequent analysis of the counterfactual, differentiated policy regimes.

The top panel reports simulated equilibrium permit prices, emissions, and benefits associated with the mandated emissions reductions. When the cost minimization model is used to simulate permit market clearing, the equilibrium permit price is remarkably close to the average damage parameter value ($1711/ton). This implies that, conditional on the assumed damages, costs, and choice model parameters, the emissions cap imposed in the NOx Budget Program is close to optimal.

The results obtained using the econometrically estimated model (reported in column 2) tell a different story. Figure 6 illustrates how the marginal abatement cost curve implied by the econometrically estimated model is much steeper than the MAC curve generated using our more stylized cost minimization model. Consequently, the simulated permit price under undifferentiated trading is much higher ($4,460/ton NOx). Figure 5 shows that this simulated price equals the average permit price that prevailed during the time that compliance decisions were made.

Table 5 also provides information regarding the distribution of permitter emissions across high and low damage sources. Facility-level data collected following the introduction of the NBP indicate that approximately 38 percent of permitted NOx emissions occurred at sources with higher than average damage parameters. The cost minimization model over-predicts the share of emissions occurring at these high damage sources (42 percent). The econometric model allocates 39 percent of permitted emissions to high damage firms.

The point estimate of the annual benefits (in terms of avoided damages) accruing from the mandated emissions reductions is approximately $1.1 billion per year. The numbers in parentheses represent the 5th and 95th percentile estimates from the Monte Carlo simulation. Strictly speaking, these should not be

interpreted as confidence intervals because our analysis does not account for all sources of uncertainty.[26] These estimates differ slightly across these two sets of simulations. There are two reasons for this. The first has to do with the differences in how permitted emissions are allocated across sources. The second is due to non-convexities in the MAC curves that result in slight differences in total emissions observed under the same emissions constraint.

The lower panel in Table 5 reports the estimated costs of complying with the observed, undifferentiated policy. The levelized annual costs associated with the compliance choices predicted by the cost minimization model add up to $468 M. By construction, our two approaches to estimating this aggregate compliance cost arrive at the same cost number.

The cost estimates in column (2) are higher. Note that the cost estimate obtained using Equation (4.4.1) is lower than the cost estimate that rationalizes the simulated compliance choices. The higher estimate is most consistent with the costs as perceived by the firms. If the difference between the two cost estimates is attributable to distortions and idiosyncracies (such as status quo bias or price discrimination on the part of pollution control equipment manufactures) that .drive a wedge between private and social costs, then Equation (4.4.1) provides a relatively more appropriate measure for use in our welfare calculations. However, it is also possible that (4.4.1) mischaracterizes the true social costs of compliance. For example, these calculations may underestimate the true cost of financing investments in pollution abatement equipment. We interpret the higher estimate as an upper bound on the aggregate compliance costs that were actually incurred under the undifferentiated regime.

Our most conservative estimate of the net benefits (i.e. the value of avoided damages less abatement costs) conferred by the NOx Budget Program is $214 million per year. The estimate obtained using the calibrated model of cost minimization is $602 million per year.[27]

5.2 Policy counterfactuals

We consider three counterfactual, differentiated policy designs: the first best differentiated tax; an emissions trading program that incorporates wedges; and a trading program that incorporates the damage-differentiated ratios. The emissions cap is held constant across all emissions trading programs.

5.2.1

Figures 7 and 8 provide an intuitive introduction to our results. Figure 7 illustrates how source-specific incentives to reduce emissions vary across the four policy regimes we analyze. The top panel corresponds to the more stylized cost minimization model. Data summarized in the bottom panel pertain to the

[26] For example, we do not capture modeling uncertainty or cost uncertainty in our analysis.

[27] Comparisons with other studies are confounded by differences in underlying modeling assumptions. This caveat notwithstanding, it is instructive to compare our estimate with Burtraw et al (2003) who estimate the net benefits from the NBP under a range of assumptions regarding program scope, market structure, and valuation parameters. These authors predict annual net benefits of $440 million ($1997) per year.

econometrically estimated simulation model. In each panel, the horizontal axis measures the source-specific damage parameters (the expected values). The vertical axis measures the cost to offset one unit of NOx emissions (measured in \$/ton).

In the undifferentiated regime, the cost to offset emissions is simply the permit price. The source-specific incentives under the undifferentiated policy are aligned horizontally at the permit prices reported in Table 5. Under the differentiated tax, the cost to offset a ton of NOx equals the source-specific marginal damage estimate. These source-specific incentives are identically distributed diagonally in the top and bottom panels of Figure 7. Under the wedge-based design, source i must pay $(\delta_i - \overline{\delta}) + \tau^w$ per ton of emissions to remain in compliance, where τ^w is the equilibrium permit price. Under the ratio-based regime, the cost per ton of emissions is $\frac{\delta_i}{\overline{\delta}}\tau^r$.

Recall that the imposed cap is close to optimal, conditional on the assumed costs, damages, and choice model parameters assumed in the calibrated model. When the emissions cap has been optimally set, the cost to offset a ton of emissions at firm i will be equal to δ_i.in all three of the differentiated policy desings. This is approximately true in the top panel of Figure 7.

Conditioning on the econometrically estimated choice model parameters, the imposed emissions cap is too strict. The market clearing permit price will therefore exceed the average marginal damage value $\overline{\delta}$. This will generate variation in the source-specific policy incentives across differentiated regimes. The incentives under the wedge-based policy lie along a parallel line above those of the differentiated tax. When trading ratios are used to equate ratios of marginal damages with ratios of marginal costs, differences in marginal abatement costs exceed differences in marginal damage parameters. These incentives form the steep line in Figure 7.

Figure 8 provides a graphical summary of the policy-differentiation-induced shift of permitted emissions from high to low damage sources. The vertical axes in these figures measure the change in emissions (in percentage terms) moving from an undifferentiated to a differentiated policy. A positive percentage change indicates that the emissions level chosen in the undifferentiated regime exceeds the emissions level chosen in the differentiated regime. Equation (7) suggests these changes should depend on the degree of heterogeneity in the source specific damage parameters (which is significant in the NBP) and the steepness of the marginal abatement cost curves. The horizontal axis measures source-specific marginal damage estimates. The circular markers correspond to the wedge-based differentiated policy. The triangular markers correspond to the differentiated policy that incorporates first best ratios. Each marker corresponds to a different electricity generating unit. Intuitively, differentiation increases (decreases) emissions at units with below-average (above-average) damage parameters.

The left panel of Figure 8 summarizes the emissions changes implied by the cost minimization model. The reallocation of emissions is fairly significant, with some low (high) damage units increasing (reducing) emissions by more than 50 percent under the differentiated policy (relative to the undifferentiated regime). Taken together, policy differentiation moves an estimated 14 to 15 percent of permitted emissions from high damage sources to low damage sources.

The econometrically estimated simulation model predicts a smaller reallocation of emissions (illustrated in the right panel of Figure 8). Intuitively, recall that steeper abatement costs imply a smaller response to policy differentiation, all else equal. Note also that the simulated reallocation of permitted emissions is more significant under the trading ratio regime as compared to the trading wedge regime. This difference occurs because the ratio design exagerrates the variation in the source-specific incentive to reduce emissions relative to the wedges (see Figure 7).

Table 6 provides a more detailed numerical summary of the results generated using the calibrated cost minimization model. All costs and benefits are expressed relative to the undifferentiated policy baseline. Note that emissions under the first-best tax define the optimal cap which is within 5 percent of the imposed cap. Appendix 1 shows that the equilibrium permit price will equal the average damage parameter when the cap is optimally set. The simulated equilibrium permit prices reported in Table 6 are fairly close to the average damage parameter value ($1711/ton NOx).

If the emissions constraint is close to the optimal, the advantage of using wedges over first-best ratios (in terms of efficiency) will be small. The estimated benefits (in terms of reduced damages from pollution) are similar across all three differentiated policy regimes. The expected annual benefits range from $161 M to $174M. In percentage terms, we estimate that policy differentiation reduces the damages associated with permitted emissions by 26-30 percent.

Abatement costs increase under policy differentiation (vis a vis the undifferentiated benchmark). Once this cost increase is accounted for, the net benefits of policy differentiation are approximately $82 M/year. The wedge-based design slightly dominates the ratio-based design, as we should expect given that the cap is not exactly optimal. In equilibrium, a majority of permitted emissions occur at low damage sources. Consequently, under the differentiated policy that incorporates wedges, the revenues paid to relatively low damage sources exceeds the compliance-based revenues collected from relatively high damage sources by an estimated $150 million. Although this transfer from the implementing agency to low damage sources is of no consequence for our welfare calculations, it could significantly impact the political palatability of wedges versus ratios.

Table 7 summarizes the results of the simulations that incorporate the econometric model of firms' compliance choices. The emissions level under the first-best tax is 34 percent higher than the emissions cap imposed in the trading simulations. This highlight s the extent to which the imposed cap is too stringent. Higher levels of emissions under the first-best tax imply higher damages. The wedge-based regime increases expected program benefits (in terms of avoided emissions) vis a vis the benchmark, but only by 3 percent. The ratio-based policy increases the expected benefits of the program by an estimated 8 percent.

When we use equation (4.4.1) to estimate compliance costs, the expected net benefits from differentiation appear *larger* under the design that incorporates ratios (versus wedges). This runs counter to observation (3) which suggest that wedges should be preferred to ratios when the cap is not set optimally. This seemingly counterintuitive result is explained by the fact that the abatement costs that rationalize the simulated compliance choices are different from the costs we use to construct the cost estimate in equation

(4.4.1). More specifically, marginal abatement costs as percieved by the firms are steeper on average. When the wedge-based policy is introduced, firms equate *perceived* marginal abatement costs with $(\tau+w_i)$. Consequently, a sub-optimal quantity of permitted emissions are reallocated from high-damage to low-damage sources. The constrained optimum is not obtained. Damage-based ratios exagerrate the difference in source-specific incentives to reduce emissions (see Figure 7). This induces a larger shift in emissions which is cost effective if costs are measured as equation (4.4.1).

In contrast, when we our welfare calculations incorporate the cost estimates that are consistent with the simulated compliance choices, the wedge-based policy welfare dominates the ratio-based regime. Notably, the ratio-based regime actually *reduces* welfare relative to the undifferentiated benchmark because too much of the permitted emissions are reallocated to low damage sources.

It is difficult to know which set of cost estimates is more appropriate for use in estimating the net gains from policy differentiation. On the one hand, it is plausible that firms' compliance choices were influenced by pre-existing distortions and idiosyncratic errors. On the othe hand, it seems unlikely that all of the differences between the abatement cost estimates constructed using equation (4.4.1) and those consistent with observed choices can be wholly attributed to artificial distortion and inefficiency. The net gains reported in the last row of Table 7 are likely to under-estimate the net benefits from policy differentiation.

5.3 Uncertainty

Thus far, this results summary has focused almost exclusively on point estimates. In this section, we characterize the uncertainty surrounding these estimates. What we are analyzing amounts to a policy lottery in which probabilities can be attached to a range of possible benefit outcomes. The simulated effects of policy are uncertain because the specific parameters of the model are uncertain. Figure 8 illustrates the distribution of our estimates of the gains from policy differentiation (i.e. the reduction in damages from the permitted emissions vis a vis the undifferentiated benchmark). Each distribution summarizes 5000 realizations of benefits, each corresponding to a different vector of source-specific damage parameters $\delta^{'}$.The top panel corresponds to the wedge-based design and the calibrated cost minimization model. The bottom panel corresponds to the wedge-based design and the econometrically estimated choice model.

The vertical black lines represent our estimate of the cost of differentiation (relative to the undifferentiated, benchmark regime). We choose to summarize our results in this way in order to clearly separate the estimated benefits of differentiation from the estimated costs. Taking our estimates of costs as given, the expected net benefits of policy differentiation are positive across all scenarios we consider.

Observation (4) provides some intuition for what drives the significant variation in these benefit realizations. Recall that the average parameters $E[\delta]$ define the terms of compliance across all realizations. The covariance between $E[\delta]$ and the the realized damage vector $\delta^{'}$varies across realizations. Realized benefits from differentiation are increasing with this covariance. In cases where undifferentiated trading

welfare dominates differentiated trading (i.e. realizations lying to the left of zero), the realized damages are negatively correlated with the source-specific average damage parameters. In these instances, the differentiated policy designs perversely shift damages from low to high damage sources.

6 Conclusion

How should market-based emissions regulations, and cap and trade programs in particular, be designed and implemented when damages from emissions vary significantly across sources? To shed light on this question, we first introduce a conceptual framework that is useful for analyzing the efficiency gains from policy differentiation. We extend the theoretical work in this area so as to consider key factors that complicate real-world implementation of differentiated policy. These include jurisdictional constraints, limited information about abatement costs, and uncertainty about how to value damages from pollution. We examine how differentiated policy designs can be modified in order to accommodate these constraints and limitations.

The conceptual framework serves as foundation for an applied analysis of the gains from policy differentiation. We consider the landmark NOx Budged Program. Prior research has shown that the damages due to NO_x emissions vary considerably according to where the emission occurs. The policy design that is currently in place fails to reflect this heterogeneity. We estimate the efficiency loss (in terms of avoided damages less increased abatement cost) associated with the decision to implement an undifferentiated policy design.

We first estimate marginal damages for each of the coal-fired boilers regulated under the NBP. These damage estimates are used to parameterize counterfactual, spatially differentiated designs. An econometric model of the compliance decisions made by firms subject to the NBP is used to simulate outcomes under both the observed and two counterfactual policy designs (Fowlie, 2010). The corresponding aggregate abatement costs and environmental damage are then tabulated and compared. Importantly, total emission levels are held constant across all of the policy scenarios we consider. What changes across regimes is the allocation of these permitted emissions across sources.

Our preferred approach to estimating the gains from policy differentiation takes advantage of the fact that we can now observe the compliance options that firms actually chose when complying with the NOx Budget program. We use these data to estimate a model of how facility-level compliance choices were made. This allows us to more accurately predict how sources in the program would have responded under counterfactual policy incentives. We find that the gains from policy differentiation in the NOx Budget Program would most likely have been small. Our estimate of the net benefits from the policy increase by only $2M per year under differentiation. The confidence interval we construct around this estimate is very large due to the parameter uncertainty that pervades the integrated assessment of source specific damages.

Another important finding comes out of a direct comparison of the econometrically estimated model of firms' compliance decisions and a more stylized cost minimization model that is calibrated to match

more standard policy simulation approaches. Consistent with earlier work by Carlson et al. (2000), we find substantive differences between the observed compliance choices and those predicted by the cost minimization algorithm. In our case, this implies that the more standard approach to simulating firm-level responses to policy incentives leads to an overestimate of the gains from differentiation.

References

[1] Abbey, D.E., B.E. Ostro, F.Peterson, R.J. Burchette, 1995. "Chronic Respiratory Symptoms Associated with Estimated Long-Term Ambient Concentrations of Fine Particulates Less than 2.5 Microns in Aerodynamic Diameter and other Air Pollutants"Journal Expo Analysis Environmental Epidemiology. 5:2., pp 137-159.

[2] Banzhaf, H. Spencer, Dallas Burtraw. Karen Palmer. 2004. "Efficient emission fees in the US electricity sector." Resource and Energy Economics. 26: 317-341.

[3] Banzhaf,H.S., B. A. Chupp, "Heterogeneous Harm vs. Spatial Spillovers: Environmental Federalism and US Air Pollution, NBER Working Paper No. 15666, January 2010

[4] Baumol, William J. and Wallace E. Oates. 1988. *The Theory of Environmental Policy.* 2nd ed. Cambridge: Cambridge University Press.

[5] Bell, Michelle L., Adrian McDermott, Scott L. Zeger, Jonathan M. Samet, and Francesca Domenici. 2004. "Ozone and Short-Term Mortality in 95 US Urban Communities, 1987-2000." *Journal of the American Medical Association,* 292(19):2372-2378.

[6] Byun, Daewon, W., and Kenneth L. Schere. 2006. "Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System." *Applied Mechanics Reviews,* 59(2):51-77.

[7] Carlson, Carlson, Dallas Burtraw, Maureen Cropper, and Karen L. Palmer. 2000. "Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?," *Journal of Political Economy*, University of Chicago Press, vol. 108(6): 1292-1326.

[8] Dales, John H. 1968. *Pollution, Property & Prices: An Essay in Policy-making and Economics.* Toronto: University of Toronto Press.

[9] Farrow, R. Scott, Martin T. Schultz, Pinar Celikkol, and George L. Van Houtven. 2005. "Pollution Trading in Water Quality Limited Areas: Use of Benefits Assessment and cost effective Trading Ratios." *Land Economics,* 81(2):191-205.

[10] Hoel, M., L. Karp, 2000. Taxes and Quotas for a Stock Pollutant with Multiplicative Uncertainty. UC Berkeley: Department of Agricultural and Resource Economics, UCB. CUDARE Working Paper 870. Retrieved from: http://escholarship.org/uc/item/9v86p5s7

[11] Krupnick, Alan, Virginia McConnell, Terrell Stoessell, Matthew Cannon, and Michael Batz. 2000. "Cost-Effective NOx Control in the Eastern United States," Discussion Papers dp-00-18, Resources For the Future.

[12] Klaassen, Ger A.J., Finn R. Forsund, and Markus Amann. 1994. "Emissions Trading in Europe with an Exchange Rate." *Environmental and Resource Economics,* 4:305-330.

[13] Lippman, Morton, and Richard B. Schlessinger. 2000. "Toxicological Bases for the Setting of Health-Related Air Pollution Standards." *Annual Review of Public Health*, 21: 309-333.

[14] Linn, Joshua. 2008. "Technological Modifications in the Nitrogen Oxides Tradable Permit Program." *The Energy Journal ,* 29(3):153-176.

[15] Mauzerall, Denise, Babar Sultan, Namsoug Kim, and David F. Bradford. 2005. "NO_x Emissions from Large Point Sources: Variability in Ozone Production, Resulting Health Damages and Economic Costs." *Atmospheric Environment,* 39: 2851-2866.

[16] Mendelsohn, Robert O. 1986. "Regulating Heterogeneous Emissions." *Journal of Environmental Economics and Management*,13:301- 312.

[17] Montgomery, W. David. 1972. "Markets in Licenses and Efficient Pollution Control Programs." *Journal of Economic Theory,* 5:395-418.

[18] Muller, Nicholas Z. and Robert O. Mendelsohn. 2007. "Measuring the Damages from Air Pollution in the United States." *Journal of Environmental Economics and Management*, 54(1):1-14.

[19] Muller, Nicholas Z. and Robert O. Mendelsohn. 2009. "Efficient Pollution Control: Getting the Prices Right." *American Economic Review* 99(5):1714-1739.

[20] Muller, Nicholas Z. 2011 "Linking Policy to Statistical Uncertainty in Air Pollution Damages." The B.E. Press Journal of Economic Analysis & Policy. Vol. 11 : Iss. 1 (Contributions), Article 32.

[21] Pope, C. Arden, Richard T. Burnett, Michael J. Thun, Eugenia E. Calle, Daniel Krewski, Kazuhiko Ito, and George D. Thurston. 2002. "Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution." *Journal of the American Medical Association,* 287(9):1132-1141.

[22] Seinfeld, J.H., S.N. Pandis. 1998 "Atmospheric Chemistry and Physics." John Wiley & Sons, Inc., NY, NY, USA.

[23] Stavins, Robert N., 2005. "Lessons Learned from SO2 Allowance Trading," *Choices*, Agricultural and Applied Economics Association, vol. 20(1).

[24] Tong, Daniel Q., Nicholas Z. Muller, Denise L. Mauzerall, and Robert O. Mendelsohn. 2006. "Integrated Assessment of the Spatial Variability of Ozone Impacts from Emissions of Nitrogen Oxides," *Environmental Science and Technology,* 40(5):1395-1400.

[25] Turner, D.B., 1994. "Workbook of Atmospheric Dispersion Estimates - An Introduction to Dispersion Modeling." 2nd Ed. Lewis Publishers, Ann Arbor, MI, USA.

[26] US Environmental Protection Agency (USEPA). 2006. *National Emissions Inventory (NEI): 2002.* Washington, DC: Office of Air Quality Planning and Standards, Emissions Inventory Group; Emissions, Monitoring, and Analysis Division.

[27] US Environmental Protection Agency. 2003. *Documentation Supplement for EPA Modeling Applications (V.2.1.6) Using the Integrated Planning Model.* Office of Air and Radiation. Washington D.C.

[28] US Environmental Protection Agency. 1999. *Regulatory Impact Analysis for the Final Section 126 Petition Rule*, Office of Air and Radiation, Washington DC.

[29] US Environmental Protection Agency. 1998a. *Regulatory Impact Analysis for the NBP*, Office of Air and Radiation, Washington DC.

[30] US Environmental Protection Agency (USEPA). 1999. *The Benefits and Costs of the Clean Air Act: 1990 - 2010.* EPA Report to Congress. EPA 410-R-99-001, Washington, DC, Office of Air and Radiation, Office of Policy.

[31] Viscusi, W. Kip and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World." *Journal of Risk and Uncertainty,* 27(1): 5-76.

[32] Woodruff, Tracey J., Jennifer D. Parker, and Kenneth C. Schoendorf. 2006. "Fine Particulate Matter ($PM_{2.5}$) Air Pollution and Selected Causes of Postneonatal Infant Mortality in California" *Environmental Health Perspectives,* 114(5):786-790.

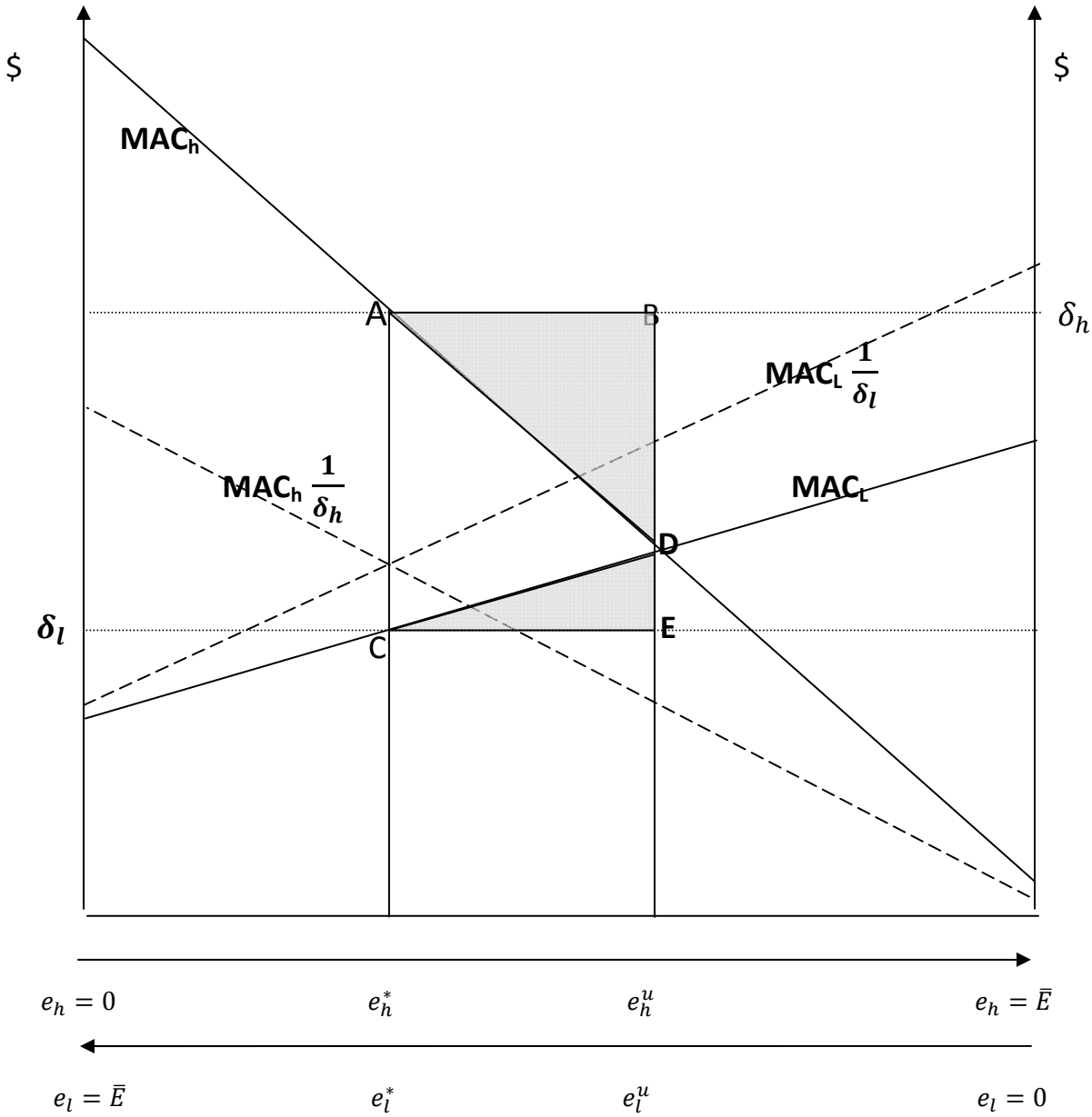

Figure 1: Emissions permit market outcomes under differentiated and undifferentiated policies: Optimal emissions constraint

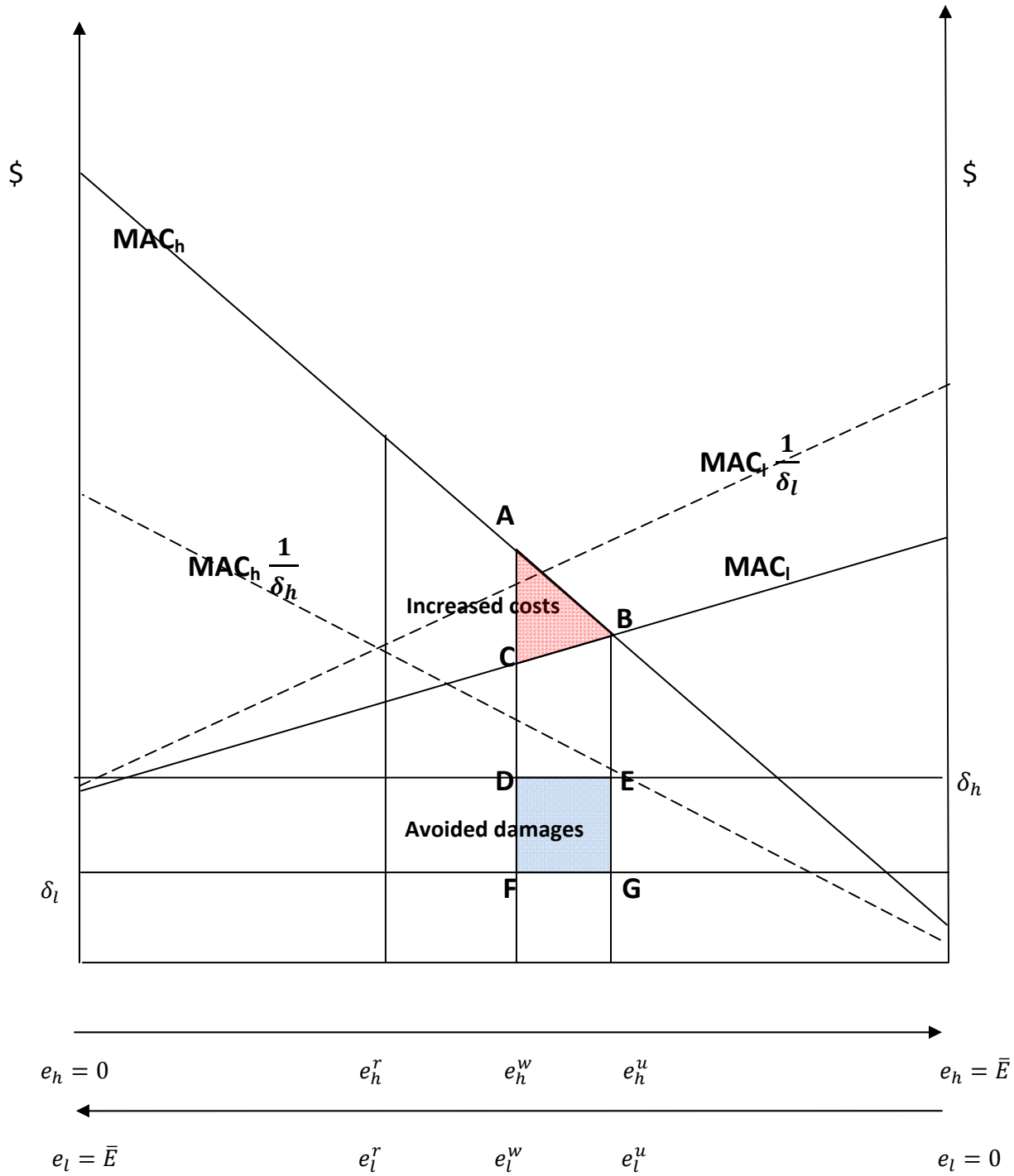

Figure 2: Emissions permit market outcomes under differentiated and undifferentiated policies: Sub-optimal emissions constraint

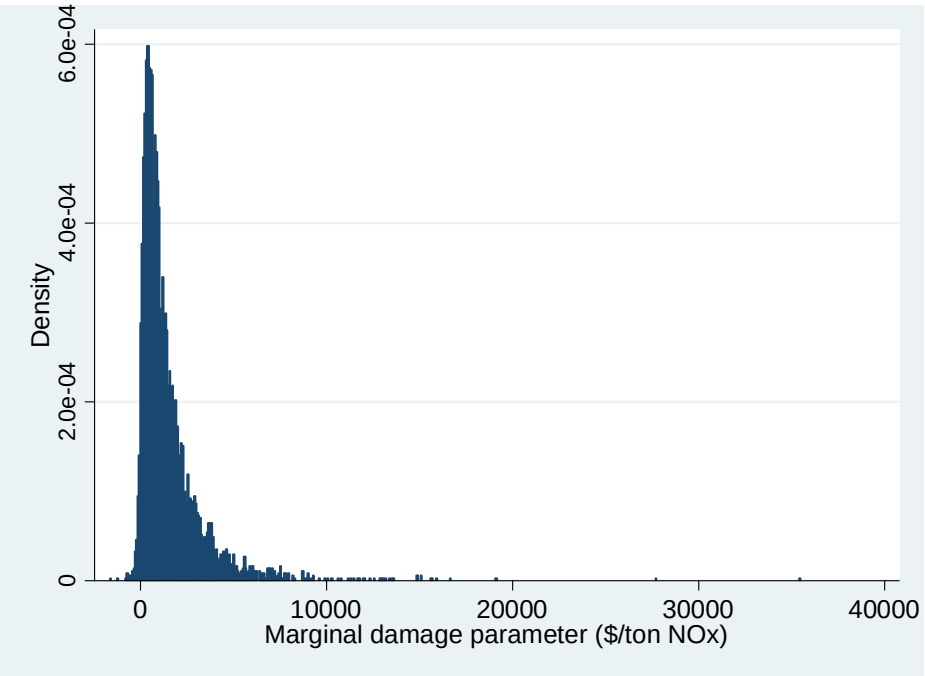

Figure 3 : Within source variation in marginal damage parameter

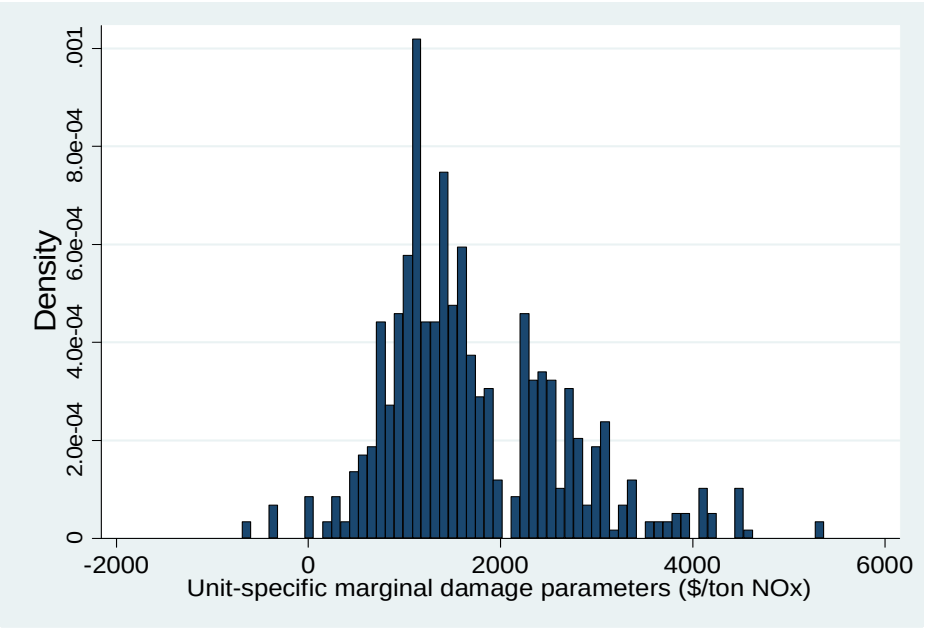

Figure 4: Between source variation in point estimates of marginal damage values

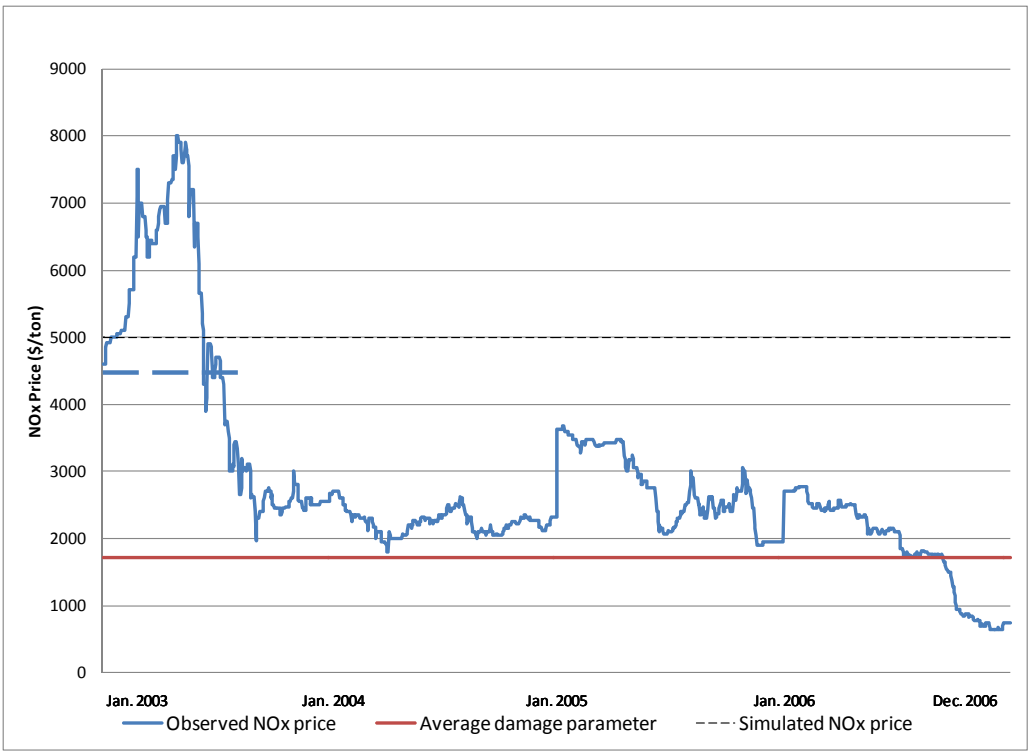

Figure 5 : NOx permit prices during the pre-compliance and compliance period

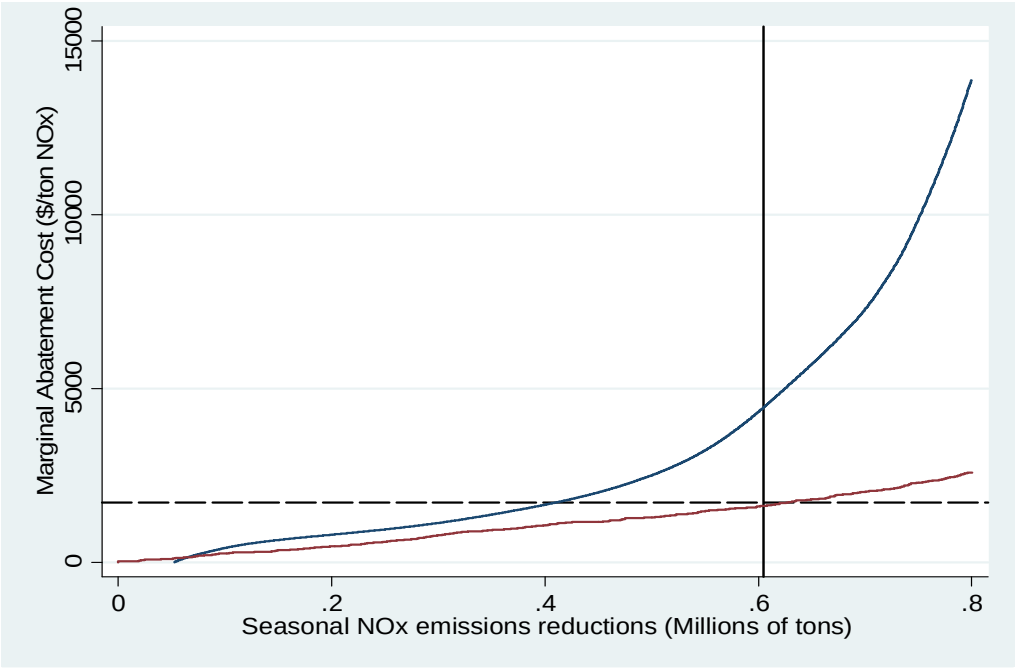

Figure 6 : Aggregate marginal abatement cost curves generated using alternative models of the facility-level compliance choice

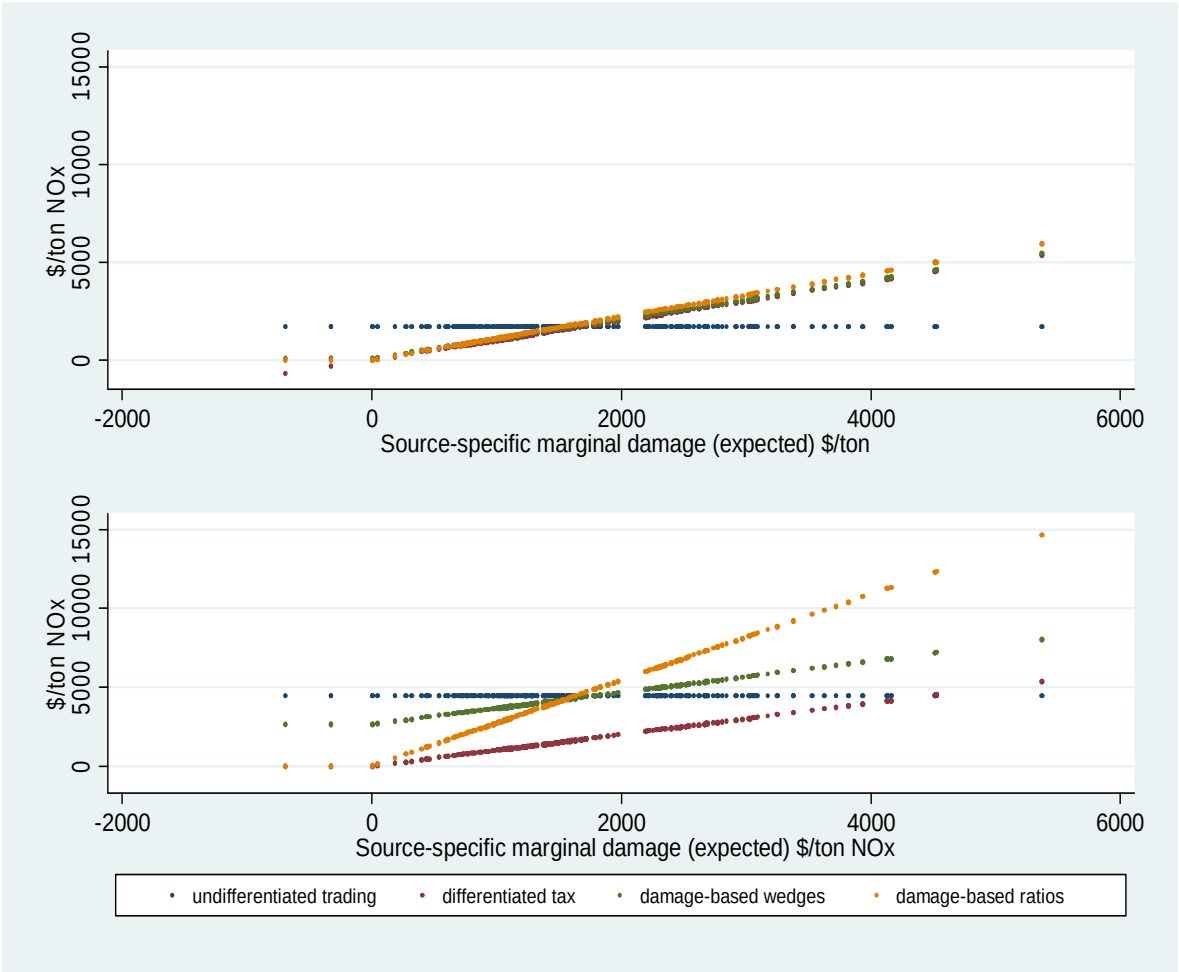

Figure 7: Reallocation of permitted emissions under differentiated NOx permit trading

Notes: The top panel illustrates the marginal emissions disincentives implied by the cost minimization algorithm. The bottom panel illustrates the marginal emissions disincentives implied by the econometrically estimated choice model. The vertical axis measures the cost of offsetting a ton of NOx under alternative policy regimes. Under the undifferentiated policy design, all firms face the same disincentive (the equilibrium permit price). Under the differentiated tax, the penalty is set equal to the source-specific marginal damage estimate. Under differentiated trading, the marginal disincentive is equal to the permit price multiplied by (added to) the trading ratio (trading wedge).

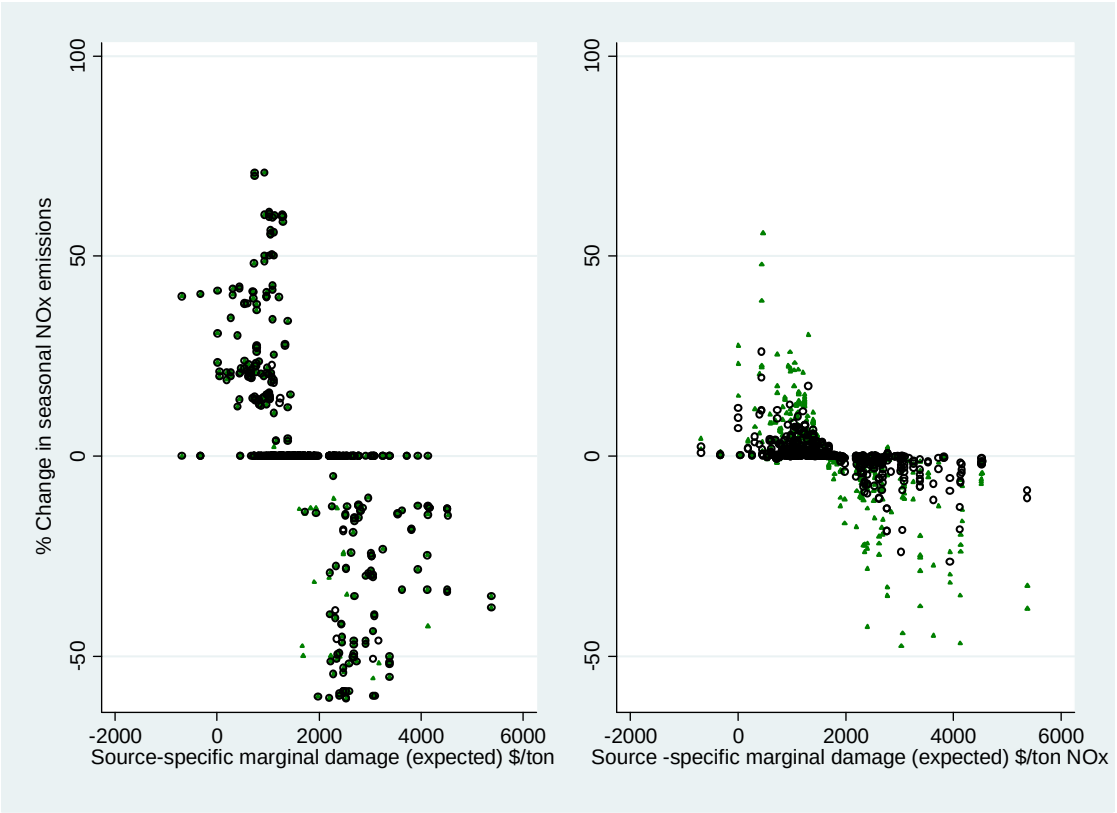

Figure 8: Reallocation of permitted emissions under differentiated NOx permit trading

Notes: The left panel summarizes emissions simulated using the cost minimization algorithm. The right panel summarizes emissions simulated using the econometrically estimated choice model. The vertical axis measures percent changes in simulated ozone season emissions in the observed, undifferentiated regime versus the simulated emissions under the counterfactual, differentiated regimes. The black circles denote emissions changes under the wedge-based design. The small triangles denote emissions changes under the ratio-based design.

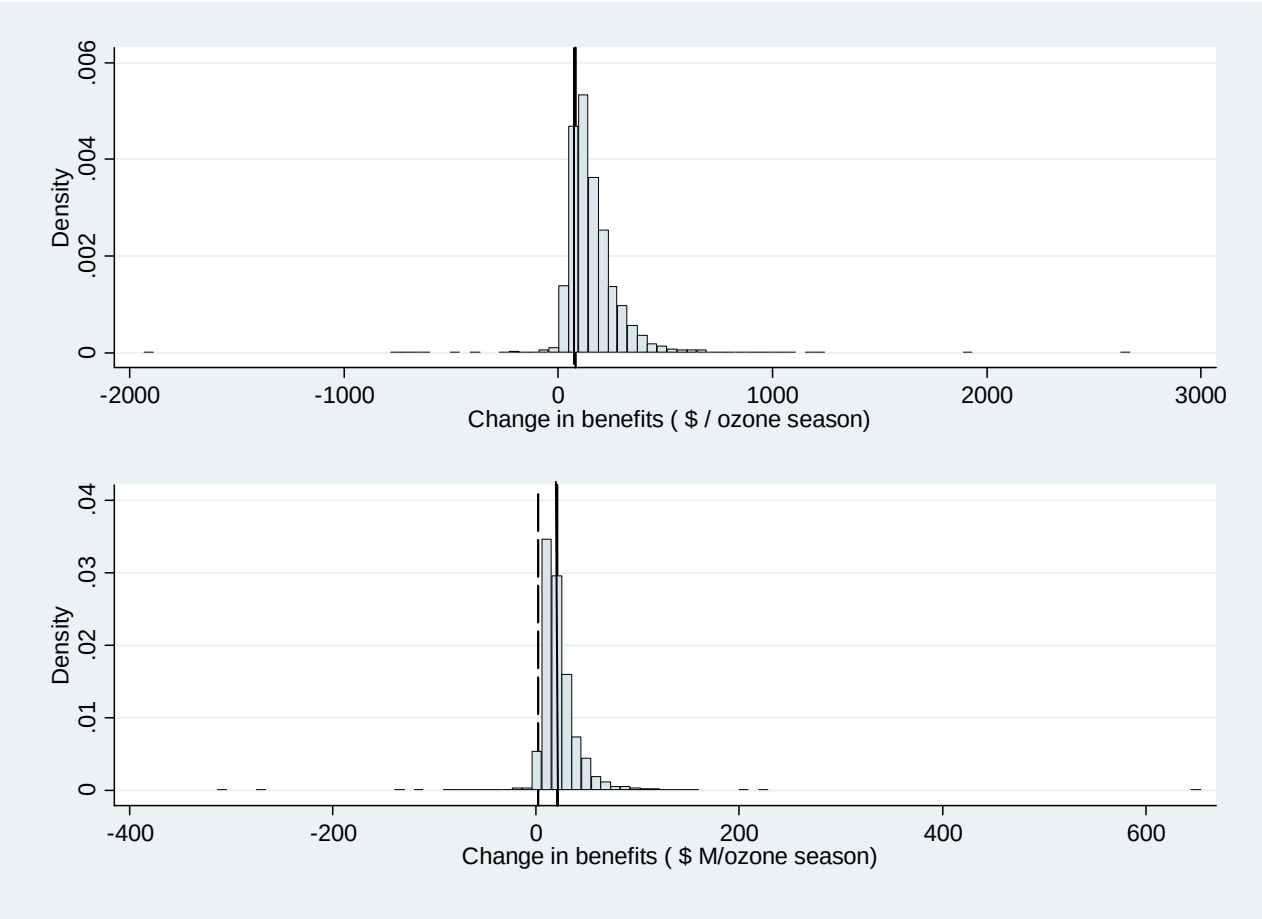

Figure 8 : Benefits of policy differentiation

Table 1 : Unit-level summary statistic

Variable	High damage	Low damage
# Units	241	391
Capacity (MW)	255.61 (234.52)	281.64 (259.84)
Pre-retrofit NOX emissions rate (lbs NOx/mmbtu)	0.55 (0.25)	0.50 (0.20)
Boiler age (years)	35.80 (10.51)	36.59 (11.53)
Summer capacity factor	65.03 (15.22)	66.07 (15.07)
Ozone season production (MWh)	780,000 (683,000)	794,000 (678,000)
Average damage parameter ($/ton NOx)	2641 (718)	1107 (405)

Notes: This table summarizes the operating characteristics of 632 coal-fired generating units regulated under the NOx Budget Trading Program. Standard deviations are in parentheses. "High damage" units are those with above average damage parameter point estimates. "Low damage" units are those with below average damage parameter point estimates.

Table 2: Compliance Cost Summary Statistics for Commonly Selected Control Technologies

NOx control technology	Capital cost ($/kW)		Variable cost (cents/kWh)	
	High damage	Low damage	High damage	Low damage
Combustion modification	6.13 (10.64)	8.12 (17.74)	1.00 (0.40)	1.00 (0.39)
Low NOx burners	17.45 (19.94)	21.98 (28.47)	0.68 (0.18)	0.65 (0.14)
SNCR	7.01 (10.09)	8.93 (11.66)	0.98 (0.37)	1.01 (0.41)
SCR	70.94 (127.99)	80.40 (155.01)	0.55 (0.19)	0.52 (0.29)

Notes: This table summarizes the ex ante predicted NOx control costs for 632 coal-fired generating units regulated under the NOx Budget Trading Program. Standard deviations are in parentheses. "High damage" units are those with above average damage parameter point estimates. "Low damage" units are those with below average damage parameter point estimates. Costs were estimated using proprietary software developed by EPRI. See text for details.

Table 3 : Observed, predicted, and correctly predicted compliance choices

Compliance choice	**SCR**	**SNCR**	**Low NOx burners**	**Combustion Modifications**	**No retrofit**	**Total**
Observed choices	187	42	53	58	292	632
Cost minimization model						
Predicted adoption rate	65	79	258	184	46	632
Correctly predicted	48	5	52	11	3	152 (24%)
Econometric model						
Predicted choices	190	16	33	22	371	632
Correctly predicted	172	8	23	18	279	500 (79%)

Notes: This table summarizes predicted and observed compliance choices for the 632 electricity generating units included in the study.

Table 4 : Econometrically estimated coefficients of the compliance choice model

Technology specific constants	High damage units	Low damage units
Post-combustion controls	-2.21 (1.66)	-3.06 (1.34)
Low NOx burners	-2.06 (0.53)	-2.33 (0.43)
Combustion modifications	-1.89 (0.85)	-2.32 (0.69)
Age* capital cost interaction	-0.17 (0.07)	-0.13 (0.06)
Manager-specific coefficients		
Annual compliance cost ($1.000,000)	-1.08 (0.81)	-0.99 (0.59)
Capital cost ($1.000,000)	-0.45 (0.43)	-0.28 (0.33)
# Units	383	269

Notes: Only point estimates are used to parameterize the simulation model. This table reports average coefficient values (averaged across facilities). Standard deviations are in parentheses. For a more detailed discussion of these econometric estimates, see Fowlie (2010).

Table 5 : Simulated outcomes under undifferentiated policy

Model of compliance choice	Cost minimization (1)	Econometric (2)
Permit price ($/ton NOx)	$1,620	$4,460
Ozone season emissions (thousand tons NOx)	656	658
Annual benefits ($M) (monetized avoided damages)	$1,070 ($184, $2,621)	$1,075 ($198, $2,636)
% permitted emissions occurring at high damage sources	41.8%	38.9%
Levelized annual abatement costs ($M) (Cost measure 1)	$468	$692
Levelized annual abatement Costs ($M) (Cost measure 2)	$468	$861
Annual net benefits ($M) (Cost measure 1)	$602 (-$284, $2,153)	$383 (-$494, $1943)
Annual net benefits ($M) (Cost measure 2)	$602 (-$284, $2,153)	$214 (-$663, $1776)

Notes: This table summarizes the results from simulating investment in NOx abatement and the associated ozone-season emissions under the observed, undifferentiated trading regime. 95 percent confidence intervals are in parentheses. "High damage" units are those with above average damage parameter point estimates. "Low damage" units are those with below average damage parameter point estimates.

Table 6 : Simulated gains from policy differentiation– Cost minimization model

Differentiated policy	Differentiated tax (1)	Wedges (2)	Ratios (3)
Permit price ($/ton NOx)	--	$1780	$1820
Ozone season emissions (thousand tons NOx)	682	652	658
% permitted emissions occurring at high damage sources	27.1%	27.6%	27.9%
Comparison with undifferentiated benchmark			
Change in annual benefits ($M)	$174.0 ($50.9, $327.6)	$161.2 ($41.9,305.1)	$174.3 ($51.1, $327.0)
% Change in annual benefits		26% (8%, 34%)	30% (11%, 37%)
Change in levelized annual abatement costs ($ M)	$51.3	$79.3	$92.4
Net gains from differentiation ($M)	$81.9 (-$41.1, $235.5)	$81.9 (-$37.4,$225.8)	$81.2 (-$41, $234.9)
Net gains from differentiation (%)	20% (-5%, 88%)	29% (0%, 65%)	37% (0%, 73%)

Notes: This table summarizes the results from simulating investment in NOx abatement and the associated ozone-season emissions under two alternative differentiated policy designs. 95 percent confidence intervals are in parentheses. "High damage" units are those with above average damage parameter point estimates. "Low damage" units are those with below average damage parameter point estimates.

Table 7 : Simulated outcomes under differentiated policy – Econometric model

Differentiated policy	Differentiated tax (1)	Wedges (2)	Ratios (3)
Permit price ($/ton NOx)	--	$4,400	$4,660
Ozone season emissions (thousand tons NOx)	879	658	658
% permitted emissions occurring at high damage sources	33.1%	37.1%	34.2%
Results reported relative to undifferentiated benchmark			
Change in annual benefits ($M)	-$306.5 (-$755.5, -$52.2)	$22.5 ($5.3, $53.5)	$60.0 ($13.5, $140)
% change in annual benefits	-22% (-31%, 46%)	3% (1%, 4%)	8% (3%, 12%)
Change in levelized annual abatement costs ($ M) (Measure 1)	-$297.5	$1.6	$8.4
Net gains from differentiation ($M) (Measure 1)	-$9.0 (-$434.9, $252.3)	$20.8 ($3.6, $51.9)	$51.6 ($5.1, $131.5)
Change in levelized annual costs ($M) (Measure 2)	-$562	$21.1	$77.6
Net gains from differentiation ($M) Measure 2	$255 ($118.9, $805.3)	$1.4 (-$15.8, $32.4)	-$17.6 (-$64.1,$62.4)

Notes: This table summarizes the results from simulating investment in NOx abatement and the associated ozone-season emissions under two alternative differentiated policy designs. 95 percent confidence intervals are in parentheses. "High damage" units are those with above average damage parameter point estimates. "Low damage" units are those with below average damage parameter point estimates.