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 program designs 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 policy 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 begin by analyzing outcomes under the observed (undifferentiated) policy design; we estimate that benefits exceed costs by almost 50 percent. We document significant variation in pollution damages across the coal plants that are responsible for most of the emissions capped by this program. This creates the potential for significant gains from policy differentiation. An econometrically estimated model of firms’ compliance choices is used to simulate outcomes under observed (undifferentiated) and counterfactual (differentiated) scenarios. A trading regime that defines terms of compliance using damage ratios results in welfare losses vis a vis the undifferentiated baseline. An alternative design based on damage differences delivers moderate welfare gains. We conclude that there are potential welfare gains from policy differentiation, but the devil is in the details.

Keywords: Market-based Policy, NO\textsubscript{x}, Budget Program, Policy Instrument Choice.

JEL Classifications: Q54, Q53, Q58
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), efficiency requires that the marginal costs of pollution abatement should vary across sources according to the degree of damage caused (Montgomery, 1972; Tietenberg, 2006). To achieve this, policy incentives must reflect variation in damages across sources. However, the vast majority of existing and planned programs are implemented as spatially uniform, "undifferentiated" trading regimes wherein all regulated emissions are penalized at the same permit price.

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 (SO\textsubscript{2}), two criteria pollutants currently subject to undifferentiated market-based regulation. Historically, these pollutants have been regulated using undifferentiated emissions trading programs. In 2008, a federal appeals court ruled that this regulatory approach fails to adequately accommodate spatial transport of pollution and the associated variation in damages across sources.\footnote{On July 11, 2008, in North Carolina v. EPA, the U.S. Court of Appeals for the D.C. Circuit vacated the Clean Air Interstate Rule (CAIR) which was to intended to provide a cost-effective, market-oriented approach to regulating a non-uniformly mixed pollutant (subject to compliance feasibility). State of North Carolina v. Environmental Protection Agency, No. 05-1244, slip op. (2008), District of Columbia Court of Appeals.} Since that time, debates surrounding market-based regulation of non-uniformly mixed pollutants have grown increasingly contentious. This paper investigates alternative approaches to regulating non-uniformly mixed pollutants from both theoretical and applied policy perspectives.

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.

To lay the foundations for a detailed analysis of the welfare gains from policy differentiation, we introduce a simple theoretical framework. As emphasized by Mendelsohn (1986), we show that the 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 can 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.

Building on this foundation, we estimate the gains from policy differentiation in the context of a landmark emissions trading program: the NOx Budget Program (NBP). Section 3 introduces this 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 damage-based restrictions on interregional trading (FR 63(90): 25902). Ex ante policy simulations projected nominal gains from policy differentiation (Krupnick et al. 2000, US EPA, 1998). Ultimately, it was decided that the potential benefits from this additional complexity would not justify the costs. 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 this paper, we revisit the decision to forego spatially differentiated NOx trading in favor of the simpler undifferentiated alternative. Our analysis is motivated by three observations:

First, in contrast to the pollution damage estimates that were used in the analysis that informed the design of the NOx Budget Program, state-of-the-art integrated assessment models now provide much richer estimates of the variation in damages from pollution. We use a stochastic integrated assessment model, AP2, 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. Moreover, we find that almost half of the variation occurs within (versus between) states. Previous analyses of spatially differentiated NOx trading considered multi-state zonal approaches to policy differentiation (Krupnick et al., 2000; USEPA, 1998). Our results suggest that differentiating emissions at the state-level will ignore a significant portion of the spatial variation in damages.

Second, policy analysts have duly noted that standard approaches to simulating permit market outcomes may fail to capture salient features of the real world decision processes that drive emissions abatement.

---

2 AP2 is the stochastic version of the APEEP model (Muller, Mendelsohn, 2007; 2009).
3 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.
decisions (Krupnick et al. 2000). Ex ante analyses of the gains from policy differentiation typically use deterministic, compliance cost-minimization algorithms to simulate firms’ regulatory compliance decisions. Engineering estimates of the capital and operating costs associated with different pollution abatement options are combined with basic assumptions regarding investment time horizons and discount rates. If firms’ environmental compliance choices deviate significantly from the predictions of these cost minimization algorithms, these policy simulation models will fail to accurately predict the gains from policy differentiation. To explore the extent of this inaccuracy, we conduct two sets of policy simulations. The first uses a fairly stylized cost-minimization-based model calibrated to mimic the simulation models that informed the design of the NOx Budget Program. The second approach takes advantage of the ex post nature of our analysis. We can observe detailed information about the compliance decisions that firms in this program actually made. We thus can incorporate an econometrically estimated model of firms’ compliance decisions. Importantly, the two modeling approaches predict substantially different firm-level responses to damage-differentiated policy incentives.

The third observation pertains to the policy relevance of the issues we address. In the years since the NOx Budget Program was implemented, debates over how to regulate non-uniformly mixed pollution have become more heated and complex. In court proceedings that challenge various aspects of the prevailing, undifferentiated policy designs, the need for more sophisticated analyses has been identified. We bring state-of-the-art integrated assessment modeling of emissions damages, together with econometric modeling of firms’ responses to market-based emissions regulation incentives, to bear on the analysis of the gains from policy differentiation.

In this working paper, we emphasize welfare comparisons across "emissions equivalent" permit trading regimes. More precisely, we compare undifferentiated and differentiated trading program designs that yield a level of emissions equal to the cap imposed in the NOx Budget Program. Although this approach has some expositional advantages, there are limitations. First, in practice, it will be difficult for the regulator to control the equilibrium level of emissions in a damage differentiated regime because the number of permits allocated need not equal the number of emissions. Second, this approach cannot accommodate the fact that the optimal level of aggregate emissions will change across policy designs. For these reasons, future versions of the paper will focus on comparisons of policy designs in which the emissions cap is set optimally conditional on the policy design and the information available at the time of policy implementation.4

Our first set of results pertain to the simulations based upon the engineering cost minimization model. Given our damage estimates and the assumptions implicit in the cost minimization model, we find that

4In future versions of the paper, we will approach the policy design problem from the perspective of a well informed regulator. We ask: how would a regulator equipped with state-of-the art modeling capabilities design and implement the NOx Budget Program. We assume this regulator has access to engineering estimates of abatement costs (based on ex ante observable plant characteristics) and source-specific damage estimates generated using stochastic integrated assessment modeling. Based on the information that the regulator could conceivably access ex ante, we identify optimal differentiated and undifferentiated policy designs. We then simulate outcomes under these alternative policy scenarios. Our preferred estimates make use of the econometrically estimated model of firms’ compliance choices. That is, we use information we can now observe regarding firms’ actual compliance choice to inform our modeling of firm behavior under differentiated and undifferentiated policy regimes.
the emissions cap imposed under the NBP was set close to optimal. Our estimates of the gains from policy differentiation are economically significant and relatively consistent across alternative differentiated policy designs. Annual damages from the permitted pollution are reduced by an average 30 percent relative to the undifferentiated policy. Net benefits of the trading program (i.e. avoided damages less abatement costs) increase by more than 20 percent. These estimates are larger than those produced by earlier studies (e.g., Krupnick et al., 2000; USEPA, 1998). The main reason is that our analysis captures significantly more spatial variation in emissions damages.

The second set of results incorporate the econometrically estimated model of firms’ compliance choices. Importantly, we find that the abatement cost curves implied by the econometric model are substantially steeper than those implied by the engineering cost minimization model.\(^5\) Particularly among more capital intensive pollution abatement options, abatement costs as perceived by firms appear to exceed our engineering estimates. This has two implications for our policy analysis. First, 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 (all else equal). Second, if the abatement costs perceived by firms accurately capture the true social cost of pollution abatement, the emissions cap imposed under the NBP is too stringent (in economic terms). In fact, we find that a differentiated policy designs which presumes an optimal cap could reduce welfare vis a vis the undifferentiated regime. An alternative policy design that maximizes welfare subject to a pre-determined emissions cap yields moderate welfare gains vis a vis the undifferentiated policy.

Finally, our analysis takes account of the significant 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).\(^6\) We find that uncertainty in our marginal damage estimates translates into significant uncertainty surrounding the estimated gains from policy differentiation.

The remainder of this paper is organized as follows. Section 2 introduces a simple model useful for analyzing the welfare implications of damage-based policy differentiation. Section 3 provides an overview of the NOx Budget Trading Program Section 4 describes how we construct our estimates of the potential gains from policy differentiation in the NBP. Section 5 summarizes our results. Concluding remarks are offered in Section 6.

---

\(^5\)Notably, Carlson et al. (2000) document very similar discrepancies in the context of the Acid Rain Program. In the policy context we consider, marginal abatement cost curves, firms operating plants in recently restructured electricity markets were reluctant to make large capital investments in pollution abatement equipment on account of low credit ratings and electricity market uncertainty (Fowlie, 2010). This could exaggerate the discrepancy between engineering cost estimates and those perceived by firms.

\(^6\)The report specifically requested the development of probabilistic, multiple-source uncertainty models based not only on available data but also on expert judgment.
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.\footnote{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.} 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: $D_i(e_i)$. 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 in source-specific emissions and additively separable. For each source, we define a marginal damage parameter $D_i'(e_i) \equiv \delta_i$. 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 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:

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

The first component in equation (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 equation (1) with respect to source-level emissions. As is well-known, assuming an interior solution, first-order conditions for total cost-minimization imply:

$$-C_i'(e^*_i) = \delta_i \quad \forall \ i.$$  \hspace{1cm} (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 $E^* = \sum_{i=1}^{N} e_i^*$. 

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, this analysis emphasizes 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.

In the emissions trading programs we consider, a quantity of tradeable emissions permits equal to the total emissions cap is 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 \((\delta_H)\); and the marginal damage for firm \((l)\), \(D_l(e_l)\), is denoted \((\delta_L)\).

We first consider an undifferentiated market-based policy design. We then contrast this with policy regimes in which the terms of compliance are designed to reflect heterogeneity in damages.

### 2.2 Undifferentiated emissions trading

Most existing and planned emissions trading programs feature undifferentiated permits: 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 \((\tau)\), to minimize the costs of complying with this emissions-based trading program.

\[
\begin{align*}
\min_{e_i, e_{si}, e_{bi}} & \quad C_i(e_i) + \tau(e_{bi} - e_{si} - A_i) \\
\text{s.t.} & \quad e_i \leq A_i - e_{si} + e_{bi} \\
& \quad e_i, e_{si}, e_{bi} \geq 0,
\end{align*}
\]

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

\[
C_h'(e_h^\mu) = C_l'(e_l^\mu) = \tau,
\]

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 $E$. We first consider a case in which the cap has been set to $E^*$. At the left origin, all emissions occur at the low damage firm (i.e. $e_l = E$) 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: $C_0'(e_l)$. At the right origin, the high damage firm emits $E^*$ (i.e. $e_h = E$) 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 $C_0'(e_h)$.

Equilibrium emissions under the undifferentiated trading regime are given by $\{e^u_l, e^u_h\}$. 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. However, this is not the optimal outcome. Total social welfare could be improved by shifting some of the permitted emissions away from the high damage source to the low damage source. In other words, an undifferentiated market design does not 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 emissions trading

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. It is straightforward to operationalize these ratios within our simple analytical framework.

Let $\bar{\delta}$ represent the average of the marginal damage across all sources in a trading program. In this simple two firm case, $\bar{\delta} = \frac{\delta_l + \delta_h}{2}$. We construct firm-specific damage ratios $r_i$, normalizing each firm’s marginal damage by the mean damage parameter: $r_i = \frac{\delta_i}{\delta}$, and therefore, $r_i \frac{r_j}{r_j} = \frac{\delta_i}{\delta_j}$. To remain in compliance, each firm must hold $r_i$ permits to offset each unit of uncontrolled emissions. The firm’s compliance constraint is modified in the following manner:

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

---

Note that $E^*$ is the socially optimal level of emissions assuming that emissions are allocated efficiently across sources. Conditional on undifferentiated trading, the welfare maximizing cap is more stringent than $E^*$. We return to this issue below.

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.
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:

\[
\frac{C'_j(e^*_j)}{C'_i(e^*_i)} = \frac{\delta_j}{\delta_i}, \quad i \neq j.
\]  

(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: \( C'_r(e_i) \frac{1}{\delta_i}, \quad i = l, h \). 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.

In Figure 1, benefits from differentiation (in the form of avoided damages from permitted emissions) 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. Thus, in this comparison of "emissions equivalent" policy regimes, the differentiated design welfare-dominates the undifferentiated design.

Imposing more structure on the model obtains an intuitive analytical expression for the gains from differentiation. 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:

\[
C_i(e_i) = \alpha_0 - \alpha_1 e_i + \beta_i e_i^2.
\]

We accommodate heterogeneity in abatement costs by allowing the parameters of this function to vary across sources.\(^{10}\)

Solving for the optimal cap in terms of the model parameters yields the following:

\[
E^* = \frac{(\alpha_{1h} \beta_l + \alpha_{1l} \beta_h - \delta_h \beta_l - \delta_l \beta_h)}{2 \beta_l \beta_h}.
\]  

(6)

Note that this optimal cap depends not only on the damage parameters \( \{\delta\} \) and the cost parameters \( \{\alpha_1, \beta\} \), but also on the correlation between the source-specific damage parameters and the source-specific cost parameters \( \beta \).

\(^{10}\)This functional form assumption implies linear marginal abatement cost curves. In practice, these cost curves may be discontinuous as pollution abatement often involves lumpy investments in emissions reducing capital equipment. The policy simulation will accommodate these non-linearities.
Suppose the emissions cap given by (6) is imposed in both the differentiated and undifferentiated regime. 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:

\[ e^r_h - e^u_h = \frac{\delta_l - \delta_h}{2(\beta_h + \beta_l)} \]

\[ e^r_l - e^u_l = \frac{\delta_h - \delta_l}{2(\beta_h + \beta_l)} \]

Intuitively, when marginal damage parameters vary across sources, policy differentiation will shift 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 straightforward to derive an expression for the net benefits from differentiation in terms of the model parameters (see Appendix 2 for a complete derivation):

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

The \( u \) superscript denotes the undifferentiated design; \( r \) denotes the differentiated design that incorporates damage-based ratios. The vectors of damage parameters and abatement cost coefficients are denoted \( \delta \) and \( \beta \), respectively.

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

(R1) The extent to which policy differentiation reduces pollution damages (via a reallocation of permitted emissions) and increases net welfare gains is increasing with the variation in damages across sources.

(R2) The extent to which policy differentiation reduces pollution damages (via a reallocation of permitted emissions) and increases net welfare gains is decreasing with the slope of the marginal abatement cost functions.

These findings should be intuitive. If damages do not vary across sources, there is no advantage to differentiated policy. Accordingly, the more heterogeneous the damages, the greater the benefits from differentiation, all else equal. 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 cost parameters 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.

In this working paper, we emphasize comparisons across emissions equivalent policy scenarios. An expository advantage of this approach: any differences in welfare across emissions equivalent differentiated
and undifferentiated designs stem solely from damage-based differentiation (versus changes in aggregate emissions).

One disadvantage of imposing emissions equivalence is that it complicates the implementation of damage-differentiated designs somewhat. Under a standard ratio-based regime, a fixed number of emissions permits is allocated. Source-specific, damage-based ratios define the number of permits that each source must hold per unit of emissions. If low (high) damage firms hold more of the allocated permits in equilibrium, aggregate emissions will exceed (fall below) the number of permits allocated. In order to achieve a pre-determined emissions (versus emissions permit) target, the damage-based ratios are instead used to define a compliance "true-up" payment (versus the number of permits held per unit of emissions). As in the undifferentiated regime, firms must hold a number of permits equal to uncontrolled emissions. In order to maintain compliance, in addition to surrendering permits, relatively high damage firms pay \((r_i - 1) \tau\) per unit of emissions. Relatively low damage firms receive a rebate of \((r_i - 1) \tau\) per permit. This damage-based true-up drives a wedge between the market clearing permit price and the price paid by firms whose marginal damage parameter exceeds or falls below average. The possible advantage of this differentiated design is that it allows the regulator to meet a pre-specified cap. One disadvantage is that it is not public revenue neutral.  

2.4 Differentiated emissions tax

Our analysis emphasizes those differentiated policy designs that would require relatively small modifications to the current emissions trading framework. However, it will also be instructive to 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 \(\delta_i\) 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.

2.5 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,

\[\text{11}\]

Under the benchmark emissions-based design, the implementing agency does not pay out or take in funds while administering program compliance. In contrast, when the damage-differentiated regime is designed to achieve a pre-determined emissions target, compliance payments to low damage firms need not equal compliance fees collected from high 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.
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 $\bar{E}$:

$$\min_{e_h, e_l} TSC = D_h(e_h) + D_l(e_l) + C_h(e_h) + C_l(e_l)$$

s.t. $e_h + e_l \leq \bar{E}$

Assuming the constraint binds, first order conditions with respect to $(e_l)$ and $(e_h)$ for constrained total cost minimization imply\(^{12}\):

$$C'_h(e_h) - C'_l(e_l) = \delta_h - \delta_l$$

Intuitively, equation (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 equalities given by equation (2) and equation (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, equation (10) is not met if the first-best ratios are used to define the terms of compliance.

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 $\{e'_h, e'_l\}$. 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 equation (10). At this point, differences in marginal damages are set equal to differences in marginal abatement costs. This allocation minimizes the total social cost associated with the emissions constraint $\bar{E}$. 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

\(^{12}\)We derive (10) using the Lagrangian method. The first-order condition with respect to $(e)_i$ is: $\delta_i + C'_i(e_i) = \lambda$. Since $(\lambda)$, the shadow value of relaxing the emission constraint, also appears in the same way in the first-order conditions with respect to $(e_h)$, $(\lambda)$ drops out of (10).
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 in addition to holding a permit, each firm must also pay \((\delta_i - \overline{\delta_i})\) per unit of uncontrolled emissions.

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

\[
\min_{e_i, e_{si}, e_{bi}} C_i(e_i) + (\tau + w_i)(e_{bi} - e_{si} - A_i) \\
\text{s.t. } e_i \leq A_i - e_{si} + e_{bi} \\
\quad e_i, e_{si}, e_{bi} \geq 0,
\]

This "wedge-based" policy design 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 \((e_{bi})\), and sales of allowances \((e_{si})\) under this spatially differentiated regime indicate that cost minimizing behavior in the part of firms delivers the socially optimal outcome in the constrained setting:

\[
C_h'(e_h) - C_l'(e_l) = (\delta_h - \delta_l).
\]

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:

\((R3)\) When the emissions cap is not set optimally, 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.
2.6 Uncertain damages

In this final section, we consider the policy implications of the considerable uncertainty that pervades the policy design process. Policy makers are confronted with uncertainty regarding both costs and damages from pollution. In our analysis of policy design, 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 here. Moreover, state of the art integrated assessment modeling provides policy makers with rich information regarding damage uncertainty which can be usefully incorporated into policy design. Uncertainty regarding costs is more difficult to formally incorporate.

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 model uncertainty.

2.6.1 Design implications of damage uncertainty

We first consider how damage uncertainty affects the design of damage-differentiated policy using a very basic framework. We assume that the joint distribution of the marginal damage parameters $f(\delta)$ is known ex ante. A risk neutral policy maker seeks to minimize the expected social costs of pollution. Importantly, we assume that policy parameters are chosen at the outset of the program before uncertainty is resolved. The terms of compliance are defined in the program design stage and are fixed for the duration.\(^{13}\)

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:

$$
\begin{align*}
\min_{e_h, e_l} \quad & TSC = C_h(e_h) + C_l(e_l) + \int \int (\delta_h e_h + \delta_l e_l) f(\delta_l, \delta_h) \\
\text{s.t.} \quad & e_h + e_l = E
\end{align*}
$$

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

\(^{13}\)We argue that these are reasonable assumptions/restrictions. Much of the variation in damage estimates is driven either by the stochastic nature of the inputs (e.g. meteorological conditions) or parameter uncertainty that is unlikely to be resolved over the duration of the program (e.g. uncertainty regarding dose response parameters).
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. This implies that the wedge parameters $w_i$ should be constructed using expected source-specific marginal damages $E[\delta_h], E[\delta_l]$.

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, linearity in 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. Were this the case, the regulator might seek to aggregate sources with similar damage estimates so as to implement a zonal policy design.

### 2.6.2 The gains from differentiation under damage uncertainty

Consider a policy regime that incorporates wedges defined using $E[\delta_l]$ and $E[\delta_h]$. By equation (14), these compliance parameters will maximize welfare in expectation (conditional on the emissions constraint). In order to illustrate how damage uncertainty leads to uncertainty about the gains from policy differentiation, it is instructive to consider the gains from differentiation conditional on a particular realization of $\delta$. Let $\delta' = \{\delta'_h, \delta'_l\}$ denote a draw from the joint distribution of marginal damages $f(\delta)$. Given this realization of damages, 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:

$$TSC^w(\delta'; \delta) - TSC^u(\delta'; \delta) = \frac{1}{2} \frac{(\delta_l - \delta_h) (E[\delta_l] - E[\delta_h])}{(\beta_l + \beta_h)} - \frac{1}{4} \frac{(E[\delta_l] - E[\delta_h])^2}{(\beta_l + \beta_h)}.$$  

The first argument in equation (15) represents the difference in realized 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). This expression helps to highlight
the implications of having realized damages $\delta'$ differ from the damage parameters that are used to parameterize the terms of compliance $E[\delta]$. When damages are uncertain, it is possible for the realized net benefits of differentiation to be negative.\textsuperscript{14} This

\textit{(R4A) The greater the correlation between realized and expected damages, the greater the realized gains from differentiation.}

The expected gain from policy differentiation is obtained by integrating over the entire distribution of damages (see Appendix 4):

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

(16)

Uncertainty in the damage parameter estimates has potentially significant implications for the gains from policy differentiation:

\textit{(R4B) Benefits from policy 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 likely that the parameters used to define the terms of compliance ($E[\delta_l]$ and $E[\delta_h]$) will inaccurately 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.

\section{The NOx Budget Program}

Section 2 provides a simple framework for thinking about the gains from policy differentiation and the various factors that determine those gains. In what follows, this framework is used to analyze outcomes in a large regional emissions trading program: 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. Over the period 2003-2008, the program established a region-wide cap on emissions of NOx from large stationary sources in twenty eastern states.

\footnote{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.}

16
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 policy design stage, 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 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.15

In 2008, a federal district court vacated the rule that was to replace the NOx Budget Program due to policy’s failure to adequately accommodate regional transport of pollution and associated spatial variation in damages.16 Since that time, debates surrounding the market-based regulation of non-uniformly mixed criteria pollutants have become increasingly contentious. In the interest of informing this important debate, we revisit the decision to forego a differentiated NBP 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).17 By exempting some units from our analysis, we are implicitly assuming that operating decisions at gas and oil-fired 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.

4 Estimating the benefits from differentiation

15 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).
16 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.
17 Natural gas and oil-fueled plants tend to have much lower uncontrolled NOx emissions rates. 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).
We use the theory model presented in section 2 as foundation for a detailed analysis of the benefits from differentiation in the context of the NOx Budget Program. This applied analysis proceeds in several steps:

1. Estimate the marginal damage parameters $\delta$ for each source in the NOx Budget Program using a stochastic integrated assessment model. These $\delta$ are used to parameterize the damage-differentiated policies under consideration (i.e. damage based ratios, damage-based wedges, and the differentiated tax regime).

2. Construct engineering estimates of source-specific, technology-specific abatement costs and emissions reduction efficiencies.

3. Simulate the source-level compliance choices under both the observed and counterfactual (differentiated) policy regimes. We do this in two ways. First, we use a deterministic cost-minimization algorithm designed to mimic standard policy simulation models. Second, we use an econometrically estimated model of firms’ compliance choices under the NBP.

4. The two sets of source-level compliance choice simulations yield corresponding sets of source-level emissions. The integrated assessment model is re-introduced for the purpose of estimating the total damages associated with the permitted level of NOx emissions under the observed and counterfactual policy regimes.

5. The abatement costs associated with simulated compliance choices are estimated. A comparison of benefits and costs under differentiated and undifferentiated regimes yields an estimate of the net benefits of policy differentiation.

Each of these steps are described in detail in the following subsections.

4.1 Estimating source-specific damages from pollution

NOx emissions affect health and environmental outcomes through two main pathways: ozone formation and particulate matter formation.\textsuperscript{18} 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).

\textsuperscript{18}NOx emissions also contribute to acid rain in some mountain regions, and exacerbate eutrophication problems.
The extent to which NO\textsubscript{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 NO\textsubscript{x} 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 NO\textsubscript{x} 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 NO\textsubscript{x} 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 ($\delta_i$) parameters capture the estimated effect of an incremental change in NO\textsubscript{x} 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\textsubscript{x}, PM\textsubscript{2.5}, sulfur dioxide (SO\textsubscript{2}), volatile organic compounds (VOCs), and ammonia (NH\textsubscript{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 $\beta_{kp}$ 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 $\alpha_k$ 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.

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

$$\delta_i = \sum_{ri} \sum_k \sum_s \alpha_k \beta_{kp}^s P_{ri} \frac{dC_{si}(e_i)}{de_i}. \tag{17}$$

Given the stochastic nature of AP2, the parameters of the atmospheric model, the population estimates $P_{ri}$, the dose response parameters $\beta_{kp}^s$ and valuation parameters $\alpha$ 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. The AP2 model is used to construct an empirical distribution for each $\delta_i$ parameter. We first make a random draw (denoted the $m^{th}$ draw) from the distributions of the parameters in (17). Next, the model is used to compute the $m^{th}$ 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.
AP2 computes the difference between damages with baseline emissions and after adding the ton of NO\textsubscript{x} to (i). This is repeated 4,999 times to estimate the empirical distribution of marginal damages for NO\textsubscript{x} emitted from facility i. This process is then repeated for each EGU in the analysis.

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 NO\textsubscript{x}. 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 NO\textsubscript{x}. In the subsequent discussion, we classify any source with estimated damages exceeding (falling below) $1711/ton NO\textsubscript{x}$ 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 $\delta$ is negative. This suggests that a decrease in NO\textsubscript{x} emissions at these sources leads to increased overall damages. This result is driven by the complex, non-linear photochemical reactions that transform NO\textsubscript{x} and VOCs into ozone. Daily ozone concentrations are non-linear and monotonic functions of NO\textsubscript{x} and the ratio of volatile organic compounds (VOCs) and NO\textsubscript{x}. At sufficiently low ratios, the conversion of NO\textsubscript{x} to ozone is limited by the availability of VOCs. In these VOC limited conditions, reductions of NO\textsubscript{x} can increase peak ozone levels until the system transitions out of a VOC-limited state (Seinfeld and Pandis, 1998). 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.

### 4.1.2 Parameterizing a damage-differentiated policy

The unit-specific damage parameter estimates summarized by Figure 4 are used to define the terms of compliance in the differentiated policies we consider. 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 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 $\delta_i$ 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 Ex ante expected, source-specific NOx abatement costs

The NBP mandated a dramatic reduction in average NOx emissions rates. 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. This was a fairly aggressive policy implementation time-frame. The compliance choices made during this four year period are modeled as static decisions.

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 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

---

19 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/mbtu).

20 We 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).
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\textsuperscript{21} 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 boiler-specific cost estimates used in this analysis (EPRI, 1999b). This cost estimation exercise is described in detail in Fowlie (2010). Importantly, the boiler-level and plant-level parameters (including boiler technology type, plant vintage, plant capacity) that generate variation in these cost and emissions estimates are plausibly exogenous to the compliance choices we are interested in modeling.

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. Variation in costs and emissions reduction efficiencies is explained. These estimates will be used in our modeling of firm-level compliance decisions.

\section*{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 couple engineering estimates of abatement costs and potential emissions reductions

\textsuperscript{21}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.
(IPM, ISIS). 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 inaccurately 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 simple cost minimization model of firms’ compliance choices. This model is calibrated to mimic standard policy simulation models. The second approach replaces the cost minimization algorithm with an econometrically estimated model of the compliance choice.

4.3.1 Cost-minimization algorithm

We use a simple algorithm to find the combination of NOx control options that minimizes the levelized annual cost of meeting the emissions cap. Let \( j = 1...J_i \) index the NOx control technology options available to the \( i^{th} \) 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 (i.e. the amount of NOx the \( i^{th} \) 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:

\[
\min_j C_{ij} = v_{ij}Q_i + l_i K_{ij}
\]

\[
\text{where } v_{ij} = (V_{ij} + \tau m_{ij})Q_i.
\]

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).\(^{22}\) 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.\(^{23}\) 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.\(^{24}\) 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

\(^{22}\) 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.

\(^{23}\) For more complete documentation of this model and its assumptions, see http://www.epa.gov/airmarkt/progregs/epa-ipm/. This discount rate is slightly lower than the 6 percent assumed by Carlson et al. (2000).

\(^{24}\) 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 \( (Q_n) \) to proxy for expected ozone season production.
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 \( \tau \) parameter in (??) is replaced with the source-specific expected damage parameter \( \delta_i \). In the differentiated regime that incorporates wedges, the variable compliance costs are redefined to be: \( (V_{ij} + \tau m_{ij} + w_{ij}m_{ij})Q_i \). 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 equilibrium permit market outcomes is as follows. The cap is set equal to the seasonal NOx emissions associated with the compliance choices we actually observe (597,000 tons).\(^{25}\) Beginning with an initial permit price \( \pi^0 \), 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.\(^{26}\)

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 compliance choices that firms actually made. The cost minimization model correctly predicts compliance choices at only 24% of the EGUs. 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 compliance choices observed in the NOx Budget Program depart markedly from those predicted by the cost minimizing algorithm. Carlson et al. (2000) document very similar discrepancies in the context of the Acid Rain Program using a very similar model of firm-level compliance choices. If the cost minimization-based policy simulations incorrectly predict how firms respond to the observed (undifferentiated) policy, these models should not be relied upon to accurately simulate firms’ response to 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 idiosyncrasies that determine firms’ environmental compliance choices.

### 4.3.2 An econometric model of the compliance decision

\(^{25}\)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.\(^{26}\)If this iterative procedure arrives at a point where it is oscillating 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.
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. We argue that this choice model can be used to generate more realistic estimates of the gains from policy differentiation as compared to simulations based on the calibrated cost minimization model.

The decision maker at unit $i$ is assumed to choose the compliance strategy that minimizes the unobserved latent value $C_{ij}$:

$$C_{ij} = \alpha_j + \beta_m^v v_{ij} + \beta^K K_{ij} + \beta^K A K_{ij} \cdot Age_{ij} + \varepsilon_{ij},$$

where $v_{ij} = (V_{ij} + \tau m_{ij}) Q_i$.

Note that this specification is structurally very similar to equation (3). The primary difference is that parameters are econometrically estimated versus calibrated to an existing simulation model. 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 $\alpha_j$ 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 $\varepsilon_{nj}$ are assumed to be iid extreme value and independent of the covariates in the model. The variable cost coefficient ($\beta^v$) and the capital cost coefficient ($\beta^K$) 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.27 The econometric model is 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 $\beta_m$ 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.

---

27 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.
Estimates of the parameters of the distribution of $\beta^v$ and $\beta^K$ 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 most likely have made in counterfactual policy scenarios.

Table 4 summarizes the choice model parameter estimates. 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 ($\beta^K$ and $\beta^v$). The ratio $(\beta^K + \beta^K A_{\text{Age}}) : \beta^v$ 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.\(^{28}\)

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. One difference is that the simulated emissions are defined probabilistically as $e_i = \sum_j P_{ij} e_{ij}$, where $P_{ij}$ is the predicted probability of unit $i$ choosing compliance strategy $j$. 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 econometric model correctly predicts that a significant portion of the mandated emissions reductions is achieved using more capital intensive compliance strategies.\(^{29}\)

\(^{28}\)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.

\(^{29}\)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),
4.4 Pollution damages

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. In this context, rather than systematically perturbing NO\textsubscript{x} emissions one source at-a-time, NO\textsubscript{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).

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.

4.5 Estimating the costs of compliance

To assess the net welfare impacts of policy differentiation, the gains (in the form of reduced damages) must be weighed against any increase in compliance costs. Accurate measurement of the costs of complying with environmental regulation is notoriously difficult. Complying with environmental regulations can 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).

Ideally, we want to estimate not only the abatement costs that were actually incurred under the undifferentiated policy, but also the costs that would have been incurred under the counterfactual, differentiated policies we consider. 30 We take two different approaches to constructing our cost estimates. The first approach uses our ex ante engineering estimates of boiler-specific abatement costs to estimate the costs that would actually be realized under different policy scenarios. In other words, we assume that the ex ante engineering estimates of costs are a good proxy for the costs that would actually materialize. The second approach derives the cost estimates that are most consistent with simulated choices. When

\footnote{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).}
the cost minimization algorithm is used to simulate source-level compliance choices, the two approaches to cost estimation yield identical results (by design). However, when the econometric model is used to simulate compliance choices, these two approaches to estimating the abatement costs associated with simulated choices yield very different results.

### 4.5.1 Engineering estimates of compliance costs

To estimate the abatement costs associated with the simulated compliance choices, we use the same cost assumptions and accounting parameters that were used to parameterize equation (2). When the cost minimization algorithm is used to simulate compliance choices, the levelized annual cost of compliance under policy regime \( r \) is defined to be:

\[
LAC^C_M^r = \sum_i V_{ir}Q_i + l_iK_{ir},
\]

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:

\[
LAC^{EST}_r = \sum_i \sum_j P_{irj}(V_{ir}Q_i + l_iK_{ir}),
\]

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

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).

### 4.5.2 Compliance costs consistent with the compliance choice model

The second approach to estimating aggregate costs is more derivative. Note that both the cost minimization model and the econometrically estimated choice model can be used to simulate permit market outcomes, and corresponding equilibrium permit prices, over a range of emissions limits or caps. Plotting simulated permit prices against the corresponding emissions reductions traces out an aggregate marginal abatement cost (MAC) curve. The integral under this curve yields an estimate of total abatement cost.

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 imposed under the NBP. The lower MAC curve is generated using the model that assumes strict cost minimization and undifferentiated permit trading. The more steeply sloped MAC curve is generated using the econometric model of the compliance choice (under undifferentiated trading).
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 $\beta_k$ and $\beta_v$ 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 to satisfy 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. A similar approach is used to estimate the costs associated with the damage-based ratio, wedge, and tax regimes.

### 4.5.3 Caveats

Neither of these approaches to estimating compliance costs are ideal. Both rely to a significant extent on the ex ante expected engineering cost estimates $V_{ij}$ and $K_{ij}$ to estimate the true costs of installing and operating the full range of feasible compliance options. These 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.

When the econometric model of firms’ compliance choices is used to simulate policy outcomes, we will find that the cost estimates that are most consistent with the econometric model exceeds the estimate obtained using the engineering cost estimates. The former is more consistent with the costs as perceived by the firms. If the difference between the two is attributable to distortions and idiosyncrasies (such as status quo bias or price discrimination on the part of pollution control equipment manufacturers) that drive a wedge between private and social costs, then equation (4.5.1) provides a more appropriate measure for use in our welfare calculations. However, it is also possible that engineering cost estimates may underestimate the true cost of financing investments in pollution abatement equipment. We will interpret the higher estimate as an upper bound on the aggregate compliance costs that were actually incurred under the undifferentiated regime.

---

31 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.
5 Results and synthesis

This section is divided into three sections. 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 policy regime. Results generated using the calibrated cost minimization algorithm are reported in column (1). Results generated using the econometrically estimated choice model are reported in column (2). The simulation outcomes in this table 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. The equilibrium permit price required to incentivize the level of abatement to meet the cap is significantly higher in the simulation model that incorporates the econometric model of firms’ compliance choices. 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 approximately equals the average permit price that prevailed during the period of time that compliance decisions were made.

Table 5 also provides information regarding the distribution of permitted 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. 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 reason is that the simulated aggregate emissions are not exactly equal across the two sets of simulations. Although the emissions cap is held constant across scenarios, this cap is never exactly met due to non-convexities in the MAC curves. This leads to varying degrees of over-compliance.

\[^{32}\text{Strictly speaking, these should not be interpreted as confidence intervals because our analysis does not account for all sources of uncertainty. For example, we do not capture modeling uncertainty or cost uncertainty in our analysis.}\]
The lower panel in Table 5 reports the estimated costs of complying with the observed, undifferentiated policy. In column (1), the two approaches to estimating this aggregate compliance cost yield the same cost number by construction. In the second column, the cost estimates that are most consistent with the econometric choice model are higher than those constructed using engineering estimates of levelized annual costs. Subtracting costs from simulated benefits yields an estimate of net benefits. Our most conservative estimate of the net benefits conferred by the NOx Budget Program is $319 million per year. The estimate obtained using the calibrated model of cost minimization is $643 million per year.\footnote{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.}

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.

5.2.1 A graphical introduction

Figures 7 and 8 provide an graphical introduction to our results. Figure 7 illustrates how source-specific policy incentives vary across the four regimes we analyze. The top panel corresponds to the calibrated cost minimization model. The bottom panel corresponds to the 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).

The source-specific policy incentives (measured in $/ton NOx) are defined as follows. In the undifferentiated regime, the cost to offset emissions is simply the simulated permit price (see Table 5). Under the differentiated tax, the cost to offset a ton of NOx equals the source-specific marginal damage estimate $\delta_i$. Under the wedge-based design, source $i$ must pay $(\delta_i - \overline{\delta}) + \tau^w$ per ton of emissions to remain in compliance, where $\tau^w$ is the equilibrium permit price. Under the ratio-based regime, the cost per ton of emissions is $\frac{\delta_i}{\overline{\delta}} \tau^r$.

If the imposed cap is optimally set, the cost to offset a ton of emissions at firm $i$ will be equal to $\delta_i$ across all three differentiated policy regimes. This is approximately true in the top panel of Figure 7. Conditional on the assumed costs, damages, and the calibrated parameters of the choice model, the cap in the NBP is approximately optimal. The incentives under the undifferentiated policy lie along the horizontal line at $1620/ton$. The incentives associated with the differentiated policy regimes lie approximately along what would be a 45 degree line if the horizontal and vertical axes were symmetrically scaled.
If the imposed cap is not optimally set, policy incentives can differ significantly within a source, across differentiated policy regimes. Conditional on the assumed costs, damages, and the econometrically estimated choice model parameters, the emissions cap imposed under the NBP is too strict. This generates variation in the source-specific policy incentives across differentiated regimes. The incentives under the differentiated tax are exogenous to the model and thus unaffected by how we model firms’ compliance choices. The incentives under the wedge-based policy, \((\delta_i - \bar{\delta}) + \tau'\), lie along a parallel line above the tax incentives \(\delta_i\). The incentives generated by the trading ratio regime, \(\frac{\delta_i}{\delta'} \tau\), form the steep line in Figure 7. Intuitively, differences in these marginal incentives across sources with different damage parameters are exaggerated when the cap is set too stringently.

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, steeper abatement costs imply smaller responses to policy differentiation, all else equal. Note also that the simulated reallocation of permitted emissions is more significant under the trading ratio regime compared to the trading wedge regime. This difference occurs because the ratio design exaggerates the variation in the source-specific incentive to reduce emissions relative to the wedges (see Figure 7).

5.2.2 Counterfactual simulations using the engineering cost minimization model

Table 6 provides a more detailed numerical summary of the results generated using the calibrated cost minimization model. Recall that key parameters (i.e., the discount rate and investment horizon) are taken from standard policy simulation models. This calibrated model poorly predicts outcomes under the
observed differentiated regime. These simulation results are primarily intended as a basis for comparison against the preferred simulations based on the econometrically estimated choice model.

The top panel reports the simulated market clearing permit price where applicable and the associated emissions. 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). Simulated emissions under the first-best tax define the optimal cap which is within 5 percent of the imposed cap.

The second panel of Table 6 reports government revenues. In the tax regime, the government collects tax payments from all sources. Recall that the differentiated ratio and wedge-based regimes are designed to deliver the level of emissions observed under the existing cap. Compliance terms are differentiated outside the permit market via compliance "true up" payments. Because the majority of permitted emissions occur at low damage sources, the payments collected from relatively high damage sources do not offset payment owed to relatively low damage sources. 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 these damage-differentiated designs. Moreover, our analysis assumes away any deadweight loss to taxation.

The third panel reports simulated changes in avoided damages relative to the undifferentiated policy baseline. The first row reports results generated using the integrated assessment model. The differentiated tax is associated with the smallest increase in benefits from pollution reduction (an estimated 12 percent increase). This is because the aggregate level of emissions increases under the tax vis a vis the undifferentiated baseline. Among emissions equivalent scenarios, the ratio-based regime is associated with the largest increase (an estimated 18 percent increase). This is because the ratios provide the strongest incentive to reallocate emissions from high to low damage sources.

The second row reports a linear approximation of these damages. Throughout the paper, we have assumed a linear, additively separable damage function. As an additional robustness check for this assumption, we construct an estimate of total damages by simply multiplying the simulated marginal damages $\delta_i$ by the corresponding simulated emissions change $\Delta \hat{e}_i$. We do this 5,000 times to match the 5,000 damage estimates for each EGU. We are careful to align the 5,000 draws when we construct measures of the aggregate change in avoided damages: $\sum_i \Delta \hat{e}_i \cdot \delta_i$. A comparison of the two sets of numbers in the second panel of table 6 provides an alternative test of the structure we impose on the aggregate damage function. These numbers suggest that the complicated processes captured by the integrated assessment model are not perfectly, but approximately, captured by a linear additively separable damage function.

The fourth panel of Table 6 reports estimated changes in abatement costs (relative to the undifferentiated policy) and the net gains from policy differentiation (i.e. the value of avoided damages less costs). Abatement costs increase under the damage-differentiated trading programs as permitted emissions are reallocated from lower cost (high damage) sources to relatively high cost (lower damage) sources.

Estimated net benefits are quite similar across the three policy regimes we consider. When net benefits are constructed using the simulated damages, the ratio-based regime welfare dominates. This is an
unexpected result which contradicts result (R4) above. However, the welfare ranking is as we would expect when the linear approximation to damages is used. Note that (R4) is predicated on the assumption of a linear, additively separable damage function. When this assumption fails to hold exactly, we see the trading ratios dominating both wedges and the damage-differentiated tax. Although this is a noteworthy finding, it seems hard to imagine that the benefits associated with defining terms of compliance that capture non-linearities and interactions in the damage function could outweigh the costs (in terms of additional complexity in both design and implementation).

5.2.3 Counterfactual simulations using the econometrically estimated choice model

Table 7 summarizes our preferred results which are generated using the econometric model of firms’ compliance choices. We believe the econometric model offers more realistic predictions of how firms would actually have responded to the differentiated policy incentives we consider.

In the top panel of Table 7, we find that the optimal emissions level (given our damage estimates, engineering cost estimates, and choice model parameters) is 34 percent higher than the emissions cap imposed. Thus, we should expect to see more dramatic differences across the damage-differentiated policy regimes we consider. Consistent with Figure 8, note that the reallocation of emissions under the damage-differentiated policies (vis a vis the undifferentiated regime) is less significant when the econometric choice model (versus the calibrated cost minimization algorithm) is used to simulate firms’ response to differentiated policy incentives.

The second panel reports government revenues. Tax revenues are higher as compared to Table 6 because abatement costs most consistent with the econometric model exceed engineering cost estimates. Thus, the aggregate emissions level under the tax increases (as does the tax revenues collected). Under the wedge-based regime, the net payment owed by the government falls (relative to Table 6) because a smaller share of emissions are reallocated from high to low damage sources. In contrast, the net payment owed by the government increases because these compliance payments are a function of the equilibrium permit price (which rises to $4,660 from $1820).

The third panel summarizes the estimated changes in pollution damages induced by policy differentiation. Note that significantly higher levels of emissions under the first-best tax (as compared to the undifferentiated trading program) imply significantly higher damages overall. The wedge-based regime increases expected program benefits (in terms of avoided emissions) vis a vis the benchmark, but only by 2 percent. The ratio-based policy increases the expected benefits of the program by less than 5 percent. Note that these point estimates of the benefits from policy differentiation are smaller (in levels and in percentage terms) as compared to the results reported in Table 6. By result (R2), gains from differentiation are decreasing with the steepness of the marginal abatement cost curves. Figure 6 shows that the marginal abatement cost curves implied by the econometrically estimated choice model are steeper than those associated with the calibrated cost minimization model.
Net gains from policy differentiation are reported in the lower panel of Table 7. In this case, the welfare ranking is consistent regardless of which damage estimates are used. And this welfare ranking is intuitive. The differentiated tax delivers the greatest gains, primarily due to the fact that abatement costs are significantly reduced as the level of emissions increases. 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. Taking the imposed emissions cap as given, point estimates of the gains from the wedge-based policy are small and positive.

5.3 Uncertainty

Thus far, this summary has focused almost exclusively on point estimates. In this section, we discuss the significant uncertainty surrounding these estimates. Tables 5, 6 and 7 summarize the outcomes of 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. The 95 percent confidence intervals provide a sense of this range. It should be noted, however, that our analysis accounts for only a subset of the sources of uncertainty. So these cannot be interpreted as confidence intervals in a strict sense.

Observation (R4B) 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

---

34 When we use equation (4.5.1) 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 (R3) 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 (4.5.1). More specifically, marginal abatement costs as perceived 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 exaggerate 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 (4.5.1).
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 and uncertainty about how to value heterogeneous damages from pollution. We examine how differentiated policy designs can 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 Budgeted Program (NBP). Prior research has shown that the damages due to NO\textsubscript{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 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. We then simulate plant-level outcomes under both observed (undifferentiated) and counterfactual (differentiated) policy designs. The abatement costs and environmental damages that correspond to these simulated compliance choices are then tabulated and compared. Importantly, total emission levels are held constant across all of the emissions trading regimes we consider. What changes across designs is the allocation of these permitted emissions across sources.

Our preferred approach to modeling firms’ compliance choices under observed and counterfactual policy designs takes advantage of the fact that we now observe the compliance options that firms actually chose when complying with the NOx Budget program. These data are used to estimate a model of how facility-level compliance choices were made. We use this model to simulate how sources in the program would most likely have responded under counterfactual policy incentives.

There are several important findings. The first 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 some fairly simulation models that are used to inform policy-making. 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, abatement costs as perceived by firms significantly exceeded ex ante engineering estimates. Steep marginal abatement costs curves imply smaller gains from policy differentiation, all else equal. We thus conclude that more standard approaches to simulating firm-level responses to policy incentives significantly overestimate of the gains from differentiation.

Second, if we assume that the compliance costs that are implied by our econometric estimates capture true social costs of pollution abatement, we conclude that some forms of policy differentiation could reduce welfare vis a vis the undifferentiated policy that was implemented. Intuitively, this is due to the fact that the imposed cap was too stringent, conditional on our damage and cost estimates. We introduce an alternative policy design based on damage differences. We find that this alternative design delivers moderate gains ($21 M per year).
Finally, our analytical approach allows us to examine how uncertainty in our marginal damage estimates translates into uncertainty surrounding the estimated gains from policy differentiation. The confidence intervals we construct around our estimated gains from policy differentiation are very large due to the significant parameter uncertainty that pervades the integrated assessment of source specific damages. This uncertainty further complicates the regulation of non-uniformly mixed pollution. Although uncertainty is not grounds for policy inaction, it does undermine the political palatability of any proposed policy design change.

References


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: Source-specific Policy Incentives

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.
### Table 1: Unit-level summary statistic

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

**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

<table>
<thead>
<tr>
<th>NOx control technology</th>
<th>Capital cost ($/kW)</th>
<th>Variable cost (cents/kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High damage</td>
<td>Low damage</td>
</tr>
<tr>
<td>Combustion modification</td>
<td>6.13</td>
<td>8.12</td>
</tr>
<tr>
<td></td>
<td>(10.64)</td>
<td>(17.74)</td>
</tr>
<tr>
<td>Low NOx burners</td>
<td>17.45</td>
<td>21.98</td>
</tr>
<tr>
<td></td>
<td>(19.94)</td>
<td>(28.47)</td>
</tr>
<tr>
<td>SNCR</td>
<td>7.01</td>
<td>8.93</td>
</tr>
<tr>
<td></td>
<td>(10.09)</td>
<td>(11.66)</td>
</tr>
<tr>
<td>SCR</td>
<td>70.94</td>
<td>80.40</td>
</tr>
<tr>
<td></td>
<td>(127.99)</td>
<td>(155.01)</td>
</tr>
</tbody>
</table>

**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

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

**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

<table>
<thead>
<tr>
<th>Technology specific constants</th>
<th>High damage units</th>
<th>Low damage units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-combustion controls</td>
<td>-2.21 (1.66)</td>
<td>-3.06 (1.34)</td>
</tr>
<tr>
<td>Low NOx burners</td>
<td>-2.06 (0.53)</td>
<td>-2.33 (0.43)</td>
</tr>
<tr>
<td>Combustion modifications</td>
<td>-1.89 (0.85)</td>
<td>-2.32 (0.69)</td>
</tr>
<tr>
<td>Age* capital cost interaction</td>
<td>-0.17 (0.07)</td>
<td>-0.13 (0.06)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Model of compliance choice</th>
<th>Cost minimization (1)</th>
<th>Econometric (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permit price ($/ton NOx)</td>
<td>$1,620</td>
<td>$4,460</td>
</tr>
<tr>
<td>Ozone season emissions (thousand tons NOx)</td>
<td>655.8</td>
<td>658.2</td>
</tr>
<tr>
<td>Annual benefits ($M) (monetized avoided damages)</td>
<td>$1,061 ($184, $2,615)</td>
<td>$1,075 ($198, $2,636)</td>
</tr>
<tr>
<td>% permitted emissions occurring at high damage sources</td>
<td>41.8%</td>
<td>38.9%</td>
</tr>
<tr>
<td>Levelized annual abatement costs ($M) (Cost measure 1)</td>
<td>$417.6</td>
<td>$692.3</td>
</tr>
<tr>
<td>Levelized annual abatement Costs ($M) (Cost measure 2)</td>
<td>$417.6</td>
<td>$755.9</td>
</tr>
<tr>
<td>Annual net benefits ($M) (Cost measure 1)</td>
<td>$643.3 (-$233, $2,197)</td>
<td>$382.7 (-$494, $1943)</td>
</tr>
<tr>
<td>Annual net benefits ($M) (Cost measure 2)</td>
<td>$643.3 (-$233, $2,197)</td>
<td>$319.1 (-$557, $1880)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Differentiated policy</th>
<th>Differentiated tax (1)</th>
<th>Wedges (2)</th>
<th>Ratios (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permit price ($/ton NOx)</td>
<td>--</td>
<td>$1780</td>
<td>$1820</td>
</tr>
<tr>
<td>Ozone season emissions (thousand tons NOx)</td>
<td>682</td>
<td>658.4</td>
<td>652.7</td>
</tr>
<tr>
<td>% permitted emissions occurring at high damage sources</td>
<td>27.1%</td>
<td>27.6%</td>
<td>27.9%</td>
</tr>
<tr>
<td>Government revenues ($M)</td>
<td>$999.6</td>
<td>-$153.8</td>
<td>-$173.4</td>
</tr>
</tbody>
</table>

Comparison with undifferentiated benchmark

| Change in annual benefits ($M) | $124.8 ($6.8, $303.7) | $157.4 ($36.8, $360.5) | $194.2 ($48.4, $439.5) |
| Change in annual benefits (linear approximation) | $127.7 ($14.7, $325.2) | $160.3 ($41.9, $304.8) | $174.1 ($51.0, $327.0) |
| Change in levelized annual abatement costs ($M) | $29.3 | $65.2 | $77.1 |
| Net gains from differentiation ($M) | $95.5 ($-22.5, $274.4) | $92.1 ($-28.3, $295.1) | $117.1 ($-28.5, $362.4) |
| Net gains from differentiation (linear measure) ($M) | $98.4 ($-14.6, $295.9) | $95.1 ($-23.2, $239.5) | $96.9 ($-26.0, $249.9) |

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

<table>
<thead>
<tr>
<th>Differentiated policy</th>
<th>Differentiated tax (1)</th>
<th>Wedges (2)</th>
<th>Ratios (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permit price ($/ton NOx)</td>
<td>--</td>
<td>$4,400</td>
<td>$4,660</td>
</tr>
<tr>
<td>Ozone season emissions (thousand tons NOx)</td>
<td>879.2</td>
<td>658.4</td>
<td>658.5</td>
</tr>
<tr>
<td>% permitted emissions occurring at high damage sources</td>
<td>33.1%</td>
<td>37.1%</td>
<td>34.2%</td>
</tr>
<tr>
<td>Government revenues ($M)</td>
<td>$1,413.6</td>
<td>-$40.3</td>
<td>-$208.3</td>
</tr>
</tbody>
</table>

Results reported relative to undifferentiated benchmark

| Change in annual benefits ($M) | -$340.2 (-$988.5 - $113.2) | $22.5 ($5.3, $53.5) | $60.0 ($13.5, $140) |
| Change in annual benefits (linear approximation) | -$306.5 (-$754.9, -$52.3) | $21.7 ($4.7, $45.6) | $57.7 ($13.7, $114.7) |

Change in levelized annual abatement costs ($M) (Measure 1)

| Net gains from differentiation ($M) | $132.5 ($-515.8, $359.5) | $22.0 ($-4.8, $53.0) | -$9.0 ($-55.0, $77) |
| Net gains from differentiation (linear benefits measure) ($M) | $166.2 ($-282.2, $525.0) | $21.2 ($4.2, $45.1) | -$11.3 ($-55.3, $45.7) |

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.